MODELLING AND SIMULATION 2017

THE EUROPEAN SIMULATION

AND

MODELLING CONFERENCE

2017

ESM®'2017

EDITED BY

Paulo J.S. Gonçalves

OCTOBER 25-27, 2017

LISBON

PORTUGAL

A Publication of EUROSIS-ETI

The 31st Annual European Simulation and Modelling Conference 2017

LISBON, PORTUGAL

OCTOBER 25-27, 2017

Organised by

ETI - The European Technology Institute

Sponsored by

EUROSIS - The European Simulation Society

IST - Instituto Superior Técnico

IDMEC

Co-Sponsored by

Ghent University

University of Skovde

Hosted by

IST

Instituto Superior Técnico
Lisbon, Portugal

EXECUTIVE EDITOR

PHILIPPE GERIL
(BELGIUM)

EDITORS

General Conference Chair

Paulo J.S. Gonçalves
Instituto Politécnico de Castelo Branco, Portugal,
and
IDMEC/LAETA, Instituto Superior Técnico, Universidade de Lisboa, Portugal

Past Conference Chairs

José Évora Gómez, SIANI, Universidad de Las Palmas de GC, Las Palmas, Spain
Jose Juan Hernandez-Cabrera, SIANI, Universidad de Las Palmas de GC, Las Palmas, Spain
Mario Hernandez-Tejera, SIANI, Universidad de Las Palmas de GC, Las Palmas, Spain
Octavio Roncal-Andrés, SIANI, Universidad de Las Palmas de GC, Las Palmas, Spain

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
Erik Lindskog, Chalmers Univ. of Techn, Gotheburg, Sweden
J. Manuel Feliz Teixeira, University of Porto, Porto, Portugal
Bert van Beek, Eindhoven University of Technology, Eindhoven, The Netherlands

Discrete Simulation Modeling Techniques and Tools
Renato Natal Jorge, FEUP - University of Porto, Porto, Portugal
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, TU Delft, The Netherlands
Morched Zeghal, Nat. Res.Council Canada, Ottawa, Canada

Agent Based Simulation
Zisheng Huang, Vrije Universiteit Amsterdam, The Netherlands
Ioan Alfred Letia, TU Cluj Napoca, Romania
Isabel Praca, Ist. Superior do Porto, Portugal

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

INTERNATIONAL PROGRAMME COMMITTEE

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

Simulation in Education and Graphics Visualization
Ana Luísa Ramos, University of Aveiro, Aveiro, Portugal
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
Filipe Pinto, Polytechnic Institute of Leiria, Portugal

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
Yan Luo, NIST, Gaithersburg, USA
Jose Machado, University of Minho, Braga, Portugal

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

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 with Petri Nets
Track Chair
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

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
Véronique Baudin, LAAS-CNRS, Université de Toulouse, Toulouse, France
Konstantinos Kirytopoulos, University of South Australia, Australia
Lode Vermeersch, Delcredere, 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 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 Grafisch Bedrijf Lammaing, Ostend, Belgium

EUROSIS-ETI Publication

ISBN: 978-9492859-00-6
EAN: 9789492859006

EUROPEAN SIMULATION AND MODELLING CONFERENCE 2017

Preface

It is my privilege and pleasure to welcome you to the 31st European Simulation and Modelling Conference - ESM'2017 held in association with IST (Instituto Superior Técnico de Lisboa) and IDMEC. As in previous editions, this conference promotes the knowledge exchange in the varied fields of Simulation and Modelling, allowing participants to share their experiences in an open forum providing an enriching environment and stimulating debates.

We have some eighty participants from 28 countries with presentations grouped in 20 main themes such as Simulation Methodology and Tools; Data Simulation; Financial Simulation; Decision Management, Production Scheduling, Supply Chain and Inventory Management, Logistics and Traffic Simulation, Sensor and Electronics Simulation; Biological Simulation Models; to mention a few. In this conference, we will have the opportunity to listen to Nicolae Vasiliu and Daniela Vasiliu who will talk about Sizing and Tuning the Damper of an Aerospace Electrohydraulic Servomechanism by Amesim, Joël Colloc giving a talk on A Fuzzy Vectorial Space that avoids to Defuzzify the Membership Functions and last but not least Helena Barbas talking about VR, AR, MR Simulations and Inspirations from “Iron Man 3”.

I would like to welcome all participants and to thank all authors for sharing their knowledge and experience. My thanks also goes to all members of the Program Committee for the reviewing work that was key to maintaining the high scientific quality of ESM'2017. I am also grateful to the Keynote Speakers who are willing to share their extensive knowlegde and experience in the field of simulation and modelling with us.

Furthremore, a special thanks to Philippe Geril from EUROSIS, whose continued dedication and hard work, as the conference organizer, has enabled the organization to maintain the standard expected for ESM’2017. Finally, I would like to express my gratitude to IST for its support and to the hosting venue for a job well done.

Last but not least, I would like to wish you all a fruitful and productive experience at the conference and an enjoyable and enriching stay in Lisbon.

Paulo J.S. Gonçalves
Instituto Politécnico de Castelo Branco, Portugal
and
IDMEC/LAETA, Instituto Superior Técnico
Universidade de Lisboa, Portugal

ESM’2017 General Chairman

CONTENTS

Preface ..... **IX**
Scientific Programme **1**
Author Listing **435**

KEYNOTES

Sizing and Tuning the Damper of an Aerospace Electrohydraulic Servomechanism by Amesim
Nicolae Vasiliu, Daniela Vasiliu, Constantin Călinoiu, Radu Puhalschi and Petru-Cristinel Irimia **5**

A Fuzzy Vectorial Space that avoids to Defuzzify the Membership Functions
Joël Colloc **13**

VR, AR, MR Simulations and Inspirations from "Iron Man 3"
Helena Barbas **25**

SIMULATION METHODOLOGY AND TOOLS

A Timed Extended Reachability Graph for the Simulation and Analysis of Bounded Timed Petri Nets
Dimitri Lefebvre **33**

Visual Estimation of Persistence in Time Series
Sihle Poswayo and Igor Litvine **38**

Looking for and finding Lost Data
William Conley **46**

Efficient Quantile Estimation via a Combination of Importance Sampling and Latin Hypercube Sampling
Marvin K. Nakayama **49**

Statistical Optimization Applied to a Variety of Non-Linear Systems of Equations
William Conley **54**

GRAPHICAL DATA SIMULATION

Multi-Master Replication in Eventually Consistent Simulation Grids
Stefan Elsen **63**

CONTENTS

Visualising of Co-occurrence Data
Igor Litvine and Oksana Ryabchenko **.....68**

Performance Comparison of Adapted Delaunay Triangulation Method over Nurbs for Surface Optimization Problems
Suyesh Bhattarai, Parag Vichare and Keshav Dahal **.....76**

Automatic Generation of a Diagnostic and Control Unit for controlling Embedded Systems Application. Case Study: HEVC Decoder
Habib Smei, Abderrazak Jemai and Kamel Smiri **.....81**

3D Filtering Color Image contaminated by mixed Noise using Sparse Representation
Alfredo Palacios-Enriquez, Volodymyr Ponomaryov and Araceli Hernandez-Fragoso **.....88**

Enhancing the Accuracy of Raster Based Algorithms for Forest Fire Spread Modelling
Yves Dumond **.....93**

FINANCIAL SIMULATION

Analysis of Relationship between Risk and (expected) Return of the Investment (Portfolio) – Simulation Experiment on the Prague Stock Exchange
Adam Borovička **.....103**

Optimal Dating of Cycles in Financial Time Series
Konrad Kapp and Igor Litvine **.....111**

Mining Patterns in Financial Time Series using Dynamic Time Warping Algorithm
Kristina Šutienė, Audrius Kabašinskas, Eimutis Valakevičius and Roland Reichardt **.....119**

DECISION MANAGEMENT SIMULATION

Generating Synthetic Individual Human Population and Activity Models
Emily Schmidt and Dhananjai M. Rao **.....127**

A Domain Specific Language for Complex Dynamic Decision Making
Souvik Barat, Vinay Kulkarni, Tony Clark and Balbir Barn **.....135**

Support for Management Processes of the Exercises of the Crisis Staffs of critical Infrastructure Entities
Jiří Barta and Josef Navrátil **.....143**

A co-simulation Framework Interoperability for Neo-campus Project
Yassine Motie, Alexandre Nketsa and Philippe Truillet.....148

PRODUCTION SCHEDULING

Modular Hybrid Modeling Based on DEVS for Interdisciplinary Simulation of Production Systems
Bernhard Heinzl, Philipp Raich, Franz Preyser, Wolfgang Kastner, Peter Smolek and Ines Leobner.....157

Simulation of a Flexible and Adaptable One-Piece-Flow Assembly Line Based on a Process Flow of Colored and Timed Petri Nets
Benedikt A. Latos, Peyman Kalantar, Philipp M. Przybysz, Susanne Mütze-Niewöhner, Christoph Holtkötter and Jan Brinkjans.....162

Production Process Evaluation and Improvement by using the Method of Discrete Event Simulation
Tola Kudret Karaca and Volkan Cakir.....167

A Simulation Based Evaluation of Dynamic Task Prioritization in Maintenance Management
Dietmar Neubacher, Nikolaus Furian, Clemens Gutschi, Tobias Elmer and Siegfried Vössner.....174

SUPPLY CHAIN SIMULATION

Modelling and Simulation for Decentralized Supply Chain Formation
Florina Livia Covaci.....185

Towards Quantitative Risk Evaluation for Supply Chains in Preparation of a Simulation Study
Birgit Mösl, Dietmar Neubacher, Nikolaus Furian and Siegfried Vössner 192

INVENTORY MANAGEMENT OPTIMIZATION

The Concept of Semi-Variance as a Tool for Safety Inventory Decisions in Case of Uncertain Demand
Katrien Ramaekers, Galina Merkuryeva and Gerrit K. Janssens.....201

Mass Customization Dynamics Simulation for Fashion and Apparel Market
Jocelyn Bellemare.....206

CONTENTS

Simulation Optimization: A Simple Approach combining Metaheuristics and Metamodels
Luiz Ricardo Pinto, Júlia Cobucci Morais, Gabriela Martins Nunes and João Flavio de Freitas Almeida **212**

LOGISTICS SIMULATION

Simulation of Logistics for Construction Management
Nikolaus Furian, Dietmar Neubacher, Siegfried Vössner, Philip Santner, Michael O'Sullivan and Cameron Walker **221**

Cooperative Decision Making Modeling in Transportation Logistics Dispatching System
Anton Ivaschenko, Ilya Syusin and Pavel Sitnikov **227**

Comparison of Cost Performance of Fixed and Flexible Collection Strategy in Return Logistics Network
Di Zhang and Uwe Clausen **233**

Evaluation in Transport Planning: A Comparison between Data Envelopment Analysis and Multi Criteria Decision Making Methods
Giuseppe Musolino, Corrado Rindone and Antonino Vitetta **238**

TRAFFIC SIMULATION

Evaluation of Car-following-Models at controlled Intersections
Laura Bieker-Walz, Michael Behrisch, Marek Junghans and Kay Gimm **247**

A Stochastic Driver Distraction Model for Microscopic Traffic Simulations
Manuel Lindorfer, Christian Backfrieder, Gerald Ostermayer and Christoph F. Mecklenbräuker **252**

Utilisation of Computer Simulation for dynamic Calculation of Train Delays
Jan Fikejz and Josef Brožek **258**

Hybrid Optimizing Models for Planning Charging Infrastructures
Hubert Büchter and Sebastian Naumann **264**

SENSOR NETWORK SIMULATION

Simulation of Optimized Nonlinear Frequency Modulation in Pulse Compression Radar
Jiří Roleček, Pavel Bezoušek and Karel Juryca **273**

An Open Architecture Framework for the Electronic Warfare Modeling & Simulation
Sang Yeong Choi, Hyeon Seo Kang, Hyeong Jun Kwon and Sug Joon Yoon **.....278**

On the Effects of the Variations in Network Characteristics in Cyber Physical Systems
Géza Szabó, Sándor Rácz, József Petö and Rafael Roque Aschoff **.....283**

Identifying the Optimal Transmission Range in Depth-Based Routing for UWSN
Mohsin Jafri, Simonetta Balsamo and Andrea Marin **.....288**

ELECTRONICS SIMULATION

Dynamic Switching of Processor Simulation Model Accuracy
Johannes Kohl, Dietmar Fey and Jürgen Bäsig **.....295**

A New Method to transform Petri Nets to Digital Circuits using Input Driven Reachability Graphs
Christoph Brandau and Dietmar Tutsch **.....301**

FLUID SYSTEMS SIMULATION

Monte Carlo Simulation of Daily Precipitation and River Flow Conditional Spatio-Temporal Fields
Nina A. Kargapolova **.....311**

Forecast of selected Quality Indicators of Wastewater flowing to the Treatment Plant using selected Black-Box Methods
Bartosz Szeląg, Krzysztof Barbusiński, Agnieszka Operacz and Jan Studziński **.....316**

ENERGY FORECASTING AND OPTIMIZATION

Neural Network Models and Electricity Demand Forecasting
Francis Bismans and Igor N. Litvine **.....325**

Optimisation of Compressed Air System's Energy Usage through Discrete Event Simulation Compressor Performance
Robbie Mulvany, Alan Arokiam, Abdelhafid Belaidi, John Ladbrook and Michael Higgins **.....328**

CONTENTS

Investigation of the Modelling Effects on the Steam Generator's Behaviour during the early Stages of a Station Blackout in a CANDU 6 Reactor
Roxana-Mihaela Nistor-Vlad, Daniel Dupleac, Ilie Prisecaru and Chris Allison **336**

AEROSPACE SIMULATION

Time Management of Heterogeneous Distributed Simulation
Clément Michel, Janette Cardoso and Pierre Siron **343**

Real-Time Simulation of Large Aircraft Fuel Systems
Stephen Wright and Alvery Grazebrook **350**

An Evaluation Framework for UAV Surveillance Applications
Michael Ettlinger, Bilge Sarp, Christopher-Eyk Hrabia and Sahin Albayrak **356**

ENGINEERING SIMULATION

Applying the Model Driven Architecture Approach to Dynamic Structure Applications
Min Zhu, Clément Foucher, Vincent Albert and Alexandre Nketsa **365**

Modelling and Optimal Control with Energy Regeneration of a 6DOF Motion Platform with permanent Magnet Linear Actuators
E. Thöndel **373**

HUMAN COMPUTER INTERFACES

Human-Computer Interface for Communication and Automated Estimation of Basic Emotional States
Svetla Radeva, Strahil Sokolov and Dimitar Radev **381**

Simulation-Based User Interfaces for Digital Twins Pre-, In-, or Post Operational Analysis and Exploration of Virtual Testbeds
Torben Cichon and Juergen Rossmann **384**

Surgery Assistant Based on Augmented Reality
Anton Ivaschenko, Alexandr Kolsanov and Aikush Nazaryan **390**

BIOLOGICAL DATA SIMULATION

The controlled Development of (Medical) Self-Diagnosis Systems with Self-Enforcing Networks
Christina Klüver and Jürgen Klüver **397**

Model of Modular IoT-based Bee-Keeping System
K. Dineva and T. Atanasova **.....404**

The Effect of Sexual Networks on Fertility Levels
Edinah Mudimu **.....407**

GRAPHICAL HUMAN BIO-ANALYSIS

Semi-automatic Brain Lesions Detection and Segmentation Method in MR Images
Carlos Segura Granados, Volodymyr Ponomaryov and Martha Hernandez-Cuellar**.....415**

Transcription Initiation Controls Skewness of the Distribution of Intervals between RNA Productions
Vinodh K. Kandavalli, Sofia Startceva and Andre S. Ribeiro**.....418**

SIMULATION IN HUMAN BIO-ANALYSIS

A Modeling Approach to Heart Failure Treatment
Alexander Lassnig, Christian Baumgartner and Jörg Schröttner **.....425**

Modeling Survival Times using Frailty Models
Liberato Camilleri, Roxanne Caruana and Alex Manche **.....428**

SCIENTIFIC PROGRAMME

KEYNOTES

SIZING AND TUNING THE DAMPER OF AN AEROSPACE ELECTROHYDRAULIC SERVOMECHANISM BY AMESIM

Nicolae Vasiliu, Daniela Vasiliu, Constantin Călinoiu, Radu Puhalschi
University POLITEHNICA of Bucharest
313, Splaiul Independentei, Sector 6
RO 060042 Bucharest, Romania
E-mail : nicolae.vasiliu@upb.ro

Petru-Cristinel Irimia
SIEMENS INDUSTRIAL SOFTWARE
13 A, Bulevardul Gării
RO 500203 Braşov, Romania
E-mail : cristi.irimia@siemens.com

KEYWORDS
Numerical simulation, fine tuning, electrohydraulic servosystems, flight control systems

ABSTRACT
The modern technical systems are including a great number of fluid power systems. The proper implementation of such a control system in order to achieve high static and dynamic performance, and a good overall efficiency too, needs the use of all the tools developed by the modern systems theory: modeling, simulation, experimental identification, and practical validation.
Each of all the above activities means a deep knowledge of all components from different fields: hydromechanics, power electronics, computer science etc. In order to facilitate the matching of all these needs, the authors of the paper studied as a typical example the modeling and simulation of an important aerospace fluid power component – the rudder electrohydraulic driving system composed by a standard electrohydraulic servomechanism, and a custom designed double effect hydraulic damper.
The simulations were performed by a complete and friendly simulation environment: LMS Imagine. Lab Amesim, (shortly – Amesim) developed by SIEMENS PLM SOFTWARE - LMS. The simulation networks were built by components from the basic libraries available in Amesim.
Typical transients generated by the set points changes and the disturbances induced by the interactions with external systems were carefully investigated. A detailed analysis of the pressure evolution in the power hydraulic lines of the fluid power system showed a strong tendency toward the cavitation phenomena which can be avoided by different specific means studied by T.J. Viersma and other aerospace researchers.
Extensive experiments performed in author's fluid power laboratories, and other well-known laboratories, were found in good agreement with the simulated results. The paper promotes the concepts of numerical simulation by Amesim as a basic tool for engineering innovation in mechatronic systems.

PROBLEM FORMULATION

The modern flight control and guidance systems developed for giant civil or fight aerospace vehicles stated many complex problems for engineers, designers, and manufacturers. First of all, the ratio between the actuation force or torque and the mass of the control device has to be as high as possible. The limited space and weight, the large range of temperature, pressure and the absolute security requirements are severe conditions for any kind of aerospace vehicle, from civil airplanes and helicopters, to missiles or other defense vehicles. On the other hand, the tightening of the environment protection conditions leaded the civil airplanes manufacturers to increase continuously the size and the capacity of these important transport means. Today, the market leader from this point of view is A380 which needs only 3 liters of kerosene for each passenger which is transported over 100 km.
The main control systems accomplishing the above conditions are the electrohydraulic ones. Most of the moving components of a flight control systems are hydraulically driven because the highest ratio force/mass can be reached. For example, the rudder of A340 needs 225 kN, and is actuated by an electrohydraulic servomechanism weighting 100 kg.
Hydraulic actuators are now electronically controlled by digital optical fiber systems, giving them a good immunity regarding the electromagnetic disturbances. Figure 1 shows the main components of A380 controlled by electro hydraulic actuators. Much more, the landing gear, the brakes, the engines thrust reverser and the wheel brakes are also hydraulically actuated.

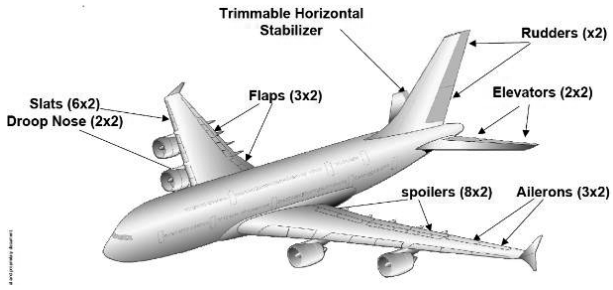


Figure 1: A380 flight control system [1].

One of the most important problems which faces the aerospace engineers who design electrohydraulic servomechanisms is the control stability. The interaction between the aerodynamic forces, the low structural stiffness and the hydraulic fluid compressibility generate oscillations which have to be limited in order to allow a

proper control of any mechanical system of the plane [2]. There are many technical solutions for solving these problems, but they are useful for moderate aerodynamic loads and mass [3], [4]. Most of them need a lot of tuning generated by the uncertain structural stiffness and natural frequencies of the driven components, connected with the hydraulic fluid compressibility. The resonant frequencies have to be as high as possible, and the amplitude of the oscillations has to be limited from fatigue considerations. The previous long experience in the domain generated different technical solution for solving this problem: use of servovalves with shaped windows in the sleeve, well controlled fluid leakages between the hydraulic cylinder chambers, hydro mechanical transient filters, hydraulic dampers in parallel with the hydraulic cylinder etc. The last solution, presented in Figure 2 was adopted by Airbus [2] for the rudders.

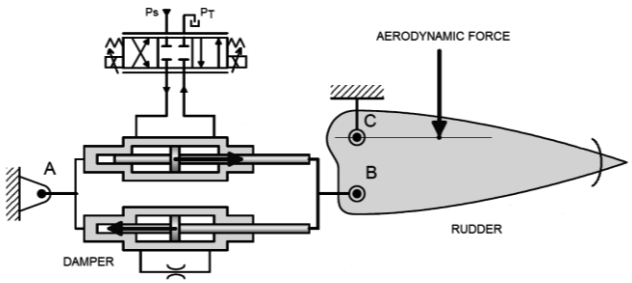


Figure 2: The electrohydraulic control system of the rudder with hydraulic damper [2].

The main advantage offered by a damper working in parallel with the hydraulic cylinder is the possibility of eliminating the steady state error in any position of the rudder. The calibrated orifice of the damper can be easy tuned by numerical simulation, and validated on the "IRON BIRD" test bench, without many corrections after the flight test. The damper design can be a classical one, using the same technology as for the hydraulic actuator. A sharp edge orifice placed in a mobile sleeve can be changed without dismantling other components of the assembly.

MODELING THE CONTROL SYSTEM BY AMESIM

In order to identify the simplest way for simulating the physical combination between the servo system and the damper, a careful comparative analysis of similar simulation software environments, including MATLAB - Simulink from Math Works, AUTOMATION STUDIO from Famic Technologies, FluidSIM from FESTO and LabVIEW from National Instruments, was performed. Finally, AMESIM environment, produced by Siemens PLM Software - LMS Company, a member of SIEMENS Group, was selected as optimal simulation tool. This complex software offers numerous advantages: rich library of hydraulic symbols and components, which allow the authors to use existing, proven models for well-known components (valves, cylinders); ability to simulate different part of the system at different levels of complexity, which allows the authors to model different parts of the system at different levels of detail, as required. AMESIM models are fully compatible with LabVIEW for real time and Hardware-in-the-Loop simulations, can be imported in LabVIEW and connected to a real-time or HIL simulation system.
The positioning system, the damper, the pressure source etc. have been modeled at a concise level, using predesigned blocks from the AMESIM Hydraulic and Mechanical libraries. Ultimately, the load of the system was considered a simple spring, without taking into account the different types of damping forces.
The simplest AMESIM model developed by the authors for the studied servo system can be seen in Figure 3.

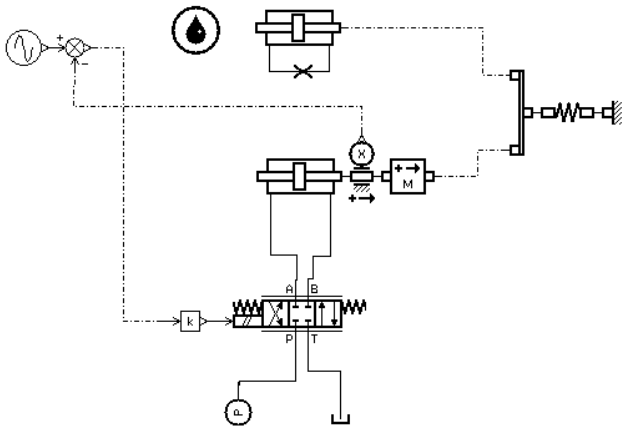


Figure 3: AMESim simulation network for a sine input signal.

The main components of the servomechanism have the followings features: hydraulic cylinder bore: 100 mm; rod diameter: 50 mm; maximum stroke: ±100 mm; nominal system pressure: 350 bar; damper cylinder bore: 40 mm; damper rod diameter: 20 mm; servovalve nominal flow: 60 l/min; damper nominal orifice diameter: 1 mm; equivalent rudder mass: 500 kg; aerodynamic force gradient, relative to the servomotor rod: 1000 N/mm.
The hydraulic fluid used in preliminary numerical simulations is a synthetic non-inflammable one, with a high bulk modulus (17,000 bar), a common density (850 kg/m^3), and a usual absolute viscosity of 51 cP at 40 °C.
A usual ratio air/gas content of about 0.1% was considered in the dynamic computation.

SINE INPUT RESPONSE OF THE SYSTEM

Three types of input signals were used in order to obtain a complete image of the dynamic behavior of the servomechanism: a constant amplitude, constant frequency sine input; an input small step; a constant amplitude, increasing frequency sine input.

The first category of simulation, performed for a sine waves of 0.5 Hz and 50% amplitude from the nominal one is presented in Figures 4 to 15.
The piston motion amplitude reaches 92% from the theoretical one (Figure 4).

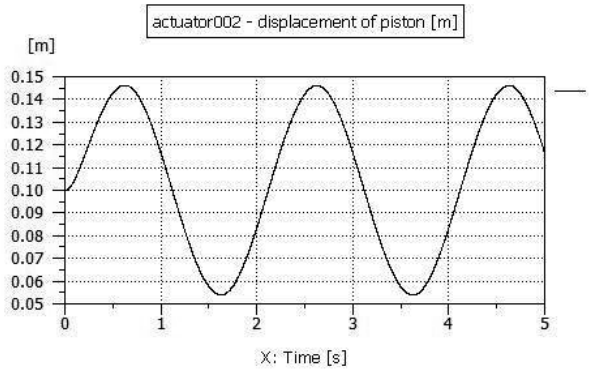


Figure 4: Displacement of the actuator piston.

The maximum piston velocity is about 0.14 m/s (Figure 5).

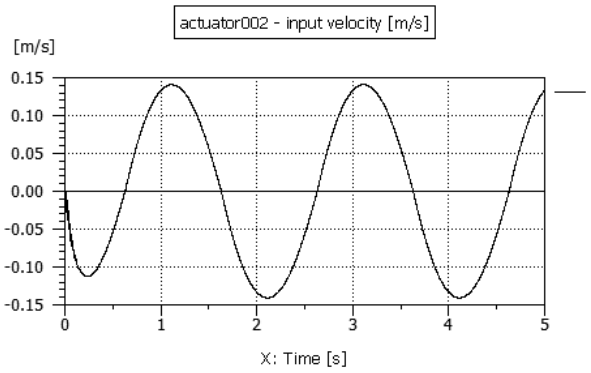


Figure 5 : Piston velocity evolution.

The pressures in the actuator chambers have periodical variations in opposition, between 135 and 215 bar (Figure 6), the average value (175 bar = 350 /2 bar) respecting the fundamental law of the critical center constant pressure supplied servovalves [3].

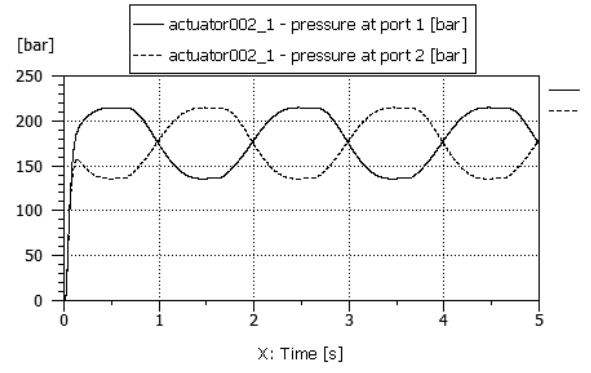


Figure 6: Pressure evolution at the actuator ports.

The force developed by the actuation rod reaches 46,500 N (Figure 7).

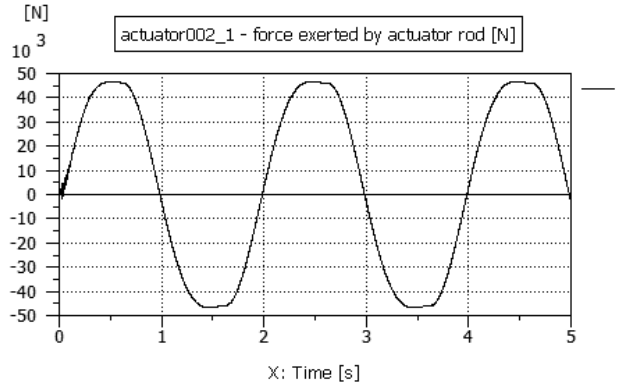


Figure 7 : Force exerted by the actuator rod.

The servovalve delivers to the actuator maximum 51,0 l/min -about 85% from the nominal flow (Figure 8). The servovalve spool displacement reaches 92% from the maximal stroke (Figure 9).

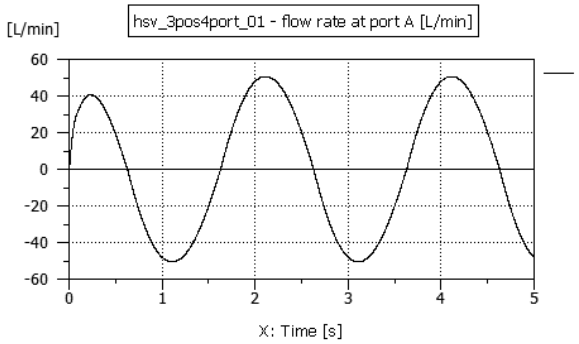


Figure 8: Servovalve flow rate at port A.

.

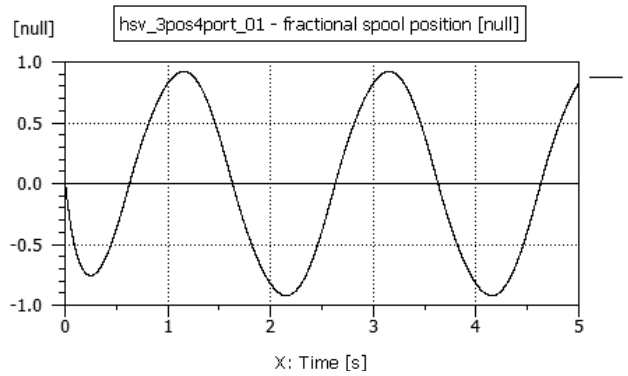


Figure 9: Fractional servovalve spool position.

The maximum flow rate generated by the damper piston reaches only 16% of the actuator one (Figure 10). The maximum force developed by the damper (about 23500 N) reaches 51 % of the force developed by the actuator one (Figure 11).

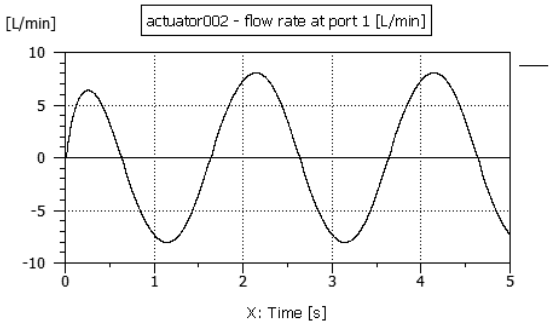


Figure 10: Damper flow rate.

.

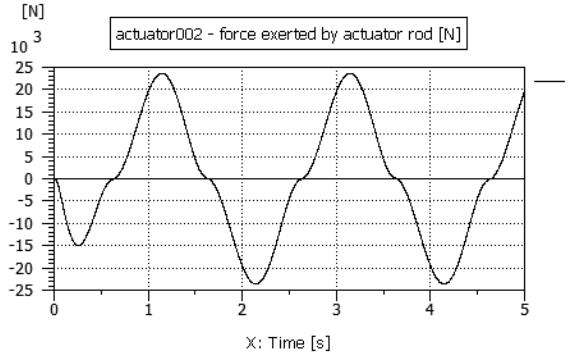


Figure 11: Damping force variation.

The overall force applied on the rudder by the combination between the actuator and the damper follows close enough the input position signal (Figure 12).

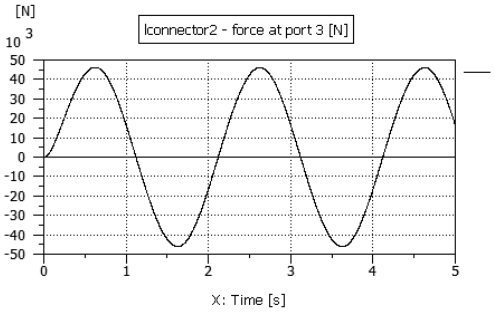


Figure 12: Variation of the force given by the combination between the actuator and the damper.

The pressures in the chambers of the damper are pulsating from zero to 250 bar (Figures13 and 14).

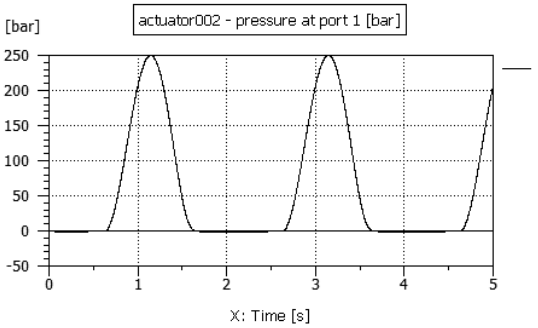


Figure 13: Pressure at port 1 of the damper.

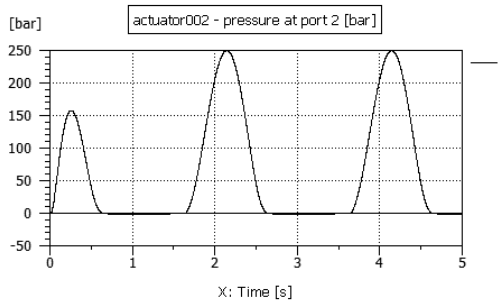


Figure 14: Pressure at port 2 of the damper.

The servomechanism closely follows a low frequency sine input (Figure 15).

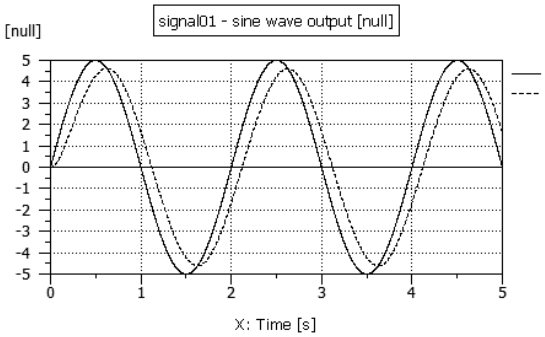


Figure 15: Sine input response of the servomechanism for f=0.5 Hz.

STEP INPUT RESPONSE OF THE SYSTEM

Another useful way to obtain practical information about the servomechanism behavior is to feed it with a small step input (Figure 16).

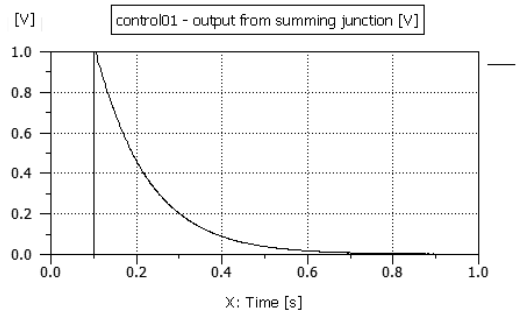


Figure 16: Output from summing junction.

The response is a first order one (Figure 17).

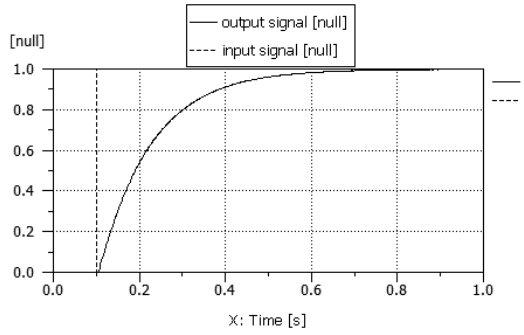


Figure 17: Input and output signal of the system.

The constant time of the actuator is small enough (0.12 s) for the size of the airplane rudder (Figure 18).

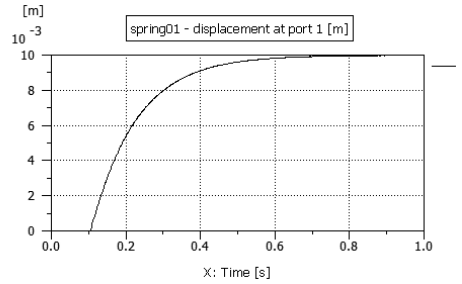


Figure 18: Piston displacement for a step input.

The pressure variations in the actuator's ports are typical for a third order system (Figure 19).

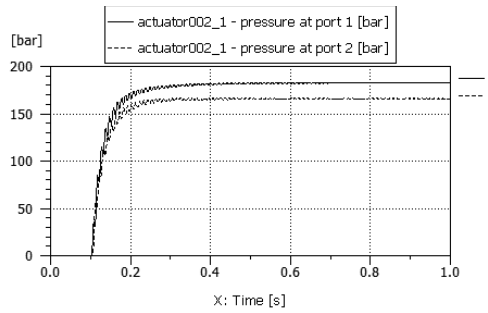


Figure 19: Pressures in the actuator's ports.

The high speed servovalve flow rate has a normal variation for a well damped control system (Figure 20).

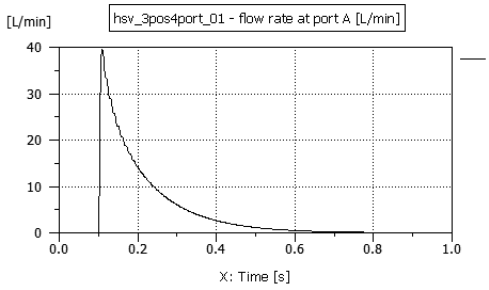


Figure 20: Flow rate at the port A of the servovalve.

According to the small step input, the fractional spool position of the servovalve is about 50% from the nominal one (Figure 21).

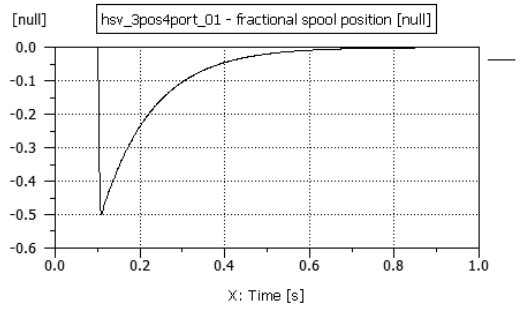


Figure 21: Fractional servovalve spool position.

The flow rate at the output port of the damper shows high frequency small amplitude variations (Figure 22). The force exerted by the actuator rod has a strong variation around the final value (Figure 23).

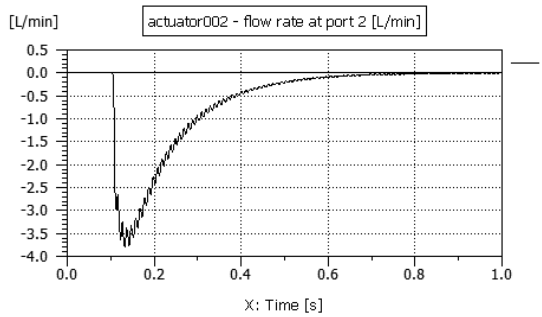


Figure 22: Flow rate at the output port of the damper.

.

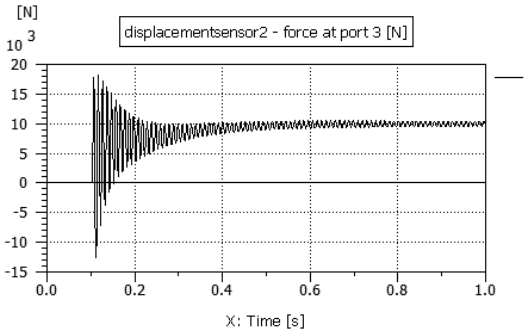


Figure 23: Force exerted by the actuator rod.

The same phenomenon occurs with the force developed by the damper rod (Figure 24).

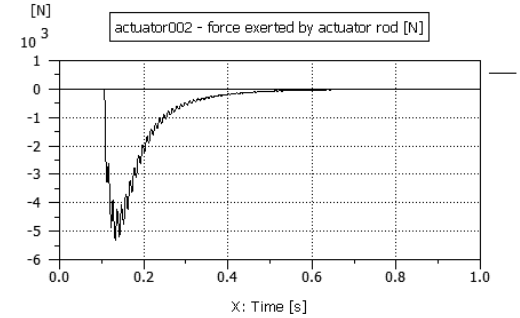


Figure 24: Force exerted by the damper rod.

However, the force at the output of the servomechanism rises continuously to the final value of about 10,000 N (Figure 25).

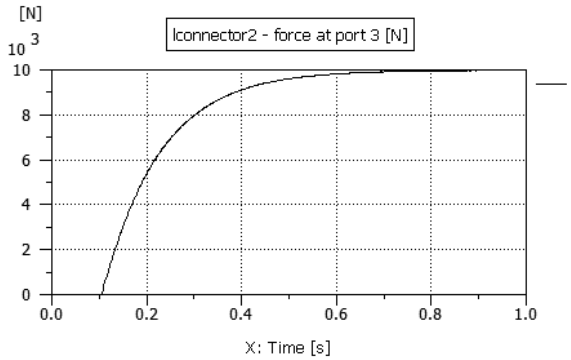


Figure 25: Force developed by the whole system at the output rod.

SYSTEM FREQUENCY RESPONSE

The integration of the servomechanism in the airplane flight control system needs the study of its frequency response. This can be easily obtained by the transferometer block from the control library of AMESim (Figure 26). The small value (1.2 Hz) of the cut frequency (Figure 27) proves that the servomechanism behaves as low pass filter for all type of input signals and for any external disturbance. This response quality is suitable for large airplanes [4].

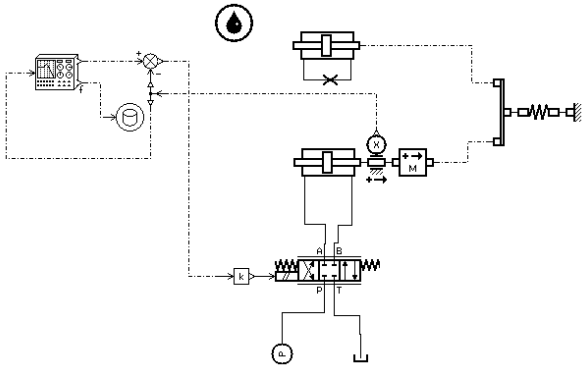


Figure 26: Simulation network of the frequency response.

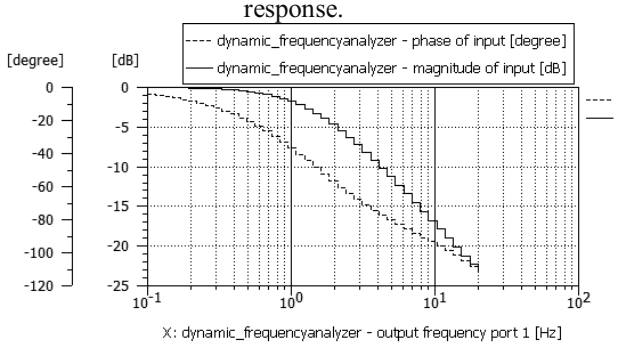


Figure 27: System frequency response.

The diameter of the damper orifice plays an important role in the dynamics of the studied system. Three common values of the orifice diameter d_o, were studied: 0.5, 0.75 and 1.25 mm. The influence of the orifice size is important for diameters smaller than 0.75 mm (Figure 28).

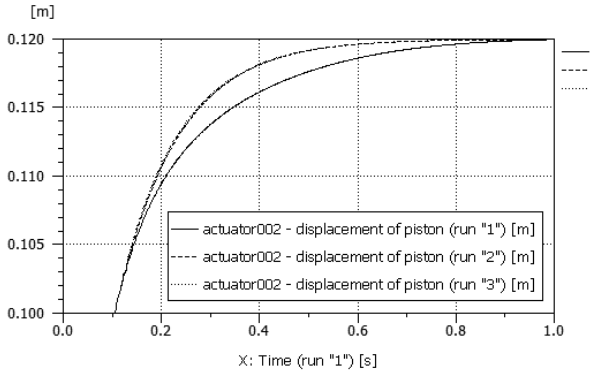


Figure 28: The influence of the damping orifice diameter on a small step input servomechanism response : the lower line corresponds to d_o=0.5mm.

CONCLUSIONS

1. The use of a hydraulic damper in paralel with the hydraulic cylinder improves the dynamics of the servosystem *without introducing a steady-state error*. The damping effect can be controlled from outside the damper by the opening of the by-pass orifice [3], [4].
This "paralel damper" solution is used now in different similar applications due to the good dynamics, and the small increase of the total price of the servomechanism. Different kind of control alghoritms are now developped in order to improve the dynamics [5].

2. All the design, test, and identification author's activity in the field of the fluid control systems pointed out that AMESim provided a strong solver and numerical core for transient simulation. As modeling a complex multi-physics system is not the main objective of engineers, it is important to have tools and interfaces which accelerate and optimize the design. From this point of view, AMESim is a complete software perfectly adapted for model creation and deployment. The wide field of application, including electric powertrain and mobile hydraulics is continuously extended by a big number of many category of users [7].

ACKNOWLEDGEMENTS

The authors are grateful for all the technical support received from the LMS COMPANY (now - a SIEMENS PLM SOFTWARE business) in different manners: free licenses, free technical training for the research team members, and many other facilities. The research stages offered by the company to young researchers generated new and valuable ideas included in many innovative Ph.D. theses [8] … [29].

REFERENCES

[1] Xavier Le tron. *A380 Flight Controls Overview.* Presentation at Hamburg University of Applied Sciences, 2007.
[2] https://www.ac-paris.fr/serail/jcms/s1_942790/ee2-lg-clg-paul-valery-portail
[3] Merritt H.E. *Hydraulic Control Systems*. John Wiley and Sons Inc., New York, London, Sydney, 1967.
[4] Guillon M. *L'asservissement hydraulique et électrohydraulique*. Dunod, Paris, 1972.
[5] Ursu I., Tecuceanu G., Toader A. and Calinoiu C. *Switching Neuro-fuzzy Control with Antisaturating Logic. Experimental Results for Hydrostatic Servoactuators.* Proccedings of the Romanian Academy, Series A, Volume 12, Number 3/2011, pp 231-238, ISSN 1454-9069.
[6] *** LMS INTERNATIONAL. *Advanced Modeling and Simulation Environment*, Release 12 User Manual, Leuven, 2012.
[7] Mare J.Ch. *AMESim for education and research in Aerospace Actuation.* LMS UK Engineering Conference, Grove, Oxfordshire, 2012.

[8] Vasiliu D. *Researches on the dynamics of the servopumps and servomotors of the automotive hydrostatic transmissions.* Ph.D. Thesis, University POLITEHNICA of Bucharest, 1997.
[9] Călinoiu C. *Contributions to the experimental identification of the aerospace fluid power systems*. Ph.D. Thesis, University POLITEHNICA of Bucharest, 1998.
[10] Muraru M.M. *Researches on the parametric synthesis of the fluid power systems.* Ph.D. Thesis, University POLITEHNICA of Bucharest, 2002.
[11] Ofrim D.V. *Researches on the digital control of the electrohydraulic servomechanisms.* Ph.D. Thesis, University POLITEHNICA of Bucharest, 2004.
[12] Vasile L.N. *Researches on the automotive hydraulic steering systems dynamics.* Ph.D. Thesis, University POLITEHNICA of Bucharest, 2005.
[13] Cazanacli Cr. *Researches on the predictive maintenance of hydro mechanical equipment from hydropower stations.* Ph.D. Thesis, University POLITEHNICA of Bucharest, 2006.
[14] Ion Guț ă D.D. *Real time simulation of the fluid power systems.* Ph.D. Thesis, University POLITEHNICA of Bucharest, 2007.
[15] Popescu T.C. *Researches on the synthesis of the fluid power systems for moving earth machines.* Ph.D. Thesis, University POLITEHNICA of Bucharest, 2008.
[16] Bățăuș M.V. *Real Time Simulation of the automatic automotive transmissions.* Ph.D. Thesis, University POLITEHNICA of Bucharest, 2010.
[17] Negoiță G.Cl. *Researches on the dynamics of the hydrostatic transmissions.* Ph.D. Thesis, University POLITEHNICA of Bucharest, 2011.
[18] Cioranu N.C. *Researches on redundant electro hydraulic control systems.* Ph.D. Thesis, University POLITEHNICA of Bucharest, 2012.
[19] Dragne F.D. *Modeling and numerical simulation of the automotive hydraulic brakes control systems.* Ph.D. Thesis, University POLITEHNICA of Bucharest, 2012.
[20] Mitroi M.A. *Researches on the modeling and real time simulation of the electrohydraulic control systems.* Ph.D. Thesis, University POLITEHNICA of Bucharest, 2013.
[21] Mihalescu B. *Researches on the experimental identification of the electrohydraulic servo valves.* Ph.D. Thesis, University POLITEHNICA of Bucharest, 2014.
[22] Puhalschi R.C. *Modeling and the real time simulation of the energy control systems.* Ph.D. Thesis, University POLITEHNICA of Bucharest, 2014.
[23] Fehér S.I. *Researches on electrohydraulic variable valve timing.* Ph.D. Thesis, University POLITEHNICA of Bucharest, 2014.
[24] Ganziuc Al. *Electrohydraulic servomechanisms with high immunity level,* 2014.
[25] Pîrăianu V.F. *Contributions to the simulation of the integrated water resources management.* Ph.D. Thesis, University POLITEHNICA of Bucharest, 2014.
[26] Dobre Al. *Mathematical modelling, numerical simulation and experimental identification of the automotive magnetorheological suspensions.* Ph.D. Thesis, University POLITEHNICA of Bucharest, 2015.
[27] Rădoi P. Fl. *Modeling, simulation, and experimental identification of the auxiliary automotive heating systems.* Ph.D. Thesis, University POLITEHNICA of Bucharest, 2015.
[28] Irimia P.C. *Researches on the power management of the off-road vehicles.* Ph.D. Thesis, University POLITEHNICA of Bucharest, 2015.
[29] Costin I.I. *Researches on the three stages electrohydraulic high flow servovalves for high power hydropower plants.* Ph.D. Thesis, University POLITEHNICA of Bucharest, 2016.

WEB REFERENCES

http://www.boschrexroth.com/
http://www.eaton.com/
http://www.sauer-danfoss.com/
http://www.moog.com/
http://www.hydac.com/
http://www.parker.com
http://www.famictech.com/
http://www.lmsintl.com/
http://www.mathworks.com/products/simulink/
http://www.acslX.com
http://www.nhancetech.com/
http://www.ni.com/matrixx/
http://www.dspace.com/
http://www.opal-rt.com/
http://www.adwin.de/
http://www.analog.com/
http://www.fluidpower.net/

BIOGRAPHIES

Nicolae VASILIU graduated in Hydropower Engineering from University POLITEHNICA of Bucharest in 1969. He became a Ph.D. in Fluid Mechanics after a research stages in Ghent State University and Von Karman Institute from Brussels. He became state professor in 1994, leading the ENERGY & ENVIRONMENT RESEARCH CENTRE from the University POLITEHNICA of Bucharest. He managed five years the Innovation Romanian Agency. He worked always for the industry, as project manager or scientific advisor, promoting the numerical simulation as an engineering tool.

Daniela VASILIU graduated in mechanical engineering in 1981 and prepared a Ph.D. thesis in the field of the dynamics of the electrohydraulic servopumps and servomotors for hydrostatic transmissions. She is currently professor in the Department of Hydraulics, Hydraulic Machines and Environmental, head of Fluid Power Laboratory of the University POLITEHNICA of Bucharest. She works in the field of modeling, simulation, and experimental identification of the electro hydraulic control systems; member of EUROSIS, FPNI, SIA, FLUIDAS etc.

Constantin CĂLINOIU graduated in Power Engineering from University POLITEHNICA of Bucharest in 1975, and in Mathematics at the Bucharest University in 1981.

After his studies he become scientific researcher in the Hydraulics Laboratory of the Romanian Aerospace Institute. In 1998 he defended his Ph.D. thesis, and became associated professor in the Fluid Power Laboratory from U.P.B. He is working mainly in modeling, simulation, and identification of the hydraulic and electrohydraulic control systems.

Radu PUHALSCHI graduated in Applied Computer Science from University POLITEHNICA of Bucharest in 2009, and a MD in Advanced Hydraulic and Pneumatic Systems in 2011. After a stage of web designer at HP Germany, he performed a Ph.D. thesis on Real-Time Simulation of hydraulic systems at the Fluid Power Laboratory from the Power Engineering Department of the University POLITEHNICA of Bucharest. Now, He is still contributing to control courses as associated professor in the same university, and he is working as control engineer in the Romanian Division of Honeywell International Inc.

Petru-Cristinel IRIMIA received the M.Sc. degree in Mechanical Engineering, Automotive Section, from the University Transilvania of Brasov in 1991. He developed a long carrer in R&D, having a deep education in finite element analysis with high level LMS software. He prepared a Ph.D. thesis within the University POLITENICA of Bucharest in the field of automotive electrohydraulic remote steering system energy management. He is serving currently as general manager of SIEMENS INDUSTRIAL SOFTWARE SRL from Romania.

A FUZZY VECTORIAL SPACE THAT AVOIDS TO DEFUZZIFY THE MEMBERSHIP FUNCTIONS

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

KEYWORDS
Fuzzy Vectorial Space, Aggregation Operator, Analogy, Time modeling, Emotion modeling, Kinematics

Abstract

In this paper, we show that the logic theory and the object oriented approach are equivalent to describe world objects. We propose a new aggregation operator to combine fuzzy membership functions without fuzzy rules and without defuzzification. This operator preserves the improvements and the features of the initial fuzzy set theory. Instead of fuzzy rules, the fuzzy membership functions can be combined in a chosen order or in any order with the aggregation operator that provides a new resultant membership fuzzy function affected by all the component functions. This approach allows to globally compare the properties of complex objects by comparing their components. In our approach the fuzzy membership functions are represented as forces and thus by vectors in a vector space. We illustrate the use of our model in building a fuzzy emotional vector space based on the Ortony Clore Collins model. Thus, we compare the advantages and drawbacks of our approach with existing works and we conclude by presenting the perspectives of this research.

Introduction

Since the fuzzy set theory was introduced by Lofti Zadeh, a lot of models was proposed to offer many useful applications. Most of the time, the models relies on a logic approach based on a generalised modus ponens using fuzzy rules Montiel et al. (2009), Lorestani et al. (2006), El-Nasr et al. (2000). These fuzzy rules combine *membership functions* and *linguistics variables* to describe subsets of characteristics values used to express knowledge chunks stored in knowledge bases. The fuzzy set theory constituted a great enhancement and offered a better flexibility to take into account the uncertainty and the approximation of necessary variable features when one wants to build knowledge bases. The main drawback of fuzzy rules comes from the boolean operating nature of inference engines that combine rules by firing them or not. To enhance this process, a lot of accurate defuzzification methods was proposed to allow fuzzy rules to be triggered by inference engines, for example Opricovic and Tzeng (2003), Martin and Klir (2007), Iancu (2012). This effort is vain. Indeed, all the fuzzy logic improvement on providing fuzzy membership functions is almost annihilated by the fact that rules are firing or not, through a boolean process by the inference engine. Besides the adequate implementation of fuzzy membership functions requires to multiply the number of rules. Moreover, another conceptual drawback is that inference rules implement a succession of deductions corresponding to a list of object characteristics rather than comparing then globally. According to cognitive psychology, the most used way of reasoning is the analogy rather than the deduction. In this paper, we propose a new aggregation operator to combine fuzzy membership functions without fuzzy rules and without defuzzification in an object oriented model. This operator preserves the improvements and the features of the initial fuzzy set theory. Instead of fuzzy rules, the fuzzy membership functions can be combined in a chosen order or in any order with the aggregation operator that provides a new resultant membership fuzzy function affected by all the component functions. This approach allows to globally compare the properties of complex objects by comparing their components. In our approach the fuzzy membership functions are represented as forces and thus by vectors in a vector space. We illustrate the use of our model in building a fuzzy emotional vector space through an example of knowledge modelling of a disease evolution in medicine. Thus, we compare the advantages and drawbacks of our approach with existing works and we conclude by presenting the perspectives of this research.

The fuzzy set theory and fuzzy membership functions

The fuzzy set theory

The knowledge and data which we own either concerning a problem or a situation in the real world is almost always imperfect: they are uncertain when we doubt

their validity; they are indistinct when it is difficult to obtain a measure or a clear explanation, they are incomplete because most of the time all the data are not (yet, or never) available. Moreover, the world is evolving and what is true at one moment could become false at the next moment. Unfortunately the indistinctness, the uncertainty, the incompleteness are inherent to our environment; we can only notice and try to supply as better models as possible. During the examination of a characteristic or a property of an object, the uncertainty is conversely proportional in the indistinctness. Indeed, more we wish to measure or to establish a property with a high degree of precision more the uncertainty: the probability to make an error of measure (to be outside the interval of tolerance) become close to 1. In order to facilitate the modeling of uncertain and indistinct knowledge, Lofti A. Zadeh of the University of Berkeley proposed, in 1965, the theory of the fuzzy subsets Zadeh (1965). The fuzzy logic can be considered as an extension of the boolean logic. This theory is defined as a generalization of the set theory. It allows to quantify the degree of membership of an object of the real world in a set named "fuzzy set" Kaufmann (1973), Kaufmann and Pichat (1977). The necessity of reasoning using indistinct and uncertain facts and to combine together such knowledge led Lofti A. Zadeh to propose his theory of the possibilities in 1978. The traditional fields of application of the fuzzy logic are all the human sciences, the medicine, the management, the economy... The proposed types of application are the decision-making support, the pattern recognition, the automatic classification of objects, knowledge bases and expert systems... In the medical domain, the fuzzy logic is especially used to aid the diagnosis of diseases. The diagnosis relies on the disease classifications and algorithms that computes values of fuzzy membership functions Bartolin et al. (1982). Unfortunately, the implementation of fuzzy membership functions to provide efficient decision support systems is not easy. In the following subsection, we briefly recall the general principles of the classic logic of Lewis Caroll as a basis of the fuzzy set theory. The study of this seminal work allows to build a strong link between semantic and object oriented models and the symbolic logic.

Lewis Carroll's Symbolic Logic

According to the theory of Charles Ludwitge Dodgson, Lewis Carroll, in "Symbolic Logic, part I" Carroll (1896): *"The Universe contains 'Things.', [For example, "I," "London," "roses," "redness," "old English books," "the letter which I received yesterday."]. Things have 'Attributes.' [For example, "large," "red," "old," "which I received yesterday."]. One Thing may have many Attributes; and one Attribute may belong to many Things. In chapter IV Lewis Carroll highlighted the importance to represent things by names: The word "Thing", which conveys the idea of a Thing, without any idea of an Adjunct, represents any single Thing. Any other word (or phrase), which conveys the idea of a Thing, with the idea of an Adjunct represents any Thing which possesses that Adjunct; i.e., it represents any Member of the Class to which that Adjunct is peculiar. Such a word (or phrase) is called a 'Name'; and, if there be an existing Thing which it represents, it is said to be a Name of that Thing."*
It is astonishing to see that the Lewis Carroll's symbolic logic is based on similar concepts that those that are used to define all semantic models Quillian (1968), entity-relationship models) Chen and Pin-Shan (1976) and furthermore current object oriented paradigm with the actual concept of class as in UML Booch et al. (1998).

Traditional sets and boolean membership functions

For any set A in a universe X, $A \subset X$ corresponding to objects $x \in X$ who belong or not to the set A because of owning a specific characteristic: the Adjunct peculiar of that so-formed class A according to Lewis Carroll logic theory. A membership function is defined : $\mu_{bool}A(x)$ where $\mu_{bool}A(x) = 1\ if\ x \in A\ and\ \mu_{bool}A(x) = 0\ otherwise$. Let us take a simple famous example: x represents the size in centimeter of men at the age of 18 years old and A to be the set of "tall men" in the set M of Men of 18 years old and $A \subset M \subset X$. The boolean membership function $\mu_{bool}A(x)$ is defined using a threshold for example $x \geq 180cm$ as $\mu_{bool}(x) = 1\ if(x \geq 180)\ and\ \mu_{bool}(x) = 0\ if(x < 180)$. The main inconvenient is that men whose size is 179cm are considered to be not tall and those whose size is 180cm are considered to be tall. Obviously, this is an unsatisfactory way to express the reality. This dissatisfaction led L.A. Zadeh to propose his fuzzy set theory Zadeh (1965) that allow to tune the degree of truth that an object belongs or not to a fuzzy set $\tilde{A}$ corresponding to a property (the value of a peculiar attribute x: size) of an object of the universe X. In the traditional set theory the operators acting on sets are the intersection $\cap$, the union $\cup$, the setminus $\setminus$ and the complement $\complement$.

Fuzzy sets and fuzzy membership functions

According to Zadeh's theory, a fuzzy membership function $\mu_{\tilde{A}}(x) : X \rightarrow [0, 1]$ is defined as a value included in the interval $[0, 1]$ indicating the membership degree for the element x to belong to a fuzzy set named $\tilde{A}$. A fuzzy set, defined as $\tilde{A}$, is a subset of a universe of discourse X, where $\tilde{A}$ is characterized by a membership function $\mu_{\tilde{A}}(x)$. The membership function allows to associate the grade of membership $\mu_{\tilde{A}}(x) \in [0, 1]$ in the fuzzy set $\tilde{A}$ to each element of the universe of dis-

course X. $\tilde{A} = \forall x \in X, (x, \mu_{\tilde{A}}(x))$. We can adopt the following notation $A = \sum_{x \in X} \mu_{\tilde{A}}(x)$ if X is finite and $A = \int_{x \in X} \mu_{\tilde{A}}(x)$ if X is infinite. The value $\mu_{\tilde{A}}(x)$ is fixed by experts of the domain. This can be more or less difficult according to the attribute type. In the previous example, if it is assumed that a man whose size is less than 150 cm is very small, less than 160 cm is small, less than 180 cm is medium, at least 180 cm is tall, and more than 190 cm is very tall. Thus these intervals of values of $[0, 150[$, $[150, 160[$, $[160, 180[$, $[180, 190]$, $[190, \infty]$ respectively correspond to subsets of X: "very small","small","medium","tall","very tall" represented figure 3.

$$S(x; \alpha, \beta, \gamma) = \begin{cases} 0 & \text{for } x \leq \alpha, \\ 2(\frac{x-\alpha}{\gamma-\alpha})^2 & \text{for } \alpha < x \leq \beta, \\ 1 - 2(\frac{x-\alpha}{\gamma-\alpha})^2 & \text{for } \beta < x \leq \gamma, \\ 1 & x > \gamma. \end{cases} \quad (1)$$

The S-function figure 1 represents the membership function "to be tall".
$\mu_{\tilde{M}}(x)$ of the fuzzy subset Medium denoted $\tilde{M}$ can be represented by a π-function represented by a bell curve figure 2,

$$P(x; \beta, \gamma) = \begin{cases} S(x; \gamma - \beta, \gamma - \frac{\beta}{2}) & \text{for } x \leq \gamma, \\ 1 - S(x; \gamma, \gamma + \frac{\beta}{2}, \gamma + \beta) & \text{for } x \geq \gamma. \end{cases} \quad (2)$$

All subsets very small (VS), Small(S) ,Medium (M), Tall (T), Very Tall (VT) can be represented in the same manner by membership functions as depicted in the three following figures.
Zadeh defined the support of a fuzzy set $\tilde{A}, supp(\tilde{A}) = \{x \in X / \mu_{\tilde{A}}(x) > 0\}$. The kernel of a fuzzy set $\tilde{A}, kernel(\tilde{A}) = \{x \in X / \mu_{\tilde{A}}(x) = 1\}$ In his fuzzy set theory L. Zadeh proposed intersection and union operators and the corresponding membership functions. Let be two fuzzy subsets $\tilde{A}$ and $\tilde{B}$ in the universe of discourse X. The membership function $\mu_{\tilde{A} \cap \tilde{B}}(x) = \{min(\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)), \forall x \in X\}$. The membership function $\mu_{\tilde{A} \cup \tilde{B}}(x) = \{max(\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)), \forall x \in X\}$.

boolean logic propositions

Boolean logic propositions express relations between facts which are true or false. In the chapter III of symbolic logic Lewis Carroll introduced different kind of propositions of relations Carroll (1896). *"A Proposition of Relation, of the kind to be here discussed, has, for its Terms, two Species of the same Genus, such that each of the two Names conveys the idea of some Attribute not conveyed by the other. For example, the proposition "Some merchants are misers" is of the right kind, since "merchants" and "misers" are Species of the same Genus "men"; and since the Name "merchants" conveys the idea of the Attribute "mercantile", and the name*

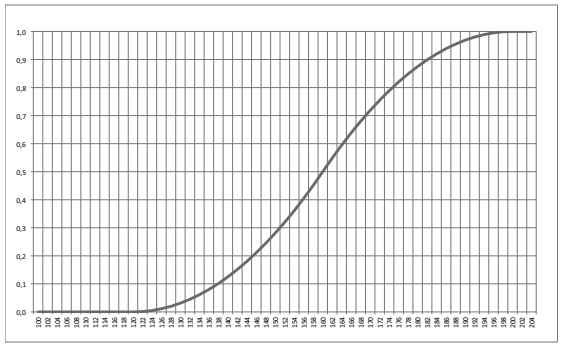


Figure 1: Fuzzy membership S-functions for the set "Tall men"

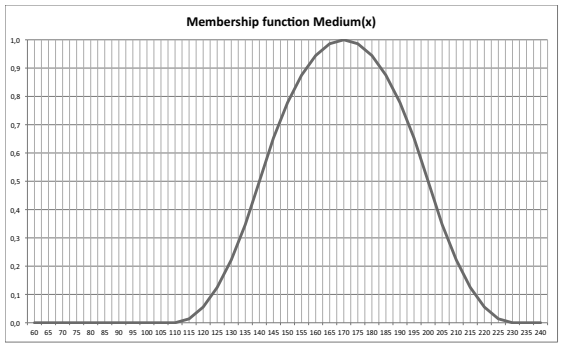


Figure 2: Fuzzy subset membership function "Medium men" (M)

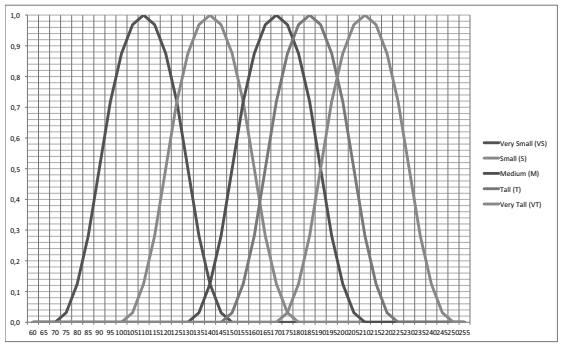


Figure 3: Fuzzy membership functions small (VS), Small(S), Medium(M), Tall(T), Very Tall(VT)

"misers" the idea of the Attribute "miserly", each of which ideas is not conveyed by the other Name. In a relation, the sign of quantity is "some", "no" or "all". A proposition of relation, beginning with "Some", is henceforward to be understood as asserting that there are some existing things, which, being members of the subject, are also members of the predicate. A proposition of relation, beginning with "No", is henceforward to be understood as asserting that there are no existing things which, being members of the subject, are also members of the predicate; a proposition of relation, beginning with "All", is equivalent to two propositions, one beginning with "Some" and the other with "No", each of which we now know how to translate. For example : "All bankers are rich men." This is equivalent to the two Propositions "Some bankers are rich men" and "No bankers are poor men." Lewis Carroll's symbolic logic is bounded to crisp sets. Here is the representation using membership functions. Let be the universe of discourse the set Men (M), the subsets of M: Bankers (B), Rich men (R), Poor men(P). The first proposition "Some bankers are rich men is represented by $\{B \subseteq M, R \subseteq M, \exists x \in B \Rightarrow x \in R\}$ or by $\{B \subseteq M, R \subseteq M, B \cap R \neq \emptyset\}$. Using the membership functions $\{\mu_{B\cap R} > 0\}$. The second proposition "No bankers are poor men" is expressed by $\{B \subseteq M, P \subseteq M, \forall x \in M, x \in B \Rightarrow x \notin P\}$ or by $\{B \subseteq M, P \subseteq M, B \cap P = \emptyset\}$. Using the membership functions $\{\mu_{B\cap P} = 0\}$.

Modus ponens

The modus ponens is a mean of representing knowledge in a natural way. x is a man, if x is a banker then x is not poor, that can be represented by the previous expression. A rule has the following structure: the proposition A is true, $A \Rightarrow B$, B is true. The knowledge is represented by rules that express facts and heuristics of a specific domain. Many expert systems rely on a base of rules. An important drawback of such systems is that the deduction is not the main reasoning mode used by experts. Furthermore, the facts are not generally boolean (true or false). It is the main reason why L Zadeh proposed fuzzy inference rules.

Fuzzy logic predicates

The same reasoning mode can be used with fuzzy sets and corresponding fuzzy membership functions. Let be $\tilde{M}$ the universe of discourse and the fuzzy sets $\tilde{B}$ to fuzzy set of Bankers, $\tilde{R}$ the fuzzy set of wealth, the corresponding membership function $\mu_{\tilde{B}}$ and $\mu_{\tilde{R}}$, $\mu_{\tilde{B}\cap\tilde{R}}(x) = min(\mu_{\tilde{B}}(x), \mu_{\tilde{R}}(x)) > 0, \forall x \in M\}$ describe the same proposition that the previous crisp membership function.

Generalized modus ponens

The generalized modus ponens is the fuzzy extension of the modus ponens. It was introduced by L. Zadeh to represent an approximate reasoning mode. The generalized modus ponens should allow to handle in the same context symbolic knowledge and numeric data while taking into account the gradual aspect of fuzzy characterization. if the datum value is very close to the rule premise, the conclusion will be very close to the rule conclusion. For example: let us suppose the fact to be a banker is represented by the seniority (in years) in this profession. Let us suppose that the wealth is expressed as being the estimation of this person's possessions (in any currency). The seniority is a variable x and the possession of the person is represented by the variable y. More his seniority x is important, more the person is considered as a banker (B) experienced and competent; higher the person's wealth is (y), more the person is considered to be rich (R). Let the set of Men M to be the universe of discourse. This proposition can be represented by the following fuzzy rule: $X \in M$, if X is B then X is R. The membership function $\mu_{\tilde{B}}$ corresponding to the fuzzy set $\tilde{B}$ (Bankers) which is the premise part of the rule, the membership function $\mu_{\tilde{R}}$ corresponding to the fuzzy set $\tilde{R}$ (Rich people) which is the conclusion part of the rule.

Fuzzy implications

There is not a unique way to generalize the implication from the boolean logic. Many different fuzzy implications were proposed by many authors : Reichenbach, Wilmott, Rescher-Gaines, Kleene-Dienes, Brouwer-Gödel, Goguen, Lukasiewicz, Mamdani Mamdani and Assilian (1975), Larsen. For example the fuzzy implication from Lukasiewicz is represented by $\mu_{\tilde{B}\Rightarrow\tilde{R}}(x,y) = min(1, 1 - \mu_{\tilde{B}}(x) + \mu_{\tilde{R}}(y))$. The membership function to $\tilde{R}$ is computed combining $\mu_{\tilde{B}}$ and $\mu_{\tilde{B}\Rightarrow\tilde{R}}$. In this specific case The universe of discourse X=Y=M $\forall y \in Y \mu_{\tilde{R}}(y) = sup_{x\in X} T(\mu_{\tilde{B}}, \mu_{\tilde{B}\Rightarrow\tilde{R}}(x,y))$. T is a t-norm used with the fuzzy implication. For example, the $T_{Lukasiewicz}(u,v) = max(u + v - 1, 0)$. Thus, $\forall y \in Y \mu_{\tilde{R}}(y) = sup_{x\in X} max(0, \mu_{\tilde{B}} + \mu_{\tilde{B}\Rightarrow\tilde{R}}(x,y) - 1)$ Bouchon-Meunier (1993),Bouchon-Meunier and Marsala (2003).

Structure of a fuzzy logic controller

An Fuzzy Logic Controller consists of a set of rules of the form previously presented. More generally, IF (a set of conditions are satisfied) THEN (a set of consequences can be inferred). A fuzzy knowledge base consists of a set of IF-THEN rules associated with fuzzy conditional statements. This type of systems are considered to be an extension of rule based expert systems like Mycin

(Buchanan and Shortliffe, 1984) Buchanan and Shortliffe (1984). In fuzzy rule-based systems, the inputs should be given by fuzzy sets, and therefore, we have to fuzzify the crisp inputs. Furthermore, the output of a fuzzy system is always a fuzzy set, and therefore to get crisp value we have to defuzzify it. Fuzzy logic control systems usually consist of four major parts: Fuzzification interface, Fuzzy rule base, Fuzzy inference engine and Defuzzification interface Iancu (2012). The whole process is presented in the figure 4 Lee (1990).

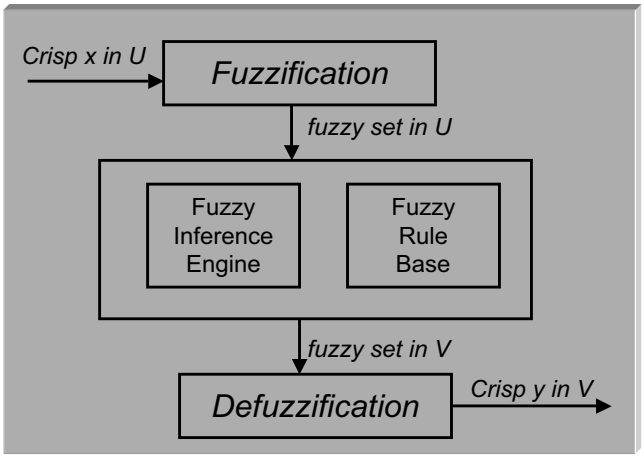


Figure 4: Fuzzy logic controller proposed by Lee 1990 Lee (1990)

Criticism of the fuzzy logic inference

The incompleteness is an important drawback. The reversibility of rules is often problematic: rules are formed by two members linked by a relation of dependence: if A is true B is also true and vice versa. When rules are not exhaustive (What is mostly the case), it is not possible to use the modus tollens. Morevover, the use of a monotonous logic does not allow to manage the exceptions. The production rules can be reversed only if we are sure to have forgotten no factor or attribute (What is impossible to be sure). This argument had beforehand been raised by John McCarthy McCarthy (1987). In fact, the inference engine can trigger contradictory rules. The partial formulation of the knowledge often leads to contradictions which could be difficult to detect. Thus, the addition of new rules in a knowledge base can create conflicts with pre-existent rules leading to inconsistency. The use of a monotonous logic does not allow to manage the exceptions McCarthy (1987). According to ordinary logic : All animals ($x \in A$) which are birds ($x \in B$) are animals able to fly ($x \in F$) is expressed by $\forall x \in A$, if $x \in B \Rightarrow x \in F$. Unfortunately, the emperor penguins P are birds and they does not fly what requires to introduce new clauses as ($x \notin P$) to take into account exceptions $\forall x \in A$, if $x \in B \; and \; x \notin P \Rightarrow x \in F$. According to McCarthy, this inconvenience is going to cause an important proliferation of clauses. Moreover, the use of fuzzy logic still increases the number of rules corresponding to the fuzzy subsets in order to correctly implement only one object property (For example the size of the men as presented figure 3). The successive inferences of the multiple rules produce and propagate an accumulation of the uncertainty concerning the obtained conclusions. Whatever is the type of production rules (crispy or fuzzy), the trigger mechanism of the inference engine is of boolean nature: each rule is activated or not. This inherent characteristic of inference engine makes lose all the profit brought by the use of the fuzzy set theory. It is the exploitation of rules and not the fuzzy logic which raises problem.

Advantages of production rules

The production rules are well suited to the deductive mode of reasoning. The independence of the knowledge with regard to the inference engine allows to separate the knowledge and the deduction. The declarative character of rules and facts is an asset. However, rules must be independent between them and a rule does not have to make reference directly to an other one. However, every rule represents only a tiny part of knowledge and the navigations in contexts or different domains necessary for most of actual problem solving make finally difficult the management and the maintenance of rules. The inference engines have the capacity to explain the made deductions from the tree of rules that were triggered to get the result. The conviviality of the rule based systems is insured by interfaces capable of justifying at any time the deductions made by the system. The existence of specific programming languages as Prolog Colmerauer et al. (1973), Colmerauer (1996) or Objective Caml Leroy and Weis (1993) is also an advantage. However, we doubt that the deduction is the most frequent and the best suited way of reasoning to elaborate knowledge bases and to implement problem solving systems. Human beings naturally prefer to use analogy rather than deduction to deal with their usual knowledge. The brain is memory and is able to do dynamically a pattern matching with the objects in the environment. According to Marvin Minsky, at any moment, our brain is able to compare the uncountable information that he has collected about objects and situations in our environment by means of our senses with the concepts that we have previously elaborated and memorized during our existence Minsky (1975),Minsky (1985). Our main research topic was to find new ways of implementing reasoning modes like the induction, the abduction and the analogy coupling fuzzy logic and object oriented models. We show in the next section that rules are not well suited to these types of reasoning.

The inconveniences of rule based systems

In the diagnosis applications, the cases which are submitted to the system are represented by comparing characteristics of complex objects with those of known clinical pictures. When using a rule based approach, many characteristics or facts is going to be thus considered successively to classify the case among those known in the knowledge base by means of a flow of triggered rules that generate in turn new facts hypotheses. In summary the classification of clinical pictures is gradually implemented by a continuation of deductions taking into account successively and not simultaneously the characteristics or the facts of the presented case.

Analogy and rule based systems

The search for an analogy between two objects and situations often requires to compare an important number of characteristics and thus a long series of deductions or inferences with several major drawbacks. - The analogy is established on a resemblance between inevitably different and unique objects or situations. It is about a resemblance between two objects and not about a perfect correspondence. Certain characteristics of these objects can be only similar, equivalent or even absent while a majority of identical or important characteristics are enough to establish the analogy. -The analogy is a parallel way of reasoning. The characteristics of an object are considered simultaneously and not successively. - The analogy is organized into a hierarchy Boulanger and Colloc (1992). All the characteristics of an object have not the same importance according to contexts, a small part of the object can have an essential meaning, for example: the keyhole of a door or the receiving site of an enzyme in a cell while other characteristics will be considered as accessories as the color of the door for a locksmith. In the structural analogy, it will be important to compare characteristics of component objects that are more important than others. Component objects have to be ranked according to their granularity and their importance in the structure of the composite object. However, some other component or characteristic objects appear to be less interesting to establish a diagnosis but could play, all together, a complementary discriminating role. -The detection of analogies by means of production rules requests excessively the inference engine. -Furthermore, if objects or situations to be compared have a fast evolution, some of their characteristics will be inappropriately taken into account in different times.

Time and rule based systems

The consideration of time raises problem in the models with production rules. Indeed, the release from initial facts supposes that they remain stable in the time, what is not the case in numerous actual domains. Undo backwards a deductive reasoning when a true fact become false is a very tricky problem. This aspect takes all its importance when the system has to manage dynamic phenomena with fast fluctuations of attribute values. The following section describes our main contribution: a fuzzy vector space to describe the state and the evolution of dynamic complex objects.

A fuzzy vector space

In this section we propose an extension of the Zadeh's membership functions which are defined on [0,1]. Our membership functions are defined on [-1,1], taking into account a full specific characteristic of an object (value=1) and its contrary (value=-1), the value 0 is considered to be neutral. The set $\mathbb{F}$ contains all the membership functions. Thus $\forall f \in \mathbb{F}\ \exists f' = -f$ defined in $[-1,1]$ that represents the opposite function of the f function. Let be three functions of $\mathbb{F}f : [-1,1] \rightarrow [-1,1], x \rightarrow f(x), g : [-1,1] \rightarrow [-1,1], y \rightarrow g(y), h : [-1,1] \rightarrow [-1,1], z \rightarrow h(z)$. We propose a fuzzy vector space among which vectors $\vec{f}, \vec{g}, \vec{h}$ express forces (as a physical metaphor) having as *mode* the "intensity" based on the value of the fuzzy membership functions $f(x), g(y), h(z)$ related to the values x, y, z corresponding with specific properties (attribute, or characteristics) of an object of the universe of the discourse. $\mathbb{R}$ is a commutative corpus. Let be a vector space E on $\mathbb{R}^3$ associated to an orthonormal coordinate system $(O, \vec{i}, \vec{j}, \vec{k})$ from the set of membership functions $\mathbb{F}$. The vector $\vec{f}$ overlaps the vector $\vec{i}$ (X axis), $\vec{g}$ the vector $\vec{j}$ (Y axis) and $\vec{h}$ the vector $\vec{k}$ (Z axis). The norm of the vector $\vec{f}$ is given by $\|\vec{f}\| = \sqrt{f(x)^2} = f(x)$ and in the same way $\|\vec{g}\| = \sqrt{g(y)^2} = g(y)$ and $\vec{h}$ est $\|\vec{h}\| = \sqrt{h(z)^2} = g(z)$. Because the system is orthonormal the dot product (or scalar product) of vectors $\vec{f}$, $\vec{g}$ and $\vec{h}$ is nil. Consequently, we can combine three forces represented by vectors $\vec{f}$,$\vec{g}$ and $\vec{h}$ with an inner additive operator named $+$. $\forall \vec{f}, \vec{g}, \vec{h} \in E$, we verify the following properties :

p1: $\vec{u} = (\vec{f} + \vec{g}) + \vec{h} = \vec{f} + (\vec{g} + \vec{h})$ associativity.

p2: $\vec{0} + \vec{f} = \vec{f} + \vec{0}$ the neutral element is the vector $\vec{0}$.

p3: $\forall \vec{f} \in E, \exists -\vec{f} / \vec{f} + -\vec{f} = \vec{0}$ the opposite vector.

p4: $\vec{u} = \vec{f} + \vec{g} = \vec{g} + \vec{f}$. Commutativity of $+$ in E.

p5: $\forall \vec{f} \in E, \exists \lambda, \mu \in \mathbb{R} / \lambda(\mu \vec{f}) = (\lambda\mu)\vec{f}$, λ and μ are scalars

p6: $\forall \vec{f} \in E, \exists e = 1 / e\vec{f} = \vec{f}e = \vec{f}$ Neutral element.

p7: $\forall \vec{f} \in E, \forall \lambda, \mu \in \mathbb{R}, (\lambda\mu)\vec{f} = \lambda\vec{f} + \mu\vec{f}$, the result is generally outside of $[-1,1]$.

p8: $\forall \vec{f}\vec{g} \in E, \forall \lambda \in \mathbb{R}, \lambda(\vec{f} + \vec{g}) = \lambda\vec{f} + \lambda\vec{g}$, the result is generally outside of $[-1,1]$. E is a vector space on $\mathbb{R}^3$.

Let be $\mathbb{R}$ the commutative corpus provided with the absolute value and E the $\mathbb{R}$-vector space previously defined, a norm on E is an application N on E with positive real numbers satisfying the following properties: Separation : $\forall \vec{f} \in E, N(\vec{f}) = 0 \Rightarrow \vec{f} = \vec{0}_E$; Homogeneity : $\forall (\lambda \vec{f} \in \mathbb{R} \times E, N(\lambda \vec{f}) = |\lambda| N(\vec{f})$. Sub-additivity or triangular inequality: $\forall (\vec{f}, \vec{g}) \in E^2, N(\vec{f}, \vec{g}) \leq N(\vec{f})$.

Calculation of the resultant vector of three fuzzy forces

Let be a membership function $f : [0,1] \to [0,1] / f(x) = \dfrac{1}{1+e^{-k(x-s)}}$ which is defined and continuous and where x is a specific characteristic of an object, s is the threshold that expresses the median value for the parameter x and k is a constant expressing the precision with which x is known. To adapt this function to the interval $[-1,1]$, a change of coordinate system and scale is necessary: $F(x) = 2f(x) - 1$ gives $F(x) = \dfrac{2}{1+e^{-k(x-s)}} - 1$. F(x) is an odd function defined and continuous on [-1,1], thus on intervals [-1,0] et [0,1] having 0 for intersection. F(x) allows to build two opposite functions $f : [0,1] \to [0,1], f(x) = \dfrac{2}{1+e^{-k(x-s)}} - 1$ and $f' : [-1,0] \to [-1,0] f'(x) = \dfrac{2}{1+e^{-k(x-s)}} - 1$. It comes that $\forall x \in [-1,1], f(x) = -f'(-x)$. In the same manner, we define $G(y), g(y), g'(y)$ and $H(Z), h(z), h'(z)$. As x, the parameters y and z correspond to different relevant characteristics of the studied object. Then according to the sign of a characteristic x, it is always possible to build the corresponding membership function f and its opposite function f' that expresses the contrary characteristic -x (for example: x: to be tall correspond to the opposite -x: to be small and y: to be joyful, -y: to be sad). The vector $\vec{f}$ is defined from f(x) and the opposite vector $\vec{f'}$ is defined from f'(x) and respectively vectors $\vec{g}$ and $\vec{g'}$ from functions g(y) and g'(y) and vectors $\vec{h}$ and $\vec{h'}$ from functions h(z) and h'(z). This adaptation allows to satisfy the P3 property (recalled in the previous subsection) necessary to build a fuzzy vector space.

Properties of f(x), g(y) and h(z) defined on $[0,1]$

Let be the vector space E on $\mathbb{R}^3$, $E(O, \vec{i}, \vec{j}, \vec{k})$, the resultant vector $\vec{u} = \vec{f} + \vec{g} + \vec{h}$. $\|\vec{u}\| = \sqrt{f(x)^2 + g(y)^2 + h(z)^2}$. Because $f(x), g(y), h(z) \in [0,1], \max(\|\vec{u}\|) = \sqrt{3}, \|\vec{u}\| \in [0, \sqrt{3}]$. The resultant function r of the membership functions f(x), g(y) and h(z) is $\forall (x,y,z) \in [0,1]^3, r(x,y,z) = \frac{1}{\sqrt{3}}\|\vec{u}\| = \frac{1}{\sqrt{3}}\sqrt{f(x)^2 + g(y)^2 + h(z)^2}, r(x,y,z) \in [0,1]$. In the vector space E, we can associate a scalar to a vector that express the importance of the component in the linear combination of vectors. $\alpha, \beta, \gamma \in R^{+*}, \vec{u} = \alpha\vec{f} + \beta\vec{g} + \gamma\vec{h}$ and the norm becomes $\|\vec{u}\| = \sqrt{\frac{|\alpha| f(x)^2 + |\beta| g(y)^2 + |\gamma| h(z)^2}{|\alpha| + |\beta| + |\gamma|}}$. Whatever the use of scalars or not we verified that $r(x,y,z) \approx \frac{1}{\sqrt{3}}\|\vec{u}\|$. The study of the properties of the opposite functions f'(x), g'(y) and h'(z) defined on $[-1,0]$ is similar. Because each function f(x) comes with its opposite function f'(x), $\alpha f(x) = -\alpha f'(-x)$, the sign of α allows to set up if the function f is agonist or antagonist and in the same manner β for g(y) and γ for h(z).

Generalization in n parameters

The model can be adapted to take into account n parameters or characteristics of objects $n \ll \infty$, $x_i \in [0,1]$ and n membership functions $f_i(x_i)$ with a scalar $\alpha_i \in \mathbb{R}^n$. A vector space E is defined on $\mathbb{R}^n$ where two vectors $\vec{f} = (x_1 \dots x_n)$ and $\vec{g} = (y_1 \dots y_n)$ have the inner scalar product (dot product) $\langle \vec{f}, \vec{g} \rangle = x_1 y_1 + \dots + x_n y_n$. E is provided with an orthonormal coordinate system $O(\vec{i_i}, \dots, \vec{i_n})$. Let be the vector $\vec{u}$ in the vector space E on $\mathbb{R}^n$ to be the sum of vectors $\vec{f_i}$ provided with the corresponding scalar α_i. Thus comes $\forall \vec{f_i} \in E, \vec{u} = \sum_{i=1}^{i=n} \alpha_i \vec{f_i}$. The norm $\|\vec{u}\| = \sqrt{\frac{\sum_{i=1}^{i=n} |\alpha_i| (f_i(x_i))^2}{\sum_{i=1}^{i=n} |\alpha_i|}}$. The resultant function r of n membership functions is: $\forall (x_i, \dots, x_n) \in [-1,1]^n, r(x_i, \dots, x_n) = \frac{1}{\sqrt{n}}\sqrt{\frac{\sum_{i=1}^{i=n} |\alpha_i| (f_i(x_i))^2}{\sum_{i=1}^{i=n} |\alpha_i|}}$. The main inconvenience is when n becomes big, the value of the resultant function r becomes very tiny. Indeed, $\lim_{n \to +\infty} r(x_i, \dots, x_n) = 0$. However, this inconvenience is not important because in practice the number n of membership functions $n \ll \infty$ and is some tens of parameters at most.

Consideration of time

We saw that the evolution of the characteristic parameters of objects requires the consideration of time. All the time t, the parameters can evolve more or less quickly or sometimes be constant. The resultant vector $\vec{u}$ undergoes modifications all the time t. The absolute time is expressed by an additional variable t (in seconds) which is calculated from 6 variables: year (aaaa), month (mm), day (jj), the hour (hh), minutes (mm), seconds (ss). An extension in the subdivisions of second may be envisaged for special applications but is mostly useless and expensive in time of computation. The other temporal units (for example, week, month, half-year) associated with the various parameters can easily be converted in the absolute time by appropriate conversion functions. Most programming languages offer such time function libraries. The variable t is used to define every membership function $f_i(x_i)$ which becomes $f_i(x_i, t) : [0,1] \to [0,1], f_i(x_i, t) = \dfrac{2}{1+e^{-k(x_t - s)}} - 1$. Some func-

tions remain constant on an interval of time while others are evolving. For all time t, the vector $\vec{u_t}$ evolves according to its component functions:$\forall \vec{f_{i,t}} \in E, \vec{u_t} = \sum_{i=1}^{i=n} \alpha_{i,t}\vec{f_{i,t}}$. The norm $\|\vec{u_t}\| = \sqrt{\frac{\sum_{i=1}^{i=n} |\alpha_{i,t}|(f_{i,t}(x_{i,t}))^2}{\sum_{i=1}^{i=n} |\alpha_{i,t}|}}$. The resultant function r of n membership functions is: $\forall (x_{i,t}, \ldots, x_{n,t}) \in [-1,1]^n, r(x_{i,t}, \ldots, x_{n,t}) = \frac{1}{\sqrt{n}}\sqrt{\frac{\sum_{i=1}^{n} |\alpha_{i,t}|(f_{i,t}(x_{i,t}))^2}{\sum_{i=1}^{i=n} |\alpha_{i,t}|}}, r(x_{i,t}, \ldots, x_{n,t}) \in [-1,1]$.

kinematic of a point in the fuzzy vector space

This part is inspired from physics and more precisely from classical mechanics. Therefore, we use a three-dimensional cartesian vector space to illustrate our proposal. Let be t in the interval $[t_0, t_n]$ and the vector space E provided with the orthogonal coordinate system $(O, \vec{i}, \vec{j}, \vec{k})$. The vector function $\vec{U}(t)$ in E is defined as $\vec{U}(t) = F(x,t)\vec{i} + G(y,t)\vec{j} + H(z,t)\vec{k}$. As previously, x_t, y_t, z_t are the characteristics of an object at the moment t and the components are defined as $x_t, y_t, z_t \in [0,1]^3, t \in [t_0, t_n], \vec{F}(t) = F(x_t,t).\vec{i}, \vec{G}(t) = G(y_t,t).\vec{j}$ *and* $\vec{H}(t) = H(z_t,t).\vec{k}$ of $\vec{U}(t)$. All of them are fuzzy membership functions of time. Because the coordinate system is orthogonal: $\forall t \in [t_0, t_n], \vec{F}(t).\vec{G}(t) = \vec{0}, \vec{F}(t).\vec{H}(t) = \vec{0}$ and $\vec{G}(t).\vec{H}(t) = \vec{0}$. The function $\vec{U}(t)$ is continuous when $t = t_0$ if and only if their components $F(x_t,t), G(y_t,t)$ and $H(z_t,t)$ are defined and continuous. According to our previous definition of fuzzy forces in a vector space E, that is the case. Indeed $F(t) : [-1,1] \to [-1,1], F(t) = \dfrac{2}{1+e^{-k(x_t-s)}} - 1$ and G(t) and H(t) are defined and continuous in the same way. So $\vec{U}(t)$ is a continuous vector function. O is the origin of the orthogonal coordinate system of the vector space where a point $M = (F(x_t,t), G(y_t,t), H(z_t,t)) = F(x_t,t)\vec{i} + G(y_t,t)\vec{j} + H(z_t,t)\vec{k} / \vec{OM} = \vec{U}(t)$ that depicts a curve which is called the hodograph of the vector function $\vec{U}(t)$ which is the position vector at the time t in the fuzzy vector space E. An example of the trajectory of the M point in the fuzzy vector space E is given on the figure 5 that uses the right-hand rule orientation.

Velocity and speed in the vector space

The derivative of $\vec{U}(t)$ is $\frac{d\vec{U}}{dt} = lim_{t \to t'} = \dfrac{\vec{U}(t) - \vec{U}(t')}{t - t'}$. $\frac{d\vec{U}}{dt}$ exists, if its components $\vec{F}(t), \vec{G}(t)$ and $\vec{H}(t)$ denotes the derivatives $\frac{d\vec{F}(t)}{dt}$, $\frac{d\vec{G}(t)}{dt}$ and $\frac{d\vec{H}(t)}{dt}$. The average velocity becomes the derivative of the position vector $\frac{d\vec{U}(t)}{dt} = \frac{d\vec{F}(t)}{dt} + \frac{d\vec{G}(t)}{dt} + \frac{d\vec{H}(t)}{dt}$. Thus, the velocity is the time rate of change of position in the vector space E and $\frac{d\vec{F}(t)}{dt}$, $\frac{d\vec{G}(t)}{dt}$, $\frac{d\vec{H}(t)}{dt}$ denote the derivative in the three components, so called dimensions with respect to time. The speed S of the point M is defined as the magnitude $S = |\vec{U}(t)| = \frac{ds}{dt}$ where s is the arc-length measured

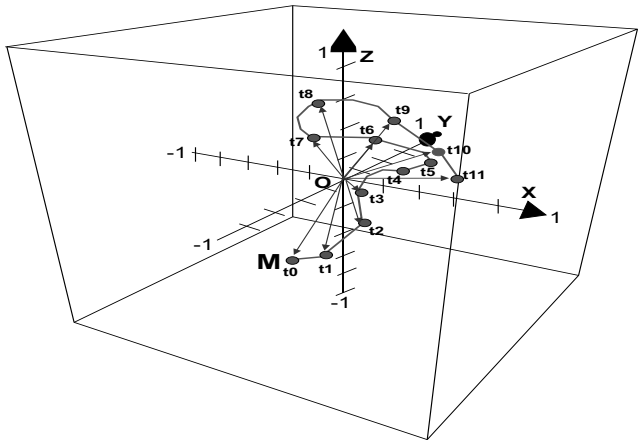


Figure 5: Hodograph of the vector function $\vec{U}(t)$ and dot M in the fuzzy vector space E

along the trajectory of the point M in the vector space E which is a non-decreasing quantity. Therefore $\frac{ds}{dt}$ is non-negative, which implies that the speed S is also positive or zero.

acceleration

The acceleration A of M relies on the rate of change of the velocity vector $\frac{d\vec{U}}{dt}$ so, $A = \frac{d^2\vec{U}}{dt^2}$. The acceleration is the first derivative of the velocity vector $\frac{d\vec{U}}{dt}$ and the second derivative of the position vector $\vec{U}(t)$. In the following, we present an example of application of our model to build an emotion fuzzy vector space.

Derivative of the sigmoid fuzzy membership function

Each component $\vec{F}(t), \vec{G}(t), \vec{H}(t)$ of the vector $\vec{U}(t)$ is defined by the same sigmoid membership function :
$\vec{F}(t) = F(x_t,t) = \dfrac{2}{1+e^{-k(x_t-s)}} - 1.$
Its derivative is calculated according to x_t, $\frac{dF(x_t,t)}{dt} = \dfrac{2k.e^{-k(x_t-s)}}{(1+e^{-k(x_t-s)})^2}$. Provided that k is a positive constant, we notice that the derivative $F'(x_t,t)$ is strictly positive.Therefore, $F(x_t,t)$ is always derivable and strictly increasing on $\mathbb{R}$. Obviously, the same properties are verified for the function $G(y_t,t)$ and $H(z_t,t)$.

Gradient of the vector U in vector space E

Consider an emotion vector space E with three parameters representing (positive/negative) couple of parameters.You can have more than three parameters but three is more convenient to depict our approach. These parameters are changing under the influence of unpredictable events occurring during an interval of time

$[t_0, t_n]$. To model the situation, we propose to use the framework presented beforehand. According to the previous subsection and figure 5, at each moment the position of the dot M and the vector $\vec{U}(t) = F(x,t)\vec{i} + G(y,t)\vec{j} + H(z,t)\vec{k}$ are defined with the membership sigmoid functions as $\vec{F}(t) = F(x_t, t)$, $G(y_t, t)$ and $H(z_t, t)$ described in subsection . Farthest the point M is from origin in the positive part of the fuzzy vectorial space, the most favorable are the conditions to feel well or to achieve correctly a task. We define a gradient named **ability** that correspond to the vector $\vec{U}(t)$ described in the equation 3. At any moment, the ability depends of the level of each component vector. In the cartesian coordinate system, the gradient of ability named $\nabla U(x_t, y_t, z_t) = \left(\frac{\partial U}{\partial x_t} + \frac{\partial U}{\partial y_t} + \frac{\partial U}{\partial z_t}\right)$.

$$\nabla U(x_t, y_t, z_t) = \frac{2k.e^{-k(x_t - s)}}{(1 + e^{-k(x_t - s)})^2} + \frac{2k.e^{-k(y_t - s)}}{(1 + e^{-k(y_t - s)})^2} + \frac{2k.e^{-k(z_t - s)}}{(1 + e^{-z(x_t - s)})^2} \quad (3)$$

k is a positive constant defining the scale and s is the pivot (neutral value), zero in our example. The calculus of sum of the partial derivative according to the direction X, Y and Z allows to calculate at each moment t the gradient of ability which is a value always defined on $[0, 1]$. To illustrate how to use it, the figure 7 shows

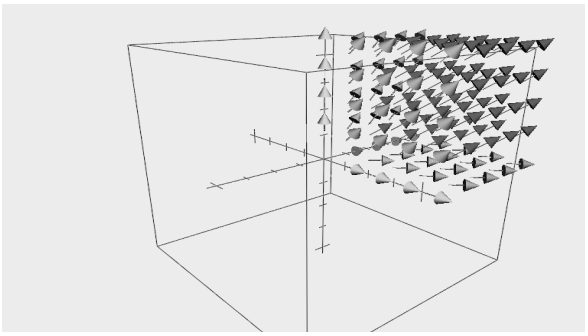


Figure 6: Gradient of vector U in FVS E: $\nabla U(x_t, y_t, z_t)$ (scale k=10)

how the attributes of the objects involved in evaluating a situation are combined in the FVS.

A Psychological model and Emotional states

Most available computationally models of emotion rely on the OCC model (Ortony et al, 1988) Ortony et al. (1988) Picard (1997). The OCC model defines events, agents and objects. Events are considered to induce emotional consequences. Agents are able of actions that have effects on the environment. Objects have imputed

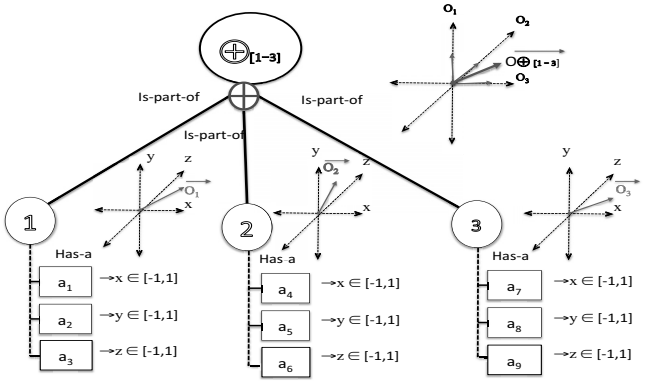


Figure 7: Composition of Object and Attributes in a FVS

properties. The OCC model represents emotions as balanced reactions to the perception of the world. That is: one can be pleased about the consequences of an event or not (pleased/ displeased); one can endorse or reject the actions of an agent (approve/disapprove) or one can like or not aspects of an object (like/dislike). Then, the events can have consequences for others or for oneself and on acting agents. Thus, the different emotional balances are depicted by couples of (positive/negative) reactions represented by variables x,y,z and their corresponding membership functions F(x), G(y) and H(z). For concision, we do not provide the complete specialization tree of the OCC model but we just summarize the couples of variables in the figure 8. A complete description of the OCC model is available in Ortony et al. (1988) and computational applications in Picard (1997).

		+	-
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 8: Couples of (positive/negative) emotion parameters of the OCC model Ortony et al. (1988)

Evaluation of the patient emotions during the treatment

This case concerns the evaluation of state of a patient at each stage of his treatment, how the patient feels and is tolerating it and the interaction with the care givers

when some acts could be painful and when the profits are a long time coming. The model OCC allows to define couples of variables: (Love/Hate) of some drugs used in the treatment, the (Gratitude/Anger) concerns the caregivers (nurses, physicians)and the consequences of events (Joy/Distress) that occur during the care route of the patient described figure 9. These attributes can be represented by fuzzy vectors that change during the time as described in figure 7. The resultant vectors allow to evaluate the patient emotional state and what is more comfortable and more efficient to bring him to his recovery. Thus each drug administration and care is associated with a variable x_t (Joy/Distress) that depicts the mood state of the patient according to events which is essential to reach the cure, y_t (Gratitude/Anger) that evaluates the prescribing physician or the nurse who makes the care and z_t (Love/Hate) that estimates the tolerability of the treatment. The FVS offers the advantage to describe the evolution of the patient's feelings during the treatment and not only a simple snapshot. The scenario presented figure 9 shows the evolution of

Scenario		
Care	**Care Giver**	**Event**
None	None	Ankle injury with a garden tool
Analgesic, Antibiotic, Antitetanic serum	Dr Girard G Practionner	First Prescription
Antitetanic injection, wound treatment	Home Nurse Sophie Platon	Treatment administration
Antibiotic tablets	Patient himself	Antibiotic allergy
Corticoid , new antibiotic prescription	Dr Durand emergency	Emergency Consultation
Care of the wound	Home Nurse Sophie Platon	Suspicion of Phlebitis
Echo Doppler	Dr Durand cardiologist	Echo-Doppler Prescription
Diag: Deep Vein Thrombosis	Dr Durand cardiologist	Heparin Prescription
Heparin Subcutaneous 2/day 10days	Home Nurse Sophie Platon	Heparin Injection
Treatment Follow-up	Dr Girard G Practionner	Healing no after-effects

Figure 9: Care route of a patient

the mood of the patient and of his/her feeling of the caregiver-patient relationship during the different stages of his care.

Future works

We will propose a model of personality and behavior that will be adapted and will complete the emotion layer presented in this paper. The ability of vectorial spaces to model the continuous flow of changing emotions is essential for dynamically expressing the mental states of the agents involved in a situation and to simulate the evolution of their relationships. We are currently using the espace vectorial space to enhance previous clinical decision support systems Colloc and Jacquet (2013),Colloc and Summons (2015). This approach was also used to simulate the process of gaming addiction and shows the emotion evolution of the gamer and his interaction with the slot machine during the game. A prototype of this simulator is developed in VBA Colloc and Summons (2017).

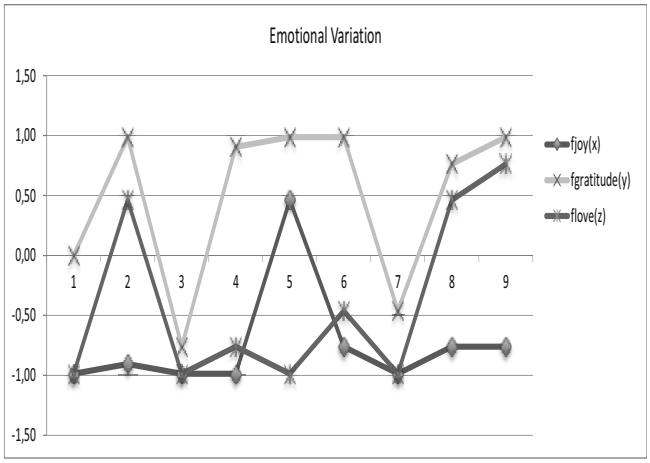


Figure 10: Evolution of the values of the emotion membership functions during the patient care route

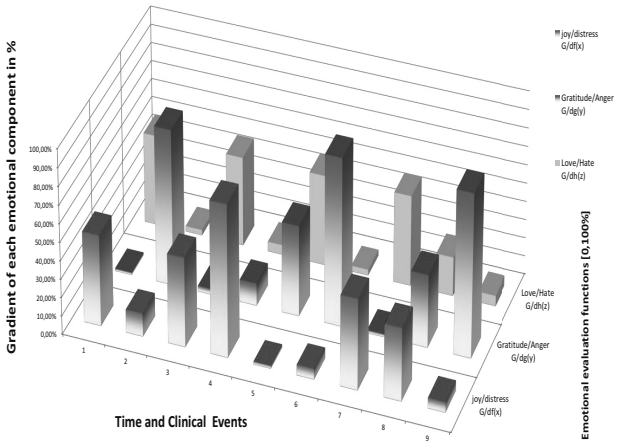


Figure 11: Gradient of values of emotion function

Conclusion

In this article we proposed a model of " fuzzy vectorial space " which constitutes a new approach to combine fuzzy logic and an object oriented model that allows time modelling of the structure and behaviour of complex objects. We have shown that it is essential to have means to aggregate the properties and features of objects as a supplement to the use of the operators of conjunction and disjunction (which can be obviously jointly used when it is semantically needed). The membership functions describes the attribute values of objects and their evolution. The fuzzy vectorial space provides a new aggregation or composition operator that can be widely used to enrich object oriented models and multi-agents systems. The fuzzy vectorial space model does

not use fuzzy rules to represent the object features but instead determines a resultant vector which represent the evolution of the object components during the time. We trust that the use of the kinematics modeled with fuzzy vectorial spaces, which is inspired by the physics, constitutes an enhancement to take into account the evolution of the structure, attributes, properties and behavior of complex objets while profiting from the advantages of the fuzzy logic. The evolution of the gradient of the resultant vector can be studied to globally evaluate the object state according to the semantics of the membership functions. This contribution offers a new pragmatic way to combine object properties and to compare them together. We have shown how to use the fuzzy vectorial space for emotion modelling with the use of the OCC psychological model. Our approach leads to develop a model of analogy allowing to compare objects and situations in complex environments where the time is an essential factor in many domains like medicine, management, logistic, cognition and human sciences.

References

Bartolin R.; Bouvenot G.; Soula G.; and Sanchez E., 1982. *Apport des sous-ensembles flous à l'aide au diagnostic biomédical : à propos de deux applications concrètes". Sem Hop Paris*, 58(22), 1631–1635.

Booch G.; J J.R.; and Jacobson I., 1998. *The Unified Modeling Language User Guide*. Addisson Wesley.

Bouchon-Meunier B., 1993. *La logique floue*. Que sais-je, n 2702.

Bouchon-Meunier B. and Marsala C., 2003. *Logique floue, principes, aide à la décision*.

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*.

Buchanan B.G. and Shortliffe E.H., 1984. *Rule-Based Expert Systems, The MYCIN Experiments of the Stanford Heuristic Programming Project*. Addison-Wesley.

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

Chen P. and Pin-Shan P., 1976. *The Entity-Relationship Model - Toward a Unified View of Data. ACM Transactions on Database Systems*, 1(1), 9–36.

Colloc J. and Jacquet A., 2013. *Système multi-agents d'aide à la décision clinique, appliqué à la prise en charge neuropsychologique des aphasiques par AVC*. Dalloz.

Colloc J. and Summons P., 2015. *An Analogical Model to Design Time in Clinical Objects*. In *RITS' Dourdan France*. 124–125.

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.

Colmerauer, 1996. *The birth of Prolog*.

Colmerauer A.; Kanoui H.; Roussel P.; and Robert Pasero, 1973. *Un système de communication homme-machine en Français*. Tech. rep., Groupe de recherche en intelligence artificielle.

El-Nasr M.S.; Yen J.; and Ioerger T.R., 2000. *FLAME - Fuzzy Logic Adaptive Model of Emotions. Autonomous agents and multi-agent systems*, 3, 219–257.

Iancu I., 2012. *A Mamdani Type Fuzzy Logic Controller*. In P.E. Dadios (Ed.), *Fuzzy Logic - Controls, Concepts, Teories and Applications*, Intech. 325–350.

Kaufmann A., 1973. *Introduction à la théorie des sous-ensembles flous, tome 1*. Paris, Masson.

Kaufmann A. and Pichat E., 1977. *Méthodes mathématiques non numériques et leurs algorithmes*. Masson.

Lee C.C., 1990. *Fuzzy Logic in Control Systems: Fuzzy Logic Controller. Part I, II, IEEE Transactions on Systems, Man and Cybernetics*, 20(2), 1182–1191.

Leroy X. and Weis P., 1993. *Manuel de Référence du langage CAML*.

Lorestani A.N.; Omide M.; Shoobaki S.B.; Borghei A.M.; and Tabatabaeefar A., 2006. *Design and Evaluation of a Fuzzy Logic Based Decision Support System for Grading of Golden Delicious Apples. International Journal of Agriculture and Biology*, 8(4), 440–444.

Mamdani E.H. and Assilian S., 1975. *An Experiment in linguistic synthesis with fuzzy logic Controll. International Journal Man-Machine Studies*, 7, 1–13.

Martin O. and Klir G.J., 2007. *Defuzzification as a special way of dealing with retranslation. International Journal of General Systems*, 36(6), 683–701.

McCarthy J., 1987. *Generality in Artificial Intelligence*. In N.Y. ACM Press (Ed.), *ACM Turing Award Lectures, the First Twenty Years 1966-1985*, Addison-Wesley Publishing Company. 257–267.

Minsky M., 1975. *A Framework for Representing Knowledge*. In *The Psychology of Computer Vision.*, McGraw-Hill.

Minsky M., 1985. *The society of mind*. Simon and Schuster, New York.

Montiel O.; Castillo O.; Melin P.; and Sepulveda R., 2009. *Medaitive Fuzzy Logic for Controlling Population Size in Evolutionary Algorithms. Intelligent Information Management*, 1, 108–119.

Opricovic S. and Tzeng G.H., 2003. *Defuzzification within a Multicriteria Decision Model. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 11(5), 635–652.

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

Picard R.W., 1997. *Affective Computing Cambridge*. MA: The MIT Press.

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

Zadeh L.A., 1965. *Fuzzy Sets. Information and Control*, 8, 338–353.

VR, AR, MR SIMULATIONS AND INSPIRATIONS FROM "IRON MAN 3"

Helena Barbas
Faculdade de Ciências Sociais e Humanas – Universidade Nova de Lisboa
Av. de Berna, 26-C
1069-061 Lisboa, Portugal
e-mail: hebarbas@fcsh.unl.pt

KEYWORDS
Digital Humanities; Human-computer interaction; Design Fiction; Iron Man 3; Female Super-heroes

ABSTRACT

Considering that all kinds of simulations are based on verisimilitude and take their referents from the real world, this presentation will focus on the tools – exosuit, helmet – holographic prototypes, and «design fiction» elements in the film "Iron Man 3". Some Sci-fi works can become prophetic a-posteriori, or more humbly show what is presently called "design fiction" objects, scenarios, or human behaviour in its context. Presented items and interfaces deserve to be explored, mainly because they have, and are, impelling the creation of real life objects and procedures.

INTRODUCTION

Simulation is the imitation of a real-world process or system; or «the production of a computer model of something, especially for the purpose of study» (O.E.D), so entering the scope of «fiction». All simulations rely on «re-presentations» of any kind of thing, or image, having the «real world» as a referent. So, they need to be as verisimilar as possible – in aristotelic and narratological terms.

In the real-world the simulation's verisimilar intention is restricted by the technical possibilities allowed by the implements being tested; their effects need to be graphically presented (2D or 3D) in order to be correctly shared (Dozortsev, 2017; Forkel, 2017).

In opposition to the linearity of narratives, Sci-fi Comics give 2D images of the worlds, functions, scenarios and characters created by their authors. Cinema supplements them with movement and, the advances in computer software, special effects techniques, and VRX, add the possibilities of translating them into 3D (or even 4D).

Some Sci-fi works can become prophetic a-posteriori, or more humbly show what is presently called "design fiction" objects, scenarios, or human behaviour in its context. Some items and interfaces seen in movies deserve to be studied either because they are influencing the creation of real life objects, or because the fictional technical hypothesis is worthy of closer critical attention.

The case study here will be "Iron Man 3".

The main curiosity about Iron Man as a superhero is that he is only too human, with no particular superpowers and, moreover, handicapped. So, in its way, the exosuit Extremis is multifunctional, varying from a prosthesis (it extends bodily functions), to a cocoon, to an autonomous robot, and a weapon.

Other «design fiction» artefacts presented in this movie will also be explored here, namely the Helmet(s) - HUD (head-up displays); how these FUI (Fantasy User Interfaces) are being adapted (or not) considering HITLs (human-in the loop situations); how the model mock-ups conform to human factors' requirements. The importance of the methods by which information is visually conveyed to the character as part of the user's interface; and how the character reacts, creates, and uses these new «design fiction» interfaces (i.e. holograms) re-presenting or exhibiting «design fiction» behaviour/gestures to deal with this new sort of data.

Considering that Marvel has already created a new female character as Iron Man's counterpart, with a similar biography and background to Tony Stark, it seems relevant (from a literary, psychological, social or economic perspective) to address the issues of female super-hero(in)s, their recent boom as male super-hero counterparts - and their fiasco. Riri Williams, the next "Iron Heart", is a 15-year-old MIT African-American girl who reverse-engineered one of Tony's exosuits in her dorm room. The film is expected to be premiered in November 2017.

Immersion/presence issues will not be discussed here, nor the implications of/from Cognitive Sciences implied by the several related experiences, to be object of a future work.

ABOUT "IRON MAN 3"

Iron Man – aka Tin Man, Iron Patriot, The Mechanic - is a fictional superhero created by writer and editor Stan Lee, developed by scripter Larry Lieber, and designed by artists Don Heck and Jack Kirby. The character made his debut appearance in *Tales of Suspense* #39 (March 1963), published by Marvel Comics.

In the basic plot, Antony Edward 'Tony' Stark is a millionaire, a playboy and a scientist. He is kidnapped, and suffers a severe chest injury. His captors intend to force him to build a weapon of mass destruction. Instead he creates a powered suit of armour (Mark I) to save his life and escape captivity.

The main curiosity about Iron Man as a superhero is that he is only too human, with no particular superpowers, and moreover, with severe health issues

The exosuit Extremis

On its own, the exosuit Extremis is a multifaceted prosthesis, extending bodily functions. It provides medical health check-ups and some first aid medical treatments. Sometimes called a cocoon, it works as a command centre, an autonomous robot. It is an armour and a weapon – and also a prison when Tony Stark uses Mark XLII to immobilize by force his enemy Aldrich Killian.

It is the suit (from Mark I to XLVI) that features the several re-incarnations of the *Invincible Iron Man*'s Comic book series and, with its continuous upgrades and add-ons, is translated into film (2008, 2010, 2013).

In "Iron Man 3" the exosuit is brain-commanded by its user (live-feed). Mark XLVI has been worn by others than Tony (his girl-friend Pepper Potts; or the American army general James Rhodes), consequently, in the plot, it can be and is hacked by friends and foes.

Concerning the exosuit's AI utilities it has memory; it has access to, and can deal with, big data in the cloud. Due to its training – kindly voiced through J.A.R.V.I.S. (Just A Rather Very Intelligent System) – it knows exactly what is necessary to Tony's survival at every moment. Nevertheless, the brain-connections are still challenging. The exosuit is not switched off when Tony goes to sleep, and from this 'glitch' results that the suit/robot is activated by Tony's nightmares: its AI system has not learned how to distinguish between human vigil and sleep states.

Another more standard of this world problem is the mechanism's need to be charged: without power it becomes very vulnerable, and presents the symptoms of dead/sleep (it is rescued by a boy, in a village).

Exosuit ersatz and exoskeletons

Announced still during Obama's presidency, The Pentagon is preparing a military version of the exosuit: «TALOS, or Tactical Assault Light Operator Suit, is a battery-powered robotic exoskeleton designed to protect the lives of soldiers on the front lines, especially those who lead the army in a mission». «In August 2018, Iron Man may go beyond fictional comic books, saving the day in real life. The U.S. Army is developing an advanced military suit that people have been likening to Tony Stark's high-tech armor» (Lanaria 2015). Also, «SOCOM expects 'Iron Man' suit testing by summer 2018» and «The Tactical Assault Light Operator Suit (TALOS), which became known as the "Iron Man" suit shortly thereafter, became its endgame.» (Douglas 2017).

Figure 1 – Iron Man Comics, Exosuit and TALOS

In this case, both Extremis and TALOS share the same power failure issues. Besides, the latter will not yet allow the wearer to fly.

Some other researchers inspired by "Iron Man 3" urge to divert the focus from its military uses: «The attitudinal framing of the exoskeleton asks that people aspire to transhumanism, and that they "imagine the possibilities in the near future of dramatically enhance[ed] human mental and physical capacities"» (Penderson, 2017).

Following this stance, in the EXOSKELETON REPORT (2015) - http://exoskeletonreport.com (i.e.) - a catalogue of applications, software, body parts addressed, and companies and firms (listed on the Stock Market) dedicated to the development of assistive wearable technology is provided, not only for medical rehabilitation purposes, but also for education, commerce and industry.

THE HELMET - HUD

Harder and more interesting to emulate is the "Iron Man 3" helmet. It is an aesthetical, cinematic, special effects and VRX feat. Done in stereo to avoid flatness (Townsend 2013) the software used allowed every piece to be dimensionalised, and adapted in/for every film sequence.

The helmet is also a HUD (head-up display) replicating the status bar in video gaming (i.e. the main character's health, items, game progression), the method by which data is visually relayed to the player as part of a game's user interface. It took its name from the monitoring in modern aircraft (pilots' helmets).

In "Iron Man 3" the VRX for the helmet's visor varies in accordance with each situation. Its creators wanted to «minimize the visual clutter» (Townsend 2013) always easily decipherable by Tony – a character that is «comfortable with dense data displays». Five icons are kept persistently in the lower part: suit status-, targeting- and optics-, radar-, artificial horizon-, and map-; sometimes augmented with goal-, person, location-, and object-sensitive awareness.

They chose the minimum pixels or phosphors possible, resorted to thin faint lines, and aimed to improve reactions to out-the-window events in a: «full-field-of-vision, very high-resolution, full-colour display» offering stereoscopic imaging. According to the VRX creators (Cantina.co), the design work was inspired by medical MRI diagnostic pattern-recognition and graph theory, namely the circular 'connectograms' used by connectomics – the study dedicated to mapping and interpreting all of the white matter fibre connections in the human brain.

The HUD's visor allows Tony to see the world around him as if he were not wearing the helmet, and lets him read his most useful data in milliseconds.

To achieve the unity/uniformity in these illusions, different programs, processes and filmic techniques were used to attain the same visual effect.

Through the several scenes three different angles are shown: Tony's point of view; the impossible camera, as the audience looking back at Tony's face; and a lateral perspective.

Figure 2 – Tony Stark's perspective

Figure 3 – Outer (impossible camera) perspective

Figure 4 – Lateral Perspective

The special effects and VRX cinematic achievement result in an (apparent) identical and proportionate 3D architectural experience of the images, multi-layered, projected over the scenarios of Tony's "real world". The closest to this experience – according to their creators – could now be provided by Google/Samsung glasses and Microsoft HoloLens, but still on flatland.

Aviation HUDs

Besides medical optical equipment, the creators were also inspired by «steampunk props, precision scopes and combat aviation systems» (Townsend 2013).

In aviation HUDs (i.e. Rockwell Collins) what is (or could) be happening out the window is almost always more important than what is shown on the display.

Presently, Enhanced Flight Vision Systems (EFVS) are heavily controlled by the FAA (Federal Aviation Administration). One of the latest releases (2017) expands its applicability to business aircraft owners and operators, but does not even come close to the experience provided by Tony's visor.

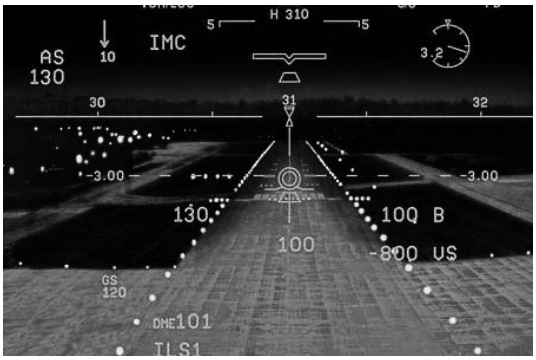


Figure 5 – EFVS - Enhanced Flight Vision System - FAA

In these real world areas, the problems faced by users with current data feeds is excessive information, making them overwhelming, unreadable – useless. Tony's speed in deciphering his HUD's details belongs to the field of gestural/behavioural «design fiction».

In real life, the more sophisticated the HUDs become, the easier they trap the pilots' attention, distracting them, slowing their reactions to out-of-the screen events. The classic report from NASA (Fisher, Haines, and Price, 1980) refers that, in their simulator study, a couple of pilots landed their planes on top of another without even noticing it.

The utility of CBTS (Computer-based training systems), either for accident reduction or in economic terms is not questioned: «Computer-based training systems (CBTS) for process operators are both a generally recognized highly efficient tool for operator training and a major global business» (Dozortsev 2017:37), and the author cautions: «the distortion of reality, or of operators perception, give way to the development of false skills».

Other wearables

From the VR, AR, MR perspectives, the HUD anticipates – or has its counterpart – in Samsung Gear, Google Oculus Rift, Konica Minolta or Microsoft HoloLens.

VR developers now have access to unprecedented physical interfaces and interactions – including wearables, curved spaces, and complex object physics.

From the medical fields (i.e. Neurosky, Mindwave) come alternative affordable solutions for health and wellness, education and entertainment.

New gadgets/biosensors can turn PC's into body or brain activity (EEC, ECG) monitors, check the attention levels of individuals. Some even allow the user to send short twitter or e-mail messages via a Google app.

The problem with all these wearables – even the ones that boast to be brain connected – is that they can only operate with their own proprietary software, and the number of applications each brand provides is very limited.

AR, VR, MR SCENES – THE HOLOGRAMS

Technically, the hologram scenes in "Iron Man 3" were created using Lidar scans, and other 3D software (i.e. Pixar Renderman). Reality scenarios were transformed into a geometry «capable of being rendered as lines» (Townsend 2013) which allowed the interaction between Tony Stark and the settings.

Tony Stark's holograms

Inside the narrative, Tony uses three-dimensional holographic horizontal interfaces, projected by J.A.R.V.I.S. into mobile vertical transparent "screens".

The Interfaces presented in the film are, likewise, a combination of actually existing digital technologies and their «design fiction» extensions. From the audience's perspective, Tony is immersed in 3D demonstrations that he can watch from all sides, and fully interact with, very easily. Tony's behaviour and swift handling of the holographic prototypes fits, again, into gestural «design fiction».

Figure 6 – 3D development – Simon Maddison

Supposedly these are inspired by Microsoft's Kinect (2010) and Holodesk (2012), that provide DIY scenes activated by natural user voice and gesture commands for Windows. In its turn, they were initially simulations of the controls in the narrative "Minority Report" (Philip K. Dick, Steven Spielberg, 2002), and of the HoloDeck in "Star Trek" (1966, 1979). Microsofts' Kinect research group (Alex Kipman) has now moved to HoloLens (2016).

Figure 7 – Microsof HoloLens experience

In real life, the illusion of touching and moving objects, mainly with computer projections, has to deal with several fields of Physics. The first one being the images' absence of density (inspiring kinetic and haptic perception research); computer and human vision issues (parallax, stereopsis and perspective to be differently addressed in humans and machines); the lack of adequate tools – either software (Faath, 2017), or cameras (Forkel 2017; Milliron, 2017).

In spite of all this, Tony's holograms are said to be emulated by Elon Musk (2017), the Tesla CEO, in the 3D printing of a rocket part – shown in a video.

Aldrich Killian's holograms

Aldrich Killian tosses a rolling ball 3D projector, a holographic device used to show the inside of his brain, his thoughts, to Pepper Potts.

Figure 8 – Previs design – Killian device by John Koltai

For the creators (Koltai, 2012), Killian's brain hologram was something new in the Marvel world. They resorted to real 3D mapping data on academic studies, once more related to the definition of the fibre pathways of the human brain. For the mind exterior and differences in depth, they resorted to three dissimilar kinds of software and image renders.

Figure 9 – Aldrich Killian's thoughts shown to Pepper Potts

Not even with very special VRX is it (yet) possible to give verisimilar information from inside the human head.

In the AR scenes, Mandarin and Pepper, Tony and Pepper, share their thoughts - "materialized" in the "external" holograms simulating the contents of their respective heads. The intent is that ideas can be shared – but the plot fails that illusion. The actors in the scenes are both receivers. For this artifice to be credible, even in «design fiction» terms, the thoughts, being of the same matter, had to become somewhat mixed: what one shows with what the other subjectively sees; and both should interfere with the (re)presented scene.

Presently, there are micro-sized projectors widely available, with several companies offering them in smartphone attachment form.

WOMEN IN IRON MAN 3 AND GIRL HERO(INE)S

The female characters in "Iron Man 3", namely the girl-friend Pepper Potts, are narrative 'clichés'. From a literary (psychological, social and economic) perspective, and in spite of the boom of female super-hero counterparts shaped in the last years – about 7 for Marvel, 121 for DC – Comics' super-heroins suffer the highest lack of credibility.

The resuscitation of "Wonder Woman" (2016) – even by a lady director, Patty Jenkins – resorting to super-powers endowed by some props, has not improved the situation.

In this scenario, Marvel has created a new character, "IronHeart", with the same background as Tony Stark: a 15 year old MIT African-American girl named Riri Williams, who reverse-engineered an exosuit in her dorm room. The film is expected to be premiered in November 2017. But the project does not raise high expectations – the student is working on an existing machine, not creating anything of her own, by herself.

From the long panoply of Comics' wonder-girls, the only one with an acceptable biography is the Scottish "Super Gran" (Forrest Wilson, 1980; TV adaptation by Jenny McDade 1985-1987).

Her theme lyrics start: «Stand back Superman, Iceman, Spiderman. Batman and Robin too […] Hang about - Look out! For Supergran». Granny Smith is hit by a ray that enhances her normal human senses into superpowers. She has some engineering skills, dealing with her flycycle, a super sleigh and an anti-gravity belt, migrating from TV to an Arcade Game for ZX Spectrum 48K (1985).

Figure 10 – Super Gran (1980-1985)

CONCLUSION

This short tour around Iron "Man 3" artefacts – the exosuit and helmet – and VRX holographic effects has shown the evolution and dissimilarities concerning Sci-Fi model mock-ups and their possible uses by humans in real-life. The interactions of Tony Stark's character with the film interfaces have become a data representation epitome that has been imitated in all fields – army, commerce and industry.

The real-life letdowns can be attributable either to the unreadiness of the tools available (software, equipment), or to HITLs issues. A latent concern is the need to translate the available 2D information into 3D, and a persistent one is to respect verisimilitude: «changes in reality must be reproduced in the interface» (Dozortsev, 2017). In non-fictional terms, out of games and films, VR, AR, MR and reality have to coincide without flaws.

Concerning female super-hero(in)s verisimilitude could also be achieved with a mere exploration of the development of more modern human skills.

REFERENCES

Blattgerste J, Strenge B, Renner P, et al (2017) Comparing conventional and augmented reality instructions for manual assembly tasks. Proc 10th Int Conf Pervasive Technol Relat to Assist Environ - PETRA '17 75–82. doi: 10.1145/3056540.3056547

Colgan A (2014) How Do Fictional UIs Influence Today's Motion Controls? In: LEAP / Blog. http://blog.leapmotion.com/fictional-uis-influence-todays-motion-controls/. Accessed 13 Sep 2017

Cuervo E (2017) Beyond Reality: Head-Mounted Displays for Mobile Systems Researchers. GetMobile Mob Comput Commun 21:9–15.

D'Elia N, Vanetti F, Cempini M, et al (2017) Physical human-robot interaction of an active pelvis orthosis: toward ergonomic assessment of wearable robots. J Neuroeng Rehabil 14:29. doi: 10.1186/s12984-017-0237-y

Dozortsev VM, Frolov AI, Novichkov AY, Pogorelov, alery P. Honeywell JSC (2017) Field Operator's Interface in Computer Simulator: Virtual Tours vs Virtual Reality. In: SCIFI-IT'2017 - The 2017 International Science Fiction Prototyping Conference. Eurosis - ETI, Brugge, pp 37–41

Fischer E, Haines RF, Price TA (1980) Cognitive Issues in Head-Up Displays. NASA Tech Pap 1711 1–28.

Forkel E, Baum J, Schumann C-A (2017) Applications of 3D-Measurement for Process Innovation. In: SCIFI-IT'2017 - The 2017 International Science Fiction Prototyping Conference. Eurosis - ETI, Brugge, pp 65–67

Matzen K, Cohen MF, Evans B, et al (2017) Low-cost 360 stereo photography and video capture. ACM Trans Graph 36:1–12. doi: 10.1145/3072959.3073645

Milliron T, Szczupak C, Green O (2017) Hallelujah: The World's First Lytro VR Experience. doi: 10.1145/3089269.3089283

Pedersen I, Mirrlees T (2017) Exoskeletons, Transhumanism, and Culture: Performing Superhuman Feats. IEEE Technol Soc Mag 36:37–45. doi: 10.1109/MTS.2017.2670224

Smith BM, Desai P, Agarwal V, Gupta M (2017) CoLux: Multi-Object 3D Micro-Motion Analysis Using Speckle Imaging. ACM Trans Graph Artic. doi: 10.1145/3072959.3073607

Thies J, Zollhöfer M, Stamminger M, et al (2016) FaceVR: Real-Time Facial Reenactment and Eye Gaze Control in Virtual Reality. doi: 10.1145/3084822.3084841

Vosmeer M, Schouten B (2017) Project Orpheus a research study into 360° cinematic VR. TVX 2017 - Proc 2017 ACM Int Conf Interact Exp TV Online Video. doi: 10.1145/3077548.3077559

Wagemakers AJ, Fafard DB, Stavness I (2017) Interactive Visual Calibration of Volumetric Head-Tracked 3D Displays. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17. pp 3943–3953

Werrlich S, Kai Nitsche bmwde, Notni G (2017) Demand Analysis for an Augmented Reality based Assembly Training. doi: 10.1145/3056540.3076190

Yandell MB, Quinlivan BT, Popov D, et al (2017) Physical interface dynamics alter how robotic exosuits augment human movement: implications for optimizing wearable assistive devices. J Neuroeng Rehabil 14:40. doi: 10.1186/s12984-017-0247-9

WEB REFERENCES

Bartenbach V (2016) EduExo: The First Robotic Exoskeleton Kit for STEM Education. In: Kickstarter. https://www.kickstarter.com/projects/1485976654/eduexo-the-first-robotic-exoskeleton-kit-for-stem. Accessed 01 Sep 2017

David (2012) Interview: Jayse Hansen (The Avengers). In: Invent. Interact. http://www.inventinginteractive.com/2012/07/08/interview-jayse-hansen/. Accessed 01 Sep 2017

Ernst D (2017) Special Operations Command expects TALOS "Iron Man" suit testing by summer 2018 - Washington Times. In: Washingt. Times. http://www.washingtontimes.com/news/2017/feb/16/special-operations-command-expects-talos-iron-man-/. Accessed 01 Sep 2017

Grady M (2016) FAA Expands Enhanced Vision Capabilities – AV web flash Article. https://www.avweb.com/avwebflash/news/FAA-Expands-Enhanced-Vision-Capabilities-228201-1.html. Accessed 01 Sep 2017

Grouchnikov K (2016) The magic of visual effects – interview with Venti Hristova. In: Push. Pixels. http://www.pushing-pixels.org/2016/04/17/the-magic-of-visual-effects-interview-with-venti-hristova.html. Accessed 01 Sep 2017

Grozdanic L (2015) Elon Musk uses Iron Man-inspired holographic 3-D user interface to print a rocket part. In: inhabitat - Green Des. Innov. Archit. Green Build. http://inhabitat.com/elon-musk-unveils-his-iron-man-inspired-hand-manipulated-3d-holographic-technology/. Accessed 01 Sep 2017

Internet Movie Database (2014) Avengers: Age of Ultron (2015) - IMDb. In: Internet Movie Database. http://www.imdb.com/title/tt2395427/faq?ref_=tt_faq_sm#.2.1.5. Accessed 01 Sep 2017

Koltai J (2012) IRON MAN 3 - Previs Design. http://www.johnkoltai.com/IRON-MAN-3-Previs-Design. Accessed 01 Sep 2017

Konica Minolta Develops the World's First Automotive 3D Augmented Reality Head-up Display - News Releases | KONICA MINOLTA. In: 2017. https://www.konicaminolta.com/about/releases/2017/0227_01_01.html. Accessed 01 Sep 2017

Lanaria V (2015) U.S. Military To Deliver Its First Bulletproof, Weaponized Iron Man Suit In 2018.

Russon M-A (2014) MIT Recreates Tony Stark's Iron Man Transparent Projection Screen. Int. Bus. Times

The Fuzzy Future of Holographic Display Interfaces [Video] - Tested. http://www.tested.com/tech/405-the-fuzzy-future-of-holographic-display-interfaces-video/. Accessed 01 Sep 2017

Saeed A (2012) Meet The Man Behind The Fantastical Technology In The Avengers, Rise Of The Planet Of The Apes, 2012 And More. In: Creators. https://creators.vice.com/en_us/article/vvz8yx/meet-the-man-behind-the-fantastical-technology-in-ithe-avengersi-irise-of-the-planet-of-the-apesi-i2012i-and-more-qa. Accessed 01 Sep 2017

(1985) Super Gran. In: ZX Spectr. Rev. http://www.zxspectrumreviews.co.uk/review.aspx?gid=5615&rid=5628. Accessed 29 Sep 2017

(2013) Simon Maddison. In: 13 CO. http://13co.io/simon-maddison/. Accessed 01 Sep 2017

(2015) Exoskeleton Report. http://exoskeletonreport.com/. Accessed 01 Sep 2017

(2017) Christopher Townsend. In: IMDb. http://www.imdb.com/name/nm0870102/?ref_=fn_al_nm_1. Accessed 01 Sep 2017

(2017) Worth the Wait — FAA's New EFVS Rule FAR 91.176 – Rockwell Collins blogs. https://blogs.rockwellcollins.com/2017/01/17/worth-the-wait-faas-new-efvs-rule-far-91-176/. Accessed 01 Sep 2017.

BIOGRAPHY

HELENA BARBAS (1951) is Professor-Lecturer of the Department of Portuguese Studies – F.C.S.H. – U.N.L. She holds a MA (1990) and a PhD (1998) in Portuguese Studies – Comparative Literature – Literature and the Arts. In 2005 she gained her "Habilitation" (2008) in Literature and Cyberarts.

In 2003 Helena attended a M.Sc. in Applied AI at F.C.T.-U.N.L. (Campus da Caparica) and from then on centred her research in the field of Digital Humanities: human-machine interaction, e-learning, interactive digital narrative, cloud computing, and serious games.

She was a member of the InStory team (2005-2007) – best Portuguese web mobile project 2006. She prepared a project on serious games, PlatoMundi, aiming to introduce e-learning and ethical issues in game playing; she is developing a new project – Numina.

In 2011 she received the (2nd.) SANTANDER Award for the Internationalization of the F.C.S.H. Scientific Production 2010, and in 2015 the «Best paper Award» for *Cloud Computing and (new) mobile storytelling in the Internet of Things*, presented at EUROMEDIA'2015, I.S.T., Lisbon, Portugal.

Homepage: http://www.helenabarbas.net

SIMULATION METHODOLOGY AND TOOLS

A TIMED EXTENDED REACHABILITY GRAPH FOR THE SIMULATION AND ANALYSIS OF BOUNDED TIMED PETRI NETS

Dimitri Lefebvre
Normandie University, UNIHAVRE, GREAH, 76600 Le Havre, France
e-mail: dimitri.lefebvre@univ-lehavre.fr

KEYWORDS
System analysis, Model desing, Event-oriented, CIME, Control systems.

ABSTRACT

This paper is about the simulation and analysis of timed discrete event systems (DESs) modeled with timed Petri nets (TPNs) The earliest firing policy is considered with respect to a set of temporal specifications. Incorporating the time in the model is important to consider many practical problems: in particular the earliest firing policy - where each enabled transition fires as soon as possible - is used to optimize the cycle time and the makespan in control and scheduling problems. In order to simulate and analyze the timed trajectories of the DES, we propose a Timed Extended Reachability Graph (TERG) that includes explicitly the time specifications into the model. The equivalence between the paths in TERG and the timed trajectories in the TPN is proved. The domains of application include but are not restricted to CIME. A set of simulations illustrates that the computational cost remains a critical issue for large systems.

INTRODUCTION

Performance in most manufacturing settings is affected by operative policies related to job scheduling and by time specifications and constraints. The selected policy has a significant impact on the systems efficiency, the operational costs and the service promise fulfilment. In order to guarantee the system efficiency, computational applications need to be developed that should be able to provide modeling aids by incorporation structural and temporal specifications in an explicit way into the model. Minimal time requirements are frequently considered as temporal specifications because they represent in a natural way the limitation of numerous technological systems. In addition, the earliest firing policy that executed each job as soon as possible is also very popular because such a policy optimizes the cycle time and the makespan in control and scheduling problems (Feng et al. 2016, Lopez and Roubellat, 2008)
In this paper, we consider timed discrete event systems with minimal time requirements and earliest firing policy. Timed Petri nets (TPNs) are used to model the systems and to incorporate the temporal specifications that correspond to the minimal durations to wait before the transitions fire. The PNs and TPNs are frequently used for performances evaluation, model checking but also for control issues, particularly for scheduling problems applied in the field of manufacturing systems (Baker and Trietsch, 2009. Leung, 2004, Jeng and al., 1998; Wang and Wang, 2012; Lei et al., 2014; Lefebvre and Leclercq 2015; Lefebvre, 2016a, 2016b).
The contribution of the present paper is to design a Timed Extended Reachability Graph (TERG) that encodes both the time specifications and the earliest firing policy in the usual reachability graph (RG). The TERG is proved to describe in an exhaustive way all feasible trajectories of the considered system and to remain of finite size as long as the RG is also finite. It is suitable for model checking, and behavioral analysis. It is also suitable for control and scheduling applications. Extending the RG to take into account the temporal specifications has been already studied. The States Class Graph (Berthomieu and Menasche, 1983; Berthomieu and Vernadat, 2003) was the first method of state space representation adapted to Timed PNs with firing time intervals associated to the transitions. For this purpose, a time domain is added to each marking thanks to a set of inequalities over the firing variables. The Zones Based Graph (Gardey et al., 2003; Lime and Roux, 2006) is another approach inspired by the Region Graph technique. This approach abstracts also automata or TPNs with time intervals associated to the transitions by using a set of zones to encode the clocks of the transitions. The zones are represented by difference bound matrices. However, in the previous approaches, the time information does not appear explicitly in the graph and the method is not directly suitable for control and scheduling issues. The explicit abstraction (including the time information) of a Timed PN with time intervals associated to the transitions has been proposed in (Klai et al., 2013) by introducing a Timed Aggregate Graph that is an exact representation of the reachability state space of a Timed PN. The states of the TAG are defined as a quadruplet that includes the marking, the set of enabled transitions, the minimum time the system must stay in the state, and the maximal time the system can stay in the state. But compared to the TERG, the TAG contains some useless information as the maximal time that is not considered in our problem. Moreover, it does not include an explicit calendar that specifies for each enabled transition the earliest firing. Such a calendar is important in many issues as control and scheduling.
The rest of the paper is organized as follows. Section 2 is about the modeling of timed DESs with TPNs. Section 3 presents the construction of TERG and details the properties of the resulting graph. Section 4 is a case study that illustrates the computational cost. Section 5 sums up the conclusions and perspectives.:

TIMED PETRI NETS

A PN structure is defined as $G = <\boldsymbol{P}, \boldsymbol{T}, W_{PR}, W_{PO}>$, where $\boldsymbol{P} = \{p_1,\ldots, p_n\}$ is a set of n places and $\boldsymbol{T} = \{t_1,\ldots, t_q\}$ is a set of q transitions of indexes $\{1,...,q\}$, $W_{PO} \in (\mathbf{N})^{n\times q}$ and $W_{PR} \in (\mathbf{N})^{n\times q}$ are the post and pre incidence matrices ($\mathbf{N}$ is the set of non-negative integer numbers), and $W = W_{PO} - W_{PR}$ is the incidence matrix. $<G, M_I>$ is a PN system with initial marking M_I and $M \in (\mathbf{N})^n$ represents the PN marking vector. Each marking M represents a state of the DES. A transition t_j is enabled at marking M if its enabling degree $n(t_j,M) = \min\{\lfloor m_k / w^{PR}_{kj} \rfloor : p_k \in {}^\circ t_j\}$ satisfies $n(t_j,M) > 0$, where ${}^\circ t_j$ stands for the set of t_j upstream places, $m_k = M(p_k)$ is the marking of place p_k, w^{PR}_{kj} is the entry of matrix W_{PR} in row k and column j. This is denoted as $M[t_j>$. When t_j fires once, the marking varies according to $\Delta M = M' - M = W(:, j)$, where $W(:, j)$ is the column j of incidence matrix and M' the new marking reached after the firing of t_j. This is denoted by $M[t_j > M'$ or equivalently by $M' = M + W.X_j$ where X_j represents the firing count vector of transition t_j (David and Alla 1992). Each transition firing represents an event (controlled or unexpected) that changes the state of the DES. A firing sequence σ is defined as $\sigma = t_{j1}\, t_{j2}\ldots t_{jh}$ where $j_1,\ldots j_h$ are the indexes of the transitions. $X(\sigma) \in (\mathbf{N})^q$ is the firing count vector associated to σ, $|\sigma| = h$ is the length of σ, and $\sigma = \varepsilon$ stands for the empty sequence. The firing sequence σ fired at M leads to the trajectory (σ,M):

$$(\sigma,M) = M(0)\,[t_{j1} > M(1)\ldots.\, M(h-1)\,[t_{jh} > M(h)\ldots\ldots(1)$$

where $M(0) = M$, $M(1),...,M(h-1)$ are the intermediate markings and $M(h)$ is the final marking (in the next, we write $M(k) \in (\sigma,M),\ k = 0,\ldots h$).

Timed PNs are PNs whose behaviors are constrained by temporal specifications (David and Alla, 1992). For this reason, timed PNs have been intensively used to describe DESs like production systems (Cassandras, 1993). This paper concerns timed PNs (TPNs) where the time specifications are similar to the one used for T-timed PNs (Ramchandani, 1973): for any $t_j \in \boldsymbol{T}$, the firing of t_j occurs after a minimal delay $d_{min\,j}$ from the date it has been enabled. $D_{min} = (d_{min\,j}) \in (\mathbf{R}^+)^q$ ($\mathbf{R}^+$ is the set of non-negative real numbers) is the vector of time specifications for the transitions. The time semantic is completely defined with infinite server as server policy, preselection as choice policy (for control application, preselection is decided by the controller), enabling memory as memory policy. Moreover, the considered TPNs systems $<G, D_{min}, M_I>$ are assumed to behave under earliest firing policy: once a transition is decided to fire, it fires **at earliest** (assumption **A2**).

A timed firing sequence σ of duration τ_h is defined as $\sigma = (t_{j1},\ \tau_1)\ (t_{j2},\ \tau_2)\ldots(t_{jh},\ \tau_h)$ where $j_1,\ldots\ j_h$ are the labels of the transitions and $\tau_1,\ldots,\ \tau_h$ represent the dates of the firings that satisfy $0 \le \tau_1 \le \tau_2 \le \ldots \le \tau_h$. The timed firing sequence σ fired at M leads to the timed trajectory (σ,M):

$$(\sigma,M)=M(0)[(t_{j1},\ \tau_1)>M(1)\ldots.>M(h-1)[(t_{jh},\ \tau_h)>M(h) \quad (2)$$

and $M(0) = M$. Under earliest firing policy, an untimed trajectory (1) can be transformed in a straightforward way into a timed trajectory (2) using Algorithm 1. This algorithm is based on a calendar $CAL(M(k)) = \{(t_{j1},\ d_{j1}),\ (t_{j2},\ d_{j2}),\ldots\}$ attached to each intermediate marking $M(k) \in (\sigma,M),\ k = 0,\ldots h$, that contains the list of the enabled transitions t_{j1}, $t_{j2},\ldots$and their minimal delay d_{j1}, $d_{j2},\ldots$ for marking $M(k)$. Note that the same transition may appear several time in the $CAL(M(k))$ with same or different delays when it is enabled several times at $M(k)$. The calendar is initialized assuming that the trajectory starts at date 0 and that no transition is enabled before 0 (line 1). The calendar is updated for each marking of the trajectory (lines 4 – 19).

Algorithm 1

(Inputs: σ, M, G, D_{min},; Outputs: σ', τ)

1. initialization: $\tau \leftarrow 0$; $CAL \leftarrow \{(t_j,\ d_{min\,j})$ st $M[t_j>\}$,
   $\sigma' \leftarrow (\varepsilon,0)$, $h \leftarrow |\sigma|$
2. for k from 1 to h
3. find in CAL the delay d_k of the earliest occurrence of the
   k^{th} transition t_{jk} in σ
4. $\tau \leftarrow \tau + d_k$, remove entry $(t_{jk},\ d_k)$ in CAL
5. $CAL_{new} \leftarrow \emptyset$, $M' \leftarrow M - W_{PR}.X(t_{jk})$
6. for all t' st $M'[t'>$
7. compute the enabling degree $n'(t',M')$ of t' at M'
8. for j from 1 to $n'(t',M')$
9 find the j^{th} occurrence (t', d'_j) of t' in CAL
10. $CAL_{new} \leftarrow CAL_{new} \cup (t,'\ \max(0,\ d'_j - d_k))$
11. end for
12. end for
13. $M'' \leftarrow M' + W_{PO}.X(t_{jk})$
14. for all t'' st $M''[t''>$
15. compute the enabling degree $n''(t'',M'')$ of t'' at M''
16. for j from 1 to $n''(t'',M'') - n'(t'',M')$
17. $CAL_{new} \leftarrow CAL_{new} \cup (t'',\ d_{min}(t''))$
18. end for
19. end for
20. $CAL \leftarrow CAL_{new}$, $\sigma' \leftarrow \sigma'\ (t_{jk},\ \tau)$, $M \leftarrow M''$
21. end for

TIMED EXTENDED REACHABILITY GRAPHS

A marking M is said ***reachable*** from initial marking M_I if there exists a firing sequence σ such that (st) $M_I[\sigma>M$ and $\boldsymbol{S}(M_I)$ is the set of all reachable markings from M_I. The considered TPN systems $<G, D_{min}, M_I>$ are assumed to be bounded (assumption **A2**). As a consequence, $\boldsymbol{S}(M_I)$ is of finite cardinality N and the transition matrix $\Omega \in (\boldsymbol{T})^{N\times N} \cup \{\varepsilon\}$ is defined st for all $(M, M') \in \boldsymbol{S}(M_I) \times \boldsymbol{S}(M_I)$, $\Omega(M, M') = t_j$ if $M[t_j>M'$ otherwise $\Omega(M, M') = \varepsilon$.

In this paper an extended reachability set $\boldsymbol{S}_E(M_I) = \{(M, CAL),\ M \in \boldsymbol{S}(M_I)\}$ is introduced that includes not only the markings but also the delays required to fire the transitions under earliest firing policy. Note that a given marking may be associated to several calendars. A Timed Extended Reachability Graph (TERG) is defined consequently. It is computed in a systematic way with Algorithm 2 that returns the extended transition matrix $\Omega_E \in ((\boldsymbol{T})^{NE\times NE}\times(\mathbf{R}^+)^{NE\times NE}) \cup$

$\{(\varepsilon,0), (\varepsilon,\infty)\}$ st for all $(S, S') \in \boldsymbol{S_E}(M_I) \times \boldsymbol{S_E}(M_I)$, $\Omega_E(S, S') = (t_j, d_j)$ if $M(S)$ $[t_j > M(S')$ and d_j is the delay to fire t_j at earliest. Otherwise, $\Omega_E(S, S) = (\varepsilon, 0)$ and $\Omega_E(S, S') = (\varepsilon, \infty)$ for $S \neq S'$.

Algorithm 2

(Inputs: M_I, G, D_{min},; Outputs: $\boldsymbol{S_E}$, Ω_E)

1. initialization: $M \leftarrow M_I$; $CAL \leftarrow \{(t_j, d_{min\,j})$ st $M [t_j >\}$, $\boldsymbol{S_E} \leftarrow (M, CAL)$, $\Omega_E \leftarrow (\varepsilon, 0)$
2. while $\exists\ S \in \boldsymbol{S_E}$ that is not explored and st $CAL(S) \neq \varnothing$
3. for each t in $CAL(S)$
4. find the delay d of the earliest firing of t in $CAL(S)$
5. compute M' such that $M(S) [t > M'$
4. compute CAL' in M' with Algorithm 1
5. $S' \leftarrow (M', CAL')$
6. if $\forall\ S'' \in \boldsymbol{S_E}, S'' \neq S', \boldsymbol{S_E} \leftarrow \boldsymbol{S_E} \cup S'$, end if
8. $\Omega_E(S, S') \leftarrow (t, d)$
9. end for
10. end while
11. Complete all non-defined entries of Ω_E st $\Omega_E(S, S) \leftarrow (\varepsilon, 0)$ and $\Omega_E(S, S') \leftarrow (\varepsilon,\infty)$ if $S \neq S'$.

Example 1: Let us consider TPN1 in Fig.1 as an example of timed Petri net that illustrates the difference between the usual RG and the TERG. The initial marking is $M_I = (2\ 0\ 0\ 0)^T = 2p_1$ and the time specifications are defined by $D_{min} = (1\ 5\ 1\ 1\ 1)^T$.

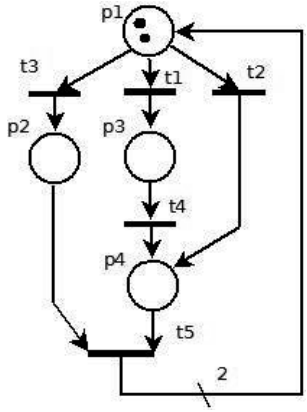


Figure 1: TPN1

$\boldsymbol{S}(M_I)$ is composed of 10 different markings and the usual RG is reported in Fig.2. $\boldsymbol{S_E}(M_I)$ has 12 different states and the TERG is reported in Fig.3. The edges from one state S to another one S' are tagged not only with the transition t such that $M(S) [t > M(S')$ but also with the duration d required to fire t at S. One can notice that the markings $1p_1 1p_4$ and $1p_3 1p_4$ appear twice in Fig. 3. The reason is that the calendars for states S_3 and S_8 differ in terms of firing delays: $CAL(S_3) = \{(t_1,0), (t_2,0), (t_3,0)\}$ and $CAL(S_8) = \{(t_1,0), (t_2,3), (t_3,0)\}$. The same holds for S_6 and S_9: $CAL(S_6) = \{(t_4,0)\}$ and $CAL(S_9) = \{(t_4,1)\}$.

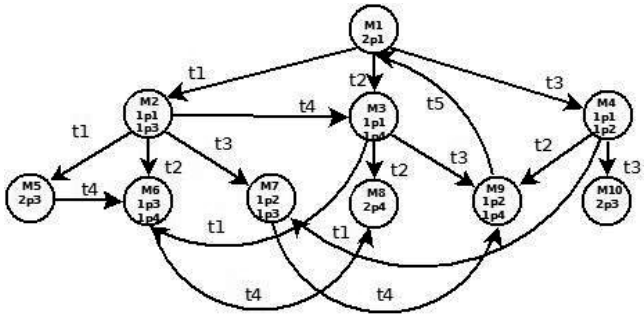


Figure 2: Usual reachability graph of TPN1

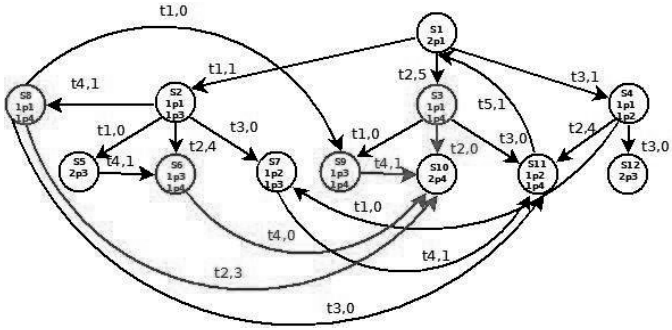


Figure 3: Timed Extended Reachability Graph of TPN1

Property 1 formalizes that the TERG defined by Algorithm 2 is an exact representation of the behaviour of a TPN system.

Property 1: Let us consider a TPN system $<G, D_{min}, M_I>$ that satisfies assumptions **A1** and **A2**. Any timed trajectory (σ, M_I) of length h and form (2) is feasible, in $<G, D_{min}, M_I>$ if and only if there exists a path $Path = S(0)\ S(1)\ \ldots\ S(h)$ in the TERG with $M(S(0)) = M_I$ and $M(S(k)) = M(k)$, $k = 1,\ldots,h$ that involves the same sequence of transitions and where the delay between two successive state $S(k)$ and $S(k+1)$ is exactly $\tau_k - \tau_{k+1}$.

Proof: Let us first consider a feasible timed trajectory $M_I [(t_{j1}, \tau_1) > M(1) \ldots. > M(h-1) [(t_{jh}, \tau_h) > M(h)$ in $<G, D_{min}, M_I>$. This trajectory is encoded in the TERG by construction: the TERG is initialized at $S(0)$ such that $M(S(0)) = M_I$ and $CAL(S(0)) = \{(t_j, d_{min\,j})$ st $M [t_j >\}$. The first part of the trajectory $M_I [(t_{j1}, \tau_1) > M(1)$ is feasible, $\tau_1 = d_{min\,j1}$ and there exists a state $S(1) \in \boldsymbol{S_E}(M_I)$ st $M(S(1)) = M(1)$ and $\Omega_E(S(0), S(1)) = (t_{j1}, d_{min\,j1})$. The calendar $CAL(S(1))$ is updated with Algorithm 1 in order to sum up the earliest occurrence of all transitions enabled at $M(S(1))$. The complete proof is obtained in an iterative way. The trajectory $M_I [(t_{j1}, \tau_1) > M(1) [(t_{j2}, \tau_2) > M(2)$ is also feasible, and $\tau_2 - \tau_1$ is the minimal delay to wait before firing t_{j2}. There exists a state $S(2) \in \boldsymbol{S_E}(M_I)$ st $M(S(2)) = M(2)$ and $\Omega_E(S(1), S(2)) = (t_{j2}, \tau_2 - \tau_1)$. The same holds for the rest of the timed trajectory.

Reciprocally, consider a path $Path = S(0)\ S(1) \ldots S(h)$ in the TERG with $M(S(0)) = M_I$ and note $\Omega_E(S(k-1), S(k)) = (t_{jk}, d_k)$ for $k = 1,\ldots,h$. Then, the first part $S(0)\ S(1)$ of $Path$ corresponds to a trajectory $M(S(0)) [(t_{j1}, d_1) > M(S(1))$ that is feasible in $<G, D_{min}, M_I>$ because (a) t_{j1} is enabled at $M(S(0))$, (b) d_1 is the minimal delay to wait before firing t_{j1}

($d_1 = d_{min\,j1}$). The path is then completed in an iterative way : *S(0) S(1) S(2)* corresponds to a trajectory $M(S(0))$ $[(t_{j1}, d_1) > M(S(1))$ $[(t_{j2}, d_2) > M(S(2))$ that is feasible in $<G, D_{min}, M_I>$ because (a) t_{j2} is enabled at $M(S(1))$, (b) d_2 is the minimal delay to wait before firing t_{j2}. The same holds for the rest of the path.□

Property 2 formalizes the obvious statement that the complexity to build the TERG of a TPN system is higher than the complexity to build its usual reachability graph. But this property ensures also that the set $\boldsymbol{S_E(M_I)}$ remains of finite cardinality as long as $\boldsymbol{S(M_I)}$ is of finite cardinality.

Property 2: If $\boldsymbol{S(M_I)}$ is of finite cardinality N, then $\boldsymbol{S_E(M_I)}$ is of finite cardinality N_E that satisfies:

$$N \leq N_E \leq kqN^2 \qquad (3)$$

Proof: Each state $S \in \boldsymbol{S_E(M_I)}$ is composed of a marking $M(S)$ and a calendar $CAL(S)$. By assumption **A2**, the set of markings $\boldsymbol{S(M_I)}$ is of finite cardinality. To prove that $\boldsymbol{S_E(M_I)}$ is also of finite cardinality, it is sufficient to prove that each marking $M \in \boldsymbol{S(M_I)}$ is associated to a set of calendars $\boldsymbol{CAL(M)}$ of finite cardinality. M enables at most q transitions and each enabled transition has an enabling degree no larger than k (assumption **A1**). Thus there are at most kq different durations in each state of the TERG. Let us consider two nodes S_1 and S_2 such that $M(S_1) = M(S_2) = M$, and the two associated calendars: $CAL(S_1)$ and $CAL(S_2)$. These two calendars contain exactly the same list of transitions, only their remaining firing delays are different. For a given transition t in calendar the remaining firing delay d satisfies: $d \in [0 : d_{min}]$. The key point is that t is enabled from the date the system enters in a new marking. As long as the number of different markings is finite, the number of possible dates is also finite, and consequently the number of possible remaining firing delays too. This last number does not exceed N. Thus N_E does not exceed $N \times kqN$. Moreover, N_E equals at least N and (3) holds.□

In some particular cases, $\boldsymbol{S_E(M_I)}$ coincides with $\boldsymbol{S(M_I)}$. Property 3 provides a sufficient condition to ensure that $N_E = N$. For this purpose, let us introduce the set of markings $\boldsymbol{M_{PR}} = \bigcup_{t \in T} \boldsymbol{M_{PR}}(t)$ with $\boldsymbol{M_{PR}}(t) = \{M \in (\mathbf{N})^n$ st $M = M_{in} - W_{PR}.X(t_{in})$ with $M_{in} \in S(M_I)$ and $t_{in} \in \boldsymbol{M_{in}}^\circ\}$ and $\boldsymbol{M_{in}}^\circ$ the set of transitions enabled at M_{in}.

Property 3: $\boldsymbol{S(M_I)}$ coincides with $\boldsymbol{S_E(M_I)}$ (i.e. $N_E = N$) if for all $M \in \boldsymbol{M_{PR}}$ and for all $t \in \boldsymbol{M}^\circ$, $\min(M - W_{PR}.X(t)) < 0$.

Proof: Let us consider a marking $M \in \boldsymbol{M_{PR}}$ and a transition $t \in \boldsymbol{M}^\circ$ such that $\min(M - W_{PR}.X(t)) < 0$. Let introduce also the marking M' reached by the firing of t: $M [t > M'$. The condition $\min(M - W_{PR}.X(t)) < 0$ means that the firing of t disables all other transitions and that the remaining firing delays of all transitions enabled at M' equal exactly the parameters in D_{min} (according to the enabling memory policy) Consequently the calendar associated to the marking M' is unique. If the property $\min(M - W_{PR}.X(t)) < 0$ holds for all $M \in \boldsymbol{M_{PR}}$ and for all $t \in \boldsymbol{M}^\circ$, then all markings in $\boldsymbol{S(M_I)}$ are associated to a single calendar and $\boldsymbol{S(M_I)}$ coincides with $\boldsymbol{S_E(M_I)}$. □

CASE STUDY

TPN2 (Fig. 4) is the timed model of a manufacturing system that produces two types of products according to two jobs (Chen, 2011). The first job is defined by transitions t_1 to t_8 and the second one is defined by transitions t_9 to t_{14}. The six resources p_{14} to p_{19} have limited capacities: $m(p_{14}) = m(p_{16}) = m(p_{17}) = 1$ and $m(p_{15}) = m(p_{18}) = m(p_{19}) = k$ The marking of places p_1 and p_8 represent respectively the number of products that can be simultaneously processed by job 1 or job 2: $m(p_1) = m(p_8) = m$. The temporal specifications are given by $D_{min} = (1\ 1\ 2\ 1\ 2\ 1\ 1\ 1\ 3\ 3\ 3\ 3\ 3\ 3)^T$.

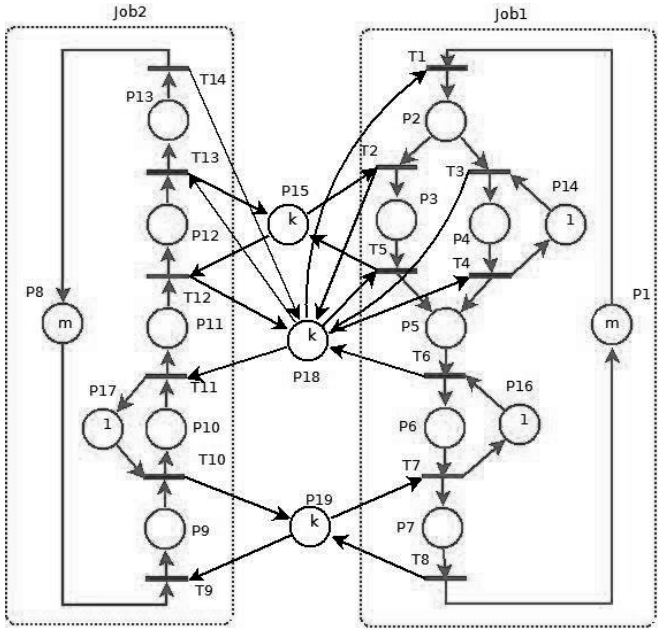


Figure 4: TPN2 model of a manufacturing system (Chen 2011)

Table 1 illustrates the variation of the size N_E of $\boldsymbol{S_E(M_I)}$ compared to the size N of $\boldsymbol{S(M_I)}$ with respect to m and k. For $k = 1$, one can notice that the size of both reachability sets tends to a limit value because the number of products that are simultaneously in progress in both jobs cannot exceed 5 when few resources are available. For $k > 1$, this value increases and one can notice that the cardinality of both reachability sets increases rapidly with respect to m and k. But N_E increases much faster than N.

Table 1: Ratio N_E / N for TPN2 in function of m and k

m / k	1	2	3
1	66/35 = 1.89	222/42=5.29	928/282=3.29
2	558/162 = 3.44	2925/432=6.77	18541/509=36
3	874/257 = 3.40	16237/1632=9.95	…
4	925/280 = 3.30	….	
5	928/282 = 3.29		

CONCLUSIONS

The contribution of this paper was to design a Timed Extended Reachability Graph that encodes both the time specifications and earliest firing policy. Such a graph was proved to describe in an exhaustive way all timed feasible trajectories of the considered DES and to remain of finite size as long as the usual reachability set of the system is finite. Consequently, it can be used for model checking, and behavioral analysis. In our opinion, it can also be used for control and scheduling applications. One of our future research directions is to study such applications. Search algorithms based on Dijkstra, A^* or other one will be used for that purpose. Nevertheless, the computation cost of the TERG should be considered with attention. This is another challenge to be addressed in the next future.

ACKNOWLEDGEMENTS

The Project MRT MADNESS 2016-2019 has been funded with the support from the European Union with the European Regional Development Fund (ERDF) and from the Regional Council of Normandie.

REFERENCES

K.R. Baker, D. Trietsch, *Principles of Sequencing and Scheduling*, John Wiley & Sons, 2009.

B. Berthomieu and M. Menasche. An Enumerative Approach for Analyzing Time Petri Nets. In IFIP Congress, pages 41– 46, 1983.

B. Berthomieu and F. Vernadat. State Class Constructions for Branching Analysis of Time Petri Nets. In TACAS 2003, vol. 2619 of *LNCS*, pages 442–457. Springer, 2003.

C. Cassandras, *Discrete Event Systems: Modeling and Performances Analysis*, Aksen Ass. Inc. Pub., 1993.

Y. Chen, Z. Li, M. Khalgui and O. Mosbahi, Design of a Maximally Permissive Liveness-Enforcing Petri Net Supervisor for Flexible Manufacturing Systems, *IEEE Trans. Aut. Science and Eng.*, 8(2): 374-393, 2011.

R. David and H. Alla, *Petri nets and grafcet – tools for modelling discrete events systems*, London: Prentice Hall, 1992.

Y. Feng, K. Xing, Z. Gao, and Y. Wu, Transition cover-based robust petri net controllers for automated manufacturing systems with a type of unreliable resources, *IEEE Trans. on Systems, Man, and Cybernetics*: *Systems*, 1–11, 2016.

G. Gardey, O. H. Roux, and O. F. Roux. Using Zone Graph Method for Computing the State Space of a Time Petri Net. In *FORMATS 2003, volume 2791 of LNCS*, pages 246–259. Springer, 2003.

M.D. Jeng, S.C. Chen, Heuristic search approach using approximate solutions to Petri net state equations for scheduling flexible manufacturing systems. *Int J FMS*, vol. 10, no. 2, pp. 139–162, 1998.

K. Klai, N. Aber, L. Petrucci, A New Approach To Abstract Reachability State Space of Time Petri Nets, 20th International Symposium on Temporal Representation and Reasoning, 2013.

D. Lefebvre and E. Leclercq, Control design for trajectory tracking with untimed Petri nets, *IEEE Trans. Aut. Contr.*, vol. 60(7), pp. 1921-1926, July 2015.

D. Lefebvre, Approaching minimal time control sequences for timed Petri nets, *IEEE Trans. Automation Science and Engineering*, vol. 13, no. 2, pp. 1215-1221, 2016a.

D. Lefebvre, Deadlock-free scheduling for Timed Petri Net models combined with MPC and backtracking, Proc. IEEE WODES 2016, Invited session "Control, Observation, Estimation and Diagnosis with Timed PNs", pp. 466-471, Xi'an, China, 2016b.

H. Lei, K. Xing, L. Han, F. Xiong, Z. Ge, Deadlock-free scheduling for flexible manufacturing systems using Petri nets and heuristic search, *Computers & Industrial Engineering*, vol. 72, pp. 297–305, 2014.

Y-T. Leung, *Handbook of Scheduling: Algorithms, Models, and Performance Analysis*, Chapman & Hall/CRC Computer & Information Science Series, 2004.

D. Lime and O. H. Roux. Model Checking of Time Petri Nets Using the State Class Timed Automaton. *Discrete Event Dynamic Systems*, 16(2):179–205, 2006.

P. Lopez, F. Roubellat, Production Scheduling, ISTE/Wiley, London, April 2008.

Q. Wang, Z. Wang, Hybrid Heuristic Search Based on Petri Net for FMS Scheduling, *Energy Procedia*, vol. 17 pp. 506 – 512, 2012.

BIOGRAPHY

Dimitri Lefebvre graduated from the Ecole Centrale of Lille (France) in 1992. He received his PhD in Automatic Control and Computer Science from University of Sciences and Technologies, Lille in 1994, and an HDR from University of Franche Comté, Belfort, France in 2000. Since 2001, he has been a Professor at Institute of Technology and Faculty of Sciences, University Le Havre, France. He is with the Research Group on Electrical Engineering and Automatic Control (GREAH) and from 2007 to 2012 he was the head of the group. His current research interests include Petri nets, learning processes, adaptive control, fault detection and diagnosis and their applications to electrical engineering.

Visual Estimation of Persistence in Time Series

Sihle Poswayo
Nelson Mandela University
Port Elizabeth 6031, South Africa
email: sihle.poswayo@nmmu.ac.za

Igor Litvine
Nelson Mandela University
Port Elizabeth 6031, South Africa
email: igor.litvine@nmmu.ac.za

KEYWORDS
Persistence, Hurst Exponent, Paired Comparisons.

ABSTRACT

In this paper we suggest to use subjective judgements to measure persistence in time series by comparing pairs of graphs with different Hurst exponent. The group of respondents consisted of 30 volunteers who were asked to identify which of two presented graphs is more jagged (that is, less persistent). The graphs were simulated using time series package of Mathematica®. The responses were processed using algorithm based on the Thurstone-Mosteller model for paired comparisons. The results of the analysis show that human eye is capable of distinguishing graphs of time series with Hurst exponents difference as small as only 0.02.

INTRODUCTION

Persistence is an important dynamic property of any time series as it provides an understanding of the behavior of this time series. The study of persistence has received significant attention since the early works of B.Mandelbrot (Mandelbrot (1969), Mandelbrot (1972)). Persistence forms the focus of many research in several fields, including Hydrology, Health Sciences, Finance and Econometrics, just to name a few. Persistence in this context, refers to the quality of a time series to keep the direction of change.

One of possible ways to understand persistence is to see it as opposite to jaggedness. More jagged series are less persistent and vice versa. Informally, it is defined as an ability of a time series to "remember" past observations to a greater or lesser extent and to follow past patterns of behaviour. Mandelbrot (1969) defined persistence as a "tendency for large positive (or negative) values to be followed by large values of the same sign".

There are various methods of studying persistence, most commonly, persistence of an observed series is defined in terms of Hurst exponent, persistence strength (less often) and the newly introduced l-parameter (Litvine and Gorshkov (2016)). The Hurst Exponent is one of the most popular parameters that is mentioned in scientific literature, although it contains many inconsistencies and contradictions. There is still much unclarity as to whether the Hurst exponent measures persistence correctly (see for example Litvine (2014)). If applied in medicine, it could result in the wrong diagnosis of illnesses such as cancer and heart disease.

Different tools exist in measuring persistence. The oldest and best-known is the so-called rescaled range (R/S) analysis popularized by (Mandelbrot and Wallis (1969a), Mandelbrot and Wallis (1969b)) and based on previous hydrological findings of Hurst (1951). Alternatives include Detrended Flactuation Analysis (DFA) (Peng et al. (1994), Matos et al. (2008)), Fractional Differencing and Periodogram Regression (Geweke and Porter-Hudak (1983)), Aggregated Variances (Beran (1994)), Gaussian Semi-parametric Estimation of Long Range Dependence (Robinson (1995)), Wavelet Analysis (Simonsen and Nes (1998)), Stabilogram Diffusion Analysis (SDA) (Gorshkov (2012)).

In this paper we suggest a visual method of assessing persistence. Human eye may be used to judge on jaggedness and therefore on persistence as well. Today, visual measurements are quite common. In engineering and other sciences, numerous visual comparisons are carried out (Beer et al. (2002)).

When studying graphs of time series, the difference between persistent and anti-persistent series is noticeably clear. In this study, we aim to answer to which extent the human eye is capable of accurately measuring persistence. We also use the method of paired comparisons (David (1988)) to process the responses of the volunteers, who participated in visual measurements of jaggedness.

Hurst Exponent

The persistence analysis has its roots in early works of the British hydrologist Harold Edwin Hurst, who introduced a parameter to investigate dependence properties of phenomena such as levels of the Nile River's volatile rain and drought conditions that had been observed over a long period of time (Hurst (1951)).

When Hurst examined capacity $R(s)$ as a function of s successive discharges for the Nile River, he divided $R(s)$ by the standard deviation $S(s)$ as a function of s successive discharges. Hence, the analysis is called rescaled-range analysis (R/S analysis). He discovered that for small values of s, the rescaled range $R(s)/S(s)$ is proportional to s^H with H being a constant between 0.5 and 1 and he judged H to be near 0.7, while Mandelbrot and Wallis (1969a) found cases where the best estimate of H is below 0.5, contradicting Hurst's claim that $0.5 < H < 1$.

The Hurst exponent H, is defined in terms of the asymptotic behaviour of the Rescaled-range as the function of the time span of a time series (Rasheed and Qian (2004) , Feder (1988)). The Hurst exponent is also defined to be a real number in the interval $H\epsilon(0, 1)$, as it corresponds to a fractal dimension between 1 and 2. The Hurst exponent (H) is a statistical measure used to classify time series. The values of the Hurst exponent vary between 0 and 1, with higher values indicating a smoother trend, less volatility, and less roughness.

Hurst exponent and the fractal dimension are independent of each other in principle (Gneiting and Schlather (2004)), nevertheless, the two notions are closely linked in most of the scientific literature. Mandelbrot (Mandelbrot (1972) , Mandelbrot (1985)) has shown that the Hurst exponent H is directly related to the fractal dimension D, therefore, the Hurst exponent can be converted into a fractal dimension using the following formula:

$$D = 2 - H \tag{1}$$

The fractal dimension D, of a surface is a measure of roughness, with $D\epsilon[m, m + 1)$ for a surface with m-dimensional space and higher values indicating rougher surfaces (Gneiting and Schlather (2004)). The fractal dimension of a line is 1, and of a geometric plane is 2. Thus, the fractal dimension of a random walk would be somewhere half-way between a line and a plane.

Another parameter that is related to the Hurst exponent is a persistence strength β and the relation between H and β is:

$$\beta = 2H + 1 \tag{2}$$

Malamud and Turcotte (1999) mentioned, that the equation 2 works most accurately for interval $0.2 \leq H \leq 0.6$.

The Hurst exponent is a useful parameter for surface analysis, also, it can be used to assess memory of a time series i.e. correlation between the increments. The impact of the present on the future can be expressed as a correlation:

$$C = 2^{(2H-1)} - 1 \tag{3}$$

where C is the correlation measure and it is not related to the auto-correlation function (ACF), H is the Hurst exponent parameter (Peters (1996)).

Classes of Hurst Exponent

There are three distinct classes of the Hurst exponent (H);

1. $0 < H < 0.5$ (anti-persistence)

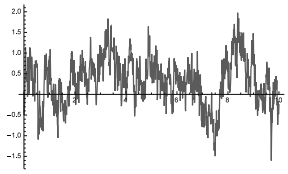


Figure 1: Example of time-series with $H = 0.25$ showing anti-persistence (more jagged)

This is an anti-persistent type of the series, the strength of this anti-persistent behaviour depends on how close H is to zero. The closer H is to zero, the closer C in equation 3 moves toward -0.5, or negative correlation (Peters (1996)). If the system has went up in the previous period, it is more likely to go down in the next period and vice versa. This kind of series is more volatile, because it would consist of frequent reversals. So, it behaves similar to mean reverting stationary process.

2. $H = 0.5$ (Random walk)

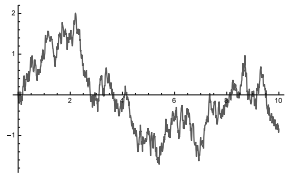


Figure 2: Example of time-series with $H = 0.5$ showing random walk

This type of a series is called a random walk because its increments are random and uncorrelated, in this case, C in equation 3 equals zero . In other words, the present does not influence the future.

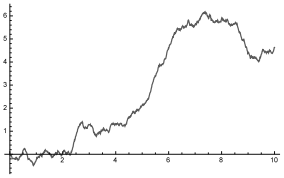


Figure 3: Example of time-series with $H = 0.75$ showing persistence (less jagged)

3. $0.5 < H < 1$ (persistence)

   Here we have a persistent, or trend-reinforcing series and it's trends are apparent. The strength of it's persistence, increases as H approaches 1 or 100% positive correlation in equation 3 (Peters (1996)). The closer H is to 0.5, the less defined its trends will be.

   The graphs become smoother and less jagged as H increases, and the range of cumulative values increase with H. If the system went up (down) in the previous period, it is more likely to continue in that manner in the next period. Mandelbrot suggested to describe such series with fractional brownian motion, or biased random walks.

Hurst exponent H, is often referred to as the index of dependence, it always lies in the interval $0 < H < 1$ and $H = 0.5$ for processes that have independent increments. Lately, particular interests focuses on the hypothesis that $0.5 < H < 1$, indicating relatively long range dependence (Torsten (2002)). Peters (1996) mentioned that persistent time series are the more interesting class because, Hurst found that they yield great quantities in nature.

Uses of Hurst Exponent

The Hurst exponent has been applied in many research fields such as hydrology (Hurst (1951)). Also, used in fractal analysis (Mandelbrot (1982) , Mandelbrot and Van Ness (1968) , Mandelbrot (1985)). The value of the Hurst exponent (H), in a time series may be interpreted as an indicator of the irregularity of the price of a commodity and currency in finance (Rasheed and Qian (2004)), it has recently become popular in health sciences as an indicator of heart failure or similar quantity (Litvine and Gorshkov (2016)).

The Hurst exponent provides a measure for predictability (Rasheed and Qian (2004)). It also measures jaggedness in time series and surfaces (Peters (1996)). It is used for quantifying the fractal features of LANDSAT images (Valdiviezo-N et al. (2014)).

Estimating the Hurst Exponent

Mandelbrot built the first formal mathematical model which is known as the Fractional Brownian Motion (FBM). He used the Hurst exponent, named after the British hydrologist, Harold Edwin Hurst, as one of the parameters (Mandelbrot (1969) , Mandelbrot (1972)). Many research methodologies have been developed through the years in an attempt to improve the mathematical models used for reliable estimation of the estimated Hurst exponent.

The Rescaled Range (R/S) is currently the most extensively used method for measuring of the Hurst exponent and persistence in time series. This method was originally developed by Hurst (1951). More recently Lo (2007) argued that the statistical R/S analysis used by Mandelbrot (1972) and Greene and Fielitz (1997) was not very reliable, and as a result, introduced an improved statistical R/S analysis.

Since then, a couple of empirical studies made use of Lo's modified R/S analysis. For example, Cheung (1993) tried to contribute alternative evidence from the perspective of long memory analysis by using Lo's modified R/S and fractional differencing test. The results of Sadique and Silvapulle (2001), achieved through R/S analysis, fractional differencing testing, and time and frequency domain versions of the score tests, contradict those of Cheung (1993).

Willinger and Teverovsky (1999) have criticised Lo's rescaled R/S test that it is inconclusive. They showed numerically that even for a long memory time series with a moderate value of the Hurst exponent like $H = 0.6$ the Lo's test cannot reject the null hypothesis of short range dependence.

Alternatives includes Detrended Flactuation analysis (DFA). Empirical tests can find that DFA analysis performs perfectly sufficient and even better than other alternative measures (Peng et al. (1994)), and also it asymptotically provides good results for stationary time series, however it appears that DFA cannot provide protection against non-stationaries. DFA has the advantage over the standard variance analysis of being able to detect long-term dependence in non-stationary time series (Peng et al. (1994)).

DFA can find values for $H > 1$, is often thought to provide evidence that DFA is superior and is considered to be a serious drawback of FA. The fact that DFA can estimate $H > 1$ reveals that, in such cases, de-trending is not actually performed (Peng et al. (1994)). If a time

series is not a fractional Brownian motion, DFA can give values of H that are out of the domain. However, todays researchers use Hurst exponent without verification whether the time series is a fractional Brownian motion or not. If the process does not satisfy the model, you may see results that are outside the domain.

Econometric and Financial empirical studies of long term dependence, often rely on the studies of Geweke and Porter-Hudak (1983), who developed a method for the calculation of the fractional differencing. These long memory measure models are now widely applied in estimating persistence characteristics of various time series (Cornelis and Yu (2008) , NyoNyo et al. (2006)).

Some of the conducted research pay attention to time series models. These include GARCH and IGARCH models which are used to estimate the behaviour of financial stock markets, for example Su and Fleisher Dongwei and Belton (1998). However, Kyaw NyoNyo et al. (2006) criticised these time series models for not being able to model long-term dependence/persistence satisfactory.

In health sciences, heart diseases are among major causes of death. That is why it is very important to develop accurate methods of early diagnostics of heart disorders. Litvine and Gorshkov (2016) used stabilogram diffusion analysis (SDA) and a new technique based on empirical persistence (EP), analysing persistence of RR-interval records for 15 subjects of which five were healthy patients, five patients with congestive heart failure and the last five patients with atrial fibrillation.

Utilising the Hurst exponent, one of the most popular parameter which is used for measuring persistence (Malamud and Turcotte (1999)). Additionally, they calculated the persistence strength (β) using the Fourier transform of the autocorrelation function of RR-intervals. After comparing all the suggested techniques objectively, The EP method showed the best results with lowest variability and certain differentiation between health and atrial fibrillation groups (Litvine and Gorshkov (2016)).

Persistence

Persistence or long-range dependence and fractal behaviour have been observed in amazing number of physical systems. Either phenomenon has been modelled by self-similar random functions, thereby implying a linear relationship between Hurst exponent, a measure of persistence or long-memory dependence and fractal dimension, a measure of roughness (Gneiting and Schlather (2004)).

(Mandelbrot (1969) , Mandelbrot (1972)) was one of the first few researchers to consider the importance of persistence for studying statistical dependence in asset returns. Since then, many empirical studies have contributed further to Mandelbrot's findings. Recent empirical financial market research has demonstrated that the Hurst exponent, a measurement tool of persistence, tends to provide a good characterization of the scaling characteristics for financial markets (NyoNyo et al. (2006) , Cornelis and Yu (2008)).

Persistent time series are defined as any time series that has Hurst exponent value of $0.5 < H \leq 1$, are fractal because they are also related to fractional Brownian motion, because in fractional Brownian motion, across time scales there is correlation between events (Peters (1996)). A persistent time series would result in a fractal dimension closer to a line as H gets closer to 1, that is smoother and less jagged than a random walk, as we saw in Figures 1, 2 and 3.

Anti-persistent time series is an opposite of persistent time series with a value of $0 < H < 0.5$, it would result in rougher surfaces with a higher fractal dimension and a more jagged series than a random walk, which is a system subject to more reversals and this precisely represents an anti-persistent time series (Gneiting and Schlather (2004)).

Paired Comparison

Generally, Paired Comparison is any process of comparing entities in pairs to judge which of each entity is preferred, or has a greater amount of some quantity property, or whether or not the two entities are identical. The method of paired comparison is classified by the modern science as a ranking procedure for ordering objects with respect to certain characteristic, which may not be defined formally (e.g. attractiveness, successfulness, superiority, etc.) (David (1988)).

This is one of few formal mathematical models, which is based on subjective preferences of judges or experts. Judges are presented with objects from the given set in pairs, as a result, the judges express and record the object that they prefer the most from the two objects all depending on which model of paired comparisons is used and the nature of objects in the experiment.

One of the reasons the objects are compared in pairs, is for the judges (respondents) so they can concentrate their attention on only two objects at a time. This kind of approach is believed to avoid the effect of so called sensory fatigue (Litvine (2004)), most of the time it leads to confusion and lack of concentration when evaluating more than two objects at a time.

Furthermore, it permits better discrimination between very small differences in the objects when observing just two objects.

This method is a simple and powerful alternative to other ranking techniques, which may require complex experiment planning procedures. The method of paired comparison has been seen by far the most superior than other ranking techniques (Litvine (2014)). That is why the method has been used by most researchers from diverse disciplines, including behavioral and social sciences (Bradley (1976), Davidson and Farquhar (1976), Litvine (1999)) .

The model that will be used for this study only deals with preferences, i.e. the judge only says which object or graph he or she prefers between the two. The other models of paired comparisons give scores to the two objects that are being compared, for example, saying how many times one object is better than the other. Some paired comparison models allow for ties in the scores, while others do not.

In paired comparisons there are continuous models, discrete models and there are distribution-free models (Litvine (2004)). In this study we use Thurstone-Mosteller model because it is one of the models that only deals with preferences, i.e. it require responses of "Yes" or "No" which is the same as in this study where only respondents are required to select "A" or "B".

The analysis of experiments involving paired comparisons has received considerable attention in statistical methodology. The study of the method of paired comparison may be traced back to a very old publication of Fechner in 1860, which he used for psychometric investigations. Followed by Thurstone (1927a), who formalized the method and provided proper mathematical background. He conceived the approach for measurement known as the "law of comparative judgment" (Thurstone (1927a), Thurstone (1927c), Thurstone (1927b)). Subsequently Mosteller (1951) was a first statistician to make a breakthrough in this area; he developed a proper mathematical foundation to the method suggested by Thurstone. This resulted as the very first and most popular model of paired comparisons called Thurstone-Mosteller model.

Paired comparisons is also used in many other applications where objects are compared in pairs for other reasons e.g. in sport statistics, the way most of games work is that players of teams play against each other two at a time, and then a score is obtained from each game i.e. Tennis or soccer. Numerous other applications of the method of comparisons are also present e.g. military, economic to name a few.

Results

This study makes use of simulated data obtained using Mathematica® to produce a sample of graphs for the questionnaire used when conducting this study. We produce 95 pairs of graphs with different Hurst exponent values not visible on the graphs and each pair is identified by an I.D. Also the questionnaires will be identified by an I.D. Both the pair and questionnaire I.D are combined to form a unique I.D. of each response.

We asked 30 statistics students in the department of Statistics at Nelson Mandela University from two different third year classes of students to fill in the questionnaires. Firstly, the questionnaire was pre-tested on a small sample of respondents, then it was administered to a wider group of respondents to extend the reach of the research. An advantage of this method is that the data collection was immediate.

The respondents were required to choose the graph that is more jagged between the two provided on the graph document and make a tick on the block under the column box, next to the graph I.D. The respondents were also required to choose one graph even if they feel like they are equally jagged, so they still need to make a choice.

There data was divided into three groups, the first one with values of Hurst exponent at a difference of 10% and 20%, the second one at a difference of 5% and 15% and the last one at a difference of 2%. The collected data consists of 2871 observations where respondents judges which graph was more jagged between the two graphs.

We observed that the respondents made about 1437 correct guesses and that is about 50% of the total. We defined the first class as those respondents with I.D number from 501-5015 and the second class from 601-6015, The first class made 48% correct guesses and thats about half of the time. Whereas the second class made made 52% correct guesses.

In the table below, we can see that there were more correct guesses at 10% and 20% than at 5% and 15% Hurst exponent difference as expected, because it is easy to judge the jaggedness of the graphs at 10% and 20% but difficult when the difference is smaller. However, the number of correct guesses at 2% Hurst exponent difference is even higher than at 10% and 20%.

In table 2 we present results of the paired comparison analysis for the case of 10% difference in H. The estimates are are close to the true Hurst exponent

Table 1: Number of correct guesses and percentages.

Difference in H	No. of correct guesses	Percentage
10% and 20%	461	49%
5% and 15%	414	45%
2%	478	56%

values, however the model we used turned to overestimate for H below 0.5 and underestimate for H above 0.5. The most important fact here is that the objects are ranked in exact order as per Hurst exponent.

Table 2: Results of Thurstone-Mosteller model at 10% H difference.

True H value	Weight	Estimated H value
H = 0.1	16.5782	0.1
H = 0.2	13.7449	0.236726
H = 0.3	11.005	0.3689424
H = 0.4	9.59786	0.436845
H = 0.5	8.26148	0.501333
H = 0.6	6.92672	0.565744
H = 0.7	5.59196	0.630154
H = 0.8	2.77521	0.766079
H = 0.9	0.	0.9

In the next three tables, we present the results of the analysis for 3 cases of Hurst exponent: $0.15 \leq H \leq 0.35$, $0.4 \leq H \leq 0.6$ and lastly $0.65 \leq H \leq 0.85$.

When Hurst exponent values were between 0.15 and 0.35, we found that the there was an instance of incorrect order, specifically case $H = 0.19$ was identified as more jugged than $H = 0.18$. Otherwise all other rankings are correct.

From the table 4 ($0.4 \leq H \leq 0.6$ we see that the model over estimated the Hurst exponent values below 0.48 and under estimated above for $H > 0.5$. Also the ranking of the objects is in a correct order.

In table 5 ($0.65 \leq H \leq 0.85$), the model overestimated the Hurst exponent values below 0.75 and underestimated for above 0.77. However, the estimated values are also close to the true values. All objects are ordered correctly.

Conclusions

In this work we asked 30 respondents to fill in a questionnaire which contains pairs of graphs and we

Table 3: Results of Thurstone-Mosteller model at 2% H difference: $0.15 \leq H \leq 0.35$.

True H value	Weight	Estimated H value
H = 0.15	19.0675	0.15
H = 0.16	17.6579	0.164785
H = 0.17	14.9434	0.193259
H = 0.18	12.2014	0.222019
H = 0.19	12.2181	0.221844
H = 0.21	10.8827	0.235851
H = 0.23	9.54743	0.249857
H = 0.25	8.19794	0.264012
H = 0.27	6.76451	0.279047
H = 0.29	5.51709	0.292131
H = 0.31	4.13785	0.306598
H = 0.33	2.80309	0.320598
H = 0.35	0.	0.35

Table 4: Results of Thurstone-Mosteller model at 2% H difference between 0.4 and 0.6.

True H value	Weight	Estimated H value
H = 0.4	16.369	0.4
H = 0.42	13.6223	0.433561
H = 0.44	12.2454	0.450383
H = 0.46	10.8827	0.467032
H = 0.48	9.51841	0.483702
H = 0.5	8.26697	0.498992
H = 0.52	6.80998	0.516794
H = 0.54	5.45602	0.533337
H = 0.56	4.20966	0.548566
H = 0.58	2.80253	0.565758
H = 0.6	0.	0.6

asked them to choose in each pair a graph which is more jagged. While misjudgments were quite common (about 50% of the responses were incorrect), after processing the responses according to the Thurstone-Mosteller model for paired comparisons, we receive highly reliable inference. We had only one incorrect ordering (for $H = 0.18$ and $H = 0.19$). Note that this was in case when graphs on the questionnaires were most similar (difference in H was only 0.02).

Surprisingly, we found that the group with pairs at 0.02 difference were judged better (in terms of proportion of incorrect judgments) than the the group at where difference between the Hurst exponent were higher (0.10 and 0.20).

We have also found that the Thurstone-Mosteller model overestimates smaller values of Hurst exponent and underestimated higher values of H. It may be

Table 5: Results of Thurstone-Mosteller model at 2% H difference: $0.4 \leq H \leq 0.6$.

True H value	Weight	Estimated H value
H = 0.65	16.2645	0.65
H = 0.67	13.5782	0.683032
H = 0.69	12.1432	0.700678
H = 0.71	10.7919	0.717295
H = 0.73	9.53128	0.732796
H = 0.75	8.12416	0.750099
H = 0.77	6.70088	0.767601
H = 0.79	5.48683	0.78253
H = 0.81	4.09088	0.799695
H = 0.83	2.65394	0.817365
H = 0.85	0.	0.85

recommended to use other models in the future (e.g. Bradley-Terry model).

REFERENCES

Beer J.A.; Stead D.; and Coggan S.J., 2002. *Technical Note Estimation of the Joint Roughness Coefficient (JRC) by Visual Comparison. Rock Mechanics and Rock Engineering*, 35, no. 1.

Beran J., 1994. *Statistics for long-memory processes*, vol. 61. CRC press.

Bradley R.A., 1976. *A biometrics invited paper. science, statistics, and paired comparisons. Biometrics*, 213–239.

Cheung Y.W., 1993. *Long memory in foreign exchange rates. Journal of Business and Econimic Statistic*, 11.

Cornelis A.L. and Yu B., 2008. *Persistence characteristics of the Chinese stock markets. International Review of Financial Analysis*, 17, no. 1, 64–82.

David H., 1988. *The method of Paired Comparisons*. Oxford University Press.

Davidson R.R. and Farquhar P.H., 1976. *A bibliography on the method of paired comparisons. Biometrics*, 241–252.

Dongwei S. and Belton M.F., 1998. *Risk, Return and Regulation in Chinese Stock Markets. Journal of Economics and Business*, 50, no. 3.

Feder J., 1988. *Fractals New York*.

Geweke J. and Porter-Hudak S., 1983. *The Estimation and Application of Long Memory Time Series Models. Journal of Time Series Analysis*, 3.

Gneiting T. and Schlather M., 2004. *Stochastic models that separate fractal dimension and the Hurst effect. SIAM review*, 46, no. 2, 269–282.

Gorshkov O., 2012. *Stabilogram diffusion analysis algorithm to estimate the Hurst exponent of high-dimensional fractals. Journal of Statistical Mechanics: Theory and Experiment*, 2012, no. 04, P04014.

Greene M. and Fielitz B., 1997. *Long-Term Dependence in Common Stock Returns. Journal of Financial Economics*, 4, 339–349.

Hurst H., 1951. *Long-term Storage Capacity of Reservoirs. Trans Amer Soc Civil Eng*, 116.

Litvine I.N., 1999. *Paired Comparisons in Science, Social Science, Economics and Healthebinaugural Address Delivered Before the University of Port Elizabeth on 24 March 1999*. University of Port Elizabeth.

Litvine I.N., 2004. *Models and methods of paired comparisons*. Lviv, Ukrain: Naulilus.

Litvine I.N., 2014. *Hurst Exponent for Linear Regression Process*. In *Proceedings of the 56th Annual Conference of SASA*. 49–56.

Litvine I.N. and Gorshkov O., 2016. *Persistence Analysis of RR-interval Series*. In *European Simulation and Modelling 2016*. 1–5.

Lo A.W., 2007. *Efficient Market Hypothesis. A Dictionary of Economics*, 17.

Malamud B.D. and Turcotte D.L., 1999. *Self-affine time series: I. Generation and analyses. Advances in Geophysics*, 40, 1–90.

Mandelbrot B.B., 1969. *Long-Run Linearity, Locally Gaussian Process, H-Spectra and Infinite Variances. International Economic Review*, 10, 82–111.

Mandelbrot B.B., 1972. *Statistical Methodology for Nonperiodic Cycles: From the Covariance to R/S Analysis. Annals of Economic and Social Measurement*, 1, no. 3, 259–290.

Mandelbrot B.B., 1982. *The fractal geometry of nature. San Francisco, CA*.

Mandelbrot B.B., 1985. *Self-Affine Fractals and Fractal Dimension. Physica Scripta*, 32, no. 4, 257. URL `http://stacks.iop.org/1402-4896/32/i=4/a=001`.

Mandelbrot B.B. and Van Ness J.W., 1968. *Fractional Brownian motions, fractional noises and applications. SIAM review*, 10, no. 4, 422–437.

Mandelbrot B.B. and Wallis J.R., 1969a. *Reply [to Comments on Noah, Joseph, and Operational Hydrologyby Benoit B. Mandelbrot and James R. Wallis]. Water resources research*, 5, no. 4, 917–920.

Mandelbrot B.B. and Wallis J.R., 1969b. *Robustness of the rescaled range R/S in the measurement of noncyclic long run statistical dependence. Water Resources Research*, 5, no. 5, 967–988.

Matos J.A.; Gama S.M.; Ruskin H.J.; Sharkasi A.A.; and Crane M., 2008. *Time and scale Hurst exponent analysis for financial markets. Physica A: Statistical Mechanics and its Applications*, 387, no. 15, 3910 – 3915.

Mosteller F., 1951. *Remarks on the Method of Paired Comparisons. Psychometrika*, 16, 3–9.

NyoNyo A.K.; Cornelis A.L.; and Sijing Z., 2006. *Persistence characteristics of Latin American financial markets. Journal of Multinational Financial Management*, 16, no. 3, 269–290.

Peng C.K.; Buldyrev S.V.; Havlin S.; Simons M.; Stanley H.E.; and Goldberger A.L., 1994. *Mosaic organization of DNA nucleotides. Phys Rev E*, 49, 1685–1689. doi:10.1103/PhysRevE.49.1685. URL `http://link.aps.org/doi/10.1103/PhysRevE.49.1685`.

Peters E.E., 1996. *Chaos and Order in the Capital Markets: A New View of Cycles, Prices, and Market Volatility*. Wiley.

Rasheed K. and Qian B., 2004. *Hurst exponent and financial market predictability*. In *IASTED conference on Financial Engineering and Applications (FEA 2004)*. 203–209.

Robinson P.M., 1995. *Gaussian semiparametric estimation of long range dependence. The Annals of statistics*, 1630–1661.

Sadique S. and Silvapulle P., 2001. *Long-term memory in stock market returns: inter-national evidence. Internation Journal of Finance and Economics 6*.

Simonsen I. Hansen A. and Nes O.M., 1998. *Determination of the Hurst exponent by use of wavelet transforms. Phys Rev E*, 58, 2779–2787. doi:10.1103/PhysRevE.58.2779. URL `http://link.aps.org/doi/10.1103/PhysRevE.58.2779`.

Thurstone L.L., 1927a. *A law of comparative judgment. Psychological review*, 34, no. 4, 273–286.

Thurstone L.L., 1927b. *The method of paired comparisons for social values. The Journal of Abnormal and Social Psychology*, 21, no. 4, 384.

Thurstone L.L., 1927c. *Psychophysical analysis. The American journal of psychology*, 38, no. 3, 368–389.

Torsten K., 2002. *Testing Continuous Time Models in Financial Markets*. Doctoral thesis, Berlin.

Valdiviezo-N J.C.; Castro R.; Cristóbal G.; and Carbone A., 2014. *Hurst exponent for fractal characterization of LANDSAT images*. In *SPIE Optical Engineering+ Applications*. International Society for Optics and Photonics, 922103–922103.

Willinger W. Taqqu M. and Teverovsky V., 1999. *Stock market prices and long-range dependence. Finance of Stochastic*, 3.

LOOKING FOR AND FINDING LOST DATA

William Conley
Austin E. Cofrin School of Business
480B Wood Hall
University of Wisconsin-Green Bay
Green Bay, Wisconsin 54311-7001 U.S.A.
Conleyw@uwgb.edu

KEYWORDS
Lost data, error or cyberattack, statistical optimization, Monte Carlo search

ABSTRACT

Important data can be lost or misplaced in many different ways and circumstances. Sometimes it is never found again. Well run organizations take steps (including spending money) to secure data and have extra copies of it stored in safe places in case the original data is lost. Another problem is sometimes data doesn't appear to be important and later it becomes important and must be found again. Also, so called cyberattacks (criminal activity) can steal and/or destroy or compromise data. Presented here is an example where important data is lost and then recovered using good management practice and statistical optimization.

INTRODUCTION

A new technical organization (Company A) has just been hired by an important client to do work for it. Winning the bid for this project was a great achievement for Company A. Therefore, at 11 am, local time, on a Monday, the first installment of data was sent to Company A by computer email. It was then forwarded to the five Company A employees. All five of the employees went to lunch a little later and when they returned at about 1 pm, all their computers failed to work and the new data was missing.

CYBER ATTACK

The president (CEO) of Company A thinks that one of its competitors, who lost out to them in the bidding for the new client, launched some kind of cyberattack to the company's computers. A few hours later Company A's computer experts have the computers up and working again, but the important data sent from their new client is gone.

This appears to be industrial sabotage. But regardless, the company CEO does not want to call its new client and say they lost the data and need more copies of it, for fear of losing this potentially long term important client's business.

Therefore, the five employees are all questioned. Did anyone write down the data? No. Does anyone remember anything about the data? Yes. Did anyone do any calculations with or about the data? Yes.

The company's statistician remembered that there were seven numbers and the median was 262 and all numbers were whole numbers of 1 to 500. She then used her calculator to calculate the sample mean, mean absolute deviation, sum of squares of their numbers, harmonic mean and the standard deviations s and s□ (called s prime). She then was going to calculate the skewness and kurtosis coefficients. However, the phone rang and she talked to a colleague for a while and then they all went to lunch. She also noticed as she was leaving her computer that the three numbers less than the median averaged 149.

Did she write the data down? No. Did she write down or remember the eight statistics she had just calculated? Yes. She wrote them on a little 3 inch by 5 inch note card.

HERE IT IS

The company still does not know the seven numbers, except for the median of 262. However, they know (according to the statistician's note card) that the:

1. Median = 262 (the median is the middle number of an odd number of numbers once they are placed in ascending order)
2. Letting x_1, x_2, x_3 be the three numbers less than the median then $x_1+x_2+x_3 = 447$ because she said their average was 149
3. Her sample mean calculation revealed $\overline{x} = \frac{\sum x_i}{n} = 255.571$
4. The harmonic mean = 190.5664
5. The mean absolute deviation M.A.D. = 91.347

6. The sum of the squares of the seven numbers = 547889.
7. Standard deviation s prime s□ = 113.811
8. Standard deviation s = 122.9310

Therefore, a nonlinear system of seven equations and six variables (remembering 262 is known to be the seventh number) is set up and solved using multi stage Monte Carlo optimization (MSMCO or statistical optimization) revealing the seven data points to be 245, 117, 85, 262, 312, 439 and 329.

DISCUSSION

The next section will do the formal solution of the seven equations and six variables (remembering the median = 262 is known). However, a few comments precede the formal calculation. First of all, not every computer security or lost data recovery problem can be dealt with effectively by solving a system of nonlinear equations. However, a few of them can be. Also, there are problems in chemistry (doing chemistry experiments to solve for an "unknown") and physics problems (many body problems), electrical circuits, business optimization problems, and economics equilibrium equations, etc. where solving for unknowns (looking for and finding the answers) is possible by solving a system of equations. However, frequently those are nonlinear and multivariate and difficult to solve with classical methods. Therefore, the statistical optimization approaches rooted in simulation are the way to deal with these nonlinear systems.

THE FORMAL PROBLEM SOLUTION

Keeping in mind that the median = 262 attempt to solve the six variable systems of equations using the statistical simulation optimization (Multi Stage Mont Carlo) technique:

$$(262. + x_1+x_2+x_3+x_4+x_5+x_6)/7. = 255.571 \quad (1)$$

$$\frac{n}{1/262.+\sum_{i=1}^{6} 1./x_i} = 190.5664 \quad (2)$$

$$(\,|262.-255.571|+\sum_{i=1}^{6}|x_i-255.571|\,)/7.=91.347 \quad (3)$$

$$262^2+\sum_{i=1}^{6} x_i^{\,2} =547{,}889 \quad (4)$$

$$(((262.-255.571)^2+\sum_{i=1}^{6}(x_i-255.571)^2)/7.)^{.5}= 113.811 \quad (5)$$

$$(((262.-255.571)^2+\sum_{i=1}^{6}(x_i-255.571)^2)/6.)^{.5}= 122.9310 \quad (6)$$

$$x_1+x_2+x_3 = 447 \quad (7)$$

subject to $1 \le x_i \le 500$ for i=1, 2, 3, 4, 5 and 6 and all x_is are whole numbers. Let the left and right hand side of equation j be L_j and R_j respectively for j = 1, 2, 3, 4, 5, 6 and 7. Then try to minimize $f(x_1, x_2, x_3, x_4, x_5, x_6) = \sum_{j=1}^{7} |L_j-R_j|$ with a thirty stage Monte Carlo solution attempt drawing 200,000 sample solutions at each of the thirty stages. (Please note that two slight alterations were made to the $f(x_1, x_2, x_3, x_4, x_5, x_6)$ function and sampling scheme to have a smoother simulation:

a. When j=4 $|L_4-R_4|$ is replaced with .001 $|L_4-R_4|$ because the right hand side value 547,889 is on the order of 1000 times larger than the other equations right hand side values.
b. Random whole numbers are read in for x_1 and x_2. Then x_3 is solved for in equation 7 as $x_3 = 447-x_1-x_2$. Then, if x_3 is greater than or equal to one, the simulation proceeds with the current random sample solution. If not, then that set of six numbers is judged to be not feasible and is discarded. Then the simulation proceeds with a new feasible solution.

An outer loop did 20 such solution attempts and the clear answer was:

$x_1 = 245$	$e_1 = .00043$
$x_2 = 117$	$e_2 = .00044$
$x_3 = 85$	$e_3 = 00006$
$x_4 = 312$	$e_4 = 00000$
$x_5 = 439$	$e_5 = 00069$
$x_6 = 329$	$e_6 = 00032$

Please see Figure 1 for a partial geometric and statistical representation of the last three stages of the thirty stage Monte Carlo simulation.

CONCLUSION

A hypothetical problem was presented where through industrial sabotage by a competitor (a cyber attack) or possibly an internal organization error important data was lost.

However, enough statistical calculations had been done by a statistician (answers recorded on paper but not the data itself) that it was possible to set up a system of equations and solve it with statistical optimization simulation (or multi stage Monte Carlo optimization). A

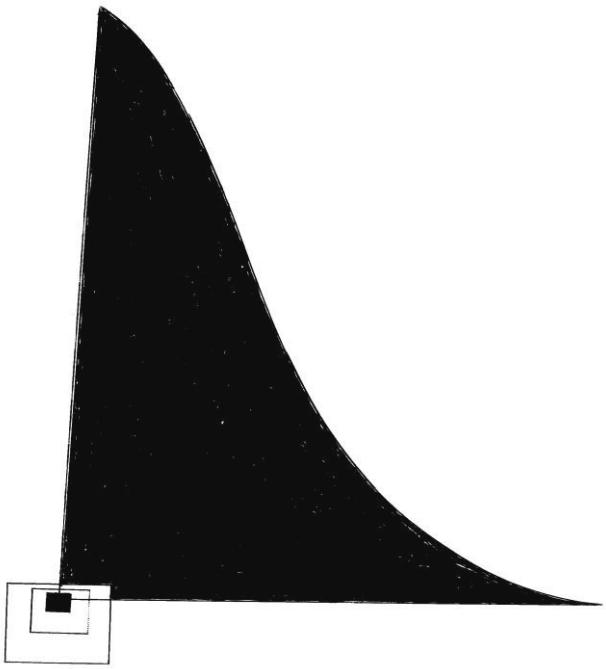


Figure 1: Multi Stage Monte Carlo at Work

real case like this would probably have the company's legal department investigating and taking appropriate legal action as warranted.

Many companies and organizations spend a lot of time, effort and money to protect and safeguard their data and industrial secrets and patents. However, given that human nature perhaps has not changed in thousands of years, industrial sabotage certainly predates the computer age.

On New Year's Eve in 1879 the inventor, Thomas Edison, invited the New York press and powerful financiers to a party at his Menlo Park laboratory to show off his improved electric light (longer lasting). The guests were astonished that the lab was lit up by dozens of electrical lights. The coverage and publicity represented serious competition to the natural gas industry (and their market for gas lights in homes).

Therefore, to protect his intellectual property from sabotage Edison had some of his security "boys" on hand. When some "hired" electricians showed up at the party to destroy the Edison light show they were forcibly escorted off the premises by his security people (Stross, 2007) and (Adair, 1996).

It is, of course, true that data and property can be lost, misplaced or damaged without any crime having been committed. However, protecting corporate, organizational or personal property in our internet connected age is an ongoing battle. Presented here is an essentially statistical optimization approach to this problem.

Reviews of statistics for management and engineers can be found in (Hayter 2002) and (Keller and Warrack 2003). Also, additional applications of statistical optimization simulation can be found in (Conley 2008), (Conley 2013a) and (Conley 2013b).

REFERENCES

Adair, G. 1996. *Thomas Alva Edison Inventing the Electric Age*. Oxford University Press. New York and Oxford, England.

Conley, W. C. 2008. "Ecological Optimization of Pollution Control Equipment from a Planning and a Simulation Perspective." *International Journal of Systems Science*, Vol., No. 1-7.

Conley, W. C. 2013a. "Three Eight Dimensional Correlation Studies." *Proceedings of the 2013 Industrial Simulation Conference*, ISC, 2013 EUROSIS, Ostend, Belgium, pp. 20-27.

Conley, W. C. 2013b. "Transportation and Minimum Content Problems." *Proceedings of European Simulation and Modeling Conference*, ESM 2013, October 23-25, Lancaster, England. EUROSIS, Ostend, Belgium, pp. 325-330.

Hayter A. J. 2002. *Probability and Statistics for Engineers and Scientists*, 2nd Edition. Duxbury Press, Pacific Grove, California.

Keller, G. and Warrack, B. 2003. *Statistics for Management and Economics*, 6th Edition. Thompson Brooks/Cole, Pacific Grove, CA.

Stross, R. 2007. *The Wizard of Menlo Park*. Crown Publishing Group, Random House, New York.

BIOGRAPHY

WILLIAM CONLEY received a B.A. in mathematics (with honors) from Albion College in 1970, an M.A. in mathematics from Western Michigan University in 1971, 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 200 publications world- wide. He is a member of the American Chemical Society, a fellow in the Institution of Electronic and Telecommunication Engineers, a senior member of the Society for Computer Simulation and named to (KME) Kappa Mu Epsilon, the national mathematics honorary.

EFFICIENT QUANTILE ESTIMATION VIA A COMBINATION OF IMPORTANCE SAMPLING AND LATIN HYPERCUBE SAMPLING

Marvin K. Nakayama
Department of Computer Science
New Jersey Institute of Technology
Newark, NJ, 07102, USA
email: marvin@njit.edu

KEYWORDS
Value-at-risk, variance reduction, risk analysis

ABSTRACT

Many application areas employ a quantile, also known as a percentile or value-at-risk, to measure risk of a stochastic system. We present efficient Monte Carlo methods to estimate a quantile through a combination of importance sampling and Latin hypercube sampling. We also give numerical results from a simple model showing that the combined methods can outperform each by itself.

INTRODUCTION

Consider a random variable Y output by a stochastic simulation. For instance, Y may be the future loss of a financial portfolio, or Y can be the *peak cladding temperature* (PCT) in a hypothesized accident at a *nuclear power plant* (NPP). For a fixed constant $0 < p < 1$, we define the p-quantile of Y, also know as its $100p$-th percentile, to be the constant ξ such that $P(Y \leq \xi) = p$. For example, the median is the 0.5-quantile.

Extreme quantiles (with $p \approx 0$ or $p \approx 1$) are often employed to assess risk. As an example, financial analysts frequently utilize a quantile, which is also called a *value-at-risk*, to measure portfolio risk. Indeed, the Basel II Accord (Basel Committee on Banking Supervision 2004) specifies capital requirements in terms of 0.99-quantiles. The U.S. Nuclear Regulatory Commission (NRC) permits NPP licensees to demonstrate compliance with federal regulations through a 0.95-quantile (U.S. Nuclear Regulatory Commission 2010).

Quantile estimation via *simple random sampling* (SRS) has been well studied; e.g., see Chapter 2 of (Serfling 1980) and (Hong et al. 2014). But SRS can produce a p-quantile estimator with large asymptotic variance, which motivates applying *variance-reduction techniques* (VRTs) to lessen the sampling error. Different VRTs have been proposed for quantile estimation, including *importance sampling* (IS) (Glynn 1996, Glasserman et al. 2000, Chu and Nakayama 2012) and *Latin hypercube sampling* (LHS) (Avramidis and Wilson 1998, Dong and Nakayama 2017). (Olsson et al. 2003) combine IS and LHS to estimate a failure probability, and they find that the combination can outperform each method by itself in estimating the probability. Our current paper combines IS+LHS for quantile estimation. We give numerical results for a simple example showing that IS+LHS can produce quantile estimators with smaller error than either technique on its own.

The rest of the paper unfolds as follows. After giving the mathematical setting, we review quantile estimation via SRS, IS, and LHS. We then develop IS+LHS quantile estimation. We follow this with results from simple numerical experiments, and end with concluding remarks.

MATHEMATICAL SETTING

For a random vector $\mathbf{X} = (X_1, X_2, \ldots, X_d) \in \Re^d$, let

$$Y = w(\mathbf{X}), \tag{1}$$

where $w : \Re^d \to \Re$ is a given computable (deterministic) function. Let G be the cumulative distribution function (CDF) of $\mathbf{X}$, which we denote by $\mathbf{X} \sim G$, so $G(\mathbf{x}) = P(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_d \leq x_d)$ for $\mathbf{x} = (x_1, x_2, \ldots, x_d) \in \Re^d$. Assume that

$$G(\mathbf{x}) = \prod_{j=1}^{d} G_j(x_j), \tag{2}$$

where each $X_j \sim G_j$, which implies the components of $\mathbf{X}$ are *independent*. We allow $G_1, G_2, \ldots, G_d$ to differ. Let F be the CDF of the output Y in (1).

We can think of the function w in (1) as a computer code that transforms a random input $\mathbf{X} \sim G$ into a response $Y \sim F$. For example, in a probabilistic safety assessment of an NPP (U.S. Nuclear Regulatory Commission 2010), a detailed computer code models the evolution of a hypothesized event, such as a loss-of-coolant accident. The random vector $\mathbf{X}$ represents uncertainties, such as the timing and size of the initiating event (e.g., a pipe break) and material properties of the core cladding. The computer code w numerically solves systems of differential equations, with coefficients determined by $\mathbf{X}$, to output a figure of merit Y, such as the PCT.

For a fixed $0 < p < 1$, we define the p-quantile of F (or equivalently, of Y) as the constant $\xi = F^{-1}(p) \equiv \inf\{x :$

$F(x) \geq p\}$. Let f be the derivative (when it exists) of F, and assume that $f(\xi) > 0$, which ensures that the p-quantile is unique. We also assume that the CDF F cannot be computed because of the complexity of the function w in (1), but we can generate an observation of $Y \sim F$ by sampling $\mathbf{X} \sim G$ and applying (1). The goal is to estimate ξ via Monte Carlo simulation. The typical approach first estimates the CDF F via multiple simulation runs, and then inverts the CDF estimator to obtain an estimator of the p-quantile $\xi = F^{-1}(p)$.

Assume that G_j^{-1} can be computed for each coordinate $1 \leq j \leq d$, and let $U_1, U_2, \ldots, U_d$ be independent and identically distributed (i.i.d.) unif$(0,1)$ random variables. We then have that $G_j^{-1}(U_j) \sim G_j$, so

$$(G_1^{-1}(U_1), G_2^{-1}(U_2), \ldots G_d^{-1}(U_d)) \sim G \tag{3}$$

by (2), from which (1) implies

$$w(G_1^{-1}(U_1), G_2^{-1}(U_2), \ldots G_d^{-1}(U_d)) \sim F. \tag{4}$$

SIMPLE RANDOM SAMPLING

We now review SRS estimation of ξ. Let n be the desired sample size, and generate $n \times d$ i.i.d. unif$(0,1)$ random numbers, which we arrange in an $n \times d$ grid

$$\begin{array}{cccc} U_{1,1} & U_{1,2} & \cdots & U_{1,d} \\ U_{2,1} & U_{2,2} & \cdots & U_{2,d} \\ \vdots & \vdots & \ddots & \vdots \\ U_{n,1} & U_{n,2} & \cdots & U_{n,d} \end{array} . \tag{5}$$

For each column $1 \leq j \leq d$ in (5), we then apply the function G_j^{-1} to each entry in the column to obtain

$$\begin{array}{lllll} \mathbf{X}_1 = (G_1^{-1}(U_{1,1}), & G_2^{-1}(U_{1,2}), & \ldots, & G_d^{-1}(U_{1,d})), \\ \mathbf{X}_2 = (G_1^{-1}(U_{2,1}), & G_2^{-1}(U_{2,2}), & \ldots, & G_d^{-1}(U_{2,d})), \\ \vdots \quad \vdots & \vdots & \vdots & \ddots \quad \vdots \\ \mathbf{X}_n = (G_1^{-1}(U_{n,1}), & G_2^{-1}(U_{n,2}), & \ldots, & G_d^{-1}(U_{n,d})). \end{array} \tag{6}$$

Because each row i in (5) has d i.i.d. unif$(0,1)$ random variables, (3) implies that each $\mathbf{X}_i \sim G$. As the n rows in (5) are independent, we also have that $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ are mutually independent.

Now let $Y_i = w(\mathbf{X}_i)$, $1 \leq i \leq n$, and each $Y_i \sim F$ by (1) and (4), with $Y_1, Y_2, \ldots, Y_n$ mutually independent. Note that $F(y) = P(Y \leq y) = E[I(Y \leq y)] = 1 - E[I(Y > y)] = 1 - P(Y > y)$, where $I(\cdot)$ denotes the indicator function, which equals 1 (resp., 0) when its argument is true (resp., false). The SRS estimator of the CDF F of Y is then defined by

$$\hat{F}_{\mathrm{SRS},n}(y) = \frac{1}{n}\sum_{i=1}^{n} I(Y_i \leq y) = 1 - \frac{1}{n}\sum_{i=1}^{n} I(Y_i > y),$$

which is *unbiased* for each y; i.e., $E[\hat{F}_{\mathrm{SRS},n}(y)] = F(y)$. The SRS estimator of the p-quantile $\xi = F^{-1}(p)$ is then

$$\hat{\xi}_{\mathrm{SRS},n} = \hat{F}^{-1}_{\mathrm{SRS},n}(p),$$

which we can compute through order statistics. Let $Y_{1:n} \leq Y_{2:n} \leq \cdots \leq Y_{n:n}$ be the ordered values of the Y_i. Then $\hat{\xi}_{\mathrm{SRS},n} = Y_{\lceil np \rceil:n}$, with $\lceil \cdot \rceil$ the ceiling function. Although *not* a sample average, $\hat{\xi}_{\mathrm{SRS},n}$ still obeys a central limit theorem (CLT) $\sqrt{n}[\hat{\xi}_{\mathrm{SRS},n} - \xi] \Rightarrow N(0, \tau^2_{\mathrm{SRS}})$ as $n \to \infty$ (e.g., p. 77 of (Serfling 1980)), where $\Rightarrow$ denotes convergence in distribution (Section 1.2.4 of (Serfling 1980)), $N(a, b^2)$ is a normal random variable with mean a and variance b^2, and the CLT's *asymptotic variance* is

$$\tau^2_{\mathrm{SRS}} = \frac{\psi^2_{\mathrm{SRS}}}{f^2(\xi)} = \frac{p(1-p)}{f^2(\xi)}. \tag{7}$$

Section 2.6 of (Serfling 1980) discusses some methods for constructing confidence interval for ξ when using SRS.

IMPORTANCE SAMPLING

Importance sampling is a VRT that is well suited to analyze rare events; e.g., see Chapter VI of (Asmussen and Glynn 2007). The basic idea is to alter the distributions driving the original system to make the rare event of interest (e.g., a large financial loss of a portfolio) occur more frequently, and then apply a correction factor to account for the change. (Glynn 1996) devises importance-sampling estimators for a quantile by first applying IS to estimate the CDF F, and then inverting the resulting CDF estimator to obtain the IS quantile estimator. We now review the details.

Let P_G and E_G be the probability measure and expectation operator, respectively, when $\mathbf{X} \sim G$, so $F(y) = P_G(Y \leq y) = E_G[I(w(\mathbf{X}) \leq y)]$. Let H be another CDF on $\Re^d$, and let P_H (resp., E_H) be the probability measure (resp., expectation) when $\mathbf{X} \sim H$. Assume that G is *absolutely continuous* with respect to H (see p. 422 of (Billingsley 1995)); i.e., if $P_H(\mathbf{X} \in A) = 0$ for some (measurable) set $A \subseteq \Re^d$, then $P_G(\mathbf{X} \in A) = 0$. The idea underlying IS comes from expressing the CDF F as

$$\begin{aligned} F(y) &= 1 - P_G(Y > y) = 1 - E_G[I(w(\mathbf{X}) > y)] \\ &= 1 - \int_{\Re^d} I(w(\mathbf{x}) > y)\, G(d\mathbf{x}) \\ &= 1 - \int_{\Re^d} I(w(\mathbf{x}) > y)\, \frac{G(d\mathbf{x})}{H(d\mathbf{x})}\, H(d\mathbf{x}) \\ &= 1 - E_H[I(w(\mathbf{X}) > y) L(\mathbf{X})], \end{aligned} \tag{8}$$

which applies a so-called *change of measure* as (8) has an expectation computed with $\mathbf{X} \sim H$ instead of G, and

$$L(\mathbf{x}) = \frac{G(d\mathbf{x})}{H(d\mathbf{x})} \tag{9}$$

is known as the *likelihood ratio* or Radon-Nykodim derivative; see p. 423 of (Billingsley 1995). Hence, (8)

implies that we can obtain an unbiased estimator of $1-F(y)$ by averaging i.i.d. copies of $I(w(\mathbf{X}) > y)L(\mathbf{X})$, where $\mathbf{X}$ is generated from H rather than G.
To do this, we assume that the joint CDF H satisfies

$$H(\mathbf{x}) = \prod_{j=1}^{d} H_j(x_j) \tag{10}$$

for $\mathbf{x} = (x_1, x_2, \ldots, x_d) \in \Re^d$, where each H_j is a CDF on $\Re$, and H_j, $1 \le j \le d$, may differ. Thus, a random vector with the joint CDF H has independent components. By (2) and (10), we see that (9) becomes

$$L(\mathbf{x}) = \prod_{j=1}^{d} \frac{G_j(dx_j)}{H_j(dx_j)} \tag{11}$$

for $\mathbf{x} = (x_1, x_2, \ldots, x_d)$. If each G_j (resp., H_j) further has a density g_j (resp., h_j), then the likelihood ratio satisfies $L(\mathbf{x}) = \prod_{j=1}^{d} [g_j(x_j)/h_j(x_j)]$. In this case, G in (2) is absolutely continuous with respect to H if $h_j(x) = 0$ implies $g_j(x) = 0$ for each coordinate $1 \le j \le d$.
We next construct an estimator of F based on (8). Assuming that each H_j^{-1} is computable, we then can sample n i.i.d. copies of $\mathbf{X} \sim H$ by applying H_j^{-1} to each entry in each column j of (5) to obtain

$$\begin{aligned}
\mathbf{X}_1 &= (H_1^{-1}(U_{1,1}),\ H_2^{-1}(U_{1,2}),\ \ldots,\ H_d^{-1}(U_{1,d})),\\
\mathbf{X}_2 &= (H_1^{-1}(U_{2,1}),\ H_2^{-1}(U_{2,2}),\ \ldots,\ H_d^{-1}(U_{2,d})),\\
&\ \ \vdots \quad\ \vdots \qquad\ \vdots \qquad\ \ddots \qquad\ \vdots\\
\mathbf{X}_n &= (H_1^{-1}(U_{n,1}),\ H_2^{-1}(U_{n,2}),\ \ldots,\ H_d^{-1}(U_{n,d})).
\end{aligned} \tag{12}$$

Because each row i in (5) has d i.i.d. unif$(0,1)$ random variables, we have that $\mathbf{X}_i$ in (12) has joint CDF H by (10). Moreover, $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ in (12) are independent because the rows of (5) are independent. Hence, an IS estimator of the CDF F of Y motivated by (8) is

$$\hat{F}_{\mathrm{IS},n}(y) = 1 - \frac{1}{n}\sum_{i=1}^{n} I(w(\mathbf{X}_i) > y)L(\mathbf{X}_i). \tag{13}$$

The IS estimator of the p-quantile $\xi = F^{-1}(p)$ is then

$$\hat{\xi}_{\mathrm{IS},n} = \hat{F}_{\mathrm{IS},n}^{-1}(p). \tag{14}$$

To compute (14), sort the $w(\mathbf{X}_i)$, $1 \le i \le n$, in ascending order, and let $i_1, i_2, \ldots, i_n$ be the indices of the sorted values; i.e., $w(\mathbf{X}_{i_1}) \le w(\mathbf{X}_{i_2}) \le \cdots \le w(\mathbf{X}_{i_n})$. Then $\hat{\xi}_{\mathrm{IS},n} = w(\mathbf{X}_{i_r})$ for the largest integer r satisfying $\sum_{l=r}^{n} L(\mathbf{X}_{i_l}) \ge (1-p)n$. ((Glynn 1996) also considers other IS estimators of F and ξ; we omit the details.)
Assuming that $f(\xi) > 0$ and $E_H[L(\mathbf{X})^3] < \infty$, (Glynn 1996) proves that $\hat{\xi}_{\mathrm{IS},n}$ obeys a CLT $\sqrt{n}[\hat{\xi}_{\mathrm{IS},n} - \xi] \Rightarrow N(0, \tau_{\mathrm{IS}}^2)$ as $n \to \infty$, where $\tau_{\mathrm{IS}}^2 = \psi_{\mathrm{IS}}^2/f^2(\xi)$ and $\psi_{\mathrm{IS}}^2 = E_H[L(\mathbf{X})^2 I(w(\mathbf{X}) > \xi)] - (1-p)^2$. (Chu and Nakayama 2012, Nakayama 2014) develop asymptotically valid (as $n \to \infty$) confidence intervals for ξ when applying IS.

LATIN HYPERCUBE SAMPLING

LHS can be thought of as an efficient way of extending stratified sampling to high dimensions, and it reduces variance by producing negatively correlated outputs. (McKay et al. 1979) originally devised LHS to estimate a mean, and (Stein 1987) further analyzes the approach. (Avramidis and Wilson 1998) use LHS to estimate a quantile, which we now review.
For each input coordinate $1 \le j \le d$ of (2) and (4), let $\pi_j = (\pi_j(1), \pi_j(2), \ldots, \pi_j(n))$ be a random permutation of $(1, 2, \ldots, n)$. Specifically, π_j is equally likely to be any one of the $n!$ permutations, and i maps to $\pi_j(i)$ in permutation π_j. Assume that $\pi_1, \pi_2, \ldots, \pi_d$ are independent, and for each $1 \le i \le n$ and $1 \le j \le d$, define $V_{i,j} = [\pi_j(i) - 1 + U_{i,j}]/n$ for $U_{i,j}$ in (5). Then arrange the $V_{i,j}$ into an $n \times d$ grid

$$\begin{matrix}
V_{1,1} & V_{1,2} & \cdots & V_{1,d}\\
V_{2,1} & V_{2,2} & \cdots & V_{2,d}\\
\vdots & \vdots & \ddots & \vdots\\
V_{n,1} & V_{n,2} & \cdots & V_{n,d}
\end{matrix}. \tag{15}$$

For each column $1 \le j \le d$ of (15) and each $1 \le k \le n$, exactly one $1 \le i \le n$ has $V_{i,j} \in ((k-1)/n, k/n]$ by construction, so each column j of (15) forms a stratified sample of the unit interval. Next apply G_j^{-1} to each entry in each column j of (15) to get

$$\begin{aligned}
\mathbf{X}_1' &= (G_1^{-1}(V_{1,1}),\ G_2^{-1}(V_{1,2}),\ \ldots,\ G_d^{-1}(V_{1,d})),\\
\mathbf{X}_2' &= (G_1^{-1}(V_{2,1}),\ G_2^{-1}(V_{2,2}),\ \ldots,\ G_d^{-1}(V_{2,d})),\\
&\ \ \vdots \quad\ \vdots \qquad\ \vdots \qquad\ \ddots \qquad\ \vdots\\
\mathbf{X}_n' &= (G_1^{-1}(V_{n,1}),\ G_2^{-1}(V_{n,2}),\ \ldots,\ G_d^{-1}(V_{n,d})),
\end{aligned}$$

the LHS analogue of (6). It is easily shown that the d entries in each row i of (15) are i.i.d. unif$(0,1)$, so $\mathbf{X}_i' \sim G$ by (3). But all of the entries in each column j of (15) share the same permutation π_j, making the rows of (15) dependent, so $\mathbf{X}_1', \mathbf{X}_2', \ldots, \mathbf{X}_n'$ are also dependent. Now define $Y_i' = w(\mathbf{X}_i')$, $1 \le i \le n$, for the function w in (1). The fact that each $\mathbf{X}_i' \sim G$ ensures that each $Y_i' \sim F$ by (4). But $Y_1', Y_2', \ldots, Y_n'$ are dependent because $\mathbf{X}_1', \mathbf{X}_2', \ldots, \mathbf{X}_n'$ are.
The LHS estimator of the CDF F of Y in (1) is then

$$\hat{F}_{\mathrm{LHS},n}(y) = \frac{1}{n}\sum_{i=1}^{n} I(Y_i' \le y) = 1 - \frac{1}{n}\sum_{i=1}^{n} I(Y_i' > y).$$

Inverting this leads to the LHS p-quantile estimator

$$\hat{\xi}_{\mathrm{LHS},n} = \hat{F}^{-1}_{\mathrm{LHS},n}(p).$$

We can compute $\hat{\xi}_{\mathrm{LHS},n}$ as $\hat{\xi}_{\mathrm{LHS},n} = Y'_{\lceil np \rceil:n}$, where $Y'_{1:n} \le Y'_{2:n} \le \cdots \le Y'_{n:n}$ are the sorted Y'_i values.
Under regularity conditions, (Avramidis and Wilson 1998) prove the LHS p-quantile estimator obeys a CLT $\sqrt{n}[\hat{\xi}_{\mathrm{LHS},n} - \xi] \Rightarrow N(0, \tau^2_{\mathrm{LHS}})$ as $n \to \infty$, where

$$\tau^2_{\mathrm{LHS}} = \frac{\psi^2_{\mathrm{LHS}}}{f^2(\xi)}, \tag{16}$$

and we define ψ^2_{LHS} below. (Dong and Nakayama 2017) devise asymptotically valid confidence intervals for ξ when applying LHS.
To give an expression for ψ^2_{LHS}, let $\chi(U_1, U_2, \ldots, U_d) = I(w(G_1^{-1}(U_1), G_2^{-1}(U_2), \ldots, G_d^{-1}(U_d)) \le \xi)$ for $U_1, U_2, \ldots, U_d$ i.i.d. unif$(0,1)$, so $E[\chi(U_1, U_2, \ldots, U_d)] = F(\xi)$ by (4). Also, for each input coordinate $1 \le j \le d$, let $\chi_j(u) = E[\chi(U_1, U_2, \ldots, U_d) | U_j = u]$. We then have $\psi^2_{\mathrm{LHS}} = \psi^2_{\mathrm{SRS}} - \sum_{j=1}^d \mathrm{Var}[\chi_j(U_j)]$, where $\psi^2_{\mathrm{SRS}} = p(1-p)$ by (7). Therefore, the asymptotic variance τ^2_{LHS} in (16) of the LHS p-quantile estimator is no larger than the SRS asymptotic variance τ^2_{SRS} in (7).

COMBINED IS+LHS

We now combine importance sampling and LHS for quantile estimation. To do this, we first apply H_j^{-1} from (10) to each entry in each column j of (15) to get

$$\begin{aligned}
\mathbf{X}'_1 &= (H_1^{-1}(V_{1,1}),\ H_2^{-1}(V_{1,2}),\ \ldots,\ H_d^{-1}(V_{1,d})),\\
\mathbf{X}'_2 &= (H_1^{-1}(V_{2,1}),\ H_2^{-1}(V_{2,2}),\ \ldots,\ H_d^{-1}(V_{2,d})),\\
&\ \ \vdots \quad\ \vdots \qquad\ \vdots \qquad\ \ddots \qquad\ \vdots\\
\mathbf{X}'_n &= (H_1^{-1}(V_{n,1}),\ H_2^{-1}(V_{n,2}),\ \ldots,\ H_d^{-1}(V_{n,d})),
\end{aligned} \tag{17}$$

which is the IS+LHS analogue of the IS-only (12). Because each row i in (15) has d i.i.d. unif$(0,1)$ random variables, we have that $\mathbf{X}'_i$ in (17) has CDF H by (10). But $\mathbf{X}'_1, \mathbf{X}'_2, \ldots, \mathbf{X}'_n$ in (17) are dependent because the rows of (15) are dependent. Our IS+LHS estimator of F inspired by (8) is then

$$\hat{F}_{\mathrm{IS+LHS},n}(y) = 1 - \frac{1}{n}\sum_{i=1}^n I(w(\mathbf{X}'_i) > y) L(\mathbf{X}'_i) \tag{18}$$

for $\mathbf{X}'_1, \mathbf{X}'_2, \ldots, \mathbf{X}'_n$ in (17), w in (1), and L in (11). Thus, (18) is the IS+LHS analogue of the IS CDF estimator in (13). We then compute the IS+LHS estimator of the p-quantile $\xi = F^{-1}(p)$ as

$$\hat{\xi}_{\mathrm{IS+LHS},n} = \hat{F}^{-1}_{\mathrm{IS+LHS},n}(p).$$

We can compute $\hat{\xi}_{\mathrm{IS+LHS},n}$ as follows. Sort $w(\mathbf{X}'_i)$, $1 \le i \le n$, in ascending order, and let $m_1, m_2, \ldots, m_n$ be the indices of the sorted values; i.e., $w(\mathbf{X}'_{m_1}) \le w(\mathbf{X}'_{m_2}) \le \cdots \le w(\mathbf{X}'_{m_n})$. Then $\hat{\xi}_{\mathrm{IS+LHS},n} = w(\mathbf{X}'_{m_r})$ for the largest integer r satisfying $\sum_{l=r}^n L(\mathbf{X}'_{m_l}) \ge (1-p)n$.

NUMERICAL RESULTS

We ran numerical experiments to assess the quality of p-quantile estimators for the following simple simulation model. In (2), we let $d = 10$, and for each $1 \le j \le d$, we chose the marginal $G_j = \Phi$, the univariate $N(0,1)$ CDF, so $\mathbf{X} = (X_1, X_2, \ldots, X_d)$ is a vector of d i.i.d. standard normals. Hence, by (3), we can generate the vector as $\mathbf{X} = (\Phi^{-1}(U_1), \Phi^{-1}(U_2), \ldots, \Phi^{-1}(U_d)) \sim G$, where $U_1, U_2, \ldots, U_d$ are i.i.d. unif$(0,1)$. We define the function w in (1) as $w(\mathbf{X}) = \sum_{j=1}^d X_j$. Thus, the CDF F of $Y = w(\mathbf{X})$ is that of the $N(0,d)$ distribution. We can then compute the true p-quantile of F as $\xi = \sqrt{d}\Phi^{-1}(p)$, and the goal is to estimate ξ via simulation.
For sample size $n = 200$, we constructed p-quantile estimators using each of the methods SRS, LHS, IS, IS+LHS for $p = 0.8$, 0.9, 0.95, and 0.99. For IS and IS+LHS, the IS marginal CDF H_j in (10) is $N(\xi/d, 1)$ for each coordinate $1 \le j \le d$, so IS changes the mean of each X_j from 0 to ξ/d. Thus, their sum Y under H has mean ξ.
For each method and each value of p, we ran 10^4 independent experiments to estimate the *root-mean-square error* (RMSE) $(E[(\hat{\xi}_n - \xi)^2])^{1/2}$ of each p-quantile estimator, generically denoted as $\hat{\xi}_n$. (As quantile estimators are generally biased, we consider RMSE instead of standard deviation because the latter does not account for bias.) We finally estimated the *relative RMSE* (RRMSE) as the estimated RMSE divided by the true p-quantile. Table 1 contains the results.

Table 1: Relative Root-Mean-Square Error of the p-Quantile Estimator for Each of the Methods SRS, LHS, IS, and IS+LHS

Method	Relative RMSE			
	$p = 0.8$	$p = 0.9$	$p = 0.95$	$p = 0.99$
SRS	0.120	0.094	0.089	0.110
LHS	0.084	0.074	0.079	0.105
IS	0.069	0.041	0.030	0.019
IS+LHS	0.061	0.038	0.028	0.018

We now note some of the table's salient features. For each p, LHS has smaller RRMSE than SRS, so LHS always produces p-quantile estimators with less error than SRS. For each p, the RRMSE of IS is much smaller than that of LHS, especially for $p \approx 1$. Moreover, IS+LHS has smaller RRMSE than IS, so we can improve IS by further incorporating LHS.
For $p = 0.8$, the ratio of the RRMSEs of SRS and LHS is $0.120/0.84 \approx 1.43$. The quantile estimators' CLTs

suggest that their RMSEs decrease at rate $n^{-1/2}$, so SRS needs a sample size that is a factor of $(0.120/84)^2 \approx 2.04$ larger than that for LHS to achieve roughly the same RRMSE. But the ratio of RRMSEs for SRS and LHS decreases to $0.110/0.105 \approx 1.05$ for $p = 0.99$. Similarly, the ratio of the RRMSEs of IS and IS+LHS decreases from $0.069/0.061 \approx 1.13$ for $p = 0.8$ to $0.019/0.018 \approx 1.03$ for $p = 0.99$. Thus, the benefit of LHS lessens for more extreme quantiles.

CONCLUDING REMARKS

Quantiles are often used to assess risk, as in finance and nuclear safety. We combined two well-known variance-reduction techniques, importance sampling and Latin hypercube sampling, to estimate a p-quantile. The combination can outperform each method by itself in terms of relative RMSE, but the benefit of IS+LHS over IS may decrease as $p \to 1$.

ACKNOWLEDGMENTS

This work has been supported in part by the U.S. National Science Foundation under Grant No. CMMI-1537322. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

Asmussen S. and Glynn P., 2007. *Stochastic Simulation: Algorithms and Analysis*. Springer, New York.

Avramidis A.N. and Wilson J.R., 1998. *Correlation-Induction Techniques for Estimating Quantiles in Simulation. Operations Research*, 46, 574–591.

Basel Committee on Banking Supervision, 2004. *Basel II: International Convergence of Capital Measurement and Capital Standards: a Revised Framework*. Tech. rep., Bank for International Settlements, Basel, Switzerland. URL `http://www.bis.org/publ/bcbs107.htm`.

Billingsley P., 1995. *Probability and Measure*. John Wiley and Sons, New York, third ed.

Chu F. and Nakayama M.K., 2012. *Confidence Intervals for Quantiles When Applying Variance-Reduction Techniques. ACM Transactions On Modeling and Computer Simulation*, 22, no. 2, Article 10 (25 pages plus 12–page online–only appendix).

Dong H. and Nakayama M.K., 2017. *Quantile Estimation With Latin Hypercube Sampling. Operations Research*. To appear.

Glasserman P.; Heidelberger P.; and Shahabuddin P., 2000. *Variance Reduction Techniques for Estimating Value-at-Risk. Management Science*, 46, 1349–1364.

Glynn P.W., 1996. *Importance Sampling for Monte Carlo Estimation of Quantiles*. In *Mathematical Methods in Stochastic Simulation and Experimental Design: Proceedings of the 2nd St. Petersburg Workshop on Simulation*. Publishing House of St. Petersburg Univ., St. Petersburg, Russia, 180–185.

Hong L.J.; Hu Z.; and Liu G., 2014. *Monte Carlo Methods for Value-at-Risk and Conditional Value-at-Risk: A Review. ACM Transactions on Modeling and Computer Simulataion*, 24, Article 22 (37 pages).

McKay M.D.; Beckman R.J.; and Conover W.J., 1979. *A Comparison of Three Methods for Selecting Input Variables in the Analysis of Output from a Computer Code. Technometrics*, 21, 239–245.

Nakayama M.K., 2014. *Confidence Intervals Using Sectioning for Quantiles When Applying Variance-Reduction Techniques. ACM Transactions on Modeling and Computer Simulation*, 24, Article 19.

Olsson A.; Sandberg G.; and Dahlblom O., 2003. *On Latin Hypercube Sampling for Structural Reliability Analysis. Structural Safety*, 25, 47–68.

Serfling R.J., 1980. *Approximation Theorems of Mathematical Statistics*. John Wiley and Sons, New York.

Stein M., 1987. *Large Sample Properties of Simulations Using Latin Hypercube Sampling. Technometrics*, 29, 143–151. Correction 32:367.

U.S. Nuclear Regulatory Commission, 2010. *Acceptance criteria for emergency core cooling systems for light-water nuclear power reactors*. Title 10, Code of Federal Regulations Section 50.46 (10CFR50.46), U.S. Nuclear Regulatory Commission, Washington, DC.

AUTHOR BIOGRAPHY

MARVIN K. NAKAYAMA is a computer science professor at the New Jersey Institute of Technology. He received a B.A. in mathematics-computer science from U.C. San Diego, and an M.S. and Ph.D. in operations research from Stanford University. He is a recipient of a CAREER Award from the National Science Foundation, and has received several best paper awards, including the Best Theoretical Paper for the 2014 Winter Simulation Conference. He served as the simulation area editor for the *INFORMS Journal on Computing* from 2007–2016, and has been on the editorial board of *ACM Transactions on Modeling and Computer Simulation* since 1998. His research interests include simulation, modeling, statistics, risk analysis, and energy.

STATISTICAL OPTIMIZATION APPLIED TO A VARIETY OF NONLINEAR SYSTEMS OF EQUATIONS

William Conley
Austin E. Cofrin School of Business
University of Wisconsin-Green Bay
Green Bay, Wisconsin 54311-7001 U.S.A.
Conleyw@uwgb.edu

KEYWORDS
Minimizing errors, Monte Carlo Methods, exact solutions, nonlinear versus linear

ABSTRACT

There is a fairly well developed and widely known theory for solving linear systems of equations which when hooked up to a modern day computer can be a powerful facility. The real cost of this approach is the all too common temptation to linearize essentially nonlinear relationships. One way to overcome this problem will be discussed here. That is using computer statistical simulation techniques to solve nonlinear systems of equations in our computer age.

Four examples of nonlinear systems of equations will be presented and solved here. They will range in size from five to twenty-five variable systems..

INTRODUCTION

The fundamental theorem of linear programming can be expressed (in a slightly oversimplified manner) to be that the optimal solution is at a "corner point" of the feasible solution space. This, along with a course in linear algebra and linear systems, points the way toward solving systems that are linear.

However, with nonlinear systems of equations, or general nonlinear optimization, the optimal solution or solutions could be anywhere in the feasible solutions spaces.

Therefore, simulation techniques such as multi stage Monte Carlo optimization (MSMCO) seem useful in dealing with difficult nonlinear examples. Four such problems are presented here. Figure 1 gives a partial geometric and statistical representation of the MSMCO simulations at work.

A FIVE VARIABLE, FIVE EQUATION EXAMPLE

Try to solve the following system of equations:

$$x_1^3 + x_2x_4 + x_5 = 17{,}192{,}756 \quad (1)$$

$$x_2^3 + x_1x_3 + x_4 = 905{,}822 \quad (2)$$

$$x_3^3 + x_2x_3 + x_5 = 6{,}557{,}117 \quad (3)$$

$$x_4^3 + x_1x_2 + x_5 = 8{,}145{,}260 \quad (4)$$

$$x_5^3 + x_1x_3 + x_4 = 3{,}356{,}396 \quad (5)$$

for an all integer solution in the range $0 \le x_i \le 300$ for i = 1, 2, 3, 4, 5.

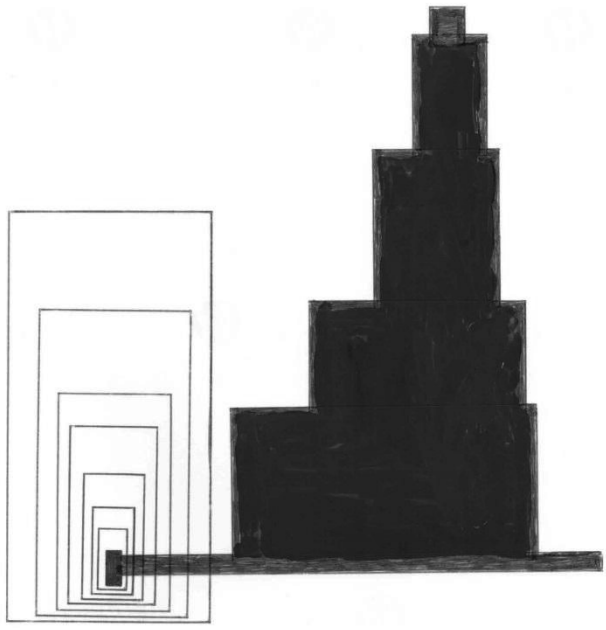


Figure 1: MSMCO Simulation Solving a System of Equations

First transform it to Minimize $f(x_1, x_2, x_3, x_4, x_5) = |L_1–R_1| + |L_2–R_2| + |L_3–R_3| + |L_4–R_4| + |L_5-R_5|$ subject to $0 \le x_i \le 300$ and all whole numbers for i=1, 2, 3, 4, 5 where L_{j-} and R_j are the left and right hand sides of equation j for j = 1, 2, 3, 4, and 5. Then select an eight stage MSMCO simulation drawing 5,000 feasible solutions at each stage with a funneling factor of F=2 to

reduce the search region in each subsequent stage by cutting the dimensions in half. Less than a second of computer run time produces the solution $x_1 = 258$, $x_2 = 95$, $x_3 = 187$, $x_4 = 201$ and $x_5 = 149$.

SIX EQUATIONS AND TWENTY VARIABLES

An electrical circuit requires four banks of five groups of resistors (each with an on/off switch) in a parallel series arrangement as described in Figure 2. The little buttons (circles) above each of the 20 resistors can be pressed or not to connect the individual resistor or disconnect it to the overall circuit. One can see that as long as at least one resistor in each of the four banks is connected to the circuit the current will flow through this four bank section of the overall electrical circuit.

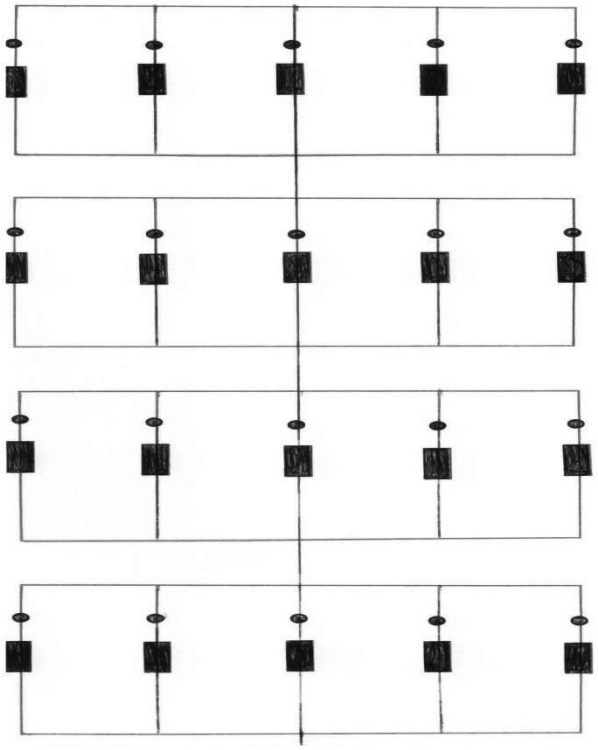


Figure 2: A Push Button Switching Resistance Network

The designers require at this point that when the following switches are turned on (connected) that the resistance will be 88 ohms, 117 ohms, 130 ohms, 140 ohms, 150 ohms, and 165 ohms in this part of the circuit. Also, each resistor cannot be larger than 700 ohms. Let the resistance values be x_i for i = 1, 2, 3 … 20 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}

The total resistance $R = R_1 + R_2 + R_3 + R_4$ for four banks if they are in series. However, if five resistors are in parallel as in this circuit in Figure 2, then

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} + \frac{1}{R_4} + \frac{1}{R_5}$$

Combining the series and parallel laws of resistance to match our Figure 2 circuit with the following switches in the on (connected) position, we get the following system of six equations in 20 unknowns.

With switches 3, 4, 5, 8, 9, 10, 13, 14, 15, 18, 19 and 20 in the off position and 1, 2, 6, 7, 11, 12, 16 and 17 in the on position the total ohms should = 88 ohms

$$\frac{1}{\frac{1}{X_1} + \frac{1}{X_2}} + \frac{1}{\frac{1}{X_6} + \frac{1}{X_7}} + \frac{1}{\frac{1}{X_{11}} + \frac{1}{X_{12}}} + \frac{1}{\frac{1}{X_{16}} + \frac{1}{X_{17}}} = 88 \text{ ohms} \quad (6)$$

And with switches 1, 3, 6, 8, 12, 13, 16 and 18 in the on position and the rest of them off, the total resistance should be

$$\frac{1}{\frac{1}{X_1} + \frac{1}{X_3}} + \frac{1}{\frac{1}{X_6} + \frac{1}{X_8}} + \frac{1}{\frac{1}{X_{12}} + \frac{1}{X_{13}}} + \frac{1}{\frac{1}{X_{16}} + \frac{1}{X_{18}}} = 117 \text{ ohms} \quad (7)$$

Also, with switches 2, 3, 7, 8, 11, 12 and 17 and 18 in on position we require

$$\frac{1}{\frac{1}{X_2} + \frac{1}{X_3}} + \frac{1}{\frac{1}{X_7} + \frac{1}{X_8}} + \frac{1}{\frac{1}{X_{11}} + \frac{1}{X_{12}}} + \frac{1}{\frac{1}{X_{17}} + \frac{1}{X_{18}}} = 130 \text{ ohms} \quad (8)$$

Then with switches 3, 4, 8, 9, 13, 14, 18 and 19 on

$$\frac{1}{\frac{1}{X_3} + \frac{1}{X_4}} + \frac{1}{\frac{1}{X_8} + \frac{1}{X_9}} + \frac{1}{\frac{1}{X_{13}} + \frac{1}{X_{14}}} + \frac{1}{\frac{1}{X_{18}} + \frac{1}{X_{19}}} = 140 \text{ ohms} \quad (9)$$

Switches 4, 5, 9, 10, 14, 15, 19 and 20 turned on must yield

$$\frac{1}{\frac{1}{X_4} + \frac{1}{X_5}} + \frac{1}{\frac{1}{X_9} + \frac{1}{X_{10}}} + \frac{1}{\frac{1}{X_{14}} + \frac{1}{X_{15}}} + \frac{1}{\frac{1}{X_{19}} + \frac{1}{X_{20}}} = 150 \text{ ohms} \quad (10)$$

While with switches 2, 5, 8, 10, 14, 15, 16 and 20 on

$$\frac{1}{\frac{1}{X_2} + \frac{1}{X_5}} + \frac{1}{\frac{1}{X_8} + \frac{1}{X_{10}}} + \frac{1}{\frac{1}{X_{14}} + \frac{1}{X_{15}}} + \frac{1}{\frac{1}{X_{16}} + \frac{1}{X_{20}}} = 165 \text{ ohms} \quad (11)$$

are required.

Using the statistical optimization approach (multi stage Monte Carlo optimization or MSMO) first the system of equations is transformed to minimize $f(X_1, X_2, X_3, \ldots X_{20}) = |L_1 - R_1| + |L_2 - R_2| + |L_3 - R_3| + |L_4 - R_4| + |L_5 - R_5| + |L_6 - R_6|$ again with the L_j and R_j as the left and right hand sides of the equations subject to $0 \leq X_i \leq 700$ ohms (Figure 1 gives a partial geometric and statistical illustration of this process).

Then a 50 stage MSMCO computer simulation draws

10,000 feasible solutions at each stage in an ever narrowing and repositioning search for better and better answers (less error). Presented here is the first and last ten stages of the printout. The answer follows the stage 50 total error of .00002 ohms. This program ran in less than 5 seconds on a desk top computer. This simulation used about six or seven places of accuracy. However, by switching to double precision arithmetic 12 to 14 decimal places of accuracy could be guaranteed if necessary.

The printout of the resistance network solution stages is $f(X_1, X_2 \ldots X_{20})$.

Stage Number	Total Error
1	568.21594
2	480.30206
3	323.37311
4	290.76581
5	75.25645
6	50.41321
7	50.00489
8	37.50713
9	16.86428
10	13.82444
□	□
□	□
□	□
41	.00034
42	.00023
43	.00017
44	.00008
45	.00008
46	. 00007
47	. 00004
48	.00004
49	.00003
50	.00002

The printout solution is:

X_1=546.48645	X_2=4.11427	X_3=66.22755	X_4=25.03539
X_5=95.81268	X_6=43.34272	X_7=229.82501	X_8=163.97571
X_9=113.46217	X_{10}=197.73457	X_{11}=204.15941	X_{12}=9.73828
X_{13}=328.54095	X_{14}=56.42705	X_{15}=458.11661	X_{16}=40.52254
X_{17}=653.32147	X_{18}=21.84053	X_{19}=9.49100	X_{20}=44.35545

AN OVER CONDITIONED SYSTEM OF EQUATIONS

Presented here is a nonlinear system of equations where the number of equations exceeds the number of variables. The theory of linear algebra might call this an over conditioned system and the electrical resistance network equations presented earlier an under-conditioned system (less equations than variables). The five equation, five variable first problem could be viewed as uniquely conditional (probably only one solution). However, those rules and linear algebra theory do not apply to nonlinear systems such as we have here.

However, it is worth mentioning that statistical optimization (MSMCO) can be tried on most any system of equations. However, as the number of equations grow in relation to the number of variables, there may be less answers to find (or no answers at all) and generally the problem is more difficult to solve (but not necessarily impossible).

Let us look at the following 8 variable 12 equation system and try to solve it for an all whole number solution by doing real valued arithmetic with MSMCO and rounding off the last stage answer to whole numbers.

This system of equations is:

$$x_8^3 + x_1x_3 + x_2 = 1{,}383{,}568 \quad (12)$$

$$x_2^3 + x_2x_4 + x_3 = 8{,}633{,}257 \quad (13)$$

$$x_5^3 + x_5x_6 + x_6 = 3{,}964{,}346 \quad (14)$$

$$x_6^3 + x_7x_8 + x_7 = 2{,}005{,}416 \quad (15)$$

$$x_4^3 + x_2x_7 + x_1 = 690{,}868 \quad (16)$$

$$x_7^3 + x_3x_6 + x_8 = 102{,}828 \quad (17)$$

$$x_1^3 + x_1x_8 + x_4 = 5{,}019{,}280 \quad (18)$$

$$x_3^3 + x_4x_5 + x_5 = 792{,}750 \quad (19)$$

$$x_1 + x_4 + x_7 + x_8 = 415 \quad (20)$$

$$x_2 + x_3 + x_5 + x_6 = 581 \quad (21)$$

$$x_1x_2 + x_2\,x_3 + x_3\,x_4 = 62{,}011 \quad (22)$$

$$x_4x_5 + x_5\,x_6 + x_7\,x_8 = 38{,}807 \quad (23)$$

with $0 \le x_i \le 250$ for i=1, 2, 3 . . . 8 and all whole numbers.

The system is transformed to minimize $f(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8) = |L_1–R_1| + |L_2–R_2| + |L_3–R_3| + |L_4–R_4| + |L_5-R_5| + |L_6-R_6| + |L_7–R_7| + |L_8–R_8| + |L_9–R_9| + |L_{10}–R_{10}| + |L_{11}-R_{11}| + |L_{12}-R_{12}|$ subject to $0 \le x_i \le 250$ and all whole numbers for i=1, 2, 3 . . . 8 and L_j and R_j are the left and right hand sides of equation j for j=1, 2, 3, . . . 12.

The MSMCO simulation used drew 1,000,000 all real valued feasible solutions at each of forty stages in the MSMCO simulation. The dimensions in each stage were reduced in size by 1.41 (or the square root of 2) and "funneled" into the fortieth stage values of $x_1 = 171.000$, $x_2 = 205.024$, $x_3 = 92{,}000$, $x_4 = 88.000$, $x_5 = 158.000$, $x_6 = 126.000$, $x_7 = 44.999$, $x_8 = 111.00$ which, when rounded, to the nearest whole numbers, produces an exact solution of $x_1 = 171$, $x_2 = 205$, $x_3 = 92$, $x_4 = 88$, $x_5 = 158$, $x_6 = 126$, $x_7 = 45$, $x_8 = 111$ with no error or $f(x_1, x_2, x_3 \ldots x_8) = 0$.

The forty stage printout is:

Stage Number	Total Error
1	2,979,655
2	2,839,862
3	2,196,819
4	1,824,941
5	1,226,064
6	859,745
7	387,499
8	387,499
9	307,404
10	217,602
11	164,887
12	117,912
13	89,859
14	63,962
15	28,556
16	28,556
17	17,463
18	17,131
19	14,388
20	7,301
21	5,991
22	3,391
23	2,666
24	1,572
25	1302
26	785
27	489
28	411
29	411
30	309
31	180
32	160
33	116
34	94
35	87
36	65
37	63
38	56
39	53
40	51

Notice stage forty did have a total error of about 51. However, once the eight variable values are rounded to the nearest whole numbers the answer is exactly correct with no error.

Another solution option would have been to just do all integer arithmetic in the whole simulation.

An additional option would be to try three solution attempts of the type done here and then any of the variables that produce the same value on all three attempts would be pinned down for a fourth MSMCO solution attempt. This modal averaging (with a mode of 3) would easily produce the exact optimal on the fourth MSMCO attempt even if the first three failed.

This repeated modal averaging was done with MSMCO by the author on the well-known (but at the time unsolved) test problems number 30, 31, 32 and 33 in the mathematics literature (Conley 1991a), (Conley 1991b and (Conley, 1993) after transforming the problems to transportation examples with 150 and 200 variables.
The true optimals were produced. This averaging and rounding technique (and additional transformations) can be effective sometimes on difficult or seemingly unsolvable optimization problems or systems of equations.

A CAPACITANCE SYSTEM

The parallel series laws for electrical capacitators are somewhat reversed from the laws for electrical resistance. For capacitators in series
$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} + \frac{1}{C_4} + \frac{1}{C_5}$ where C is the total capacitance and C_1, C_2, C_3, C_4 and C_5 are the capacitances of the individual capacitators.

Also, $C = C_1 + C_2 + C_3 + C_4 + C_5$ says that for capacitators in parallel the total capacitance is the sum of the individual capacitators values. Please see Figure 3 for a 5x5=25 capacitator bank with on/off toggle switches.

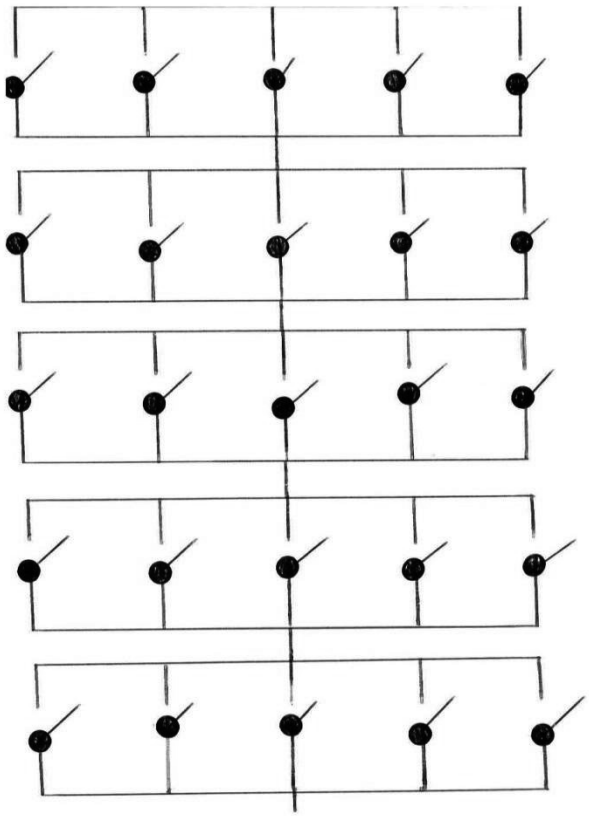


Figure 3: A Toggle Switching Capacitance Network

Therefore, for the example here if capacitators 1, 5, 6, 10, 11, 15, 16, 20, 21 and 25 are the only ones connected to the network in Figure 3 the required total capacitance is 3,500 microfarads expressed in equation (24).

Again, referring to Figure 3, the variables used here from left to right and top to bottom are 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}

Also, if only capacitators 2, 4, 7, 9, 12, 14, 17, 19, 22 and 24 are connected then a total capacitance of 3000 microfarads is required in equation (25). Additionally, if only capacitators 3, 4, 8, 9, 13, 14, 18, 19 and 23, 24 are connected a total capacitance of 2500 microfarads is required in equation (26). And, if only capacitators 1, 3, 7, 8, 12, 13, 16, 18, 21 and 23 are connected a total capacitance of 4000 microfarads is needed (equation 27). Continuing if only capacitator 2, 5, 6, 9, 11, 13, 18, 19, 22, and 24 are connected, the capacitance required is 4500 microfarads in equation (28).

Therefore, to solve this five equation twenty-five variable nonlinear capacitance system of equations, the approach is to try to minimize $f(x_1, x_2, x_3, \ldots x_{25}) = |L_{24}-R_{24}| + |L_{25}-R_{25}| + |L_{26}-R_{26}| + |L_{27}-R_{27}| + |L_{28}-R_{28}|$ where L_j and R_j are the left and right hand sides of equation j for j = 24, 25, 26, 27, and 28. A fifty stage MSMCO statistical optimization simulation is attempted drawing 10,000 feasible solutions and always storing the best answer so far, and repositioning the search about the new best answer. Also, after every stage of 10,000 feasible solutions, the next stage has its search dimensions reduced by a factor of 1.41 (the square root of 2). This allows the simulation to "funnel" (so to speak) into a useful solution.

The printout of the solution coordinates for this few seconds computer run follows here after the system of equations.

$$\frac{1}{\frac{1}{X_1+X_5}+\frac{1}{X_6+X_{10}}+\frac{1}{X_{11}+X_{15}}+\frac{1}{X_{16}+X_{20}}+\frac{1}{X_{21}+X_{25}}}=3{,}500 \text{ micro-farads (24)}$$

$$\frac{1}{\frac{1}{X_2+X_4}+\frac{1}{X_7+X_9}+\frac{1}{X_{12}+X_{14}}+\frac{1}{X_{17}+X_{19}}+\frac{1}{X_{22}+X_{24}}}=3{,}000 \text{ micro-farads (25)}$$

$$\frac{1}{\frac{1}{X_3+X_4}+\frac{1}{X_8+X_9}+\frac{1}{X_{13}+X_{14}}+\frac{1}{X_{18}+X_{19}}+\frac{1}{X_{23}+X_{24}}}=2{,}500 \text{ micro-farads (26)}$$

$$\frac{1}{\frac{1}{X_1+X_3}+\frac{1}{X_7+X_8}+\frac{1}{X_{12}+X_{13}}+\frac{1}{X_{16}+X_{18}}+\frac{1}{X_{21}+X_{23}}}=3{,}000 \text{ micro-farads (27)}$$

$$\frac{1}{\frac{1}{X_2+X_5}+\frac{1}{X_6+X_9}+\frac{1}{X_{11}+X_{13}}+\frac{1}{X_{18}+X_{19}}+\frac{1}{X_{22}+X_{24}}}=4{,}500 \text{ micro-farads (28)}$$

subject to $10 \le x \le 100{,}000$ microfarads upper limit on individual capacitators for i=1, 2, 3, . . . 25.

The printout answer for

x_1=7642.6216	x_2=11863.8330	x_3=21466.3770
x_4=15333.8799	x_5=28376.4180	x_6=18855.1875
x_7=7679.6050	x_8=4658.4243	x_9=1296.3455
x_{10}=9913.1416	x_{11}=2133.1001	x_{12}=11529.0068
x_{13}=25480.7969	x_{14}=4001.8213	x_{15}=3834.6125
x_{16}=29435.1660	x_{17}=14696.3193	x_{18}=27944.4727
x_{19}=17130.3594	x_{20}=23035.1055	x_{21}=6368.3716
x_{22}=9195.6914	x_{23}=4724.4390	x_{24}=1996.4415
		x_{25}=20980.98.24

Stages	Total Stage Error
1	2425.28760
2	2044.93896
3	900.61182
4	894.59595
5	802.41040
6	613.55273
7	440.20532
8	372.60815
9	213.61523
10	177.54297
11	127.21094
12	56.91919
13	53.88916
14	35.41992
15	26.86377
16	18.78198
17	10.80054
18	10.80054
19	7.0896
20	5.44214
21	3.096119
22	2.71631
23	1.67212
24	1.33545
25	.75439
26	.65649
27	.29590
28	.25366
29	.23584
30	.08423
31	.08423
32	.08423
33	.04126
34	.04102
35	.03076
36	.01563
37	.01270
38	.01123
39	.00806
40	.00562
41	.00244
42	.00195
43	.00146
44	.00098
45	.00049
46	.00024
47	.0000
48	.0000
49	.0000
50	.0000

CONCLUSION

A selection of nonlinear systems of equations was presented and solved using the multi stage Monte Carlo optimization (MSMCO) simulation technique.

One system had the same number of equations as variables and required whole number answers. Another had more equations than variables, while the other two had fewer equations than variables and required real value solutions.

A nice feature of the approach is that if the last stages of the MSMCO simulation produce a total error term of zero, one not only receives an answer but it automatically checks in each equation.

Additionally, if the process did not converge to an exact solution and the total error term is relatively small, it still might be a useful solution in the worlds of science, big business, and economics where sometimes approximate solutions that are arrived at quickly are preferred.

More examples of MSMCO problems (or statistical optimizations) and statistical analyses in linear and nonlinear settings are in (Anderson, Sweeney and, Williams 1999), (Anderson 2003), (Black 2014), (Conley 1981), (Conley 2009), (Hayter 2002) and (Keller and Warrack (2003). Statistical optimization is a new application area for survey sampling in our computer age.

It should be mentioned that in a sense the statistical optimization process (MSMCO) runs throughout with a random number generator. Many of these generators are a type of abstract algebra function which can repeat itself after a while.

So, for example, after generating may be one hundred thousand random numbers between 0 and 1 the function might start repeating the same sequence over again.

This worry does not matter with multi stage Monte Carlo optimization (MSMCO) because every time a better answer is found the search re-centers about that new "best answer so far. Also, after each stage of this multi stage process the search dimensions are reduced. Therefore, this free flowing ever moving (see Figure 1) and changing of the n dimensional "rectangles" (for a function of n variables) completely re-randomizes even if the generator ever starts to repeat. The same would be true if the random numbers were read in from a big file instead of being generated by an abstract algebra function. Survey sampling is useful in estimation problems. Additionally, hooked up to a computer, it can also find optimal solutions and solve systems of equations in our computer age. Just make sure that the preliminary survey sampling (in stage one) of the optimization is followed by subsequent survey samples in an ever reducing search area. This allows MSMCO to close in on the right answer.

REFERENCES

Anderson, D.R., Sweeney, D. J., Williams, T. A. 1999. *Statistics for Business and Economics*. 7th Edition. South-Western 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. 1981. *Optimization: A Simplified Approach*. Petrocelli Books. Princeton and New York.

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

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

Conley, W. C. 1993 "Multi Stage Monte Carlo Optimization Applied to a Three Hundred Point Traveling Salesman Problem." *Mathematical Modeling and Scientific Computing,* Principia Scientia, St. Louis, Madison, pp. 298-403.

Conley, W. C. 2009. "Linear and Nonlinear Input Output Analysis." Proceedings of 7th International Industrial Simulation Conference, ISC, 2009, Loughborough, U.K. June 1-3, EUROSIS-ETI, Ghent, Belgium, pp. 39-41.

Hayter A. J. 2002. *Probability and Statistics for Engineers and Scientists*, 2nd Edition. Duxbury Press, Pacific Grove, California.

Keller, G. and Warrack, B. 2003. *Statistics for Management and Economics*, 6th Edition. Thompson Brooks/Cole, Pacific Grove, CA.

AUTHOR BIOGRAPHY

WILLIAM CONLEY received a B.A. in mathematics (with honors) from Albion College in 1970, an MA in mathematics from Western Michigan University in 1971, MSc in statistics in 1973 and a PhD in mathematics (computer statistics) from the University of Windsor in 1976. He taught mathematics, statistics, and computer programming in universities for 35 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 200 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.

GRAPHICAL DATA SIMULATION

MULTI-MASTER REPLICATION IN EVENTUALLY CONSISTENT SIMULATION GRIDS

Stefan Elsen
University of Trier, Department of computer science
email: elsen@syssoft.uni-trier.de

KEYWORDS
Eventual Consistency, Grid, Distributed computation, Peer-to-peer, Game, Logic Process, Multi-Master Replication

ABSTRACT

Based on a previous work about eventually consistent simulation grids, this paper introduces novel mechanisms to synchronize distributed simulation states by utilizing inconsistency maps provided by the original simulation system. Using merge operations as a part of a multi-master replication scheme, the overall resilience of the system can be increased, and the quality of remaining inconsistent entities improved. Different strategies are presented to optimize the merge output, and quality profiles introduced to establish comparability.

INTRODUCTION

In the previous work, Elsen (2016), an approach was introduced to design eventually consistent, spatial-parallel simulations.
In this approach a continuous simulation space is distributed across an array of participating logical nodes, such that each is responsible for a bounded domain. To maintain availability in the presence of hard- or software faults the system may temporarily enter a partial inconsistent state. Simulation continues while nodes adaptively record their simulation and communication history in order to recover to consistency when feasible. To minimize communication, work, and storage demand, each node maintains a binary mask of regions affected by detected faults. As recovery removes inconsistency, the masks eventually return to an empty state.
While the outlined approach is functional, and provides some resilience against random, short-lived faults, it has proven less effective during extended simulations.
With increasing size of the connected inconsistent area, the number of affected components, and conversely the probability of contained new faults grows; as does the resulting impact on recovery.
Fig.1 depicts the average survived time-steps of a 3D simulation scenario. Independent fault probabilities and favorable migration conditions of the simulation scenario can not prevent eventual termination.

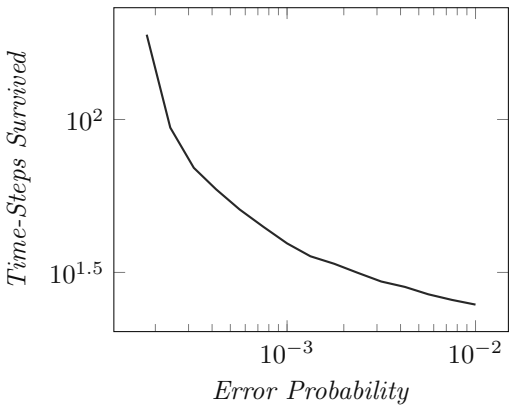


Figure 1: Measurement of average survival rounds across 30 runs, given independent node fault probability (per lockstep operation)

To counteract the problem, component reliability may be increased, and/or additional computation time allocated for recovery. However, the properties of consistency-masked simulation states lend themselves to merge operations, and thus physical redundancy, increasing reliability without changing component quality or sacrificing responsiveness. Similar in concept to Multi-Master Replication in database systems, this approach allocates multiple physical nodes for regions previously governed by each one node.
In this paper, the basic approach of replication and merging is discussed as well as its impact on system performance. It is shown that merge operations can significantly increase overall system resilience, while reducing the extent of inconsistency present in the simulation. Various strategies are discussed that help further reduce average entity inconsistency in remaining inconsistent areas.
While primarily an addition to the system outlined in the original paper, the presented merge operations are considered applicable to any redundancy based simulation architecture utilizing spatial datasets of which consistency masks can be extracted or generated.

BACKGROUND AND STATE OF THE ART

Recent years have seen increasing popularity of distributed data stores that provide simple storage and

retrieval operations based on large numbers of distributed machines, consistent of commodity hardware, as opposed to traditional high-performance, high-cost server architectures. NoSQL (commonly but not universally read 'Not only SQL'), and the accompanying classification acronym BASE (Basically Available, Soft state, Eventually consistent) are particularly popular concepts. More detailed definitions are given by Cattell (2011). These systems achieve what Bondi (2000) defined as 'horizontal scalability', meaning that more performance can be achieved simply by adding more computers to a given infrastructure.
Eventual consistency (Vogels (2009)) describes systems that provide no guarantee regarding their current consistency, but maintain availability in the presence of faults. When consistency is eventually restored, newly received updates may result in conflicts, similar to merge-conflicts in Distributed Version Control (e.g. Brindescu et al. (2014), Chacon and Straub (2014), O'Sullivan (2009)). Shapiro et al. (2011) aim to mitigate this issue by introducing conflict-free data types.

TERMINOLOGY AND METHODOLOGY

This paper is based on Elsen (2016), sharing most of its terminology.
The term *shard* is used to describe a single logical node in the grid that communicates with its neighbors and potential clients. Each shard is responsible for simulating the entities residing in its assigned space. This space segment is denoted its Shard Domain (SD). Each SD is inconsistency-masked, proving an n-dimensional map called Inconsistency Coverage (IC) per version. The IC tracks the current state of inconsistency as a consequence of past faults in the grid, and expands by the maximum entity influence radius R as the simulation progresses (see fig.2).

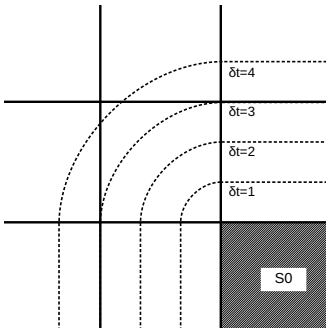


Figure 2: IC Expansion around lost shard S0

To maintain a single continuous space, a shard's SD edges border those of its neighbors. Simulated entities remain agnostic of these internal borders.
With each global time-step, all shards generate each one new Shard Domain State (SDS), which contains all simulated entities currently within the local SD, along with the respective IC.

Original Protocol

Using a wall-clock synchronized lockstep protocol, shards communicate changes in the form of Remote Change Sets (RCSs) with their neighbors in order to evolve the global state.
Only the newest SDSs are visible to user clients, allowing a shard to discard older SDSs if they are not required for correction operation. Specifically, a shard is not required to maintain any SDSs older than its newest fully consistent SDS. Newer inconsistent SDSs, however, must be maintained until correction operations can restore their state to consistency.
During each global time-step, a fixed number of lockstep-synchronized correction steps (C) are executed which may produce less inconsistent SDSs, ideally to the point of no inconsistency. On average, $C > 1$ must be true to eventually reach consistency. Increasing C can dramatically improve grid resilience but reduces responsiveness/availability due to the increased amount of time spent on maintenance.

MULTI-MASTER REPLICATION

Multi-Master Replication, when applied to this system, aims to increase the number of physical shards responsible for the same SD. It may be employed with varying density, replicating only some shards, or by providing a dedicated, actively redundant backup cluster as an on-demand migration target.
In addition to its regular neighbor links, a replicated shard maintains a connection to at least one sibling. This setup represents the minimal number of links necessary for the system to operate under uniform conditions. Changes in replication degree between neighbors require additional links to sibling neighbors in addition to local neighbors.
The number of siblings per shard directly affects system resilience, especially when covering error classes like arbitrary and byzantine. Less critical error classes like omission, timing, and crashes can be handled by the standard non-replicated setup, but benefit from additional siblings.
The remainder of this paper focuses on two siblings, but the concepts are easily expanded to higher numbers.

STATE MERGE OPERATIONS

Whenever siblings synchronize inconsistent data, a merge operation is triggered. This operation requires two simulation states as input and aims to produce a new state of reduced inconsistency. Upon completion, the result is adopted by both siblings.
The process requires samples from both input ICs to produce any improvement. By selecting entities from

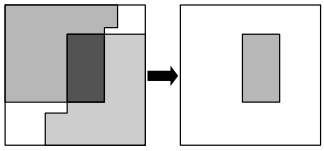


Figure 3: Example of intersecting ICs (dark=potentially inconsistent) and the resulting merged IC

the respective consistent section of either input SDS, large areas of inconsistency are potentially removed. If the two ICs do not intersect at all, then the resulting state is fully consistent. This event is denoted a *Perfect Merge* from this point on.

Constraints

In order to maintain the requirements of the original system during merge operations, a number of conditions must be met. The resulting merged IC must not underestimate the intersection between the two input ICs. If an area is considered consistent by the resulting IC, then it must have been consistent in either input SDS, and preserve any contained, consistent entities. Furthermore, the merged consistent space must not contain any entities flagged as potentially inconsistent by their respective ICs.

Resilience Evaluations

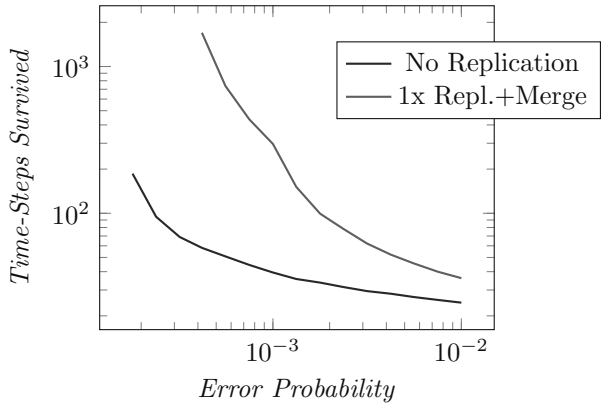


Figure 4: Survivability comparison. Omitted graph sections have exceeded the maximum runtime of 10k time-steps in more than 10% of all runs

Fig.4 shows a comparison between non-replicated, and single-merge-replicated simulation runs, using a simulation-implementation of the original protocol plus the introduced merge logic. Data persistence is maintained by a simulated distributed database.
The usage of simulation over real-life measurements allows consistency verification and close control of general

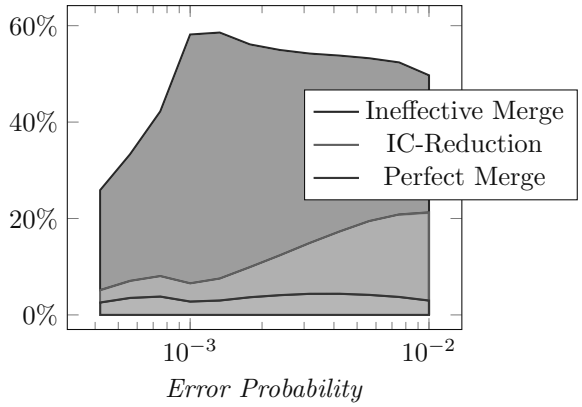


Figure 5: Merge events probabilities per shard and time-step

state, fault simulation, and recovery progression. While arguably closer to any actual application, real-life measurements remain equally hard to generalize due to flexibilities in entity behaviors, highly adjustable hard- and software component configurations, and the risk of unplanned faults occurring in large-scale setups.
Fig.5 shows attempts and outcomes of merge operations. Perfect merges appear to maintain a mostly constant probability of between 2.5% and 4.5% per shard-pair and time-step. Despite the apparently low success rate, the strategy appears to have a profound impact on grid resilience, however, real-life results are expected to vary.
At most one generation was merged each time-step, favoring the oldest available generation. Frequent hash-digest broadcasts to neighbors are required to negotiate the respective merge choice.

Strategies

The merge requirements determine the handling of consistent entities, but notably leave out entities not located in consistent sections in either input SDS. Once all consistent entities have been placed, and assuming some inconsistent locations remain, the contained entities may be inserted into the merged inconsistent space. In this regard, different strategies can cause considerable statistical divergence.

Quality Profile
In order to establish comparability between different strategies, a number of measurable metrics are combined into a profile. The applied metrics are the number of missing (M) and unwanted entities (U), entity inconsistency probability (P), and average spatial entity inconsistency (Ω). Ω is calculated as the integral of spatial divergence over time, as defined by Zhou et al. (2004), and presented in multiples of R per entity.
To retrieve the necessary values, a 16x16 2D shard grid (each 8x8R) was filled with 260k aggressively

environment-sensitive entities with both motion and sensor radius limited to $0.5R$. Only entities within the remaining (merged) inconsistent area are observed; entities of in- or output are disregarded if they lie outside this area. The metrics are determined by comparing the in- and outputs with a consistent control grid, averaged across one million merge operations.
The resulting vector is considered the *Quality Profile* of a strategy.

Individual Entity Selection
Prog.1 defines a strategy to improve merge-results by means of individual placement of inconsistent entities.

Program 1 Individual entity selection algorithm

Let:
SDS_0, SDS_1 : Input SDSs
$IC_0 \in SDS_0, IC_1 \in SDS_1 : Entity \rightarrow \{0,1\}$: IC functions, classifying inconsistency of an entity. 0 indicates full consistency.
$e_0 \in SDS_0, e_1 \in SDS_1$: Incarnations of inconsistent entity e. e_0 or e_1 exist, or both.
$L(e_0, e_1)$: Orthographic lesser of e_0, e_1.
$Insert(e)$: Insert entity e into the resulting SDS.
$Insert_*(e)$: Insert entity e into the closest inconsistent location in $min(IC_0, IC_1)$.
$i, j \in \{0,1\}, i \neq j$: Source indexes.

$\exists e_0 \wedge \exists e_1 :$
  $IC_j(e_i) = 0 \wedge IC_i(e_j) \neq 0 : Insert(e_j)$
  $IC_1(e_0) \neq 0 \wedge IC_0(e_1) \neq 0 :$
    $IC_0(e_0) = IC_1(e_1) : Insert(L(e_0, e_1))$
    $IC_i(e_i) < IC_j(e_j) : Insert(e_i)$
  $IC_1(e_0) = 0 \wedge IC_0(e_1) = 0 : Insert_*(L(e_0, e_1))$,
$\exists e_i \wedge \nexists e_j :$
    $IC_i(e_i) \leq IC_j(e_i) : Insert(e_i)$

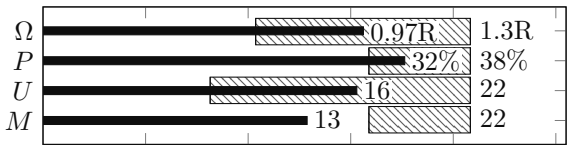


Figure 6: Averaged results of the Individual Entity Selection strategy

Results are depicted in Fig.6. The hatched areas depict the range between average minimum and maximum value of the two input SDSs, and give some impression of how the result compares to the data it was based on. The axis maximum is taken from the worse of the two input SDSs, and printed at the right edge.
Reducing inconsistency beneath the better input SDS is hard, as 96% of remaining inconsistent entities lie within the merged inconsistent area. Further information is unavailable to properly distinguish multiple incarnations of the same entity if both reside within the remaining inconsistent area.
The strategy decreases the number of missing entities, which arguably presents a noticeable improvement to a human observer. If an entity is incorrectly missing from one input, there is a chance it is present in the other, and ultimately included. Unfortunately this statement is also true for unwanted entities, resulting in increased numbers being added to the result.

Exclusive Input Selection
Based on the observation that entities of the same input state exhibit similar levels of inconsistencies within their IC, it is reasonable to preserve inconsistent entities only from one of the two SDSs, relocating them if necessary. The input SDS with least accumulated IC values is selected for inclusion, the inconsistent entities of the other are discarded. For binary ICs, the accumulated value matches the IC volume, which appears to be an acceptable but imperfect indicator.

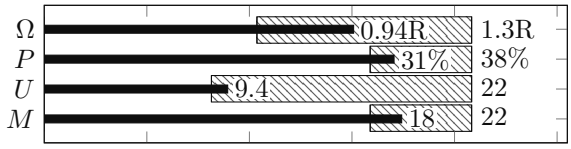


Figure 7: Averaged results of the Exclusive Input Selection strategy

Results are shown in Fig.7. This strategy generally produces better results than individual selection at the expense of increased missing entities.
It is assumed that the choice based on IC samples, on average, results in the selection of the input of lesser inconsistency. Additionally, it was observed that the viral nature of inconsistency causes similarities in entity inconsistency within one input state. As can be seen in the hatched areas of the profiles, on average the better of the two input SDS displays half the level of inconsistency of the other. Coupled with reduced numbers of unwanted entities—each inherently contributing non-zero Ω—this approach results in a slight drop of average inconsistency.

Enhanced IC
Rather than representing inconsistency as a binary state, some improvements are gained from maintaining the number of time-steps an IC location has spent inconsistent, denoted as Divergence Depth (DD) in the original paper. Here DD=0 indicates full consistency; new inconsistencies are marked with DD=1, and subsequently incremented by one each time-step.
The sampled DD value of some location allows a coarse estimation of actual entity inconsistency.

The definition of the IC function changes to:
$IC'_0 \in SDS_0, IC'_1 \in SDS_1 : Entity \to \mathbb{N}_0$,
where 0 indicates full consistency.
The introduced algorithms for individual and exclusive entity selection can be based on IC' without further adjustments.

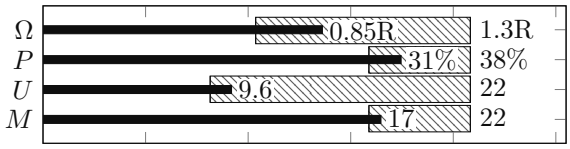


Figure 8: Averaged results of the Individual Entity Selection strategy, utilizing Divergence Depth

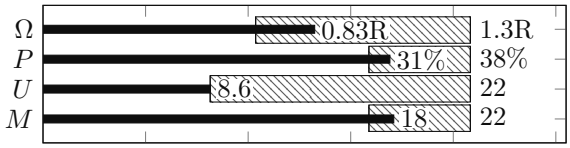


Figure 9: Averaged results of the Exclusive Input Selection strategy, utilizing Divergence Depth

Fig.8 and 9 show the results of individual and exclusive input selection when utilizing DD. Its inclusion as an additional indicator appears to generally improve the results of both algorithms, also increasing their similarity.

CONCLUSION AND FUTURE WORK

This paper enhances the concepts outlined in the previous paper by adding multi-master replication and state merge operations. It was shown that the increase in physical redundancy can positively affect the resilience of the previously original spatial-parallel simulation.
The information inherently provided by the system can be used to merge two consistency-masked simulation states into one single state, resulting in increased recovery performance and consequently system resilience.
Further research is necessary regarding partial replication, RCS merge operations, and ad-hoc migration/replication. There is reason to assume that node replication may be employed adaptively within the viewing area of users to reduce the risk of observed failure and/or inconsistency.
Real-life measurements are necessary to evaluate the simulated survivability rates.
Beyond the benefits for fault tolerance, merge operations can reduce spatial entity inconsistency within inconsistent areas, when compared to the average input pre-merge data. Various merge strategies are presented, including comparable results using quality profiles. Spatial inconsistency is hard, if not impossible, to reduce beneath the better of its two input SDSs due to the selective nature of the suggested strategies. The strategies can, however, reliably reduce the remaining spacial inconsistency to beneath the average of the two input SDSs.
Further research may reveal meaningful enhancements to the information provided by simulation and/or IC in order to improve merge quality.

REFERENCES

Bondi A., 2000. *Characteristics of Scalability and Their Impact on Performance. Proceedings of the 2nd international workshop on Software and performance*, 195–203. doi:10.1145/350391.350432. URL `http://portal.acm.org/citation.cfm?id=350432{\&}dl={\%}5Cnpapers2://publication/uuid/CC882EDE-5736-49BB-A98D-5941E4FD8F09`.

Brindescu C.; Codoban M.; Shmarkatiuk S.; and Dig D., 2014. *How do centralized and distributed version control systems impact software changes? Proceedings of the 36th International Conference on Software Engineering - ICSE 2014*, undefined, no. undefined, 322–333. ISSN 02705257. doi:10.1145/2568225.2568322. URL `http://dx.doi.org/10.1145/2568225.2568322{\%}5Cnhttp://dl.acm.org/citation.cfm?doid=2568225.2568322{\%}5Cnhttp://dl.acm.org/citation.cfm?doid=2568225.2568322`.

Cattell R., 2011. *Scalable SQL and NoSQL data stores. ACM SIGMOD Record*, 39, no. 4, 12. ISSN 01635808. doi:10.1145/1978915.1978919.

Chacon S. and Straub B., 2014. *Pro Git*, vol. 288. Apress, 2nd ed. ISBN 978-1484200773.

Elsen S., 2016. *A distributed simulation grid using eventual consistency.* In *30th European Simulation and Modelling Conference*. Las Palmas.

O'Sullivan B., 2009. *Mercurial: The Definitive Guide: The Definitive Guide.* " O'Reilly Media, Inc.".

Shapiro M.; Preguiça N.; Baquero C.; and Zawirski Z., 2011. *Conflict-free replicated data types.* Tech. rep. URL `http://link.springer.com/chapter/10.1007/978-3-642-24550-3{\_}29`.

Vogels W., 2009. *Eventually Consistent. Queue*, 52, no. 1, 14. ISSN 15427730. doi:10.1145/1466443.1466448.

Zhou S.; Cai W.; Lee B.s.; and Turner S.J., 2004. *Time-Space Consistency In Large-Scale Distributed Virtual Environments.* 14, no. 1, 31–47.

Visualising of Co-occurrence Data

Igor Litvine
Nelson Mandela University
Port Elizabeth 6031, South Africa
email: igor.litvine@mandela.ac.za

Oksana Ryabchenko
Nelson Mandela University
Port Elizabeth 6031, South Africa
email: oksana.ryabchenko@mandela.ac.za

KEYWORDS
Co-occurrence Analysis, Co-occurrence Matrices, Visualisation, Large Data

ABSTRACT

A new technique for analysis of co-occurrence matrices, based on visualisation and simulations is suggested. This approach is tested on some examples from real and simulated data.

Introduction

Co-occurrence analysis was initially introduced in language studies (known as lexical co-occurrence). In these publications researchers justify that co-occurrence of words in natural languages may be of either semantic ("car"-"vehicle") or associative ("car"-"driver") nature.

Co-occurrence analysis (COA) is based on counting of pairs of words (and/or terms) that appear simultaneously within a defined window (in selected clusters or systems). Traditionally, COA is used in building bibliographical catalogs (Leydesdorff and Vaughan (2006), Buzydlowski et al. (2002)). Nowadays one can see the increasing use of these techniques for analysis of marketing strategies (shopping cart analysis) (Karakatsanisa et al. (2017)) and revealing relations between market variables (Altuntas et al. (2016)).

This technique is very popular in analysing of big volumes of publications (e.g. Davis et al. (2017)). The researchers employ multiple methods (literature collection, Topic Modelling, and Co-occurrence Mapping of entities) to demonstrate how automated analyses of large bodies of literature may assist with important research tasks.

Co-occurrence matrices have many common features with contingency tables (aka cross-tabulation), which are widely used in Statistical studies for analysis of interrelations between two variables. While contingency tables and co-occurrence matrices have some principal differences, still cross-tabulation may be considered as a special case of co-occurrence tabulation. In the next section we discuss the challenges which arise when contingency tables are used for analysis of large data.

Statistical Inference on Large Samples

Everybody knows that statistical analysis based on a small sample is not reliable. Firstly, because small random sample may not be representative, in other words it may exhibit behaviour very different from what the population properties suggest. If a person has five daughters and not a single son this cannot serve as a proof that he could not have had a son - the sample is too small for such conclusions.

Secondly, many statistical tests are asymptotic (e.g. Chi-square test for independence) and therefore require reasonable number of observations to make the asymptotic theorems work. On the other hand, very big samples are no good for statistical testing either. The tests performed on large samples are very sensitive to small differences between a hypothetical value and true population value of the studied parameter.

Consider an example. Suppose, the following table gives cross-tabulation of two variables X (which may take two values 0 or 1 and Y (which also takes only two values A and B):

Table 1: Example 1

$X \setminus Y$	A	B
0	101	99
1	99	101

In the above example, we have a sample of size 400 to built this 2×2 cross-tabulation. The sample size is more than sufficient to apply Pearson's test on independence of X and Y. The value of the Chi-square test statistic is 0.04 (p-value is 0.84) and we cannot reject the null hypothesis that the variables are independent. However, if we build a new sample simply joining 100 of the initial samples (see table 2), we get $\chi^2 = 40$ and now $p = 0.0455$ and we can reject the null hypothesis at significance of 5%. Further continuing this process making the sample even 10 times bigger

we get $\chi^2 = 400$, $p \approx 0$ and we have to reject the hypothesis on the independence very confidently.

Table 2: Example 1 (continued)

$X \setminus Y$	A	B
0	10100	9900
1	9900	10100

This result is counter-intuitive. If the initial sample was representative, then the multiplied sample should be also representative, and is expected to yield the same inference. Hence, in case of large samples, statisticians employ so-called practical significance tests. For example, in case of testing for independence a Cramer's $V = \sqrt{\chi^2/(n\ df)}$ is used. However, Cramer's V is also not a flawless approach (e.g. it may be highly biased, overestimating the association).

Visualisation

Visualisation may be considered as an alternative approach for assessing association. One should remember that human eye is quite a precise measuring device. Measurements based on visualisation are used very widely in science, engineering and technology.

The association data (in other words the contingency tables) may be visualised in various ways, for example (a) 3d bar charts/histograms; (b) surface plots; and (c) heat maps. It is important to note that in all the above cases the images for original and multiplied samples will look exactly the same (see figure 1). This means that eye measuring based on visualisation will not depend on the sample size and may be used both for small and large data sets.

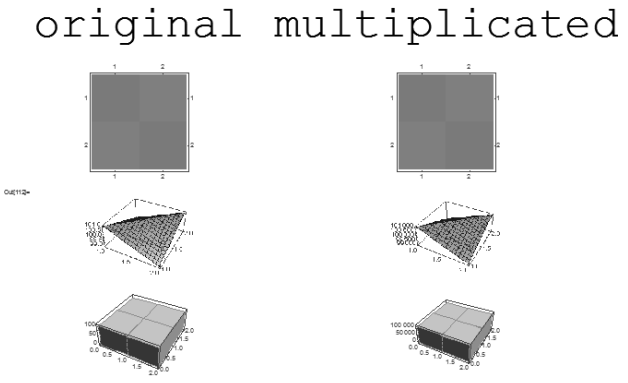


Figure 1: Visual images for example 1

Co-occurrence Analysis

The usual technique used in such studies is based on so-called hierarchical clustering. A distance between words (to be more precise - keywords) is defined. The pairs of keywords that have similar co-occurence properties are joined in clusters, then similar (close) clusters are joined in bigger clusters, and so on. This technique allows to find unobvious relationships between keywords that may assist in future research. Typical output of the hierarchical clustering (in a form of a heat map) is shown on figure 2.

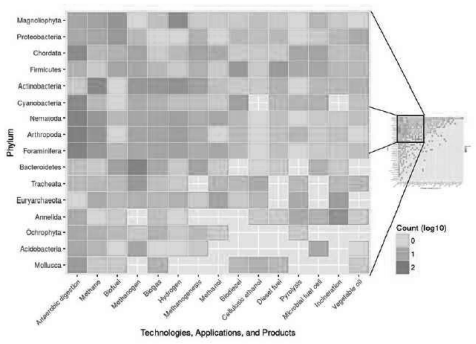


Figure 2: Typical output of Hierarchical Clustering

Statistical Methodology for Analysis of Co-occurrence Matrices

In this paper we present a technique alternative to the Hierarchical Clustering. The main idea of this approach is to manipulate with rows and columns of a co-occurrence matrix to obtain an image which will be interpretable from a statistical viewpoint. In other words the image should resemble some kind of a shape from which a statistician can (via visual inspection) draw conclusions about the population (for example inference on potential statistical associations).

Example: one dimension

Consider two histograms (figures 3 and 4). On the first site they look distinctively different. However, closer look suggests that the bars in the first chart were simply re-shuffled. The question we may ask now is whether we can restore the picture on the first figure from the second one? If so, what we need to know to successfully complete the task?

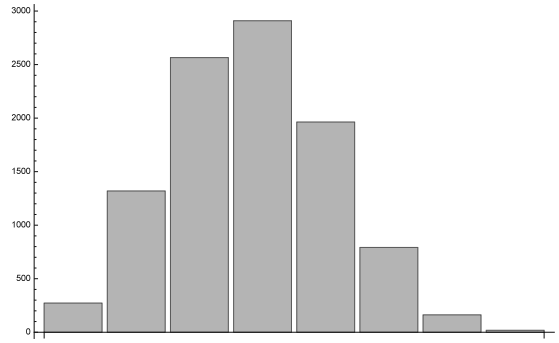


Figure 3: Histogram (a)

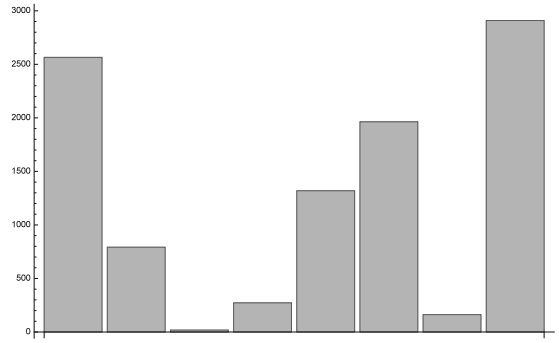


Figure 4: Histogram (b)

Fully known distribution

The task may be completed relatively easy if we fully know the distribution which was used to obtain the random sample. In our case the distribution was Binomial(7,0.4) (seven trials, probability of success is 0.4). The sample size was 10000. Knowing that, we can suggest the following procedure:

1. calculate the expected frequencies
2. match the calculated frequencies with the observed (e.g. using the least squares method)
3. order the bars accordingly

In the above example the frequencies are:

Table 3: Frequencies

	0	1	2	3	4	5	6	7
expected	279.9	1306.4	2512.7	2903.0	1935.4	774.1	172.0	16.4
observed (mixed)	2565	792	18	272	1319	1963	162	2909
observed (matched)	272	1319	2565	2909	1963	792	162	18

Known type of the distribution, unknown parameters

The task is complicated just a bit if we know the type of the distribution, but do not know the parameters. In this case we have to estimate the parameters for every possible permutation and select the permutation where the fit is the best (e.g. comparing the χ^2 statistics). Fortunately, in this case the number of the permutations is not so big ($8! = 40320$) and a computer can manage this task. On a moderate Desktop PC we received the following results (in under 20 seconds):

Table 4: Optimal permutations: fitting to Binomial distribution with unknown parameters

	0	1	2	3	4	5	6	7	p	n	χ^2
optimum 1	272	1319	2565	2909	1963	792	162	18	0.401	7	2.33
optimum 2	18	162	792	1963	2909	2565	1319	272	0.599	7	2.33

As one may see from the above table, we have two optimal solutions. The first one is exactly the same as the one received in the previous section. The other one is the mirror image of it (the frequencies are reversed and $p_2 = 1 - p_1$).

Also we can see that the value of the test statistics is $\chi^2_5 = 2.33$ in both cases. This value corresponds to the test p-value $p = 0.80$. Also, for sample of such large size (10 000) it is common to use Cramer's V-test for practical significance:

$$V = \sqrt{\frac{\chi^2_5}{n\ df}} = \sqrt{\frac{2.33}{10\ 000\ 6}} = 0.0062$$

Values of V less than 0.17 are regarded as an indication of a small practical significance and we may accept the following hypothesis: "the frequency table represents shuffled frequency table of a sample from Binomial distribution".

We should note that in the above example the frequencies are very easy to match. This is because: (a) sample size is large (10 000); (b) the number of cases (categories) is relatively small (eight).

Broken Glass Model

A scientific quest may be seen as a pursuit to sort the wheat from the chaff (see Matthew 13:24-30). Regularities in real processes (we shall refer to them as laws of nature) are hiding among chaotic disturbances of various kinds. If in the early years of Science revealing the laws of nature was relatively simple (myth says, that Archimedes was taking a bath when he discovered his

famous law, aka law of buoyancy), nowadays the Scientific research towards discovering new orderliness is complicated due to the following factors:

- need for complex sophisticated and very expensive equipment (e.g. Large Hadron Collider, Square Kilometre Array Radio-telescope, etc.);
- need to process large amounts of data and information of different sorts (e.g. pictorial or textual data).

Here we suggest a model which may be used for describing the challenges of the modern Science. Suppose one observes a heap of broken glass. One may see these pieces as a chaotic mix of unrelated debris. On the other hand, a question may be asked if these pieces are parts of a single object which we cannot immediately identify due to the "disturbance" (in this case the disturbance may be in the fact that the glass object was dropped on a hard surface, some pieces are missing, some pieces do not belong here, etc.).

The analysis of the glass debris may aim at answering the following questions: (a) can we accept/reject a hypothesis that the pieces (all of them or majority of them) are parts of a single disintegrated object? (b) If we accept the above hypothesis, what object was it (e.g. glass sheet, vase, sculpture, etc.)? (c) are these all the pieces or some are missing? Clearly the most natural way to proceed is to try to rearrange the trashes in such a way that some kind of an object is assembled. If we are successful, (d) can we formally test that this is the same object that was crashed?

Similarly, if we have lots of data which seem chaotic we may suggest an investigation along the following objectives:

1. rearrange the data in such a way that it becomes less chaotic and, as a result, more informative;
2. understand what process could have generated the data;
3. if the above is successful, find the missing "pieces" to complement the body of knowledge;
4. perform a formal statistical test that the results are indeed making scientific sense and not a result of a coincidence.

Statistical Analysis of Co-occurrence Matrices

Let's define what we want to achieve via visualisation of co-occurrence matrices. Firstly, we should understand that changing positions of rows and columns of co-occurrence matrices does not change anything that the matrix can tell us about the data observed (same is true for cross-tabulation). However (as we could see in the one-dimensional example above), a re-ordering of the clusters can make understanding of the data much more convenient. So, we want:

1. Re-order the rows and columns in the co-occurrence matrix in such a way that the visual image becomes **interpretable** (we shall give simple examples of interpretable matrices in what follows).
2. Design statistical tests to verify if such interpretable form could be a result of a chance or this is the result of the regularities existing in the data.
3. Design tests that can confirm that the data is purely chaotic and cannot, in principle, be reorganised into interpretable form.

Example 2: Co-occurrence in Evgeny Onegin

To test our algorithms for co-occurrence matrix construction we used the poem of Alexander Pushkin (probably the most distinguished Russian poet), translation of Ch.Johnston (http://lib.ru/LITRA/PUSHKIN/ENGLISH/onegin_j.txt). We present here some outcomes of this analysis as they look quite amazing.

The poem is 385 verses long (each verse has 14 lines). The length of the poem in words is 32945. Total number of different words used is 6806. Remarkable fact is that 3964 words were used only once. The heat map of the co-occurrence matrix looks quite special (figure 5). One can see from the image that the main color is light blue indicating low frequencies (the matrix is sparse). This is due to the fact that over 58% of the words used were used only once. Secondly we can see that the image resembles a source of light in the top left corner with radiation beams coming from it.

Such type of image cannot be observed in scientific research publications. In the next section we discuss three types of images which most often can be obtained in co-occurrence analysis of publications in scientific journals.

Simple Examples of Interpretable Co-occurrence Matrices

Single-focus co-occurrence
The images on figure 6 show some simulated co-occurrence data by means of heat map (red for higher co-occurrences and blue for lower co-occurrences).

The images on figure 6 were obtained from the images on figure 7 by rearranging rows and columns. While interpretation of the images on the figure 7 is not

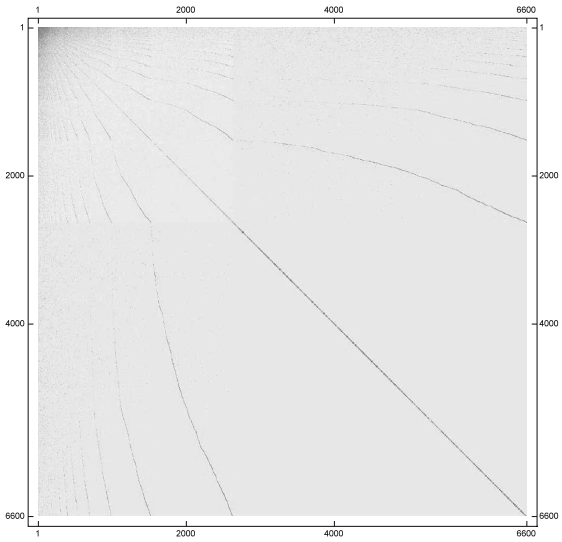


Figure 5: Co-occurrence matrix of Evgeny Onegin

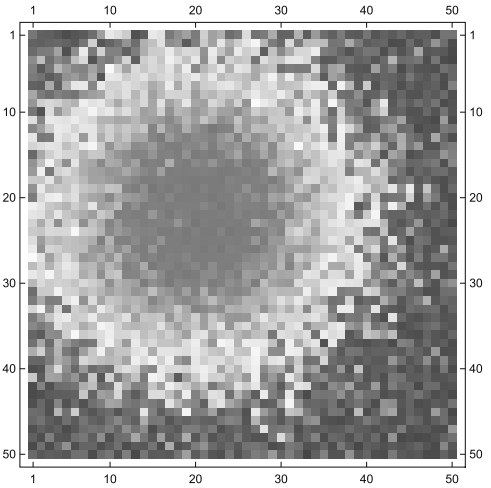


Figure 6: Focused set, no associations (heat map)

easy, the interpretation of the re-ordered matrices may be as follows: the papers included in the study are focused around certain keywords (in other words, some combinations of the keywords are much more likely than others). On the other hand, there is no statistical association between the key words (in other words, presence of one keyword does not increase or decrease the chances of presence or absence of another keyword).

For comparison we provide in (figure 8) co-occurrence matrices which cannot be re-ordered into interpretable format (while they consist of exactly the same cells).

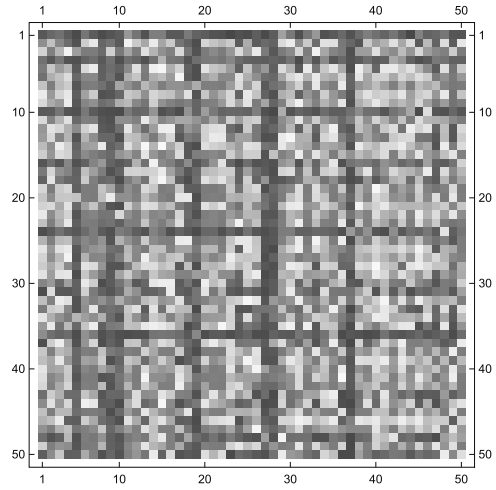


Figure 7: Focused set, no associations (before re-ordering)

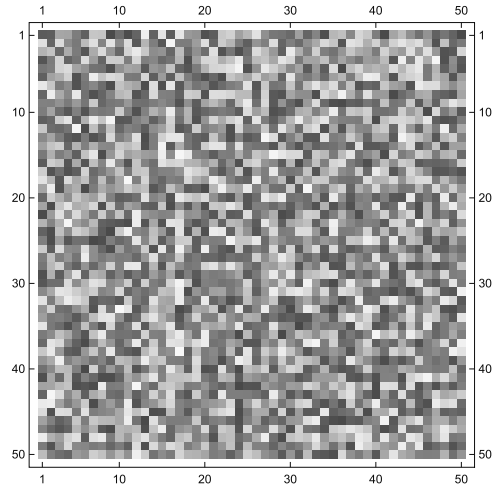


Figure 8: Chaotic co-occurrence: Single Focus

Double-focus co-occurrence
The images on figure 9 show co-occurrence data by means of heat map.

The images on figure 9 were obtained from the images on figure 10 by rearranging rows and columns. While interpretation of the images on the figure 10 is not easy, the interpretation of the re-ordered matrices may be as follows: the papers included in the study are focused around two sets of certain keywords (in other words, all publications may be subdivided into two sets with single focus).

Formally, statistical association between the keywords exists. However, after splitting the publications according to focal areas, the association will vanish.

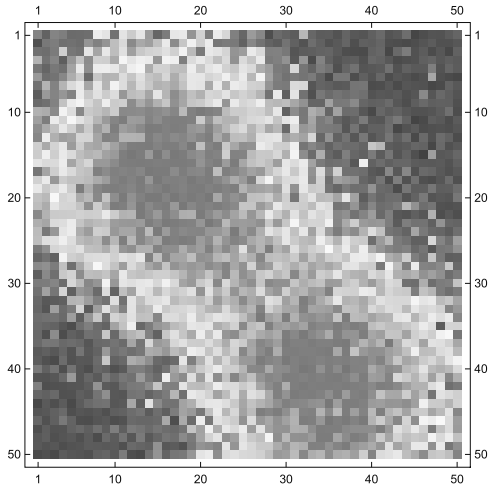


Figure 9: Set with two foci

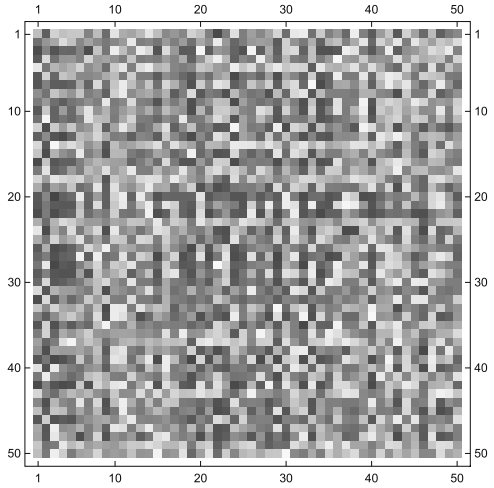


Figure 10: Set with two foci before re-ordering

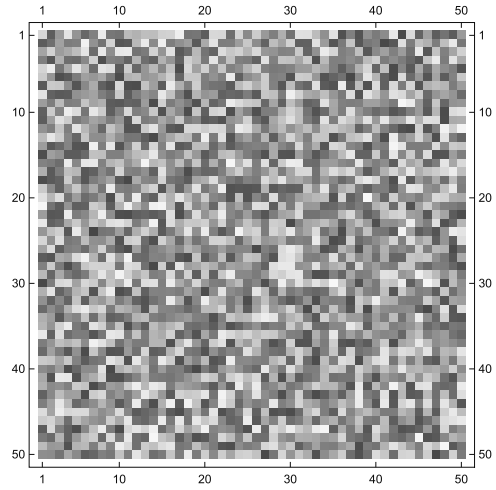


Figure 11: Chaotic Co-occurrence: Double Foci

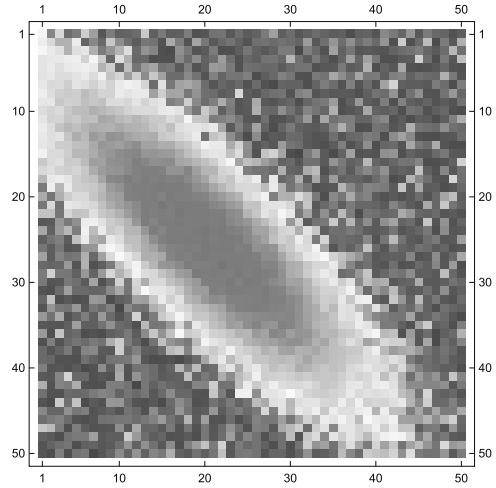


Figure 12: Focused set, strong association

For comparison we provide in (figure 11) co-occurrence matrices which cannot be re-ordered into interpretable format (while they consist of exactly the same cells).

Single focus co-occurrence, strong association
The images on figure 12 show co-occurrence data by means of heat map.

The images on figure 12 were obtained from the images on figure 13 by rearranging rows and columns. While interpretation of the images on the figure 13 is not easy, the interpretation of the re-ordered matrices may be as follows: the papers included in the study are focused around a set of certain keywords.

Statistical association between the key words exists in the following sense: if a keyword belongs to the focus area, then some other keywords are more likely to be present in the same publications. In other words, presence of one keyword does increase or decrease the chances of presence or absence of other keywords).

For comparison we provide in (figure 14) co-occurrence matrices which cannot be re-ordered into interpretable format (while they consist of exactly the same cells).

Discussion

Even with a naked eye one may see the difference between figures 7, 10, 13 (original datasets) and 8, 11, 14 (chaotic datasets). This gives us confidence that making same distinction will be possible with some formal algorithm (e.g. SVM, or other type of machine learning).

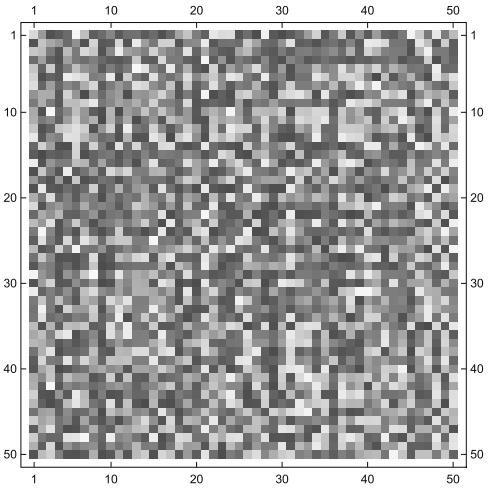


Figure 13: Focused set with strong association before re-ordering

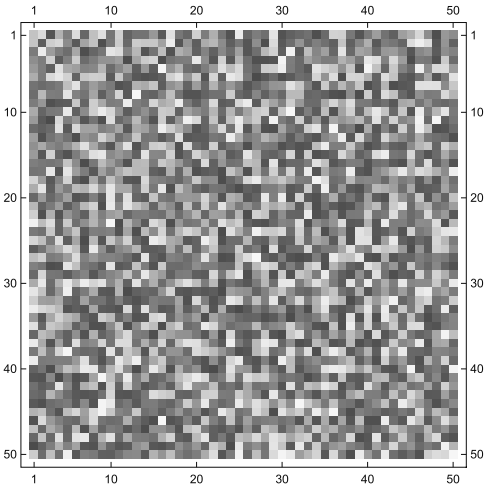


Figure 14: Chaotic Co-occurrence: Association

Application of tests for association does not require re-ordering into interpretable format. As we know, practically all tests for association are not sensitive to reordering rows and/or columns.

On the other hand, associations in the second and the third case are distinctively different. It may be suggested, therefore, that firstly, the dataset should be subdivided into single-focus datasets and then testing for association.

Of course many other interpretable images are possible. However, at this stage, we limit our study to these three only.

Proposed Technique

The statistical technique for co-occurrence analysis that we advocate here may be summarised as follows:

1. Firstly, a researcher has to come up with a hypothesis of what kind of image she/he is expecting to obtain as a result of reordering of rows and columns in the matrix. This is usual step in statistical analysis and modelling of any kind. For example, a researcher may suspect one of the three images defined in this paper.

2. Perform statistical tests to verify if the available matrix may be (in principle) reorganised into hypothetical image. We should note here that many statistical tests performed on matrices are totally insensitive to the order of rows and columns of a matrix. For example the classical Pearson's chi-square test for independence may be used to differentiate between the two single-focus images.

3. When the target image is defined and confirmed one may use a battery of available clustering algorithms. Different algorithms may yield different images, so the researcher should select the result which complies with the hypothesis most. If all of the clustering algorithms fail to produce expected image, a custom-made algorithms should be designed specifically to suit given hypothesis.

4. Interpret the image as per the subject area circumstances.

Real Data Example

The data for this example was collected in the following way. The co-occurrence data was derived from literature collected for a computer-assisted survey of valorization pathways for waste biomass (our thanks to Dr CB Davis, University of Groningen, for making these data available). Academic literature was collected from both Web of Science and Scopus. The abstracts for each article were analysed, and terms corresponding to organisms and chemicals were identified. This information was then recorded in a document-term matrix which recorded for which document (i.e. academic abstract), which terms (related to chemicals or organisms) were found. The cross-product of the document-term matrix was then calculated to generate a co-occurrence matrix indicating how many times a particular term was found in an abstract with another term. In this matrix the rows correspond to chemicals and the columns correspond non-chemicals.

Firstly, we removed columns and rows with totals of zero or one. This was done to reduce the dataset by removing obviously insignificant entries. We were left

Table 5: Outcomes of the classifying process

Method	Accuracy
LogisticRegression	0.995
NaiveBayes	1.
NearestNeighbors	0.82
NeuralNetwork	0.485
RandomForest	1.

with 1625 chemicals and 557 other keywords.

Secondly, we came up with the hypothesis that we deal with the case of "single Focus, no association". This hypothesis was tested in two ways.

1. We performed tests on practical significance for associations. The results were negative. So, if we have to chose between the three models, then we have to reject the other two models as both of them should have been yielding positive outcome for the association testing.
2. Then we had to test if the co-occurrence matrix is not a set of chaotic entries which cannot lead to an interpretable image. For this we generated (using random simulations) two sets of matrices: (a) sixty matrices of the shape as per image 7 and sixty matrices as per image 8. Fifty matrices from each set were used for training the classification algorithm and the other ten (from each set) for testing. The results of the process are given in the table 5. As one may see, only the ANN method gave poor results, while the Naive Bayes and Random Forest were the very best. So we used the Naive Bayes and Random Forest methods to classify the co-occurrence matrix we studied. Both methods classified the matrix as non-chaotic.

Then we performed the re-ordering, using various clustering techniques and and also try-and-error process. The best result is presented on the figure 15 (note that we removed low frequency entries for better visualisation).

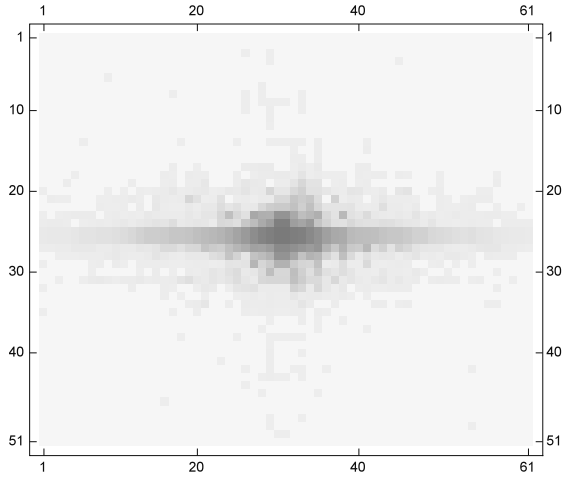


Figure 15: Interpretable image of first kind (low frequency entries removed)

Conclusions

The principal difference between the Hierarchical Clustering and the Broken Glass technique is that we first identify what kind of image we may get after re-ordering the rows and columns.

Our main result is that we can confidently distinguish between basic interpretable images and totally chaotic data.

The classification methods that are recommended for the above tasks are: Logistic Regression, Naive Bayes and Random Forest. The nearest Neighbour method is still good, but least preferred. ANN method may not be recommended.

REFERENCES

Altuntas S.; Dereli T.; and A. K., 2016. *Assessment of corporate innovation capability with a data-mining approach: industrial case studies. Computers & Industrial Engineering*, 102(2016), 732–764.

Buzydlowski J.; White; and Lin X., 2002. *Term Co-occurrence Analysis as an Interface for Digital Libraries. Lecture Notes in Computer Science*, 2539, 133–144.

Davis C.; Aid G.; and Zhu B., 2017. *Secondary Resources in the Bio-Based Economy: A Computer Assisted Survey of Value Pathways in Academic Literature. Waste and Biomass Valorisation*, Submitted.

Karakatsanisa I.; AlKhaderb W.; MacCroryc F.; Atif Omarb M.; Aunga Z.; and Lee Woona W., 2017. *Data mining approach to monitoring the requirements of the job market: A case study. Information Systems*, 65, 1–6.

Leydesdorff L. and Vaughan L., 2006. *Co-occurrence matrices and their applications in information science: Extending ACA to the Web environment. Journal of the American Society for Information Science and Technology*, 57(12), 1616–1628.

PERFORMANCE COMPARISON OF ADAPTED DELAUNAY TRIANGULATION METHOD OVER NURBS FOR SURFACE OPTIMIZATION PROBLEMS

Suyesh Bhattarai*
Parag Vichare
Keshav Dahal
Artificial Intelligence, Visual Communication and Networks Research Centre
University of the West of Scotland
Paisley, UK
Suyesh.bhattarai@uws.ac.uk

KEYWORDS
Delaunay Triangulation, NURBS, surface, optimization.

ABSTRACT

Traditionally NURBS (Non-Uniform Rational Basis Spline) are used as the basis for defining free-form surfaces as they can define non-regular surfaces with minimal control points. However, they require parameters such as knot vectors and weights to configure a surface. Similarly, DT (Delaunay Triangulation) is proven and used widely for meshing, rendering and surface reconstruction applications, but its capability in freeform surface design for optimization is untested. Thus, this paper proposes Adapted Delaunay Triangulation (ADT) method which can generate a surface from scattered data points without any parameters. The paper presents a comparison of the performance of ADT method and NURBS fitting method for surface generation from scattered 3D coordinate points. This method was suggested so that the generated surface could be used in Stochastic Optimization Algorithm (SOA) methods and computational fluid dynamics applications (CFD) simultaneously. Data points that other 3D point clouds fitting methods would ignore as outliers are included in ADT method. Small change in each data point during optimization cycle should show a distinctive change in its output as SOA approaches depend on such differences for its optimal performance. Special consideration has been made for fast processing and rendering of the surface with minimum complexity (removing parameters such as knots and weights) and storage requirements as SOA methods demand generation of numerous surfaces to solve any problem.

INTRODUCTION

Design optimization applications rely heavily on rendering surfaces and use various techniques for generation of these surfaces. Regular 2D planer facets can be created with straight or curved lines and the whole geometry for computer aided engineering (CAE) applications is created by merging these facets. These facets are usually built from well-defined base points. Generating surfaces from scattered points adds more complexity as undefined nature of the geometry may result in undesired, self-intersecting facets. Existing methods fit these points into Bezier or B-Spline surface, generating free-form surfaces (Narvaez, Narvaez, & Branch, 2010; Pizo & Motta, 2009). Usually, this method leaves out a number of points for configuring C1 or higher continuity surface. By contrast, other reported method by Boissonnat (J. D. Boissonnat, 1984) incorporates scattered point data in order to generate 3D elements for meshing solid geometry such as convex hull. Boissonnat and Cazals (Boissonnat & Cazals, 2002) and Amenta et. al. (Amenta, Bern, & Kamvysselis, 1998) reconstructed existing 3D surfaces from given sets of points, but with the assumption that i) the reference surfaces are smooth; ii) resulting surface will not have any open boundaries (such as solid models) and iii) normals to the surfaces are known.

Geometric design optimization for CAE application requires large scattered point cloud where every point has unique significance thus, cannot be filtered out for configuring geometry. Such examples include developing a freeform surface that could result in any shape as an optimization output. Most structural optimization methods study the strain and stress profile on the existing geometry and evaluate the most optimal design from the strain/stress graph (Madsen, Shyy, & Haftka, 2000; Papadrakakis, Lagaros, & Tsompanakis, 1998) but, it neglects the possibility of having an entirely new design unrelated to the existing one as discovered in the study by Linden (Linden, 2002). Especially for fluid dynamics studies, where a change in the interacting surface changes the overall nature of the flow, each change in the scattered point cloud is of importance.

This paper studies the previous works conducted in this area in section 2, explains the proposed Delaunay based method in section 3, compares the method against the widely used NURBS method in section 4, discusses the advantages and applications of the proposed method in section 5 and provides the conclusion in section 6.

SURFACE CONSTRUCTION APPROACHES

Most surface generation work has been concentrated in surface reconstruction from a given set of scattered data points. The data points are obtained from vision based laser scanning sensor and are used to reconstruct these surfaces for rendering,

graphics and pattern recognition. Research on configuring surfaces from point cloud has been classified by Boissonnat and Cazals (Boissonnat & Cazals, 2002) as:

1. Local projections (J. D. Boissonnat, 1984; Levin, 2004) develop surface as a function defined in a local reference domain. The surface is considered a graph of the function and approximated by triangulating in a moving projection plane or using least square function approximation techniques. These methods are fast but provide stretched and discontinuous surfaces with non-uniform and very sparse datasets.
2. Sculpting methods (J.-D. Boissonnat, 1984; J. D. Boissonnat, 1984) are based on removal of non-boundary facets from spatial arrangement, such as the convex hull. This method has performed well when the sampling is dense but reconstructed surface may not pass through all the sample points and may have additional holes.
3. Implicit methods (Boissonnat & Cazals, 2002; Hong-Kai, Osher, & Fedkiw, 2001; Ohtake, Belyaev, & Seidel, 2003) estimate a tangent plane from the sample data and uses distance to the plane as distance function. The zero-set of this function is then sampled at grid points and the surface is generated from these points. These methods require uniform and dense sampling for practical uses.
4. Deformable models (Amenta et al., 1998; Gary Wang, Dong, & Aitchison, 2001; Hoppe, DeRose, Duchamp, McDonald, & Stuetzle, 1992; Leal, Leal, & Branch, 2010) form an initial shell to which deformations are applied to minimize a function of energy and get closer to surface. Its performance depends largely on the initial guess which should be sufficiently close to the actual surface. These methods converge to local minima and could be significantly different from the true surface.

Sculpting and Deformable models based methods have an underlying assumption that all the surfaces are smooth and do not contain noise (Boissonnat & Cazals, 2002). Their performance has been commendable for surfaces without sharp edges and ample point density. But, these methods may fail to be robust and may require prohibitively large amounts of time to generate output for scattered point (J. D. Boissonnat, 1984; Hoppe et al., 1992). Thus, we will compare the performance of the proposed method based on local projections and NURBS based on implicit methods for our research problem of generating a surface from scattered points, inclusive of all the points.

Unlike polygons, NURBS are resolution independent and provide smooth curves and excellent continuity with fewer control points. But there are other parameters that greatly affect the topology of NURBS such as weights, knots and the degree of the curve (Narvaez et al., 2010). All these values must be perfectly coordinated to achieve the desired topology. NURBS requires a grid of control points that form the individual curves that can be moulded together to form a surface. This topology cannot be extended but can be patched with another such surface. In order to generate a NURBS surface from a set of scattered points, we first align the points cloud into a rectangular mesh. This mesh acts like the grid for the provided data set. The NURBS surface is generated using these points as the control points. The weights of each grid point are fixed as one and the degree of the spline curve is fixed as three to reduce variable parameters. First, three knot-vectors are defined as zero and the last three as one with uniformly spaced values in the remaining knots at the centre to ensure that the curves pass through the start and end points.

ADAPTED DELAUNAY TRIANGULATION (ADT) METHOD

As summarized in the previous section, while NURBS surfaces have got distinguished advantages, they demand considerable computing resources for preparing geometry for CAE and CFD applications. This provides a scope for developing a light weight geometry preparation method for engineering analysis applications. The method proposed in this paper is for the specific purpose of real time applications on mechanical design optimization problems. The method utilizes Delaunay triangulation algorithm to generate a surface as a patch of triangular surfaces with straight and sharp edges.

Algorithm

1. Define limits for the 3D points cloud

$$x_{min} \le x \le x_{max};\ y_{min} \le y \le y_{max};\ z_{min} \le z \le z_{max} \quad (1)$$

2. Define the number of points desired

$$[n] = \{1, \ldots\ldots\ldots, n\} \quad (2)$$

3. Generate the 3D points cloud with n points.

$$f(x) = \text{random}(\{x: x_{min} \le x \le x_{max}\}) \quad (3)$$
$$S = \{f(x_i, y_i, z_i), \forall\ i \in n\} \quad (4)$$

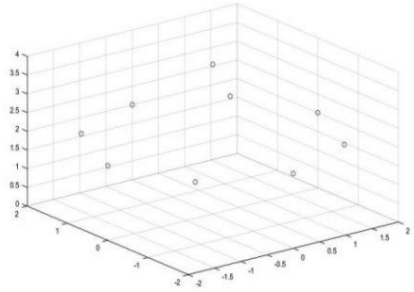


Fig.1. Generated points (Step 3)

4. Evaluate the spread of the coordinates by calculating their standard deviation

$$v(x) = \text{stdev}(\{f(x_i), \forall\ i \in n\}) \quad (5)$$
$$V = \{v(x), v(y), v(z)\} \quad (6)$$

5. Choose the coordinate axis with the minimum (or maximum) value of standard deviation to obtain depth axis of the surface. The chosen axis is the axis perpendicular to the generated surface

floor.axis := axis with min{V} (7)

6. If the values of standard deviation are equal follow the priority order of Z axis first and Y axis second.

floor.axix := z-axis, if $v(x) = v(y) = v(z)$ (8)
:= y-axis, if min{V} = $v(x) = v(y)$

7. Create a set of 2D points with the remaining two coordinate axis values.

$P = \{f(x_i, y_i, z_i), \forall\ i \in n\} - \{f(u_i)\}$ (9)
where, u := z, if floor.axis is z axis
:= y, if floor.axis is y axis
:= x, if floor.axis is x axis

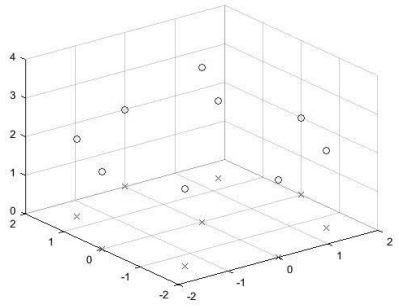


Fig.2. 2D projection (Step 7)

8. Apply 2D Delaunay algorithm to the generated set P.
9. Obtain the triangulation information (set of points that form a triangle) from 2D Delaunay output.

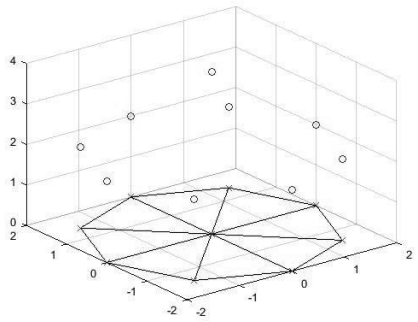


Fig.3. 2D Triangulation (Step 9)

10. Form a surface with the same triangulations in 3D space with respective x, y and z coordinates.

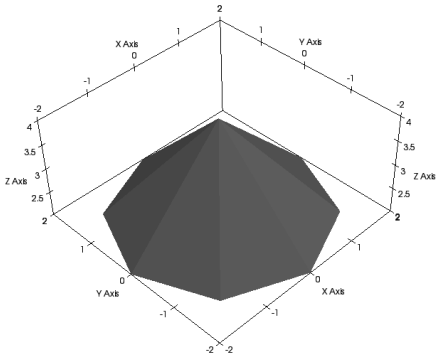
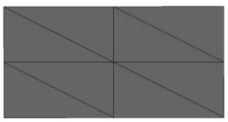
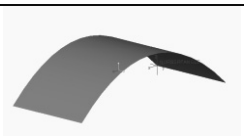
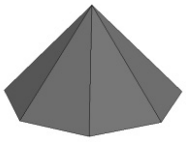
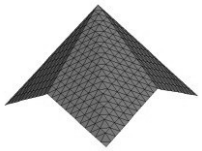
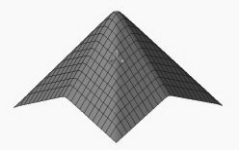
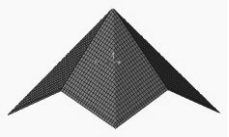


Fig.4. Surface Generation (Step 10)

This algorithm is basis of the method suggested in this paper.

Table 1: Comparison of ADT and NURBS

ADT method		NURBS method	
	Number of points: 9 File size: 298 bytes Time taken: 0.0035 sec		Number of points: 9 Number of control points: 3 Number of knots: 6 File size : 1,376 bytes Time taken: 0.0118 sec
	Number of points: 9 File size: 298 bytes Time taken: 0.0026 sec		Number of points: 9 Number of control points: 3 Number of knots: 6 File size : 1,376 bytes Time taken: 0.0123 sec
	Number of points: 9 File size: 304 bytes Time taken: 0.0045 sec		Number of points: 9 Number of control points: 3 Number of knots: 6 File size : 1,376 bytes Time taken: 0.0156secs
	Number of points: 625 File size : 25,963 bytes Time taken: 0.0073 sec		Number of points: 625 Number of control points: 25 Number of knots: 28 File size : 33,128 bytes Time taken: 0.0312secs
	Number of points : 10,000 File size : 490,737 bytes Time taken : 0.0847 sec		Number of points: 10,000 Number of control points: 100 Number of knots: 103 File size : 509,084 bytes Time taken: 0.3814secs

EXPERIMENTATION

The outputs generated from the different input values with the ADT based method and the NURBS based method is compared in this section. The comparison in Table 1 shows that there is a distinct advantage of using the ADT method over the NURBS method for optimization applications with minimal processing of the random data fed as input.

DISCUSSION

The advantages of the two methods employed in this paper can be briefed as following from the information collected from the above table. A typical case study with 10,000 points is considered for the comparison below.

1. Speed of generation: The most important factor while generating surfaces during optimization is the speed in which the geometry is created. The experiment shows that ADT method is 4.45 times faster than NURBS based method. This provides a massive advantage over the NURBS based method while generating multiple geometries.

2. Storage memory: The other important factor in optimization problems is the memory requirement and with the ADT method we get a 3.61% reduction in the total memory requirement for 10,000 points. And for geometry with 625 points, we get a 21.62% reduction in the memory requirement for ADT method. Such reduction in memory requirements enable running the simulation for even more geometries and allow more exhaustive search in SOA.

3. Geometric Continuity: The image generated from ADT method is made from joining together of flat triangular surfaces and hence provides a C0 continuity with respect to the adjacent surface. Whereas the NURBS method fits in the surface so that the continuity is maintained at C1 or above as specified by the codes. The ADT surface will look patched and pixelated while the smoothness of the NURBS surface adds to the aesthetic appeal for such surfaces.

4. Ability: The C0 continuity of ADT method allows for the geometry to incorporate sharp corners and a sudden change in the gradient of the surface topology, but NURBS being a fitting method does not allow for sharp corners and sudden change in the gradient of the surface. With designs requiring sharp edges and corners, two or more NURBS surfaces will have to be patched together. For designs requiring a smooth transition, ADT will require dense point cloud in such area of the geometry.

5. Pre-processing of Input data: ADT method takes the entire dataset as a whole and processes it all together to form the surface so pre-processing of the input data is not required. For the NURBS method, the data must be pre-processed and arranged in a grid to fit the basis spline curves. This pre-processing of data increases the complexity of this method.

6. Input data inclusion: The ADT method includes all the points on the surface and hence has no outliers. Every point lies on the surface of the geometry generated from ADT method. This ensures that a single change in the input data shows some drastic change in the output surface. Whereas in NURBS method, the surface is fitted based on predefined degree equation and hence some points do not lie on the surface of the geometry. This reduces the impact of changing a single point on the entire geometry. In SOA applications, it is desirable to have definitive changes in the geometry from a change in a coordinate point.

7. Variables: The ADT method only requires the coordinate values of the point cloud to generate a surface. But the NURBS method requires additional parameters such as weights, knots and degrees to generate the surface. These additional parameters may require some changes depending on the nature of the points cloud. In SOA applications, these parameters increase the complexity of the problem and may fail to provide a suitable surface as an output.

8. Robustness: The ability of Delaunay methods have been proven from studies carried out in the past. It is able to handle a large number of scattered points. These points need not be arranged in a grid, but the distribution must be fairly uniform to avoid holes and unwanted features. The NURBS method needs the points to be arranged in a proper grid and hence is less robust as it might be difficult to form grids from some groups of scattered points. The other parameters, such as weights and knot vectors, need their values to be well defined to achieve the desired NURBS surface. When dealing with numerous scattered point cloud sets, the same parameter values might not yield the best results.

9. Compatibility: The ADT method generates the surfaces in VTK format and this format can be used with any open source rendering and simulation packages. The NURBS format was developed for industrial use and is mostly associated with commercial software packages. It makes the ADT method easier to access for the general public.

10. Applications: This overall comparison shows that ADT method is ideal for use in SOA applications to determine the initial design of any surface whose performance can

be determined from CFD simulations. The NURBS method output is smooth and aesthetically pleasing and can hence be used in imaging and rendering applications. It can also be used to generate a geometry based on the final output from ADT method and run simulations on it.

CONCLUSION

Current developments in graphics and surface rendering are demanding smoother surface finish and aesthetics for graphical interfaces. Such applications require considerable computational power at hand to process limited graphical information on the screen. Other applications require generating numerous geometries with constraints of time, computational power and storage capacity. The proposed ADT method is robust and provides about 4 times faster and simpler construction with 3-20% less memory requirement to generate surfaces that are compatible with multiple simulation packages and can be used together with SOA. The proposed method is dependent only on the coordinate points and hence provides consistent outputs for the same data while allowing sudden changes in the gradient and sharp corners that other freeform methods cannot. These qualities make this method very desirable for applications where the performance of the surface is dependent on its geometry, especially where a small change in one portion of the geometry may call for major changes in the remaining portion such as fluid flow over the surface. This method was suggested to be used together with computational fluid dynamics simulation software and stochastic optimization algorithms to produce an optimal surface for geometric design problems.

ACKNOWLEDGEMENT

The first author likes to acknowledge the support provided by the EU Erasmus Mundus project SmartLink (552077-EM-1-2014-1-UK-ERA) to carry out this research at the University of the West of Scotland, UK.

REFERENCES

Amenta, N., Bern, M., & Kamvysselis, M. (1998). A new Voronoi-based surface reconstruction algorithm. Paper presented at the Proceedings of the 25th annual conference on Computer graphics and interactive techniques.

Boissonnat, J.-D. (1984). Geometric structures for three-dimensional shape representation. ACM Trans. Graph., 3(4), 266-286. doi:10.1145/357346.357349

Boissonnat, J.-D., & Cazals, F. (2002). Smooth surface reconstruction via natural neighbour interpolation of distance functions. Computational Geometry, 22(1), 185-203. doi:http://dx.doi.org/10.1016/S0925-7721(01)00048-7

Boissonnat, J. D. (1984). Geometric Structures for Three-Dimensional Shape Representation. ACM Transactions on Graphics, 3(4), 21.

Gary Wang, G., Dong, Z., & Aitchison, P. (2001). Adaptive Response Surface Method - A Global Optimization Scheme For Approximation-Based Design Problems. Engineering Optimization, 33(6), 707-733. doi:10.1080/03052150108940940

Hong-Kai, Z., Osher, S., & Fedkiw, R. (2001). Fast surface reconstruction using the level set method. Paper presented at the Proceedings IEEE Workshop on Variational and Level Set Methods in Computer Vision.

Hoppe, H., DeRose, T., Duchamp, T., McDonald, J., & Stuetzle, W. (1992). Surface reconstruction from unorganized points. SIGGRAPH Comput. Graph., 26(2), 71-78. doi:10.1145/142920.134011

Leal, N., Leal, E., & Branch, J. W. (2010). Simple Method for Constructing NURBS Surfaces from Unorganized Points. In S. Shontz (Ed.), Proceedings of the 19th International Meshing Roundtable (pp. 161-175). Berlin, Heidelberg: Springer Berlin Heidelberg.

Levin, D. (2004). Mesh-Independent Surface Interpolation. In G. Brunnett, B. Hamann, H. Müller, & L. Linsen (Eds.), Geometric Modeling for Scientific Visualization (pp. 37-49). Berlin, Heidelberg: Springer Berlin Heidelberg.

Linden, D. S. (2002). Antenna Design Using Genetic Algorithm. Paper presented at the GECCO.

Madsen, J. I., Shyy, W., & Haftka, R. T. (2000). Response surface techniques for diffuser shape optimization. AIAA Journal, 38(9), 1512-1518. doi:10.2514/2.1160

Narvaez, N. E. L., Narvaez, E. A. L., & Branch, J. W. (2010). Automatic Construction of NURBS Surfaces from Unorganized Points. DYNA, 78(166), 8.

Ohtake, Y., Belyaev, A., & Seidel, H.-P. (2003). A multi-scale approach to 3D scattered data interpolation with compactly supported basis functions. Paper presented at the Shape Modeling International.

Papadrakakis, M., Lagaros, N. D., & Tsompanakis, Y. (1998). Structural optimization using evolution strategies and neural networks. Computer Methods in Applied Mechanics and Engineering, 156(1), 309-333. doi:http://dx.doi.org/10.1016/S0045-7825(97)00215-6

Pizo, G. A. I., & Motta, J. M. S. T. (2009). 3D Surface Generation from Point Clouds Acquired from a Vision-based Laser Scanning Sensor. Paper presented at the 20th International Congress of Mechanical Engineering, Gramado, Brazil.

Automatic Generation of a Diagnostic and Control Unit for Monitoring Embedded System Applications. Case Study: HEVC Decoder

Habib SMEI
Université de Tunis El Manar, Faculté des Sciences de Tunis, Laboratoire LIP2, 2092, Tunis, Tunisie.
Institut Supérieur des Etudes Technologiques de Rades, Rades, Tunisie.
E-mail : habib.smei@isetr.rnu.tn

Abderrazak JEMAI
Université de Tunis El Manar, Faculté des Sciences de Tunis, Laboratoire LIP2, 2092, Tunis, Tunisie.
Université de Carthage, INSAT, Tunis, Tunisie.
E-mail : abderrazak.jemai@insat.rnu.tn

Kamel SMIRI
Université de Tunis El Manar, Faculté des Sciences de Tunis, Laboratoire LIP2, 2092, Tunis, Tunisie.
Université de Manouba, Institut Supérieur des Arts Multimédias Manouba, Campus Universitaire de Manouba, 2010, Manouba, Tunisie.
E-mail : smiri_kamel@yahoo.fr

KEYWORDS

MPSoC, co-design flow, Performance Estimation, Run time controller, SADF, SDF3, Zedboard platform.

ABSTRACT

In this paper, we present a new approach for a generation of a run time controller of an embedded system application.
This generation is based on design time modeling as an input and generates automatically two versions of the controller. A software controller that can be run as a software supervisor in an Asymmetric MultiProcessing (AMP) mode and a Hardware one (IP hardware) that can be implemented in an FPGA.
In experimentations, we use the High Efficiency Video Coding (HEVC) decoder as an embedded application case study and the Zedboard platform as a experimental platform. Experimentations show that the our generation approach is able to generate controllers (software version and hardware version) for any application modeling with SADF-FSM with multiple scenarios of executions.

1. INTRODUCTION

A Co-Design flow of embedded systems [1] consists of a set of steps beginning with a requirements specification and ending with the integration of the software and hardware into silicon chip.

The specification step is usually represented as a software application coded with a high level language such as C / C++, Matlab, Sytemc. This specification is run on a host machine in order to test its functionalities and further understand the specificities of the whole system.

Once this specification is tested, a profiling step [2] [3] is usually performed to define a profile of each entity of the system and establish a call graph function as well as other information such as execution time, size and type of data Exchanged.

All this information is a parameter whose designers use to define architectural choices either in the modeling phase and partitioning phase of the Co-Design flow.

Once the modeling is done, a verification of the established model is required. Several methods, languages and tools are available to designers to do this verification. The choice of tools and approaches depends essentially on the nature of the application to be modeled (for example, data flow oriented or control flow oriented), but also depends on the experience of the design team and the availability of tools grasped.

Usually the verification step is performed by simulation, hardware emulation or by formal methods (e.g. Model Checking, Theorem Proving, ...). But when applications become increasingly complex (e.g. Multimedia applications) and architectures are increasingly powerful (e.g. MPSOCs), the use of these traditional verification methods on complex embedded systems appears increasingly inefficient. In fact, the use of traditional methods of verification with these new systems consumes either an intolerable processing time (e.g. logic simulation), or an exaggerated memory size (e.g. Model Checking). For this purpose, designers have recourse to new modeling and verification methods, namely Model of Computation (MOC) [4][5].

A MOC is a description (modeling) based on mathematical rules to calculate and analyze the semantics of the system to be modeled and to treat its competitive behavior.

Among these MOCs, the SDF model (Synchronous Data Flow) is widely used in multimedia applications [6][7]. This is one of the most appropriate performance estimation methods used at the system level of a Co-Design flow.

In addition to the modeling verification, SDF offers a set of formalism rich in concepts for better application modeling. SDF offers other services such as tasks parallelization and their mapping on the hardware architectures (assignment of tasks to processors elements). SDF is often used to model multimedia applications with real-time constraint.

Once the modeling is validated, the following steps in the design flow have taken place. These include steps such as partitioning, software and hardware synthesis,

integration of two components (software and hardware) and testing of the entire system.

At run time, the system requires real-time monitoring and control in order to check its running states and provide solutions for remedies and correction in the event of a breakdown, for example.

This control system, if it exists, is designed in the majority of cases once the system is realized and put into operation. It is therefore designed and realized with methods and tools other than those used at design time (at the design and modeling step).

The basic idea of the work that we propose in this article is to take into account the control of the system (at run time) since the design stages. Consequently, the model used in design stages will not be left apart; it will accompany the system itself at the run time to control it.

In our case, we use the SADF-FSM model, a variant of the SDF for the performance estimation phase and model phase of the HEVC decoder application, our case study in this work.

We have developed a generic tool integrated in our Co-design flow (LIP2 Co-design flow) [8] in order to generate automatically a control component for the system being designed. This controller called **DCU** (**D**iagnostic and **C**ontrol **U**nit) will be used at the run to meet several needs.

It can be used as a simple controller for the execution time of various tasks of the system running and therefore it will monitor the execution time, as it can be used to monitor the energy consumption of the system and to intervene in case of problem. It can also be used for safety purposes or for other purposes.

This automatic DCU generator tool takes as input the model realized with SADF-FSM. As output, it produces the controller in two vesrions. A software version as a C program that can be compiled and executed in parallel to the application to be controlled (in an AMP mode for example) and an IP Hardware component generated in VHDL that can be directly placed in an FPGA programmable area.

In our case, the hardware component was implemented in the programmable part (PL part) of the Zedboard platform, our experimental hardware platform.

The automatic DCU generator tool has been tested on the HEVC decoder, the HM test Model Application [9] and also on three other models provided in the official web site of SDF3 [10].

In all tests, a DCU (with software and hardware version) was generated and tested on Zedbaord platform.

The results show that in all cases, the DCU can detect the delay caused by actors and notify the operating system in order to make the appropriate action.

The rest of this article is organized as follows: Section 2 presents a discussion of related work that deals with the same problematic of the article. În Section 3, we give an overview of the HEVC Standard and HEVC h.265 decoders. In Section 4, we detail our approach to generate the DCU controller. Section 5 shows experimental results. Finally, we conclude with a conclusion in which we quote some perspectives.

2. RELATED WORK

Modeling systems with Synchronous Data Flow graphs is studied in several works that deal with performance estimation of MPSoC applications. S. Stuijk [11] is the first person who proposed an SDF graph to map a multimedia application on NoC-MPSoC platforms in order to minimize the resource usage. The flow starts with an application-aware SDF that is incrementally transformed to handle resource sharing on a multi-tile architecture. In [12], they use timed SDF graph to model NoC architectures with predefined guaranteed of bandwidth and a maximum latency. Various predictable arbitration mechanisms of the NoC have been considered, such as Round-robin, TDMA, and Static sharing. In [13], SDF graphs are used to estimate the system's performance after the migration of a task from software to hardware. In [14], authors proposed an approach to analyze applications that are modeled by SDF graphs and executed on MPSoC platform by using a Resource Manager (RM) actor. RM is a task responsible for resources access (critical or not). The designer reserves for the RM a whole execution node (CPU, memory, bus,...) which increases the cost of the total MPSoC system.

Wiggers and al. [15] proposed a solution that consists in mapping purely software tasks and their communication channels on the target processors. They exploited the SDF graphs to compare the throughput obtained with the target throughput of the application. In [16], authors developed a generic communication assistant module for multi-processor and multi-application systems, where SDFs are used to estimate the worst-case performance of a system before implementation.

In [17], [18], authors use FSM-SADF in order to model applications with multiple behaviors. Each scenario (behavior) is modeled by an SDF, and the FSM represents the possible orders in which active scenarios occur. Real-time calculus also focuses on stream based on applications and has mode-based approaches [19], but handling cyclic dependencies is limited to Marked Graphs without modes [20].

[21] deals with a stochastic version of the SADF model. Also [22] deals with scenarios of SDF behavior, but in their case only homogeneous SDF graphs are considered (graphs in which all consumption and production rates are equal to one), and only the execution times of affixed collection of actors can vary with scenarios. In earlier work [23], authors introduced techniques to find linear upper bounds on transient behavior of an SDF, which also allows the behavior of an SADF to be analyzed, but without exact results. In [23], the FSM specification of possible scenario orders could not be taken into account and one would have to assume any scenario order to be possible.

3. CASE STUDY: HEVC DECODER

The HEVC [24] [28] is the acronym of "High Efficiency Video Coding". This is the latest video coding format used by JCT-VC. This standard is an improvement of the H.264 / AVC standard [24] [25]. It was created at the end of January 2013 [27], at which a first version was finalized. It was developed jointly by the ISO / IEC Moving Picture Experts Group (MPEG) & ITU-T Video Coding Experts Group (VCEG).

The main objective of this standard is to significantly improve video compression compared to its predecessor MPEG-4 AVC / H.264 by reducing bitrate requirements by as much as 50% compared to H.264/ AVC with equivalent quality. The HEVC supports all common image definitions. It also provides support for higher frame rates, up to 100, 120 or 150 frames per second.

The HEVC standard defines the process of encoding and decoding the video.

As input, the encoder will process an uncompressed video. It performs the prediction, transformation, quantification and entropy coding processes to produce a bitstream conforming to the H.265 standard.

The decoding process is divided into four stages (figure 1). The first stage is the entropy decoding for which relevant data such as reference frame indices; intra-prediction mode and coding mode are extracted. These data will be used in the following stages. The second is called reconstruction step, which contains the inverse quantization (IQ), inverse transform (IT) and a prediction process, which can be either intra-prediction or motion compensation (inter-prediction). In the third stage, a de-blocking filter DF is applied to the reconstructed frame. Finally a new filter called Sample Adaptive Offset (SAO) is applied in the fourth stage. This filter adds offset values, which are obtained by indexing a lookup table to certain sample values [28].

In HEVC, the coding structure is based on a quaternary tree representation allowing partitioning into multiple block sizes that easily adapt to the content of the frames. Three types of coding structures are defined; ALL intra or intra-only (AI), Low Delay (LD) and Random Access (RA).

In the first structure, all pictures are encoded as intra, which yields very high quality, and no delay; this mode is mainly aimed for studio usage.

In the low delay one, the first is an intra frame while the others are encoded as generalized P or B pictures (GPB). This structure is conceived for interactive real-time communication like video conferencing or similar real-time uses were no delay for waiting for future frames are permitted.

Finally, the random access structure is similar to hierarchical structure and intra pictures are inserted periodically, at the rate of about one per second. It is designed to enable relatively frequent random access points in the coded video data. This coding order has an impact on latency, since it requires frame reordering: for this reason the decoder might have to wait to have decoded several frames before sending them to output. This is the most efficient mode for compression but also requires most computational power.

Each of these three modes have a low complexity variant were some of the tools are disabled or switched into a faster version. As an example low complexity uses CAVLC instead of CABAC

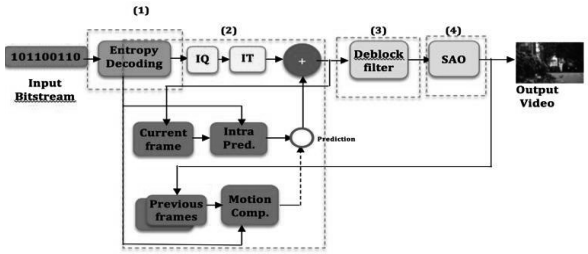


Figure 1: Functional structure of the HEVC decoder

4. DCU GENERATION FLOW

a. Introduction

In the Co-Design flow, SDF is used for the performance estimation and the modeling step. This phase takes as input parameters informations about all tasks of the system (actors) such as the Worst Case Execution time (WCET), the number and the size of data exchanged (size of the communication channels), the size of local objects to each actor (size of local variables), ...

The SDF model established will be validated by the SDF3 tool [10] in order to determine the maximum throughput assured taking into account functional constraints. In the case of a video decoding application, the constraint may be the number of decoded images per second.

SDF Model can be coupled with an architecture model for the mapping of software tasks onto a hardware architecture taking into account the functional constraints of the application and the non-functional constraints related to the architecture (such as the number of processors, their types, memory space available and the communication channels).

Once the SDF model is validated, the designer can proceed to the next steps of the Co-design flow.

Currently, in our LIP2 Co-design flow, the SDF model is used only for modeling and performance estimation at Design time.

In this work we propose an extension of use of the SDF model to accompany the designer in the advanced phases of the Co-design flow to produce a functional implementation in order to allow a control of the system at Run time (figure 2).

In its first version, the controller generated (DCU) is used only to monitor the execution times of the various actors. It should be noted that this DCU could be used for other purposes. It can be used to control energy consumption of the system to notify the operating system of exceedances of the authorized thresholds. It could also be used for security purposes and intervenes to signal a security problem such as an attack.

The DCU can also be used to monitor inter-component communication (IP, processors, system memory,

communication system) and report communication problems (such as a bottleneck).
The DCU could also intervene to put the system in safe mode to ensure availability in the event of a software or hardware failure.
It could also intervene before failure, in preventive mode. For example, it can foresee a saturation of the RAMFS exchange file system of the RAMDISK, it intervenes to prohibit the tasks that are the source of this saturation (to block the task that causes the saturation that to stop the whole system).

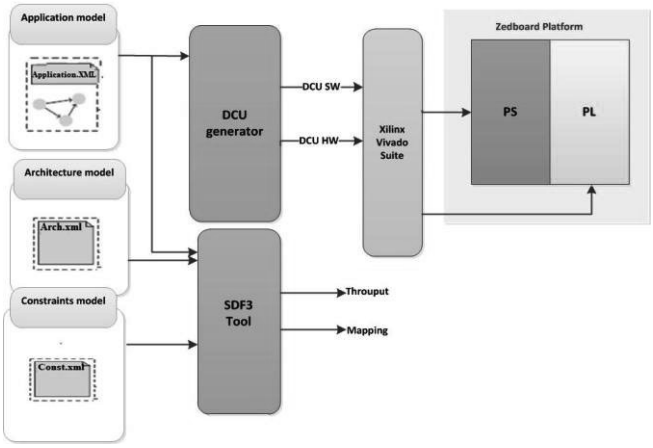


Figure 2: DCU Generation flow

b. DCU Generation Approach

The process of generating the controller consists of a set of steps (Figure 3).
The generator takes as input the SDF model of the application (SADF-FSM in our case) (figure 4).
To model an application with an SDF graph, we must have at our disposal information that concern the application, its functions (actors), the exchanged data (tokens) and their sizes, the size of the data used in each function and constraints to which the application is subject. This information can be extracted by performing a fine profiling step of the application. We have thus performed a profiling step that combines manual profiling with the use of available profiling tools such as Valgrind [29], Gprof [30], memprof [31]. In order to perform a good evaluation of the standard, the JCT-VC developed a document with reference sequences and configurations that should be used with each codec operation [32].
The SDF model includes:

- An XML description of all the actors of the application, in terms of execution time (WCET)
- The number of tokens exchanged between actors (input / output)
- The size of the data exchanged
- The size of the local variables of each actor
- The execution scenarios (which correspond to the possible configurations of the application)

The generator reads this file to automatically extract the information related to the execution times of the various actors (WCET of each actor) as well as the possible execution scenarios.
This information will be used to generate the controller in software version and in hardware version, based on templates (one for the software version and one for the hardware version).
The software version consists of a set of C-based controller files with a Makefile file that will be used to generate the executable. This version can be used in AMP (Asymitric MultiProcessors) mode. In this case the controller will be non-intrusive. It will run on a separate processor under an operating system other than that manages the application (figure 5).
The hardware version (figure 6) consists of a VHDL file with input and output signals from the various actors in addition to the start, stop and frequency signals.
As output, the controller generates information of the time spent by the active actor and an interrupt signal if the expected time for the active actor is exceeded.
The hardware version (VHDL) will be synthesized to generate the RTL file and to be implemented in the programmable space of the Zedboard (PL).
A configuration file is also generated that includes scenarios execution of the application and the WCET time for each actor. This file will be used by the DCU at run time.

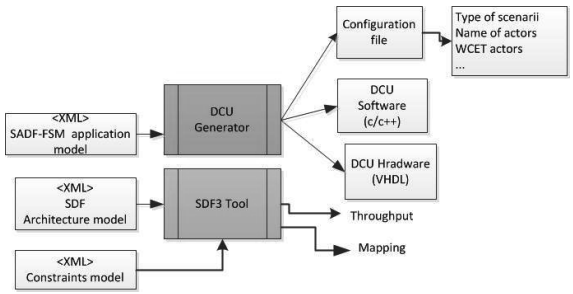


Figure 3: DCU Generation steps

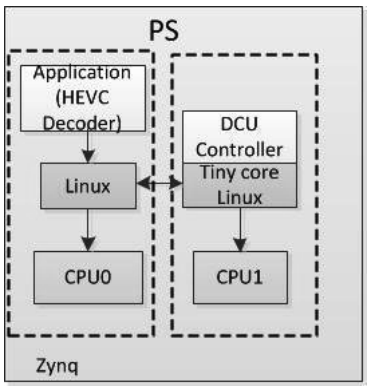


Figure 4: DCU Software on AMP Mode

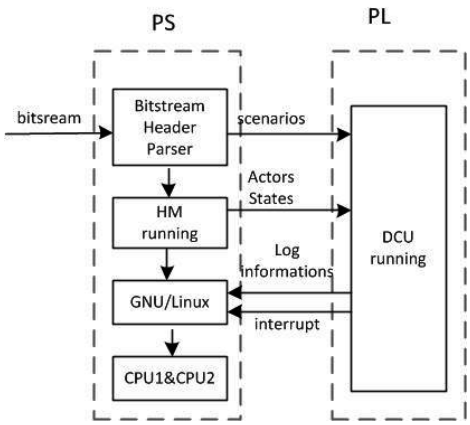


Figure 5: DCU Hardware

5. Experimental results

a. Zedboard Platform

Zedboard platform [33] (table I) is an evaluation platform based on a Zynq-7000 family [34]. It contains on the same chip two components.

A dual-core ARM Cortex MPCore based on a high-performance processing system (PS) that can be used under Linux operating system or in a standalone mode and an advanced programmable logic (PL) from the Xilinx 7th family that can be used to hold hardware accelerators in multiple areas.

The two parts (PS and PL) interact between them by using different interfaces and other signals through over 3,000 connections [34]. Available four 32/64-bit high-performance (HP) Advanced eXtensible Interfaces (AXI) and a 64-bit AXI Accelerator Coherency.

TABLE 1. ZEDBOARD TECHNICAL SPECIFICATIONS

Component	Characteristics
Processeur ZYNQ-7020 AP SOC XC7Z020-7CLG484CES	2 ARM Cortex A9 cores at 667 MHz
Memory	512 MB DDR3, 256 MB Quad-SPI Flash et SD Card
Communication	10/100/1000 Ethernet, USB OTG et USB UART
Extension	FMC (Low Pin Count) et 5 Pmod headers (2*6)
Display	HDMI output, VGA output et 128*32 OLED
Input / Output	8 switches, 7 push butons et 8 leds
Current and Voltage	3.0 A (Max) et 12V DC input
Certification	CE and RoHS certifier

b. The SDF3 tool

The open-source SDF3 tool set [10] used in this work, offers an SADF graph generation algorithm that constructs random SADF graphs which are connected, consistent, and deadlock-free. This generation algorithm can be used to benchmark novel SADF analysis, transformation, and implementation algorithms. If desired, the user can restrict relevant properties of the generated graph (e.g., limit port rates, or construct only acyclic or strongly connected graphs). All algorithms and technics implemented in SDF3 can be accessed through a set of command line tools as well as a C/C++ API.

The SDF3 tool has a conservative approach, which means that it will assume the worst-case scenario at any stage of the analysis. The constraint that the SDF graph must respect is the bitrate that must not be below 25 frames per second. Therefore, 25 iterations per second because the decoder decompresses an image by iteration. The value of the constraint will therefore be 0.025 iterations per time unit (ms).

c. Modeling

The decoding process of a bitstream depends on the coding configuration (AI, RA, LD, …) and on the class of the bitstream (Class A, B, …). FSM-based-SADF is the most suitable tool to describe this application because of its various configurations. To model an application using the SDF3, it is necessary to describe it by an XML file. This file contains a lot of information such as actor's names (table II presents actors names of the HEVC Decoder), memory sizes, execution times, scenarios, and transitions between scenarios. Thanks to profiling step, we have all data available to start creating our XML file. But it is essential to adapt this information to the syntax imposed by the SDF3 because there are many rules to follow:

- Sizes must be in bytes.

- A same time unit must be used along the description.

- The execution times introduced are just for a single iteration.

- The extension of SDF, FSM-based-SADF, which we exploited in our work, requires that the values introduced in the XML description of the application are values that describe the worst case.

Once all information is set in the SADF model (XML file) of the application, SDF3 can generate the throughput available and the application graph (figure 6).

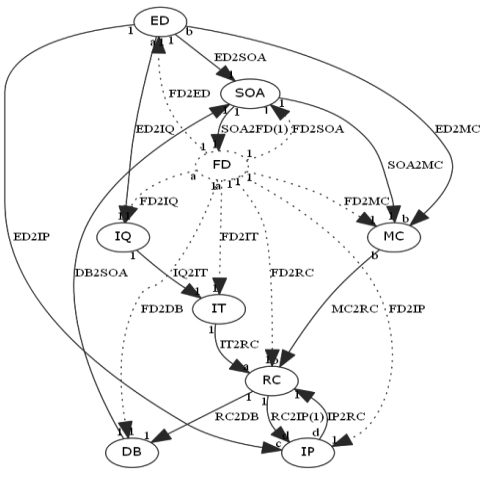


Figure 6: SADF Model of HEVC Decoder (generated from SDF3)

TABLE 2. LIST OF HEVC DECODER ACTORS

ACTOR	TASK
ED	Entropy Decoding
IQ	Inverse Quantification
IT	Inverse Transform
IP	Intra Prediction
MC	Motion Compensation
RC	Reconstruction
DB	Deblocking Filter
SAO	SAO Filter

For experimental results, we have tested the generator tool with four SDF models (Three models from the SDF3 web Site and our SADF-FSM Model of the HEVC Decoder).

In all tests, the generator tool generates two DCU versions for each SDF Model.

For the software version, we have implemented the DCU controller on a Host server and the application HEVC Decoder (HM Version) on a client host and we have used the Socket technics o test its functionality.

For the hardware version, we have used ModelSIM to simulate the behavior of the DCU hardware controller and all tests were successfully performed (figures 7 and 8).

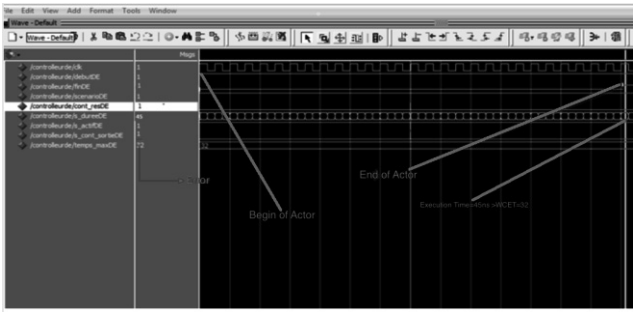


Figure 7: Abnormal termination of actor (total execution time >WCET)

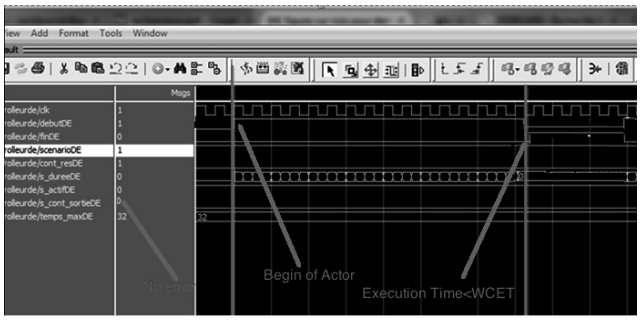


Figure 8: Normal termination of actor (total execution time <WCET)

Now we are working on the implementation of DCU hardware on the FPGA (PL) on the Zedboard platform.

6. CONCLUSION

In this paper, we presented an approach to the automatic generation of a controller called DCU that will track the execution of an embedded application at run time. This component was generated from the modeling step at Design time.

This allowed us to use the same model that served at the design stage at run time.

All tests have been validated successfully. The DCU can be used on various issues.

It can be used to control energy consumption of the system to notify the operating system of exceedances of the authorized thresholds. It could also be used for security purposes and intervenes to signal a security problem such as an attack or to monitor inter-component communication

The DCU could also intervene to be in safe mode to ensure availability in the event of a software or hardware failure.

We are planning in future work to use the controller for energy control and for security purposes. We will extend in this case annotate the SDF model to represent others proprieties.

Also, we plan a joint use of two versions of the DCU (Software version and hardware version) and switches when needed from one version to another.

REFERENCES

[1] Ehrlich P., Radke S. 2013. "Energy-aware software development for embedded systems in HW/SW co-design", Design and diagnostics of electronic circuits & systems (ddecs), IEEE 16th international symposium.

[2] H. Kai., Z. Xio-xu., X. 2015. Si-wen., and al. "Profiling and annotation combined method for multimedia application specific MPSoC performance estimation", springer-verlag berlin heidelberg,

[3] H. Smei, K. Smiri and A. Jemai. 2017. "Profiling of HEVC Decoder application in a codesign flow" - 6th International Colloquium in Applied Research and Technology Transfer.

[4] A. Jantsch, I. Sander. 2005. "Models of computation and languages for embedded system design". IEE Proc. Comput. Digit. Technol. 152(2), 114–129.

[5] E.A. Lee, S. Neuendorffer. 2005. "Concurrent models of computation for embedded software". IEE Proc. Comput. Digit. Technol. 152(2), 239–250.

[6] E. Lee and D. Messerschmitt. 1987. "Synchronous data flow". IEEE Proceedings, 75(9):1235{1245, Sept.

[7] S. Stuijk. 2007. "Predictable Mapping of Streaming Applications on Multiprocessors", Ph. D. thesis, Eindhoven University of Technology.

[8] H. Smei and A. Jemai. 2016. "Pipelining the HEVC Decoder on ZedBoard Plateform" - International Design & Test Symposium IDT 2016 .

[9] BSD Licence HEVC decoder (HM). 2017. Reference web site: https://hevc.hhi.fraunhofer.de, code available Online at https://github.com/bbc/vc2-reference. Accessed.

[10] S. Stuijk, M. Geilen, and T. Basten. 2006. "SDF3: SDF For Free," in Int. Conf. on Application of Concurrency to System Design, ACSD 06, Proc. IEEE, pp. 276–278.

[11] S. Stuijk. 2007. 'Predictable Mapping of Streaming Applications on Multiprocessors', Ph. D. thesis, Eindhoven University of Technology.

[12] Arno Moonen, Marco Bekooij, and Jef van Meerbergen. 2004. 'Timing analysis model for network based multiprocessor Systems', Proceedings of Progress Symposium on Embedded Systems, pp. 122-130.

[13] Imed Bennour, Dorsaf Sebai, Abderrazak Jemai, 'Modeling SW to HW task migration for MPSOC performance analysis' DTIS , 2010.

[14] [10] Kumar, B. Mesman, B. Theelen, H. Corporaal, Y. Ha. 2008. "Analyzing composability of applications on MPSoC platforms",Journal of Systems Architecture, ISSN 1383-7621, ElsevierScience.

[15] Maarten H. Wiggers, Nikolay Kavaldjiev, Gerard J. M. Smit, Pierre G. Jansen. 2005. "Architecture Design Space Exploration for Streaming Applications Through Timing Analysis". Centre for Telematics and Information Technology, University of Twente, Enschede, Technical Report TR-CTIT-05-36.

[16] A. Shabbir, A. Kumar, S. Stuijk , B. Mesmana, H. Corporaal. 2010. "CAMPSoC: An Automated Design Flow for Predictable Multi-processor Architectures for Multiple Applications". Journal of Systems Architecture - Embedded Systems Design 56(7): 265-277.

[17] M. Geilen, 2010. "Synchronous dataflow scenarios", ACM Trans. Embedded Computing Systems, vol. 10, no. 2, pp. 16:1-16:31.

[18] Stuijk, S. et al.. 2011. "Scenario-Aware Dataflow: Modeling, Analysis and Implementation of Dynamic Applications". 11th International Conference.

[19] L. T. X. Phan, S. Chakraborty, and P. S. Thiagarajan. , 2008. "A multi-mode real-time calculus". In RTSS '08: Proceedings of the 2008 Real-Time Systems Symposium, pages 59-69, Washington, DC, USAIEEE Computer Society

[20] L. Thiele and N. Stoimenov. 2009. "Modular performance analysis of cyclic dataow graphs". In EMSOFT '09: Proceedings of the seventh ACM international conference on Embedded software, pages 127{136, New York, NY, USA.

[21] B. D. Theelen, M. Geilen, T. Basten, J. Voeten, S. V. Gheorghita, and S. Stuijk. 2006. "A scenario-aware data model for combined long-run average and worst-case performance analysis". In MEMOCODE, pages 185-194.

[22] P. Poplavko, T. Basten, and J. van Meerbergen. 2007. "Execution-time prediction for dynamic streaming applications with task-level parallelism". In DSD '07: Proceedings of the 10th Euromicro Conference on Digital System Design Architectures, Methods and Tools, pages 228-235, Washington, DC, USAIEEE Computer Society.

[23] M. Geilen. 2009. "Synchronous data flow scenarios". Transactions on Embedded Computing Systems, Special issue on Model-driven Embedded-system Design, to be published.

[24] T. Wiegand et al. 2003. "Overview of the H.264/AVC Video Coding Standard", IEEE Transactions on Circuits and Systems for Video Technology, vol. 13, No. 7, pp. 560-576.

[25] D. Marpe and T. Wiegand, J. Sullivan. 2006. Microsoft Corporation, IEEE Communications Magazine.

[26] Article about HEVC-[Online] : http://en.wikipedia.org/wiki/High_Efficiency_Video_Coding

[27] Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to About 1.5 Mbit/s—Part 2. 1993. Video, ISO/IEC 11172-2 (MPEG-1), ISO/IEC JTC.

[28] Bingjie Han , Ronggang Wang, Zhenyu Wang, Shengfu Dong, Wenmin Wang, Wen Gao. 2014. "HEVC decoder acceleration on multi-core x86 platform" IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP).

[29] Valgrind Web site: http://valgrind.org/.

[30] gprof Web site: https://sourceware.org/binutils/docs-2.16/gprof/

[31] MemProf Web site: https://wiki.gnome.org/Apps/MemProf.

[32] F. Bossen. 2012. Common test conditions and software reference configurations, 9th Meeting of the JCT-VC in Geneva (Switzerland).

[33] Zedboard Plateform web site. www.zedboard.org. Accessed, September 2016

[34] Xilinx, Inc. (2016). Zynq-7000 All Programmable SoC Technical Reference Manual. http:// www.xilinx.com/support/documentation/user_ guides/ug585-Zynq-7000-TRM.pdf.

[35] Philippe J.M., Carbon A., Brousse O., Paindavoine M. 2015. "Exploration and design of embedded systems including neural algorithms", Design, Automation & Test in europe conference & Exhibition.

3D FILTERING COLOR IMAGE CONTAMINATED BY MIXED NOISE USING SPARSE REPRESENTATION

Alfredo Palacios-Enriquez
Volodymyr Ponomaryov
Instituto Politecnico Nacional
Santa Ana 1000, San Fco Culhuacan
Coyoacan, Mexico City
email: alfredoepe@yahoo.com.mx
email: volodymyr.ponomaryov@gmail.com

Araceli Hernandez-Fragoso
Colegio de Estudios de Posgrado
de la Ciudad de Mexico
16 de Septiembre 10, Acolman
Cuautitlan Izcalli, Mexico
email: ariiitta@gmail.com

KEYWORDS
Image Denoising, Multimedia Tools, Computer Vision, Additive Noise, Impulsive Noise, Mixed Noise, Sparse Representation, PSNR, SSIM.

ABSTRACT

Filtering image has different applications, such as computer vision, multimedia tools, telemedicine, and satellite imaging, where the objective is restoring the lossed information due to the presence of noise. The noise present in an image can be modeled as a stochastic process. There are diverse reasons why the noise appears, such as: non-uniform lighting, random fluctuations in an object's surface orientation and texture, sensor limitations, non-ideal transmission, and interference. Noise affects not only the performance of an image in a specific problem but also its perceived quality. In this paper, a novel framework is presented for denoising colour images corrupted by a mixture of additive and impulsive noise. The proposed method could be described in three stages: 1)impulsive noise filtering; 2)additive noise filtering; and 3)post-processing. In the first stage, a pixel contaminated by impulsive noise should be detected, the detection is performed using the values of the pixels in local and interchannel way. Also, the restoration of corrupted pixels is performed using a filter based on the summing of distance vectors. In the next stage, filtering of additive noise is based on behavior of the additive noise on the discrete cosine transform domain, sparse representation and 3D-processing. Finally, the post-processing stage increases filtering quality using a Wiener filter.

INTRODUCTION

The presence of diverse kind of noises in an image causes degradation and losses of information in the image. In order to enhance the quality of a image is necessary to implement filtering techniques that minimize the effects produced by noise. Commonly, the filtering techniques of noise are used as a pre-processing stage in different applications, like: video transmission, remote sensing, object recognition, images for medical applications, etc., where the main objective is to reduce noise while preserving the fine details and features.
The noise can be described as a stochastic process with known statistical characteristics, and in some cases, and it is necessary to obtain these characteristics during the processing stage (Eddins et al. 2009). The additive and impulsive noises are the most common types of noise, although these are not the only types. Images may be corrupted by different reasons such as: interference and imperfections in the channel or in the reception equipment. Also, the digital cameras can introduce a noise because of failure in their sensor CCD, electronic interference or errors in data acquisition (Young et al. 2007). In an image, it can be exist a mixture of noise, this mixture is represented by the sum of different types of noises. The model of mixed noise more used is a combination of additive Gaussian noise and impulsive noise. This type of noise can be represented as follows:

$$Y(i,j)\begin{cases} y(i,j)+n_{add} & ;\text{ with probability } 1-p_k \\ n_{imp} & ;\text{ with probability } p_k \end{cases} \quad (1)$$

where $y(i,j)$ is the original, n_{add} is a random process with Gaussian probability density $N(0,\sigma^2)$, n_{imp} is modelled via Uniform probability distribution, and $Y(i,j)$ is the noisy image.

RELATED WORKS

At present, there are exist different methods of image restoration contaminated with some type of noise, where it is obtained a filtered pixel using diverse characteristics within of the vicinity. These methods of denoising image can be classified depending on the noise present in the image.
Recently, different techniques to filtering an image contaminated by additive noise have been propose based on search of similar blocks to a reference block, such as: Non-Local Means (NLM) (Buades and Coll 2005) and BM3D (Dabov et al. 2007). The similarity between two

blocks is obtained via a similarity measure. The filtering technique of NLM utilizes the auto-similarities that can exist in an image to calculate the weights that are used in denoising the contaminated image. The BM3D technique uses a 3D filtering based on an enhanced sparse representation in transform domain. The NLM, BM3D and other techniques of suppress additive noise (Bilateral Filter, Anisotropic Diffusion) present problems in the presence of impulsive noise, because the impulses are considerate as edges or fine details. So, the pixels contaminated by impulsive noise should be restoring previously.

There are different techniques for suppress of impulsive noise, mostly based on Median Filter or a variant (Lukac 2003). Other filtering techniques can be described in two stages: 1) detection of impulses; and 2) restoration of corrupted pixels. The restoration stage uses different filtering techniques (Malinski and Smolka 2016, Rosales-Silva et al. 2012).

These methods have been designed for a particular kind of noise, either additive or impulsive. The techniques proposed to filtering an image corrupted by mixed noise, usually, perform the filtering of impulsive noise in a first stage, and the filtering of additive noise is applied in a second one.

In this paper, a technique of restoration of a corrupted color image by a mixture of additive noise, with Gaussian probability distribution, and impulsive noise with a uniform probability distribution on DCT domain (3D-FMN-DCT) is developed.

PROPOSED METHOD

The proposed method to restore color image corrupted by mixed noise can be described in three stages: a) impulsive noise filtering; b) additive noise filtering; and c) post-processing. (See Fig. 1). In the following sections each a stage will be described.

Impulsive Noise Filtering

In the proposed filter, in order to reduce the impulsive noise present in an image, two processes are developed: firstly, the detection and tagged of corrupted pixels by impulsive noise in each a RGB channel is performed. After, the filtering of tagged pixels as impulses is performed in each a RGB channel in independent form, obtaining an approximation of pixel taking into consideration the correlation that exists between the RGB channels.

Detection of corrupted pixel by impulsive noise

This step is of the utmost importance, because a bad detection can generate errors and artifacts in the restored image, such as: blurring, degradation of edges and fine details. The detection is realized in exhaustive form in each a RGB channel, i.e., each pixel of an image is analyzed to know if it is contaminated or not. When a pixel is corrupted by impulsive noise, this one differs respect their neighbors. So, a way to recognize if a pixel is an impulse is comparing the central pixel and their neighbors within a vicinity of size $v \times v$.

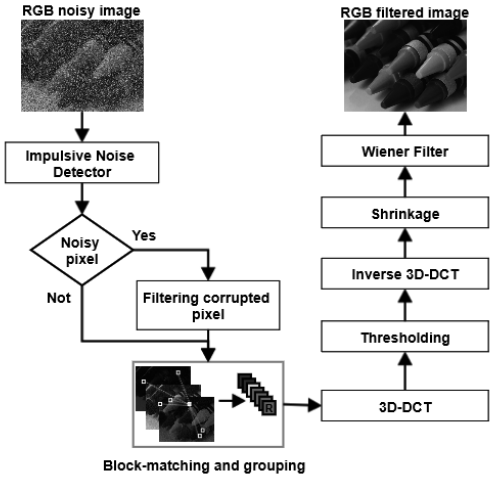


Figure 1: Block-diagram of proposed method to filtering color image contaminated by mixed noise.

Two pixels are similar when the absolute difference (AD) between they is near to zero, if it compares a pixel and an impulse, then the value of absolute difference is big. La difference absolute is defined as follows:

$$AD_{k,l} = |W(i,j) - W(k,l)|, \tag{2}$$

where $W(i,j)$ is the central pixel, $W(k,l)$ are the neighbors around of pixel central with $i \neq k$ and $j \neq l$.

After, it is necessary to consider more of one value of $AD_{k,l}$, to avoid that an edge pixel could be considerate as an impulse. Now, the values $AD_{k,l}$ are ordered in ascending form and the vector AD_v is obtained. Next, only there ones are taken p values of AD_v and a trimmed sum $SD_{i,j}$ is calculated. In order to distinguish between a corrupted pixel and a noise-free pixel, the value $SD_{i,j}$ is compared with a fixed threshold as follows:

$$Y_{tagged}(i,j) = \begin{cases} P_{noise} & ; \text{if } SD_{i,j} > T \\ Y(i,j) & ; \text{otherwise} \end{cases}, \tag{3}$$

where $Y_{tagged}(i,j)$ is the image that contains the tagged pixels as corrupted pixels by impulsive noise. The process of detection of corrupted pixels by impulsive noise is applied to each a RGB channel, therefore a tagged image is obtained for each channel $(R_{tagged}, G_{tagged}, B_{tagged})$.

Restoration of corrupted pixel impulsive noise

Once, the tagged image is obtained, the next step consists of replacing the noisy pixels using a filtering technique. The VMF proposed by Astola (Astola et al. 1990) realizes the filtering based on minimizing the sum of vector distances. In this work, this idea is taken to find a

first approximation of the corrupted pixel. The filtering process is realized to each RGB channel in independent form.

Let explain the process to Red channel. Firstly, a first approximation is obtained of a tagged pixel as noisy with the information contained in the G and B channels. Let W_R, W_G and W_B be the centered windows at position (i, j), for R, G and B channels, respectively. The absolute difference between pixels at position (i, j) is defined as follows:

$$\begin{aligned} D_{RG}(i,j) &= |W_R(i,j) - W_G(i,j)|, \\ D_{RB}(i,j) &= |W_R(i,j) - W_B(i,j)|, \end{aligned} \tag{4}$$

where $D_{RG}(i, j)$ and $D_{RB}(i, j)$ represent the absolute difference between the $R-G$ and $R-B$ channels, respectively. The positions of the minimum values in $D_{RG}(i, j)$ and $D_{RB}(i, j)$ are obtained to calculate the approximation $\widehat{R}(i, j)$, as follows:

$$R_1(i,j) = \frac{W_R(m,n) + W_R(p,q)}{2}, \tag{5}$$

where (m, n) and (p, q) are the corresponding positions of the minimum values.

The calculating of second approximation ($\widehat{R}_2$) is based in the value of the sum of absolute differences in the R channel (SD_R). The SD_R value of a pixel in the position (i, j) within W_R is defined as follows:

$$SD_R = \sum_{k,l=1}^{v} |W_R(i,j) - W_R(k,l)|, \tag{6}$$

where $i = 1, 2, \ldots, v$. It is worth mention that the SD value is not calculated to pixels tagged as noisy.

Let be (m, n) the position of the minimum value, so the $\widehat{R}_2$ value can be defined:

$$\widehat{R}_2(i,j) = W_R(m,n). \tag{7}$$

Finally, the restored pixel is obtained as follows:

$$\widehat{R}(i,j) = \frac{\widehat{R}_1(i,j) + \widehat{R}_2(i,j)}{2}. \tag{8}$$

Additive Noise Filtering

The additive noise filtering is based on sparse representation and 3D filtering on DCT domain. The techniques that use sparse representation to suppress additive noise are based in the behavior of noise in the domain of some fixed bases like: Fourier, Cosine, Wavelet, etc. Further, the filtering based on shrinkage method allows reducing the additive noise, whereas the edges and fine details suffer less deterioration when such reconstruction is performed. The proposed additive noise filtering stage is performed on DCT domain and can be divided in two steps: 1) grouping using block-matching, and 2) 3D-filtering.

Block-matching and grouping

The Block-matching technique is based in search of blocks that present a high similarity within a noisy image. The Grouping process (Dabov et al. 2007) consists in taking all 2D blocks to form a 3D structure called Group. A block is similar to another one when the similarity value is higher that a fixed threshold. In this stage, the similarity measure based on the Sum Absolute Difference (SAD).

The three channels from a color image are divided in non-overlapped blocks of the same size as the reference block. The more similar blocks and the reference block are grouped into a 3D array ($A_{3D}(i, j, k)$),where the blocks are sorted depending of grade of similarity. i.e., $A(i, j, 1) = A_1(i, j)$ is the reference block, $A(i, j, 2) = A_2(i, j)$ is the block with the major similarity grade, and $A(i, j, k) = A_k(i, j)$ has the minor similarity grade.

3D filtering

The 3D filtering uses the assumption that the noise can be expressed via sparse representation in the transform domain (Fevralev et al. 2011). In other words, if a coefficient has an absolute value approximating to zero, this means that it possibly belongs to noise, so a natural decision consists of eliminating this coefficient. The simplest way is by means of a threshold chosen according to some reason. The discrete Cosine Transform (Pogrebnyak and Lukin 2012) is applied to each a block belonging to 3D array ($A_{3D}(i, j, k)$), obtaining the 3D array $A_{3D-DCT}(p, q, r)$ in DCT domain. Once that the hard-thresholding is applied, the 3D array $\widehat{A}_{3D-DCT}(p, q, r)$ on DCT domain is obtained, this 3D array is considered as free noise.

Following, the Inverse DCT is applied to $\widehat{A}_{3D-DCT}(p, q, r)$, obtaining the estimation of the 3D array $\widehat{A}_{3D}(i, j, k)$. Now, it is necessary to obtain an approximation of the reference block from $\widehat{A}_{3D}(i, j, k)$, this process is called *shrinkage*. The output of shrinkage process $\widehat{A}(i, j)$ is defined as follows:

$$\widehat{A}(i,j) = \frac{\sum_{l=1}^{k} A_l(i,j) w_l}{\sum_{l=1}^{k} w_l} \tag{9}$$

where w_k are the weights defined as $w_k = 1 - SAD(A_1(i, j, A_k(i, j))$.

Finally, in order to obtain a filtered image $\widehat{Y}(i, j)$, the additive noise filtering is applied to each RGB channel of an image.

Post-processing

In the previous filtering stages are produced some artefacts undesirables, so in the filtered image that artefact should be corrected. A Wiener Filter (Lim 1990) is employed to increase the quality of filtered image. The

Wiener Filter is defined as follows:

$$\widehat{Y}_{Wiener}(i,j) = \mu + \frac{\sigma_W^2 - v^2}{\sigma_W^2}\left[\widehat{Y}(i,j) - \mu\right] \qquad (10)$$

where μ is the local mean, σ_W^2 is the local variance and v^2 is the average of all the local estimated variances.

EXPERIMENTAL RESULTS

The experimental results were performed using a set of color test images proposed in (Malinski and Smolka 2016). Mentioned set contains images with different texture and fine details structure that can guarantee robustness of investigating techniques.
The evaluation criterias used for measure performance are Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM). PSNR is based on comparisons using explicit numerical criteria and several references are possible such as the ground truth or prior knowledge expressed in terms of statistical parameters and tests; SSIM is based on human judgment and operates without reference to explicit criteria, and this measure was introduced in (Wang et al. 2004).
The proposed method is applied to test color images corrupted by different noise levels, and in our case a mixture of additive and impulsive noise with uniform distribution is used. In the table 1, there are shown the average PSNR and SSIM values obtained for all images.

Table 1: The average PSNR and SSIM values.

%	10	20	30	40	50
σ			**PSNR**		
10	30.09	29.51	28.69	28.07	27.23
20	28.93	28.55	27.79	27.24	26.30
30	27.72	27.34	26.48	25.89	24.79
40	26.21	25.83	24.87	24.27	23.17
50	23.99	23.65	22.72	22.26	21.36
σ			**SSIM**		
10	0.9421	0.9358	0.9261	0.9177	0.9053
20	0.9214	0.9161	0.9043	0.8940	0.8737
30	0.9008	0.8933	0.8739	0.8564	0.8229
40	0.8730	0.8611	0.8308	0.8069	0.7638
50	0.8148	0.7966	0.7562	0.7299	0.6860

Our design framework appears to demonstrate ability to eliminate a mixture of impulsive noise with random distribution (the values of the random impulses are distributed between 0 and 255) and additive noise.

Comparison with state-of-art techniques
There are different techniques for mixed noise suppression. In order to evaluate the proposed method, we compare it with the better existing state-of-art techniques: Wiener (Lim 1990), Bilateral (Zhang et al. 2014) and WESNR (Jiang et al. 2014). The filter Wiener was designed to decrease the additive noise, so it is necessary to perform, previously, the filtering of impulsive noise to compare with our technique. So, the suppression of impulsive noise is realized by proposed impulsive noise filtering stage. The same case is applied using the Bilateral filter. In figure 2, there are shown the visual results obtained to image pic027 corrupted by different values of % and σ in the case of impulsive and additive noise, respectively. The figure 2 show that the

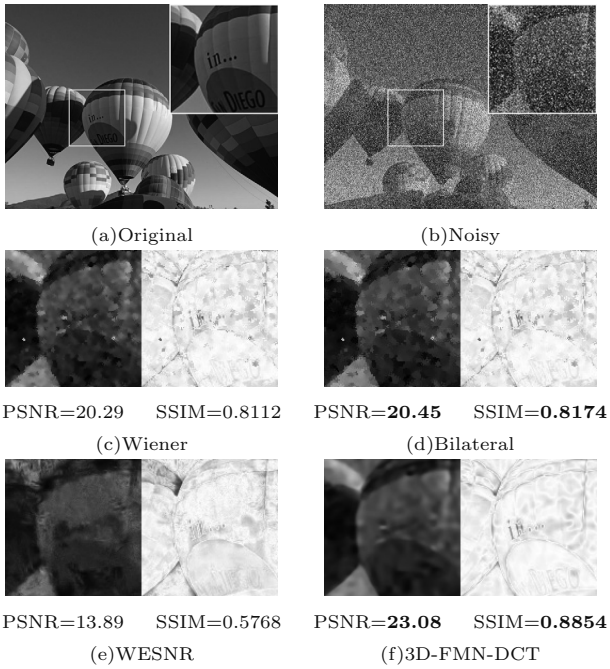


Figure 2: Filtered and inverted error images of the image *pic027* filtered by Wiener, Bilateral, WESNR and 3D-FMN-DCT techniques for a mixture of noise of Additive Noise ($\sigma = 45$) and Random Impulsive Noise ($\% = 45$).

3D-FMN-DCT technique obtains a better preservation of smoothed zones with changes in illumination and exposes a better preservation of details in the area of edges and fine details compared with the other methods.

CONCLUSIONS

A novel technique to decrease the effects produced by mixed noise (impulsive-additive) is presented. The denoising approach consists of three principal stages: impulsive noise filtering; additive noise filtering; and post-processing. The designed technique can achieve good performed in filtering without previous knowledge of the noise characteristics, such as percentage value of impulsive noise and/or variance of Gaussian additive noise.
The impulsive noise filtering stage uses a detection of contaminated pixels by impulsive noise and, after, only these pixels are filtered using local and interchannel information. The proposed detection stage appears to demonstrate good ability to distinguish between a pixel corrupted by impulsive noise and a pixel that belongs to

an edge or a detail zones, even in the presence of additive noise. In the additive noise suppression stage, 3D filtering is used with image sparse representation in the DCT domain, selecting pixels with high similarity to the 3D array. In the post-processing stage, the effects produces by previous filtering stages are corrected. The artifacts corrected in the filtered image are: the mosaicing produced by additive noise filtering using a non-sliding window. Finally, the restoration of pixels degraded by the filtering stages are performed. Future work should be devoted to implementing the current filtering approach to restore videos. For videos, more information could be used, such as the correlation between frames.

ACKNOWLEDGEMENT

Authors would like to thank to Instituto Politecnico Nacional (Mexico) and Consejo Nacional de Ciencia y Tecnologia (Mexico) (grant 220347) for their support in realizing this work.

REFERENCES

Astola J.; Haavisto P.; and Neuvo Y., 1990. *Vector median filters. Proceedings of the IEEE*, 78, no. 4, 678–689. ISSN 00189219. doi:10.1109/5.54807. URL `http://ieeexplore.ieee.org/document/54807/`.

Buades A. and Coll B., 2005. *A non-local algorithm for image denoising. Computer Vision and Pattern*, 2, no. 0, 60–65. ISSN 1063-6919.

Dabov K.; Foi A.; Katkovnik V.; and Egiazarian K., 2007. *Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering. IEEE Transactions on Image Processing*, 16, no. 8, 2080–2095. ISSN 1057-7149. doi:10.1109/TIP.2007.901238.

Eddins S.L.; Gonzalez R.C.; and Woods R.E., 2009. *Digital image processing using Matlab*. Gatesmark Publishing, Knoxville, TN, 2nd ed. ISBN 978-0-9820854-0-0.

Fevralev D.V.; Ponomarenko N.N.; Lukin V.V.; Abramov S.K.; Egiazarian K.O.; and Astola J.T., 2011. *Efficiency analysis of color image filtering. EURASIP Journal on Advances in Signal Processing*, 2011, no. 1, 41. ISSN 1687-6180. doi:10.1186/1687-6180-2011-41.

Jiang J.; Zhang L.; and Yang J., 2014. *Mixed Noise Removal by Weighted Encoding With Sparse Nonlocal Regularization. IEEE Transactions on Image Processing*, 23, no. 6, 2651–2662. ISSN 1057-7149. doi:10.1109/TIP.2014.2317985.

Lim J.S., 1990. *Two-dimensional signal and image processing*. Prentice Hall. ISBN 0139353224.

Lukac R., 2003. *Adaptive vector median filtering. Pattern Recognition Letters*, 24, no. 12, 1889–1899. ISSN 01678655. doi:10.1016/S0167-8655(03)00016-3.

Malinski L. and Smolka B., 2016. *Fast adaptive switching technique of impulsive noise removal in color images. Journal of Real-Time Image Processing*, 1–22. ISSN 1861-8200. doi:10.1007/s11554-016-0599-6.

Ng P.E. and Ma K.K., 2006. *A switching median filter with boundary discriminative noise detection for extremely corrupted images. IEEE Transactions on Image Processing*, 15, no. 6, 1506–1516. ISSN 10577149. doi:10.1109/TIP.2005.871129.

Pogrebnyak O. and Lukin V.V., 2012. *Wiener discrete cosine transform-based image filtering. Journal of Electronic Imaging*, 21, no. 4, 043020. ISSN 1017-9909. doi:10.1117/1.JEI.21.4.043020.

Rosales-Silva A.J.; Gallegos-Funes F.J.; and Ponomaryov V.I., 2012. *Fuzzy Directional (FD) Filter for impulsive noise reduction in colour video sequences. Journal of Visual Communication and Image Representation*, 23, no. 1, 143–149. ISSN 10473203. doi:10.1016/j.jvcir.2011.09.007.

Wang Z.; Bovik A.; Sheikh H.; and Simoncelli E., 2004. *Image Quality Assessment: From Error Visibility to Structural Similarity. IEEE Transactions on Image Processing*, 13, no. 4, 600–612. ISSN 1057-7149. doi:10.1109/TIP.2003.819861.

Young I.T.; Gerbrands J.J.; and van Vliet L.J., 2007. *Fundamentals of Image Processing (v.2.3)*. Delft University of Technology. ISBN 9075691017.

Zhang Y.; Tian X.; and Ren P., 2014. *An adaptive bilateral filter based framework for image denoising. Neurocomputing*, 140, 299–316. ISSN 18728286. doi:10.1016/j.neucom.2014.03.008.

BIOGRAPHIES

ALFREDO PALACIOS-ENRIQUEZ received the M.Sc. degree (2015) from the Instituo Politecnico Nacional (Mexico). His research interests include image/video processing.

VOLODYMYR PONOMARYOV received the Ph.D. degree in 1974 and D.Sci. in 1981. His research interests include signal/image/video processing, real-time filtering, etc. He has published of more than 500 international journal and conference scientific papers, and 23 patents of ex USSR, Russia and Mexico, and five scientific books in international editorials.

ARACELI HERNANDEZ-FRAGOSO graduated (2015) from Colegio de Estudios de Posgrado de la Ciudad de Mexico. She is student of M.Sc degree in teaching and education management from the same institution.

ENHANCING THE ACCURACY OF RASTER-BASED ALGORITHMS FOR FOREST FIRE SPREAD MODELLING

Yves Dumond
Laboratoire LISTIC, Université Savoie Mont Blanc
Campus Scientifique
F-73376 Le Bourget-du-Lac Cedex
e-mail: yves.Dumond@univ-smb.fr

KEYWORDS
Forest fire, simulation model, vector-based approach, raster-based approach, geographical information system.

ABSTRACT

We introduce in this paper a raster-based algorithm for forest fire growth modelling. The computation is performed on a grid of 2-D cells representing the landscape. During the simulation, the different rates of spread are calculated on the edges of a graph with the proviso that local fires are modelled by ellipses. In this graph, the neighbourhood of a given vertex is defined by a set of intra-cell edges whose orientation and direction depend on a local ellipse. The neighbourhood concerned has therefore a basically dynamic character. Next, we compare the results of this algorithm with those of another one involving static neighbourhoods. In both cases, fire spread is determined using the Dijkstra's shortest path algorithm. We conclude that the facts show that the model involving dynamic neighbourhoods provides more accurate results.

INTRODUCTION

Simulation is a corner-stone in fighting operations against forest fires. Such a technique has reached a level of maturity in North-America where it is confidently employed for operational purposes (Finney, 2004), (Peterson et al., 2009), (Tymstra et al., 2009), (Noonan-Wright et al., 2011). In Europe and the Mediterranean basin, its use is less common and related works are mostly confined to the academic world.

However, regardless of the area in which they are used, forest fire simulation tools must at least meet the two following requirements:

- The obtained fire contours must be accurate enough to fit operational purposes, that is to provide reliable assistance for means dispatching on the theater of operations. On this point, it should be noted that simulation results with an accuracy level of some meters is generally utopian. This is because it is extremely difficult to carry out a precise assessment of various parameters, e.g. current state of vegetation on wide areas, meteorological parameters, etc. Nevertheless, the deployment of fighting means is always implemented with a substantial time safety margin which allows a certain degree of imprecision to be tolerated. In fact, the major concern is the anticipation on the direction taken by the head of the fire.
- The calculations must be fast enough to allow numerous simulations within a given operation.

The work presented in this paper has been carried out in the context of a partnership between the University Savoie Mont Blanc and a Fire Department in the South of France (SDIS 06). We report on experiments in implementing simulation tools for forest fires, our approach being clearly oriented towards an effective use in the field.

FIRE SPREAD MODELLING

Research works on the topics started during the 1970s with the emergence of computing technologies. Since that time many works have been conducted (Sullivan, 2009). Three overall approaches can be distinguished: *physical models* (Richards, 2005) which aim at identifying key parameters, e.g. outdoor temperature, sun exposure, wind, slope, fuel bed properties, air moisture, nature of the soil, etc., and then to define numerical models, e.g. partial differential equations; *empirical models* (FCFDG, 1992) which are carried out with the view of establishing statistical laws based on thousands of field measurements; *semi-empirical models* (Rothermel, 1972), (Lopes et al., 2002), (Vakalis, 2004), (Kalabokidis et al., 2013) which are a middle ground between the two previous approaches, i.e. statistical laws are the basis for the elaboration of formulae which are used to calculate fire rates of spread.

It is established practice that semi-empirical models are, at least for the time being, best suited for the development of operational tools (Sullivan, 2009). They can be divided into two types, namely vector-based and raster-based approaches.

Vector-based models (Finney, 2004) assume the hypothesis that starting from an ignition at a given point, the fire spontaneously evolves to a specific shape, generally an ellipse or a double-ellipse. Note that other works use ovoids, tear drops, or lemniscates as the reference shape. The implementation of this principle leads to discretize at time t the global fire contour by a set of points which are considered as independent ignition points. The envelope curve (Glasa and Halada,

2007) of all the resulting shapes then provides fire contour at time $t + \delta t$.

In *raster-based models* (Hernándes Encinas et al., 2007), (Peterson et al., 2009), (Tymstra et al., 2009), the focus is on the division of the geographic space into a grid of cells. Usually, these have a square form, but in some cases, hexagonal cells are used (Trunfio et al., 2011). Then, starting from a set of burning cells which defines the initial fire contour, the model specifies how fire spreads from cell to cell until the simulation ends.

Beyond these two archetypes, reference shapes can also be used to improve the accuracy of raster-based models, thus defining hybrid approaches. Such a solution has been chosen in the present work. However, since our algorithm explicitly uses a grid of cells, we consider that it falls into the category of raster-based models.

BACKGROUND DATA

The algorithm described in this paper assumes that the landscape is represented by a 2-D grid of square cells of size 25m $\times$ 25m. For each cell, we have:

- The altitude at the center of the cell and its geographic coordinates.
- The local *slope*.
- The local *aspect,* namely the orientation of the cell with respect to the four points of the compass. Both the slope and the aspect are provided by the underlying geographical information system.
- The *local* wind. Here, it is important to stress the difference between upper level winds, called *laminar* winds, with the winds just above ground level, called *local* winds. The calculation of the latter requires a specific simulation process using as input data the characteristics of a given laminar wind and the relief of the area concerned. Thus, for each dominant wind, we have elaborated several dedicated wind maps according to regular variations in direction. The issue of variation in intensity, under the assumption of a fixed direction, did not prove to be relevant. Indeed, in the case of simulations involving different intensities, the resulting local wind vectors have been clearly found to be homothetic. Moreover, the wind maps have been calculated using the WindNinja software system (Butler et al., 2014).

MODELLING FIRE GROWTH WITH ELLIPSES

On the basis of the previous data, consider a cell $\mathcal{C}$ which belongs to the 2-D grid. The calculation of fire spread inside $\mathcal{C}$ has to take into account the morphological properties of the landscape. For that purpose, we reason in a plane, denoted $P_{\mathcal{C}}$, having the slope and the aspect of $\mathcal{C}$. In order to simplify its equation, this plane is defined to pass through the origin $O(0, 0, 0)$. Furthermore, the vertical projection on $P_{\mathcal{C}}$ of $\mathcal{C}$ and of any point A of $\mathcal{C}$ are respectively denoted $\mathcal{C}^*$ and A^*. Note that if the slope of $\mathcal{C}$ is different from zero, $\mathcal{C}^*$ is a parallelogram but not a square.

Although wind is the predominant factor in fire spread modelling, ascending slopes have an attractive effect which must be taken into account too. Many works have addressed the combined effects of wind and slope on fire growth (Van Wagner, 1988), (Weise and Biging, 1997), (Viegas, 2004). The algorithm presented in this article is based on a model introduced in (McAlpine et al., 1991). The slope is thus represented as an "equivalent-wind" vector. Hence, this makes it possible to define the conjugate action of wind and slope as a vectorial sum.

Now, consider an ignition point I located on one of the borders of $\mathcal{C}$. As stated above, the local winds and the morphological properties of the landscape are key elements in the model. For this reason, the calculations are made in the cell $\mathcal{C}^*$ (Fig. 1):

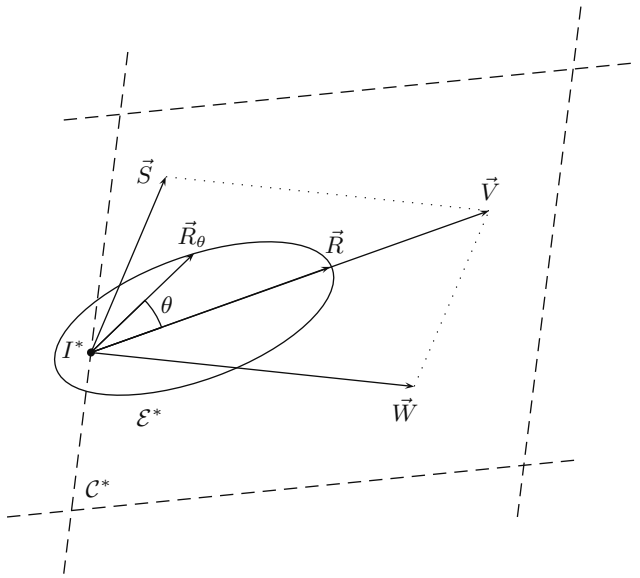


Figure 1: Ellipse construction in the cell $\mathcal{C}^*$

The ignition point in $\mathcal{C}^*$ is therefore the vertical projection of I, i.e. the point I^*. Let $\vec{W}$ be the local wind vector in $\mathcal{C}^*$. The norm and the direction of $\vec{W}$ are provided by the wind map selected according to current meteorological conditions. The local slope vector $\vec{S}$ is colinear to the strongest slope line of $\mathcal{C}^*$. Its norm is given by the following formula (Finney, 2004):

$$|\vec{S}| = 15.275 \times \beta^{-0.0225} \times \tan^2(\alpha) \quad (1)$$

where β depends on the fuel bed and α is the angle defining the strongest slope. For the Mediterranean shrub, we use the expression $10 \times \tan^2(\alpha)$ truncated at the value of 10 $m.s^{-1}$ for slopes steeper than 45 degrees.

Hence, to reify the combination of the effects of wind and slope, we set down: $\vec{V} = \vec{S} + \vec{W}$. Thus, $\vec{V}$ is called the *wind-slope resultant vector.* Moreover, the approach described in (McAlpine et al., 1991) specifies that if I^* is an ignition point, the fire spontaneously evolves to an ellipse $\mathcal{E}^*$ such that:

- I^* is located at the rear focus of $\mathcal{E}^*$.
- the direction of the major axis of $\mathcal{E}^*$ is given by that of $\vec{V}$.
- The distance between the rear focus and the forward extremity of $\mathcal{E}^*$ is defined by the fire maximum rate of spread, here represented by the vector $\vec{R}$.
- The length-to-breadth ratio LB of $\mathcal{E}^*$, from which its eccentricity can be immediately deduced, is calculated as a function of the norm of $\vec{V}$.

The different elements referenced above easily find their counterpart in the present setting. First, we have a formula for the calculation of the highest rate of spread. The latter has been designed by J-C Drouet (unpublished work) for Mediterranean shrub in the South of France. It is as follows:

$$|\vec{R}| = 180 \times e^{(0.06 \times T_s)} \times \tanh(\tfrac{100 - S_w}{150}) \times (1 + 2 \times (0.8483 + \tanh(\frac{|\vec{V}|}{30} - 1.25))) \quad (2)$$

where the parameters T_s and S_w respectively denote the temperature in the shade given in Celsius degrees and the soil water content given in millimeters (with $0 \leq S_w \leq 100$). Note that this formula also involves the norm of the wind-slope resultant vector. Moreover, it has been elaborated in order to deliver a zero speed when S_w is equal to 100 mm. This matches facts observed in the field in the area concerned.

Second, we use a formula due to (Alexander, 1985) to obtain the length-to-breath ratio of $\mathcal{E}^*$, namely the ratio between the respective length of its major and minor axes. This data, denoted LB, is calculated as a function of the norm of $\vec{V}$:

$$LB = 1 + 0.0012 \times (2.237 \times |\vec{V}|)^{2.154} \quad (3)$$

It is suggested in (Finney, 2004) that LB must be truncated at a value of 8. Furthermore, the eccentricity of $\mathcal{E}^*$ can be immediately deduced from LB by:

$$\epsilon = \sqrt{1 - \frac{1}{LB^2}} \quad (4)$$

The ellipse $\mathcal{E}^*$ being fully characterized, the fire rate of spread in a direction forming an angle θ with $\vec{R}$ is provided by the cosine correction of (Albini and Chase, 1980):

$$|\vec{R}_\theta| = \frac{(1 - \epsilon)}{(1 - \epsilon * cos\theta)} \times |\vec{R}| \quad (5)$$

NEIGHBOURHOOD OF AN IGNITION POINT

Now, we have all the elements required for the discretization of the part of $\mathcal{E}^*$ which overlaps $\mathcal{C}^*$ by a set of vectors $\vec{e}_i$ of origin I^* and with the proviso that $\vec{R}$ is one of them. The number of these vectors, say s, is defined statically and a particular care is paid to an increased density at the head of the ellipse. Note that for sake of clarity, the number of vectors occurring in Figure 2 has been drastically reduced in comparison with what is actually implemented.

Now, for all i, consider the line of orientation vector $\vec{e}_i$ passing through I^* and its intersection with one of the three other borders of $\mathcal{C}^*$. This provides a set of points hereafter denoted J_i^* :

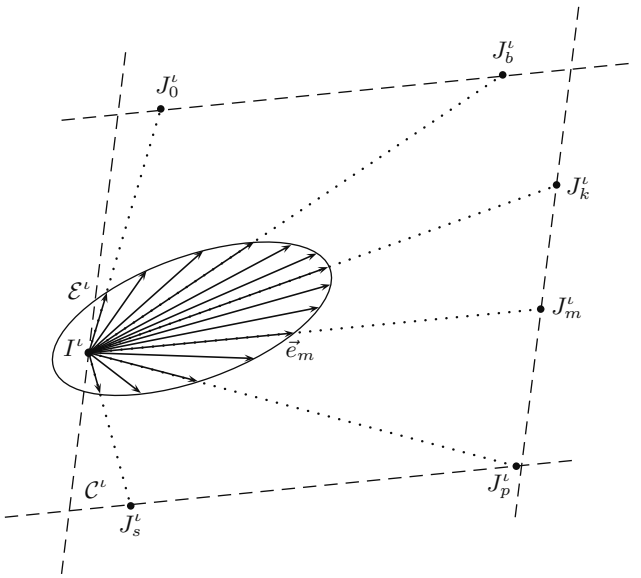


Figure 2: Discretization of $\mathcal{E}^*$

The vertical projection of the different points J_i^* on the plane $P_\mathcal{C}$ in turn provides a set of points J_i. Note that we have not considered the vertical projection of the ellipse $\mathcal{E}^*$ since such an operation does not preserve foci. Furthermore, since angles are not preserved either, the distribution of the angles between the edges IJ_i may become less regular than that between the edges $I^*J_i^*$. If the case occurs, additional points are inserted on the borders of $\mathcal{C}$ alongside with the points J_i, thus resulting in a new set of points K_i.

The different edges IK_i constitute the out-neighbourhood of I in $\mathcal{C}$. Fire propagation from I will therefore be considered in $\mathcal{C}$ through this neighbourhood.

PROPAGATION PHASE

Now, we turn to fire propagation in $\mathcal{C}$. In our approach, the four borders of every cell of the 2-D grid are divided in a fixed number of equal segments (Fig. 3). Each segment has *at most one* potential ignition point with a corresponding ignition time. Although this point obviously belongs to the segment in question, its location is not predefined, i.e. it may vary from segment to segment. This makes a significant difference with common raster-based algorithms.

Besides, we say that a segment is ignited when the corresponding point is ignited. When this occurs, a propagation phase which implements fire spread towards other segments is performed. Every segment of the grid is assumed to be in one of the following states:

- *Burned*: the segment concerned has been formerly ig-

nited and the corresponding propagation phase has been achieved. Thus, both the ignition point and the ignition time are final.

- *Ignition*: the segment is the latest to have been ignited. It was previously in the state "pre-ignition". The propagation phase related to this segment is in progress. When the latter is terminated, the segment switches to the state "burned". At a given moment, only *one* segment in the 2-D grid is in the state "ignition". This is a direct result of the use of the Dijkstra's shortest path algorithm to handle global fire spread on the grid.
- *Pre-ignition*: at least one ignition point and therefore one ignition time, have been calculated for the segment in question. However, as long as the segment has not switched to the state "ignition", these data may be modified.
- *Outer*: no ignition point has been determined for the segment concerned, which means that it is outside of the fire contact area.

The actions carried out during the propagation phase are specified hereafter. Let K_ℓ be a vertex belonging to the out-neighbourhood of I in $\mathcal{C}$. We study fire spread along the edge IK_ℓ. We denote σ_I and σ_{K_ℓ} the segments I and K_ℓ respectively belong to.

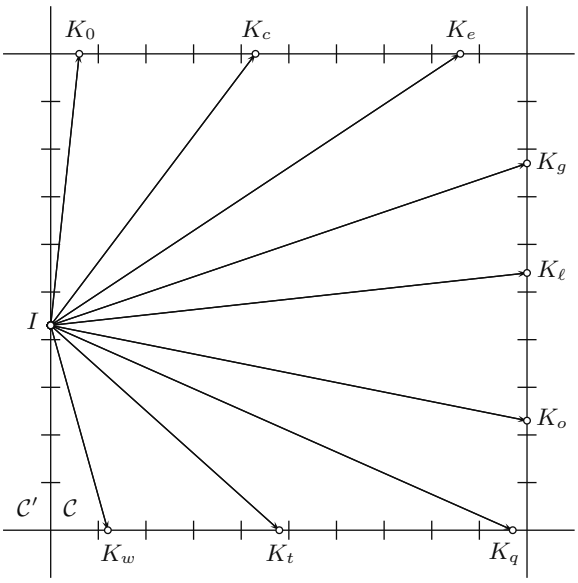


Figure 3: Some vertices of the out-neighbourhood of I in the cell $\mathcal{C}$

Different cases must be considered depending on the state of σ_{K_ℓ} which, by definition, cannot be "ignition":

- "burned": K_ℓ cannot be a new ignition point and thence no further action is required.
- "pre-ignition": the segment σ_{K_ℓ} already has an ignition point, say K'. Therefore, we must calculate the potential ignition time of K_ℓ and then compare it to that of K'. Let ω be the angle between the vector $\vec{R}$ with an origin at I^* and the line segment $[I^*, K_\ell^*]$. The equation (5) provides the fire rate of spread along $[I^*, K_\ell^*]$ as a function of ω. Hence, we immediately get the spreading duration $\tau(I^*, K_\ell^*)$ between the two points. Since we can postulate that $\tau(I^*, K_\ell^*) = \tau(I, K_\ell)$, the ignition time of K_ℓ is obtained by adding that of I to this duration. If the calculated value is earlier than the ignition time of K', then K_ℓ becomes the new ignition point of σ_{K_ℓ}. Otherwise, K' holds its status. It should be emphasized that in the general case, the respective locations on σ_{K_ℓ} of the two points K' and K_ℓ are different.
- "outer": K_ℓ becomes the first ignition point of σ_{K_ℓ} and it is equipped with the corresponding ignition time which is calculated as in the previous case.

The previous sequence of actions is obviously repeated for all the vertices of the out-neighbourhood of I. Together, these sequences constitute the propagation phase of σ_I in $\mathcal{C}$.

It is important to note that regardless of the location of the segment which has caused the ignition of σ_I, fire propagation starting from I must be considered both in $\mathcal{C}$ and in the adjacent cell $\mathcal{C}'$. In the case of the latter, the principles are clearly identical to those introduced for $\mathcal{C}$. It is just sufficient to mention that the different calculations are based on the ellipse defined in the 3-D cell $\mathcal{C}'^*$.

Finally, note that if an ignition point is located at the corner of four cells then the propagation phase must be performed in the four adjacent cells using each time the corresponding ellipse.

With respect to the process of fire propagation just described, three important comments can be made:

- Having only one ignition point for each segment allows to prevent explosion in the number of points to consider.
- Depending upon the number of segments per cell and that of elements in neighbourhood of I, some segments of $\mathcal{C}$ whose state is either "pre-ignition" or "outer" may not contain any element K_i. Nevertheless, due to the number of edges involved, this did not turn out to be a problem.
- One must stress the fundamentally dynamic character of the neighbourhoods generated by the above process. In fact, the model presented here was designed with this main objective. To understand the interest of this approach, suppose the location of I^* and the orientation of $\mathcal{E}^*$ are such that the main rate of spread vector $\vec{R}$ with an origin at point I^* is inside $\mathcal{E}^*$ (as in Fig. 1). With neighbourhoods involving fixed edges, the projection of $\vec{R}$ on $P_\mathcal{C}$ generally does not fit the direction of any edge. This inevitably results in under-estimating the propagation of the head of the fire. On the contrary, in the present work, there exists *by construction* a vertex K_j such that the edge IK_j exactly matches the direction of the projection of $\vec{R}$. Therefore, the progression of the head of the fire, which is a key issue, is taken into account more accurately.

GLOBAL FIRE SPREAD CALCULATION

In the present setting, fire spreads over a directed graph, denoted $\mathcal{G}$, whose vertices are segments of the 2-D grid. Its set of vertices can be split into three sub-sets:

- $\mathcal{B}$ for the segments in the state "burned".
- $\mathcal{I}$ for the unique segment in the state "ignition".
- $\mathcal{A}$ for the segments in the state "pre-ignition". This set of segments defines the fire vicinity.

It is worth noting that all the segments of the 2-D grid which are in the state "outer" are not, at least temporarily, in the graph. Indeed, it follows from the previous section that $\mathcal{G}$ grows dynamically. The global process of fire spread simulation, which uses the principles of the Dijkstra's shortest path algorithm, goes through two stages:

- *Phase 1:* the initial fire contour is drawn on the screen and the meteorological parameters required by the simulation are captured. The different cell borders forming the fire contour are identified. An ignition point is assigned to every corresponding segment which consequently switches to the state "pre-ignition". This point is located at the center of the segments concerned and the related ignition time is given the value 0. The segments which are inside the fire contour, i.e. not on its border, are assigned the state "burned". All the other segments are assigned the state "outer".

- *Phase 2:* ignition times clearly induce an order among the elements of $\mathcal{A}$. Therefore, $\mathcal{A}$ admits a sub-set bringing together all the minimum elements. Let us choose at random one of them, say σ. Then, σ is switched to the state "ignition." It is consequently removed from $\mathcal{A}$ and put in $\mathcal{I}$. The propagation phase of σ is thus performed. This results in the fact that some segments of $\mathcal{A}$ can have their ignition conditions, i.e. point and time, modified. Moreover, some segments being in the state "outer" can be introduced in the graph: indeed, their first ignition point and ignition time are calculated on this occasion. When the ignition phase of σ is finished, the latter switches to the state "burned" and it is therefore removed from $\mathcal{I}$ and inserted in $\mathcal{B}$.

  Another segment is then the chosen in $\mathcal{A}$ according to the same criteria and the sequence of actions described above is repeated until the earliest element of $\mathcal{A}$ is posterior to the dead-line of the simulation.

FIRE SPREAD WITH STATIC NEIGHBOURHOODS

For comparison purposes, we briefly introduce below an alternative raster-based algorithm for forest fire spread modelling (Dumond, 2016). The underlying model and the different elements of calculation are the same as those described in the present work except for the fact that, in that case, static neighbourhoods are used.

Thus, considering the 2-D grid, a directed graph $\mathcal{S}$, over which fire spread is calculated, is defined as follows: the vertices of $\mathcal{S}$ are the centers and the corners of the cells of the grid; the out-neighbourhood of a given vertex depends on its "type", i.e. "center" or "corner" and it is defined by numerous edges, as shown in Figure 4:

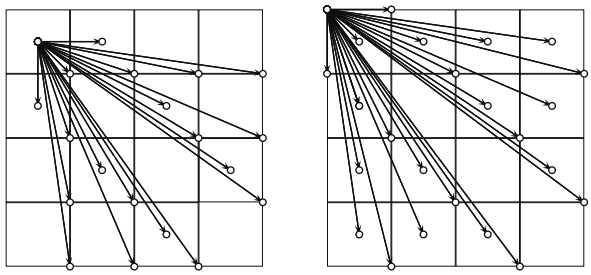


Figure 4: Southeast out-neighbourhood of vertices of respective type "center" and "corner"

It should be noted that a given vertex belongs to the out-neighbourhood of several vertices. Moreover, many edges cross several cells. Under this basis, they can be split into different segments, this term having a meaning different from that used in the previous algorithm.

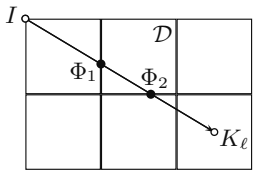


Figure 5: An edge with four segments

Without loss of generality, consider an edge IK_ℓ made of four segments (Fig. 5). The calculation of the spreading duration between the points I and K_ℓ requires the determination of four different spreading durations, i.e. one for each segment the edge IK_ℓ is made of. Let us for instance pay attention to the segment $[\Phi_1, \Phi_2]$. This segment is inside a given cell, say $\mathcal{D}$. As for the first algorithm, the calculations are made in 3-D inside the cell $\mathcal{D}^*$ and using the corresponding ellipse. This process is obviously repeated for each segment. Then, it is enough to add together the corresponding values to get the spreading duration between I and K_ℓ.

The overall fire spread propagation is then calculated with the same principles as those used in the previous case, namely with the Dijkstra's shortest path algorithm.

To close the description of this second algorithm, recall that the use of long edges is sometimes criticized since it is supposed to induce distortions. This is clearly true if the rate of spread along all the edge is calculated only under the basis of the physical properties of the cell in which the ignition occurs. On the contrary, in the model proposed here, the calculation is decomposed so as to use local data only. That being said, it must also be acknowledged that there is no assurance that in a chaotic relief the longest edges necessarily match the fastest ways.

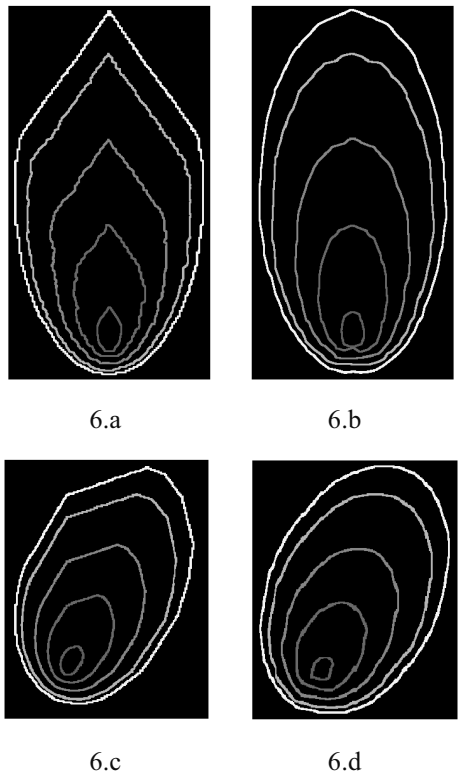
COMPARISON OF THE TWO ALGORITHMS

We consider hereafter two kinds of test which we respectively refer to as *in vitro* and *in vivo* tests. In both cases, they are carried out with the following provisos:

- In the algorithm involving dynamic neighbourhoods, ellipses are discretized by a set of 48 vectors. This does not count the edges which may be added to regularize the distribution of directions in the out-neighbourhood of some ignition points. Furthermore, cell borders are divided into 20 segments.
- The static out-neighbourhoods used by the other algorithm are those described in Figure 4.

In vitro tests

In the *in vitro* tests, simulations are performed under the assumptions of a single ignition point, a flat land and a constant wind. In that case, the use of a geographical information system is not required. The main objective is in fact to compare shapes provided by simulation with ellipses predicted by the model. Whenever these shapes are or evoke ellipses, particular interest is given to characteristics such as orientation and length-to-breath ratio. In case of static neighbourhoods involving a limited number of edges, e.g. 8 or 16, the shapes obtained may widely differ from ellipses (Cui and Perera, 2008), (Peterson et al., 2009). The following are the results obtained with the two algorithms introduced in this article:



6.a 6.b

6.c 6.d

Figure 6: In vitro tests

Static and dynamic neighbourhoods have been respectively used for the simulations whose results correspondingly appear in the first and in the second column. The pictures on the first row show propagations along the y-axis. On the second row are the results of propagations in a direction forming an angle of 20° with the y-axis. Among others, this direction entails the distortion shown in picture 6.c. In each cases, the orientation and the length-to-breath ratios are correct. However, the shapes obtained using dynamic neighbourhoods, whilst not perfect, are clearly much more satisfactory. This suggests that the corresponding algorithm could estimate fire contours more accurately.

In vivo tests

In the *in vivo* tests, comparisons are made using an implementation of the two algorithms on the same geographic information system. Thus, comparisons can be conducted between the respective contours obtained from the two algorithms or between results of simulations and field surveys. Though it is essential, the second task is clearly the trickiest part of the work. Indeed it is subject to rigorous conditions:

- It is mandatory to have field surveys established accurately. In addition to precise contours, current meteorological conditions must have been recorded.
- Fire growth must not have been hampered by the actions undertaken by fire fighters.

It is generally difficult to meet these requirements. Aerial photographs taken at short and regular time intervals would be ideal. Such data are rarely, if ever, available. In some cases, extracts from archives can be used (Fig. 7).

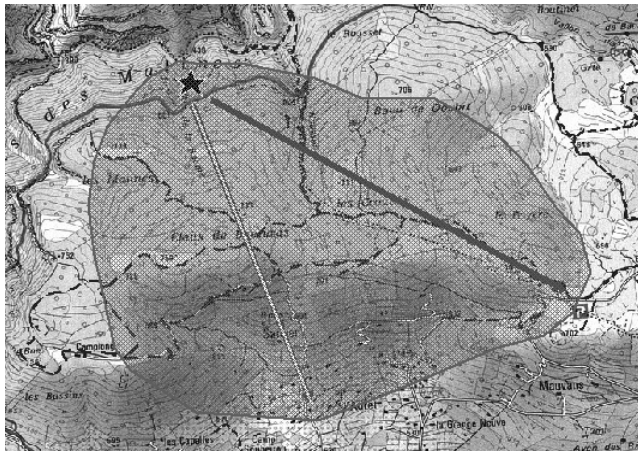


Figure 7: Contour of a past fire (Courtesy of SDIS 06). In this case, the fire traversed approximately 2490 meters along its main axis in around 3 hours.

Valid studies can thus be carried out but in a limited number of cases. Such a work has been done for the algorithm using static neighbourhoods and a certain amount of correlation between simulations and field data has been established (Dumond, 2016). Nevertheless, the global trend is clearly to

a certain under-rating of fire contours, which makes perfect sense considering the approximations induced by this algorithm.
In the present setting, we can in addition consider the contours provided by the algorithm involving dynamic neighbourhoods. The situation set out in Figure 8 is, in this respect, very representative of the results we have achieved in this new study. As one would expect, the contours obtained thanks to a dynamic neighbourhood (in yellow) encompasses that obtained with a static neighbourhood (in blue). The difference remains however limited. Furthermore, since approximations cannot be fully eliminated, the largest of both contours remains within the actual fire area.

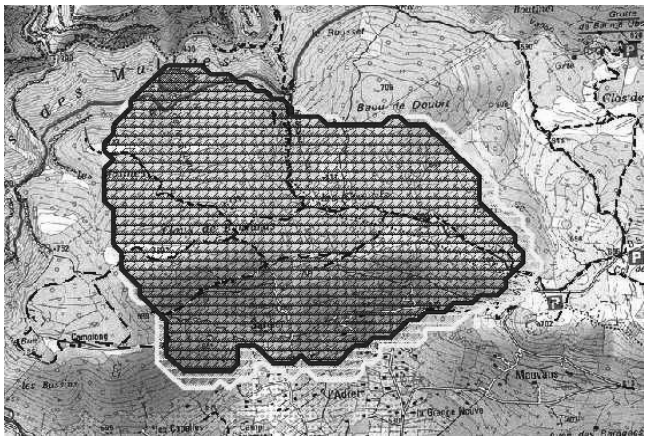


Figure 8: Results of simulations

For the tests whose results are given above, the physical parameters were as follows: laminar wind intensity=37 $km.h^{-1}$; temperature in the shade= $32°C$; soil water content=$40mm$.

CONCLUSION

We have introduced in this paper an algorithm for forest fire growth modelling whose specificity lies in the use of a dynamic neighbourhood for every ignition point considered. Two main conclusions can be drawn from this study:

- Fire shapes provided by the new algorithm are almost systematically between those obtained with static neighbourhoods and field surveys. This tends to prove its relevance and the interest of improving the calculation of rates of spread in the vicinity of fire's head. It may be noted, however, that the shift remains subdued. This is probably due to the significant number of edges used, at least in the present setting, for the definition of static neighbourhoods. Indeed, this allows considering many different possible paths. Nevertheless, a highest level of accuracy seems established for the new algorithm.
- Under the aforementioned hypotheses, namely ellipses discretized by a set of 48 vectors and cell borders split into 20 different segments, the algorithm with dynamic neighbourhood has been demonstrated to be approximately twice as fast as the other. For example, a simulation corresponding to a fire propagation of 3 hours on the ground takes less than two minutes on a laptop equipped with a Core I7 processor. Again, the difference between the two algorithms seems to be due to the structure of the static neighbourhoods whose longest edges require the calculation of up to 6 different rates of spread.

In conclusion, the accuracy of fire growth prediction has been enhanced by the algorithm proposed in this paper. Hopefully this can achieve beneficial effects at an operational level.

ACKNOWLEDGEMENT

The author wishes to thank:

- The Fire Department of the Alpes-Maritimes (SDIS 06) for its logistic support.
- The Rocky Mountain Research Station (Missoula, Montana) of the US Forest Service which graciously provides the WindNinja software system.
- Mr Bryan Lovell whose help has been extremely precious in improving the English style of the text.

BIOGRAPHY

Yves Dumond obtained his PhD in Computer Science from the University of Nice Sophia-Antipolis (France) in 1988. He is currently associate professor at the University Savoie Mont Blanc, where he has worked since 1990. His topics of interest include concurrency theory and knowledge engineering.

REFERENCES

Albini, F.A. and Chase, C.H. (1980). Fire containment equations for pocket calculators. *Research paper, RP-INT-268*, US Department of Agriculture, Forest service, Intermountain Forest and Range experiment station, Ogden, Utah, USA, 18 p.

Alexander, M.E. (1985). Estimating the length-to-breadth ratio of elliptical forest fire patterns. In *Proceedings of the 8th Conference on Forest and Fire Meteorology*, 29 April-2 May, Detroit, Michigan, USA, pp. 287–304.

Butler, B., Forthofer, J. and Wagenbrenner N. (2014). An update on the WindNinja surface wind modeling tool. *Advances in Forest Fire Research*, Domingos Xavier Viegas Editor, Coimbra University Press, pp. 54–60.

Cui, W. and Perera, A.H. (2008). A study of simulation errors causes by algorithms of fores fire growth models. *Forest Research Report No. 167*, Ontario Forest Research Institute, Sault ste. Marie, Ontario, Canada, 17 p.

Dumond, Y. (2016). Forest fire spread modeling in practice. In *Proceedings of the 30th European Simulation and Modelling Conference*, 26-28 October, Las Palmas, Gran Canaria, Spain, pp. 341–347.

Finney, M.A. (2004). FARSITE: fire area simulator-model development and evaluation. *Research paper RMRS-RP-4 Revised*, US Department of Agriculture, Forest Service, Rocky Mountain Research Station, Ogden, Utah, USA, 52 p.

Forestry Canada Fire Danger Group (1992). Development and structure of the Canadian Forest Fire Behaviour Prediction System. *Information Report ST-X-3*, Ottawa, Canada, 63 p.

Glasa, J., and Halada, L. (2007). Enveloppe theory and its application for a forest fire front evolution. *Journal of Applied Mathematics, Statistics and Informatics*, vol. 3, no. 1, pp. 27–37.

Hernándes Encinas, A., Hernándes Encinas, L., Hoya White, S., Martín del Rey., A., and Rodríguez Sánchez, G. (2007). Simulation of forest fire fronts using cellular automata. *Advances in Engineering Software*, no. 38, pp. 372–378.

Kalabokidis, K., Athanasis, N., Gagliardi, F., Karayiannis, F., Palaiologou, P., Parastatidis, S. and Vasilakos, C. (2013). Virtual Fire: a web-based GIS platform for forest fire control. Ecological Informatics, no. 16, pp. 62-69.

Lopes, A.M.G., Cruz, M.G., and Viegas, D.X. (2002). Firestation – an integrated software system for the numerical simulation of fire spread on complex topography. Environmental Modelling & Software, no. 17, pp. 269–285.

McAlpine, R.S., Lawson, B.D., Taylor, E. (1991). Fire spread across a slope. In *Proceedings of the 11th Conference on Fire and Forest Meteorology*, 16–19 April 1991, Missoula, Montana, USA, pp. 218–225.

Noonan-Wright, E.K., Opperman, T.S., Finney, M.A., Zimmerman, G.T., Seli, R.C., Elenz, L.M., Calzin, D.E., and Fielder, J.R. (2011). Developing the US Wildland Fire Decision Support System, Journal of Combustion, vol. 2011, 14 p.

Peterson, S.H., Morais, M.E., Carlson, J.M., Dennison, P.E., Roberts, D.A., Moritz, M.A., and Weise, D.R. (2009). Using HFire for spatial modeling of fire in shrublands. *Research paper PSW-RP-259*, US Department of Agriculture, Forest Service, Pacific Southwest Research Station, Riverside, California, USA, 44 p.

Richards, G.D. (2005). An elliptical growth model of forest fire fronts and its numerical solution. *International Journal for Numerical Methods in engineering*, vol. 30, no. 6, pp. 1163–1179.

Rothermel, R.C. (1972). A mathematical model for predicting fire spread in wildland fuel. *General Technical Report INT-115*, US Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station, Ogden, Utah, USA, 79 p.

Sullivan, A. (2009). Wildland surface fire spread modeling, 1990-2007. 1: Physical and quasi-physical models. 2: Empirical and quasi-empirical models. 3: Simulation and mathematical analogue models. *International Journal of Wildland Fire*, vol. 18, no. 4, pp. 349–403.

Trunfio, G.A., D'Ambrosio, D., Rongo R., Spataro W. and Di Gregorio S. (2011), A New Algorithm for Simulating Wildfire Spread through Cellular Automata. *ACM Transactions on Modeling and Computer Simulation*, vol. 22, pp. 1–26.

Tymstra, C., Bryce, R.W., Wotton, B.M. and Armitage, O.B. (2009), Development and structure of Prometheus: the Canadian wildland fire growth simulation model. *Information Report NOR-X-417*, Natural Resources Canada, Canadian Forestry Service, Northern Forestry Centre, Edmonton, Alberta, Canada, 88 p.

Vakalis, D., Sarimveis, H., Kiranoudis, C., Alexandridis, A., and Bafas, G. (2004). A GIS based operational system for wildland fire crisis management I. Mathematical modelling and simulation. Applied Mathematical Modelling, vol. 28, Issue 4, pp. 389–410.

Van Wagner, C.E. (1988). Effect of slope on fires spreading. Canadian Journal of Forest Research, vol. 18, no. 6, pp. 818–820.

Viegas, D.X. (2004). Slope and wind effects on fire propagation. International Journal of Wildland Fire, vol. 13, pp. 143–156.

Weise, D.R., and Biging, G.S. (1997). A qualitative comparison of fire spread models incorporating wind and slope effects. Forest Science, vol. 43, no. 2, pp. 170–180.

FINANCIAL SIMULATION

Analysis of Relationship between Risk and (expected) Return of the Investment (Portfolio) – Simulation Experiment on the Prague Stock Exchange

Adam Borovička
Department of Econometrics
University of Economics, Prague
W. Churchill Sq. 4, Prague
Czech Republic
E-mail: adam.borovicka@vse.cz

KEYWORDS
Monte Carlo simulation, portfolio, relationship of risk and return, stock

ABSTRACT

A relationship between risk (volatility) and (expected) return is a big question discussed in a number of papers and studies that do not provide the uniform results. In this paper, this interesting problem is developed through highly liquid stocks traded on the Prague Stock Exchange. In contrast to the most current studies provided a deterministic analysis, a stochastic procedure is proposed in order to analyze this problem satisfactorily. In the first phase of this procedure, the investment portfolios from selected stocks listed in PX Index with different level of risk are made by means of non/linear mathematical programming model. In terms of the second phase, Monte Carlo simulation experiment with these portfolios are performed for several specified time periods. Probability distribution of returns for each stock is determined by the appropriate nonparametric statistical using the historical observations. On the basis of simulation and probability analysis, a form of relationship of risk (volatility) and (expected) return for both investment portfolios and their individual components is declared. The results show that this relationship is ambiguous. It is particularly influenced by time period, or price development on the capital market.

INTRODUCTION

Risk can be primarily comprehended as a fear that the investment will not produce an expected return. Thus, it is obvious that a risk is related to the return. The question of a relationship of the risk and (expected) return of the investment instruments (mostly stocks) is discussed in many papers and studies. Most are dedicated to a verification of a validity of CAPM model which is namely developed by Sharpe (1964). Some resources confirm its functionality – positive linear dependence of the return on the risk (e.g. Fama and MacBeth 1973; Koutmos et al. 1993; Omet et al. 2002; Salman 2002;), some prove its invalidity (e.g. Fama and French 1992; Fama and French 1996; Jegadeesh, 1992; Pamane and Vikpossi 2014). These articles actually declare a beta coefficient (based on the returns' covariances) as a bad measure of the (systematic) risk. However, this is not the only concept of how risk of the investment can be understood, or measured. Only a few papers deal with a relationship of the risk expressed as a (conditional) volatility and return (e.g. Guo et al. 2007). Even with regard to the failure of beta coefficient as a risk measure, the risk is often comprehended as a volatility (instability over time) expressed by some statistical characteristics from the historical returns – standard deviation, variance, semivariance, average absolute negative deviation etc. These measures (with the others) can produce an interesting view of a distribution (development) of returns in the past.

Now I can generate some important and interesting question. Is there any general relationship between risk and return of the investment? Does a higher risk ensure a higher return? Maybe even more interesting, does a higher risk produce a greater ability to overcome the average (expected) return? Are the answers to these questions valid for any time periods? Is a situation the same for both the entire investment portfolio and its individual components? Answers to these questions can be greatly helpful for an investment decision making. So, the main aim of this article is to find the answers to these questions.

Most articles make simple regression analysis to study a relationship of the risk (volatility) and return (e.g. Guo et al. 2007). Return of the investment instrument (or investment portfolio) should be comprehended as a stochastic element characterized by some probability distribution. This fact is not often taken into account in the analyses. I propose a more complex approach including this typical element of uncertainty for the capital market. Then a proposed procedure is designed as stochastic. It consists of two phases. In the first one, the investment instruments are chosen. Several times periods with various price development are determined. All data are selected and appropriate characteristics are calculated for each period. Then two investment portfolios (with the lowest/highest risk measured by an average absolute negative deviation) for each period is determined via a model of (non)linear mathematical programming. In the second phase, the probability distributions of returns of the investment instruments, or portfolios are determined by an appropriate nonparametric statistical test. Then the scenarios of returns of the investment over the tracked time period are generated. From this Monte Carlo simulation experiment, necessary final values, returns of the investments can be calculated. Due to these characteristics and probability analysis, the form of a relationship of the risk and return can be specified.

This procedure is applied to the Prague Stock Exchange that is the main stock exchange in the Czech Republic. One reason for this choice is that an investment on the Czech Stock Exchange is more and more popular. And there is no such an analysis which could help many Czech or foreign investors. The results of this analysis show that a relationship of the risk and return is ambiguous. It is namely based on a price development on the capital market, as well as a selection of the investment instruments or portfolios.

Let me summarize the main contributions of this article. From a theoretical point of view, it is a proposed procedure for an analysis of a relationship between investment risk and return. The benefit of this analysis is an inclusion of the stochastic returns. Moreover, this procedure can be used for both individual investment instruments and the investment portfolio. From a practical point of view, it is a unique analysis on the Prague Stock Exchange for both portfolios and particular stocks which can be a support for many investors.

The article has the following structure. After the Introduction, the methodical approach (stochastic procedure) for an analysis of the specified problem is proposed. In the next section, the Prague Stock Exchange is necessarily described. Followed by a practical part, where all analyses are performed, so stocks selection, stock portfolios making, Monte Carlo simulation experiments and final results discussion. Finally, the article is summarized and some ideas for future research are outlined.

PROPOSED STOCHASTIC PROCEDURE

For the reasons stated in Introduction, to analyse a relationship of the risk and return of the investment (stock or stock portfolio in this paper), the stochastic procedure is proposed. This procedure consists of two phases.

First phase

At first, the capital market and stocks must be selected. For an analysis complexity, a relationship of the risk and return is studied in several time periods with various price development. Then all necessary data for each period, namely the time series of prices, are collected. Then the (capital) return (daily, weekly, monthly, yearly) of the stocks are calculated. The risk of the stock can be expressed as standard deviation, variance, semivariance, average absolute negative deviation etc. I do not prefer a usual measure variance or standard deviation. The main reason is that the negative and also positive deviations are included. However, the positive deviations from mean are desirable. For investment portfolio risk, similar situation occurs with the covariances of returns in the Markowitz concept (Markowitz 1952; Markowitz 1959). This drawback is partly solved by semivariance approach (Markowitz 1959). Moreover, the risk is a nonlinear function which can cause a problem in finding a solution of the mathematical model for a portfolio selection. To eliminate all negative aspects of these risk measures, I propose to apply average absolute negative deviation concept. It is computed for the *i*-th stock as follows

$$r_i = \frac{\sum_{x_{ij} < \bar{x}_i} (\bar{x}_i - x_{ij})}{m},$$

where $x_{ij}\ (i = 1,2,...,n;\ j = 1,2,...,m)$ is the *j*-th historical return of the *i*-th stock, $\bar{x}_i$ is the average return of the *i*-th stock and *m* is a number of historical returns which are lower than the average return. Then the risk of the entire portfolio can be calculated as the following weighted sum

$$r_p = \sum_{i=1}^{n} r_i x_i,$$

where $x_i\ (i = 1,2,...,n)$ is a share of the *i*-th stock in the portfolio and *n* is a maximum possible number of stocks in the portfolio. This indicator is a representative measure of risk. It indicates the average negative deviation from the average value. It is comprehensible, easily applicable in the real situations. The portfolio return is analogically specified as a weighted sum of the returns of the individual stocks.

To be able to analyse a relationship of the return and risk in the level of stock investment portfolios, the portfolios with different level of risk for each period must be made. A portfolio selection is performed by a minimizing (maximizing) linear mathematical programming model that can be formulated in the following form

$$\begin{aligned} &\min/\max \quad z = \sum_{i=1}^{n} r_i x_i \\ &\sum_{i=1}^{n} x_i = 1 \\ &\mathbf{x} \in X \end{aligned},$$

where $\sum_{i=1}^{n} x_i = 1$ is a "portfolio" condition, $\mathbf{x} = (x_1, x_2, ..., x_n)$ is a vector of variables representing a share of the stock in the portfolio and *X* is a set of other possible conditions (e.g. conditions for a portfolio diversification).

Second phase

At the beginning of the second phase, a probability distribution of returns of each stock from the portfolios for each time period must be determined for a simulation experiment. For this purpose, the nonparametric statistical tests are applied. In case of a small number of historical returns (observations), Durbin-Watson test is applied. Otherwise, Kolmogorov-Smirnov test can be used. Then the returns of each stock are generated from a particular probability distribution. For the representative results, the number of scenarios should be greater, for instance 1000 scenarios describing a possible development of stock returns for each portfolio related to a determined time period. Now, the investment strategy is specified. The strategy can be represented by a one-shot investment at the beginning of the time period or by a continuous investing in each stated time interval (e.g. month as in the practical part) during the entire time period. Of course, we can work only with the returns. But in my opinion, the analysis is more telling with some concrete amount of invested money.

For instance, if the stated interval in terms of the watched time period is month, at the end of each month a value of the stock is calculated on the basis of generated monthly returns.

This final value of the stock at the end of the chosen time period is compared with the investment to obtain a return of this investment (in percent) that can be calculated as

$$return\_inv_i = \frac{final_i}{investment_i} - 1\,(100\%)\,, \quad (1)$$

where $final_i$ represents a final value of the *i*-th stock and $investment_i$ is an invested amount of money in the *i*-th stock. The final value $final_i$ is determined as an average of the final values over all scenarios. The return of the *j*-th entire portfolio (in percent) is formulated as follows

$$return\_port_j = \frac{\sum_{i=1}^{n} final_i}{\sum_{i=1}^{n} investment_i} - 1\,(100\%)\,, \quad (2)$$

where *n* is a number of stocks in the portfolio. On the basis of these returns' characteristics (and level of risk), the relationship between risk and return of the investment can be analyzed.

Further, the average (expected) return is calculated for each stock, or stock portfolio. It is determined by the formulae (1) and (2), but the final value is calculated from one possible scenario represented by the average (monthly) return. The relationship of the final and average (expected) return is expressed as the following proposed ratio for the *i*-th stock with positive, or negative return

$$ratio_i = \frac{return\_inv_i}{exp\_return\_inv_i}, \text{ or } ratio_i = \frac{exp\_return\_inv_i}{return\_inv_i},$$

where $exp\_return\_inv_i$ is an average (expected) return of the *i*-th stock. Ratio for *j*-th portfolio is analogically calculated as

$$ratio_j = \frac{return\_port_j}{exp\_return\_port_j}, \; ratio_j = \frac{exp\_return\_port_j}{return\_port_j},$$

where $exp\_return\_port_j$ is an average (expected) return of the *j*-th stock portfolio. To consider a shape of returns distribution satisfactorily, the ratio must be improved by the probability that the final return is better than expected. Then the probability ratio is proposed as follows

$$prob\_ratio_i = ratio_i * prob_i, \; prob\_ratio_j = ratio_j * prob_j,$$

where $prob_i$, or $prob_j$ is a mentioned probability for the *i*-th stock, or the *j*-th stock portfolio. It is calculated as a ratio of the number of scenarios where the final value of investment is higher than expected and the total number of scenarios. The value of ratio expresses a power of the investment to overcome the average (expected) return, or suppress the loss. A higher ratio indicates a greater ability to overcome the average (expected) return. This indicator allows to easily analyse a relationship of risk and ability of overcoming of the average (expected) return of the investment.

ANALYSIS ON THE PRAGUE STOCK EXCHANGE

In the practical part of this article, the analysis of a relationship between risk and return of the investment on the Prague Stock Exchange is performed by means of the proposed stochastic procedure.

At first, the Prague Stock Exchange is briefly introduced. Secondly, the return and risk data of the selected stocks is collected, or calculated. Subsequently, a probability distribution of the returns of each stock is determined by the appropriate parametric statistical test. Further, the stock portfolios are made. Then the Monte Carlo simulation experiment with these portfolios are performed to analyze a declared relationship. The results are closely discussed.

Prague Stock Exchange

Prague Stock Exchange is one of two stock markets in the Czech Republic. It plays a significant role in the securities trading in this country. This stock market has been developing since 1993 when the capital market regenerates after the socialist regime collapse (Veselá 2011).

The Prague Stock Exchange consists of three markets – Prime, Standard and Start. The particular markets especially differ in demands on information duty from the issuers or other authorized persons. Prime and Standard are regulated markets. Start is unregulated. The strictest conditions hold to issue on the Prime market. Trading on these markets takes place in the automated trading system that is electronical system controlled by orders and quotes (Investování v ČR). Trading is mediated by authorized brokers, exchange members respectively (Princip obchodování 2017).

The traded products are stocks, bonds, exchange traded funds, warrants, investment certificates, futures and options (O finančních nástrojích 2017). Three indices on this stock market are calculated – PX, PX-TR and PX-GLOB. PX Index is an official price index of the Prague Stock Exchange, therefore it is a point of interest in this analysis. It is a price index with a weighted ratio of the most liquid stocks calculated in a real time (Index PX 2017). Three fifths of a market capitalization are held by the finance sector. Less than a fifth is created by the segment of electricity suppliers. Other components of the index are the segments of consumer goods, technology and telecommunication, basic industries and consumer services. It consists of 13 issues, which are listed in descending order of the share in the index: Erste Group Bank, Komerční banka, ČEZ, Moneta Money Bank, VIG, O2 C.R., Unipetrol, Pegas Nonwovens, Philip Morris ČR, CETV, Stock, Fortuna, Kofola ČS. PX-TR Index takes into account a dividend returns (Burzovní indexy – PX 2017). PX-GLOB is a price index including all traded stocks (Popis indexů 2017).

Stock data

As mentioned above, the basis of PX Index is formed by 13 issues. However, all stocks are not included in the analysis because some issues are on the Prague Stock Exchange shortly. Then the sufficient time series of returns are not available. This is a stock of Moneta Money Bank, VIG, Stock, Fortuna and Kofola ČS. Furthermore, Stock, Fortuna and Kofola has very small share in PX Index. Approximately 80 % of basis are preserved. But that is not so important for the analysis. Remaining 8 stocks partake of the analyses - Erste Group Bank (ERSTE as abbreviation for future use), Komerční banka (KB), ČEZ, O2 C.R.,

Unipetrol, Pegas Nonwovens (Pegas), Philip Morris ČR (PM) and CETV.

For an analysis complexity, a longer time period 2007-2017 is selected. This period includes various development of the stock prices. Therefore, it is divided into 3 subperiods – period of "drop", "rise" and "calm". The first period lasts from November 2007 to February 2009. During this financial crisis, prices went down sharply with a few smaller positive corrections. The period of "rise" is denoted as a phase from March 2009 to March 2011. This is very positive period on the capital markets. The stock prices significantly uprised. The period of "calm" is set from April 2011 to June 2017. This period is characteristic by a variable development of the stock prices. However, there are no huge falls or rises. These periods are chosen to reflect a major market development which is clearly displayed in the graph on this web page (Vývoj indexu PX 2017). Of course, it is not possible to foreclose that the price of some particular stocks can develop slightly different.

In order to have a sufficient number of observations, the return of stocks is calculated monthly from the closing prices available in the data bank of Patria online (Databanka Patria online 2017). Risk connected with the investment in the stock is measured by the average absolute negative deviation from the monthly returns. Both mentioned characteristics are calculated for each stock in each time period. Of course, the stock portfolio for the entire 10-year period must be also made for a subsequent analysis. So, these characteristics are also calculated for the entire period (see Table 1).

Table 1: Stocks' characteristics in the entire 10-year period

Stock	**Av. return [%]**	**Risk [%]**
ERSTE	*0.35*	9.90
KB	*0.43*	6.26
ČEZ	*-0.81*	4.86
O2 C.R.	*-0.12*	5.20
Unipetrol	*0.17*	4.72
Pegas	*0.55*	5.34
PM	*0.57*	4.62
CETV	*-0.63*	13.86

As we can see, the average monthly return is not so high related to the risk. Of course, it depends on a time period (see below). But this proportion is usual for the stocks. The values of both characteristics for all three periods are in the following table (Table 2).

Table 2: Stocks' characteristics in period of "drop", "rise" and "calm"

Stock	**Av. return [%]**			**Risk [%]**		
	Drop	***Rise***	***Calm***	***Drop***	***Rise***	***Calm***
ERSTE	*-10.03*	*6.68*	*0.46*	13.80	9.04	8.26
KB	*-4.95*	*4.48*	*0.23*	12.30	5.36	4.29
ČEZ	*-3.81*	*1.24*	*-0.85*	7.45	4.07	4.95
O2 C.R.	*-2.38*	*0.48*	*0.17*	4.93	4.10	5.33
Unipetrol	*-5.83*	*2.04*	*0.83*	7.86	4.61	3.19
Pegas	*-5.49*	*2.77*	*1.10*	11.59	4.69	3.38
PM	*-2.60*	*2.70*	*0.54*	7.31	5.63	3.18
CETV	*-12.51*	*6.54*	*-0.48*	19.41	14.34	11.54

Now, a probability distribution of each stock for each time period is determined via Anderson-Darling test because of smaller number of observations in period of "drop" and "calm". Nonparametric Kolmogorov-Smirnov statistical test can be applied to period of "calm" due to 75 observations. The list of probability distributions is in Table 3.

Table 3: Probability distribution of stocks' returns

Stock	**Drop**	**Rise**	**Calm**
ERSTE	*Weibull*	*Gumbel*	*Weibull*
KB	*Gumbel*	*Gumbel*	*Logistic*
ČEZ	*Gumbel*	*Lognormal*	*Beta*
O2 C.R.	*Weibull*	*Logistic*	*Student*
Unipetrol	*Logistic*	*Weibull*	*Logistic*
Pegas	*Logistic*	*Gumbel*	*Logistic*
PM	*Logistic*	*Weibull*	*Weibull*
CETV	*Normal*	*Gumbel*	*Logistic*

The results show the expected fact that stocks' returns are usually not normally distributed. It turns out that a determination of a proper known distribution can be sometimes difficult. Then the selected distribution describes the returns only approximately.

Portfolio selection

In the next step of the procedure, the portfolios for each period (including a 10-year) are made. We find the stock portfolio with minimum and maximum risk. Some other demands may be placed on the portfolio. If the minimum number of stocks in the portfolio will not be declared, then the portfolio will contain only one stock with minimum, or maximum risk. On the other side, too high number of stocks should be unnoticed and badly controllable for the investor. Based on my own experiences with the investment, the portfolio should consist of at least 3 stocks. This condition can be expressed via the maximum share of one stock in the portfolio. Moreover, this specification simplifies a formulation of the following mathematical model. So, the maximum share is determined as 40 %. Another conditions should be added, but it is not desirable for this analysis. For a portfolio selection, the following mathematical model minimizing/maximizing risk of the portfolio is formulated as

$$\min/\max \quad z = \sum_{i=1}^{8} r_i x_i$$

$$0 \le x_i \le 0.4 \qquad i = 1,2,...,8,$$

$$\sum_{i=1}^{8} x_i = 1$$

where $r_i\,(i = 1,2,...,8)$ is a risk of the i-th stock, $x_i\,(i = 1,2,...,8)$ denotes a share of the i-th stock in the portfolio. The values of i correspond upwardly to the stocks in the order from Table 1 ($i = 1$ for ERSTE, $i = 2$ for KB,…, $i = 8$ for CETV).

Due to the formulation of the risk objective function, it is a linear model of mathematical programming that can be

easily solved by well-known simplex method implemented in Lingo optimization software. Minimizing form of the model is solved for 3 periods (marked as *Pdrop_MIN* for period of "drop", *Prise_MIN* for period of "rise" and *Pcalm_MIN* for period of "calm") and also for a whole 10-year (marked as *Pwhole_MIN*). Similarly, the maximizing form of the model is used for a portfolio selection in all periods (marked analogically *Pdrop_MAX*, *Prise_MAX*, *Pcalm_MAX* and *Pwhole_MAX*). So, we have 8 mathematical models with different risk data from Table 1 and 2. As expected, the optimal solution for all models are found. Further, based on the model specification, it is possible to expect that the final portfolio for each case will be composed from 3 the least/most risky stocks. All portfolios are displayed in the following table (Table 4).

Table 4: Portfolios for all time periods

Portfolio	**Stock Share**	**Stock Share**	**Stock Share**
Pwhole_MIN	ČEZ 20 %	Unipetrol 40 %	PM 40 %
Pwhole_MAX	ERSTE 40 %	KB 20 %	CETV 40 %
Pdrop_MIN	ČEZ 40 %	O2 C.R. 20 %	PM 40 %
Pdrop_MAX	ERSTE 40 %	KB 20 %	CETV 40 %
Prise_MIN	ČEZ 40 %	O2 C.R. 40 %	Unipetrol 20 %
Prise_MAX	ERSTE 40 %	PM 20 %	CETV 40 %
Pcalm_MIN	Unipetrol 40 %	Pegas 20 %	PM 40 %
Pcalm_MAX	ERSTE 40 %	O2 C.R. 20 %	CETV 40 %

The values of average monthly return and risk of these portfolios are summarized in Table 5.

Table 5: Return and risk of all portfolios

Portfolio	**Av. return [%]**	**Risk [%]**
Pwhole_MIN	*0.14*	*4.71*
Pwhole_MAX	*-0.03*	*10.76*
Pdrop_MIN	*-2.76*	*6.39*
Pdrop_MAX	*-10.01*	*15.74*
Prise_MIN	*1.09*	*4.19*
Prise_MAX	*5.83*	*10.48*
Pcalm_MIN	*0.77*	*3.22*
Pcalm_MAX	*0.02*	*8.99*

Monte Carlo simulation experiment

In the next step, Monte Carlo simulation for each stock investment portfolio is performed. 1000 scenarios of return for each month for each stock of the portfolios are generated. Then we have 1000 scenarios of a development of stocks' returns for each portfolio related to a determined time period. The monthly returns are generated from a proper probability distribution (see Table 3) by means of Crystal Ball in MS Excel environment. As mentioned earlier, a continuous monthly investment is considered. It is 10 000 CZK per month throughout the entire period. The first investment is made at the beginning and the others at the end of each month (except the last month). This monthly invested amount is rather symbolic to operate with it better. On the Prague Stock Exchange, the trading takes place in the standardized units whose value can be higher. But the amount of investment is not essential for our analysis. At the end of each period, we have a final value of investment in each stock, in each portfolio as well. The expected final value of investment in stocks and portfolios is stated on the basis of average monthly returns. All important characteristics (return and probability ratio) of each portfolio and its individual components are digestedly shown in the appropriate tables below. The amounts of money are expressed in CZK.

Firstly, the investment situation for a 10-year time period is analyzed. The portfolio characteristics are displayed in the following table (Table 6).

Table 6: Portfolio characteristics for a 10-year period

Portfolio **(Probability)**	**Final Investment Return**	**Expected Investment Return**	**Ratio (Pr. Ratio)**
Pwhole_MIN (0.585)	1 573 824 *1 160 000* 35.67 %	1 321 683 *1 160 000* 13.94 %	2.57 (1.5)
Pwhole_MAX (0.575)	1 870 529 *1 160 000* 61.25 %	1 202 231 *1 160 000* 3.64 %	16.82 (9.67)

Regardless of distributions of stock returns, it is obvious that a riskier investment portfolio produces a higher return. It is a confirmation of a relationship "higher risk-higher return". Both portfolios overcome their expected return. Ratio "final return/expected return" is higher for a riskier stock portfolio. To take into account a distribution of the returns, a probability ratio is calculated. Then the values of this probability ratio confirm a validity of a positive relationship between risk and ability to overcome the average (expected) return.

And how do particular stocks of portfolios behave? See the characteristics of the stocks in the next table (Table 7).

Table 7: Stocks' characteristics for a 10-year period

Portfolio	**Stock (Prob.)**	**Final**	**Expected**	**Ratio (Pr. Ratio)**
Pwhole_MIN	ČEZ (0.495)	-29.89 %	-35.53 %	1.19 (0.59)
	Unipetrol (0.645)	41.04 %	10.63 %	3.96 (2.55)
	PM (0.565)	62.1 %	41.98 %	1.48 (0.84)
hol e_M	ERSTE (0.645)	139.05 %	23.52 %	5.91 (3.81)

	KB (0.505)	50.74 %/	29.86 %	1.7 (0.86)
	CETV (0.345)	-11.29 %	-29.35 %	2.6 (0.9)

The risk level of the stocks in a less risky portfolio (*Pwhole_MIN*) is very similar as we can see in Table 1. Stock ČEZ with the highest risk produces a loss as the only one. It is caused by its negative average monthly return for this period. It was not possible to expect that this stock (with the highest risk) automatically produces the highest return. On the contrary, the opposite was expected. However, a relation of the final return (loss) to the expected return (loss) is essential. The loss is approximately about 19 % less than its expected value. Stock PM has the lowest risk and also produces the highest return. It is expected on the basis of its the highest average monthly return. The (probability) ratio of PM is higher than the ratio of ČEZ stock. This fact confirms an invalidity of a relationship of the risk and ability to overcome expected return. A comparison of stocks Unipetrol and PM do not affect a validity of this relationship because Unipetrol has a higher risk and also a higher (probability) ratio. It means that a higher risk produces a greater ability of overcoming of the expected return in this case.

In terms of a riskier portfolio (*Pwhole_MAX*), the validity of a studied relationship at a level of particular stocks is also not generally confirmed. The stock with the second highest risk ERSTE produces the highest level of return and also (probability) ratio than the stock with the highest risk CETV. Analogically, the results for other time periods could be discussed. Due to a limited number of pages of the conference paper, only the most important aspects can be accented. Let us see the following table (Table 8) with both portfolio characteristics in the period of "drop".

Table 8: Portfolio characteristics for the period of "drop"

Portfolio **(Probability)**	**Final Investment Return**	**Expected Investment Return**	**Ratio (Pr. Ratio)**
Pdrop_MIN (0.455)	125 887 *160 000* -21.32 %	127 383 *160 000* -20.39%	0.96 (0.44)
Pdrop_MAX (0.45)	79 031 *160 000* -50.61 %	77 245 *160 000* -51.72 %	1.02 (0.46)

As expected in this period, both portfolios produce a loss. A riskier portfolio has a higher loss. It is expectable result according to the characteristics in Table 5. Without knowledge about distribution of the returns, it is possible to say that a higher risk does not ensure a higher return. More interesting view produce an inclusion of a probability analysis, or a ratio of expected and final return. A less risky portfolio (*Pdrop_MIN*) has a higher loss than expected. But a riskier portfolio (*Pdrop_MAX*) has a lower loss than expected. This portfolio has a greater ability to overcome its negative return. This fact is also confirmed by a probability ratio. In other word, a higher risk produces a higher chance to overcome average (expected) return.

As in the previous case, the necessary characteristics of each stock in both portfolios are available in the following table (Table 9).

Table 9: Stocks' characteristics for the period of "drop"

Port-folio	**Stock (Prob.)**	**Final**	**Expected**	**Ratio (Pr. Ratio)**
Pdrop_MIN	ČEZ (0.525)	-29.96 %	-26.39 %	0.88 (0.46)
	O2 C.R. (0.53)	-18.01 %	-17.61 %	0.98 (0.52)
	PM (0.415)	-19.47 %	-22.5 %	1.16 (0.48)
Pdrop_MAX	ERSTE (0.545)	-55.13 %	-51.24 %	0.93 (0.51)
	KB (0.565)	-33.26 %	-29.93 %	0.9 (0.51)
	CETV (0.415)	-61.44 %	-60.31 %	0.98 (0.41)

This table confirms the fact that in the period of negative development on the capital market the stocks with a higher risk produce a higher loss. And does a higher risk produce a higher chance to overcome expected (average) return in this time period? The answer is NO. The riskiest stocks of portfolios ČEZ and CETV have the lowest level of probability ratio. Their ability to overcome of their average (expected) return is in the lowest level.

Let us focus on the period of "rise". The portfolio characteristics are in Table 10.

Table 10: Portfolio characteristics for the period of "rise"

Portfolio **(Probability)**	**Final Investment Return**	**Expected Investment Return**	**Ratio (Pr. Ratio)**
Prise_MIN (0.44)	293 299 *250 000* 17.32 %	290 028 *250 000* 16.01%	1.08 (0.48)
Prise_MAX (0.37)	556 619 *250 000* 122.65 %	582 207 *250 000* 132.88%	0.92 (0.34)

On the one side, a positive relationship between risk and return is confirmed. This fact is based on a higher average return of a riskier stock portfolio. On the other side, the ability of this portfolio to overcome average (expected) return is lower than for a less risky portfolio that even produce a higher return than expected.

And what about a behavior of the individual components of these portfolios in the period of positive price development on the stock exchange?

As we can see in Table 11, the riskiest stock in a portfolio *Price_MIN* produces the highest return. For a riskier portfolio (*Prise_MAX*) this fact does not hold, because the riskiest stock is CETV. As you can see, the ability of overcoming of the average (expected) return is also not confirmed.

Table 11: Stocks' characteristics for the period "rise"

Port-folio	Stock (Prob.)	Final	Expected	Ratio (Pr. Ratio)
Prise_MIN	ČEZ (0.435)	20.21 %	17.84 %	1.13 (0.49)
	O2 C.R. (0.285)	6.03 %	6.49 %	0.93 (0.27)
	Unipetrol (0.43)	34.11 %	31.41 %	1.09 (0.47)
Prise_MAX	ERSTE (0.385)	161.75 %	157.81 %	1.03 (0.4)
	PM (0.445)	43.32 %/	44.01 %	0.98 (0.44)
	CETV (0.315)	123.21 %	152.39 %	0.81 (0.26)

Finally, the simulation (probability) analysis is performed to the period of "calm". Traditionally, the characteristics of both portfolios are listed first (see Table 12).

Table 12: Portfolio characteristics for the period of "calm"

***Portfolio* (Probability)**	**Final Investment Return**	**Expected Investment Return**	**Ratio (Pr. Ratio)**
Pcalm_MIN (0.325)	939 116 *750 000* 25.22 %	1 021 697 *750 000* 36.23 %	0.7 (0.23)
Pcalm_MAX (0.365)	754 479 *750 000* 0.6 %	770 262 *750 000* 2.7 %	0.22 (0.08)

A riskier portfolio does not produce a higher return. This fact is namely caused by the stock CETV that has a negative average return and very high risk. The less risky portfolio has also a greater chance to overcome an expected return according to the (probability) ratio.
The characteristics of particular stocks are displayed in the following table (Table 13).

Table 13: Stocks' characteristics for the period of "calm"

Port-folio	Stock (Prob.)	Final	Expected	Ratio (Pr. Ratio)
Pcalm_MIN	Unipetrol (0.27)	18.18 %	39.10 %	0.47 (0.13)
	Pegas (0.335)	44.78 %	55.83 %	0.8 (0.27)
	PM (0.4)	22.47 %	23.54 %	0.95 (0.38)
Pcalm_MAX	ERSTE (0.435)	23.58 %	19.64 %	1.2 (0.52)
	O2 C.R. (0.365)	3.88 %/	6.74 %	0.58 (0.21)
	CETV (0.275)	-24.02 %	-16.26 %	0.68 (0.19)

Most stocks produce a lower return than expected. As expected, a higher risk does not mean a higher final return (see for instance the riskiest stock CETV). The least risky stock PM in the portfolio *Pcalm_MIN* has the greatest ability for overcoming of an average (expected) return. The relationship "higher risk-greater ability to overcome an average (expected) return" is also negated. This relationship does also not hold for a highly risky portfolio.

Results discussion

This empirical analysis based on the proposed stochastic procedure proves that a higher risk may not lead to a higher return of the investment. This result is expected according to the main statistical characteristics (mean, average absolute negative deviation) in the selected time periods. For instance, it is expectable that in the period of "drop" a riskier stock portfolio will produce higher losses, as well in the whole 10-year period. Then it is obvious that this type of relation will hold rather in the period of positive prices development on the capital market, or in stable conditions. But this conclusion does not hold for the individual stocks. Even in one period, it is not possible to confirm a general positive relationship of the risk and return.
I think that the most significant contribution of the stochastic analysis is a study of relationship of risk and ability to overcome an average (expected) return. For investment portfolios, this relationship is confirmed for longer 10-year period and period of "drop" where a riskier portfolio can mitigate the losses better. This relationship cannot be possible to confirm for a set of particular stocks in any time period.
Finally, we can say that any studied relationship does not generally hold. It depends on the time period, or prices development (and related returns distribution) on the capital market, and as well as the set of selected stocks (and on which capital market are traded), or the investment portfolios. Of course, a choice of the capital market itself can affect an analysis. But as mentioned above, some specific patterns for some time period on the Prague Stock Exchange hold. This result can be helpful in the investment decision making process. The (potential) investor knows which stocks would be more or less suitable in particular market situation. In addition, the investor has not to apply a chosen investment strategy. This analysis was also performed for a one-shot investment and the results were very similar. Furthermore, the proposed procedure can be also applied to another investment instruments (e.g. open unit trusts).
The last note concerns a nature of the obtained results. The analysis is based on the historical data. History may not repeat. A potential investor should treat the results with caution. Monte Carlo simulation is based on a knowledge about the returns distributions which are sometimes determined with difficulty. Moreover, a generation of scenarios of the returns is a random process. Scenarios can vary greatly. But a sufficient number of scenarios uprates the results that we still should understand as average (approximate).

CONCLUSION

This article deals with an analysis of a relationship of the risk (volatility) and return of stocks, or stock investment portfolio which is very interesting investment issue in a long run. For this purpose, the stochastic procedure is proposed. This approach uses a principle of linear mathematical programming and mainly Monte Carlo simulation. The simulation experiments are performed for the selected stocks and stock investment portfolios in different time periods on the Prague Stock Exchange. All questions specified in the Introduction are satisfactorily answered. It has been shown that a relationship of the risk and return is ambiguous and must be handled with care when investing. The proposed methodical concept using a simulation approach has proved successful in analyzing.

For the future research, the concept of conditional volatility estimated from some econometric model could be included as in (Guo et al. 2009). This concept would take into account a risk (volatility) variability over time. Greater emphasis may be placed on choosing a time period. For instance, Fiore and Saha (2005) identified a different investment behavior in summer months which would be interesting to analyze separately on the Czech Stock Exchange.

Acknowledgements

The research project was supported by Grant No. IGA F4/57/2017 of the Internal Grant Agency, Faculty of Informatics and Statistics, University of Economics, Prague.

REFERENCES

Fama, E. F.; and J. MacBeth. 1973. "Risk Return and Equilibrium: Empirical Tests." *Journal of Political Economy 81*, No. 3 (May-Jun), 607-636.

Fama, E. F.; and K. R. French. 1992. "Cross-section of Expected Stock Returns." *Journal of Finance 47*, No. 2 (Jun), 427-463.

Fama, E.F; and K. R. French. 1996. "Multifactor Explanations of Asset Pricing Anomalies." *Journal of Finance 51*, No. 1 (Mar), 55-84.

Fiore, C; and A. Saha. 2015. "A Tale of Two Anomalies: Higher Returns of Low-Risk Stocks and Return Seasonality." *The Financial Review 50*, No. 2 (Apr), 257-273.

Guo, H.; R. Savickas; Z. Wang; and J. Yang. 2009. "Is the Value Premium a Proxy for Time-Varying Investment Opportunities: Some Time Series Evidence." *Journal of Financial and Quantitative Analysis 44*, No. 4 (Feb), 133-154.

Jegadeesh, N. 1992. "Does Market Risk Really Explain the Size Effect?" *Journal of Financial and Quantitative Analysis 10*, No. 3 (Sep), 337-351.

Koutmos, G.; C. Negakis; and P. Theodossiou, P. 1993. "Stochastic Behaviour of the Athens Stock Exchange." *Applied Financial Economics 3*, No. 2 (Jun), 119-126.

Markowitz, H. (1959). *Portfolio Selection: Efficient Diversification of Investments*. John Wiley & Sons, Inc., New York.

Markowitz, H. 1952. "Portfolio selection." *Journal of Finance 7*, No. 1 (Mar), 77-91.

Omet, G.; M. Khasawneh; and J. Khasawneh. 2002. "Efficiency Tests and Volatility Effects: Evidence from the Jordanian Stock Market." *Applied Economics Letter 9*, No. 12, 817-821.

Pamane, K; and A. E. Vikpossi. 2014. "An Analysis of the Relationship between Risk and Expected Return in the BRVM Stock Exchange: Test of the CAPM." *Research in World Economy 5*, No. 1 (Mar), 13-28.

Salman, F. 2002. "Risk-Return-Volume Relationship in an Emerging Stock Market." *Applied Economics Letters 9*, No. 8, 549-552.

Sharpe, W. F. 1964. "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk." *Journal of Finance 19*, No. 3 (Sep), 425-442.

Veselá, J. 2005. *Burzy a burzovní obchody – výchozí texty ke studiu*. Oeconomica, Praha.

WEB REFERENCES

Burzovní indexy - PX [online], available at: *https://www.pse.cz/indexy/hodnoty-indexu/segmentace/?ID_NOTATION=325088&ISIN=XC0009698371*, [cit. 18-07-2017].

Databanka Patria online [online], available at: *https://www.patria.cz/kurzy/vyzkum/databanka.html*, [cit. 01-07-2017]

Index PX [online], available at: *https://www.pse.cz/indexy/popis-indexu/index-px/*, [cit. 18-07-2017].

Investování v ČR [online], available at: *http://www.akcie.cz/radce-investora/investice-zaklady/cz/*, [cit. 18-07-2017].

O finančních nástrojích [online], available at: *https://www.pse.cz/pruvodce-burzou/o-financnich-nastrojich/*, [cit. 18-07-2017].

Popis indexů [online], available at: *https://www.pse.cz/indexy/popis-indexu/*, [cit. 18-07-2017].

Princip obchodování [online], available at: *https://www.pse.cz/pruvodce-burzou/uvod-do-burzovniho-sveta/princip-obchodovani/*, [cit. 18-07-2017].

Vývoj indexu PX [online], available at: *https://www.pse.cz/indexy/hodnoty-indexu/graf/?ID_NOTATION=325088&ISIN=XC0009698371&TIMESEL=1&PERIOD=10Y&PERIOD_FROM=&PERIOD_TO=&CHART_TYPE=line&BENCHMARK=&INDICATOR=&MAV=&TURNOVER=0&EVENT=&SCALE=lin*, [cit. 28-07-2017]

AUTHOR BIOGRAPHY

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. He primarily deals with the questions of multiple criteria decision making, fuzzy optimization problems and stochastic processes in the field of the capital market. He has publications and submissions in international journals and conferences. He also teaches at Czech Technical University, Faculty of Nuclear Sciences and Physical Engineering.

Optimal Dating of Cycles in Financial Time Series

Konrad Kapp
COEF, Nelson Mandela University
Port Elizabeth 6031, South Africa
email: konrad.p.kapp@gmail.com

Igor Litvine
COEF, Nelson Mandela University
Port Elizabeth 6031, South Africa
email: igor.litvine@nmmu.ac.za

KEYWORDS
Time series, financial cycles, large data, optimisation

ABSTRACT

The study of cycles in the context of economic time series has been active for many decades, if not centuries; however, it was only in recent decades that more formal approaches for identifying cycles have been developed. In (Litvine (2016)) a new approach is proposed for dating cycles in financial time series for purposes of optimising buy-sell strategies. In this approach, cycle dating is presented as an optimisation problem. A method was also introduced for solving this problem, known as the hierarchical method (using full evaluation, or HR-FE2). However, this method may be impractical for large data sets as it may require unacceptably long computation time. In this paper, a novel procedure, known as buy-sell adapted extrema importance identity sequence retrieval (BSA-EIISR), is introduced as an alternative to the hierarchical method in order to alleviate the problem of long computation times. It was found that BSA-EIISR outperforms HR-FE2 by a large margin in terms of computation time. In many cases, BSA-EIISR also outperforms HR-FE2 in terms of objective function value, but it was found that there are still certain cases where HR-FE2 outperforms BSA-EIISR in terms of objective function value.

Introduction

The Merriam-Webster dictionary defines the word "cycle" as "a repeating series of events or actions". The keyword here is "repeating": this implies that there is an identifiable regularity in these events, which in turn implies that inference may be made with regards to future events with some degree of confidence.

There are many fields and professions where predictive power with regard to future events is important, such as medicine, engineering and city planning, for example. In economics and finance, policymakers and portfolio managers can make better decisions when informed of highly probable events that might occur in the future. This is one of the reasons why business cycles have been of interest to economists for nearly two centuries - business cycles often occur with surprising regularity, allowing economists to anticipate and act on these regular occurrences in the economy.

When one attempts to discover cycles in a time series, one generally wants to find segments where the time series has an overall upward trend, which is followed by a segment where the time series has an overall downward trend. This is equivalent to finding alternating peaks and troughs that separate the upward and downward trend segments (also known as turning points). The process of finding these peaks and troughs will be referred to as cycle dating. An example of this is shown in figure 1.1. Note, that we will refer to these troughs and peaks as buy-sell points.

In (Litvine (2016)) a methodology was proposed to date cycles by means of maximising an objective function. This approach allows a numerical measure of the effectiveness of a given dating process, which makes comparisons between different methods easier. The optimisation problem is presented as follows:

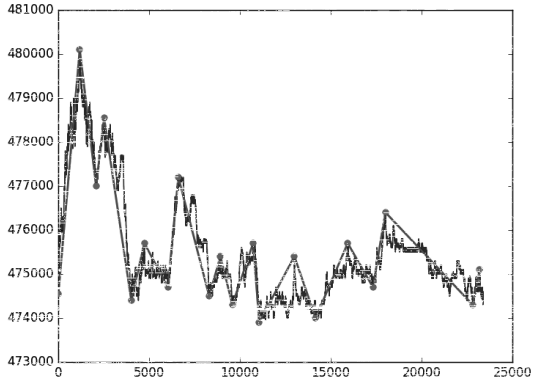


Figure 1: An example of dated cycles, using the approach proposed by Litvine (2016), applied to intraday share price data (per second). The peaks and troughs ("sell" and "buy" points, respectively) that the algorithm extracted from the time series are shown as red dots. 20 buy-sell points are found here, equivalent to nine and a half full cycles.

Maximise
over $x_i^*, i = 1, 2, ..., K$ (where K is even)

$$f_{profit}(X^*) = \sum_{i=1}^{K/2} (y(x_{2i}^*) - y(x_{2i-1}^*)) \quad (1)$$

subject to

$$x_i^* < x_j^*, \ i < j,$$

where $y(x)$ is the asset price at time x (or alternatively, *time index* x), and X^* is a set of the time indices of buy-sell points, i.e. $X^* = \{x_1^*, x_2^*, \ldots, x_K^*\}$.
Note that, as shown in figure 2, the cycle dating procedure can also represent a time series relatively well with remarkably few of the original observations. This makes sense, since maximising equation (1) implies that we obtain peaks and troughs that are separated generally by larger distances. These peaks and troughs are therefore more likely to be more prominent ones, and thus tend to preserve the general shape of the time

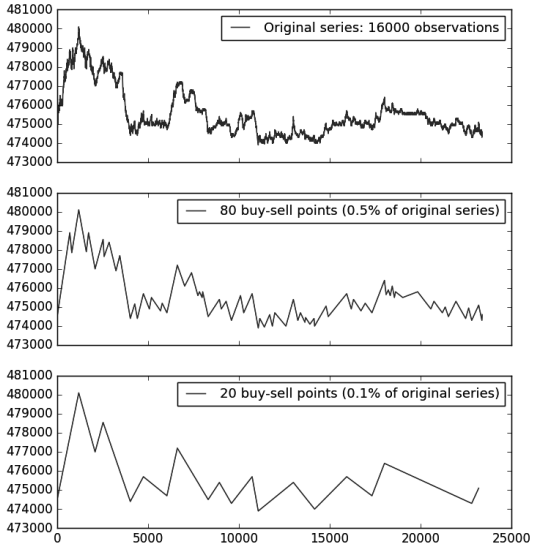


Figure 2: An example of how the cycle dating approach can also effectively represent a time series by much less observations than the original.

series. It may also be specified how many buy-sell points should be found (given by K in equation (1)): in the case of more buy-sell points, more information from the original time series is preserved, and for less buy-sell points, less is preserved, implying that only the most prominent peaks and troughs are obtained.

Applications

The cycles that are found via this technique are cycles that are defined by times, where it was optimal to buy or sell a given asset. These cycles may therefore be used to analyse the frequency and price levels of optimal times to buy and sell, as well as a benchmark for traders' historical performance by comparing their obtained profit with potential profit.
As shown in figure 2 and discussed thereafter, this method may also be used as a time series compression technique, which may reduce an excessive amount of data to acceptable levels for fur-

ther analysis. This methodology may therefore be considered as a possible alternative to existing time series compression techniques, such as those discussed by Pratt and Fink (2002) and Keogh et al. (2004), for example.

Problem statement

The approach proposed in (Litvine (2016)) is a divide-and-conquer technique, known as the Hierarchical Method (HR), to optimise the objective function given by (1). However, it was not known whether this method provides an optimal solution to the objective function, or whether there exists faster, more efficient methods.

Time to obtain a solution is crucial when considering that some practitioners may wish to apply these methods to high sampling frequency time series of asset prices. These time series often contain tens of thousands of observations. Consider per-second data of share prices for a single trading day: such a time series may contain at least 27 000 observations if share prices are sampled at a period of a second or less. Applying methods that are inefficient may therefore prevent practitioners and researchers from doing analysis on these time series data sets in reasonable time.

For the reasons above, we therefore present a novel method, the buy-sell adapted Extrema importance identity sequence retrieval (BSA-EIISR), as an alternative to the Hierarchical (HR) method.

BSA-EIISR

In Wu and Huang (2009), a method known as Extrema importance identity sequence retrieval, is proposed for extraction of prominent peaks and troughs in a given time series. We adapted this method for maximisation of the objective function given by (1). This adaptation is described in the current section.

In the approach presented here, buy and sell points are found separately: first all the buy points, then all the sell points. Alternatively, all the sell points, then all the buy points may be found. For simplicity, we will focus only on finding all buy points first, then sell points, but note that it may also be done the other way around.

We break the optimisation problem represented by the profit objective function into two sub-problems. In the first sub-problem, the best troughs to use as buy points are found. Since the objective function may also be written as follows,

$$f_{profit}(X^*) = \sum_{i=1}^{K/2} y(x^*_{2i}) - \sum_{i=1}^{K/2} y(x^*_{2i-1}) \qquad (2)$$

the optimal troughs are in the solution to the second summation in equation (2):

Maximise
over x^*_{2i-1}, for $i = 1, 2, ..., \frac{K}{2}$

$$f_{troughs}(X^*) = -\sum_{i=1}^{K/2} y(x^*_{2i-1}), \qquad (3)$$

subject to $x_i < x_j$, $i < j$.

Once the optimal troughs have been found, all sell points are found directly by extracting the the highest peak between each pair of buy points:

Maximise
over x^*_{2i}, for $i = 1, 2, ..., \frac{K}{2}$

$$f_{peaks}(X^*) = \sum_{i=1}^{K/2} y(x^*_{2i}) \qquad (4)$$

subject to

$$\begin{cases} x^*_{2i-1} < x^*_{2i} < x^*_{2i+1} & \text{if } i < \frac{K}{2} \\ x^*_{2i-1} < x^*_{2i} \leq Y_{size} & \text{if } i = \frac{K}{2} \end{cases}$$

where all x^*_j, for j odd (the troughs found previously), are given constants.

The algorithm presented here to solve the optimisation problem, using this general approach, will mainly be concerned with optimising the first sub-problem (finding optimal troughs), since the second sub-problem is fairly trivial to solve once the first has been solved. Figure 3 demonstrates how the second sub-problem is solved.

Sub-problem 1 (equation 3)

The method for solving sub-problem 1 is based on the Extrema importance identity sequence retrieval (EIISR) algorithm invented by Wu and

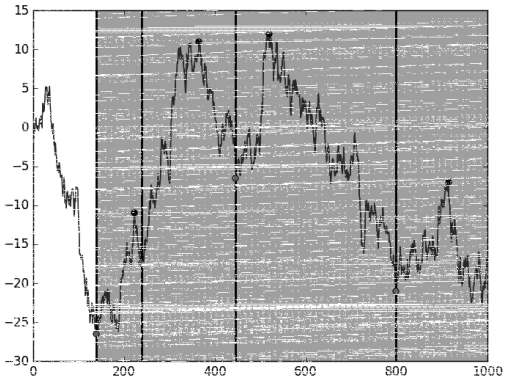


Figure 3: The points extracted in sub-problem (4) (black dots) are the highest points in each of the yellow shaded segments, demarcated by the vertical black lines. Each segment is defined by two successive troughs (red dots), found in sub-problem (3), and the last segment by the last trough and the end of the time series. Sub-problem two is therefore fairly trivial to solve.

Huang (2009). This algorithm starts by finding all points that are lower than points on either side of them in the time series (local troughs) and putting their time indices in a set V_1. All the points in the set V_1 are local troughs. All troughs in the set V_1 are then checked, and those troughs that are lower than both observations two observations away either side of it in the time series, are put into set V_2. The same is repeated for set V_2: all troughs in V_2 that are lower than points three observations away in the time series, on either side of it, are put into set V_3. This continues until the stopping criterion is met, which will be discussed shortly.

Since the troughs in the set V_i are the only points that will be investigated after iteration i, it is clear that we have a decreasing sequence of sets

$$\forall i \in \mathbb{N} : V_i \supseteq V_{i+1}. \quad (5)$$

V_i will be referred to as the *clearance set* after iteration i, and the sequence in equation (5) as the sequence of clearance sets. Note that the clearance set V_i contains all those troughs that are lowest in a window of radius i. Note that all troughs x_j in V_i satisfy

$$y(x_j) \leq y(x_{j\pm k}), \text{ where } k = 1, 2, ..., i \quad (6)$$

The main purpose of sequentially finding troughs that are lowest in wider and wider windows is to make the set from which troughs are chosen for the final solution, as small as possible. This sequential process makes it easy to determine which window size yields a set of troughs that is approximately equal in size to the required number of troughs, $\frac{K}{2}$. If the number of troughs in the final set is not exactly equal to $\frac{K}{2}$, we evaluate all solutions corresponding to all combinations of $\frac{K}{2}$ troughs extracted out of the two final sets of troughs (with the highest point between each, selected as sell points) to find the best solution. This is explained in more detail below. If the final set contains exactly $\frac{K}{2}$ troughs, no further selection is needed, thus, the final set then consists of the optimal troughs.

Stopping conditions

Assume that the method for sub-problem 1 stops after F iterations (our final clearance set is therefore V_F). If $|V_i| \leq \frac{K}{2}$, all the troughs in the clearance set at iteration i and after will be unconditionally included when selecting the $\frac{K}{2}$ most prominent troughs. Also, the troughs in V_i are the only points that the algorithm will check at least once in subsequent iterations (after iteration i). It therefore follows that if $|V_i| \leq \frac{K}{2}$, then no more information on which troughs to extract could be obtained by continuing the optimisation process, since there is no need to distinguish between troughs that will definitely be extracted anyway. The first part of the optimisation procedure is therefore terminated when $|V_i| \leq \frac{K}{2}$. If $|V_F| = \frac{K}{2}$, all the troughs in V_F are taken as the troughs in the final solution. Otherwise, if $|V_F| < \frac{K}{2}$, then, since the algorithm was stopped as soon as $|V_i| \leq \frac{K}{2}$, it follows that $|V_{F-1}| > \frac{K}{2}$. Therefore we will also choose from the troughs in V_{F-1}, since V_{F-1} is then the smallest set that contains at least all the troughs that will be extracted (because V_F does not contain enough troughs). However, all the troughs con-

tained in V_F must be kept because the troughs with the highest prominence are contained in this set; therefore $N_{extract} = \frac{K}{2} - |V_F|$ troughs must be selected from a subset of the set V_{F-1} and added to the troughs in V_F for a full solution of $\frac{K}{2}$ troughs. Since all the troughs in V_F are in the final solution, and $V_F \subset V_{F-1}$, the subset of V_{F-1} that $N_{extract}$ troughs is extracted from, is defined as

$$V_{extract} = V_{F-1} \setminus V_F$$

Each combination of size $N_{extract}$ extracted from $V_{extract}$ is combined with all the troughs in V_F to form a full solution for each combination. There will therefore be $\binom{|V_{extract}|}{N_{extract}}$ different solutions that must be evaluated to obtain a final solution (the solution with the highest fitness out of the $\binom{|V_{extract}|}{N_{extract}}$ solutions generated from the troughs in V_F and the combinations out of $V_{extract}$).
If the final iteration occurs after sufficient iterations F, we can generally assume that $V_{extract}$ will be a relatively small set. Therefore the different $\binom{|V_{extract}|}{N_{extract}}$ combinations of points to be tested will not be very large and can all be tested without a great sacrifice in computation time. Note that, as stated already, this assumes that the number of iterations F is sufficiently large. If F is relatively small (such as when K is relatively large compared to the series size) then BSA-EIISR may take very long to finish computation due to the large number of combinations that must be evaluated. This, in our view is the main shortcoming of the algorithm.

Algorithm 1 BSA-EIISR procedure for extracting optimal troughs

procedure ExtractTroughs(V_{final}, $V_{final-1}$, K)
  if $|V_{final}| < \frac{K}{2}$ **then**
    $V_{extract} \leftarrow V_{final-1} \setminus V_{final}$
    $N_{extract} \leftarrow \frac{K}{2} - |V_{final}|$
    $\mathcal{P}_{N_{extract}} \leftarrow$ AllCombinations($V_{extract}$, $N_{extract}$)
    for each $V_{comb} \in \mathcal{P}_{N_{extract}}$ **do**
      $V_{cand} \leftarrow V_{final} \cup V_{comb}$
      $X \leftarrow$ FindPeaks(V_{cand})
      if $f_{profit}(X) > f_{profit}(X_{best})$ **then**
        $X_{best} \leftarrow X$
        $V_{best} \leftarrow V_{cand}$
      end if
    end for
    return V_{best}
  else
    return V_{final}
  end if
end procedure

Results

In this section, we present results from running the two algorithms on a real time series data set. We will first describe the data set, after which we explain the methodology and the graphical methods that were utilised to communicate results. Note that in all cases, 100 buy-sell points were found in each time series (i.e. $K = 100$).

Data

The data set that the algorithms were applied to is that of intraday Microsoft (MSFT) share prices, from the 1st of April 2015 to 31 May 2015. The data is sampled at intervals of one second, which implies that there are 23401 observations for a trading day that starts at 9:30AM and ends at 4:00PM.
Note that neither the sampled, nor the full-sized (23401 observations) time series never spanned over more than one trading day.

Graphical Comparisons

We used two kinds of scatter plots to present results: an ordinary scatter plot of fitness vs. computation time, and a scatter plot of relative fitness and computation times between BSA-EIISR and the Hierarchical method (which we will refer to from now on as HR). The latter plot made it easy to see how many times one method improved on another in terms of computation time and fitness. This plot we will refer to as a *baseline plot*.
The baseline plot has, on the x-axis, where computation time is plotted, the values $\frac{t(E_i)}{t(H_i)}$, where $t(Z_i)$ is the time taken by the method represented by Z on the ith time series. Here, E represents BSA-EIISR and H represents HR. Since

$t(H_i)$ is the denominator, the lower HR's computation time, and the higher EIISR's computation time was, the higher this value is. Therefore, a higher value implies favourable computation time for HR. If it is above 1, then HR was faster than EIISR, and if it is below 1, EIISR was faster than HR.

On the y-axis, where fitness is normally plotted, the value $\frac{g(H_i)}{g(E_i)}$ is plotted, where $g(Z_i)$ represents the fitness which results from running algorithm Z on the ith time series. Therefore, if HR has higher fitness, and BSA-EIISR lower fitness, then this value is higher, and vice versa. Once again, a higher value implies favourable performance by HR, this time in terms of fitness. A value of 1 also implies equal fitness, as is the case with computation time.

The baseline plot may therefore be divided into quadrants: a value plotted in the top right represents an instance where HR outperformed BSA-EIISR in both computation time and profit fitness. On the other hand, a value plotted in the bottom left represents an instance where BSA-EIISR outperformed HR in both criteria. The other two quadrants represent cases where one method outperformed another in terms of only one criteria.

Computation time and fitness

Results on computation time and fitness of the two methods are now presented.

Scatter and baseline plots are shown in figure 6, where the algorithms were applied to 42 time series, each of size 8000. The scatter plot, shown in figure 4, clearly shows that BSA-EIISR was significantly faster (at least five times faster) than HR in all cases. However, it is not easy to see which algorithm had higher fitness for which time series on this plot. The baseline plot, in figure 5, shows that BSA-EIISR outperformed HR in terms of fitness in most cases, while HR did, at times, marginally outperformed BSA-EIISR in terms of fitness.

In figure 9, scatter and baseline plots are shown for when the methods are applied to a full trading day, which consists of 23401 time series observations. Here, the baseline plot clearly shows that HR performs mostly better than BSA-EIISR

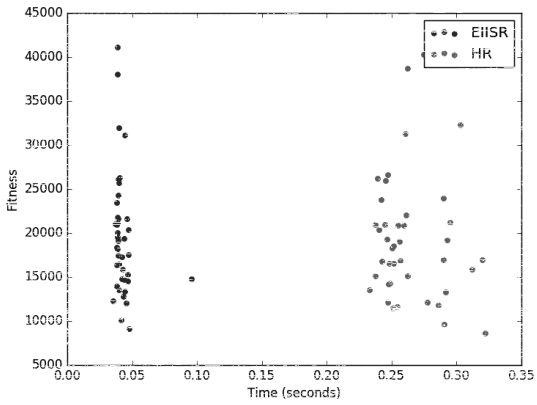


Figure 4: Scatter plot

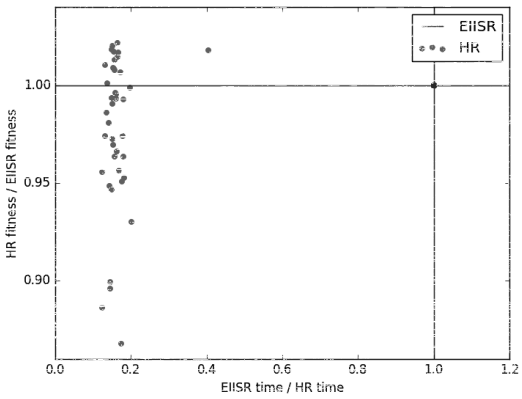


Figure 5: Baseline plot

Figure 6: Scatter and baseline plots for series size 8000

in terms of fitness, while still performing slower. This relative increase in HR's fitness is most likely due to the fact that both at the start and the end of a trading day, volatility is usually much higher than at other times (Webb and Smith (1994)). This causes situations as that shown in figure 10, to arise often. As can be seen in figure 10, very large up/down movements happened in the first ten seconds of trading. The red dots show where HR managed to identify prominent peaks and troughs. Since BSA-EIISR uses the rule in equation (6) for including prominent troughs in clear-

ance sets, the trough with the red dot in figure 10 would not have been included in later clearance sets and was thus ignored by BSA-EIISR. This is due to the fact that only a few observation to the right of this trough, another trough is lower. This situation is more likely to occur when we have violent up and down movements, which we are more likely to observe at the start and end of a trading day due to higher volatility. Since a full trading day will always include these situations, and the sampled, smaller time series will most likely not include them (since they generally occur over such a short period of time at the start of day), HR often got higher fitness values than BSA-EIISR since it was able to identify these prominent peaks and troughs, while BSA-EIISR did not, for the reasons explained above.

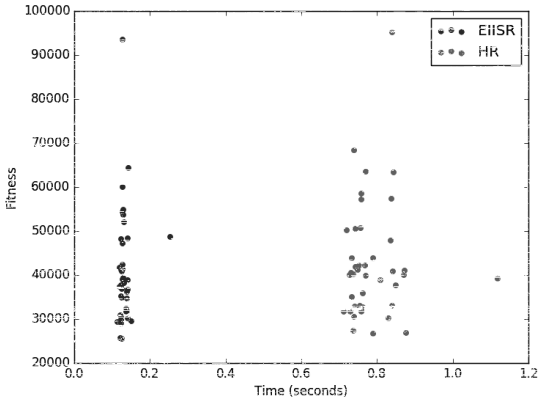


Figure 7: Scatter plot

Scaling of computation time

Some results on how computation time scales with series size are shown for both methods in figures 11 and 12. Each plotted value is the average of computation times for 42 time series of the corresponding size.
It is clear in both figures that there is a strong evidence of a linear relationship between series size and computation time. Indeed, the ratio for the computation time at 23401 observations over the computation time at 5000 observations is

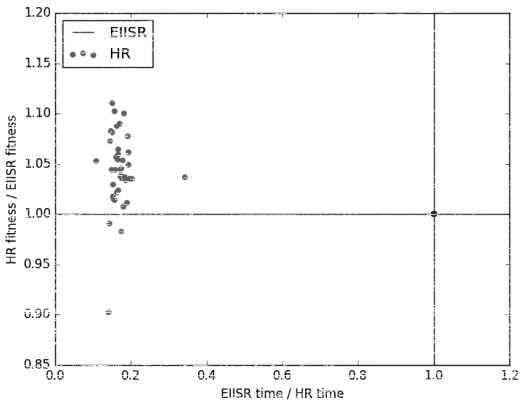


Figure 8: Baseline plot

Figure 9: Scatter and baseline plots for series size 23401 (full trading day)

around 4.5 for both methods, meaning that each method's computation time increased by around 4.5 times from 5000 observations to 23401. The slightly larger value for computation time at series size 2000 for BSA-EIISR is likely due to the fact that K was quite high relative to the series size. This would have made it more likely that not exactly $\frac{K}{2}$ troughs were in the final clearance set at the end of sub-problem 1, which would have required some extra computations to find the optimal set of $\frac{K}{2}$ troughs, explaining the longer computation time.
Considering the discussion above, it seems likely

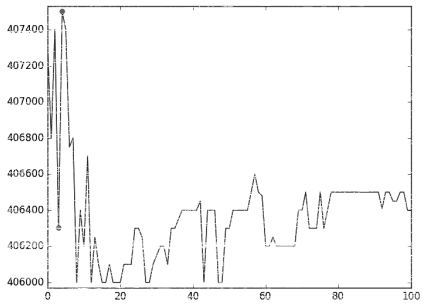


Figure 10: Share prices in first 100 seconds of trading day

that when applying these methods to significantly larger time series sizes than 23401, computation time would still not be an issue, with BSA-EIISR most likely keeping its superior computation speed relative to HR's.

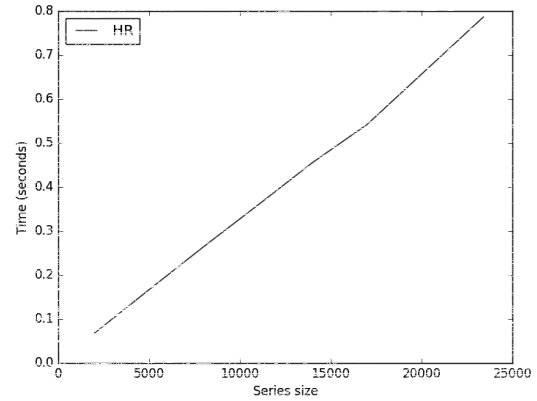


Figure 11: Computation time vs series size for HR

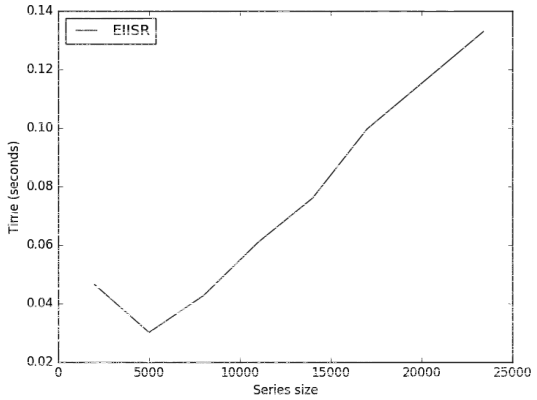


Figure 12: Computation time vs series size for EIISR

Conclusion

In this paper, we introduced a new method, BSA-EIISR, which may be used to find cycles and compress time series data sets. The motivation for this method was the long computation times that the existing methods took to solve the optimisation problem. We found that BSA-EIISR does indeed take much less time to find an optimal solution, and in many cases also outperforms HR-FE2 in terms of objective function value. However, it was found that there are certain cases where HR-FE2 still outperforms BSA-EIISR in terms of objective function value. Future directions for research could focus on statistical analysis of the set of buy-sell points extracted from the time series. The effect of the objective function value on this analysis could also be investigated. On the level of the actual programming of the algorithms, parallel implementations of these methods were not attempted here. Therefore, future research could focus on the viability of these methods for parallel implementation, either using GPU hardware or multi-core CPUs.

REFERENCES

Keogh E.; Chu S.; Hart D.; and Pazzani M., 2004. *Segmenting time series: a survey and novel approach.* In M. Last; A. Kandel; and H. Bunke (Eds.), *Data mining in Time Series Databases*, World Scientific Publishing Company. ISBN 978-981-4486-54-5, 1–21.

Litvine I., 2016. *Economic and Financial Cycles in South Africa.* Ph.D. thesis, University of Lorraine. Nancy, France.

Pratt K.B. and Fink E., 2002. *Search for patterns in compressed time series. International Journal of Image and Graphics*, 02, no. 01, 89–106. doi:10.1142/S0219467802000482.

Webb R.I. and Smith D.G., 1994. *The effect of market opening and closing on the volatility of eurodollar futures prices. Journal of Futures Markets*, 14, no. 1, 51–78. ISSN 1096-9934. doi:10.1002/fut.3990140106. URL `http://dx.doi.org/10.1002/fut.3990140106`.

Wu X. and Huang D., 2009. *Representing financial time series based on important extrema points.* In *Intelligent Information Technology Application, 2009. IITA 2009. Third International Symposium on Intelligent Information Technology Applications.* vol. 1, 501–504.

MINING PATTERNS IN FINANCIAL TIME SERIES USING DYNAMIC TIME WARPING ALGORITHM

Kristina Šutienė, Audrius Kabašinskas, Eimutis Valakevičius

Department of Mathematical Modelling
Kaunas University of Technology
Studentų 50, 51368 Kaunas, Lithuania
E-mail: kristina.sutiene@ktu.lt

Roland Reichardt

University of Applied Sciences Düsseldorf
Münsterstraße 156
40476 Düsseldorf, Germany
E-mail: roland.reichardt@hs-duesseldorf.de

KEYWORDS
Dynamic time warping, shape-based similarity, hierarchical clustering, pension funds.

ABSTRACT

In the paper, the application of dynamic time warping algorithm for financial time series is considered. The overview of published researches has shown that this algorithm in financial area is not widely applicable. Over the past two decades, the financial markets have become increasingly difficult to analyse using traditional methods, so there is a need of developing new techniques or adapting methods from the other areas that would let to explore financial time series. Therefore, the application of dynamic time warping algorithm is investigated applying it for time series of pension fund market and well-known indices. The possible added value of this algorithm is presented in three cases: subsequence matching, bivariate alignment, multivariate pattern matching and clustering. Notwithstanding the application to pension funds data only, this work offers valuable insights into dynamic time warping application in finance area.

INTRODUCTION

It is a widely held view that by analysing financial time series, one can observe trends, cyclical fluctuation, seasonal effects, or volatility and hence to forecast the movement of financial variable. Although these properties of financial data have been extensively studied for a reasonable period of time, new prospects have been opened up by increasing availability of data sets and the application of computer-intensive methods for their analysis. The analysis of huge amount of data has generated new challenges, since the dynamics of financial variable exhibits some quite nontrivial statistical features and nonstationarity (Cont 2001; Sewell 2011). Classical financial models that typically assume homoskedasticity and normality cannot describe properties, such as heavy tails, volatility clustering, or nonlinear dependence observed in empirical finance, which sometimes referred to as "stylized facts". A recent study (Champagnat et al. 2013) presented the importance of an adequate modelling of the heavy-tail behaviour of financial variable for a proper risk estimation. Similarly, a number of studies (Chavez-Demoulin et al. 2014; Su and Hung 2011) have reported specialized methods or models for risk estimation when a financial variable exhibits a set of stylized empirical facts. Numerous studies (Epaphra 2017; Banumathy and Azhagaiah, 2015) have attempted to investigate the volatility in financial time series by creating a model for its pattern, while the others (Shah and Roberts 2013; Zhao et al. 2016) have examined the relationship among financial time series, which in most cases is nonlinear, in order to develop multivariate models or to improve forecasts by introducing dependencies or causal relations. While developing multivariate models or analysing big data cases, the clustering of time series becomes relevant (Guam and Jiang 2007; Durante et al. 2014). Therefore, there is a large volume of published studies that investigate the time series similarity measured in different ways. If financial time series are significantly varying in time (e.g., asset returns, index values), matching of them using improved alignment based on shape makes sense, thus Dynamic Time Warping (DTW) may be used as a similarity measure.

The idea of DTW is to find the optimal alignment between two series under certain constraints. DTW was introduced in 1960, but it became popular only after twenty years when it was successfully applied in speech recognition (Sakoe and Chiba 1978). Later on, it was used as a similarity measure in data mining for pattern recognition (Muller 2007). The main advantage of DTW is its functionality to warp two sequences nonlinearly in the time dimension producing so called "warping path". Moreover, the algorithm does not require these sequences to be of the same length. This implies that DTW can be used as shape matching tool for pattern mining in financial time series. Consequently, the nonlinear similarity distance matrix obtained by DTW could be used for other application, such as time series forecasting, classification, etc. Up to now, far too little attention has been paid to DTW application in finance. A recent paper (Tsinaslanidis et al. 2014) presents the experimental study where DTW is used to measure and match the similarity across daily returns of different classes of months; thus showing DTW applicability for determining the seasonality of the market. Through a simulation process the authors have also shown that DTW based similarity measure takes lower (greater) values when Pearson's and Spearman's correlation coefficients are great (low) in absolute terms. In a study (Kia et al. 2013) investigating the exchange rate dynamics, the authors introduced a methodology that uses K-nearest neighbours (KNN) and DTW to improve the fluctuation prediction and to have

better evaluation parameters compared to other researches. The experimental study was performed with USD/JPY exchange rate time series from 1971 to 2012 that were partitioned into 30 element fragments based on the monthly behaviour of the time series. Then, each of obtained fragments were divided with 7:3 ratio, and KNN was used to determine three nearest neighbours based on DTW similarity measure. The authors achieved 10% improvement in directional prediction comparing to the work (Yao and Tan 2000) which is one the most cited papers in the field of financial prediction. The paper (Bagheri and Peyhani 2014) presented a hybrid DTW - Wavelet Transform method for automatic pattern extraction in time series, especially for Foreign Exchange Market (FX). The authors pointed out that the presented method is a very effective for price forecasting and pattern extraction. FX market was also analysed in the paper (Wang et al. 2012) where DTW was employed to study the topology of similarity networks among 35 major currencies, measured by the minimal spanning tree approach. The study (Lee and Oh 2011) demonstrated the development of trading system for making investment decisions, where DTW was used to determine similar patterns in the frequency of stock data ascertain the optimal timing for trade. The patterns retrieved by the algorithm were verified by executing simulation under specific strategies.

Considering all of this evidence, it seems that there is a relatively small body of literature that is concerned with DTW application in finance field. This work therefore focusses on the pattern mining tasks based on DTW in financial time series. The research attempts to assess the effect of DTW application for subsequence matching in a certain time series and for clustering multiple financial time series. For the purpose of analysis, the time series of pension funds (PFs) operating in Lithuania have been selected. Recently, the pension reform replaced traditional pay-as-you-go system with advanced funding system introducing compulsory and voluntary pension funds. It is not surprising that time series of these emerging market pension funds are nonstationary and exhibit some features referred to as "stylized facts", thus highlighting the need of more enhanced methods for their analysis. In the experimental study, PFs are described by equity curve, which day by day shows the gross return from the initial time moment being set up. Since observations are fixed daily, time series are volatile and have a distortion in the time axis. For this reason, we claim that one-to-one mapping (e.g., Euclidean distance measure) would be inadequate since it is very sensitive to small distortions in the time axis. Euclidean and similar distance metrics could more suitable when data are aggregated in time, as quarterly or yearly time series. PFs to be analysed in this paper are characterized by different investment strategy resulting in diverse underlying dynamics of time series describing the historical performance of funds. Since PFs' participants can migrate between funds by deciding which fund to select, the analysis of their performance can provide added value for decision making. The following are some issues could be addressed by participant: identify fund of the same risk profile but with better growth pattern; discover time periods with similar dynamic patterns; identify funds with less commissions but the same growth pattern, etc. On the other hand, the comparative analysis of PFs dynamics, pattern matching of PFs with well-known indices could be also relevant for pension accumulation companies. Finally, we assume that the matching of PFs time series using DTW can be taken as an alternative to cross-correlations estimated between time series that are of particularly importance of dynamics for financial markets. To our knowledge, DTW application to analyse pension funds of different risk profile has not been published yet.

The authors of the paper pursue the research to develop the model, which would recommend to select the optimal pension fund in the second pillar of Lithuanian pension system by providing some guidelines to participants. The preliminary concept of model has been presented in the study (Kabasinskas et al. 2014). The analysis of return-risk performance of PFs has been already published by authors (Sutiene et al. 2014). This research continues the work in this field in order to reach definitive findings.

DATA: PENSION FUND MARKET TIME SERIES

The application of DTW algorithm is investigated applying it for time series of pension funds operating in Lithuania. At the end of 2016, five Pension Accumulation Companies managed these funds. Approximately 1.0 million participants (76% of employees) accumulated their pension in them. The most of participants chose the funds managed by "Swedbank investiciju valdymas" (40.44%) and "SEB investiciju valdymas" (22.07%); then other companies followed. 63% of PFs assets were also managed by these two companies (Bank of Lithuania 2016). For this reason, PFs managed by these companies and characterized by different investment strategy were selected for the experimental study to explore their historical dynamics and pattern similarity: two conservative investment funds (SW1, SEB1; 0% stocks), one low risk investment fund (SW2; up to 30% stocks), three medium risk investment funds (SW3, SW4, SEB2; 30-70% stocks), and two high risk investment funds (SW5, SEB3; 70-100% stocks). The daily data retrieved cover the time period of 05 May 2011 to 04 May 2017, starting from the latest fund history available. The dynamics of net asset value, $NAV(t)$, a conventional measure of the value of pension assets, describes the evolution of PFs. The equity curve $f(t) = NAV(t) / NAV(0)$, which in each t shows the gross return from the beginning to time t, is used as the fundamental variable in this research to describe the performance of funds. The notation $NAV(0)$ denotes net asset value at 05 May 2011. The resulting curves for each of PFs to be analysed are depicted in (Fig. 1). To evaluate the performance of any investment, it is compared against an appropriate benchmark. The manager of PFs uses its own composite benchmark index consisting of one or more universally adopted and widely used indices of financial instruments. The manager may select but not necessarily these indices intentionally. That is why the other well-known indices are included to reflect changes in markets or

parts thereof. The indices, such as Financial Times Stock Exchange 100, S&P 500, Dow Jones Industrial Average, Nasdaq, iShares iBoxx, Merrill Lynch, Nasdaq Bonds, Merrill Lynch Bonds, shortly FTSE, GSCP, DJI, NDAQ, HYG, ML, NDAQ.B, ML.B respectively, are used in the study in order to benchmark PFs' performance (Fig. 1). In total, data set consisted of 16 time series.

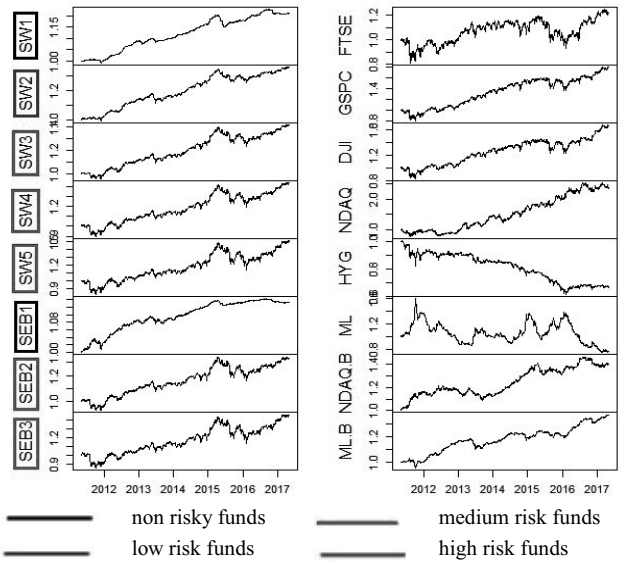


Figure 1: Equity Curves *f(t)* of PFs and indices

A visual inspection of historical performance in Fig. 1 points to some evidence about funds' dynamics: investment in stocks is led by higher day-to-day fluctuations compared with low risk funds. Any two time series may have very similar overall shape of dynamics, but those shapes are not exactly aligned in time. For this reason, the matching of patterns in time series requires nonlinear alignment in order to achieve more sophisticated similarity measuring.

METHODOLOGY: DYNAMIC TIME WARPING

Suppose we have two time series or two fragments of one time series that can be of different length. Let's denote them as $X = \{x_i\}_{i=1}^{n}$ and $Y = \{y_j\}_{j=1}^{m}$. To align two series using DTW, the $n \times m$ cross-distance matrix $d(i,j) \geq 0$ between x_i and y_j is determined. Various distance measures can be used for this purpose. At the core of the algorithm, the warping path $\phi(k) = (\phi_x(k), \phi_y(k))$, $k = \overline{1,T}$, $\max(n,m) \leq T \leq n+m-1$ is constructed, where time warping functions $\phi_x(k) \in \overline{1,n}$, $\phi_y(k) \in \overline{1,m}$ remap the time indices of series X and Y respectively. The path is valid if the following constraints are satisfied (Giorgino 2009):

- *Monotonicity* is introduced to preserve their time ordering and escape meaningless loops:
  $\phi_x(k+1) \geq \phi_x(k)$ and $\phi_y(k+1) \geq \phi_y(k)$;
- *Boundary* condition ensures that the time series' heads and tails are constrained to match each other:
  $\phi_x(1) = \phi_y(1) = 1$, $\phi_x(T) = n$, $\phi_y(T) = m$;
- *Continuity* constraint implies that arbitrary time compressions and expansions are allowed, and that all elements must be matched:
  $|\phi_x(k+1) - \phi_x(k)| \leq 1$, $|\phi_y(k+1) - \phi_y(k)| \leq 1$.

The classical formulation of DTW alignment has various modifications that have been published in the literature (Sakoe and Chiba 1978), (Rabiner and Juang 1993), (Myers et al. 1980). The authors usually classify them according to the bounds imposed on the slope, slope weighting, local continuity constraint type, and etc. For example, it can be set up how many elements repeatedly can be matched or how many skipped (Giorgino 2009).

Given ϕ, the average accumulated distortion between the warped time series X and Y is estimated

$$d_\phi(X,Y) = \sum_{k=1}^{T} d(\phi_x(k), \phi_y(k)) N_\phi(k) / M_\phi ;$$

where $N_\phi(k)$ is per-step weighting coefficient and M_ϕ is the corresponding normalization constant. DTW optimally aligns two series so that their distance is minimized $D(X,Y) = \min_\phi d_\phi(X,Y)$, i.e. deformation of the time axes of X and Y stretches time series to match each other as close as possible.

EXPERIMENTAL STUDY

The study covers three experiments. The first case demonstrates subsequence matching with DTW applied for pattern analysis. A single time series is decomposed into query sequence and the remainder one. Pattern matching is then performed on the extracted fragments of time series. Similar application was performed in the study (Kapler 2012). Second case considers the nonlinear alignment of two time series, such as any pension fund and index. The implementation of DTW was extended to multivariate case (the third case). DTW distance matrix obtained from the algorithm was applied as similarity criteria while performing the hierarchical clustering. This technique does not require to set the number of clusters in advance, that is considered as its advantage (Rani and Sikka 2012). Equity curves of PFs and indices were clustered according to the DTW measure. In total, 16 time series were aligned and grouped into clusters. To determine clustering quality, the clustering validity index – Silhouette value was introduced. The results of clustering under different corresponding constraints of DTW alignment were compared to hierarchical clustering with Euclidean distance and presented in this section. Experiments were performed using R software and packages dtw, tsclust (Giorgino 2009), (Montero and Vilar 2014).

Case 1: Univariate analysis

Let's apply DTW algorithm for funds SW1 (conservative fund) and SW5 (most risky fund) separately to search for subsequent matching in their historical dynamics. Most recent 40 days were selected for the experiment (query pattern). The algorithm was set up with 10 number of

matches to find over all horizon to be analysed. The visual demonstration of patterns obtained is depicted in Fig. 2-3.

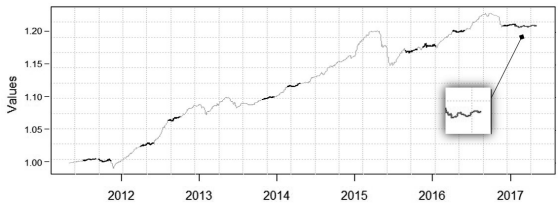


Figure 2: 10 Patterns (black) Obtained Matching to Recent 40 days (red) for SW1

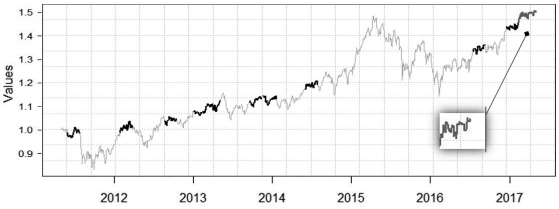


Figure 3: 10 Patterns (black) Obtained Matching to Recent 40 days (red) for SW5

Fig. 2-3 demonstrate the time series fragments (black) that were obtained as most similar to the given query pattern (red). Shape matching allows to discover similar shape based patterns in time series, thus sensitive to short-term oscillations. Such analysis allows to observe the time periods when the pattern recurs over the whole period, how the extracted fragments overlap among different funds, hence apply this information in prediction using strategies that are based on pattern recognition. It is worth noting that DTW alignment justifies for more volatile curve, i.e. for funds that invest in stocks. If the curve is stable, it is better to choose point-to-point matching method, such as Euclidean distance.

Case 2: Bivariate analysis

To demonstrate DTW alignment for two time series, the conservative pension fund SW1 has been selected that is benchmarked with ML.B index. Sakoe-Shiba slope-constrained pattern was implemented by choosing slope parameter $p = 0.5$. The use of Sakoe-Chiba Band limits the scope of the warping path, that's why it is rather monotonically increasing in Fig. 4-5.

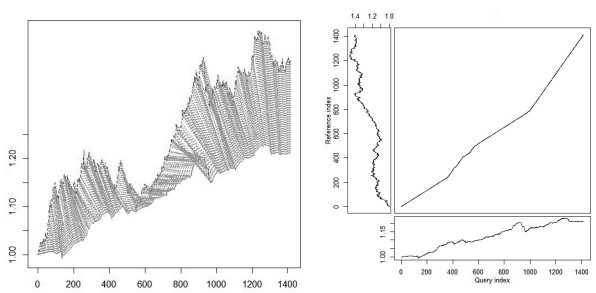


Figure 4: DTW Alignment (left) of SW1 (black) and ML.B (red), and Warping Path (right)

Comparing SW1 dynamics with ML.B, it can be seen that both time series exhibit rather similar pattern but it is shifted in the time axis except at the end of horizon to be analysed (Fig. 4). Under the same settings, the experiment was repeated to align most risky pension fund SW5 along with FTSE index. Both time series are rather volatile. At the beginning of period the behaviour of time series was coinciding in time but at the second half of horizon FTSE moved forward to SW5 (Fig. 5). This finding suggests the idea that FTSE index could be used as causal indicator or predictor to reflect possible movements in SW5.

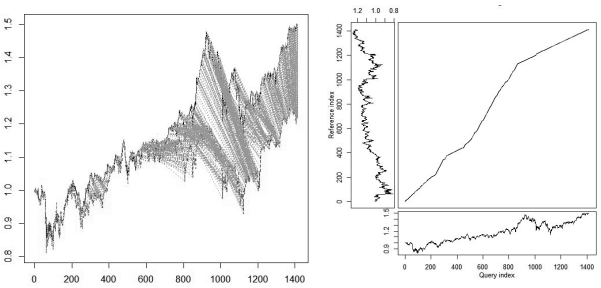


Figure 5: DTW Alignment (left) of SW5 (black) and FTSE (red), and Warping Path (right)

The usage of some well-known index to benchmark PF's performance can provide very useful information about fund's management. It may be the case to use index as causality tool for nonstationary time series to predict possible behaviour in future.

Case 3: Multivariate Analysis

The nonlinear alignment using DTW algorithm of PFs equity curves implementing various types of step patterns was pursued, then the hierarchical clustering of them was applied. The resulting dendrogram is depicted in Fig. 6 (on the left). In case of five clusters fixed, Silhouette value was improved by 21% compared with the case when the hierarchical clustering was performed using Euclidean distance. The clusters obtained in the dendrogam fully reflect the investment profile of PFs. One can see that there are three groups of PFs that are fairly similar: (SW1, SEB1) – conservative funds, (SW5, SEB3) – high risk funds, and (SW2, SEB2, SW3, SW4,) – low and medium risk funds. It is therefore likely that these two pension accumulation companies resulted in similar performance over time.

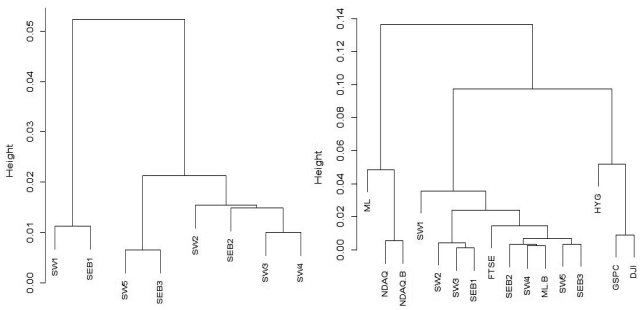


Figure 6: Dendrogram of PFs and Indices

The same procedure was repeated for equity curves of PFs in conjunction with indices. The dendrogram is given in Fig. 6 (on the right). The outcome of clustering PFs together with indices shows that funds exhibit similar patterns to the dynamics of FTSE and ML.B indices only. In this case, Silhouette value was improved by 28%. The current results highlight the importance of a benchmark to be chosen.

CONCLUSIONS

The time series of financial markets, especially emerging ones, exhibit some quite nontrivial statistical features. Thus, there is a need of developing new models or adapting methods from other research areas. The findings of this research provide insights for DTW application for similarity search in financial time series. The mining of similar patterns for one time series can have implications for the understanding possible behaviour in future. It is inferred that DTW is preferred to point-to-point similarity measure for more volatile curve of time series because of its flexibility to existing distortions. Despite the exploratory nature of the paper, this study offers some insight into DTW usage as causality tool for nonstationary time series to predict possible movements in future according to some indicator to be chosen. The implementation of DTW is also demonstrated for clustering time series. The results under different corresponding constraints of DTW alignment compared to hierarchical clustering with Euclidean distance have shown the superiority of DTW as similarity measure because of improved value of cluster validity index.

Further studies need to be carried out in order to validate DTW application for mining patterns. If it gives the promising results for analysis of PFs dynamics, DTW can be used as for planning a long-term pension accrual by selecting optimal fund.

REFERENCES

Bagheri, A., Peyhani, H.M. and Akbari, M. 2014. "Financial Forecasting using ANFIS Networks with Quantum-Behaved Particle Swarm Optimization". *Expert Systems with Applications*, 41(14), 6235-6250.

Bank of Lithuania. 2016. "Review of Lithuania's 2nd and 3rd pillar pension funds and of the market of collective investment undertakings. Technical report, Lithuania.

Banumathy, K. and Azhagaiah, R. 2015. "Modelling Stock Market Volatility: Evidence from India". *Managing Global Transitions*, 13(1), 27-42.

Champagnat, N., Deaconu, M., Lejay, A., Navet, N. and Boukherouaa, S. 2013. "An empirical analysis of heavy-tails behavior of financial data: The case for power laws". HAL Id: hal-00851429 https://hal.inria.fr/hal-00851429

Chavez-Demoulin, V., Embrechts, P. and Sardy, S. 2014. "Extreme-quantile tracking for financial time series". *Journal of Econometrics*, 181(1), 44-52.

Cont, R. 2001. "Empirical Properties of Asset Returns: Stylized Facts and Statistical Issues". *Quantitative Finance Volume,* 223-236.

Durante, F., Pappadà, R. and Torelli, N. 2014. "Clustering of financial time series in risky scenarios". *Advances in Data Analysis and Classification*, 8(4), 359-376.

Epaphra, M. 2017. "Modeling Exchange Rate Volatility: Application of the GARCH and EGARCH Models". *Journal of Mathematical Finance*, 7, 121-143.

Giorgino, T. 2009. "Computing and Visualizing Dynamic Time Warping Alignments in R: The dtw Package". *Journal of Statistical Software*, 31(7), 1-24.

Guam, H.S. and Jiang, Q.S. 2007. "Cluster financial time series for portfolio". In *Proceedings of International Conference on Wavelet Analysis and Pattern Recognition*, Beijing, 851-856.

Kabasinskas, A., Sutiene, K., Valakevicius, E. and Maggioni, F. 2014. "Stochastic programming framework for Lithuanian pension payout modelling". *Croatian operational research review*, 5(2), 387-399.

Kapler, M. 2012. "Evaluating Sample Trading Strategies using Backtesting library in the Systematic Investor Toolbox". Forum in www.SystematicInvestor.wordpress.com

Kia, A.N., Haratizadeh, S. and Zare, H. 2013. Prediction of USD/JPY Exchange Rate Time Series Directional Status by KNN with Dynamic Time Warping as Distance Function. *Bonfring International Journal of Data Mining*, 3(2), 12-16.

Lee, S.J. and Oh, K.J. 2011. "Finding the Optimal Frequency for Trade and Development of System Trading Strategies in Futures Market using Dynamic Time Warping". *Journal of the Korean Data and Information Science Society*, 22(2), 255-267.

Montero, P. and Vilar, J.A. 2014. "TSclust: An R Package for Time Series Clustering". *Journal of Statistical Software*, 62(1), 1-43.

Muller, M. 2007. *Information Retrieval for Music and Motion*. Chapter 4. Springer-Verlag, 69-84.

Myers, C., Rabiner, L. and Rosenberg, A. 1980."Performance Tradeoffs in Dynamic Time Warping Algorithms for Isolated Word Recognition." *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 28(6), 623-635.

Rabiner, L. and Juang, B.H. 1993. *Fundamentals of Speech Recognition*. Ch. 3. Prentice-Hall, Upper Saddle River, USA.

Rani, S. and Sikka, G. 2012. "Recent Techniques of Clustering of Time Series Data: A Survey". *International Journal of Computer Applications*, 52(15), 1-9.

Sakoe, H. and Chiba, S. 1978. "Dynamic Programming Algorithm Optimization for Spoken Word Recognition." *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 26(1), 43-49.

Sewell, M. 2011. "Characterization of financial time series". Research Note RN/11/01, University College London, London.

Shah, N. and Roberts, S.J. 2013. "Dynamically Measuring Statistical Dependencies in Multivariate Financial Time Series Using Independent Component Analysis". *ISRN Signal Processing*, 1-14.

Su, J.B. and Hung, J.C. 2011. "Empirical analysis of jump dynamics, heavy-tails and skewness on value-at-risk estimation". *Economic Modelling*, 28(3), 1117-1130.

Sutiene, K., Kabasinskas, A., Strebeika, D., Kopa, M. and Reichardt, R. 2014. "Estimation of VAR and CVAR from financial data using simulated alpha-stable random variables". In *Proceedings of 28th European Simulation and Modelling Conference - ESM'2014*, October, 2014, Portugal, 159-163.

Tsinaslanidis, P., Alexandridis, A., Zapranis, A. and Livanis, E. 2014. "Dynamic Time Warping as a Similarity Measure: Applications in Finance". In *Proceedings of Hellenic Finance and Accounting Association*, Volos, Greece.

Wang, G.J., Xie, C., Han, F. and Sun, B. 2012. Similarity measure and topology evolution of foreign exchange markets using dynamic time warping method: evidence from minimal spanning tree. *Physica A: Statistical Mechanics and its Applications*, 391, 4136-4146.

Yao, J. and Tan, Ch.L. 2002. "A case study on using neural networks to perform technical forecasting of forex". *Neurocomputing*, 34(1), 79-98.

Zhao, Z.Y., Xie, M. and West, M. 2016. "Dynamic dependence networks: Financial time series forecasting and portfolio decisions". *Applied Stochastic Models in Business and Industry,* 32(3), 311-332.

DECISION MANAGEMENT SIMULATION

GENERATING SYNTHETIC INDIVIDUAL-HUMAN POPULATION AND ACTIVITY MODELS

Emily Schmidt
CSE Department, Miami University
Oxford, OH 45056, USA.
email: schmidee@miamiOH.edu

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

KEYWORDS
Demographics, Contact Networks, Individual-based Model, Model Generation, Epidemiology

ABSTRACT

Simulation-based analysis is rapidly gaining importance for developing public policy for dealing with a broad spectrum of issues ranging from natural catastrophes to seasonal communicable diseases. Simulation-based analysis requires the use of high fidelity models for conducting in-depth studies of different scenarios. However, generating valid, large-scale human models from raw demographic, geographic, and other statistical data is challenging. This paper proposes a "first principles" approach to generating models using summary statistics from authoritative sources. Our generated model is called Synthetic, Individual-human Population and Activity Model (SIPAM). SIPAM characterizes population demographics, schools, businesses, and typical daily activities of individuals at a given level of detail (currently 1 km, based on data availability). This paper discusses: ① our method for generating a SIPAM, ② the exhaustive set of verification experiments (10 million replications with ~1.04 billion activities), and ③ case studies of seasonal influenza epidemic for comprehensive validation. The experiments establish that the proposed method yields valid synthetic models that are amenable to a variety of analysis.

INTRODUCTION

Population growth and increasing urban densities pose many challenges for dealing with a broad spectrum of issues ranging from natural catastrophes, seasonal communicable diseases, to routine city planning. An example of such a issue was witnessed around the Gulf Coast of United States during the 2005 hurricane Katrina – it was tragically discovered the evacuation plans were not adequate to evacuate the impacted area (Daniels, 2007). The root issue is that responses to particular catastrophes are usually based on historical examples of similar events and not on current scenarios. Conventional statistical analyses are not conducive for "what-if" type analyses that is required for policy assessments. Furthermore, they do not yield sufficiently detailed and intuitive information about individuals in the population and their stochastic behaviors. Consequently, innovative approaches are needed to *proactively* address these growing challenges.

Simulation-based analyses are rapidly growing in importance to meet the aforementioned needs. Catalyzed by rapid advancement in computational infrastructures, simulations enable systematic and multifaceted analysis required for policy development (Giridharan and Rao, 2016). Simulations fundamentally rely availability of a valid, comprehensive, and robust model. Important model characteristics include: ❶ Realism: The model must be realistic and mirror geographic, demographic, and behavioral characteristics; ❷ Reusability & accuracy: Investments into model development and validation are effectively amortized only when models can be reused or easily adapted for different types of analyses; and ❸ Computational costs: Time and resources required for model generation, validation, simulation, and analysis need to be balanced with realism, accuracy, and effective use (Chen and Zhan, 2008).

Generating realistic, reusable, and cost effective models for comprehensive, multifaceted analysis of human populations and day-to-day activities is a daunting task (Barrett et al., 2009) due to the following diverse challenges: ① identifying data necessary for model generation, ② finding suitable data sources, ③ preprocessing and streamlining data for modeling, ④ developing and validating methods for model generation, and ⑤ verification of generated models.

Overview & advantages of proposed approach

In this paper we propose a novel method for generating realistic synthetic models of individuals in a population along with their associated daily activities. Our model is called Synthetic Individual-human Population and Activity Model (SIPAM). A SIPAM is generated from anonymous summary statistics that can be readily obtained from authoritative data sources, such as: USCB (2011), USBLS (2015), and demographic databases (CIESIN, 2016). SIPAM preserves demographic, socioeconomic, and geospatial characteristics consistent with the resolution of available data – that is, aggregate characteristics of a SIPAM are statistically indistinguishable from the census data used to create it.

The generated model includes temporospatial activities (for a full week) for each individuals based on their age and employment status (working, unemployed, retired). The activities are structured around a variety of buildings, such as: homes, schools, businesses, etc. The buildings are generated as part of the model using census data. These are key aspects of the SIPAM that ensure it is realistic. The data is supplied

via a comprehensive input configuration file to enable effective reusability and extensibility. The activities are scheduled in configurable time intervals or time-blocks (currently set to 10 minute blocks) thereby striking a balance between realism versus computational costs for model generation and simulation. The models can be used for a variety of simulation-based analyses using different approaches, including: cycle-based simulations, agent-based simulations, or discrete event simulations. Our exhaustive suite of verification and validation experiments (see Section EXPERIMENTS) establish that the proposed method generates valid, realistic SIPAMs that are amenable to a variety of analyses.

BACKGROUND & RELATED WORKS

Realistic and comprehensive modeling and simulation-based analysis of specific aspects of human populations has been an active area of research in diverse fields, including: computational epidemiology, transportation, and economics. Humans are either modeled as collection of interacting individuals or groups. Group models represent a collection of collocated individuals modeled as an indivisible entity. Different types of group models have been proposed by several investigators including Balcan et al. (2010), Rao et al. (2009), and Keeling (2005). In their models, the groups are organized based on their geographic locations resulting in a logical structure similar to a Voronoi tessellation. Interactions between groups is modeled implicitly based on their adjacency or via explicit mobility networks. The benefit of using a group model is it reduces the computational cost because the model consists of fewer number of entities which significantly reduces computations preformed at each time step. Furthermore, such models do not require detailed, voluminous data about the population, which can be hard to obtain.

The primary disadvantage of aggregate or group models is that information about each individual is not preserved. The models do not preserve heterogeneity that may be present within the group. However, such information maybe vital for certain types of analyses and design of public policies. Consequently, several researchers have proposed the use of individual-based models, where each individual is independently modeled. Such models essentially embody contact-networks which define temporal interactions that occur between individuals. Longini et al. (2005) discuss the generation and use of synthetic, individual-based models for containing influenza epidemics. They use a variety of data sources to generate individuals and their temporospatial activities. Recently Bhatele et al. (2017) discuss enhancements to modeling and simulation of individual-based model of the United States generated as part of prior work by Barrett et al. (2009). Their model is fine-grained and has been generated from a variety of public and proprietary data sources.

The advantages of individual-based models are: ① they yield temporospatial characteristics for epidemics because they explicitly model the location and contacts between individuals in the populations, ② since they are typically less prescriptive, they can be used for analyzing epidemics whose parameters are not well established, and ③ they enable vivid visualization. The drawbacks of individual-based models are: ① they are computationally demanding for model generation and simulation (Bhatele et al., 2017), ② it is hard to fit such models to surveillance data due to the significant differences in resolution, and ③ the volume of data generated by the models can pose bottlenecks for analysis.

Similarities & differences to our approach

The proposed method for generating a Synthetic, Individual-human Population and Activity Model (SIPAM) is a novel combination of group models and individual-based models. The design rationale is to preserve advantages of the two types of models while minimizing their drawbacks. SIPAM is similar to individual-based models in that daily activities for each individual is explicitly generated. This enables tracking interactions between individuals for various analysis. However, spatial resolution for buildings and people is limited (currently to 1 km^2) similar to aggregate or group models. Another objective is to enable generation of SIPAMs using freely available summary data sets. This enables their ready use by the community for different applications and geographic regions.

MODEL GENERATION METHOD

The proposed method for generating a Synthetic Individual-human Population and Activity Model (SIPAM) is summarized in Figure 1. The method consists of four main phases, namely: ❶ Population/family generation, ❷ Building generation, ❸ Daily schedule generation, and ❹ Output generation. The statistical data required for SIPAM generation is supplied via a text-based configuration file. The generated SIPAM can be used for a different types of simulation-based analyses by combining it with a suitable *policy* model implemented during simulation. Policies essentially modify attributes of individuals and their daily schedules. For example, in this study an epidemic policy is used to simulate progression of seasonal influenza in the generated synthetic population. Simulation of epidemics using a synthetic model has been enabled via a cycle-based simulator called HAPLOS, developed as part of this investigation. The details of each of the four phases are described in the following subsections.

Phase 1: Population/Family Generation

The first phase in generation of a SIPAM involves creation of households of different "sizes" (*i.e.,* number of individuals), with each individual meeting the specified age and sex distributions observed in a census. In our example models, we have obtained household size, individual age, and sex distribution summaries from the USCB (2011). Census data is summarized as probabilities in the input configuration file as shown in Figure 2. The population generation process beings with creating families of different sizes

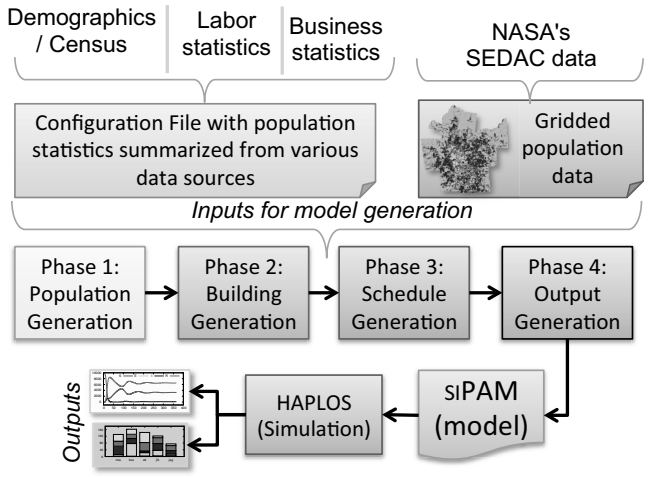


Figure 1: Overview of proposed method for generating a Synthetic Individual-human Population and Activity Model (SIPAM)

Total_Population=1979202	Male_Probablity=0.492782305
Family_Size_1_Probablity=0.27	Age_5-Younger_Probablity=0.065
Family_Size_2_Probablity=0.34	Age_5-13_Probablity=0.114
Family_Size_3_Probablity=0.16	Age_14-17_Probablity=0.052
Family_Size_4_Probablity=0.14	Age_18-24_Probablity=0.097
Family_Size_5_Probablity=0.06	Age_25-44_Probablity=0.263
Family_Size_6_Probablity=0.02	Age_45-64_Probablity=0.261
Family_Size_7_Probablity=0.01	Age_65-Older_Probablity=0.148

Figure 2: Fragment of input configuration file with demographic data of desired synthetic population

based on the probability values specified in the input configuration. A standard uniform, discrete random number generator (`std::discrete_distribution`) is used to determine family sizes. This random number generator produces integers on the interval $[0, n)$, where the probability of each individual integer i is defined as $w_i \div S$, where $S = \Sigma w_i \quad (0 \leq i \leq n)$, that is the probability of the ith integer divided by the sum of all probabilities.

Next, the given number of individuals of different ages are generated for each family. The age and gender of each person in a family is determined based on the demographic probabilities specified in the input configuration file (see Figure 2), with an added restriction that the first member of a family must be an adult. A uniform, discrete random number generator is used to determine age and sex for each person. For persons older than 64 years, the fraction of unemployed adults is used to assign a "retired" status. The family generation process is repeated until the number of persons in the model exceed the specified population.

It must be noted that currently our method does not track other demographic properties such as race, ethnicity, etc., for each person. Consequently, the aforementioned method yields households that are statistically indistinguishable from the census data. However, if additional demographic properties are desired in the generated SIPAM, then Iterative Proportional Fitting (IPF) statistical method along with Public Use Microdata Sample (PUMS) can be used (Beckman et al., 1996). PUMS essentially provides an n-dimensional table of typical family configurations and marginals. The IPF statistical procedure iteratively adjusts cell values to estimate family size and configurations such that marginal totals remain fixed (Beckman et al., 1996). The resulting fitted tables can be used for generating synthetic households.

Phase 2: Building Generation

The second phase of model generation involves creation of variety of buildings associated with different activities. Our method currently generates the following types of buildings – ① *business*: general buildings where people may visit or work, ② *medical*: hospitals where people work, visit, or are interned, ③ *school*: further subdivided into elementary, middle and high schools where people may visit, study, or work, ④ *daycare*: children are interned while adults may visit or work, ⑤ *transport hub*: used as temporary holding area for commuters using public transport, and ⑥ *building*: used a general purpose placeholder for homes, apartments, etc.

Building generation commences with assigning homes to each family generated in Phase 1. Location of homes is determined based on population counts at a given spatial resolution. Our building generation method uses gridded population counts to appropriately distribute homes and families to reflect spatial population characteristics. We use gridded ASCII data format from NASA's Socioeconomic Data and Applications Center (SEDAC) (CIESIN, 2016). SEADC provides gridded population data at a resolution of 30 arc-seconds or approximately 1 km^2 as shown in Figure 3. Furthermore, a transport hub is assigned to each grid.

In the next step, sufficient number of day care centers as well as elementary, middle, and high schools are generated to accommodate children and school-age persons generated in Phase 1. The school capacities and occurrence probabilities are supplied via configuration data as shown in Figure 4. The sizes and occurrence probabilities have been determined from summary tables published by USCB (2012). The locations of the schools are determined based on population densities in the gridded population, with higher population areas having more buildings.

Similarly businesses are generated to accommodate all working adults. The size and occurrence for businesses in the generated SIPAM is determined based on probabilities (computed from data published by USCB (2012)) specified in the input configuration data as shown in Figure 4. The businesses are also spatially distributed to various grids based on relative population densities in the grids. Moreover, as buildings are generated, references to them are stored in different datastructures to enable rapid access during schedule generation. In this context, it must be noted that unlike homes that have persons in a family assigned to them, people are not yet assigned or "mapped" businesses and schools. Instead, this mapping occurs during schedule generation in Phase 3.

Figure 3: Example of gridded population data (persons per km^2) for city of Cincinnati, Ohio, USA (total population: ~2 million). Data freely available from CIESIN (2016)

Business_Size_0-4_Pr=0.478	School_Size_0-4_Pr=0.0008
Business_Size_5-9_Pr=0.135	School_Size_5-9_Pr=0.002
Business_Size_10-19_Pr=0.085	School_Size_10-19_Pr=0.005
Business_Size_20-99_Pr=0.092	School_Size_20-99_Pr=0.013
Business_Size_100-499_Pr=0.049	School_Size_100-499_Pr=0.009
Business_Size_500_Pr=0.161	School_Size_500_Pr=0.001

Figure 4: Fragment of input configuration file with building probabilities from USCB (2012)

Phase 3: Schedule Generation

The third phase deals with generation of weekly schedules for each individual depending on their age and employment status. A weekly schedule includes two distinct sub-schedules, namely: 5 weekday (working days) schedules and 2 weekend (non-working day) schedules. The daily schedules are rounded to 10 minute activity blocks, *i.e.,* each day is represented by 60 minutes$\times 24$ hours $\div 10 = 144$ blocks. The primary motivation for organizing schedules into blocks is to strike a tradeoff between model complexity, accuracy, memory, and runtime of simulations.

The schedule associated with a person is determined primarily on their age and employment status. The employment probabilities based on age have been determined based on total statistics reported by USBLS (2015). The travel distances to work and time taken to travel have been estimated based on type of transportation to work reported by USCB (2014). The probabilities and associated travel distances are specified as part of the input configuration file. Currently, we have included the following restrictions for travel: ① distance traveled by walking is limited to 2 miles, ② maximum distance traveled by public transport is limited to 10 miles, ③ school attending children are limited to travel to school in a 10 mile radius.

The schedules generated for each person is generated from one of the following templates:

- Young child template (age $<$ 5 years): Young persons are assigned the same schedule as an adult in their family. If the adult works then the adult will be assigned to take the individual to a daycare prior to leaving to work location. The child will be assigned to the closet daycare (generated in Phase 2) to their home location that is not filled to capacity. When the adult in the family is no longer at work, they are scheduled to retrieve the child from the daycare. The child will then follow the adults schedule till the next day.
- School Age Child Template (5–17 years of age): Each person is assigned to attend a school (generated in Phase 2) during weekdays only. School is assigned based on the person's age and the nearest school to their home that provides the necessary grade level and is not at capacity. The schedule of young school children (*i.e.,* 5–13 years of age) is generated based off a designated care-giving adult in their family. However, school children older than 13 years are assigned an independent schedule with limited travel radius.
- Working adult template ($>$18 years): Weekday schedules are anchored around working at a business location generated in Phase 2. The distribution of working hours model the data from USBLS (2015). If the weekly work hours exceed 40, then the maximum number of hours/day is set to 10 hours. Otherwise the maximum number of works hours/day is assumed to be 8 hours. During periods when an adult is not at work, they are scheduled to visit other buildings or return home. The capacity for visitors for businesses will range between 1–500 visitors/hour depending on the number of employees. Furthermore, the schedules are further modified if the person is also designated as a child-care adult. In this scenario, the schedules are updated to include travel to-and-from daycare location.
- Non-working adult template ($>$18 years): The non-working adult template is meant for unemployed or retired adults (designated in Phase 1). This differs from the working adult template by removing the need of having to be at a job. Thus their schedules will be much more random in terms of locations visited during the day. The time spent at each of these locations will then be distributed through out the week randomly.

Similar to the above weekday travel patterns, weekend travel patterns are also specified via configuration file parameters. Weekend travel patterns are determined based on age of a person, with younger children spending more time at or near home. Adults are set to spend 50% of daytime traveling with a fixed radius. Currently, the schedules do not include spe-

cial scenarios such as vacations, holidays, etc. Such schedules need to be suitably incorporated into the model during simulation depending on analysis needs.

Phase 4: Output generation

The previous three phases operate using in-memory datastructures to enable rapid model generation. The final phase of model generation stores the resulting model to disk. A custom textual file format has been used to store the resulting model. The format has been chosen to ease the use of the model for conducting simulations and performing various analysis. Furthermore, summary statistics on the synthetic population demographics and buildings are also generated to aid verification and validation.

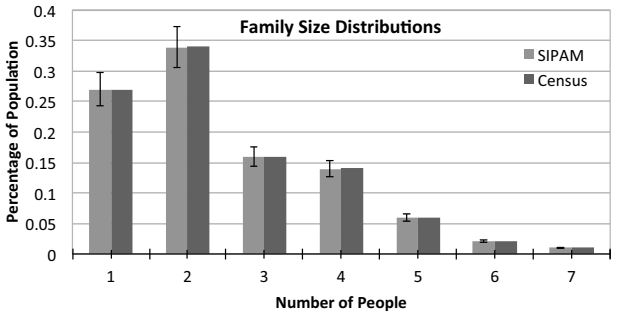
EXPERIMENTS

The proposed method for generating a Synthetic Individual-human Population and Activity Model (SIPAM) has been validated using both real-world data and a number of test data sets. The real-world dataset for validation was based on the greater metropolitan area of Cincinnati, Ohio, a typical city in mid-west United States. Figure 3 shows the gridded population data from NASA's SEDAC data set that is freely available from CIESIN (2016). The metropolitan area has a population of ~2 million with a significantly varying population spread over 1440 × 960 grid from CIESIN (2016). In addition to the large real-world data set a smaller test city with 22,000 residents spread on a 500 × 500 grid using a triangular distribution has also been used. The primary motivation of using a smaller test city was to enable extensive validation of schedule generation, which is a time consuming for large models.

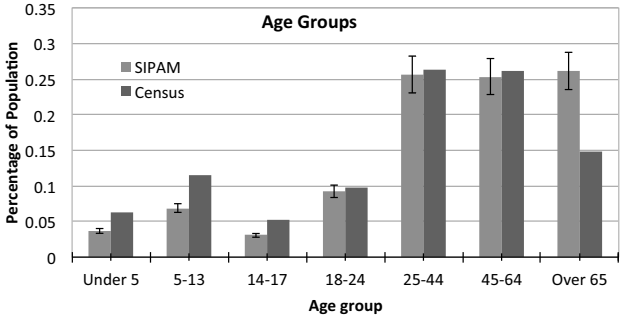
Validation of the model generation method and the generated SIPAM has been conducted in two steps. First, the key characteristics of the generated model has been validated using real-world data set as discussed in the following subsection. Next, the schedule generation has been verified to ensure that valid schedules are generated in a broad range of scenarios. Importantly, the SIPAM as also been validated using an influenza epidemic scenario as discussed in subsection Full SIPAM validation

Validation of population & building generation

The core method for generating synthetic populations and buildings has been validated using Cincinnati, Ohio as the reference. Since the model generation method is probabilistic, each run will yield a slightly different SIPAM. Consequently, 10 different models were generated (from exactly the same input configuration) and collectively analyzed for validation. The chart in Figure 5(a) shows a comparison of family size distributions. The chart in Figure 5(b) shows a comparison of individuals in different age groups. The error bars in the chart show the 95% Confidence Intervals (CIs) computed from variance observed in the 10 replications. As



(a) Family size distribution



(b) Population age distribution

Figure 5: Comparison of key demographics of SIPAM versus census data

illustrated by the charts the family size, the primary parameter, is statistically indistinguishable between the SIPAM and input census data. The age groups also show consistent distribution in most cases except for age $>$ 65 years. Our analysis suggests that the source of this discrepancy (for age $>$ 65 years) is with using a discrete distribution which rounds values more disparately.

The chart in Figure 6 shows a comparison of size of business (*i.e.,* number of employees) between the synthetic model and the supplied census data. The error bars in the chart show the 95% Confidence Intervals (CIs) computed from variance observed in the 10 runs. As illustrated by the charts the business sizes is statistically indistinguishable between the SIPAM and input census data. We did observe that for small models the variance in business sizes was larger. However, as the size of the model increases, the synthetic model produces statistically identical distributions. The experimental analysis establishes that the proposed method produces valid synthetic models.

Verification of schedule generation

Verification of generated synthetic schedules was conducted in two steps. First, the number of generated schedules for different categories of individuals in the SIPAM was compared with the summary census data. The chart in Figure 7 shows a comparison of different schedule types. The error bars in the chart show the 95% Confidence Intervals (CIs) computed from variance observed in 10 replications of model. Note that

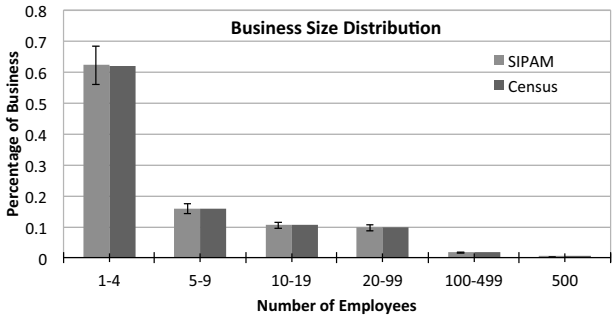


Figure 6: Comparison of business size distribution of SIPAM to census data

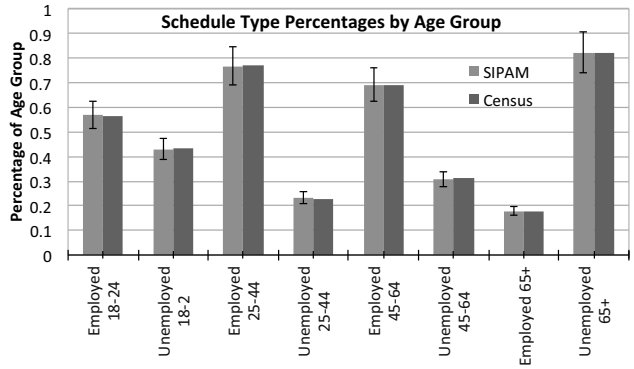


Figure 7: Comparison of different schedule types for different age groups in SIPAM to census data

variance is expected due to the stochastic nature of model and schedule generation. As illustrated by Figure 7, the distribution of different types of schedules is statistically indistinguishable between the SIPAM and input census data. This chart provides baseline verification of the proposed schedule generation method.

Next, a series of independent unit tests were developed to test the accuracy and viability of the generated schedules. The tests included checks to ensure school children are scheduled to attend school on weekdays and are not scheduled to go to work. The schedules for young children were follow the same timelines as the assigned child-care adult in the family. Checks were added to ensure employed adults are scheduled to work at least once and meet their total work hours. In additional several general checks such as being at home for 6 to 8 hours every 24 hours etc. was included in the tests. The tests also checked for consistency of schedules for members in a family.

The resulting tests covered various combinations of 16 different scenarios. The tests were used to validate schedules generated from 10 million replications with different parameter settings. The tests covered ~1.04 billion unique schedules for over 280 million families. The tests were useful to identify unique, conflicting scenarios which required inclusion of necessary logic to resolve the conflicts. With the schedule conflicts resolved, the final model generation passed all of the billion tests, thereby verifying the schedule generation method discussed in Section Phase 3: Schedule Generation

Full SIPAM validation

Having verified the proposed model generation method, we pursed validation of the complete synthetic model through case studies of an epidemic. For this experiment, we chose to model a seasonal influenza (*i.e.,* flu) epidemic and compare our results against validated results from similar models proposed by Nsoesie et al. (2012). Progression of the influenza epidemic in an individual has been characterized using classical compartmental models, with the following 4 compartments: Susceptible (S) → Exposed (E) → Infective (I) → Recovered (R). Transitions between the SEIR states is governed by probabilistic transitions. The incubation period (*i.e.,* E → I) has been set to $\{1, 2, \text{or } 3\}$ days with probabilities of $\{0.3, 0.5, \text{or } 0.2\}$ respectively. Likewise, the infection period (*i.e.,* I → R) has been set to $\{3, 4, 5, \text{or } 6\}$ days with probability $\{0.3, 0.4, 0.2, \text{or } 0.1\}$ respectively.

Infection transmission rate from an infective to a susceptible person (*i.e.,* S → E transition) has been characterized using the following equation proposed by Nsoesie et al. (2012):

$$Pr(w(i,j)) = 1 - (1-\tau)^{w(i,j)} \quad (1)$$

where, i denotes infected person, j denotes susceptible person, $w(i,j)$ denotes the contact time between person i and j, and τ is the disease transmission probability per unit of time. Contacts between persons arise when they are collocated in a building. As the time of collocation $w(i,j)$ increases, the probability of infection also increases.

Human Population and Location Simulator (HAPLOS)

Simulation of epidemics using a synthetic model has been enabled via sequential simulator called HAPLOS. It is a conventional cycle-based simulator in which activities of each individual are performed in each time step. Recollect that the time step of a SIPAM has been set to 10 minute increments to strike a tradeoff between model complexity, accuracy, memory, and runtime of simulations. In each time step, a person's location is suitably updated based on their generated weekly schedules. Next, epidemic progression between S → E → I → R compartments is simulated. Finally, contacts between susceptible and infective individuals collocated in each building simulated. Periodically HAPLOS saves the full state of the model to enable visualization, analysis, and validation.

Epidemic scenarios & model validation

Model validation using influenza epidemic as a case study has been conducted using HAPLOS and a SIPAM with a population of 22,000 residents spread (using a triangular distribution) on a 500 × 500 grid. The model was simulated with 5 randomly selected individuals to be exposed to the infection. The model was simulated with 5 different values of τ (see Equation 1). Results from multiple stochastic simulations for each value of τ have been averaged for analysis and

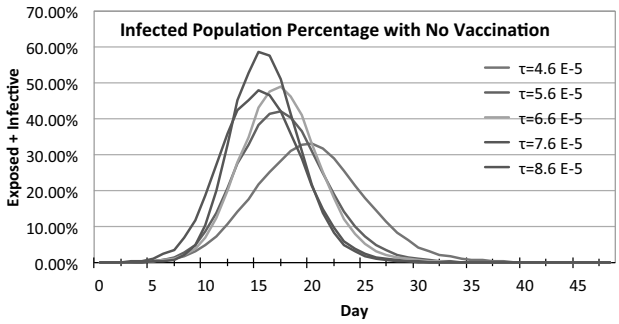


Figure 8: Fraction of exposed + infective population (averaged from multiple simulations) for different values of τ

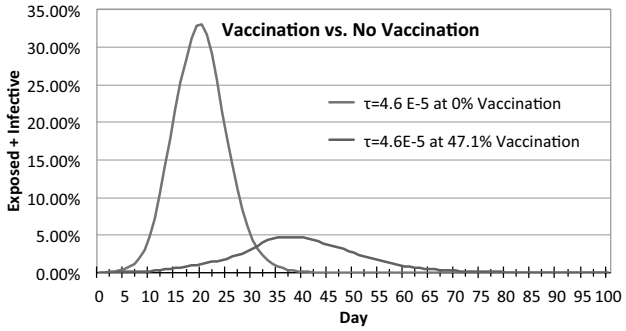


Figure 9: Comparison (Fraction of exposed + infective population) of epidemic progression with and without vaccination

plotted Figure 8. As illustrated by the chart, the epidemic curves follow the expected characteristic trend (see Nsoesie et al. (2012) for analytical details), with higher values of τ causing earlier infection peaks. Importantly, the characteristic curves establish the overall validity of the synthetic population and schedules.

Experiments were also conducted to validate initial settings and model behavior for conducting simulation-based analysis of different scenarios. Specifically, we explored the impact of vaccinating 47.1% of the population, analogous to the current vaccination rates. At this vaccination rate the epidemic is expected to have a significantly reduced peak, but with a slightly extended epidemic period (Nsoesie et al., 2012). Vaccination of population is modeled by randomly initializing 47.1% of the population to the Recovered (R) SEIR state. The chart in Figure 9 compares the fraction of infective population with and without vaccination. Consistent with expectations, vaccination decreases the peak infection while extending the epidemic period.

Summary of experimental results

The proposed method for generating a Synthetic Individual-human Population and Activity Model (SIPAM) and the generated model has been extensively validated using a variety of approaches. The experimental results in Figure 5 and Figure 6 essentially validate Phase 1 and Phase 2 of model generation that deal with generating synthetic populations, homes, schools, and businesses. Generated schedules were verified to have distribution consistent with census data as shown in Figure 7. Moreover, extended verification of schedules was conducted by generating 10 million randomized SIPAMs and verifying ~1.04 billion unique schedules. Having verified the each phase of model generation, the resulting SIPAM has been further validated using an seasonal influenza case study as discussed in Section Full SIPAM validation. The comprehensive set of verification and validation experiments establish that the proposed method generates valid synthetic models.

CONCLUSIONS

The need to analyze, design, and proactively implement sophisticated public health policies has rapidly grown due to population growth, urbanization, and emergent communicable diseases. Simulation-based methods are gaining broad applicability in this area due to their advantages. Simulations requires valid, realistic temporospatial models that effectively characterize human demographics and daily activities. The current state-of-the-art models models are broadly classified into two categories, namely: ① individual-based models in which each person and their activities are explicitly represented; and ② group-based or aggregate models in which a set of collocated humans are modeled as an indivisible entity. These methods have their respective advantages and disadvantages (see BACKGROUND & RELATED WORKS). Immaterial of the type of model being used, generating realistic, reusable, and cost effective models for comprehensive, multifaceted analysis of human populations and day-to-day activities is a challenging task Barrett et al. (2009). The challenges arise due to a myriad of issues, including: availability of data sources, computational costs, verification, and validation of complex models.

This work discussed a novel, "first-principles" method for generating a realistic Synthetic Individual-human Population and Activity Model (SIPAM). The design rationale underlying our modeling approach is to preserve advantages of both individual and group models without succumbing to their drawbacks. SIPAM includes daily activities for individuals but buildings and population densities are grouped into small regions. The size of the region is determined by resolution of input data, which is currently at 1 km^2.

Statistical analysis on generated models for various regions shows that SIPAM preserves demographic, socioeconomic, and geospatial characteristics – that is, aggregate characteristics of a SIPAM are statistically indistinguishable from the census data used to create it. The model and method have been validated using 10 million model replications with different parameter settings. The tests covered ~1.04 billion unique schedules for over 280 million families.

In addition to exhaustive verification, the model has also been validated using influenza epidemics as case studies. The case studies also explored impact of vaccination policies. Simulations for the case studies were conducted using a custom

cycle-based simulator called Human Population and Location Simulator (HAPLOS). HAPLOS has been developed as part of this investigation. The experimental data shows that SIPAM faithfully reproduces epidemic characteristics consistent with other validated models. The comprehensive set of verification and validation experiments establish that the proposed method generates valid synthetic models. Our investigations also establish that SIPAM strikes an effective balance between key model characteristics discussed in Section , namely: realism, reusability & accuracy, and computational costs.

A conspicuous advantage of our method is that uses anonymous summary census statistics that can be readily obtained from authoritative data sources. Since our method relies only on small subset of census data it can be readily used for other countries and geographic regions. The resolution of the model can increased or decreased via different gridded population data represented via simple ASCII text files. The activities are scheduled in configurable time intervals or time-blocks (currently set to 10 minute blocks) thereby striking a balance between realism versus computational costs for model generation and simulation. The models can be used for a variety of simulation-based analyses using different approaches, including: cycle-based simulations, agent-based simulations, or discrete event simulations.

FUTURE WORK

This paper presented the current checkpoint in our ongoing investigations on generating realistic models. We are continuing to refine and extend our methods to incorporate more detailed real world data when available. Currently, we are investigating the use of freely available street maps to further refine locations of homes, schools, businesses, and other buildings. The use of additional demographic data from NASA's SEDAC data sets is also being explored. Furthermore, we are working on performance improvements to our model generation routines to reduce time and memory footprint. We are optimistic that with these ongoing enhancements will yield improved, realistic, cost optimal synthetic models that can be used for a broad range of analysis in a variety of fields.

ACKNOWLEDGEMENTS

This work was funded in part by Miami University's Center for Analytics & Data Science (CADS).

REFERENCES

Balcan D.; Gonalves B.; Hu H.; Ramasco J.J.; Colizza V.; and Vespignani A., 2010. *Modeling the spatial spread of infectious diseases: the GLobal Epidemic and Mobility computational model*. *Journal of computational science*, 1, no. 3, 132–145. ISSN 1877-7503. doi:10.1016/j.jocs.2010.07.002.

Barrett C.L.; Beckman R.J.; Khan M.; Kumar V.A.; Marathe M.V.; Stretz P.E.; Dutta T.; and Lewis B., 2009. *Generation and analysis of large synthetic social contact networks*. In *Proceedings of the 2009 Winter Simulation Conference (WSC)*. ISSN 0891-7736, 1003–1014. doi:10.1109/WSC.2009.5429425.

Beckman R.J.; Baggerly K.A.; and McKay M.D., 1996. *Creating synthetic baseline populations*. *Transportation Research Part A: Policy and Practice*, 30, no. 6, 415–429. ISSN 0965-8564. doi:10.1016/0965-8564(96)00004-3.

Bhatele A.; Yeom J.S.; Jain N.; Kuhlman C.J.; Livna Y.; Bisset K.R.; Kale L.V.; and Marathe M.V., 2017. *Massively Parallel Simulations of Spread of Infectious Diseases over Realistic Social Networks*. In *2017 17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGRID)*. 689–694. doi:10.1109/CCGRID.2017.141.

Chen X. and Zhan F.B., 2008. *Agent-based modelling and simulation of urban evacuation: relative effectiveness of simultaneous and staged evacuation strategies*. *Journal of the Operational Research Society*, 59, no. 1, 25–33. ISSN 1476-9360. doi:10.1057/palgrave.jors.2602321.

CIESIN, 2016. *Gridded Population of the World, Version 4 (GPWv4): Population Count*. NASA Socioeconomic Data and Applications Center (SEDAC). doi:10.7927/H4X63JVC. Center for International Earth Science Information Network, Columbia University.

Daniels R.S., 2007. *Revitalizing emergency management after Katrina*. *Public Manager*, 36, 16–20. ISSN 2381-4160.

Giridharan N. and Rao D.M., 2016. *Eliciting Characteristics of H5N1 in High-Risk Regions Using Phylogeography and Phylodynamic Simulations*. *Computing in Science Engineering*, 18, no. 4, 11–24. ISSN 1521-9615. doi:10.1109/MCSE.2016.77.

Keeling M.J., 2005. *Models of foot-and-mouth disease*. In *Proceedings of the Royal Society B: Biological Sciences*. 1569. ISSN 0962-8452, 1195–1202. doi:10.1098/rspb.2004.3046.

Longini I.M.; Nizam A.; Xu S.; Ungchusak K.; Hanshaoworakul W.; Cummunings D.A.T.; and Halloran M.E., 2005. *Containing pandemic influenza at the source*. *Sience*, 309, no. 5737, 1083–1087. doi:10.1126/science.1115717.

Nsoesie E.O.; Beckman R.J.; and Marathe M.V., 2012. *Sensitivity Analysis of an Individual-Based Model for Simulation of Influenza Epidemics*. *PLOS ONE*, 7, no. 10, 1–16. doi:10.1371/journal.pone.0045414.

Rao D.M.; Chernyakhovsky A.; and Rao V., 2009. *Modeling and analysis of global epidemiology of avian influenza*. *Environmental Modelling & Software*, 24, no. 1, 124–134. ISSN 1364-8152. doi:10.1016/j.envsoft.2008.06.011.

USBLS, 2015. *U. S. Bureau of Labor Statistics: Employment status of the civilian noninstitutional population by age, sex, and race*. URL `https://www.bls.gov/cps/cpsaat03.htm`.

USCB, 2011. *U. S. Census Bureau: Statistical Abstract of the United States*. URL `https://www2.census.gov/library/publications/2011/compendia/statab/131ed/tables/pop.pdf`.

USCB, 2012. *U. S. Census Bureau: Number of Firms, Number of Establishments, Employment, Annual Payroll, and Estimated Receipts by Enterprise Employment Size for the United States, All Industries, Establishments the United States*. URL `http://www2.census.gov/econ/susb/data/2012/us_6digitnaics_2012.xls`.

USCB, 2014. *U. S. Census Bureau: Means of transportation to work*. URL `http://factfinder.census.gov/bkmk/table/1.0/en/ACS/14_1YR/B08301`.

A DOMAIN SPECIFIC LANGUAGE FOR COMPLEX DYNAMIC DECISION MAKING

Souvik Barat
Tata Consultancy Services Research
Pune, India

Vinay Kulkarni
Tata Consultancy Services Research
Pune, India

Tony Clark
Sheffield Hallam University
Sheffield, United Kingdom

Balbir Barn
Middlesex University
London, United Kingdom

KEYWORDS
Organisational decision making, Conceptual model, Simulation model, Domain specific language.

ABSTRACT

Effective decision making of organisation requires deep understanding of various organisational aspects such as its goals, structure, business-as-usual operational processes in the context of dynamic, socio-technical and uncertain business environment. Decision making approaches adopt a range of modelling and analysis techniques for effective decision making. The current state-of-practice of decision-making typically relies heavily on human experts using intuition aided by ad-hoc representation of an organisation. Existing technologies for decision making are not able to represent all constructs that are needed for effective decision making nor do they comprehensively address the analysis needs. This paper proposes a meta-model to represent organisation and decision artifacts in a comprehensive, relatable and analysable form that serves as a basis for a domain specific language (DSL) for complex dynamic decision making. The efficacy of the proposed meta-model as regards specification and analysis is evaluated using a real-life scenario.

INTRODUCTION

Modern organisations need to meet their stated goals by adopting appropriate courses of action in increasingly dynamic environment subjected to a variety of change-drivers. It calls for the precise understanding of various aspects such as organisational goals, organisation structure, operational processes and past data (Shapira 2002). The large size of modern enterprises where the necessary information is both heterogeneous and distributed make compilation of information and subsequent analysis difficult. Furthermore, the socio-technical nature of enterprise (McDermott et al. 2013), inherent uncertainty (Rumsfeld 2011), non-linear causality in business interactions, and high business dynamics chiefly contribute towards organisational decision making being a complex dynamic decision making (CDDM) endeavour.

The common industry practice of organisational decision-making relies on human experts who typically use tools such as spreadsheets, word processors, and diagram editors (Locke 2009). Though adequate for capturing and collating the required information, these tools provide limited analysis support thus putting the decision making onus solely on intuition and interpretation by human experts. Moreover, these tools can only capture a static snapshot of enterprise and are largely devoid of the capability of supporting the dynamism. As a result, decision making with these tools tends to be time-, effort- and intellectually-intensive.

The state-of-the-art specification and analysis techniques approach the decision making problem in two ways namely, data-centric approach and model-centric approach. The data-centric approach makes use of sophisticated AI-based pattern recognition and predictive analysis techniques on relevant past data to predict future outcomes. This approach has worked well when the past data is comprehensive and the future is typically a linear extrapolation of the past. However, the two conditions are increasingly not being met for modern large enterprises thus leading to inappropriate decisions for emerging business contex (HBR 2014).

The model-centric approaches, in contrast, characterise the real organisation using representative models which span across a wide spectrum. At one extreme of the spectrum are enterprise specifications that provide a well-defined structure for the organisational aspects of interest and rely on a variety of visualisation techniques to help humans gain the desired understanding of the organisation. For instance, ArchiMate (Iacob et al. 2003) is one such specification. At the other extreme of the spectrum are machine interpretable and simulatable specifications such as i* (Yu et al. 2005), BPMN (OMG 2011), and System Dynamics (SD) model (Meadows and Wright 2008). Principally they adopt reductionist view (Beckermann 1992) to help analyse enterprises where the mechanistic world view holds. On the other hand, the languages and specifications advocating actor model of computation (Hewitt 2010) and agent-based systems (Macal and North 2010) support emergentism (O'Connor and Hong 2002) through bottom-up simulation. They fare better in analysis of systems with socio-technical elements.

However, the above mentioned techniques and technologies capture only a fragment of what ought to be captured and analysed for effective CDDM (Kurt et al. 2016). For example, enterprise modeling languages are incapable of specifying uncertainty as well as emergent behaviour (Barat et al. 2016), and actor/agent languages are inadequate to express complex goal structure, organisational hierarchies, and behavioural uncertainty in a relatable form (Bonabeau

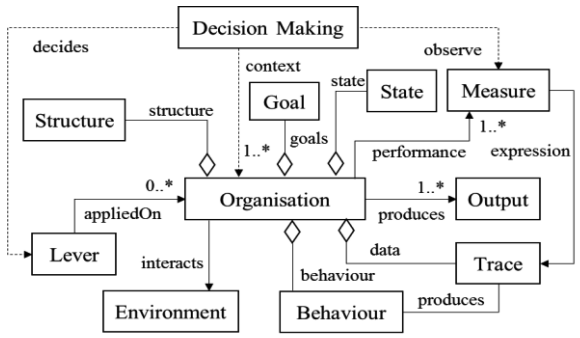


Figure 1: Structural description of decision making

2002). Moreover, as none of the EM specifications and actor based languages are designed for decision making purpose they are found lacking in expressing the necessary decision making concepts in an intuitive and closer-to-the-problem manner (Bonabeau 2002, OMG 2016).

This paper discerns a structure on the required information for CDDM and presents a conceptual meta-model, OrgML, to better support CDDM so as to overcome some of the present limitations. The OrgML is capable of specifying the relevant aspects of organisation and their inter-relationship in a closer-to-the-problem-domain and yet formal machine interpretable form. The key research contributions of this paper are three-fold: (i) a meta-model that represent the structure on the required information for CDDM, (ii) definition of OrgML, *i.e.*, a DSL that captures the above structure in a precise form, and (iii) a systematic derivation mechanism from OrgML to a simulatable form. We claim that the proposed structured representation of decision making problem is an advancement over recent standardisation initiative on Decision Model and Notation (OMG 2016). We also claim that OrgML as a Domain Specific Language (DSL) for CDDM is an advancement over existing enterprise modeling and actor languages for the very same purpose.

The rest of the paper is structured as follows: section 2 describes the modelling and analysis requirements of CDDM, and it briefly evaluates the state-of-the-art techniques and technologies for the same. Section 3 introduces OrgML, establishes the conceptual relationships with foundational concepts, and proposes an approach to use OrgML effectively. Section 4 presents the validation of our claims through a case-study from real life; this section also demonstrates how OrgML is an advancement over existing EM techniques, actor languages, and some of our earlier work. Section 5 provides conclusions and future work.

PROBLEM FORMULATION

This section formulates a structure on the required information for CDDM inspired by some of the decision making models from management sciences literature, and evaluates the state-of-the-art of modelling and analysis techniques for supporting this structure thus establishing a background for our research.

CDDM Structure and Requirements

Table 1: Requirements of CDDM

	Requirement	Description
Aspect	Why	Goals, objectives and intentions of multiple stakeholders
	What	Structural Specification with complex hierarchy and interactions
	How	Behavioural specification with interactions
	Who	Stakeholders and human actors of the system
	Where	Information about location
	When	Temporality in behaviour and adaptation
Socio-technical Characteristics	Modular	A system can be decomposed into multiple parts.
	Compositio-nal	Multiple parts should be composed to a consistent whole.
	Reactive	Must respond appropriately to its environment
	Autonomou s	Possible to produce output without any external stimulus.
	Intentional	Intent defines the behaviour
	Adaptive	Adapt itself based on context and situation
	Uncertain	Precise intention and behaviour are not known a-priori.
	Temporal	Indefinite time-delay between an action and its response
DC	Measure	Ability to specify what needs to be measured
	Lever	Ability to specify possible courses of action
Analysis	Machine Interpretable	Models that are interpretable by machine (*i.e.,* support for simulation/execution)
	Top-down and Bottom-up	Support for top-down and bottom-up modelling and simulation to support reductionist view and emergentism

The philosophical basis of our solution is largely inspired by the decision making models from management sciences such as rational model (Simon 1955), Incremental model (Cyert and March 1992), Carnegie model (Mintzberg et al. 1976) and Garbage Can (Cohen et al. 1972). Though adopting different methodological styles, these models agree on the core concepts of decision making namely, objective or *goal*, course of action or *lever*, and performance indicator or *measures* (Yu 2012). Further they rely on *contextual information* as the basis to analyse achievability of a goal or efficacy of a lever for achieving goal.

Therefore, we argue that the activity of decision making largely depends on two key factors: (i) the ability to capture the core decision making concepts and contextual information in a formal manner, and (ii) the ability to perform *what-if* and *if-what* analyses on the information captured. The former requires completeness and the latter expects the efficacy. We argue that comprehensive information about six interrogative aspects namely *why*, *what*, *how*, *when*, *where*, and *who* as recommended in Zachman framework (Zachman 1987) ensures completeness, and a suitable processor, say a simulator or an interpreter, of the specification ensures efficient and effective analysis.
The class diagram in Figure 1 overlays a structure on the information necessary for CDDM. It depicts the relevant concepts borrowed from management sciences namely, *Goal*, *Measure* and *Lever*, and the concepts necessary to capture contextual information namely, *Goal*, *Structure*, *State*, *Trace*, *Behaviour* and *Environment*.

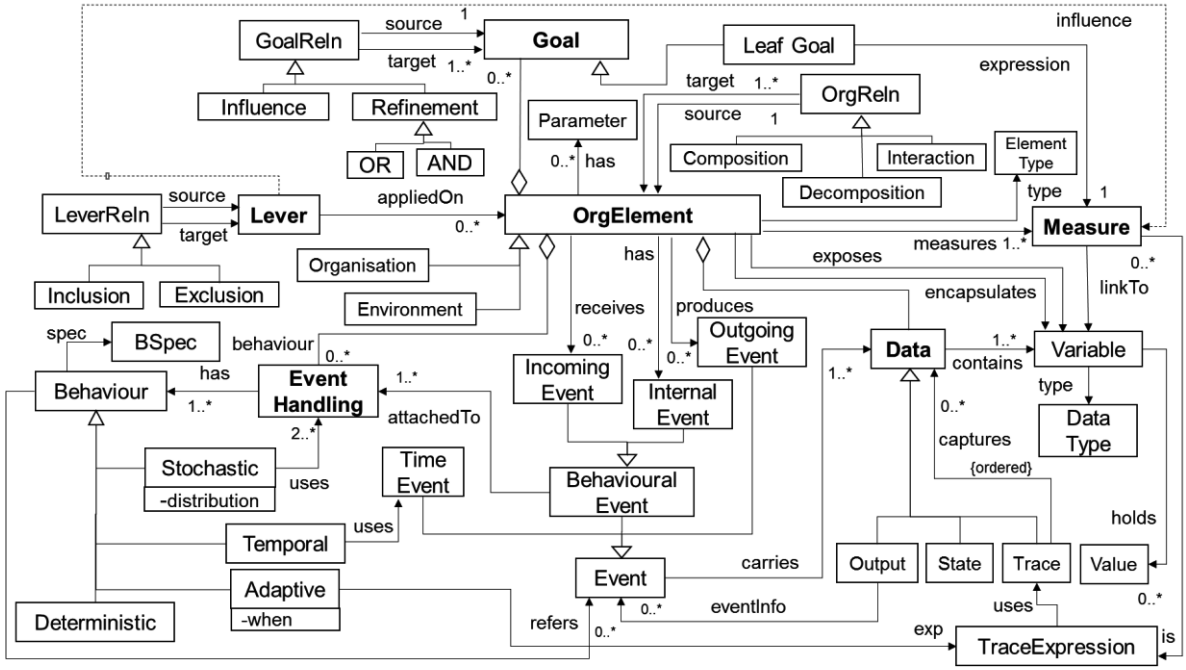


Figure 2: OrgML – a metamodel to represent organisation

An *Organisation* relies on its Structure and Behaviour to produce Output so as to achieve the stated Goals while operating in an Environment. Behaviour induces State changes thus producing Trace (i.e. historical record comprising of State and Output) over a period of time. Goal is a conditional expression over Measures which are views over Trace. A Lever is possible modification to Goal and/or Behaviour, and/or Trace.

The complex dynamic decision making (CDDM) for large and complex organisation with volatile operating environment calls for additional requirements on the specification described in Figure 1. A large organisation often contains complex hierarchy with large number of socio-technical elements as part of its structure. The constituent socio-technical elements are mostly autonomous, reactive and goal-directed. Also, the organisation elements exhibit uncertainty, temporality and adaptability. Thus, complex structure, socio-technical characteristics, and inherent uncertainty form additional specification requirements for CDDM. The specification requirements also demand the necessary constructs to represent decision making related concepts namely Goal, Measure, Lever. A-priori assessment of decisions is suggestive of simulation capability. Further, the simulation can be approached using *top-down* or *bottom-up* manner. Therefore a CDDM may need any of these two approaches or it may demand a middle-out approach. Table 1 enumerates the specification and analysis needs for CDDM.

Review of state of the art and practices

In (Barat et al. 2016), we systematically evaluated the suitability of EM techniques to support CDDM. The evaluation concluded with a critical observation that the existing EM techniques are capable of satisfying the expected requirements of CDDM described in Table 1 only in parts. In particular, we found the EM techniques that support necessary aspects of CDDM (such as Zachman Framework (Zachman 1987) and ArchiMate (Iacob et al. 2003)) are not machine interpretable and thus not amenable for rigorous analyses. Similarly, prevalent general purpose conceptual enterprise model, MEMO (Frank 2002), supports most of the specification needs except decision making related constructs. In contrast, specifications capable of precise analyses, such as BPMN (OMG 2011), i* (Yu et al. 2006) and System Dynamic (SD) (Meadows and Wright 2008) models, on their own, are not capable of representing all necessary aspects. For instance, BPMN analyses and simulates the process aspect, i* analyses the high level goals and objectives, and SD model simulates complex dynamic behaviour of the system. On the other hand, the multi-modelling and co-simulation environments, such as DEVS (Camus et al. 2015) and AA4MM (Siebert et al. 2010), collectively support the analysis needs for all aspects depicted in Table 1. However, they are not capable of expressing many socio-technical characteristics such as autonomy, uncertainty and temporal behaviour. Moreover, they are not suitable for bottom-up construction of a system that results into emergent behaviour.

The general purpose actor languages and frameworks, such as Scala Actors (Haller and Odersky 2009) and Akka (Allen 2013), are capable of specifying and analysing a range of socio-technical characteristics and emergent behaviour. However, we find the effective use of these general purpose languages in the context of decision making is a hard proposition as they do not have the language support to represent goal, measure and lever explicitly. Moreover, a large and complex organisation may need both the top-down and bottom-up analysis (Bonabeau 2002), which is not supported in general purpose actor languages.

PROPOSED SOLUTION - OrgML

This section presents the core contribution of this paper. It first introduces OrgML meta-model, then conceptually correlates the model with requirements of CDDM, and proposes an approach on using OrgML for CDDM.

OrgML meta-model

OrgML is a meta-model to represent the information required for decision making in a structured and machine interpretable form as shown in Figure 2. It extends the concepts depicted in Figure 1 along two dimensions – (i) to capture the specification requirements described in Table 1, and (ii) to enable the top-down / bottom-up modelling and simulation.
The core element of OrgML is *OrgElement*, which is a parametric entity that can have: a set of *Goals* to represent its intention or objective, a set of *EventHandling* units to represent behaviour, and *Data* to capture *state* and *trace* of OrgElement. An OrgElement is an event-centric abstraction with *<eventName, eventHandlingSpecification>* tuples constituting its behaviour. Amongst the events being processed, *IncomingEvent* (*i.e.*, events received) and *OutgoingEvent* (*i.e.,* events produced) are exposed to the environment whereas *InternalEvents* are not. The OrgElement Data is a set of typed *Variables* that can hold Values. An OrgElement may expose Variables to other OrgElement thus relaxing data hiding. An OrgElement can have a set of *Measures* to represent key performance indicators. Thus, OrgElement is the basic building block for specifying organisations for decision making. Detailed description of OrgML follows:

Goal: OrgML enables a goal to be decomposed into sub-goals, sub-sub-goals etc. to the desired level of detail. At each level, a goal can be related to other goals through relationships, for instance, meeting a goal g1 disables the related goal g2 to be met, or meeting a goal g1 subsumes the related goal g2 etc. We group such dependency relationships under *Influence* kind of relationship. The other kind of relationship is *Refinement* which enables hierarchical decomposition of goals. OrgML supports all the key relationships from i* goal specification (Yu et al 2006). A *LeafGoal* is a conditional expression over Measures. The conditional expression language includes temporal operators. As Measures have values, it is possible to compute if a LeafGoal is met or not. The semantics of Influence and Refinement relationships determine evaluation of the overall goal as a bottom-up traversal of the graph.

Structure: An OrgElement supports structural composition, decomposition and interactions through *OrgReln*, and represents type using *ElementType*. Organisational units, stakeholders and system of an organisation can be represented using ElementType. One can also consider the *Active-* and *Passive-* Structure defined in ArchiMate (Iacob et al. 2003) as the basis for type definition. Further, we consider Organisation and its *Environment* are two specialised OrgElement. An Organisation or Environment cannot be composed with other OrgElement. However, they can interact with other OrgElement and decompose into finer level of granularity. For example, an Organisation can interact with Environment, Organisation can be decomposed into Organisational units, an Environment of an Organisation can be visualised as multiple competitors, *etc*.
Data: Data of OrgElement is of three kinds namely *Output*, *State*, and *Trace*. The *Output* represents outcome of executing an OrgElement. It represents: (i) Variables of the

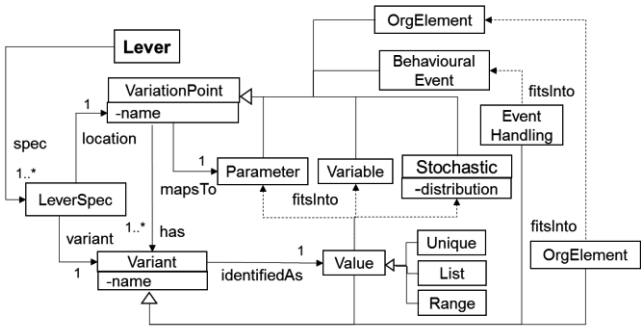


Figure 3: Lever specification of OrgML

OrgElement , (ii) Events produced internally, and (iii) Events communicated to other OrgElements. The *State* represents Variables of the OrgElement. *Trace* is collection of Data of the OrgUnit from a point in time in the past till now.

Behaviour: Elements Event and EventHandling describe the behaviour of an OrgElement. The Event can be classified into three categories namely *OutgoingEvent*, *TimeEvent*, and *BehaviouralEvent*. The *OutgoingEvent* specifies the Data from OrgElement accompanying the Event. The *TimeEvent* is an Event that represents Time information. The *BehaviouralEvent* specifies the Event definition and its implementation using EventHandling element. The BehaviouralEvent is further classified into IncomingEvent and InternalEvent wherein the former represent the Events that are consumed by the OrgElement and the latter represent the Events that are internal to the OrgElement. The EventHandling element is the primitive behavioural unit for describing the Behaviour of an OrgElement which can be of four kinds namely *Deterministic*, *Stochastic*, *Temporal* and *Adaptive*. We consider standard language constructs such as *assignment*, *expression evaluation*, *loop*, *recursion*, *message passing*, *etc.*, to express Deterministic Behaviour. The Stochastic Behaviour specifies the uncertainty along two dimensions – uncertainty in raising an OutgoingEvents and uncertainty in responding to an IncomingEvent. Specification of both requires use of special constructs providing probability distribution guided choice. The Temporal Behaviour uses TimeEvent to express temporal relationships within behavioral specification. Adaptive Behaviour describes adaptation rules. Essentially, it express a Behavior that activates when a specific condition is matched – it uses TraceExpression, *i.e.*, an expression over Trace element, to define the conditions. We consider element *BSpec* as a placeholder for behavioural specification.

Measure: *Measure* are a set of *TraceExpression* that essentially represents Variables of interest.

Lever: *Lever* represents possible courses of action that can be applied on OrgElement. A lever specification contains two kinds of specification: (i) lever usage specification and (ii) lever definition. Lever usage specification is illustrated in Figure 2 using *LeverReln* and its specialisation. The Lever inclusion and exclusion relationships can be defined using *LeverReln*.

The lever definition specifies a modification to either structure or data or behaviour of an OrgElement or a combination thereof. We adopt the concept of *Variation*

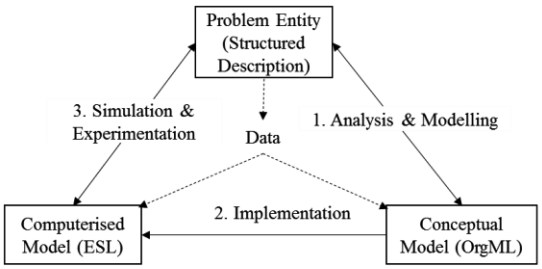


Figure 4: Overview of modelling and simulation approach

Point and *Variant* of variability modelling (Kulkarni 2012) to define lever specification as depicted in Figure 3. Essentially, a lever is set of *LeverSpec* where each LeverSpec describes the change specification using two named elements namely *VariationPoint* and *Variant*. The VariationPoint describes the location of a change, and the Variant describes the changed element. A predefined set of core elements of OrgML can act as VariationPoint and Variant. Further, there is a notion of compatibility between VariationPoint and Variant. The element Parameter, Variable, Stochastic Behaviour, Behavioural Event and OrgElement of OrgML can act as VariationPoints wherein the element Value can fit into Parameter, Variable, Stochastic Behavior; the element EventHandling can fit into BehaviouralEvent; and an OrgElement can fit into OrgElement. We consider VariationPoint as a Parameter, and Variant as a Value to realise fitsInto relationship.

Analysis of OrgML as a decision making aid

Conceptually, the elements of OrgML refines the structure defined in Figure 1 and enables the characteristics described in Table 1. Event definition, Data, and OrgElement structure specify the *what* aspect, OrgElement help specify the *who* aspect, Goal specification specifies the *why* aspect, and Behaviour specifies the *how* and *when* aspects. The concept of OrgElement ensures desired modularity and encapsulation: the Event helps to specify reactive nature, InternalEvent and TimeEvent collectively specify the autonomous behaviour, Stochastic Behaviour helps in specifying required uncertainty, the Temporal Behaviour and TimeEvent specify the temporal behaviour, and Adaptive Behaviour is capable of specifying the adaptive nature of an OrgElement. We argue that the *Composition* relationship of OrgElement and *Influence* relationship of Goal specification together help in bottom-up design, whereas the *Decomposition* relationship of OrgElement, *Goal Refinement Relationship*, and an ability to share Variables using *exposes* relationship help in top-down design.

The proposed meta-model is grounded with a set of existing concepts. The modularisation and unit hierarchy are taken from component model concepts. Goal-directed reactive and autonomous behaviour can be traced to actor behaviour (Hewitt 2010). Defining states in terms of a type model is borrowed from UML. An event driven architecture is introduced for reactive behaviour. The concept of intentional modelling (Yu et al. 2006) is adopted to enable specification of goals. The behavioural classification and uncertainty is defined from the *uncertainty theory* by Donald Rumsfeld (Rumsfeld 2011) wherein the *known knowns* behaviour can be specified using Deterministic Behaviour and the *known unknowns* (KU) can be specified using Stochastic Behaviour.

Table 2: Guideline for constructing OrgML model from existing specifications

OrgML Concept	Possible sources (concept mapping from existing languages)
OrgElement	UML Class Diagram:: Class that represents Organisational elements such Organisation, Organisational Unit, Environment. ArchiMate:: Business Actor, Business Role, Business Object, Application Component, System Software
Data	UML Class Diagram:: Class that represents entities ArchiMate:: Data Object, Artifacts. BPMN:: Data Object
Goal	i* specification:: Goal. ArchiMate:: Meaning
Behaviour	UML State Machine:: State, Transition ArchiMate:: Business Service, Business Process, Business Function, Application Function, Infrastructure Function. BPMN:: process definition
Event	UML State Machine: Transition. BPMN:: Event. ArchiMate:: Business Interaction, Business Event
Measure	i*:: Task, Leaf level Goal. BPMN:: KPI
Lever	Description about possible courses of action

Enabling CDDM using OrgML

We adopt a simplified modelling method recommended by Robert Sargent in (Sargent 2005) to capture organisation specification using OrgML and enable required analysis. Method uses three distinct representations namely problem entity, conceptual model and computerized model, to systematically transform a real-life problem into analyzable model and perform analysis/simulation as shown in Figure 4. The problem entity is a description of real environment, conceptual model is a purpose specific representation of the problem entity, and computerised model is an executable/simulatable model of conceptual model. In our approach, we consider textual description, i*, class diagram, state-machines, BPMN, and ArchiMate as the possible specification aids to describe a problem entity, OrgML is a specification aid for conceptual model, and an actor-based language named as *Enterprise Simulation Language* (ESL) (Tony et al. 2017) as computerised model.

ESL, like standard actor languages (Haller and Odersky 2009, Allen 2013), supports the notion of *event*, *state*, *reactive*, *autonomous* and *adaptive* behavior. ESL supports further concepts such as *Time*, *Stochastic* behaviour (using *probability* distributions) and *Temporal* behavior as shown in

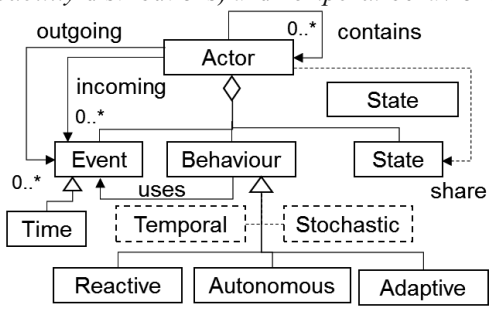


Figure 5: Characteristics of ESL

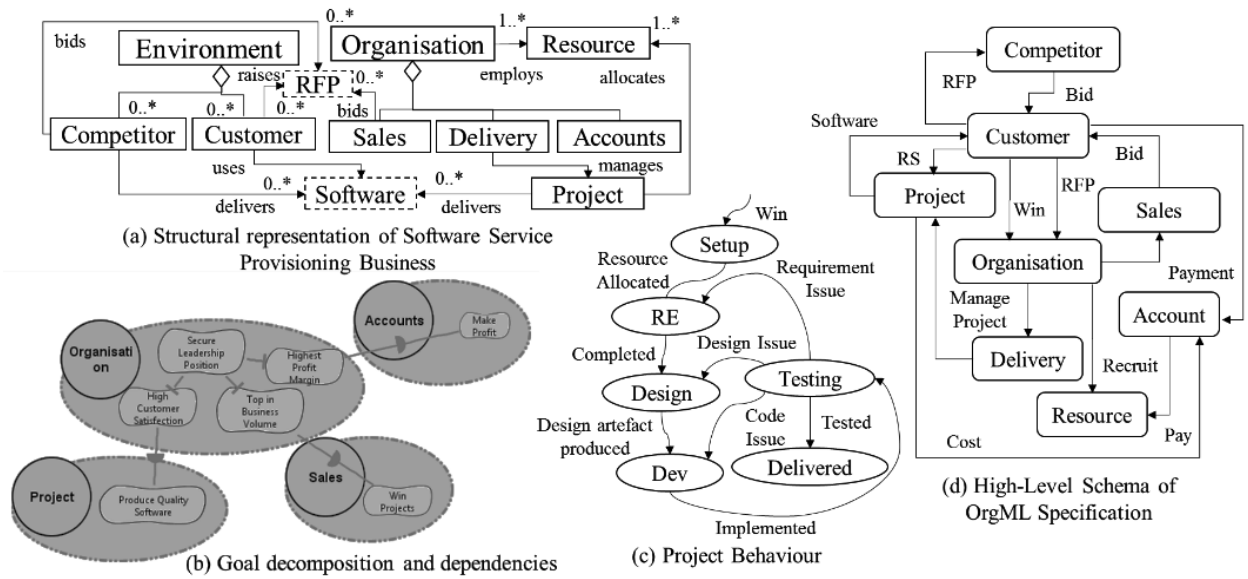


Figure 6: Specifications of Software Service Provisioning Organisation Case Study

Figure 5. These additions (depicted with dotted box and lines in Figure 5) are beneficial in making ESL a transformation target for OrgML.

Step 1 *(Analysis and Modelling)* manually converts a *problem entity* specification into an OrgML specification by identifying primitive elements, such as *Organisation*, *Organisational Unit*, and *Stakeholders*, and specifying their goals, behaviour, and measures. For example, the goals represented using i* model can be translated into the OrgML goal specification wherein the *'decomposition'* relationships of i* can be translated into OrgML *Refinement* relationships, A basic guideline for constructing OrgML specification from existing specification languages is depicted in Table 2.

Step 2 *(Implementation)* uses a fixed set of transformation rules to transform OrgML specification into an ESL specification and is depicted in Table 3. We use a java based ESL simulator to simulate converted ESL specifications. A simulation of transformed specification progresses with primitive Time events (of ESL) wherein each translated OrgElement performs its own behaviour by reacting to IncomingEvents and InternalEvents. It updates state variables, persists trace information, produce OutgoingEvents, and evaluates goal expressions. An OrgElement adapt to new behaviour when adaptation logic is satisfied. The simulator displays identified measures in the form of graphics and animation.

EVALUATION

We present an evaluation of OrgML and the proposed modelling approach using an established real-life case study of a software service-provisioning organisation (Kulkarni et al. 2015a, Kulkarni et al. 2015b)

The case study, primarily, focuses on the decision making of a *Software Service-Provisioning Organisation* (SSPO) that aims to *Secure Leadership Position* by focusing on three sub-goals namely *High Customer Satisfaction*, *Top in Business Volume* and *Highest Profit Margin*. The SSPO aim to achieve its goals by provisioning high-quality bespoke software in a dynamic and competitive environment that contains multiple customers who outsource their development activities to service provisioning organisations, and a set of competitors who have similar goals as SSPO.

Internally, SSPO contains three autonomous organisational units namely, *Sales*, *Delivery*, and *Account*, a dynamic organisational unit, terms as *Project*, which is formed on demand, and a set of skilled *Resources*. A simplified structure of the case study is depicted using a class diagram in Figure 6 (a). Each organisational unit of SSPO has its goals that contribute to the organisational goal as shown in Figure 6(b); and each organisational unit has its own behaviour to achieve their individual goals. A simplified behaviour of Project unit is illustrated using a state-machine in Figure 6 (c). In general, Customers raise *request for proposal* (RFP) for software provisioning; SSPO organisation and *Competitors* bid RFPs; and Customers evaluate bids and select one organisation for service provisioning. Once a bid is won by SSPO, it forms a development *Project* by allocating appropriate *Resources*, executes the project using standard software development process (as depicted in Figure 6 (c)), and finally delivers *Software* to the respective *Customer*.

Table 3. OrgML to ESL Transformation Rule

OrgML	ESL
OrgElement	Actor
Data	Actor Variables
Goal	Expression over Actor variables
Event	Event
EvenHandling	Event Handling specification
Measure	Expression over Actor variables
Lever	ESL specification
Parameter	Input parameter of ESL Actor
Trace	Actor Variable
Deterministic	Standard Behavioural specification
Stochastic	Probabilistic Behaviour
Temporal	Behaviour with temporal expression
Adaptive	Behavioural with guarded expression
Deterministic	Behavioural specification

The case study explores the possibility of achieving high level goal, *i.e.*, *Secure Leadership Position*, and sub-goals *i.e.*, *High Customer Satisfaction*, *Top in Business Volume* and *Highest Profit Margin*, of SSPO for a given environment that contains a set of Customers with varying outsourcing needs, and a set of Competitors with specialised bidding strategies and project execution strategy. The case study further explores various Levers, such as a *competitive bidding strategy*, *increase of resource strength*, *recruitment of skilled resources*, and *reskilling of existing resources*, as possible options to increase the possibility of achieving its goals.

In our evaluation, we considered an extended form of SSPO specification depicted in Figure 6 (a)-(c) as the *problem entity* specification (extended figures are not included due to space limitation). In our approach, we first constructed OrgML specification of SSPO from problem entity specification using the guideline described using Table 2. An OrgML specification specify SSPO, organisational units (*i.e.*, Sales, Delivery, Account, and Projects), and the identities that describes Environment (*i.e.*, Competitor and Customer) as parameterised OrgElement. The goals (as depicted in i* model) are translated into the OrgML goal specification. Three Measures are identified from LeafGoals. The Measures are: *Customer Satisfaction Index*, *Business Volume* and *Profit Margin*. The behavioural specifications that are represented as state-machines are translated into OrgML Events and Event specifications. A schema of converted OrgElements and their interactions of constructed OrgML specification is depicted in Figure 6 (d). Next, we converted SSPO OrgML specification into SSPO ESL specification using transformation rules depicted in Table 3 to simulate SSPO using ESL simulator and validated simulation results with earlier experimentation presented in (Kulkarni et al. 2015a).

We used SSPO case study to validate multiple enabling techniques and technologies that have potential to serve as effective aids for CDDM. In (Kulkarni et al. 2015b), we evaluated a multi-modelling and co-simulation approach that used three prominent EM techniques namely i*, BPMN and Stock-and-Flow in a coordinated manner. In (Kulkarni et al. 2015a) we experimented ESL to encode SSPO case study and evaluate decision alternatives. This paper proposes OrgML supported with a method as an affective modelling and analysis aid for CDDM. The effectiveness of four alternatives, i.e., EM based approach (from (Kulkarni et al. 2015b)), pure actor language based approach (from existing literature such as (Bonabeau 2002), ESL based approach (from (Kulkarni et al. 2015a)), and OrgML based approach are summarised in Table 4. As shown in the table, an EM based approach and an actor language based approach are complementary in nature. The former one supports aspect (*i.e., why*, *what*, *how*, *etc.*) specification and a top-down simulation approach, whereas actor language based approach is more effective for representing socio-technical characteristics and bottom-up simulation approach. But, it is not convenient for aspect specification. ESL bridge the gaps between two class of specifications with explicit support for uncertainty, temporal behaviour, and the bottom-up and top-down combination. Therefore we argue that ESL is technically complete for CDDM requirements described in Table 1. However, ESL is a general purpose simulation language and it is not convenient to specify most of the aspects (*e.g.*, *why*, *who*, *where*, *when*), and decision making constructs namely *goal*, *measure* and *lever*. Hence OrgML is further improvement towards the infrastructure that we envision for CDDM. It helps in expressing the most of the requirements in a convenient and machine interpretable for leading simulation based CDDM.

Table 4. Evaluation Summary

Requirement	EM Spec.	Actor Lang.	ESL	Org ML	Enabling OrgML Concepts
Why	√	⊥	⊥	√	Goal
What	√	√	√	√	OrgElement
How	√	√	√	√	EventHandling
Who	√	⊥	⊥	√	OrgElement
Where	√	⊥	⊥	⊥	OrgElement
When	√	⊥	⊥	⊥	Time Event
Modular	√	√	√	√	OrgElement
Compositional	⊥	√	√	√	Composition Relationship
Reactive	⊥	√	√	√	IncomingEvent, OutgoingEvent
Autonomous	×	√	√	√	InternalEvent
Intentional	√	√	√	√	Goal
Adaptive	⊥	√	√	√	Adaptive Behaviour
Uncertainty	×	⊥	√	√	Stochastic Behaviour
Temporal	⊥	×	√	√	Temporal Behaviour
Measure Spec	⊥	⊥	⊥	√	Measure
Lever Spec	⊥	⊥	⊥	√	Lever
Top-down/ Bottom-up	Top-down	Bottom-up	Hybrid	Hybrid	Composition Relationship, Shared State Variable
Legends: √ : Supports adequately, ⊥ can be specified with difficulties, × : not supported					

CONCLUSION

Currently, there is a gap in technologies available for decision making notably in the precision of aspects such as precision of organisational goals, structure, operational processes, environment and expressing uncertainty. To address this gap, we have presented OrgML - a meta-model structure over information required for decision making. The model content is inspired by key concepts advocated in various decision making models from management sciences. The formal structure is achieved by integrating appropriate concepts from Enterprise Modelling, actor model of computation, uncertainty theory and variability modeling.

A rationale for how OrgML can enable a decision making approach that can potentially overcome some of the limitations of current state of art and practice of CDDM has been presented. The principal benefits are derived from an extended form of actor model of computation; is composable; is capable of specifying uncertainty in behavior; and is simulatable. A systematic approach on how to derived an simulatable specification from an OrgML model instance using a model map has been outlined. The derivation process and the efficacy of the OrgML model was evaluated using a real-life scenario. We acknowledge this paper does not

discuss in detail the language constructs of ESL, the transformation rules from problem entity specification to OrgML, and automatable transformation rules to translate OrgML to ESL. We restricted the principal objectives of this paper to: establishing the core concepts of CDDM, defining a specification language that can act as an effective aid for specifying organisations for CDDM and enabling a simulation based approach to CDDM.

Our research is an example of technical action research (Wieringa and Moralı 2012) in that it presents a validation of a design science artifact wherein we have demonstrated how OrgML and the proposed approach is relevant and effective for practitioners to adopt in CDDM. As part of future research, we intend to further validate this in real business scenarios as well as proposing further extensions to OrgML for introducing game theoretic approach in simulation.

REFERENCES

Allen, J. 2013. Effective Akka. O'Reilly Media.

Barat, S., Kulkarni, V. Clark, T. and Barn, B. 2016. "Enterprise modeling as a decision making aid: A systematic mapping study." The Practice of Enterprise Modeling, pp. 289-298.

Beckermann, A., Hans F., and Jaegwon K. 1992. "Emergence or reduction?: Essays on the prospects of nonreductive physicalism". Walter de Gruyter.

Bonabeau, E. 2002. "Agent-based modeling: Methods and techniques for simulating human systems." Proceedings of the National Academy of Sciences 99, no. suppl 3: 7280-7287.

Camus, B.; Bourjot, C.; Chevrier, V. 2015. "Combining DEVS with Multi-agent Concepts to Design and Simulate Multi-models of Complex Systems." <hal-01103892>

Clark, T.; Kulkarni, V.; Barat, S.; Barn, B. 2017. "ESL: An actor-based platform for developing emergent behaviour organisation simulations". In: International Conference on Practical Applications of Agents and Multi-Agent Systems. pp. 311–315.

Cohen, M. D.; March, J. G.; Olsen, J. P. 1972. "A Garbage Can Model of Organisational Choice". Administrative Science Quarterly 17 (1): 1–25.

Cyert, R.; March, J. G. 1992. A Behavioral Theory of the Firm (2 ed.). Wiley-Blackwell. ISBN 0-631-17451-6.

Frank, U. 2002. "Multi-perspective enterprise modeling (memo) conceptual framework and modeling languages." In System Sciences, 2002. HICSS. pp. 1258-1267.

Haller, P. and Odersky, M., 2009. "Scala actors: Unifying thread-based and event-based programming". Theoretical Computer Science, 410(2), pp.202-220.

HBR 2014, "9 Habits That Lead to Terrible Decisions." https://hbr.org/2014/09/9-habits-that-lead-to-terrible-decisions.

Hewitt, C., 2010. "Actor model of computation: scalable robust information systems". arXiv preprint arXiv:1008.1459.

Iacob M, et al. 2003. "State of the art in architecture support, ArchiMate deliverable D3.1." Enschede, The Netherlands: Telematica Instituut.

Kulkarni, V.; Barat, S.; and Roychoudhury, S. 2012. "Towards business application product lines." Model Driven Engineering Languages and Systems, pp. 285-301.

Kulkarni, V.; Barat, S.; Clark, T.; Barn, B. 2015a. "Using simulation to address intrinsic complexity in multi-modelling of enterprises for decision making." In Proceedings of the Conference on Summer Computer Simulation, pp. 1-11.

Kulkarni, V.; Barat, S.; Clark, T.; Barn, B. 2015b. "Toward overcoming accidental complexity in organisational decision-making." Model Driven Engineering Languages and Systems (MoDELS 15).

Kurt, S. et al. 2016. "Enterprise Modelling for the Masses–From Elitist Discipline to Common Practice.". The Practice of Enterprise Modeling, pp. 225-240.

Locke, E. 2009. Handbook of Principles of Organisational Behavior: Indispensable Knowledge for Evidence-Based Management. ISBN: 978-0-470-74095-8

Macal, C.M. and North, M.J., 2010. "Tutorial on agent-based modelling and simulation". Journal of simulation, pp.151-162.

McDermott, T.; Rouse, W.; Goodman, S.; Loper, M. 2013. "Multi-level Modeling of Complex Socio-Technical Systems." Conference on Systems Engineering Research (CSER'13).

Meadows, D.H. and Wright, D. 2008. Thinking in systems: A primer. chelsea green publishing.

Mintzberg, H.; Raisinghani, D.; and Theoret, A. 1976. "The structure of unstructured decision processes". Administrative Science Quarterly, 21, 246-275

O'Connor, T., and Hong Y. W. 2002. "Emergent properties."

OMG. 2011. Business Process Model and Notation. http://www.omg.org/spec/BPMN/2.0/, formal/2011-01-03

OMG. 2016. Decision Model and Notation (DMN), www.omg.org/spec/DMN/1.1/

Rumsfeld, D. 2011. Known and unknown: a memoir. Penguin.

Sargent, R.G., 2005. "Verification and validation of simulation models". In Proceedings of the 37th conference on Winter simulation (pp. 130-143).

Shapira, Z. 2002. Organisational Decision Making. Part of Cambridge Series on Judgment and Decision Making. ISBN: 9780521890502

Siebert, J.; Ciarletta, L., and Chevrier, V. 2010. "Agents and artefacts for multiple models co-evolution: building complex system simulation as a set of interacting models." Autonomous Agents and Multiagent Systems: Volume 1, pp. 509-516.

Simon, H. A. 1955. "A behavioral model of rational choice." The quarterly journal of economics 69, no. 1: 99-118.

Wieringa, R.; Moralı, A. 2012. "Technical action research as a validation method in information systems design science." In International Conference on Design Science Research in Information Systems, pp. 220-238. Springer Berlin Heidelberg.

Yu, E.; Strohmaier, M; Deng, X. 2006. "Exploring intentional modeling and analysis for enterprise architecture". In EDOC Conference Workshops.

Yu, P. 2012. "Multiple-criteria decision making: concepts, techniques, and extensions". Vol. 30. Springer Science & Business Media.

Zachman, J.A. 1987. "A framework for information systems architecture." IBM systems journal, 26(3):276-292.

BIOGRAPHIES

Souvik Barat is senior researcher at Tata Consultancy Services Research (TCSR). He holds a Masters Degree in Computer Science and Engineering from Indian Institute of Technology Madras.

Vinay Kulkarni is distinguished Scientist at Tata Consultancy Services Research (TCSR). He holds a Masters Degree in Electrical Engineering from Indian Institute of Technology Madras.

Tony Clark is Head of the Department of Computing and a Professor of Software Engineering at Sheffield Hallam University in the UK.

Balbir Barn is Deputy Dean and Professor of Computer Science at Middlesex University London

SUPPORT FOR MANAGEMENT PROCESSES OF THE EXERCISES OF THE CRISIS STAFFS OF CRITICAL INFRASTRUCTURE ENTITIES

Jiří Barta and Josef Navrátil
Department of Crisis Management
University of Defence in Brno
Kounicova 65, 662 10 Brno
Czech Republic
E-mail: jiri.barta@unob.cz

KEYWORDS
Computer simulation, computer assisted exercises, CAXManager application, critical infrastructure, crisis scenario, practical exercise.

ABSTRACT

Preparedness of the crisis management authorities is essential to ensure management processes in dealing with emergencies and crisis situations. Due to increasing numbers as well as impacts of anthropogenic and natural disasters, emphasis is placed on the preparedness of public authorities and business entities, particularly critical infrastructure entities, as disruption of their activities would bring a serious impact on the state security system, population, property and environment.
The article focuses on the activities and procedures in dealing with simulated exercises of crisis management authorities at different levels of public administration and critical infrastructure. Emphasis is placed on the use of a simulation tool in the preparation phase of the exercise which is used to support decision-making processes and exercise recording. Exercise recording is subsequently used to evaluate the efficiency and benefits of the exercise.
Based on the experience from previous exercises, the CAXManager software tool has been designed and programmed. Within the exercise it allowes in particular the exercise operator entering, executing and checking the level of fulfilment of scenario tasks. For the participants CAXManager application shows performance in individual tasks, their gradual delegation to individual levels of progress and finishing the task being addressed in the final phase.
Currently, CAXManager software is in the testing phase. Several user-unfriendly situations have been detected during testing which have been recorded and discussed with programmers. Substantial remarks and suggestions have continually being incorporated into the program to improve its user-friendliness and functionality for evaluating the exercise.

INTRODUCTION

The Czech Republic, as a Member State of the European Union, has introduced Council Directive 2008/114/EC (Council Directive 2008) defining critical infrastructure and conditions and principles of its protection. Protection of national critical infrastructure was included in Act No. 240/2000 Coll. On Crisis Management, as amended (Act 2000b). Based on the analyses carried out (Oulehlová 2017; Alcaraz and Zeadally 2015), pressure has been put on the implementation of preventive measures and increase in the preparedness to deal with emergencies or crisis situations emerged or influencing the critical infrastructure bodies (Luděk and Ráček 2011, Řehák et al. 2016).
Due to the extent of potential problems with critical infrastructure bodies, effective cooperation between all stakeholders (primarily the critical infrastructure entities, public administration authorities and components of the integrated rescue system) is necessary in dealing with crisis situation (Ristvej et al. 2016; Vašková et al. 2017). In particular, public authorities and critical infrastructure entities should be prepared for a joint co-ordinated response to emergencies and crisis situations (Hrídel and Karták 2013; Malachová and Oulehlová 2017a).

CURRENT SITUATION ANALYSIS

This chapter deals with the issue of tactical and staff exercise and increase in the level of development of professional education and skills of crisis staff members. Attention is paid both to the possibilities of software support to the exercise as well as to evaluation of the already executed national staff exercise with involvement of all interested parties, including critical infrastructure entities.

Responsibilities in preparation of crisis staffs

Emergency situation preparedness or crisis situation preparedness significantly contribute to providing efficient and rapid response and damage minimization. Acquiring practical skills and knowledge enables successful response to the emerged situation. An integral part of crisis preparedness is continuous and structured education of the relevant staff. Theoretical knowledge should be followed by practical exercise which will contribute to mastering acquired knowledge. According to the aim of the exercise, it can be an improvement, screening, co-operation or testing exercise.
Usefulness and practicality of processed crisis documentation is as well as the human resources tested during exercises. According to the Crisis Act (Act 2000a), liaison officer and company management are responsible for preparation and implementation of practical exercise implementation.

Computer simulation enters the area of staff practical training only at a slow pace. This is probably due to lack of confidence in new technologies as well as relatively high acquisition costs for building complex simulation centers (Barta, Vašková and Urbánek 2016). For this reason, complex simulators for staff practical training or crisis staff training are available, with exceptions, only in military facilities (Hubáček and Řezáč 2013; Hubáček and Vráb 2013; Ristvej et al. 2016).

SIMEX 2016 exercise

In the framework of a research project aimed at software support of practical training of crisis staff members with a focus on energy critical infrastructure, an analysis of possible simulation tools for exercise was carried out (Barta, Vašková and Urbánek 2016; Barta and Řezáč 2016). On the basis of the analysis of the utility properties of the individual simulation tools, SIMEX simulator was created. It belongs to a group of constructive simulators, where participants use simulated equipment and entities in a simulated environment. Constructive simulation is often referred to as "war" because it is similar to strategic war games in which players command individual soldiers, military groups and technique in the playing environment (Hubáček and Vráb 2013; Urbánek et al. 2013).
The SIMEX simulator was employed during practical exercise of crisis management authorities of the South Bohemian Region, subordinated municipalities with extended powers and components of the integrated rescue system. The exercise topic was a large-scale disruption of the natural gas supply due to a simulated accident on the high-pressure gas pipeline in the gas distribution system of the gas company E.ON Distribution. The topic of the exercise was chosen to correspond with a probable event which would require regional level coordination and was approved by the Security Council of the South Bohemian Region (Malachová and Oulehlová 2017b).
The exercise was carried out on 24-25 May 2016 as a multilevel staff exercise at the regional level. The simulated crisis situation was situated to January 2016. Scope of the exercise was set at a level enabling involvement of several critical infrastructure entities. These included especially E.ON Czech Republic, Ltd., E.ON Services and E.ON Distribution which actively participated in dealing with the emergency.

CLASSIFICATION OF SIMULATION APPROACHES

The SIMEX simulator was used for creating the plan of the staff exercise to simulate the intervening activities. SIMEX simulation system consists of software equipment for simulation, terrain database, 3D models and effects, software equipment of the communication system as well as a DIS Logger. Environment in the simulator is created by the combination of terrain database created from the detailed geographical data, model of weather and other dynamic environmental models. Terrain database is created based on real geographic data and includes landscape elements essential for performing the simulation. 3D models and effects allow displaying terrain, entities, effects of burning, extinguishing, etc. The database contains all common objects in the countryside (bodies of water, roads, built-up areas, vegetation, relief, type of soil and other objects).
Individual objects have predefined features influencing simulation of their own entities in relation to their purpose. Weather editor enables to set basic parameters (date and time, air temperature, velocity and direction of the wind, type and intensity of precipitations, humidity and pressure of air, type of cloud cover, light intensity etc.). Some of the parameters are mutually interlinked based on the actions happening in the atmosphere known from meteorology. Dynamic models of environment enable to modify the countryside with objects and phenomena which can change their form in the course of time. There are accidents simulated in great detail as well as a vast database of forces and means. DIS Logger is designed for recordings the simulations and their subsequent re-playing during the exercise evaluation.
Simulated communication system Astra provides simulation of communication between the workplace of the instructor, scenarios and trainees. Exercise has had allocated telephone numbers and roles in the communication scheme.

RESULTS AND DISCUSSION

The chapter describes evaluation of the national multi-level SIMEX 2016 exercise (Malachová and Oulehlová 2017b) where a natural gas supply failure was simulated due to an accident on the distribution system of the high-pressure gas pipeline.

SIMEX 2016 exercise objectives

One of the objectives of the exercise was to check the computer-supported design of the SIMEX exercise simulator design for crisis management authorities. Simulator that was used in the exercise was designed to provide support for the collective exercise of crisis preparedness at various levels of crisis management (regional and municipal). Simulator support was also aimed at integrating components of the integrated rescue system, critical infrastructure entities, their suppliers and customers, and possibly representatives of other entities operating in the area (Malachová and Oulehlová 2017b).
Procedures of the preparation of the individual components of the SIMEX simulator and their implementation into the real environment of the exercising entities within SIMEX 2016 exercise were tested. At the same time, preparation of map data for a predefined exercise area was checked. It has shown that the map data available from the Czech Office for Surveying, Mapping and Cadastre (administrator of the map materials of the Czech Republic) have to be supplemented with additional elements to increase the simulation accuracy. Extensive adaptation of map data in GIS system and implementation of required entities and functionalities that were contained in the exercise scenario were carried out.
As part of the implementation of the SIMEX simulator to support management processes in dealing with emergencies or crisis situations, it was necessary to map and take into account the following facts:

- Who exactly will take part in the exercise (functions, management level, organization);
- What the means will be (simulator, communication system, something else);
- What roles the participant will have in scenarios;
- Where (workplace) he will be during the exercise.

Evaluation of the results of the crisis management bodies' exercise using simulation also verified usefulness of the individual simulator elements, the simulator communication system as well as SIMEX simulation tool as a whole, which was designed to support the exercise of crisis staff members.

Evaluation of the simulator

All interested parties expressed their opinion on the exercise carried out using computer support to deal with an emergency situation. Assessment came from the level of the regional crisis staff, critical infrastructure entity of E.ON Distribution, the VRGroup SIMEX simulator manufacturer and independent observers from the University of Defence.
When evaluating SIMEX simulator used within SIMEX 2016 exercise, it was necessary to look at a simulator from three planes according to its main areas of support for control processes in dealing with emergency situation.

Simulation system
It was a basic function of the SIMEX simulator which created a virtual world for participants where the simulated crisis situation and activities of the deployed forces and resources of the components of the integrated rescue system took place. Simulator and its individual workstations were controlled by trained operators. Simulator operator entered information from the scenarios, followed the instructions of the participants in the simulated virtual environment and reported on the state of tasks or the situation development (Spálenková, Řezáč and Oulehlová 2016; Malachová and Oulehlová 2017a).
At this stage of the exercise, a great demand emerged on the size of the supportive team of operators who provided the simulator operation. For each workplace it was necessary to evaluate what was needed to simulate for the exercise. Feedback of this exercise phase showed that it was necessary to restrict the number of simulator operators.

Communication system
The simulated Astra communication system replaced standard communication channels between scenario workplace and participants. This is an isolated, configurable system enabling voice and email communication. The exercise workplace had an assigned phone number in the communication scheme. Communication was recorded during the exercise (synchronously with the simulation) and it was possible to use it for the subsequent evaluation of the exercise (Spálenková, Řezáč and Oulehlová 2016).
Individual elements of the simulated Astra communication system were telephones and email accounts, located directly at the participants' workplaces. Their use was similar to common means of communication; however, participants did not mostly use them.

Based on the exercise evaluation, it was found out that simulated Astra communication system needs to be transformed and adapted closer to the needs of participants. Reasons for not using the Astra communication system were technical problems with the Internet connection quality. Because of the recording of the communication for the purpose of exercise evaluation, participants provided fragmented, incomplete or limited information. Their communication was restricted to essential basic information.
Some entities were using their internal set communication channels (such as a fire rescue service), which made the exercise more challenging for the participants to have the same information provided by the Astra system as well. Some participants had the problem that the real numbers they had memorized in accordance with the connection plans (part of the crisis documentation) were different from those used in the simulated Astra communication system.

Means for support exercise evaluation
During the exercise, not only synchronous recording of activities and communication, but also recording on the timeline was performed via simulation. Time line recorded other important events as the start and end of scenarios, the external stimuli of participants as well as the steps of the given procedures. These records formed a significant set of data for the subsequent evaluation of the activity of the participants (Spálenková, Řezáč and Oulehlová 2016).
Analysis and evaluation of the exercise was performed immediately after the exercise. During evaluation, the controversial or discussed sections of the audio recording were replayed as well as visualization of the simulated development of the situation or the actions of forces and means deployed in the simulation. This generated an exercise feedback and increased the efficiency and the overall impact of the exercise.
Evaluation also brought the requirement to display fulfilment of the individual phases of the scenarios from the point of view of both participants and the exercise operator. Participants thus would have a better insight into the progress of the scenarios development and the degree of the performance. This means that relevant operators have to record:

- Taking over the scenario;
- Way of dealing with the scenario;
- Delegating of the scenario;
- Ending the scenario.

By doing so, both the participants and the operators will see all scenarios and the level of dealing with them. At the same time, time may be assigned to perform the tasks or the sequence of tasks.

CAXManager application design

The CAXManager application was developed based on the gained knowledge from the SIMEX 2016 exercise evaluation. The goal of the application was to test the behaviour and work of individual exercise entities of crisis management bodies during a major emergency or crisis situation. From the point of view of the topic, exercise aimed at extensive power and gas outage, major floods, windstorms

and so on, where multiple scenarios can be developed and load the application more, were theoretically preferred.
Application can be logged into according to a defined role in the exercise. The main roles in the exercise are:

- Exercise operator represents the main entity in the simulation. It must always be running. The password defined in the configuration file is required for the access. At the beginning, the operator can choose from three options:
  - New - create a new exercise using the wizard. Basic information about the exercise was defined as well as the components involved in the scenario, exercise participants, types of scenarios and scenarios available for the practical exercise;
  - Load - already created exercise that can be edited in different ways is loaded
  - Start - an exercise that has already been run or a new exercise is started.

  Operator in the simulation performs a role of an observer with the possibility to start and finish scenarios. At the end of the exercise it is possible to get a final overview of the whole simulation.
- Participant can log into the application at the time of exercise beginning or realization. For access, a unique password generated solely for the current exercise is required. Participant deals with individual scenarios by performing certain steps which he or she receives directly or are handed to him by other participants. Participant finishes his scenario by completing the task or passing it to another participant. Number of participants is determined by the exercise operator while creating scenarios during the exercise preparation phase.
- Scenario members (scenario openers) task is to distribute scenarios among corresponding participants, to check the reception, fulfilment and termination of scenarios.

Actual start-up and subsequent communication between individual entities is provided through special e-mail communication. Email communication has been chosen for its functionality, reliability and ease of use.

Communication protocol
Communication of the individual CAXManager application installations took place via a single email box that was set in the configuration file - communication.xml. The application uses POP3 protocol to receive messages and SMTP protocol to send messages. Implementation of receiving and sending messages was in the Connector.cs class. Receiving and sending the messages take place by intervals and is set in the corresponding configuration file.
The best functionality has proven to be on one email configuration (POP3) and an individual email box (SMTP) for each user for sending emails. These settings also worked in testing more extensive exercises with extensive and data-intensive e-mail communication between participants. In other configurations, proper functionality used to be compromised by the spam filter on the email server.

Email communication
For the proper functioning of the communication, email format was important. The message type was defined in the email subject, see Table 1, and the message itself was placed in the body of the email.
Type of messages:

- Specialised
  - INIT – an initialization message that contained XML with a completely defined exercise;
  - ATTACHMENT – individual scenario attachments (body contains no information).
- Standard – the message has a unified XML format.

Table 1: Types of messages in CAXManager application

Message name	Activities within exercise simulation
EVENT	Message with the scenario
STARTSIM	Simulation start-up
ENDSIM	Simulation finish
ENDEXERCISE	Exercise finish
CREATENEWEVENT	Creating a new scenario during the course of the simulation
PAUSESIM	Simulation pause
UNPAUSEDSIM	Continuation of simulation
CONNECTUSER	Connect a user
DISCONNECTUSER	Disconnect a user
CHECKCONNECTION	Check the connection of all users
SETTIME	New simulation time setting

By connecting the CAXManager application to the simulation process, parallel activity of the participants has been achieved in the work both with the simulator and the communication tools. This has led to the operator's better overall overview of individual scenarios and their fulfilment process.

CONCLUSION

Increasing the level of preparedness of the staff to deal with emergencies or crisis situations is a very important process. An significant advantage for developing knowledge and skills are exercises of all stakeholders at the appropriate levels. In recent years, many planned tactical and staff exercises of public administration authorities and components of the integrated rescue system have taken place in the Czech Republic. The topics were based on potential risks affecting the given territory or giving response to an unexpected and yet not approached extraordinary event.
In 2016, a computer-assisted exercise on the large-scale disruption of the natural gas supply called SIMEX 2016 was

carried out in the South Bohemian Region. The primary objective of the exercise was to examine crisis plans to address emergencies or crisis situations. SIMEX simulator was tested in the exercise and the basic functionalities for the newly developed CAXManager application were determined based on the evaluation.

The application has currently being tested and will be modified based on the comments which have arisen. The updated version of the application will be used in the multi-level exercise of crisis management authorities scheduled for November 2017 in the South Bohemian Region. Long-term practice shows that large-scale damage caused by extensive emergency or crisis situations can be eliminated by repeating and testing crisis preparedness through joint exercises of the crisis management authorities, components of the integrated rescue system and critical infrastructure entities.

ACKNOWLEDGEMENT

Results presented in this article were obtained as a part of the solution of the project by Technology Agency of the Czech Republic with the topic Research and Development of Simulation Instruments for Interoperability Training of Crisis Management Participants and Subjects of Critical Infrastructure (research project No. TA04021582).

REFERENCES

Act. 2000. Act No. 240/2000 Coll., On crisis management and amending certain Laws (Crisis Law) from 28 June 2000. In: Collection of law No. 118/2011, no. 44, pp. 1114 – 1135. ISSN 1211-1244.

Alcaraz, C. and S. Zeadally. 2015. Critical Infrastructure Protection: Requirements and Challenges for the 21st Century. *International Journal of Critical Infrastructure Protection.* (8), 53–66. ISSN 1874-5482.

Barta. J. and D. Řezáš. 2016. Simulační prostředky využitelné pro výcvik rozhodovacích procesů krizového managementu kritické infrastruktury. In: Riešenie krízových situácií v špecifickom prostredí. Žilina: EDIS-vydavateľské centrum ŽU, 2016, s. 38-46. ISBN 978-80-554-1213-9.

Barta, J.; M. Vaskova and J. Urbanek. 2016. Evaluation of Simulation Programs Applicable to the Support of Decision-Making Processes in Crisis Management of Critical Infrastructure. *International Journal of Education and Learning Systems*, vol. 2016, no. 1, p. 74-80. ISSN 2367-8933.

Council Directive. 2008. Council Directive 2008/114/EC of 8 December 2008 On the identification and designation of European critical infrastructures and the assessment of the need to improve their protection. *In Official Journal of the European Union.* L 345/75-82.

Hrídel, J. and Š. Karták. 2013. Web-Based Simulation in Teaching. In *European Simulation and Modelling Conference 2013, ESM 2013.* Lancaster, United Kingdom: Lancaster University. p. 109 – 113. ISBN 978-90-77381-79-3

Hubáček, M. and D. Řezáč. 2013. Simulation Technology and Training of Rescue Services. *In The Science for population protection,* vol. 5, no. 3/2013, pp. 21-38. 2013. ISSN 1803-635X.

Hubáček, M. and V. Vráb. 2012. The Use of Constructive Simulation for Policemen Training. *The Science for Population Protection*, vol. 4, no. 3, pp. 1-16. ISSN 1803-635X.

Ludík, T. and J. Ráček. 2011. Process Methodology for Emergency Management. *IFIP Advances in Information and Communication Technology*, Heidelberg: Springer 2011, 359, p. 302-309. ISSN 1868-4238.

Malachová, H. and A. Oulehlová. 2017a. Training of the crisis management bodies focused on gas supplies breakdown of great scale using simulation tools. *The Science for Population Protection*, vol. 9, no. 1, p. 123-132. ISSN 1803-635X.

Malachová, H. and A. Oulehlová. 2017b. Analysis of the Gas Distribution System Operator's Activities on Declaring the State of Emergency. In: *Safety and Reliability - Theory and Applications - Čepin & Briš (Eds).* London: Taylor & Francis Group, p. 51-58. ISBN 978-1-138-62937-0.

Oulehlová, A. 2017. Identification of the Electricity Blackout Impacts on the Environmental Security. *In: Risk, Reliability and Safety Innovating Theory and Practice.* London: Taylor & Francis Group, p. 2175-2182. ISBN 978-1-138-02997-2.

Ristvej J; R. Ondrejka; L. Šimák; T. Loveček; K. Hollá; M. Lacinák; L. Šurinová and M. Jánošíková. 2016 Simulation Technologies in Risk Prevention within Crisis Management. In *European Simulation and Modelling Conference 2016, ESM 2016.* Gran Canaria, Spain: University of Las Palmas. p. 327 – 330. ISBN 978-90-77381-95-3

Řehák, D.; J. Markuci; M. Hromada and K. Barcova. 2016, Quantitative evaluation of the synergistic effects of failures in a critical infrastructure systém. *International Journal of Critical Infrastructure Protection*, vol. 14, p. 3-17.

Spálenková, M.; D. Řezáč, and A. Oulehlová. 2016. Cvičení orgánů krizového řízení a slož ek integrovaného záchranného systému v Jihočeském kraji - narušení dodávek zemního plynu velkého rozsahu: SIMEX 2016. In: *Sborník 9. mezinárodní vědecké konference Bezpečnost regionů.* Brno: Vysoká škola Karla Engliše, a. s., s. 331-337. ISBN 978-80-86710-87-7.

Urbánek, J. F. et al. 2013. *Crisis Scenarios*. Brno: University of Defense. 240 pp. ISBN: 978-80-7231-934-3.

Vašková, M.; J. Barta and J. Johanidesová. 2017. Possibilities of critical infrastructure protection. *In: Risk, Reliability and Safety: Innovating Theory and Practice.* London, United Kingdom: Taylor & Francis Group, p. 556-562. ISBN 978-1-138-02997-2.

WEB REFERENCES

Emergency Response Coordination Centre, 2014. Home Page, Brussels, Belgium 〈〉,. Available from: http://ec.europa.eu/echo/about/ERC_en.htm

Sesame. 2011. Securing the European Electricity Supply Against Malicious and accidental thrEats. D9.2 Sesame Newsletter. Version: 2.0, p. 35. Available from: https://www.sesame-project.eu/publications/deliverables/d9-2-sesame-newsletter/at_download/file

BIOGRAPHY

JIŘÍ BARTA, Ph.D. was born on 16th June 1977 in Vyškov, Czech Republic. He graduated from Military University of Ground Forces in Vyškov, Faculty of Economic and Management in 2001. From 2003 to 2004 he worked as a lecturer in the Civil Protection Department of the Military University of Ground Forces in Vyškov.

Since 2004 he is a senior lecturer at the University of Defence in Brno, Czech Republic. His research fields are Civil Protection, Interoperability, Security Management and Crisis Scenarios. He has carried out many national research and development projects. He is the author of more than 65 scientific articles, 2 patents and a co-author of three monographs of collective expertise.

E-mail: jiri.barta@unob.cz

A co-simulation framework interoperability for Neo-campus project

Yassine Motie, Alexandre Nketsa, Philippe Truillet
LAAS-CNRS, University of Toulouse, CNRS, IRIT, UPS, Toulouse, France
email: Yassine.Motie@irit.fr, alex@laas.fr, philippe.truillet@irit.fr

KEYWORDS
Complex systems; interoperability; Mediation; co-simulation; FMI; Neo-campus

ABSTRACT

It is common accepted that complex systems or cyber-physical systems need co-simulation for their study. Further more, they are made of heterogeneous sub-systems that have to exchange data. Usually each sub-system is modeled using specific tools, environments and simulators. The simulators have to interoperate to realize all the simulation of the system. It is known that interoperativity is a broad and complex subject. Interoperability is a strong commitment as the communication solution in heterogeneous systems. This paper describes a co-simulation framework interoperability based FMI (Functional Mock up Interface) standard for the structural part and data mediation for semantic part. We present a case study for Neo-Campus project that shows how the framework helps to build the semantic interoperability of a cyberphysical system.

INTRODUCTION

The Neo-Campus project (Gleizes et al. 2017), supported by the University of Toulouse III, aims to link the skills of researchers from different fields of the University to design the campus of the future. Three major areas are identified : facilitating the life of the campus user, reducing the ecological impact and controlling energy consumption. The campus is considered as a smart city where several thousand data streams come from heterogeneous sensors placed inside and outside the buildings (CO2, wind, humidity, luminosity, human presence, energy and fluid consumption , ...). We distinguish :

- Raw data : These are the energy consumption data (water, electricity, gas).
- Activity-specific data : These are post-processed data resulting from the merging of raw data (pedagogical activities, room occupancy, ..).
- Incident-specific data : These are the failures identified on campus (heating of computer equipment, network failures, ...)
- The ambient data : This concerns the context in which the scenario takes place (temperature, weather, CO2 level in the air).

We have built a knowledge base from sensors data that provides real-time data and relationship between them in order to be used for simulation for example.

Each expert working in Neo-Campus project interact with data differently using a specific field within simulation, hence the need to build a co-simulation framework in order to ease the collaboration.

The paper is organized as follows, first we present the related works on interoperability and our approach to build the co-simulation framework interoperability. We continue with the Neo-Campus use case and give a conclusion and future work.

RELATED WORKS

Interoperability can be defined as the ability of two or more entities to communicate and cooperate despite differences in the implementation language, the execution environment, or the model abstraction (Kohar et al. 1996). Interoperability is a complex problem. There are many ways to deal with it. We have identified two main approaches : (1) based on levels of conceptual interoperability models (LCIM), (2) based on structural and semantic interoperabilities. In (Diallo et al. 2011), LCIM is used as the theoretical backbone for developing and implementing an interoperability framework that supports the exchange of XML-based languages. They defined 7 levels of LCIM with the goal to separate model, simulation, and simulator in order to better understand how to make models interoperate. The authors of (Rezaei et al. 2013) presented an overview of the development of interoperability assessment models. They proposed an approach to measure the interoperability and used four interoperability levels to define a metric. (Li et al. 2013) used comparisons between reusability and interoperability, composability and interoperability to show the importance of interoperability. The authors proposed to use models to build interoperability. Thus they decomposed interoperability into: (1) technical (or structural) interoperability (communication ports) (2) and substantive (or semantic) interoperability (contents meaning) It means that simulation sub-systems can talk to each other and exchange data, but to understand

each other correctly and co-simulate effectively requires substantive interoperability. We agree with this decomposition but instead of using models, which imposes an important customization of the exported model, we take advantage of a known simulation standard, FMI (Functional Mock up Interface) to build the structural part. Our semantic interoperability will be based on data mediation.

CO-SIMULATION FRAMEWORK INTEROPERABILITY APPROACH

Our approach for the design of a co-simulation framework interoperability is based on :

1. a **co-simulation**
2. a **software components** approach which is defined by (Szyperski 1996) as a unit of composition with contractually specified interfaces and explicit context dependencies only
3. a **Structural interoperability** using the standardized interface FMI (Functional Mock-up interface)
4. a **semantical interoperability** using mediation for adaptation of the data

Co-simulation

Co-simulation is defined as the coupling of several simulation tools (Hessel et al. 1999) where each tool handles part of a modular problem where data exchange is restricted to discrete communication points and where subsystems are resolved independently between these points. This allows each designer to interact with the complex system in order to retain its business expertise and continue to use its own digital tools.

From the literature, we have constructed a global scheme of co-simulation in order to have an overall view of it (cf. Figure 1). The models are described by their interfaces without any access to their contents. Aiming at securing model exchanging between designers and ensuring privacy and robustness, we went for a black box model interoperability. This last makes it then possible to exchange and use the information between those components.

Component approach

We chose a component approach in order to overcome the limits of the white box approach which consists on redeveloping of all the heterogeneous subsystems in the same language (Kossel et al. 2006). This imposes that

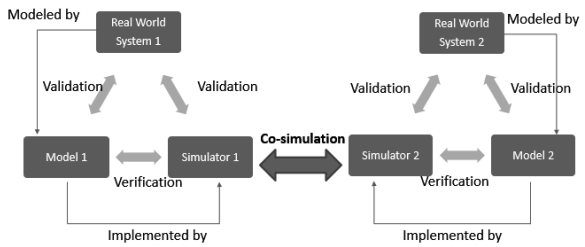


Figure 1: Overview of co-simulation of two subsystems

models developped by different designers are transparent and accessible (Allain et al. 2002). The component approach makes it possible to co-operate prefabricated pieces, perhaps developed at different times, by different people, and possibly with different uses in mind. The main reason is to improve the flexibility, reliability, and reusability of our framework due to the (re)use of software components already tested and validated avoiding risks of robustness. A component is an autonomous deployment entity which encapsulates the software code showing only its interfaces. An interface can be described as a service abstraction, that defines the operations that the service supports, independently from any particular implementation (Lea and Marlowe 1995). Each component should provide the way how it can be generated (plug-out) either from a white box model, or another black box provided by a simulator. (cf. Figure 2) represents a communication between two software components; componentA (as a piece of software) and componentB (UML notation). This component

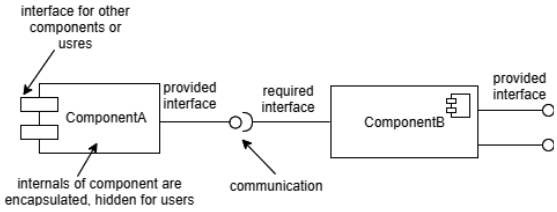


Figure 2: interaction between two Components

approach should allow data mapping between models, applications and several building simulators (Simulink, Dymola, Saber...) using it. A standard interface would be used to warranty the compatibility between these tools.

Structural interoperability based FMI

Structural interoperability requests to define communication ports and has to guarantee the possibility of connection by respecting data types and the direction of the ports.

Structural Interoperability Standard - FMI

We chose the standard FMI (Blochwitz et al. 2011) see (cf. Figure 3) which uses a master-slave architecture as a simulation interoperability standard.

FMI : Functional Mockup Interface is a Standard interface for the solution of coupled time dependent systems, consisting of continuous or discrete time subsystems. It provides interfaces between master and slaves and addresses both data exchange and algorithmic problems. Simple and sophisticated master algorithms are supported. However, the master algorithm itself is not part of FMI for Co-Simulation and should be defined (Consortium et al. 2010) (Enge-Rosenblatt et al. 2011). FMI supports different working modes, in particularly:

1. FMI for model exchange (when modeling environment can generate C code of a dynamic system model that can be utilized by other modeling and simulation environments)

2. FMI for co-simulation (when an interface standard is provided for coupling of simulation tools in a co-simulation environment)

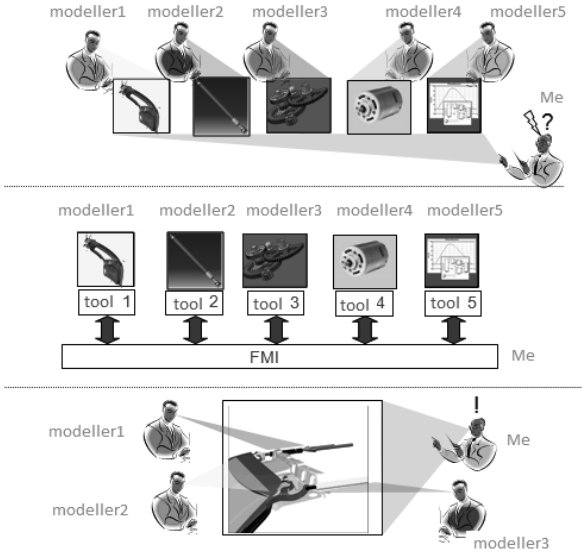


Figure 3: Interoperability using FMI

The use of FMI can be summarized in 4 steps:

- **The design step** :The package of the model of simulation in one component FMU which summarizes in the modeling (creation of the model of simulation), And transformation (publication of the FMU which contains an xml file and the model code or its file), dealing with the main challenge, which is the gap between the semantics of the source formalism of the various calculation models (state machines, Discrete event, data flow or timed automata) and the semantics of FMI (Tripakis 2015).

- **The composition step** : The model of the subsystem is joined to the complex system by establishing the connection graph of the simulation components

- **The deployment step** : The FMUs are made available to the slave simulators. This can be made offline (manually by the user) or online (automatically by the master and where the user specifies in which network the instances of the FMUs are transferred)

- **The simulation step** : The master is responsible for the life cycle of FMUs instances during the execution of the simulation

This choice is motivated by the fact that FMI is a standard and therefore minimizes the customization of the exported model. It is a tool independent that facilitates the exchange of models between different tools and which therefore minimizes the effort of integration by proposing approaches that are specific to it.

Semantical interoperability based data mediation

The interoperability is the ability to share information between systems and applications in meaningful ways. While most system engineering or system applications stop at the structural level, assuming that if you can read it you're going to understand it. Additional level of interoperability and the next that really matters for us is a semantic one which needs common information model to be defined for exchanging the meaning of information. Then the content of the information exchanged is unambiguously defined.

This approach uses a model that describes the information shared by taking their semantics into account in the form of contexts of use. The mediation model leads to define three types of integration rules:

1. constraint rules that reduce the objects to be considered according to predicates,

2. merge rules that aggregate instances of classes of similar,

3. and join rules that combine information from multiple object classes based on one or more common properties.

We distinguish two types of mediation:

1. Schema mediation which provides better extensibility and often better scalability (object interfaces, rule-based language).

2. Context mediation seeks to discover data that is semantically close, it is able to locate and adapt information to ensure complete transparency. We can therefore take advantage of the robustness of the

mediation schema approach and combine it with the semantic approximation techniques of context mediation. This semantic mediation will be used to correlate, aggregate and dispatch data with respect to the control that we want to enforce on data produced or consumed.

For example, one component may produce data with meaning T0, while another may consume data with T1. It may be that there is no direct T0T1 bridge, but there are separated T0T' and T'T1 bridges. A mediator is required to assemble the bridges to complete the T0T1 translation. Our solution is based on formal interface descriptions. When a simple component has an input or output event as a kind of interaction, our interface description will list those events including their (typed) parameters. It could be done automatically if a generator tool, based on language mapping, processes an interface description and produces a proxy (environment side stub) and a driver (component-side stub) for the component. Proxy and driver communicate through a mediation channel, using a protocol for message exchange, see (cf. Figure 4). At the present version, mediators are hand-coded.

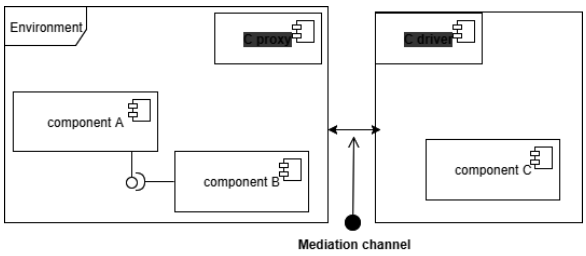


Figure 4: components mediation

The co-simulation framework based component

Simulator

We chose the CosiMate environment which offers a complete simulation environment dynamically linking heterogeneous simulators and which can be extended to different simulation environments on different platforms. CosiMate provides synchronization methods that take into account the different behaviors of the languages and the simulators used. Thus, when a simulation is performed on a network, CosiMate considers the intrinsic constraints of the communication medium. CosiMate adapts to the network configuration, offering a co-simulation based on a multi-client multi-server to avoid unnecessary communications between simulators instantiating local routers for each computer in the co-simulation. And when different parts of the system are co-simulated at different levels of abstraction, it is necessary to add adapters (wrappers) to the compatibility of data exchange between models at different levels. CosiMate meets our needs by offering two co-simulation working modes:

- **Event mode :** The router (that manages the data exchange and synchronizes the simulator) does not deal with any notion of time. This mode of communication makes it possible to establish a connection between the event simulators (HDL simulators, UML models) and sequential simulators (code C for example). The data is transmitted once available on the CosiMate bus. CosiMate is flexible enough to support different communication protocols.The data is transmitted once available on the CosiMate bus, the valid transmission of the router between the sender and receiver (does not check if the recipient has read the data). CosiMate is flexible enough to support different communication protocols.
- **Synchronized mode :** : The router synchronizes the models taking the minimum time. This mode is suitable for simulation engines using solvers (such as Matlab / Simulink).

Cosimate-FMI

Cosimate allows the co-simulation between FMI and also non-FMI models. There are simulators which are not supported by Cosimate, thus the need to wrap them as an FMU component in order to plug them to our cosimate bus

System

The co-simulation of a complex system can thus be based on the joint simulation of all its subsystems. It also makes it possible to simulate the whole system by coordinating and exchanging data calculated and interpreted by each subsystem, in order to obtain a result which does not modify the functionality of the implementation of the future system. Among our different simulated models, we distinguish:

1. Functional simulation allowing the validation of the aspects of the system which are independent of time, and here we can dissociate the sequential simulators and the event simulators.
2. Temporal simulation which aims to exchange data in time windows. Knowing that if data is not consumed in a given time window, it may be lost. A synchronization model is therefore necessary to coordinate the parallel execution of our different simulators.

We mention the master-slave model (cf. Figure 5), comprising a master simulation and one or more slave simulations. In this case, the slave simulators are executed using procedure calls, which results in an inability to execute them simultaneously. The distributed model overcomes the limitations of the master-slave model, which

relies on a co-simulation bus used as a communication protocol (cf. Figure 6). The complexity of this model focuses on the co-simulation platform: managing access to the co-simulation bus and coordinating the data by the bus controller. Another great difficulty comes from the integration of time (Yoo and Choi 1997) which is different between embedded software systems, hardware and the surrounding environment.

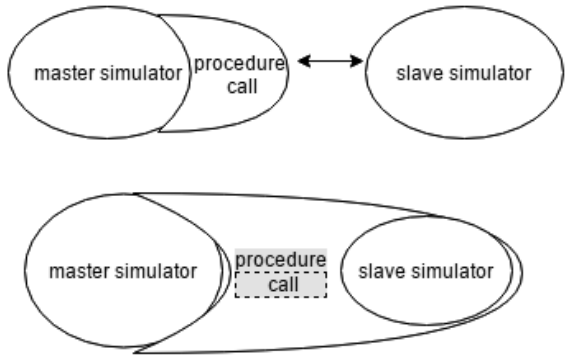


Figure 5: Example of a master-slave co-simulation platform

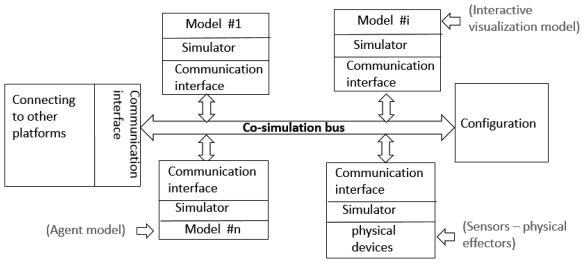


Figure 6: Example of a distributed co-simulation platform

NEO-CAMPUS CASE STUDY

System description

As described in the introduction there are several simulators and sensors scattered around the campus (cf. Figure 7)

Different simulators used

We therefore have: In one hand several simulators on a different fields using different kind of data. One is working with **Matlab Simulink** on the energetic consumption using a black box neural network Heat pump model to ensure a comfortable desired in the rooms by heating and cooling when it is necessary. It is interacting with the outdoors getting from sensors the Electric power (Kw) and the temperature coming from the building and generating with a specific time step. The second simulator is working with **powersims** toolbox of Simulink (Khader et al. 2011) using Maximum Power Point Tracking making it possible to follow the maximum power point of a non-linear electric generator. It is interacting with the outdoors getting the values of the Photovoltaic current and the voltage and generates Converter control setpoint. The third simulator using **Contiki** (Dunkels et al. 2004) which is an operating system for networked, memory-constrained systems with a focus on low-power wireless Internet of things devices using Cooja which allows large and small networks of Contiki motes to be simulated in order to evaluate the performances (energy, delays) of IOT networks using the protocol CCN (content centric Networking) applied on a network of sensors. Developed in C++ under linux OS. The way it interacts with the outdoors is using interest (requests sent by users containing the name of the data such as the temperature) and generating the value of this data. In the other hand the **collection of sensors data** is stored in a NoSQL **database** (mongodb).

Co-simulation engine

The design approach of Neo-campus is necessarily scalable and adaptive, which directs our work towards the development of global and open simulation environment. As we said before we adopted the component approach and described the general FMI's way of working. This last follows a master slave architecture, and we mentioned that a master algorithm needs to be defined in order to synchronize the simulation of all subsystems and to proceed in communication steps, that the data exchange between subsystems is connected via MPI, TCP/IP, Sockets, and that the mapping between outputs to inputs has to be initialized. Cosimate, as described, makes us save the efforts of dealing with synchronization between our subsystems. To perform this integration, CosiMate provides libraries to make the cus-

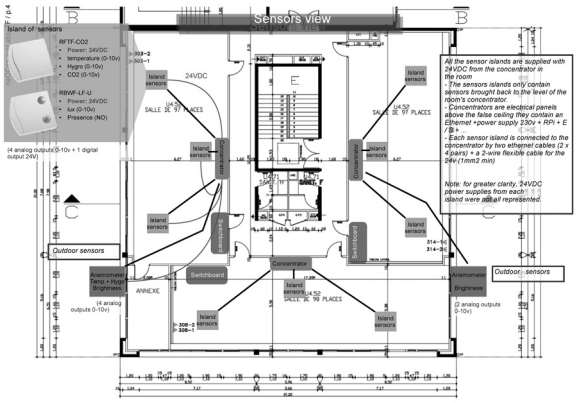


Figure 7: View of neo-campus sensors

tomization easier. The libraries contain (1) I/O ports compiled and described for the simulator/language used. (2) I/O ports description. This description depends on the environment in which the ports are to be used: for example, a header file for C/C++ language is provided. In our case and according to the different simulators mentioned previously, we constructed an FMU's component for each one either by using FMI toolbox like for MATLAB/Simulink or PSIM, or a wrapper using FMI Library from Modelon for the simulator using Contiki. We made it easy to connect all the simulators knowing that for example, Simulink and PSIM are supported by Cosimate but Contiki is not. But the CosiMate FMI connector can load and run all FMI models compatible with FMI 1.0/2.0 for the Co-simulation mode. As we said before each of our FMU files is a zip file that contains a file named modelDescription.xml and one or more platform-dependent shared libraries. The XML files are used to describe how a model running in a simulation environment is connected to the CosiMate bus. We should mention that CosiMate allows execution in the native simulation environment, users can easily work in their familiar environment controlling, debugging, and monitoring simulations as if they are running in a stand alone mode integration. We can also use remote procedure: if the model is to be run on a remote machine. The CosiMate Spy tool is used to monitor and control the co-simulation components and processes. It acts as a reader of the CosiMate bus without modifying data exchanges or simulations synchronization during the co-simulation.

Mediator components

One of the problems encountered is the mediation part, since we want to achieve a semantic interoperability we offered the possibility for each simulator to decide of the way it wants to receive information and depending of the components it's talking to (if it already knows them) to convert its output. For that we encapsulate a mediator with each component before connecting it to the cosimate bus. We added some procedures which allow us to copy and later restore the complete state of an FMU component providing a mechanism for rollback (inspired from the optional functions of the API of FMI 2.0). For our sensor network and as we said that it uses a database to store raw data. This has led us to develop a java simulator (using Mongodb Java Driver) that bridges between the mangodb database and our cosimate bus. We encapsulate the database and our simulator using JFMI (a java wrapper for FMI). So this virtual encapsulation offers capabilities of data mediation and distributes query processing. So the other simulators have no need to know about the database type and location and data can be accessed easily see (cf. Figure 8).

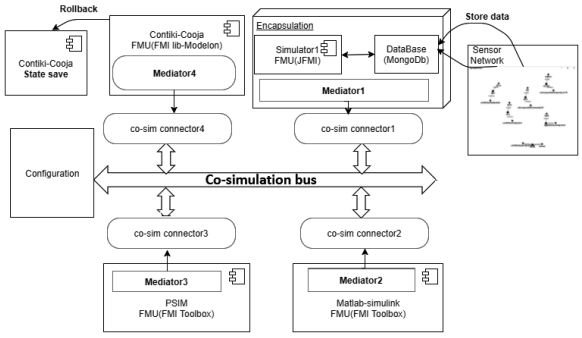


Figure 8: components mediation

CONCLUSION

We have implemented our architecture and our modeling works well, we took as example the 4 simulators including Contiki which is not compatible with Cosimate and for which we had to generate a slave FMU. Our database was encapsulated using JFMI. We were able to solve not only these structural problems but we added mediators to our platform in order to achieve semantic interoperability. It is necessary to mention that our framework allows the integration of all types of simulators and that for the non FMI and even if they are not supported by cosimate the use of a wrapper is enough to envelop them with a c code in order to connect them to the cosimate bus.
This work allowed us to first make an inventory of the practices of the various actors of the neOCampus project. In order to allow the various experts to communicate and collaborate, we realized that it was preferable for them to keep their own practices by allowing them to build or improve their own "expert" simulator. Thus, the objective is a completely open system, easy to use, accepting all types of simulators.
As future work, we would like to build a tool for the generation of mediator. Moreover, we would like to approach semantic interoperability using ontology for the comparison purposes.

REFERENCES

Allain S.; Chateau J.P.; and Bouaziz O., 2002. *Constitutive model of the TWIP effect in a polycrystalline high manganese content austenitic steel. steel research international*, 73, no. 6-7, 299–302.

Blochwitz T.; Otter M.; Arnold M.; Bausch C.; Elmqvist H.; Junghanns A.; Mauß J.; Monteiro M.; Neidhold T.; Neumerkel D.; et al., 2011. *The functional mockup interface for tool independent exchange of simulation models.* In *Proceedings of the 8th International Modelica Conference; March 20th-22nd; Technical Univeristy; Dresden; Germany.* Linköping University Electronic Press, 063, 105–114.

Consortium M. et al., 2010. *Functional Mock-up Interface for Co-Simulation. Accessed March*, 1, 2013.

Diallo S.Y.; Tolk A.; Graff J.; and Barraco A., 2011. *Using the levels of conceptual interoperability model and model-based data engineering to develop a modular interoperability framework.* In *Proceedings of the Winter Simulation Conference*. Winter Simulation Conference, 2576–2586.

Dunkels A.; Gronvall B.; and Voigt T., 2004. *Contiki-a lightweight and flexible operating system for tiny networked sensors.* In *Local Computer Networks, 2004. 29th Annual IEEE International Conference on*. IEEE, 455–462.

Enge-Rosenblatt O.; Clauß C.; Schneider A.; and Schneider P., 2011. *Functional Digital Mock-up and the Functional Mock-up Interface-Two Complementary Approaches for a Comprehensive Investigation of Heterogeneous Systems.* In *Proceedings of the 8th International Modelica Conference; March 20th-22nd; Technical Univeristy; Dresden; Germany*. Linköping University Electronic Press, 063, 748–755.

Gleizes M.P.; Boes J.; Lartigue B.; and Thiébolt F., 2017. *neOCampus: A Demonstrator of Connected, Innovative, Intelligent and Sustainable Campus.* In *International Conference on Intelligent Interactive Multimedia Systems and Services*. Springer, 482–491.

Hessel F.; Le Marrec P.; Valderrama C.A.; Romdhani M.; and Jerraya A.A., 1999. *MCImultilanguage distributed co-simulation tool.* In *Distributed and Parallel Embedded Systems*, Springer. 191–200.

Khader S.; Hadad A.; and Abu-Aisheh A.A., 2011. *The Application of PSIM & MATLAB/SIMULINK in power electronics courses.* In *Global Engineering Education Conference (EDUCON), 2011 IEEE*. IEEE, 118–121.

Kohar H.; Wegner P.; and Smit L., 1996. *System for setting ambient parameters.* US Patent 5,554,979.

Kossel R.; Tegethoff W.; Bodmann M.; and Lemke N., 2006. *Simulation of complex systems using Modelica and tool coupling.* In *5th Modelica Conference*. vol. 2, 485–490.

Lea D. and Marlowe J., 1995. *Interface-based protocol specification of open systems using PSL.* In *European Conference on Object-Oriented Programming*. Springer, 374–398.

Li X.; Lei Y.; Wang W.; Wang W.; and Zhu Y., 2013. *A DSM-based multi-paradigm simulation modeling approach for complex systems.* In *Simulation Conference (WSC), 2013 Winter*. IEEE, 1179–1190.

Rezaei R.; Chiew T.k.; and Lee S.p., 2013. *A review of interoperability assessment models. Journal of Zhejiang University SCIENCE C*, 14, no. 9, 663–681.

Szyperski C., 1996. *Independently extensible systems-software engineering potential and challenges. Australian Computer Science Communications*, 18, 203–212.

Tripakis S., 2015. *Bridging the semantic gap between heterogeneous modeling formalisms and FMI.* In *Embedded Computer Systems: Architectures, Modeling, and Simulation (SAMOS), 2015 International Conference on*. IEEE, 60–69.

Yoo S. and Choi K., 1997. *Optimistic timed HW-SW cosimulation.* In *in Proc. of APCHDL97*. Citeseer.

PRODUCTION SCHEDULING

MODULAR HYBRID MODELING BASED ON DEVS FOR INTERDISCIPLINARY SIMULATION OF PRODUCTION SYSTEMS

Bernhard Heinzl
Philipp Raich
Franz Preyser
Wolfgang Kastner
Institute of Computer Aided Automation
Vienna University of Technology
Treitlstraße 1-3
A-1040 Vienna
Austria
E-mail: bernhard.heinzl@tuwien.ac.at

Peter Smolek
Ines Leobner
Institute for Energy Systems and Thermodynamics
Vienna University of Technology
Getreidemarkt 9
A-1060 Vienna
Austria

KEYWORDS
hybrid DEVS, Modular Hybrid Modeling, Production Systems, Industrial Energy Efficiency, Cubes

ABSTRACT

In order to analyze and optimize complex industrial production systems, methods for hybrid simulation are necessary that incorporate discrete as well as continuous aspects while at the same time providing modularity across engineering domains. To this end, this paper presents an overview of an approach based on hyPDEVS – an extended DEVS formalism – that facilitates reuse of hybrid components. The application of this approach is demonstrated on a model of an industrial oven as well as a complete case study example of a typical production plant. A prototype in-house simulator implementation shows sufficient performance to enable simulation-based optimization with a high number of iterations and allows to be integrated into existing industrial automation software infrastructure in order to facilitate energy-aware plant operation.

INTRODUCTION

Energy and resource efficiency in the industrial sector has become increasingly important in recent years because of its economical and ecological impact and, at the same time, significant potential for savings. In order to support decision-making processes during planning and for energy-efficient operation of production facilities, many software tools employ dynamic simulation (Herrmann et al. 2011). Yet covering such systems as a whole, incorporating aspects from different engineering domains (production machinery, logistics, energy infrastructure, building) remains challenging as it requires combining discrete as well as continuous simulation models as part of a hybrid simulation approach.
In the research project *BaMa - Balanced Manufacturing*, a software tool is being developed that enables analyzing and improving energy efficiency through monitoring, prediction and simulation-based optimization. One of the sub-goals is to integrate simulation functionality into existing automation systems. Several fundamental requirements on the simulation software have been identified, the most important ones being hybrid simulation capabilities, open interfaces, computational performance (to be able to perform simulation-based optimization with possibly thousands of iterations), transparent simulation engine (for trustworthiness of simulation results), and ease of use for non-experts.
These requirements rule out most of the available off-the-shelf solutions. Furthermore, the usability requirements demand not only hybrid modeling, but also modular implementation of reusable components, which is difficult to achieve with common techniques for hybrid simulation like co-simulation. Instead, other approaches for a tighter integration of discrete and continuous models have to be explored. One possible direction, which we present in this paper, follows a hybrid model description based on the *Discrete Event System Specification* (DEVS) (Zeigler et al. 2000).

BACKGROUND

Related Work

Some approaches for simulating hybrid systems employ coupling of different simulation environments as part of a so-called *co-simulation* (see (Heinzl 2016) for a more detailed definition), typically by using some middleware that handles orchestration and coordination between the software tools. Apart from the computational overhead, co-simulation usually introduces significant complexity into the modeling process, in particular regarding model development, maintenance and overall usability of existing components. Co-simulation frameworks are often highly customized to a particular application and/or

simulation tools with low reusability of model parts.
The *Functional Mock-up Interface* (FMI) is a tool-independent standard with the goal of facilitating co-simulation of dynamic models. FMI support sophisticated coupling strategies, e.g. communication step size control or higher order signal extrapolation. However, FMI is focused predominantly on continuous models based on the Modelica language and is not well-suited for event-driven simulations.
Regarding modeling and simulation based on DEVS, there are a number of software tools available (Franceschini et al. 2014). However, only a minority of them offer hybrid simulation like for example PowerDEVS (Kofman et al. 2003). As a drawback, PowerDEVS does not support the Parallel DEVS (PDEVS) formalism, which made it unsuitable for our particular applications, as our evaluations showed (Preyser 2015).

Hybrid DEVS Formalism

The classic *Discrete Event System Specification* (DEVS) is a formal model description (accompanied by a simulator execution algorithm) for modeling and simulation of discrete-event systems (Zeigler et al. 2000). Based on DEVS, a family of extensions was proposed in the subsequent years, including *Parallel-DEVS* (PDEVS) with improvements for handling concurrent events, *DEV&DESS* (Discrete Event and Differential Equation System Specification) combining the description of discrete-event and continuous systems and, in recent years, *Hybrid PDEVS* (hyPDEVS) (Deatcu and Pawletta 2012, Heinzl 2016).
All these DEVS-based formalisms allow to build models from components in a hierarchical manner by distinguishing between *atomic* and *coupled* components. More formally, a hyPDEVS *atomic* is specified by the tuple (see also (Heinzl 2016, Deatcu and Pawletta 2012))

$$M_{hp} = \langle X, Y, S, f, c_{se}, \lambda_c, \delta_{state}, \delta_{ext}, \delta_{int}, \delta_{conf}, \lambda_d, ta \rangle \,, \tag{1}$$

with the sets of input events X, outputs Y and states S, all of which may contain discrete as well as continuous values. The remaining entries in eq. (1) are functions specifying the continuous and discrete behavior: rate of change function f describing ordinary differential equations (ODEs), continuous and discrete output functions λ_c and λ_d, state event condition function c_{se} for localizing state events, and transition functions δ_{state}, δ_{ext}, δ_{int} and δ_{conf} for state event, external, internal and concurrent (confluent) transitions, respectively.
In addition to atomic hyPDEVS, the formalism also specifies *coupled* hyPDEVS models, which are comprised of an external interface (input/output), subcomponents (which must again be hyPDEVS components) and coupling relations:

$$CM_{hp} = \langle X, Y, D, \{M_d | d \in D\}, EIC, EOC, IC \rangle \,. \tag{2}$$

The set M_d denotes the sub-components with corresponding index set D. Three distinct sets describe the connections between sub-components: EIC for external input couplings, EOC for external output couplings, and IC for internal couplings.
The hyPDEVS formalism itself does not specify how to handle numerical integration of the ODEs in the simulation engine. For this, several approaches are available, see (Deatcu and Pawletta 2012) for more details.

MODULAR HYBRID MODELING

For implementing a simulation model of a dynamic system, many approaches and software tools follow a *component-based* paradigm where existing model components are composed into larger models in a bottom-up manner. Connections between these components represent dependencies and interactions, thereby capturing the dynamic behavior of the overall system.
The relationship between application models, components and simulation software is illustrated in Figure 1 as a conceptual layered architecture. A simulation model of a particular system (*Application Model*, Layer V) is composed of instances of *Model Components* (Layer IV), which are implemented in some kind of simulation software (*Simulator*, Layer III), that typically also handles numerical computation and event handling using a *Simulation Engine* (Layer II). The simulation engine in turn typically builds on some formal description (*Formalism*, Layer I) that defines the semantics of the used modeling constructs (e.g. events, output function, etc.) in order to facilitate correctness and consistency.

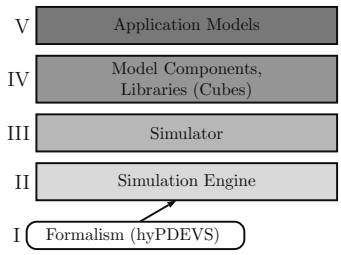


Figure 1: Conceptual layered architecture for component-based modeling and simulation.

One of the advantages of component-based modeling is that it facilitates *separation of concerns*, which makes it easier to manage the complexity of the overall system and distribute model development. It also promotes *reuse* by enabling to build libraries of validated model components that can be instantiated, which is crucial in an attempt to reduce the effort necessary for non-experts to build a new application model.
However, when it comes to hybrid models, a modular design of reusable components is particularly challenging. In order to achieve true modularity (and thus reusabil-

ity) for hybrid models, it is crucial to employ an integrated approach across the entire architecture stack (see Figure 1). This approach starts at the model description formalism (Layer I) upon which the software can build and which has to encapsulate both continuous and discrete model aspects within the same component boundary. After evaluating different formalisms and approaches (Preyser 2015, Heinzl 2016) including DEVS-based simulation because of its sound formal basis and extensive available research, we chose the hyPDEVS formalism. More details regarding the implementation are described later.

Combining modular hybrid modeling with interdisciplinary considerations and requirements regarding ease of understanding demands a closer look at model components as the base unit for bottom-up model development, what these components represent and how they are applied in practice. Especially the underlying modeling formalism is often difficult to comprehend for non-experts, as experience showed. In an effort to assist communication between domain experts as well as between experts and non-expert application engineers, we provide a conceptual abstraction from the underlying modeling formalism, the so-called *Cube*. A Cube denotes a meaningful part of the overall system that typically represents some well-defined physical component in the real world and which interacts with it surroundings by exchanging energy, material and information. A Cube can for example represent a machine tool, a conveyor belt, a compression chiller or a thermal zone inside a building. The neutral denotation as a "Cube" independent from the terminology of a particular engineering domain allows to incorporate aspects of various disciplines within the same abstract concept.

EXAMPLE CUBE: OVEN

As an example of a Cube, we present a model of a conveyor oven that accepts entities (workpieces, etc.), moves them through a temperature-controlled area and outputs the entities on the other end.

Semi-formal Description

Instead of directly formalizing the model in hyPDEVS, a preceding step involving a semi-formal description not only allows the model engineers to make quicker prototyping iterations, but such a model is also easier to understand.

A Cube model is specified by its outer structure (i.e. interfaces) and its inner behavior. Figure 2 shows an overview of the Cube's inputs, outputs, internal parameters and variables. Regarding the internal behavior – which involves continuous as well as discrete aspects – Figure 3 presents the *discrete* behavior in the form of a state diagram, governing the material flow described as discrete entities.

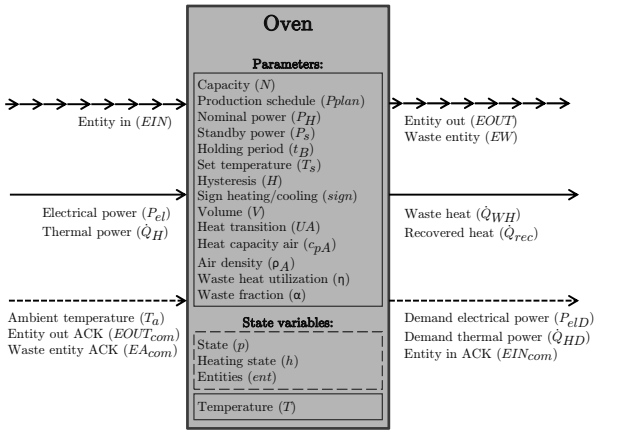


Figure 2: Interfaces, internal parameters and variables of the oven Cube model.

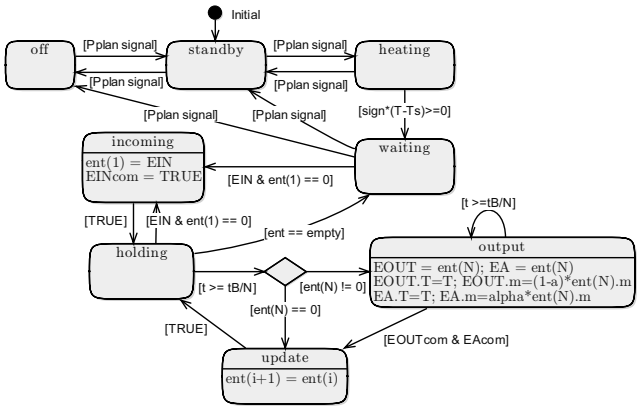


Figure 3: State diagram depicting the discrete internal behavior of the oven Cube model.

States are provided for `off`, `standby` and `heating` mode, determined via a signal from a production schedule (`Pplan signal`). A more detailed explanation of the state machine is given in (Heinzl 2016).

In addition to the discrete behavior, continuous aspects can be modeled using differential and algebraic equations. These aspects typically involve energy flows described by balance equations. For example,

$$\frac{dT}{dt} = \frac{\dot{Q}_H - (T - T_a) \cdot UA}{c_{pA} \cdot \rho_A \cdot V + \Sigma_{E \in ent} E.c_p \cdot E.m} \tag{3}$$

describes the dynamic of the internal temperature T as an energy balance equation, with $\dot{Q}_H$ as the heating power input, T_a the ambient temperature, UA the heat transition coefficient and the thermal mass of the oven $c_{pA} \cdot \rho_A \cdot V$. The term $\Sigma_{E \in ent} E.c_p \cdot E.m$ denotes the thermal mass of all entities $E \in ent$ inside the oven.

Translation to hyPDEVS

For implementing the oven Cube model, the semi-formal model description – which is independent from any

DEVS-based implementation – has to be translated into a hyPDEVS compliant model by providing the respective function specified by hyPDEVS, see eq. (1).
Unfortunately, this translation is not a trivial process, as several DEVS-related modeling considerations have to be taken into account. For example, one has to take care of how to safely hand entities from one atomic to another, which is not covered natively by hyPDEVS. Though the formalized hyPDEVS model is omitted here for reasons of brevity, more details can be found in (Heinzl 2016, Raich et al. 2016).

IMPLEMENTATION

A first implementation of basic Cube components and a simplified application example (which is presented in (Raich et al. 2016, Heinzl 2016)) was carried out using the MatlabDEVS Toolbox (Deatcu and Pawletta 2012) for MATLAB. This implementation served as a proof-of-concept for the Cube approach as well as DEVS-based hybrid simulation.
After verifying the feasibility of the approach, development for a stand-alone simulator could be initiated (implemented in C++), starting with the simulation engine based on hyPDEVS (cf. layer II in Figure 1). This simulation engine was then embedded into a larger simulator architecture (layer III) that provides necessary infrastructure for creating and executing simulation models. Building on this simulator infrastructure, Cube models (layer IV) could be re-implemented, which were then verified against the existing MATLAB implementation and other independent implementations (e.g. in Modelica). The result is a library of components intended to be reused across a multitude of application models (level V) for future use cases. One such use case is presented in the following section.

APPLICATION EXAMPLE

To demonstrate how the DEVS-based modular hybrid modeling approach and the Cube models can be applied to simulate a real-world production facility, we derived an application example.

Description

This example, shown in Figure 4, is based on a real production plant of an industrial bakery that produces baked goods in different variants, fresh as well as frozen, and features a processing line with different paths, an energy supply system, thermal zones and a building hull. Discrete material flow is incorporated as well as continuous energy flow and information signals.
In particular, the model includes a building Cube containing four thermal zones, each representing a distinct part of the facility: production hall, cold storage, plant room and office. The thermal zones all have independent conditioning and exchange heat with each other (as well as with the environment) via heat transfer. Energy system and network Cubes model conversion and distribution of energy inside the system, in the form of heating, cooling and electric energy.
For producing baked goods, respective ingredients are pulled from a storage (top right in Figure 4) and processed along the production line, including mixing, splitting (Cube `Rex`), baking, cooling and packaging.
The description for the oven Cube was presented in the previous sections. The same model can also be used for the cooler, freezer, proofer and VGS stations. Other Cube models concerning building and energy system are described in detail in (Smolek et al. 2016).

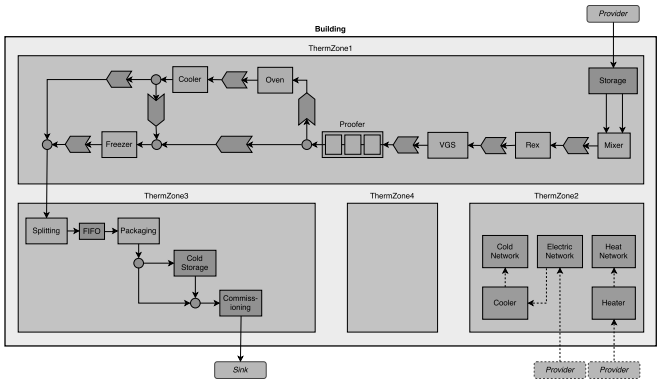


Figure 4: Example application of a production facility consisting of a processing line, an energy system and thermal building Cubes.

For defining simulation scenarios, the user can specify production schedules, which are read as input parameters. In particular, an entry in a production schedule constitutes a command for state change inside a Cube (see for example the oven state diagram in Figure 3). An adequate production schedule is especially crucial for the oven and similar Cubes, as they need time to prepare (e.g. switching temperature).

Testing and Results

This section presents some exemplary simulation results in order to demonstrate the application of the implemented example model.
Figure 5 presents some of the relevant simulation results for a typical production schedule over one day, producing two different products in two resp. three batches. The trajectory for the oven temperature reveals the heating and re-heating intervals as well as the cool-off period after receiving the `off` signal at 17:00. In addition, the simulation yields overall energy consumption, which is important in order to assess energy efficiency of various scenarios.
Without going into detail regarding the numerical results, the point of the case study is to demonstrate what

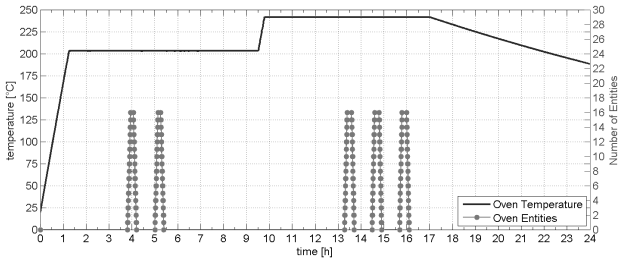


Figure 5: Simulation results for the example scenario: Oven temperature and entities.

kinds of results can be obtained from the simulation: continuous (temperature, energy, etc.) as well as discrete behaviour (persistent entities), spanning multiple domains of engineering (production machinery, energy system, building), all from a single simulation model.

DISCUSSION

The prototype implementation has shown that all of the requirements described in the introduction can be fulfilled in principle: Open interfaces can be achieved since we have full control over the source code. The hyPDEVS formalism provides a transparent specification for the simulation engine and the Cube concept allows to increase ease of use for non-experts. However, still a lot of software development is necessary, for example regarding graphical user interface and model editor as well as visualization.

Results concerning execution performance were also more than satisfactory. More extensive simulation scenarios with simulation time of up to 7 days and roughly 500,000 entities have a running time of less than 10 seconds on conventional state-of-the-art computer hardware. Execution time is sufficiently short to allow performing simulation-based optimization techniques.

One disadvantage of DEVS-based modeling in engineering applications involves the "roughness" and genericity of the formalism, see also (Raich et al. 2016). This makes initial model implementation time-consuming and prone to inconsistencies. However, as model engineers can potentially draw on ever-growing libraries of ready-to-use components for building their models, necessary effort for building new components is expected to decrease in the future.

CONCLUSION AND FUTURE WORK

We presented an approach for modular modeling of hybrid components as well as an implementation of a case study that demonstrates the feasibility as an alternative approach for tackling complex interdisciplinary systems. Compared to co-simulation, hyPDEVS provides a tighter integration between continuous and discrete model aspects.

Future work includes enhancing the simulator with a framework for interconnection between the simulation and optimization algorithms. Furthermore, implementing additional Cube libraries will allow to build more elaborate simulation models. In the end, the software will enable interdisciplinary investigations of energy efficiency alongside production planning and scheduling in order to optimize operation strategies in production facilities.

REFERENCES

Deatcu C. and Pawletta T., 2012. *A Qualitative Comparison of Two Hybrid DEVS Approaches. SNE - Simulation Notes Europe*, 22, no. 1, 15–24.

Franceschini R.; Bisgambiglia P.A.; Touraille L.; Bisgambiglia P.; and Hill D., 2014. *A Survey of Modelling and Simulation Software Frameworks Using Discrete Event System Specification.* In *OASIcs-OpenAccess Series in Informatics.* vol. 43.

Heinzl B., 2016. *Hybrid Modeling of Production Systems: Co-Simulation and DEVS-Based Approach.* Diploma thesis, TU Wien, Vienna, Austria.

Herrmann C.; Thiede S.; Kara S.; and Hesselbach J., 2011. *Energy Oriented Simulation of Manufacturing Systems - Concept and Application. CIRP Annals - Manufacturing Technology*, 60, no. 1, 45–48.

Kofman E.; Lapadula M.; and Pagliero E., 2003. *PowerDEVS: A DEVS–based Environment for Hybrid System Modeling and Simulation. School of Electronic Engineering, Universidad Nacional de Rosario, Tech Rep LSD0306*, 1–25.

Preyser F., 2015. *An Approach to Develop a User Friendly Way of Implementing DEV&DESS Models in PowerDEVS.* Diploma thesis, TU Wien, Wien.

Raich P.; Heinzl B.; Preyser F.; and Kastner W., 2016. *Modeling Techniques for Integrated Simulation of Industrial Systems Based on Hybrid PDEVS.* In *2016 Workshop on Modeling and Simulation of Cyber-Physical Energy Systems (MSCPES).* IEEE, 1–6.

Smolek P.; Leobner I.; Gourlis G.; Mörzinger B.; Heinzl B.; and Ponweiser K., 2016. *Hybrid Building Performance Simulation Models for Industrial Energy Efficiency Applications.* In *Proceedings of the 11th Conference on Sustainable Development of Energy, Water and Environment Systems (SDEWES 2016).* Lisbon, Portugal.

Zeigler B.P.; Prähofer H.; and Kim T.G., 2000. *Theory of Modeling and Simulation: Integrating Discrete Event and Continuous Complex Dynamic Systems.* Academic Press. ISBN 978-0-12-778455-7.

Simulation of a Flexible and Adaptable One-Piece-Flow Assembly Line Based on a Process Flow of Colored and Timed Petri Nets

Benedikt A. Latos
Peyman Kalantar
Philipp M. Przybysz
Susanne Mütze-Niewöhner
Institute of Industrial Engineering and Ergonomics
RWTH Aachen University
Bergdriesch 27, 52062 Aachen
Germany
E-mail: b.latos@iaw.rwth-aachen.de

Christoph Holtkötter
Jan Brinkjans
Miele & Cie. KG
Carl-Miele-Straße 29
33325 Gütersloh
Germany

KEYWORDS
Production, Discrete simulation, Decision-making, Process-oriented, Stochastic

ABSTRACT

The growing demand for individualized consumer goods and the increase in production volume variations brings significant challenges for manufacturing enterprises so that engineering teams are assigned the task to find innovative production solutions to meet these requirements. Especially personnel-intensive assembly departments have to be designed to meet the challenge of enabling such a flexible production whilst still maintaining good working conditions. When comparing different assembly concepts, simulation can help to explore the behavior and implications of different innovative assembly concepts. By doing so, set up cost can massively be reduced since the system can be launched in a more efficient manner. This paper presents the development of a simulation model based on a colored Petri Net flow chart. The area of application of this model is a manufacturer of white goods situated in Germany. The process-oriented and actor-integrated simulation model provides the possibility of simulating diverse scenarios when comparing various concepts in an early assembly line planning phase.

INTRODUCTION

It is essential for manufacturing companies to adopt to frequent market changes in order to remain competitive. The need for continuous product innovation, increase in product variants, shorter product lifecycles, fluctuating demands and handling unexpected requirements in recent years are a result of these increased dynamics (Spath et al. 2013). Assembly typically represents the point of variant formation. Hence, focusing on its processes is necessary when designing a strategy to increase the production flexibility (Petersen 2005). In addition, current developments (e.g. demographic changes and migration) imply that employers will have to deal with these additional diversities among their employees with respect to different performance levels in a short and long-term perspective. Therefore, there is a need for adaptable assembly systems, which tolerate interpersonal performance differences.

The final assembly of a manufacturer of white goods in Germany currently is organized as a typical clocked assembly system. The ability of the current concept to cope with the mentioned challenges is questionable. Therefore, the evaluation of alternative concepts for enabling long-term flexible and adaptable manufacturing is required (Latos et al. 2016; Latos et al. 2017).
The assembly organizational form of One-Piece-Flow can be regarded as an extended flow-line approach where the working person and the product exhibit concurrent movement characteristics. The assembly object is transferred along the assembly stations with either manual or electrically supported transportation equipment. The line is typically designed in a U-shape to minimize the walking distances for working persons. This principle implies that buffers within the line remain unnecessary, as working persons balance varying process times (Bullinger et al. 2009). In general, a single person is responsible for the complete assembly of a product, which represents a holistic product responsibility approach (Arzet 2005). The main advantages of such a system can be stated as enabling the assembly of diverse product variants. Moreover, a flexible output of the system can be achieved by adjusting the number of working persons in the system (Lotter and Wiendahl 2012; Schlick et al. 2010).
In this paper, the development of a One-Piece-Flow assembly simulation model in the field of the German manufacturer of white goods, which was conducted as participative innovation process, is presented. The assembly line is modeled with Colored Petri Nets. The model is implemented in FlexSim simulation software. In the following, simulation approaches are briefly explained and Petri Nets are introduced as well as the selected approach and the assumptions which have been used for modeling are presented. In the simulation analysis chapter and conclusion, the simulation results and their interpretation which can be used for decision-making are discussed.

ASSEMBLY SIMULATION APPROACHES

According to German VDI 3633 simulation is a method for reproducing a system in an experimental model in order to gain findings that can be transferred to reality (VDI3633). In general, simulation approaches can be classified as activity-oriented or actor-oriented (Duckwitz et al. 2010; Tackenberg

et al. 2010). Actor-oriented approaches focus on the personnel's abilities and decision-making processes, whereas in activity-oriented approaches processes are driven by tasks. Both approaches can further be divided according to the degree of persons' consideration in the model (Duckwitz et al. 2010; Licht et al. 2007).

When considering different assembly concepts, one may regard monetary, capacity-related and flexibility-oriented decision criteria (Lotter and Wiendahl 2012). Nevertheless, it is crucial to follow a systematic assembly planning method for being able to reach a substantial decision. First, different assembly concepts can be developed. Next, a model can be developed for capturing the characteristics of the assembly concept. After that, simulation can be implemented and verified and finally, after validating the simulation results, the decision-making process follows (Banks 1998).

In order to model the process flow of an assembly line as a discrete event system Petri Nets were used. Petri Nets can be considered as a method that manipulates events according to certain rules. The great feature of Petri Nets is including explicit conditions that enable to model complex control schemes. A Timed Petri Net is a six-tuple (P, T, A, w, x, V) where (Cassandras & Lafortune 2008):

- P is the finite set of places (round nodes in the graph).
- T is the finite set of transitions (rectangular nodes in the graph).
- $A \subseteq (P \times T) \cup (T \times P)$ is the set of arcs from places to transitions and vice versa in the graph.
- w: $A \rightarrow \{1, 2, 3,...\}$ is the weight function on the arcs.
- x is a marking of the set of places.
- $V = \{v_j : t_j \in TD\}$ is the clock structure.

We can model deterministic or stochastic systems by relating the values of V to a set of known values and outcome values of a probability distribution function respectively (Cassandras & Lafortune, 2008).

Colored Petri Nets surpass the drawbacks of simple Petri Nets such as modeling the characteristics of the semi-finished products without adding another structure to the net and by giving attributes to tokens. Colored Petri Nets enable the analysis of temporal structures of manufacturing processes. Moreover, complex processes can be modeled with a high level of clarification which provides the opportunity for visual analysis. By adding another tuple "C" to the Petri Net, the Colored Petri Net can be defined where C is a non-empty set of colors and CS is a set of all color tokens that can be stored in a place S (Viswanadham & Narahari, 1987). Colors of token represent their attributes. Attributes represent data types, show characteristics of semi-finished parts, indicate process related properties of tokens or simplify the visualization of complex processes by differentiating between tokens (Jensen & Kristensen, 2015).

Model Conception

Here a washing machine was considered as a product with available variants for simulating its assembly process in a One-Piece-Flow assembly line. The modelling, simulation and evaluation was conducted according to the simulation procedures of Banks (1998). The processing logic of the simulation is stochastically modeled with Timed and Colored Petri Nets. Basic assumptions for modeling the process flow with a Petri Net were:

- Only one person at a time can work at one working station.
- Overtaking maneuvers should be possible during assembly. The model applies a performance index for each person that allows initiating overtaking procedures. If a person arrives at a working station where another person with a lower performance index is already assembling, the assembling person will finish the task, then wait and let the latter to overtake.
- There is no shortage in assembled materials.
- For each task, the simulation uses proper task execution times. As real assembly times are distributed, a Gaussian distribution of the assembly times is implemented in the model. In addition, other distributions can be applied.

Ultimately, the simulation model calculates the assembly times as follows in equation (1):

$$t_{ij} = OEI_j * SEI_{ij} * t_i \sim N(\mu_i, \sigma_i^2) \qquad (1)$$

t_{ij} = assembly time for employee j at work station i
OEI_j = overall efficiency index of employee j
SEI_{ij} = stationwise efficiency index of employee j at station i
μ_i = standard time at station i from working plan
σ_i^2 = standard deviation of standard time at station i

The simulation model is based upon the two-dimensional block layout of the assembly line which was created by using an open access factory planning PowerPoint tool (Kampker et al. 2012). The model defines stopping points within the layout and imports the corresponding MTM-UAS standard times from a database. This allows for simulating the assembly of different product variants. Moreover, variants that happen to not require one process step feature a standard time of zero for that specific process. This enables the simulation of mixed model production programs.

Unrealistic accumulations of working persons in spatially close areas can be avoided by assigning a maximum number of tolerable persons for certain areas of the assembly line.

The activities of a workstation including overtaking maneuvers as a subsystem are modeled with Timed Petri Nets. The Colored Petri Net is used to facilitate visual analyses of the system, as suggested by Gradisar & Music (2013). The Petri Net graph is shown in Figure 1.

The Petri Net model then was hard coded in FlexSim 2017 in a process flow. Features of the software were used to create a 3D model (Figure 2) and integrate the process flow into it. The model allows to set different parameters for each simulation run. To do so, input variables like the number of persons, traveling speed, performance indices and starting positions of persons as well as shift duration were defined to simulate different scenarios. In addition, the model incorporates fixed inputs and disturbance variables.

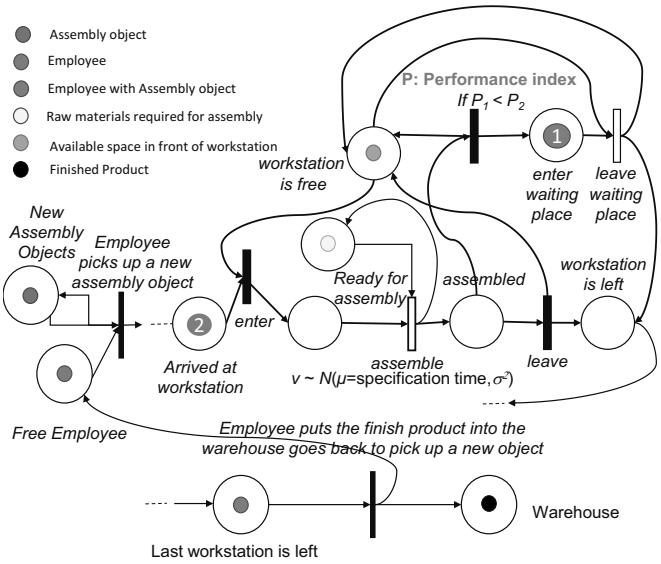


Figure 1: Colored Petri Net model of the assembly line

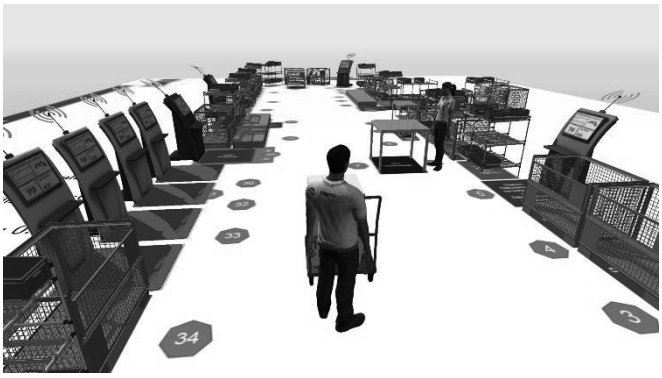


Figure 2: 3D simulation model in FlexSim

For verification, validation and testing techniques informal, static as well as dynamic methods described by Banks (1998) were employed. For instance, an expert evaluation was used as informal approach, whereas a semantic analysis of the source code was conducted as static method. As a dynamic technique, sensitivity analyses were performed with respect to the parametrization of the model.

SIMULATION ANALYSIS RESULTS

The system behavior of the assembly line can be explored, as the input variables can be varied within multiple simulation runs. However, this paper only presents a range of simulation results. Moreover, the data has been modified by applying a fictitious multiplication factor to the results by means of industrial data protection.

In theory, normalized cycle time is calculated by dividing the overall work content by the number of persons in a One-Piece-Flow system (Arzet 2005). This determines the output of the system. By varying the normalized cycle time, one can meet the actual cycle time, which is induced by customer demands.

Yet, it is trivial to understand that the output of the system will not increase proportionally if one increased the number of workers incrementally. In fact, the normalized output time cannot be reduced to a minimum because no arbitrary high number of employees can be assigned to a One-Piece-Flow system. Therefore, the output also will flatten at some point.

With simulation it is possible to determine the number of workers in the system from which on the output does not continue to grow proportionally without the physical existence of the assembly line.

Therefore, the number of employees in the line was gradually increased and (n = 200) simulation runs were conducted for every scenario. The output of the system was tracked in order to determine the point from which on the output flattens because of the increase in waiting times due to interpersonal disturbances (Figure 3). By doing so, statistical conclusions concerning the output can be drawn. For instance, confidence intervals for the output regarding a certain assembly scenario can be calculated.

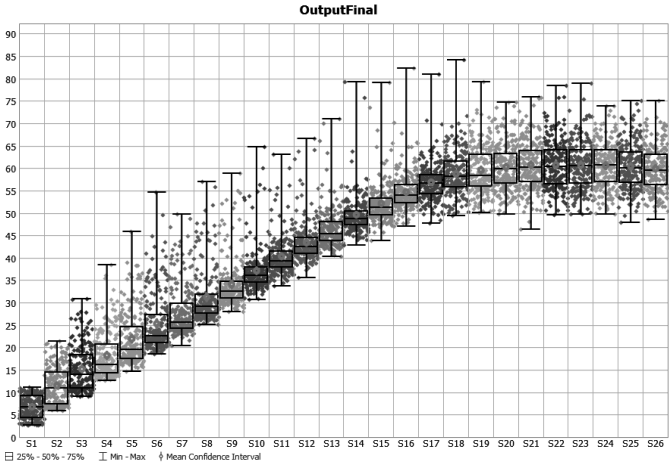


Figure 3: Exploration of possible outputs with respect to the number of persons in the assembly line (modified data, simulation analysis conducted with FlexSim)

The simulation approach enables the consideration of capacity-examinations in an early planning phase. This supports the decision process when comparing different assembly concepts. Furthermore, the employees' ratio of different activities or the composition of throughput times can be tracked and analyzed. Simulation may help to iteratively approximate towards the ideal operating point of a production system. This facilitates to launch the system in a more efficient manner.

Evaluation of Assembly Time Distribution Types

Fischer et al. (2005) state that an important consideration during the design and simulation phase of an industrial assembly station should be the stochastic nature of assembly times. Therefore, the question to be answered is which type of distribution should be used (Fischer et al. 2005). Possible types of continuous distributions, which can be applied in simulation, are presented in (Banks 1998). Fischer et al. (2005) state that the problem during a data collection phase of a simulation project is that the required parameters for the often-suggested Normal- or Gamma-distribution are usually not available. In contrast, the parameters of the beta distribution were easy to estimate, since they are characterized by an optimistic, pessimistic and a most common time value (Fischer et al. 2005; Schlick and Demissie 2015). Neumann (1975) recommends choosing a

distribution for execution times that fulfils three conditions: Firstly, the distribution should be steady; secondly, the resultant activity execution times should be bounded above and below, and thirdly, the realization of activity execution times should be concentrated around a certain value. Fischer et al. (2005) conclude that the beta distribution complies with these conditions and possesses the additional advantage of modelling stochastic activity times with the PERT methodology by using an easy accessible estimation of an optimistic value (OT), a pessimistic value (PT) and a most common value (see Figure 4 and also Neumann 1987). Furthermore, the parametrization of shaping parameters enables to model non-axially symmetric execution times (Banks 1998). Finally, Krajewski and Ritzman (1999) also suggest that this approach is suitable for modelling strongly scattered operation durations (see also Fischer et al. 2005).

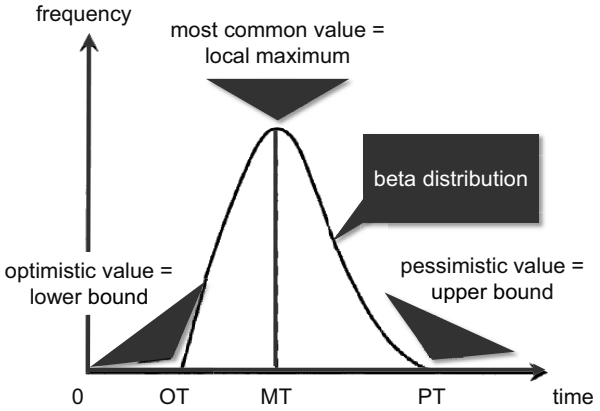


Figure 4: Modelling beta distributions according to the PERT methodology (in the style of Fischer et al. 2008)

For modelling assembly times in this simulation approach, a parametrization of beta distribution was chosen which has a longer tail to the right side since assembly times will rather tend to take longer due to disturbances. The symmetric normal distribution has the drawback that a situation where an employee assembles faster than standard time has the same probability as a situation where he needs longer within a certain absolute value of standard deviation above and below the mean. Hence, the beta distribution was parametrized in FlexSim in the following way:

$$t_i = \text{beta}(0.75 * \mu_i, 2.00 * \mu_i, 2, 5, \text{default}) \qquad (6)$$

μ_i = standard time at station *i* from working plan

In order to analyze whether beta distributed assembly times do effectively affect the simulation results, a comparative study was conducted and the testing hypothesis were formulated as follows:

H_0: There is no difference between the mean values of normal and beta distributed assembly times.

H_a: There is a difference between the mean values of normal and beta distributed assembly times.

In this paper, exemplary results for the saturation point of the assembly system with respect to the number of employees working in it are presented. On the one hand normally distributed assembly times were assumed and (n = 200) simulation runs were conducted. This procedure was repeated with the assumption of beta-distributed assembly times according to equation (6) on the other hand. Figure 5 illustrates the results as total output per shift.

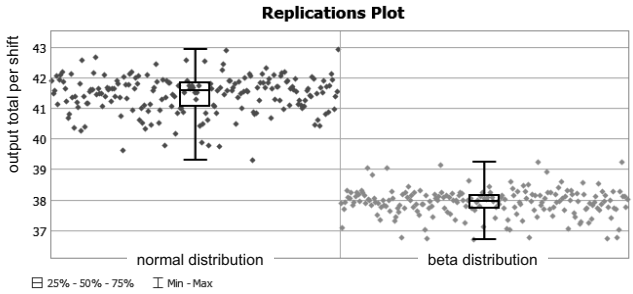


Figure 5: Comparison of normal and beta distributed assembly times (modified data, simulation analysis conducted with FlexSim)

From Figure 5 a significant difference between the two distribution times may be expected. This can be shown by means of statistical testing methods. A selection of a suitable method was conducted according to the approach of Albers et al. (2009). Due to the consideration of one variable and the level of measurement, a parametrical test of two independent samples was conducted as a two-sided t-test.
The calculated p-value ($p < 0.0001$) is smaller than the chosen significance alpha level of 0.05, thus the null hypothesis H_0 must be rejected, whereas the alternative hypothesis H_a should be accepted. The risk of rejecting H_0, even though it may be true, is less than 0.01 %. In conclusion, the application of beta distributed assembly times and the chosen parametrization in the simulation does affect the simulation results in terms of lower output values. Hence, it is important to further evaluate in field studies which distribution type describes assembly times in a One-Piece-Flow system more accurately.

CONCLUSION

This paper presents the approach of simulating a One-Piece-Flow assembly line at a manufacturer of white goods in Germany in an early planning phase. The simulation helps to predict the system's behavior and enables capacity-examinations without its physical existence.
It was shown that the chosen parametrization of beta parametrization leads to a lower system output in comparison to the parametrization of assembly times as a normal distribution.
The simulation model makes it possible to examine diverse questions in an early planning phase. Moreover, the model can be used in further planning steps to predict the system's behavior with an updated data basis in a more accurate way. However, one must consider that simulation cannot model human interactions adequately. Especially psychological influences on working persons are not explicitly modelled. By employing variances in simulation, uncertainties can at least be considered. In the end, the results are based on assumptions that can be discussed controversially.

FUTURE RESEARCH

Simulation enables to apply distributions of process times; hence, results that are more realistic can be achieved in

comparison to calculations which employ fixed standard times. However, further research should evaluate whether a beta distribution approximates assembly times in a One-Piece-Flow system more accurately than a normal distribution. This may be achieved with time studies for such assembly systems.

Finally, it is possible to parametrize and validate the model more precisely with empirical values after implementing the assembly line. In this regard, the simulation results may also be practically validated as well as a simulation tool can be used as an operative planning tool to forecast the output before shift start. If an automated tracking system was implemented, the production manager could simultaneously track the position of the product and therefore display the assembly progress within the simulation model in real-time. With regard to the field of industry 4.0, the assembly status would always be accessible, transparent and even predictable. Future research will focus on different working modes of a One-Piece-Flow system, as it can also be run whist dividing the overall work content into segments that are assigned to different employees respectively. Moreover, it is possible to evaluate whether these segments should be assigned to working groups that may support each other in terms of cooperative work when a group happens to have lots of products in the waiting queue of their line segment. Simulation will further be explored as a method for designing work systems in a participative way. In this context, virtual reality (VR) technology can be used to directly visualize planning results of workshops with employees in a VR-environment. Ergonomic assessment methods will be integrated into the simulation model in order to analyze the ergonomic implications for working persons in an early planning phase. Furthermore, web based simulation approaches will be evaluated to enable small and medium-size companies to tackle the digital transformation of their planning processes.

ACKNOWLEDGEMENT

The foundation of this case study concerning the composition of diverse innovation teams was developed within the research project “derobino” (2011-2015) funded by the Federal Ministry of Education and Research in Germany according to Grant No. 01HH11007. The authors would like to express their gratitude for this support.

REFERENCES

Albers, S.; Klapper, D.; Konradt, U.; Walter, A.; Wolf, J. 2009. Methodology of Empirical Research (in German). 3rd ed.. Wiesbaden: Gabler

Arzet, H.: Grundlagen des One-piece-flow. 2005. Guideline for planning and realizing operator-based production systems (in German). Berlin: Rhombos.

Banks, J. 1998. Handbook of Simulation. Principles, Methodology, Advances, Applications and Practice. New York: Wiley.

Bullinger, H.; Spath, D.; Warnecke, H.; Westkämper, E. 2009. Handbook of Enterprise Organization. Strategies, Planning, Implementation (in German). 3rd ed.. Berlin: Springer

Cassandras, C. G., & Lafortune, S. 2008. Introduction to Discrete Event Systems. 2nd ed.. Springer.

Duckwitz, S.; Tackenberg, S.; Karahancer, S.; Schlick, C. M. 2010. A Meta-Model for Actor-Oriented, Person-Centered Simulation for the Management of Development Projects. *In: Proceedings of the 17Th International Conference on Industrial Engineering and Engineering Management*

Fischer J., Stock P., Zülch G. 2005. Simulation of Disassembly and Re-assembly Processes with Beta-distributed operation Times. In: Zülch G., Jagdev H.S., Stock P. (eds) Integrating Human Aspects in Production Management. IFIP International Conference for Information Processing, vol. 160. Springer, Boston, MA

Gradisar, D., Music, G. 2013. Petri Net modelling for batch production. 7th IFAC Conference on Manufacturing Modelling, Management. Saint Petersburg. 1566-1571.

Jensen, K., & Kristensen, L. 2015. Colored Petri nets: a graphical language for formal modeling and validation of concurrent systems. Communications of the ACM, Vol. 58. No. 6, 61-70.

Kampker, A.; Burggräf, P.; Mecklenborg, A.; Kreisköther, K. 2012. Efficient Layout Planning in Interdisciplinary Teams. Complexity-suiting Tools Enable Sustainable Project Success (in German). In: VDI-Z. 154. Vol. 2012, No. 11/12, 74-75

Krajewski, L. J.; Ritzman, L. P. 1999 Operations Management. Reading, MA.: Addison Wesley.

Latos B. A., Holtkötter C., Brinkjans J., Mütze-Niewöhner S. 2016. Simulationsbasiertes Vorgehen zur Planung einer Montagelinie nach dem Prinzip One-Piece-Flow (in German). IAW Spectrum, Vol. 12, 8-9.

Latos B. A., Holtkötter C., Brinkjans J., Przybysz, P. M., Mütze-Niewöhner S., Schlick C. M. Partizipatives und simulationsgestütztes Vorgehen zur Konzeption einer flexiblen und demografierobusten Montagelinie bei einem Hersteller von weißer Ware (in German). *In Proceedings of 63. Kongress der Gesellschaft für Arbeitswissenschaft : FHNW Brugg-Windisch*, Switzerland, 15.-17. February 2017

Licht, T.; Schmidt, L.; Schlick, C.; Dohmen, L.; Luczak, H. 2007. Person-centred simulation of product development processes. In: Int. J. Simulation and Process Modelling, Vol. 3, No. 4, 204-218

Lotter, B.; Wiendahl, H.-P. 2012. Assembly within Industrial Production. A Handbook for practitioniers (in German). 2nd ed.. Berlin: Springer

Neumann, K.: Operations Research Verfahren. Vol. 3: Graph Theory, Network Analysis (in German). Munich, Vienna: Carl Hanser Verlag, 1975.

Neumann, K. 1987. Network Analysis. In: Fundamentals of Operations Research, part 2. (in German). ed.: GAL, Tomas. Berlin et al.: Springer, 165-260.

Petersen, T. 2005. Forms of Organization in Assembly. Theoretical Foundations, Organizational Principles and Configuration Approaches (in German). Aachen: Shaker

Schlick, C.; Bruder, R.; Luczak, H. 2010. Industrial Engineering (in German). 3rd ed.. Berlin: Springer

Schlick, M.; Demissie, B. 2015. Product Development Projekcts. Dynamics and Emergent Complexity. Heidelberg: Springer

Spath, D.; Müller, R.; Reinhart, G. 2013. Future Assembly Systems. Economic, Changeable and Reconfigurable (in German). Stuttgart: Fraunhofer

Tackenberg, S.; Duckwitz, S.; Schlick, C. 2010. Activity- and actor-oriented simulation approach for the management of development projects, In: International Journal of Computer Aided Engineering and Technology, Vol. 2 No. 4, 414-435

VDI-Richtlinie 3633. Simulation of Logistic-, Material Flow- and Production Systems, Foundations (in German). VDI-Handbuch Materialfluss und Fördertechnik, Band 8. Berlin: Beuth Verlag.

Viswanadham, N., & Narahari, Y. 1987. Colored Petri net models for automated manufacturing systems. In Proceedings of IEEE International Conference on Robotics and Automation, Raleigh, NC, 1985-1990.

WEB REFERENCES

FlexSim Software Products, Inc. (FSP). 2017. FlexSim 2017 manual, URL: https://www.flexsim.com/

PRODUCTION PROCESS EVALUATION AND IMPROVEMENT BY USING THE METHOD OF DISCRETE EVENT SIMULATION

Tolga Kudret Karaca
Science Institute, Istanbul Arel University, Turkoba Mahallesi Erguvan Sokak No:26/K 34537 Buyukcekmece Istanbul, Turkey. E-mail: tolga.karaca.tk@gmail.com

Volkan Cakir
Department of Industrial Engineering, Istanbul Arel University, Turkoba Mahallesi Erguvan Sokak No:26/K 34537 Buyukcekmece Istanbul, Turkey. E-mail: volkancakir@arel.edu.tr

Key Words
Simulation, discrete event simulation, bottleneck analysis, Arena, label printing

ABSTRACT

The label printing and packaging industry has a rapidly growing and evolving market in the world. The most important aim of the companies is to accomplish the demand and expectations of the customers in the market. Label printing houses have to improve their production processes to produce fast and good quality products. In this paper, system performance was analyzed and remedial solutions were sought for a company, which is active under variable demand conditions. Aim of this study is identifying bottlenecks, balancing production lines and improving system efficiency by using simulation. A conceptual model was created including two main production lines and a discrete event simulation model was created at Arena software. After validation and verification steps scenario analysis were conducted and solutions for production problems and improvement suggestions were made to the management of the company.

1. INTRODUCTION

Etisan co. is one of the most popular and biggest label printing houses in Turkey. It was founded in 1971. Shortly after, it started producing a variety of products including all kinds of adhesive and sleeve shrink labels. The Company has one R200 letterpress printing machine, two Gallus EM 410S Flexo printing machines, two Rotoflex cutting machines and a quality control machine in the main line. The company also has a HCI Sleeve shrink agglutination sheeter and quality control machines in sleeve shrink line. The main purpose and motivation of the study is increasing the throughput of the factory as company management requested. Because of the complex structure of label demands, simulation is the preferred method for label printing process. The products were grouped according to their production technic, order arrival times, order quantities, deadlines, production times; and time to failures, failures down times were analyzed. Studies found during literature survey, that are using simulation at the label printing production environment are "JIT Performance In a Printing Shop (Patterson, et al., 2002)", "Product Development Process Modeling Using Advanced simulation (Cho & Eppinger, 2001)", "Simulation Modeling And Analysis Of A New Mixed Model Production Lines (Hasgül & Büyüksünetçi, 2005)", "Discrete Event Simulation As An Aid In Conceptual Design Of Manufacturing Systems (Jägstam & Klingstam, 2002)", and "Process And Quality Improvement Using Work Methods And Simulation (Olcar, 2014)"

2. METHODOLOGY

Simulation refers to broad collection of methods and applications to mimic the behavior of real systems usually on a computer with appropriate software (Kelton, et al., 2010). Discrete event simulation was preferred as a suitable method for label printing process because of the complex structure of demand and production. Suggestions and solutions that provide the increasing productivity were intended by using simulation method. Steps in Simulation study are described at Figure 1 (Banks, 1999).

In this study, all production data during one year was investigated. Model was constructed according to principles of 24-hour non-stop 3-shift production system. To perform simulation study, Rockwell Automation Arena 14.0 Simulation Software was used. The Arena simulation program works with an object-oriented design to create a visual model. The simulators are machine, operator, and material handling systems and so on. They work with visual objects called modules to identify system components. The Arena works with SIMAN simulation language in a structure (Takus & Profozich, 1997). We consider simulation to include both the construction of the model and the experimental use of the

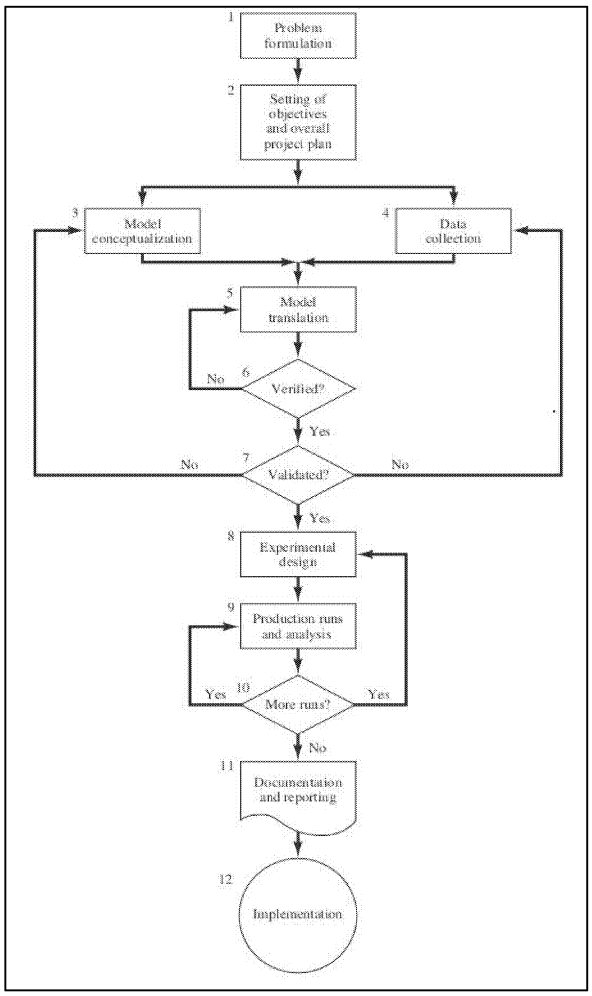


Figure 1: Simulation Steps

model for studying a problem. Thus, we can think of simulation modeling as an experimental and applied methodology, which seeks to:

- Describe the behavior of a system.
- Use the model to predict future behavior, i.e. the effects that will be produced by changes in the system or in its method of operation (Shannon, 1998).

3. APPLICATION

3.1. Problem Definition

The first step in model building is to examine the problem itself (Altiok & Melamed, 2007). According to first observations, demand cannot be met adequately; product delivery times were failed. The main reason for these problems was imbalanced production between printing and the quality control lines. The bottom line is long queues, errors and delays of orders causing lost customers and fines. In turn, any delay experienced in deliver time and unmet customer demands cause loss of new orders and decrease of competition power at the sector.

3.2. Setting of Objectives

For the simulation study, a set of targets can be created based on problem definitions. Such targets represent the main purpose of working by establishing criteria (El-Haik & Al-Aomar, 2006).

3.3. System Definition and Model Conceptualization

Conceptual modeling is the abstraction of a simulation model from the part of the real world it is representing 'the real system'. A non-software specific description of the computer simulation model (that will be, is or has been developed) describing the objectives, inputs, outputs, content, assumptions and simplifications of the model (Robinson, 2011).

In pre-press phase raw material (Self-adhesive material and PET, PVC Sleeve Shrink), UV letterpress or UV Flexo inks, other printing supplies are generally stored in the company warehouse. Graphic artwork and plates preparation processes are made in facility. Raw material needs are met within maximum one day. Production processes are investigated as three processes:

1- Pre-press process

2- Printing process

3- Cutting, slicing and Quality Control process (Figure A and Figure B)

Orders were separated into main groups according to their printing technics. Arrival times, order amounts and printing technics were analyzed.

Consecutive orders were analyzed because sometimes the same products (variant) and sometimes, different product orders come into production department and this totally effects preparation and setup times, because of the printing technics and color settings. At the planning department, work orders are created for every product order and orders should be separated into three categories as a new, revised or repeating order. Because of the short time durations of graphic study for orders and it does not cause delays for production processes, graphic process was excluded from the model. When orders arrives, two work orders are created by planning department; one for the warehouse in order to prepare raw materials and the one for the cliché department for plates. Because of the short preparation time of R200 letterpress plates, it was assumed zero and it does not cause delays. Therefore, preparation process of letterpress plates was excluded from the model. On the other hand, Flexo plate process was taken into consideration. Because, plate preparation time of this machine depends on how many plates will be exposed.

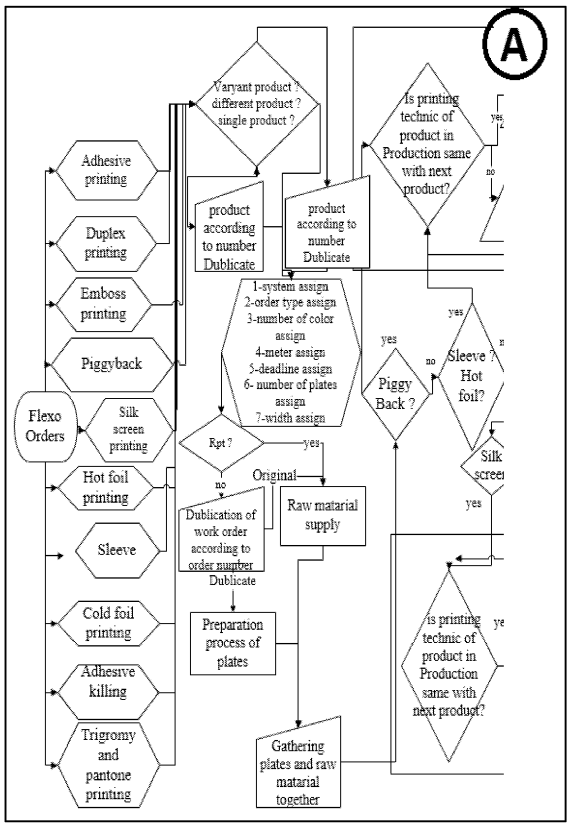


Figure 2: Conceptual Model Part A

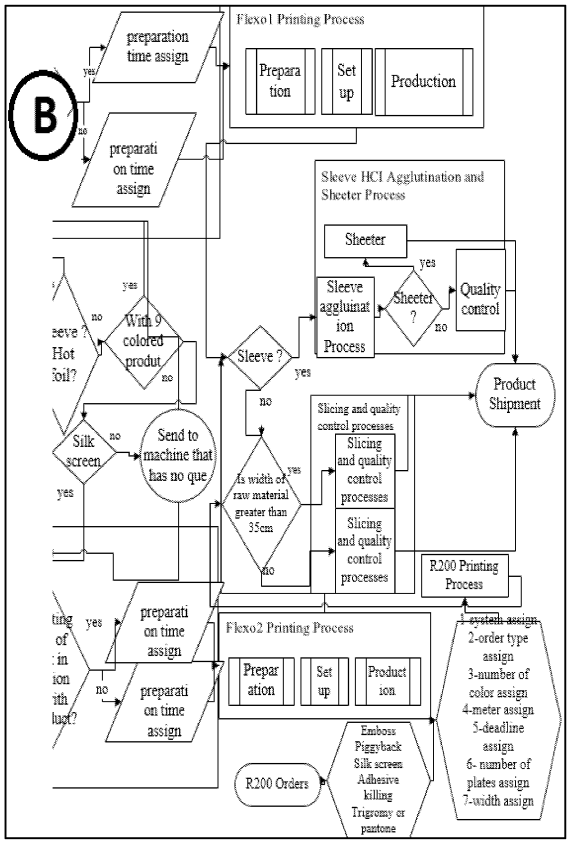


Figure 3: Conceptual Model Part B

A work order is scheduled only when its raw materials and plates are ready. There are two Flexo machines in Flexo line. One of them has nine color units and one die cut unit and the other one has eight color units plus one die cut unit. Beyond that, two machines have different printing capabilities in terms of printing techniques. R200 machine has one letterpress machine. Lower-quality or low-volume jobs are printed on R200 letterpress machine. Gallus R200 Letterpress machine has eight color units, and varnish unit, also a die cut unit. Management of company decides which jobs are printed on Letterpress machine or Flexo machines according to their costing:

- Nine colors labels, sleeve shrinks, hot foil labels (generally all of them accepted nine colors labels), and Piggyback labels can be printed on Flexo 1 machine.

- Eight colors labels, Silk screen labels with eight colors (generally printed on Flexo 2 because of its silk screen printing performance)

- Eight colors (plus extra varnish) labels can be printed on R200 letterpress. The products of the same order that are completed in pre-press are printed sequentially.

When partial work orders of same order are come to production line, they are taken consecutively otherwise orders follow FIFO principle. Printing process is separated into three phases:

- Machine Preparation

- Set-up

- Production

Machine Preparation is the operation that is performed on machine configuration. Preparation times vary according to printing technique used, number of colors, changes of colors and color units. Since this is a complex and very time-consuming job, it needed the most attention during data preparation stage as will be explained in the next part of this paper.

Following the preparation phase, set-up phase begins. At this phase, the crosses of each plate on rollers of machine are placed on top of each other, and colors are set according to confirmed customer cromolyn. It is observed that in some cases, this operation consumes a lot of effort and time of the workers and it is totally varying whether if job is a repeating, new, revised job and the number of colors. Production time depends on the velocity of machines and the size of work order (in meters). Velocity of machines also depends on number of the colors and printing technique used. During our observation, some problems are observed causing downtimes. These causes were grouped as follows:

- Production downtimes due to problems caused by raw materials.
- Production downtimes due to problems caused by plates die cut punch or knife.
- Production downtimes due to problems caused by failures.
- Production downtimes due to problems caused by job confirmation.
- Production downtimes due to problems caused by Test cases.

After printing operation, finished rolls are taken to the quality control area for slicing and quality control processes. Sleeve shrinks are taken to HCI agglutination and quality control machines. In some cases, orders go to sheeter machine, if the customer specifically requests it. After printing, self- adhesive rolls are taken to the queue of Rotoflex machines according to their width for slicing. Rolls that are greater than 330 mm width are taken into the queue of Rotoflex 430, and the others are taken to queue of Rotoflex 330. Slicing and quality control times vary according to lengths of rolls and velocity of machines. Extracting defective products, shaving roll edge, burr debugging are decreasing the capacity. Downtimes in HCI line were neglected, since they are extremely rare according to supervisor and not observed during data collection.

3.4. Data Collection and Data Analysis

Data collection is necessary for estimation of model input parameters. The assumptions can be formulated from distributions of random variables. Even if the data is incomplete, it may be possible to specify the parameter ranges, and all of the input parameters in this range may be simulated using some of them. Data collection is also necessary for verification (Altiok & Melamed, 2007). The procedure includes the following steps:

- Step 1: Plot the data

  Use a histogram and summary statistics to determine the general characteristics of the underlying distribution:

- Step 2: Select a family of distributions.

  Use the results of step 1 to select a set of "reasonable" distributions. Fit each distribution to the data observed and estimate the distribution parameters.

- Step 3: Select the best distribution.

  Determine which of the fitted distributions best represents the data observed using one or more appropriate statistics.

- Step 4: Check the distribution quality.

  Determine the distribution goodness of fit using:

  Chi-square test

  Kolmogorov–Smirnov test

  Anderson–Darling test

Arena Input Analyzer coming with Rockwell Automation Arena 14.0 simulation package was used in order to determine the most appropriate probability distributions for collected data. While determining the distribution for each process and transaction an elaborated input analysis including Kolmogorov-Smirnov and Chi-square tests were used. The appropriated distributions were used at the simulation, only if test results were complimented with the appropriate P-value ANOVA analysis was conducted on in order to understand whether there is a difference between days order units. In addition, orders were analyzed separately as orders include variant non-variant and first-time product orders. Order quantities were defined as probability distributions of each order type for each system. Number of colors and demand deadlines were analyzed and for each product of each order. Work orders were separated into number of colors and order quantities were analyzed (in meters). Number of plates that should be exposed for revised and first-time orders and their exposure times were analyzed. Production times for each machine were analyzed and distributions were defined. Printing machine's preparation times were analyzed in two categories:

- Because of color changes, preparation time analysis in the same configuration on the machine.
- Because of color changes, preparation time analysis in the different configuration on the machine.

Set up times were found for each order type (revise, repeat, first-time order) by the number of colors used. After setup times, machine production rates were examined based on each system and the number of colors used. Interval times of failures and maintenance time lengths were analyzed.

3.5. Model Translation

Briefly, a model is constructed with the appropriate simulation language (Shannon, 1998). Model was constructed by using Rockwell Automation Arena 14.0 Simulation Software.

3.6. Verification

Verification is ensuring that the model is correctly transferred to the computer environment and that the application is correct (Sargent, 2011). In order to verify the model, we simply check to see if the model is

behaving as expected (El-Haik & Al-Aomar, 2006). During, the study period with the help of the variables used, model was observed and errors were extracted. The model outputs were examined in detail. Also for model validation, another expert approval has been obtained (El-Haik & Al-Aomar, 2006). Verification was done clearly for this model.

3.7. Validation

Validation is the process of ascertaining that a model is an acceptable representation of the "real world" system. The validation process is concerned with establishing the confidence that the end user will have on the model. Some critical questions to ask during this stage are; does the model adequately represent the relationships of the "real" system? And, is the model generated output "typical" of the real system? (Centeno, 1996). WIP plot diagrams were created and analyzed for 10 runs. At the end of the analysis of WIP, warm-up period of 15 days came to stable condition accepted as warm-up period. Epsilon value was accepted as 10 and numbers of finished work orders were analyzed.

The worst-case error was considered as; "worst-case error" $\leq \varepsilon$ (Table 1)

The worst-case error =

$$\left|\text{Real system value} - (\text{Average} - \text{Halfwidth})\right| \quad (1)$$

Table 1: For 100 Runs Total Finished Work Orders and Confidence Interval

20:41:04 Category Overview Haziran 22, 2015

Values Across All Replications

Etisan

Replications: 100 Time Units: Minutes

User Specified

Counter

Count	Average	Half Width	Minimum Average	Maximum Average
Biten toplam isemri sayisi	397.14	7,87	291.00	489.00

For 100 runs, the worst-case error conditions were provided:

$$398,2 - (397,14 - 7,87) = 8,93$$

$$8,93 < 10$$

Total output (pcs) T-test

Moreover, T-Test was performed for total outputs (finished work orders).

$$t_0 = \frac{\bar{y}-\mu_0}{s/\sqrt{n}} \quad (2)$$

From 2 we can find:

$$t_0 = |-0,26668|$$

$$0,26668 \leq 1,98$$

$$\hat{\delta} = \frac{|\bar{y}-\mu_0|}{s} \quad (3)$$

$$\beta(\hat{\delta}) \quad (4)$$

From 3 and 4 we can find:

$\alpha = 0,05$ With double-sided tail test and from (OC)

Curve; β (0, 25) = 0, 20 then

1-β (power of test) = 0,80 value was accepted enough for power of test (Banks, et al., 2005).
In T -test of total number of finished work orders. It was accepted that model represents actual system with probability of % 95 for 100 runs.

Time spent in the system for work orders (min)

T-test

In addition, the time spent in the system for work orders were analyzed.

$|\bar{y} - \mu_0|$ =1000 is accepted as critical value

$$t_0 = |-0,68|$$

$$0,68 \leq 1,98$$

β (0, 28) = 0, 20

1-β (power of test) = 0, 80

In T-test of time spent in the system for work orders, it was accepted that model represent actual system with probability of % 95 for 100 runs.

4. SCENARIOS

In the output analysis, the analysis of the simulation outputs is performed to understand the system behavior. These outputs are used to obtain determinations about the behavior of the real system. At this stage, visualization tools can be used to help. The aim of visualization is to provide a better understanding of the actual system under investigation and to facilitate the examination of the large numerical data set produced in the simulation run (Wainer, 2009). To find a solution for order delay and bottleneck problems, two scenarios were created:

Scenario A: The least deadline ranking method was used where orders to be delivered first were processed first instead of FIFO.

Scenario B: The least deadline ranking method with consideration of order quantities ranking method was used in mixture model. At this scenario shortest delivery date work orders has priority in plate exposure process for printing processes and smallest quantity work orders has priority in quality control and slicing machines of Rotoflex line and HCI .

The assumptions were made for testing the difference between two averages is performed as follows (Kelton, et al., 2010) :

- Samples are independent.
- There is no relationship between the elements in the samples
- Variances of populations where the samples were picked up from are unknown and not equal.

$$CI_{\overline{Y}_1-\overline{Y}_2} = \overline{y}_1 - \overline{y}_2 \mp t\alpha_{/2,v} \cdot \sqrt{\frac{s_1^2}{R_1} + \frac{s_2^2}{R_2}} \qquad (5)$$

From 5 we perform T-test

T-test performed to Scenario A and Scenario B for Flexo line (Number of delayed work orders)

$$CI_{\overline{Y}_1-\overline{Y}_2} = (14.35, 39.22)$$

Scenario B, with the probability of 95%, it shows an improvement in terms of delayed work.

T-test performed to Scenario A and Scenario B for Flexo line (the time-minute spent in the system)

$$CI_{\overline{Y}_1-\overline{Y}_2} = (1797.034, 3542.186)$$

Scenario B, with probability of 95%, it showed provided an improvement for the time spent in the system (Table 2). Time spent of work orders in the system, number of delayed work orders was analyzed, and as we can see on Table 2, there is no significant increase in outputs of system. However, with available resources we can easily recognize decreases in number of delayed work orders and time spent.

5. CONCLUSION

An improvement was observed at the mixture strategy; number of delayed work orders decreased from 25% to 10% in Flexo line, and 19% to 8% in R200 line. In addition, %24 reduction in time spent in the system was observed for Flexo line work orders.

Table 2: Comparison of the Performance Measurement Basic Models and Scenarios

Performance Parameters	Basic Model	Scenario A	Scenario B
Number of Total Outputs	397.14	408.49	405.48
Number of Outputs of Flexo line	293.95	303.32	303.38
Number of Orders deliveried on time Flexo line	220.41	245.49	272.34
Number of Delayed orders Flexo line	73.54	57.83	31.0400
Number of Outputs of R200 line	103.19	105.17	102.10
Number of Orders deliveried on time R200 line	82.63	97.43	93.1500
Number of Delayed orders R200 line	20.56	7.74	8.9500
Duration of stay in the system for Flexo work orders	7485.45	8339.44	5669.80
Duration of stay in the system for R200 work orders	6803.64	4832.72	4224.19
Flexo 1 Scheduled Resource Usage	0.9947	1.0201	1.0335
Flexo 2 Scheduled Resource Usage	0.7750	0.7960	0.8244
R200 Scheduled Resource Usage	0.6543	0.6616	0.8110
Rotoflex 1 Scheduled Resource Usage	0.7186	0.7524	0.7441
Rotoflex 2 Scheduled Resource Usage	0.8150	0.8264	0.8114
Flexo 1 Number of Waiting	5.5289	5.5347	6.5647
Flexo 2 Number of Waiting	3.8688	4.1896	4.2237
Quality control and slicing queue 1 Number of Waiting	7.7485	8.3406	3.7737
Quality control and slicing queue 2 Number of Waiting	9.8617	9.2639	4.4180
R200.Queue Number of Waiting	3.5849	4.4951	3.5630
Flexo 1 line. Queue Waiting Time	2713.94	2832.32	3127.57
Flexo 2 line. Queue Waiting Time	2276.79	2413.32	2350.73
Quality control and slicing queue1 Waiting time	3668.30	3813.45	1645.87
Quality control and slicing queue1 Waiting time	3622,77	3289.83	1320.51
R200 Machine line.Queue Waiting Time	2685.64	3225.43	2388.69

A % 37 reduction for time spent in system for R200 line work orders was observed. This simulation study with integrated planning process showed that a significant amount of improvement is possible without any investment.

As a result, especially in the study, planning phases became very important. Because work orders were being taken to productions processes without knowledge of about priority and deadlines of the orders. Work orders were being taken to the processes randomly or with the first in, first out method. The proposed plan has been implemented by the management and FIFO methodologies abandoned. Mixed sorting methodology in which orders are sorted by date of delivery in plate preparation process and printing process has been used. In addition, the method in which orders with less quantity are sorted first has been used in ROTOFLEX line.

REFERENCES

Altiok, T. & Melamed, B., 2007. *Simulation Modeling and Analysis with Arena.* New Jersey: Elsevier Inc..

Arena, 2012. *Getting Started With Arena®.* s.l.:s.n.

Banks, J., 1999. *Discreete Event Simulation.* Georgia: ed. P.A. Farrington, H.B. Nembhard, D.T. Sturrock, G.W. Evans pp. 7-13..

Banks, J., Carson, J. S., Nelson, B. L. & Nicol, D. M., 2005. *Discrete-Event System Simulation.* 5th Edition. USA: Prentice Hall, pp. 3-4-8-9.

Centeno, M. A., 1996. *An Introduction to Simulation Modeling.* Miami, Florida 33199, U.S.A., ed. J. M. Cbarnes, D. J. Morrice, D. T. Brunner, and J. J.Swain, p. 15.

Cho, S. & Eppinger, D. S., 2001. *Product development process modelling using advanced simulation.* Pennsylvania , s.n.

El-Haik, B. & Al-Aomar, R., 2006. *Simulation-Based Lean Six-Sigma and Design for Six-Sigma.* New Jersey: John Wiley & Sons.

Hasgül, S. & Büyüksünetçi, A. S., 2005. *Simulation Modeling And Analysis Of A New Mixed Model Production Lines.* Eskişehir, M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, eds., p. 1408.

Jägstam, M. & Klingstam, P., 2002. *A handbook for integrating discrete event simulation as an aid in conceptual design of manufacturing systems.* New Jersey, ed. E., p. 1.

Kelton, W., Sadowski, R. P. & Sweets, N. B., 2010. *Simulation With Arena.* 2 Edition dü. Boston: Mc Graw Hill.

Olcar, Z., 2014. *Process And Quality Improvament Using Work Methods And Simulation,* Istanbul: İstanbul Arel Universitesi.

Patterson, B. M., Ozbayrak & Mustafa, 2002. *Simulation of JIT Performance In a Printing Shop.* U.K., E. Yücesan, C.-H. Chen, J. L. Snowdon, and J. M. Charnes, eds., p. 1914.

Robinson, S., 2011. *Choosing The Right Model: Conceptual Modeling For Simulation.* Warwick Business School, S. Jain, R.R. Creasey, J. Himmelspach, K.P. White, and M. Fu, eds., p. 1433.

Sargent, R. G., 2011. *Verification and Validation of Simulation Models.* Syracuse, NY 13244, U.S.A., S. Jain, R.R. Creasey, J. Himmelspach, K.P. White, and M. Fu, eds., p. 188.

Shannon, R. E., 1998. *Introduction to The Art and Science.* Texas, D.J. Medeiros, E.F. Watson, J.S. Carson and

Sokolowski, J. A. & Banks, C. M., 2009. *Principles of Modeling and Simulation: A Multidisciplinary Approach.* Hoboken: John Wiley & Sons.

Takus, D. A. & Profozich, D. M., 1997. *Arena® Software Tutrial.* Pennsylvania 15143, U.S.A., ed. S. Andradóttir, K. J. Healy, D. H. Withers, and B. L. Nelson, p. 541.

Wainer, G., 2009. *Discrete-Event Modeling And Simulation: A Practitioner's Approach.* USA: CRC Press.

Warwick, S. R., 2011. *Choosing The Right Model: Conceptual Modeling For Simulation.* Stewart Robinson Warwick Business School University of Warwick Coventry, CV4 7AL, UK, S. Jain, R.R. Creasey, J. Himmelspach, K.P. White, and M. Fu, eds., p. 1432.

AUTHOR BIOGRAPHIES

Tolga Kudret KARACA obtained his B.Sc. in economics from Anadolu University, Eskisehir at 2008. He finished his M.Sc. in engineering management from Arel University, Istanbul in 2015. He currently studies Ph.D. in industrial engineering at Kadir Has University, Istanbul. His research interest areas are simulation, production management and risk analysis. He is currently working as main production and planning manager at Etisan Labeling and Packaging Company in Istanbul. tolga.karaca.tk@gmail.com.

Volkan ÇAKIR obtained his B.Sc. in electronics engineering from Turkish Air Force Academy, Istanbul at 1992. He obtained his M.Sc. in industrial engineering from Middle East Technical University, Ankara in 2001. He received his Ph.D. in engineering management at the Old Dominion University, Norfolk, Virgina in 2011. His research interest areas are simulation, statistical quality control, system dynamics and risk analysis. He is currently an assistant professor and head of the Industrial Engineering Department at Istanbul Arel University. volkancakir@arel.edu.tr.

A SIMULATION-BASED EVALUATION OF DYNAMIC TASK PRIORITIZATION IN MAINTENANCE MANAGEMENT

Dietmar Neubacher
Nikolaus Furian
Clemens Gutschi
Tobias Elmer
Siegfried Vössner
Graz University of Technology
Department of Engineering- and Business Informatics
Kopernikusgasse 24, 8010 Graz, Austria
email: dietmar.neubacher@tugraz.at

KEYWORDS
Discrete Event Simulation, Maintenance Management, Hiarachical, Decision Support System

ABSTRACT

Reliability and availability of machines are crucial for an efficient manufacturer and great maintenance operations are the key. Due to the advancement of condition monitoring and predictive analytics more information is available and the field shifts from corrective to preventive task fulfillment. Nevertheless, well trained experts are rare and responsible for different machines at the same time. Most experts have already identified that the sequence of task execution could heavily influence the production output. However, in practice only simple heuristics are used, yet they still improve the performance significantly. The question arise if a prioritization of work orders could be further enhance using bottleneck detection methods. To show the potential of these techniques, a simulation based study is conducted. First, the reliabilities of selected bottleneck detection methods are tested on a simplified line model. Second, the performances of all prioritization policies are evaluated and compared using a real industrial use case. Contrary to the expectations, a simple heuristics has shown a good performance. Nevertheless, applying the Active Period Method has significantly outperformed the other approaches and the bottleneck ranking has proven to be a very good indicator to prioritize work orders.

INTRODUCTION

Manufacturers have to secure high-quality products and low costs to stay competitive. These competing goals can be achieved if machines are efficiently utilized and in perfect condition. The complexity of automatized production facilities requires much effort in maintenance planning. However, as new trends strive towards flexible production systems, the complexity is going to rise dramatically and practicable methods are needed to support decision makers (Neubacher et al. 2016).
The advancement of condition-monitoring, forecasting and prediction has significantly shifted the field from Corrective Maintenance (CM) to Preventive Maintenance (PM). Wang (2002) compared several maintenance policies and stated their advantages and disadvantages. Additionally, it is mentioned that a efficient maintenance strategy incorporates more than one policy. Though, huge and highly responsive *firefighting* units have become obsolete, as tasks can more often be scheduled in advance and capacities can be aligned.
As a result of the high degree of automation, the ratio of machines to human workforce is often high. Consequently. workers have to inspect and monitor many machines at the same time. The prioritization of workorders becomes a crucial task, as the sequence of execution could have a significant impact on the performance of the system. A random execution might potentially extend the downtime, cause losses and decrease the overall efficiency of production facilities. The importance of task prioritization is already recognized in industrial communities and most of them have internal policies to determine the optimal sequence. Since there are so many aspects that have to be considered when determining priorities, policies make usually use of heuristic rules or common sense derived from human expert knowledge (Yang et al. 2006).
These heuristics perform very good for simple production lines, but as the complexity rise, they can lead to non-optimal or even bad decisions. Therefore, this paper first investigates common bottleneck detection methods and proposes to use these approaches to prioritize tasks. The reliability of these bottleneck detection methods is evaluated on a simply simulation model. Additionally the prioritization policies will be tested using a simulation of a real production line layout[1]. Finally the paper concludes with a recap and a recommendation for task prioritization policies.

[1]This paper is based on the master's project of Elmer (2017)

MAINTENANCE PLANNING

According to EN 13306:2010-10-01 (2010), the term maintenance is defined as a *combination of all technical, administrative and managerial actions during the life cycle of an item intended to retain it in, or restore it to, a state in which it can perform the required function.* Following this definition, there is way more to consider beside the actual repair task. An efficient maintenance planning and control systems includes various areas. *Job Planning and Execution* manages all preventive and corrective maintenance tasks and considers task scheduling, capacity and cost planning. *Asset management* handles historical and actual data or information about all assets. *Material Management* is necessary to keep track of spare parts and other components that are needed to ensure the availability of assets. *Resource management* manages personal and operating resources to carry out relevant function. Finally, *Analysis and Reporting* is used to enhance the maintenance by documenting repair tasks and enables statistics for further planning.
However, all activities have to be aligned in order to secure low maintenance costs and less downtime. Especially the costs for not-availability can have a significant impact and are hard to predict. To measure the availability, two common metrics are used, the Mean Time Between Failure (MTBF) and Mean Time to Repair (MTTR). During the normal operation a machine can be either in *production*, *Starved*, *Blocked* or in a *Failure Mode*. With exception of the last, all remaining states are considered to be active. Thereby, the MTBF summarize only the active states and represents an average period between two failures.

$$MTBF = \frac{\sum Operation\ Time}{Number\ of\ Machine\ Failures} \quad (1)$$

In contrast, the MTTR expresses the average period of a downtime and does inherently consider often more than just the repair activity itself. More precisely, an unpredicted breakdown often leads to a much longer downtime as a planned shutdown. In case of the sudden disruption, an operator needs to recognize and diagnose the failure. Additionally, the needed spare parts have to be gathered and afterwards the machine needs to be tested.

$$MTTR = \frac{\sum Repair\ Time}{Number\ of\ Repairs} \quad (2)$$

The sum of MTTR and MTBF is the *Planned Production Time*, which is used to define the inherent availability of a single machine (Ebeling 2009). The most important key performance indicator in maintenance management is the Overall Equipment Effectivenenss (OEE), which allows an objective evaluation of production systems. In order to calculate this factor, the availability is multiplied by the *Performance Efficiency* and a *Quality Rate*. While the first factor considers losses during the production process, the second is a ratio of parts that passes the quality inspection and those not.

Despite the straightforward calculation for a single machine, this indicator cannot be aggregated along entire production lines. Especially, because such systems often consist of many different machines, which are linked to each other by rigid or flexible transportation devices. As a result of this linkage and cascading effects, a single machine breakdown can amplify its impact on the performance of the production line. In general a bottleneck machine limits the throughput and every stop of this entity will directly influence the performance. Commonly, a static bottleneck can be identified as the machine with the longest cycle time. But, resulting from the previously mentioned cascading effects, a single machine breakdown could drain or fill buffers and the bottleneck will be forced to stop. Furthermore, due to stochastic behavior and breakdowns the actual bottleneck machine varies over time (Roser et al. 2003). As a result of this situation, a decision which maintenance task should be executed first has to be reevaluated every time a new work order arises, or the bottleneck shifts. Depending on the number of machines and frequency of failure occurrence, this procedure can be very complicated and time consuming. But still, the task assignment is just one area of maintenance planning and other aspects have to be considered as well. But an efficient task prioritization with a focus on bottleneck situations can support the decision making process significantly.

PRIORITIZATION POLICIES

This paper aims for a practicable, quantitative method to assign priorities and support decision makers. But first, it is crucial to identify a measurable criteria to evaluate the performance of prioritization policies. According to Wang (2002) common criteria are the minimization of costs, failure rate and downtime, or the maximization of availability and reliability. As some production facilities are working with a make-to-order policy a maximization of throughput does not necessarily make sense. Therefore, we added a maximization of the degree of logistic or demand fulfillment. Nevertheless, for the final experiments the improvement of throughput is used to compare the performance of different prioritization policies.

In this section three prioritization policies are explained and their advantages and disadvantages are stated. First the *First-In, First-Out* policy is mentioned. Despite this is no real prioritization strategy, it is frequently used, if no other approach is available. Second, some *Heuristic Methods* are explained, because they are often used by decision makers in practice. Finally, the proposed *Dynamic Bottleneck Detection Methods* are explained and compared regarding their practicability to prioritize maintenance work orders.

First-In, First-Out (FIFO)

This method originated as a service policy in queuing theory, which states that requests are processed in the order they arrive. Despite it is not a real policy in maintenance management, it is often common practice that corrective repair actions are executed using this policy if no other rules are applied. For the remainder of this section, the FIFO Method will be used as a baseline to compare the performance of other policies. Furthermore, if other prioritization lead to the same priority for different tasks the work orders will executed following the FIFO principle.

Heuristic Methods

In practice, maintenance units already know that a FIFO strategy is no efficient way to schedule all maintenance tasks. Based on insight and understanding of the system, most workers are capable to prioritize tasks even without a sophisticated prioritization method. The following approaches are easy to apply since all required data is easily accessible. Although, these methods do not provide much information about the actual performance of machines and cannot be used to reliably detect bottlenecks.

Part-Out-Part-Out Time
This is often called the *true cycle time* and defines the timespan from a part leaving the machine until the exit of the following part. Therefore, this true cycle time does also take machine failure, blocking and starving activities into account. Despite this very simple mechanism, it is often very complicated to calculate the true cycle time of a machine, due to the stochastic effects within the production line. In order to have statistical confidence, an approach is to record a set of production cycles and define an expected value as a suitable true cycle time. A big drawback of this methodology is that it does not differ between internal (blocking and starving) or external (failure) reasons.

Availability
This method is commonly used in practice, as most IT-Systems do already provide the required informations for decision makers. A machine is available if it is capable to produce during the planned production period. That means that the only state that reduces the availability is a breakdown. Equation (1) and (2) can be used to calculate the availability (3).

$$Availablility = \frac{MTBF}{MTBF + MTTR} \tag{3}$$

The biggest drawback of this method is that it does not take into account any effects caused by the dynamic behavior of the production line. For instance, if a machine is frequently blocked because its successor has a much higher cycle time, the effective performance of this machine decreases significantly, but the availability does not. Therefore, a change of the availability is a good indicator to evaluate the effectiveness of maintenance management, but not to estimate the performance of the machine itself.

Redundancy
If there are many machines executing the same procedure within a production system, the number of such redundant machines can be used to prioritize maintenance tasks. The big advantage is that there is still a production flow, if a redundant machine is switched off, in contrast to a complete stop, when a single machine is down. An simulation based experiment conducted by Neubacher et al. (2016) shows that it is possible to estimate the impact of a single machine downtime on the performance, even for flexible production systems. However, this approach is very time consuming and consequently not practicable to prioritize every maintenance task in real-time. Despite this drawback, the results confirmed that a basic prioritizing can be done by using the cycle time and the number of redundant machines.
In practice these heuristic methods are often combined together. However, with the rise of the digitalization much more information is available and more data can be processed in real-time.

Dynamic Bottleneck Detection

The aim of the following section is dynamically identify the bottlenecks and assign the corresponding priority to work orders for these machines. The basic concept behind is that an enhancement of a non-bottleneck machine will have a lower impact on the overall system improvement, as a change on the bottleneck machine. The following equation shows a mathematical formulation for the bottleneck definition. $\Delta TP_{sys,i}$ is the system's throughput increment which is caused by an improvement of machine i and ΔTP_i is the single machine's throughput increment.

$$\Theta_{max} = \max\left(\frac{\Delta TP_{sys,1}}{\Delta TP_1}, \frac{\Delta TP_{sys,2}}{\Delta TP_2}, \cdots, \frac{\Delta TP_{sys,n}}{\Delta TP_n}\right) \tag{4}$$

According to (4) a bottleneck can be identified as the machine with the highest sensitivity value Θ_{max} (Chang et al. 2006). Nevertheless, this bottleneck can change over time and priorities have to be adapted accordingly.

Blocking & Starving Probability
The previous mentioned static bottleneck detection method is very precise, but it has the big disadvantage that it is not possible to determine a bottleneck based on real-time production data. Therefore the first proposed method uses the probability that a machine will be blocked or starved to identify the limiting entity of a production system. Generally, a bottleneck will cause

all preceding (upstream) machines to be blocked and all succeeding (downstream) machines to be starved. According to Kuo et al. (1996), arrows can be assigned to indicate the direction in which to bottleneck is located. For instance, if the blockage of a machine is greater than the starvation probability of the next downstream machine, the bottleneck is located downstream the line.

$$TB_j > TS_{j+1} \qquad j = 1, \cdots, n-1 \tag{5}$$

In contrast, if the blockage is smaller the starvation prob of the downstream machine, the bottleneck is located upstream the line.

$$TB_j < TS_{j+1} \qquad j = 1, \cdots, n-1 \tag{6}$$

Following this mechanism, the bottleneck can be identified, if two arrows are pointing towards a machine. In case of multiple bottlenecks within a production system, Kuo et al. (1996) introduced a *Bottleneck Severity* to define a primary bottleneck. In case of redundancy, the method uses average values of the blocking and starving probability of the machines within an Operation Sequence (OS).
For complex production systems Li et al. (2007) introduced another bottleneck detection method, which is also based on the blocking and starving probability. In contrast to the method of Kuo et al. (1996), they analyzed trends and identified the bottleneck as a turning point between blockage and starvation. For cases of multiple bottlenecks, they used a bottleneck index I to identify the primary bottleneck. The advantages and disadvantages of this method are explained during the simulation based comparison in the following section.

Active Period Method
The underlying idea of this method is that the longer an entity is working without interruptions, the more likely it is the bottleneck of the system. Therefore, the actual bottleneck of a production system is the machine which has the longest uninterrupted active period at this point in time. Therefore this method is suitable to detect and monitor the momentary bottlenecks as well as the shifting bottlenecks at any time (Roser et al. 2001).
The first step of this method is to classify all possible states of an entity into active or inactive. As illustrated in table 1, this method can be used for many different purposes and systems. However, for manufacturing machines the active states are processing, repair, changing tools, and service. Only the starved and blocked states are classified as inactive. The greatest benefit of this method is that it does not need any information about the structure of the system.
Most conventional methods calculate the percentage of production time, but this approach tracks the duration of an active state of a machine. Remarkably, this period is not interrupted by a repair or a tool change, only by waiting activities, such as blocking and starving.

Table 1: States of System Entities

Entity	Active States	Inactive States
Machine	Processing, Repair, Service, Tool-change	Starved, Blocked
AGV	Moving to location, Recharging, Repair	Waiting
Human Worker	Working, Recovering	Waiting
Supply	Obtaining new part	Blocked
Output	Removing Part	Waiting

To determine the actual bottleneck of a system the active durations of all machines will be compared. If there are more entities active in a system, the bottleneck for the current period has to be the machine with the longest uninterrupted active period.
Over a certain period of observations, a machine i has n active periods of which each period has a duration of $a_{i,j}$. This results in a set of durations A_i for each machine:

$$A_i = \{a_{j,1}, a_{j,2}, \cdots, a_{j,n}\} \tag{7}$$

The Average Active Period (AAP) of a machine i over this period of observations is calculated as follows:

$$\overline{a_j} = \frac{\sum_{J=1}^{n} a_{i,j}}{n} \tag{8}$$

Following this steps, the machine with the longest average period $\overline{a_j}$ is identified as the current bottleneck of the system. Roser et al. (2001) stated that this mechanism works reliable for steady state production systems. To determine shifting bottlenecks, the method has to be enhanced to identify overlapping phases of two active entities. During this period no unique bottleneck is in the system and both machines will be denoted as shifting bottlenecks. In order to explain this concept the following example is used.
Table 2 shows a production data for a period of 9 observations. I order to transform this production data, i is set to be the number of machines and j to be the number of observations. This production data is transformed in a state matrix A, with i columns and j rows:

Table 2: Production Data for 9 Observations

t	M1	M2	M3
0	Produce	Starved	Starved
1	Repair	Starved	Starved
2	Produce	Produce	Starved
3	Blocked	Tool Change	Produce
4	Blocked	Produce	Produce
5	Produce	Starved	Starved
6	Produce	Starved	Produce
7	Produce	Produce	Starved
8	Blocked	Produce	Produce

$$A = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 \end{pmatrix} \quad (9)$$

In the next step the values from (9) have to be accumulated to get the current duration $b_{i,j}$ of the active period of each machine.

$$b_{i,j} = a_{i,j} * (b_{i,j-1} + 1) \quad (10)$$

Bye applying (10) on (9) a accumulated state matrix B is calculated.

$$B = \begin{pmatrix} 1 & 2 & 3 & 0 & 0 & 1 & 2 & 3 & 0 \\ 0 & 0 & 1 & 2 & 3 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 & 2 & 0 & 1 & 0 & 1 \end{pmatrix} \quad (11)$$

In order to determine the bottleneck of the system, the machine with the longest active duration $b_{i,j}$ has to be identified in each time step. If a machine has the longest accumulated active period, the value $c_{i,j}$ will be set to 1 in the bottleneck matrix C.

$$C = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (12)$$

As illustrated in the bottleneck matrix C, in the third and eighth time step, two machines are determined as bottlenecks. In order to distinguish between a sole and a shifting bottleneck, further steps are necessary.
The proposed approach uses the *Accumulated State Matrix B* (9) and (3) to calculated the AAP. In this case the machine with the longest active duration $\overline{a_j}$ is assigned a priority value of 1 and the other machines according to their active period related to the longest $\overline{a_j}$.

System Sensitivity Analysis
In terms of accuracy, this method is the best of all investigated methods. Hence, it is often used to validate other bottleneck detection methods, but seldom used in practice. The big disadvantage of this approach is that it requires the use of simulation and multiple runs. Consequently, every time a bottleneck needs to be detected, an experiment has to be conducted and requires actual system parameters (e.g. buffer levels, cycle times,...). Nevertheless, a simulation model is created in order to evaluate the reliability of the previous explained bottleneck detection methods.

Additional Factors

In order to transform the bottleneck detection to a prioritization algorithm several additional factors have to be considered. However, it should be noted that a variation of these values can influence the prioritization and consequently the production system's performance significantly. A comprehensive analysis of the impact of each parameter is beyond the scope of this paper, but in order to give a short insight the most relevant effects are mentioned.

Time of Observation
For all scenarios the period of observation will have a great impact. If the sample is very long, the bottleneck detection will focus on long term bottlenecks and strive towards a static bottleneck detection for stochastic systems with less variance. The first simulation runs indicated that T_{obs} is a very critical factor and if the period is to short, uncritical machines might be preferred.

Priority Increase Factor
If a production line is tremendously understaffed, some work orders could decay in the queue. Therefore, dynamic priority increase factors are common for many services, but - to the knowledge of the author - the influence on the prioritization policy for maintenance tasks has not been investigated yet. Therefore, we add p_{inc} as a linear priority increase for outstanding repair orders.

$$P_{real} = P_{calc} + p_{inc} * (t_{assign} - t_{breakdown}) \quad (13)$$

The first experiments have shown that this value is strongly dependent on the period of observations. If T_{obs} is very long, non-bottleneck machines will receive not enough priority and the overall performance drops. Several experiments were conducted to find an optimal gradient p_{inc}. Following the findings of Neubacher et al. (2016), the impact of an hour of downtime is likely to vary considerably, depending on the number of redundant machines. Nevertheless, in the course of the following comparison the same gradient is set for all machines.

Prioritization Frequency
This factor determines how often the priority is evaluated. In practice, it could be performed every time a new breakdown occurs, or a work order has to be scheduled. But, as the execution of the algorithm takes a short amount of time, an update frequency f_{Prio} is introduced to speed up the simulation.

Degree of Logistic Fulfillment
This factor is important for centralized organizational structures. If there are more production lines running at the same time, work orders from different lines have to be weighted. In practice, manufacturers defined priority lines and favor work orders from this production lines over all others. Most of the time this *all or nothing* principle does also prefer low-priority tasks on priority lines, over important tasks on non-priority lines. In order to compare work orders from different production lines, we propose a quantitative weighting parameter for every work order. Each manufacturing line has a capacity which is planned for a specific period. The *Degree of Logistic Fulfillment* $r_{i,j}$ will set a focus on production lines, which are more likely to miss the planned capacity. However, this paper will focus on the task prioritization within a single production line and $r_{i,j}$ will not be used.

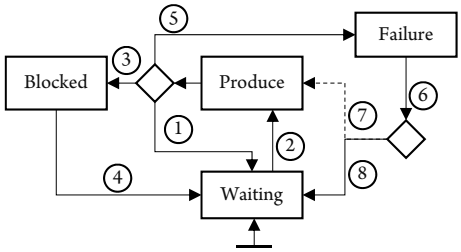


Figure 1: Activity Diagram of a Machine, according to Neubacher et al. (2016)

HIERARCHICAL SIMULATION MODEL

The simulation model is based on the hierarchical control model from Neubacher et al. (2016). The big advantage of this Discrete Event Simulation (DES) approach is the flexibility through hierarchical control structures and the possibility to include minor agent-based aspects if needed (Furian et al. 2015).

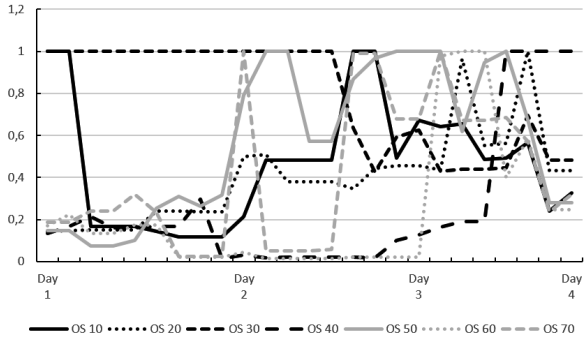
As illustrated in figure 1, a machine can perform several activities. If a failure activity starts, a repair request is sent to the corresponding Control Unit (CU) and will be queued according to the priority of the machine. In the experimental model, there are only two CUs implemented. A CU of an OS controls the material handling and dispatching within the sequence. On the highest level, a Line-CU handles the material flow between machines and dispatches incoming repair orders. Every time a new order arrives, it will be ranked according to the actual priority of the machine. If a worker is idle, the highest ranked work order will be automatically assigned and the worker starts to repair the machine.

The worker and the material are simple state-machines and can either be busy or idle. Furthermore, this experimental model is limited to a single production line, but in ongoing research projects higher hierarchical levels are implemented and allow centralized service units to be shifted between production lines.

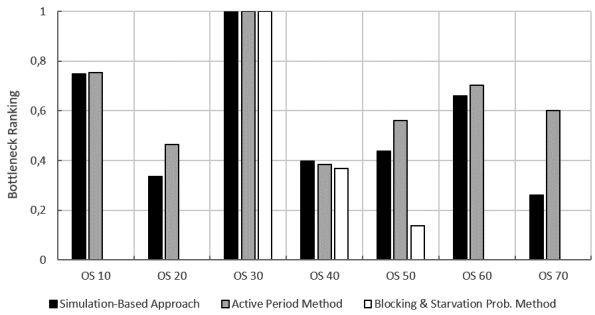
This conceptual model is used to create several production lines. In order to evaluate the reliability of the bottleneck detection methods, a simplified production line with seven machines is used. In the next step an industrial use cases with realistic data is used to evaluate the performance of the prioritization policies. In advance of all experiments the production data was sampled and used for every prioritizing policy. Finally, for every experiment $n = 15$ different production data sets are evaluated.

Comparison of Bottleneck Detection Methods

The Blocking & Starving Probability (BS) method is actually pretty fast, but has great disadvantages. First, the entire structure and information of the production line has to be taken into account. Second, as shown in



(a)



(b)

Figure 2: Results of the Simple Production Model
(a) Dynamic Bottleneck Situation
(b) Performance of Bottleneck Detection

figure 2(b) the BS method was only capable to identify the relevant bottlenecks. But during the experiment also OS 10 and OS 60 were bottlenecks for a significant amount of time and not even recognized. Furthermore, there are only values for three bottlenecks, which makes it impossible to prioritize all other machines.

The Active Period (AP) method works great for static and dynamic systems. Despite the effort for implementation, this method can deal with production data and does not need any information about the systems structure. Figure 2(b) also indicates that this method is very precise and determines the bottleneck rating of all machines almost perfect. During the conducted experiment the bottleneck situation shifted several times, which is illustrated in figure 2(a). The AP method was able to reproduce this shifts in contrast to the BS approach.

Comparison of Prioritization Policies

The three Heuristic Methods (HM) are very good for static systems and there is almost no effort for implementing this basic prioritization policies in companies. Nevertheless, for dynamic and stochastic systems HM are not suitable and cannot find an optimal solutions. The bottleneck detection methods seem to be better, but there is a certain gap between the AP and the BS

Table 3: Comparison of Prioritization Policies

	HM	BS	AP	SysSens
Accuracy for static systems	+	+	+	+
Accuracy for dynamic systems	-	o	+	+
Effort for implementation	+	o	+	-

method. Especially for very dynamic bottleneck situations, the BS method is not able to detect all relevant bottlenecks and makes it almost useless for prioritization. The AP approach is capable to identify bottleneck shifts and compared to other mechanisms, it is easy to implement. Finally, the *System Sensitivity* is the most precise method, but the efforts to execute a simulation study every time a decision has to be made, limit its practicability.

Industrial Case

Finally we abstracted a real car engine manufacturing line and build a simulation model to identify the efficiency of these prioritization policies. We also used the real production data and cycle times in order to create a valid model of a real industrial use case. This investigated manufacturing line consists of 26 OS and contains 48 machines. Relevant production data was sampled from a historical set of one year and the model was validated using historical throughput. Following the information of the project partner, every breakdown that exceeds 10 minutes will be treated as a repair order. All stops shorter are assumed to be minor faults, which are directly handled by the machine operator. The simulation period is 90 days with a warm-up of 7 days. To achieve statistical confidence 15 runs are performed.
In Figure 3 an excerpt of the first three days after the warm-up period is illustrated. Following the static bottleneck ranking (see figure 3(a)), *M12* seems to be the most critical machine. But in course of the simulation, other machines have a much higher priority than the initial bottleneck. As illustrated in figure 3(b), during the first day *M4* has already a higher priority than *M12*. At the end the *M12* is only ranked fourth. Noteworthy, the initially last prioritized *M22* also passed the initial bottleneck and is ranked third at the end of day three. Similar behavior occurs in many simulation runs, also for completely different production lines. Such highly dynamic bottleneck situations lead to great challenges for prioritization policies. To evaluate their performance, some final parameters for the experimental are set.
The HM uses a combination of *Availability* and *Redundancy*, which has proven to be the best. After interviews with decision makers on the manufacturing line,

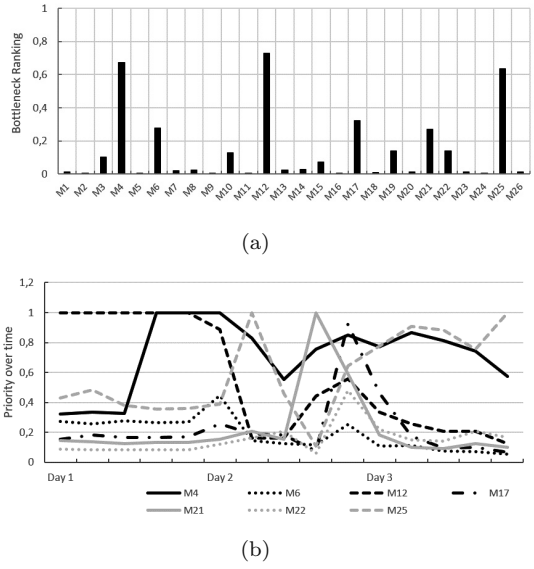


Figure 3: Bottleneck Detection and Prioritization (a) Static Bottleneck (b) Dynamic Priorities

this heuristic comes close to a real policy. Moreover, previous simulation runs have shown that using just one of these methods always leads to a poor performance.
For the bottleneck prioritization policies the period of observation is set to seven days and the priority increase factor to 0,1 per hour. Furthermore the prioritization will be reevaluated every 0,2 hours. The same Time Between Failure (TBF) and Time to Repair (TTR) sequence is used to compare the strategies. Therefore, each run starts with a simple FIFO strategy to determine the baseline throughput. Afterwards the same sequence is used to calculate the output and work orders are dispatched using the priorities from HM, BS, and AP. After all policies are simulated, a new set of TBF and TTR is sampled and the evaluation is repeated.

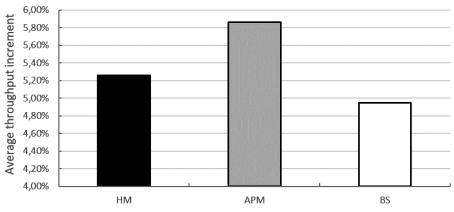


Figure 4: Average Throughput Increment of Different Task Prioritization Policies

The result of the experimental runs are presented in figure 4. In all runs a task prioritization strategy exceeded the performance of a plain FIFO policy. Despite its very simple rules, the HM has proven to be very efficient and the throughput is increased by 5,26% on average. However, the BS method could not meet these expectation. Nevertheless, in comparison to a simple

FIFO strategy, the throughput is also 4,95% higher. The drawback of the BS approach is that it does not give any feedback about the criticality of most machines and non-critical entities are maintained to late. By improving the throughput by 5,86% the AP method shows the best results. The only drawback is that this method is sensitive to changes in the parameters T_{obs} and p_{inc}. However, the algorithm works pretty fast and it is possible to optimize these parameters quickly. Nevertheless, this method has many other benefits and can be used for many different production lines.

CONCLUSION

According to Wang (2002) a good maintenance strategy incorporates various policies in order to efficiently react to unpredicted breakdowns and be proactive when needed. Independent of these policies, available work force has to be assessed and incoming orders have to be assigned. Simple FIFO strategies are very inefficient, as some machines are more crucial than others. In favor of a constant work flow and a high throughput, work orders have to be prioritized and efficiently planned. The bottleneck is the most crucial machine in a production line, but as shown in the conducted experiments, the initial is not always the primary bottleneck. In practice there are other machines that block the real bottleneck quite fast and simultaneously have a much lower reliability. These machines are crucial and conventional bottleneck detection methods are not able to identify them.
In this paper several task prioritization policies are compared and evaluated using simulation models of production lines. In the first step bottleneck detection methods were evaluated regarding their capability to identify the actual bottleneck at any point in time. After this analysis an industrial use case was implemented in order to test the performance of task prioritization policies on a real use case. In comparison to a simple FIFO policy all strategies have shown a significant throughput increase. A heuristic approach that is often used in practice produced a significant increase and confirms that most maintenance units achieve good results using this policy. The prioritization according to the *Blocking & Starving Probability* also boosts the performance of the production system, but simultaneously had the least increase of all evaluated prioritization policies.
Finally, linking the task priority to the average active period of a machine has shown the best results and lead to an increase of 5,86% in production throughput. Additionally, this method is generic and can be applied on many different production lines, as it is completely independent of the line structure. Hence, this method is the only one, which can deal with the flexible production systems. Adding to the statement of Roser and Nakano (2015), this method is not only practicable to reliably detect bottlenecks, rather it can be used to prioritize maintenance work orders most efficiently.

REFERENCES

Chang Q.; Ni J.; Bandyopadhyay P.; Biller S.; and Xiao G., 2006. *Supervisory Factory Control Based on Real-Time Production Feedback. Journal of Manufacturing Science and Engineering*, 129, no. 3, 653–660.

Ebeling C.E., 2009. *An Introduction to Reliability and Maintainability Engineering.* Waveland Pr Inc.

Elmer T., 2017. *Dynamic Task Prioritization in Maintenance Management.* Master's thesis, Graz University of Technology.

EN 13306:2010-10-01, 2010. *Maintenance - Maintenance terminology.*

Furian N.; OSullivan M.; Walker C.; Vössner S.; and Neubacher D., 2015. *A conceptual modeling framework for discrete event simulation using hierarchical control structures. Simulation Modelling Practice and Theory*, 56, 82–96.

Kuo C.T.; Lim J.T.; and Meerkov S., 1996. *Bottlenecks in serial production lines: A system-theoretic approach. Mathematical Problems in Engineering*, 233–276.

Li L.; Chang Q.; Ni J.; Xiao G.; and Biller S., 2007. *Bottleneck Detection of Manufacturing Systems Using Data Driven Method.* In *2007 IEEE International Symposium on Assembly and Manufacturing.* 76–81.

Neubacher D.; Furian N.; Gutschi C.; and Vössner S., 2016. *A Hierarchical Control Simulation Model to Support Maintenance Planning in Flexible Production Systems.* In *Proceedings of the 2016 European Simulation and Modeling Conference.* 420–425.

Roser C. and Nakano M., 2015. *A Quantitative Comparison of Bottleneck Detection Methods in Manufacturing Systems with Particular Consideration for Shifting Bottlenecks*, Springer International Publishing.

Roser C.; Nakano M.; and Tanaka M., 2001. *A practical bottleneck detection method.* In *Proceedings of the 2001 Winter Simulation Conference.* 949–953.

Roser C.; Nakano M.; and Tanaka M., 2003. *Comparison of bottleneck detection methods for AGV systems.* In S. Chick; P. Sanchez; D. Ferrin; and D. Morrice (Eds.), *Proceedings of the 2003 Winter Simulation Conference.* vol. 2, 1192–1198.

Wang H., 2002. *A survey of maintenance policies of deteriorating systems. European Journal of Operational Research*, 139, no. 3, 469 – 489.

Yang Z.; Chang Q.; Djurdjanovic D.; Ni J.; and Lee J., 2006. *Maintenance Priority Assignment Utilizing On-line Production Information. Journal of Manufacturing Science and Engineering*, 129, no. 2, 435–446.

SUPPLY CHAIN SIMULATION

MODELLING AND SIMULATION FOR DECENTRALIZED SUPPLY CHAIN FORMATION

Florina Livia Covaci
Business Information Systems Department
"Babes-Bolyai" University
Str.Teodor Mihali, Nr.58-60, 400591, Cluj-Napoca, Romania
E-mail:florina.covaci@ubbcluj.ro

KEYWORDS

Supply Chain Modelling, Agent-based Simulation, Decentralized supply chain

ABSTRACT

During the supply chain formation process there are multiple commercial parties involved at different levels in the network. Parties that are on the same level are usually competitors, and the companies that are linked vertically in supply chains have a trading relationship as suppliers-consumers. All the parties that are connected in the supply chain are individual entities that have self-interest and they may compete for gaining a commercial advantage. Given such complex relations the current work aims to find a mechanism for an allocation in supply chain network for the participants. This paper focuses on modelling and simulation of the supply chain formation in a network of potential participants in order to establish and enforce contract parameters between each pair of component consumer/supplier and to facilitate end-to-end contract parameters. The Supply Chain Formation problem is translated in terms of an acyclic graph where the nodes are represented by suppliers/consumers. The initial graph is then transformed in a cluster graph and by message exchange between the agents the contract parameters values are propagated along the network from the underlying suppliers towards the root.

INTRODUCTION

The dynamism of supply networks that the new industrial revolution promises requires innovative mechanisms for modelling of automated supply chain formation. Supply chain formation process consists of business interactions that can be quickly and flexibly formed and easily dissolved to better respond to rapidly changing market conditions. As a result, some make-to-order supply chain management models have emerged. Make-to-order represents a production approach in which a confirmed order for a product is received, and then the product is tasked to be built. This approach is widely used for highly configured products (Holweg and Pil 2004). These dynamic supply chains are often short-lived – quickly formed per project and then dispersed when the service is no longer needed. They are characterized by high-speed, automated formation using software agents (Lo and Kersten 1999; Buyya et. al. 2000).

The Supply Chain Formation (SCF) problem has been tackled in multiple works in the multi-agent systems literature. Multiple contributions can be found in the literature 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) based on the application of Pearl's belief propagation algorithm (Pearl 1988). This inference scheme was applied to the SCF problem, noting that the passing of messages is comparable to the placing of bids in standard auction-based approaches. It used iterative stages of message passing as a means for estimating the marginal probabilities of nodes being in given states: 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.

Although the LBP approach that has been used in (Winsper and Chli 2010), (Winsper and Chli 2012), (Winsper and Chli 2013) as a solution for supply chain formation has a lot of advantages over other approaches, it has the limitation that the proposed model uses only cost for pairwise agents and doesn't addresses issues such quality or delivery constraints. The current approach goes further and proposes the use of 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.).

The current work considers the problem of supply chain formation as a form of coordinated commercial interaction. The considered supply chain scenario represents a network of production and exchange relationships that spans multiple levels of production or task decomposition. The agents are characterized in terms of their capabilities to perform tasks, and their interests in having tasks accomplished. A central feature in the considered scenario is hierarchical task decomposition (Walsh and Wellman 2003). In order to perform a particular task, an agent may need to achieve some subtasks, which may be delegated to other agents. These may in turn have subtasks that may be delegated, forming a supply chain through a decomposition of task achievement. Constraints on the task assignment arise from the underling suppliers' network as exemplified in Figure 1

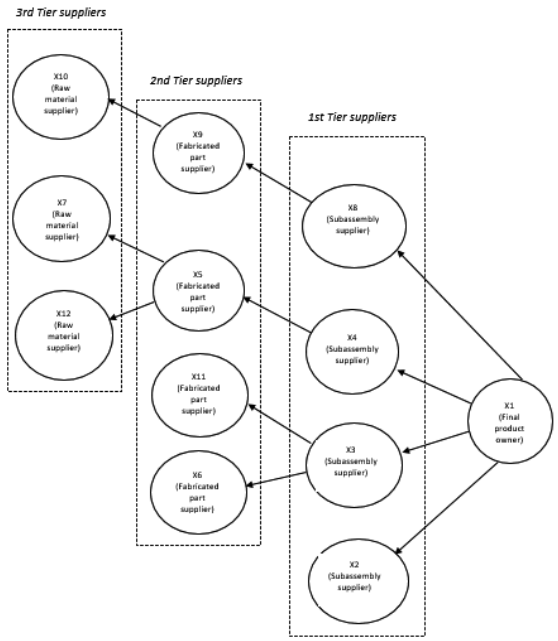


Fig. 1 Example of supply chain with hierarchical task decomposition

The paper is structured as follows: section 2 provides the background for supply chain formation, section 3 describes the model proposed for automated supply chain formation, section 4 provides simulation results and finally section 5 provides conclusions and future work.

BACKGROUND

The Supply Chain Formation (SCF) problem has been widely studied by the multi-agent systems community using computational agents that act in behalf of the participants during the SCF process and making possible to form SCs in a fraction of the time required by the manual approach (Davis et. al. 1983), (Walsh and Wellman 2000), (Collins et.al. 2002), (Walsh and Wellman 2003), (Norman et. al. 2004),(Cerquides et.al. 2007), (Giovannucci et. al. 2008), (Winsper and Chli 2010), (Mikhaylov et. al. 2011) (Winsper and Chli 2012), (Winsper and Chli 2013).

SCF methods can be classified in three categories depending on the architecture they follow. A first division is to separate SCF into centralized and decentralized architectures. Furthermore, we can separate the decentralized methods into two further categories depending on whether the communication between participants is either direct or mediated.

In a centralized approach (Walsh and Wellman 2000), (Collins et.al. 2002), (Cerquides et.al. 2007), (Giovannucci et. al. 2008), (Mikhaylov et. al. 2011), participant agents inform a central authority of their preferences (encoded as offers). After collecting the offers of all participant agents, the central authority determines the resulting SC.

Decentralized SCF appears as an alternative to centralized SCF in order to overcome some of its limitations as: participants might be reluctant to share this information with any central authority, given the hardness of the SCF problem centralized optimal solvers might suffer from scalability issues, the existence of a central authority introduces a single point of failure for the SCF process.

One approach to decentralized SCF is that of mediated SCF. 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 2000), (Walsh and Wellman 2003). The authors proposed a market protocol with bidding restrictions referred to as simultaneous ascending (M+1)st price with simple bidding (SAMP-SB), which uses a series of simultaneous ascending double auctions. SAMP-SB was shown to be capable of producing highly-valued allocations solutions which maximize the difference between the costs of participating producers and the values obtained by participating consumers over several network structures, although it frequently struggled on networks where competitive equilibria did not exist. The authors also proposed a similar protocol, SAMP-SB-D, with the provision for de-commitment in order to remedy the inefficiencies caused by solutions in which one or more producers acquire an incomplete set of complementary input goods and are unable to produce their output good, leading to negative utility.

Another approach to decentralized SCF is Peer-to-Peer (P2P), where each participant agent communicates directly with the participant agents representing its potential buyers and sellers. Therefore, the SCF process takes place between participant agents with no intervention of any third party, thus preserving participants' privacy since they only need to share their preferences with local trusted parties rather than communicating them to a central authority and it offers better scalability for large scenarios due to the fact that each participant is responsible of a small part of the computation.

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.

MODELLING FOR SUPPLY CHAIN FORMATION

In order to model the supply chain formation process, the present work translates the possible trading relationships into a graph and proposes a mechanism to find allocations in the supply chain network.

The current approach uses a message passing mechanism for the values of the contract parameters that the agents are sharing. 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.

The following paragraph provides a formal description of the supply chain formation problem in terms of a directed, acyclic graph (X, E) where $X = \{X_1, X_2, ..., X_n\}$ denote set of participants in the supply chain represented by agents and a set of edges E connecting agents that might have a commercial relationship.

Notation v_i represents the expectation of a participant in the supply chain on issue i of the contract and U(v) the utility that a participant obtains by receiving the actual value $v = (v_{i1}, v_{i2}, ..., v_{ik})$.When a supplier (seller) negotiates with a consumer (buyer), both parties are interested in obtaining those contract values $v = (v_{i1}, v_{i2}, ..., v_{ik})$ that maximize their utility functions U(v) . This means that during the supply chain formation process, the agent sends a messages to its neighbors regarding the states of his variables that is maximizing its utility function.

Each agent interacts with its neighbors agents such that the utility of an individual agent U(v), is dependent on its own state and the states of these other agents.

Solving the problem stated above provides means for finding an allocation that maximize each agent utility in the supply chain within the underlying partners' constraints.

An allocation is a sub-graph $(X',E') \subseteq (X, E)$. For $X_i, X_j \in V'$, an edge between X_i, X_j means that agent X_j provides goods to agent X_i. An agent is in an allocation graph if it acquires or provides goods.

In order to find the best possible assignment in the supply chain network, this paper proposes 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.

A cluster graph is a data structure that provides a graphical flow chart of the process of manipulating factors (KaskK et. al. 2005).

For the simplicity of illustration of all the exchanged messages the present section considers a simple example as the one in Figure 2.

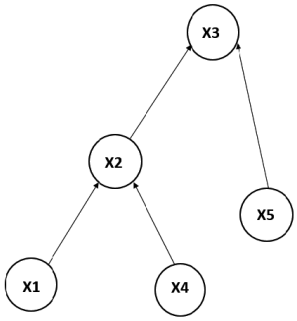


Fig. 2 Simple supply chain graph example

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. The node X3 represents the end consumer and it requires a complex good/service. As the required product is a complex one it might not be provide by a single supplier and it might require some subcontractors in the underlying levels of the supply chain. X3 can buy the good/service from X5 or might buy it from X2. But X2 cannot provide the required good as it might be a complex one and he needs to subcontract a part or a subassembly either from X1 or X4.

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. The present example will consider five contract parameters: A,B,C,D,E,F. For the simplicity of the illustration of all the exchanged messages it will be considered that each utility function of the participants depends on two contract

parameters but it can be generalized to multiple contract parameters. Let be the utility functions in Table 1 attributed to each participant:

Table 1 Utility functions of the participants and corresponding factors

X1: U(A,B)=θ_1
X2: U(B,C)=θ_2
X3: U(C,D)=θ_3
X4: U(E,B)=θ_4
X5: U(D,F)=θ_5

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 3.

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.

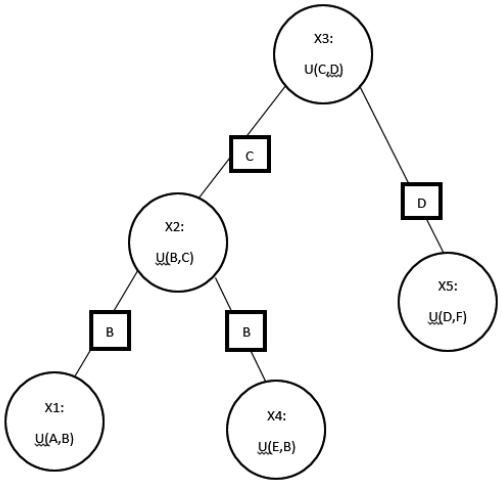


Figure 3 Initial graph transformed to cluster graph

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 (3), which is his preferred value from the set of values for issue B.

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.

The following figures illustrate a full message exchange over cluster graph in Figure 3.

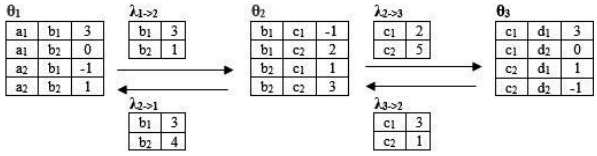


Figure 4 X1-X2-X3 Message exchange

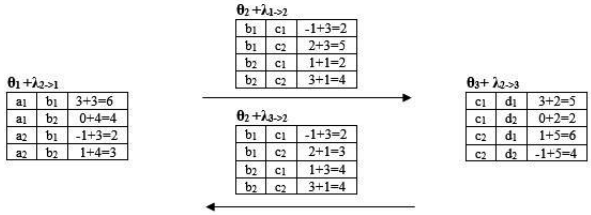


Figure 5 Computation performed during the X1-X2-X3 message exchange

According to the messages exchanged and the computations in Figure 4 and 5, the allocation X1-X2-X3 is unfeasible. The best possible completion for X1 is (a_1,b_1) with the score of 6, (c_2,d_1) for X3 with the score of 6 and (b_2,c_1) or (b_2,c_2) as they have the same score of 4 for X2. The cluster node X2 agrees with the cluster node X3 on the state of the variable they share (C) and that is c_2, but X2 cannot agree with X1 on the state of variable B, as X2 prefers the state b_2 and X1 prefers the state b_1.

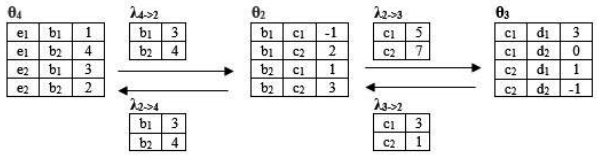


Figure 6 X4-X2-X3 Message exchange

The message exchange between X4-X2-X3 provide a feasible solution because all of the nodes agree on the states of the variables they share. The best possible completion for X4 is (e_1,b_2) with the score of 8, for X2 (b_2,c_2) with the score of 4, and for X3 (c_2,d_1) also with the score of 8.

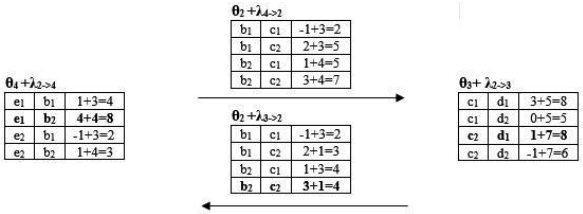


Figure 7 Computation performed during X4-X2-X3 message exchange

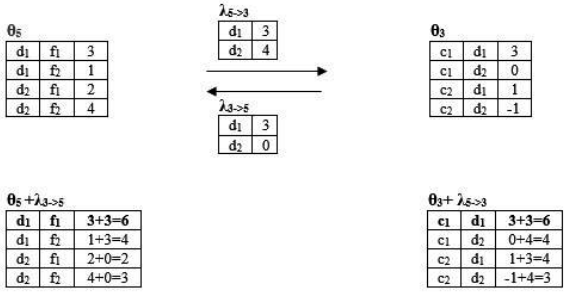


Figure 8 Message exchanged and computations performed for possible allocation X5-X3

Another feasible allocation is X5-X3 because the two cluster nodes also agree on the states of the variable they share. The best possible completion for X5 is (d_1, f_1) with the score of 6 and for X3 is (c_1, d_1) also with the score of 6, meaning that they agree on state d1 of the variable D, that they share.

Analyzing the two feasible solutions obtained in the above example, it can be stated that the optimal allocation in the supply chain is X5-X3 as the values (c1,d1) provide the highest utility of the end consumer X3, with a score equal to 3, according to his utility function.

SIMULATION FOR SUPPLY CHAIN FORMATION AND EVALUATION

In order to implement the mechanism proposed in the section above, PeerSim simulator was selected because of its performance regarding scalability and because it is based on components that allows prototyping a new protocol, combining different pluggable building blocks. PeerSim is a single-threaded peer-to-peer simulator beeing developed in a modular and scalable way (Montresor and Jelasity 2009).

The simulation process starts by reading the configuration file, given as an input parameter that defines the protocols to experiment. Then, both nodes and protocols are created and initialized. After the initialization phase, by default, every instance of the protocols running on each node is executed once per simulation cycle. The current implementation uses a transport protocol for message exchange that provides the communication primitives and allows adding more realism to the simulation.

The mechanism proposed in section above is implemented using existing building blocks in PeerSim, the nodes beeing scheduled according to a list of pending messages. PeerSim models the set of nodes as a collection through the class peersim.core.Network. The peers in the proposed model denote the set of supply chain formation participants represented by agents. The overlay network is represented by a adjacency list where every peer n of the network is connected to a set of neighbors *N(n)*. Furthermore, each node $v \in V$ has a utility functions and a list of numeric attributes that represents the contract parameters.

The utility functions are calculated by means of a weighted sum, the weights measuring the importance of a given issue for a certain participant in the chain. The list of values for contract parameters is periodically sent to its neighbors via message exchange mechanism. This process of message exchange is run in as many cycles as configured in the configuration file which is given as a parameter of the simulation.

The simulation was run having the following values in the configuration file:

```
# number of network nodes
SIZE 18
# number of simulation cycles
CYCLES 25
network.size SIZE
# Initializers to use
include.init init netInit
# Controls to use
include.control neighObs valueObs
#Supply chain protocol
protocol.supplychain SupplyChainFormation
# Network generation
init.netInit WireRegRootedTree
# The linkable protocol to operate on
init.netInit.protocol lnk
# Number of outgoing edges to generate from each
node
init.netInit.k 3
```

The simulator sets up the network nodes, and the protocols in them. The supply chain environment was generated using WireRegRootedTree class, which uses the Linkable protocol of PeerSim and adds connections that define a regular rooted tree. The simulations that have been runned used multiple supply chain network architectures with two up to five tiers.

At the beginning of the simulation all the agents in the network are being initialized with random preferred values for the states variables and also for the weights used at computing every agent utility. Each node *v* has a vector of numeric values that are the preferred states of the negotiated issues. Each node that is not a leaf is receiving messages from its neighbors, composing new messages and sending them upward to its neighbors.

Each node will assess messages received from the corresponding neighbor according to his own utility function.

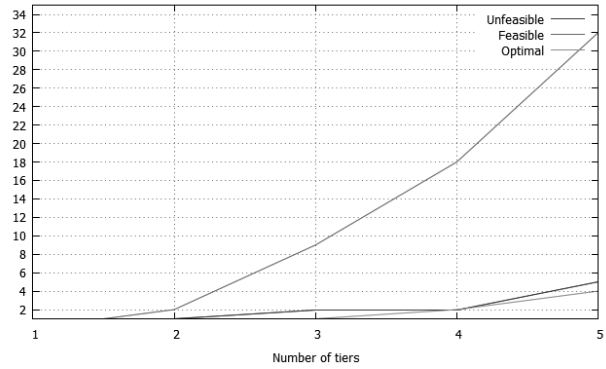


Figure 9 Dynamic of allocations relative to the number of tiers in the network

Analyzing the results of the simulation in the Figure 9 it can be observed that as the complexity of the network increases, the number of the feasible solutions grows fast, meanwhile the number of optimal and unfeasible solutions has a slower growth.

CONCLUSIONS AND FUTURE WORK

The present paper proposes an automated mechanism for supply chain formation. As opposed to the previous decentralized approaches, the current approach translates the SCF optimization problem not as a profit maximization problem but as a means for maximizing a utility function. Hence, it incorporates multiple contract parameters and enforces the propagation of the negotiated values from the underlying suppliers to the upper levels in the supply chain. The current work replaces the process of bidding in auctions with message passing between agents allowing participants to share their beliefs about the optimal structure of the supply chain only with relevant participants and thus preserving the self-interest of participating agents. Also this approach operates in a decentralized and distributed manner, with participants acting only on the basis of local information. As a future work we consider doing simulation on various network topologies in different economic hypothesis and also increasing the complexity of the proposed mechanism by adding constraints regarding complementary inputs needed for consumer nodes.

REFERENCES

Bishop, C. M., et al.,Pattern recognition and machine learning, Springer New York, New York, (2006)

Buyya, R., Abramson, D., Giddy, J.. An Economy Driven Resource Management Architecture for Global Computational Power Grids. Proceeding of International Conference on Parallel and Distributed Processing Techniques and Applications, (2000)

Cerquides J., Endriss U., Giovannucci A., and Rodriguez-Aguilar J. A., Bidding languages and winner determination for mixed multi-unit combinatorial auctions, IJCAI, Morgan Kaufmann Publishers Inc., pp. 12211226, (2003)

Cerquides, J., Endriss, U., Giovannucci, A., and Rodrıguez-Aguilar, J. A. 2007. “Bidding languages and winner determination for mixed multi-unit combinatorial auctions”. In IJCAI, pages 1221–1226. Morgan Kaufmann Publishers Inc.

Collins, J., Ketter, W., Gini, M., and Mobasher, B. 2002. “A multi-agent negotiation testbed for contracting tasks with temporal and precedence constraints”. International Journal of Electronic Commerce, 7:35– 58.

Davis, R. and Smith, R. G. 1983.”Negotiation as a metaphor for distributed problem solving”. Artificial intelligence, 20(1):63–109.

Giovannucci, A., Vinyals, M., Rodriguez-Aguilar, J. A., and Cerquides, J. 2008. “Computationally-efficient winner determination for mixed multi-unit combinatorial auctions”. In Proceedings of the 7th international joint conference on Autonomous agents and multi-agent systems Volume 2, pages 1071–1078. International Foundation for Autonomous Agents and Multiagent Systems.

Holweg, M. and Pil, F. The Second Century: Reconnecting Customer and Value Chain through Build-to-Order, Cambridge, MA and London, UK: The MIT Press, (2004)

KaskK., DechterR. LarrosaJ., DechterA “Unifying cluster-tree de compositions for reasoning in graphical models”. Artificial Intelligence, 166(1-2), (2005)

Lo G., Kersten G.E.. Negotiation in Electronic Commerce: Integrating Negotiation Support and Software Agent Technologies. Proceedings of 5th Annual Canadian Operational Research Society Conference, (1999)

Mikhaylov, B., Cerquides, J., and Rodriguez-Aguilar, J. A. 2012. “Solving sequential mixed auctions with integer programming”. In Advances in Artificial Intelligence, pages 42–53. Springer.

Montresor, A., Jelasity, M. PeerSim: A scalable P2P simulator, In Proc. of the 9th Int. Conference on Peer-to-Peer, pages 99-100, Seattle, WA, (2009)

Norman, T. J., Preece, A., Chalmers, S., Jennings, N. R., Luck, M., Dang, V. D., Nguyen, T. D., Deora, V., Shao, J., Gray, W. A., et al. 2004. “Agent-based formation of virtual organizations”. Knowledge based systems, 17(2):103–111.

Pearl, J. 1988. ”Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference”, 1st ed. San Francisco, CA, USA: Morgan Kaufmann, vol.1.

Penya-Alba, T., Vinyals, M., Cerquides, J., and Rodriguez-Aguilar, J. A.. A scalable Message-Passing Algorithm for Supply Chain Formation. 26th Conference on Artificial Intelligence (AAAI 2012), 2012

Walsh W. E. and Wellman M. P., Decentralized supply chain formation: A market protocol and competitive equilibrium analysis, Journal of Articial Intelligence Research (JAIR),vol. 19, pp. 513-567,(2003)

Walsh, W. E., Wellman, M. P., and Ygge, F., Combinatorial auctions for supply chain formation, Proceedings of the 2nd ACM conference on Electronic commerce,pp. 260-269, (2000)

Winsper M. and Chli M., Decentralized supply chain formation using max-sum loopy belief propagation, Comput. Intell.,vol. 29,pp 280-309, (2012)

Winsper M., Using min-sum loopy belief propagation for decentralised supply chain formation , PhD thesis, Aston University, (2012)

Winsper, M. and Chli, M., Decentralised supply chain formation: A belief propagation-based approach, Agent-Mediated Electronic Commerce,(2010)

Winsper, M. and Chli, M., Decentralized supply chain formation using max-sum loopy belief propagation, Computational Intelligence, vol.29(2), pp. 281-309, (2013)
Winsper, M. and Chli, M., Using the max-sum algorithm for supply chain formation in dynamic multi-unit environments, Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems, vol.3, pp. 1285-1286, (2012)

AUTHOR BIOGRAPHY

FLORINA L. COVACI is currently a teaching assistant at Business Information Systems Department within "Babes-Bolyai" University. She has a working experience of more than 10 years in .NET related technologies and her areas of competence include Project Management and IT Service Management. She is *Prince2 Practitioner* and *ITIL Foundation and Service Design* Certified. Her current research interests are multi-agent systems and inference in graphical models.

TOWARDS QUANTITATIVE RISK EVALUATION FOR SUPPLY CHAINS IN PREPARATION OF A SIMULATION STUDY

Birgit Mösl
Dietmar Neubacher
Nikolaus Furian
Siegfried Vössner
Graz University of Technology
Department of Engineering- and Business Informatics
Kopernikusgasse 24, 8010 Graz, Austria
email: birgit.moesl@tugraz.at

KEYWORDS
Supply Chain Risk Management, Conceptual Modeling, Decision Support System, Hierarchical Control Conceptual Modeling (HCCM), Discrete Event Simulation

ABSTRACT

Over the last years Supply Chains (SCs) are subject to major changes. They developed from classical intra-SCs to inter-SCs and further to flexible and more complex SC-networks. Because of globally distributed SC networks, dynamic environments and increasing customer expectations managers are faced with new challenges.
The increasing complexity of SCs results in a new dimension of risks (variable, uncertain, global), that has to be considered and managed. Hence, managers are required to assess these new risks and make decisions in a very complex and dynamic environment. Therefore, methods and tools are needed to support managers in evaluating scenarios and decision making. Modeling and simulation are often used for latter purposes as they are techniques to describe, shape and investigate problems. Especially they are used to improve the problem understanding, test and compare different scenarios and make what-if analyses without influencing the real world.
The focus of this paper is on the evaluation of risks along the SC, especially on the quantification of risks for further evaluation using simulation. Special attention has been given to risks related to countries. Appropriate risk indicators have been identified and discussed. The determined information provide a good basis for further simulation work.

INTRODUCTION

Today supply chains are more and more globally situated and therefore companies are confronted with the arising complexity, and also the associated risks. (Serdarasan 2013) Managers are required to make decisions in a dynamic and not-clear environment. To set up a new supply chain or to change a supply chain is not a simple and transparent task and therefore tools are needed to support the decision making process. Compared to domestic SC, global SC implicate other issues, as for instance currency risks, transport modes and times. (Serdarasan 2013)
An important step within a simulation study is the process of defining a conceptual model. One phase of conceptual modeling is about the definition of input factors. In this phase it has to be clarified which factors influence the considered system. In this paper the focus is on risks along a SC and therefore the utilization of conceptual and systematic methods to identify, evaluate and categorize risks and their influencing factors is explained. In order to be able to measure the risks, it is necessary to quantify the influence factors, which means it is necessary to find appropriate indicators. A guideline for this process is proposed as a first result of an ongoing research project.
This paper is structured in the following way: The next section is looking at the term 'SCM'. Further more the Supply Chain Operations Reference (SCOR) model is outlined briefly. The following section deals with the definition of risk and their occurrence along a SC. In the section 'Conceptual Model' the HCCM Framework is introduced. In the sub-sequential section the process of identifying risks along a specific SC is explained, as well as the issue of quantification is examined. The paper concludes with a short outlook.

SUPPLY CHAIN MANAGEMENT (SCM)

Particular attention has been paid to Supply Chain Management (SCM) in the last decades. Only recently there was no clear distinction between 'logistics' and 'SCM' for most experts in applied sciences and researchers. (Lambert and Cooper 2000)
In previous phases of the logistics development the company level was focused, with SCM this changed and the focus shifted to the level of a 'jointly operated network'. Consequently inter-organizational aspects become more crucial. (Beckmann 2012)

SCM includes more than the transport *of goods, services and information*, it includes also information technology, cash flows and tasks which are related to adjusting and aligning activities between the different SC partners. Thus, various factors influencing the system exist subject to a variety of boundary conditions. (Lummus et al. 2001)

To handle complicated and/or complex systems frameworks can be very helpful. The process reference model Supply Chain Operations Reference, short SCOR, is also referred to as a framework to endorse supply-chain activities and processes independently from the specific industry. SCOR defines a set of metrics and allows companies to evaluate their current business performance and subsequently to benchmark it with best-practices. (Stewart 1997, Beckmann 2012, Huan et al. 2004, Lambert et al. 2005, APICS Supply Chain Council 2016)

The six SCOR level-1-processes, are: Plan*, Source*, Make*, Deliver*, Return and Enable. The initial four processes are marked with (*). Figure 1 illustrates a SC based on the SCOR model. (Lambert et al. 2005, APICS Supply Chain Council 2016)

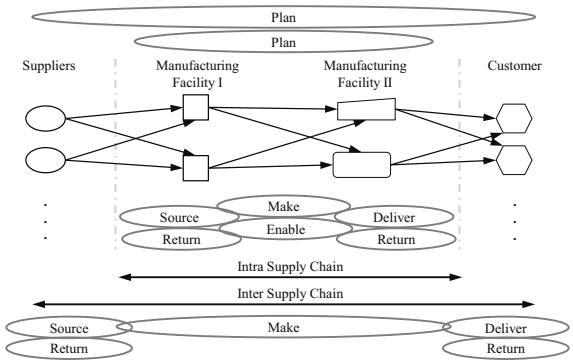


Fig. 1: SCOR model-based SC, based on Huan et al. (2004)

But to answer questions like the following some more aspects should be mentioned. Where should the production facility be located? Which supplier or sourcing strategy should be chosen? These and a lot of other questions are issues of SC management and demand answers, respectively decisions. (Thierry et al. 2010, Tako and Robinson 2012)

Decisions are always associated with a certain degree of uncertainty and also with risks. Identifying, measuring and evaluating risks are important tasks that give further insights and allow the management to make their decision based on facts and their judgment. Thereby it has to be mentioned, that it is not about risk avoidance, rather it is about risk awareness. The underlying risks in context with SC are explained in the following section.

RISKS IN SUPPLY CHAINS

Production has changed in the last decades and is becoming more and more complex. That means, in former times the production flow, from the raw material via the manufacturer to the end customer was rather simple. Today companies are faced with global competition and price pressure. Further, rapid developments in the area of information technologies, for example using Enterprise Resource Planning (ERP) systems, have changed access and availability of data. The simple production flow changed to more complex, longer and global SCs with a higher number of participants, influenced by two major trends: globalization and consolidation of firms, which results in a higher unpredictability for the market players. Hence, risks have also changed and new ones have to be added. Risks can be directly connected with the end product or by interruptions somewhere along the SC, for instance due to hurricanes, epidemics, terrorist attacks or other random events. (Tang and Nurmaya Musa 2011, Manuj and Mentzer 2008b, Harland et al. 2003)
Manuj and Mentzer (2008a) demonstrated that risk is understood in different ways in the literature, depending on the domain. But also the kind of industry influences the view of risk. Despite this variety Manuj and Mentzer (2008a) found out that there are three elements:

- potential losses
- likelihood of those losses
- significance of the consequences of the losses

The literature review in the study of Manuj and Mentzer (2008b) demonstrates that there is no suitable definition of risk in context of global SCs. Therefore, they developed a definition, which reads as follows:

> *. . . the distribution of performance outcomes of interest expressed in terms of losses, probability, speed of event, speed of losses, the time for detection of the events, and frequency.*

Which types of risks can arise in a global SC? How can risks be classified? These questions and more come up when dealing with the issue risk. Therefore, some classifications need to be made.
In general, risks can be of quantitative or qualitative nature. Quantitative risks are measurable, for example by stock levels, delivery times, delivery reliability. Qualitative risks are not measurable directly and for example refer to reliability, know-how and so on. (Manuj and Mentzer 2008a)
The following classification of risk along the entire SC can be found in the literature: supply risks, operational risks, demand risks and security risks. This four types of risks are directly connected to the SC and they influence supply and demand, as illustrated in figure 2. These risks can be completed by macroeconomic risks, policy

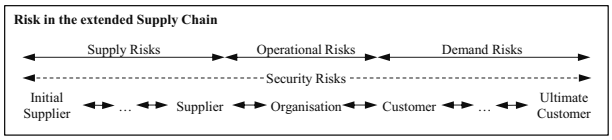


Fig. 2: Risk in the extended SC (Manuj and Mentzer 2008a)

risks, competitive risks and resource risks. (Manuj and Mentzer 2008a;b)

The assessment of risks is an important step in risk management and it means to assess the probabilities of various risks and corresponding weight of the consequences, upon occurrence. The probability depends on the one hand on how exposed the incident is, in other words how dangerous an event can be in terms of risks. And on the other hand it depends on the chance that the trigger will be activated. An activation can be triggered by individuals or by organizations, but also by things beyond control. It is also a question of the power of the company, in what extent the company is able to influence their environment, e.g. organizations. The weight of the consequences can be roughly calculated if there exist directives or laws and therefore, the consequences resulting from non-compliance are relatively well known. It is important to have in mind not only quantitative measurable consequences or losses, but also other intangible values ('soft'-facts) like reputation and image, also taking into account the influence of media and social media. (Harland et al. 2003, Tummala and Schoenherr 2011)

For managing risks along the SC Tummala and Schoenherr (2011) adapted the Risk Management Process (RMP) to the SC resulting in their suggested conceptual framework Supply Chain Risk Management Process (SCRMP). The framework consists of three phases, from risk identification to risk control & monitoring.

Phase one of the SCRMP contains 'risk identification', 'risk measurement' and 'risk assessment', which should be realized in the mentioned sequence. The task of 'risk identification' is to find out all risk factors influencing the SC and get a complete 'picture' of them. In the next step 'risk measurement' the consequences of the identified risks and their order of magnitude are set. Consequences are of different nature, for example time (e.g. delays), money (e.g. exceeded costs), performance (e.g. poor quality). A widely known classification of consequences is made by Crockford (1986) and is shown in table 1.

The before explained matters are the basis for the next step, focused in this paper: determining input factors for the conceptual model. Risks along a SC are considered, especially their identification, quantification, evaluation and their role as input factors. For that reason, the next section is about conceptual modeling.

Tab. 1: Risk classification - consequences (Crockford 1986)

Consequence	Frequency	Severity	Predictability
trivial	very high	very low	very high
small	high	low	reasonable, with infrequent occurrence
medium	low	medium	reasonable, with frequent occurrence
large	very low	high	minimal

CONCEPTUAL MODELING

The procedure to abstract the real system to a conceptual model, which can be translated into a computer model plays a crucial role in every simulation study. The challenge is to model the real world in a sufficient way, but as simple as possible. Therefore, it has to be decided what has to be included in the model and what excluded. The result of this process is a conceptual model.
HCCM is a framework for conceptual modeling, suggested by Furian et al. (2015). It can be seen from figure 3 that after understanding the problem situation the next step is 'identification of modeling and general objectives', which is a part of building a conceptual model. The modeling objectives refer to the specific purpose of the simulation study, in other words the company's aim. The general objectives include the specification of the simulation tool, for example run-time or flexibility to changes. Phase three contains 'defining input factors' and 'defining output responses'. Following Furian et al. (2015) and Robinson (2011) input factors are experimental factors, which are altered with respect to various experiments and simulation runs with the aim to meet the modeling objectives. Outputs are the results of the experiments and they are used to assess if the modeling objectives are met or not. The next phase 'model content' consists of model structure, model individual behavior and model control. For more detailed information see Furian et al. (2015).

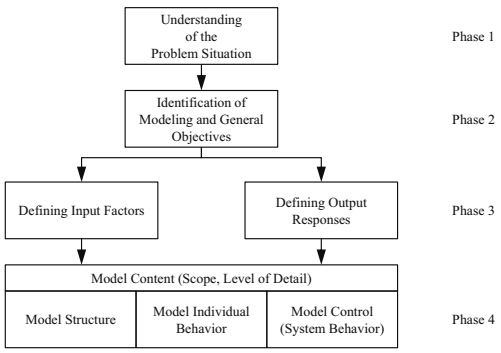


Fig. 3: Structure of the HCCM framework (Furian et al. 2015; p.ȧ9)

Simulation is a modeling technique, which is used to

model systems of interest that are usually too complex to be studied by analytic solutions. (Law 2015, Hillier and Lieberman 2005) Simulation can be a useful tool for a great variety of fields, as for example 'Designing and operating transportation systems such as airports, freeways, ports and subways' or 'Analyzing supply chains', which is focused in this paper. (Law 2015)
In order to be able to simulate a certain scenario the conceptual model has to be defined as mentioned previously. Therefore the process of finding input factors and quantifying them is explained in the next section based on a case study.

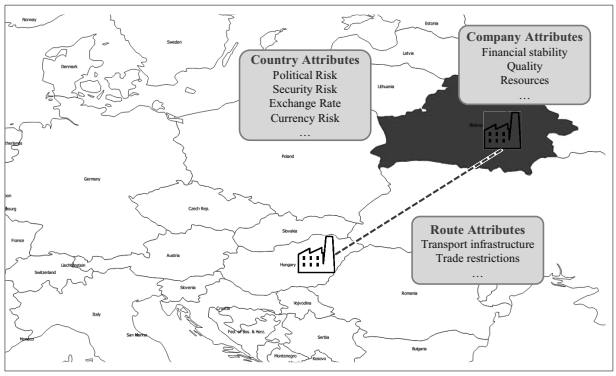


Fig. 4: Dimensions

CONCEPTUAL SC MODEL

Within a project in cooperation with a middle-sized company, operating world-wide and active in the field of micro-controller-based control units and operating elements for ergonomic solutions the question of risks along a SC was investigated.
When analyzing the system, it seems that there are many different factors influencing the supply chain. Especially impacts with respect to the countries arose as an important factor. Therefore, the dimension 'country' and its connection to related factors has been investigated more closely.
Subsequently three dimensions could be identified: 'Country', 'Company' and 'Route/Link'. In figure 4 these dimensions and some of their attributes are illustrated. In this paper the term 'dimension' is for this purpose defined and understood as a summary for different entities and their related risks and opportunities. The dimension 'Country' describes all inputs, outputs and influences directly related to the country level, for instance political issues, governmental regulations, currency or trade restrictions. These dimensions are assigned to levels, which are understood in this paper as a kind of a view level, in other words a high level refers to a macro level. 'Country' is positioned on the highest level. The next dimension 'Company' is situated on a more detailed level than 'Country' and involves all matters associated with the individual supplier, for example quality, technical performance, know-how, financial situation, flexibility or capacity. The third dimension 'Route/Link' can not be assigned to one of these two levels, as it depends on the kind of usage. 'Route/Link' covers issues that belong to the connection between two entities. Depending on the type of entity, the dimension is located on the different levels or between them, this means 'Route/Link' can connect two or more suppliers, two or more countries, but also suppliers and countries. It expresses all topics concerning the linkage, as for example the distance between the entities, transport mode, border crossings.
The input factors of the current system are on the one hand, various risks related to the before introduced dimensions. On the other hand, the design of the network and the mass flows (quantities of parts, assemblies or products) between the nodes of the network. The output factors are probability distributions of the various risks, network designs and mass flows that can be compared to each other.

Thus, the input factors have to be identified. For that, in a first step a criteria matrix was used, based on Vester (2015). In the investigated system, the by Vester (2015) suggested areas of life are described by the processes of the SC. Therefore, the criteria matrix was adjusted to the application and the variables are put in context to these processes. The variables are related to the risk triggers, which were listed in the first column of the criteria matrix. The risk triggers were based on the risk triggers of the SCRMP (see (Tummala and Schoenherr 2011)). They were completed by elements of a supplier evaluation. In the first row the processes of the SCOR model were listed. With this matrix the importance of the various risks to the individual process steps was evaluated by choosing one of three possible values: '0' for 'not relevant', '0.5' for 'partly relevant' and '1' for 'fully relevant'. This matrix should ensure completeness and relevance of the risks and therefore it was accomplished for all processes.

Next, to investigate the influence of the various risks on each other an impact matrix was prepared, following Vester (2015). This matrix gives information on the inter-dependencies of the variables. The impact of variable A on variable B, C, etc. was evaluated by the following scale: 0 - no dependency, 1 - very small dependency (big change of A results only in a little change of B), 2 - medium dependency (proportional change), 3 - strong dependency (little change of A results in a big change of B). Variables with a big impact on others have a high active sum (AS), variables strongly influenced by others have a big passive sum (PS). Therefore, the variables can be distinguished in active elements, critical elements, reactive elements, buffering elements and neutral elements.

The gained insights from the criteria and impact matrix has been discussed with experts and reevaluated. Some

risks were assessed as not important or not applicable, some others were added. An emerging issue during the discussions was 'cultural matching'. This also shows that with global sourcing totally new issues arise, which can result in unexpected difficulties. For example problems due to language differences, or due to distinctions in code of behavior, can be a big issue. The individual risk triggers are assigned to different risk categories. The result of the discussion, i.e. if the variable should be included or excluded in the further risk evaluation, is documented. This procedure has been suggested also by Robinson (2008).
An adapted list with risks was prepared. The various risks were grouped according to different categories and also assigned to the different mentioned dimensions.
The next step was the question of 'risk measurement', that means how can the before identified risks be assessed, as objectively as possible, using indicators. At this point some questions arises: Which indicators are available? Which indicator can represent which risk factor or risk type? Besides, more technically questions turned up as: In which format is the data available? How should the data be stored and processed?
For finding appropriate indicators various sources are available. Some companies provide aggregated and edited data for various issues, but they are almost only pay-for-services. For example, ControlRisks (2016) provides information on security risks, political risks, a risk map and so on. Other sources provide their data for free, certainly the data are not as well-structured as provided by pay-for-services. The focus was on free available data in this case.
Indeed some of the found indicators can be used for various risk triggers. Obviously risk triggers regarding the dimension 'Route' often depend on the distance between the individual SC members.
It has to be mentioned, that available data change and improve and thus, it is suggested to repeat the research on indicators within a reasonable time. However, it is important to consider the timeliness of the data. Especially for some risk triggers this can be essential, for instance the regional stability.
Now the individual risks were considered in the context of a SC setting. From purchasing the individual parts on different markets, to the delivery to the end customer, there are a lot of steps in between and consequently many potential risks arise. Risks appear not only linear, but rather overlap.
Based on the previously defined quantitative indicators risks belonging to the distinctive dimensions can be described and measured. Next, these indicators have to be scaled, based on an objective scaling matrix. Furthermore, the indicators can be weighted using a weighing matrix. Afterwards the product streams can be taken into account by estimated quantities and results obtained by a value stream analysis, for example. Also constraints can be considered, as for example a certain part has to be made by supplier C in country X. All these assumptions and simplifications result in a probability distribution of several risks included in the setting.
A simplified example is shown in figure 5. In the illustration the suppliers and manufacturers are shown as squares. The ellipses represent the countries, where the parts are purchased. The arrows stand for the connections between the SC members, the distance and the risk associated with the route. The thickness of the arrows symbolize the value stream, which results from the quantity and price per piece. The value stream is an suitable indicator, as it makes a difference if there is a huge amount of parts needed, but the price per piece is very low or there are only few, but very expensive ones, needed.

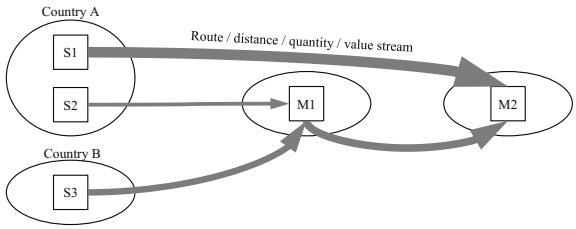


Fig. 5: Design Model

By varying the settings of the considered SC, for instance by using another supplier or purchasing from a different country, distinctive results are gained and can be compared to each other. The outcomes can serve as a basis for discussions and at the end, decisions.

OUTLOOK

The before examined design model needs to be translated into a computer model, which can be used for simulations in the end.
In preparation of a computer model, an important step is to gather the data of the risk indicators. It is supposed to use a database to store the data from the different sources. As there is no standardized interface and each data set looks different, this is not a simple task. Data administration is an important issue in this case. Using a data base for the adminstration of the risk indicators is also suggested by Deleris and Erhun (2005), who points out also the importance of currency of the data and therefore, the need of regular updates.
As a next step in order to be able to simulate various scenarios the model needs to be detailed in depth and according to Furian et al. (2015) the model content has to be specified. Subsequently the translation into a computer model is required for the simulation runs and the evaluation. No final decision has yet been taken on the simulation software. In the moment the following possibilities are in focus: the open source library HCDESLib[1]

[1]available at https://github.com/nikolausfurian/HCDESLib

and the commercial software AnyLogic. The results of the simulation runs will support the process of decision making, as the gained results for the different scenarios can be evaluated and compared to each other. Therefore this will be a helpful tool in a complex environment to gain more insights.
Within the before mentioned project the next steps should be realized. A simulation prototype should be generated to be able to compare different scenarios and demonstrate the benefit of using modeling and simulation. Furthermore it is planned to deduce a framework.

CONCLUSION

The process of conceptual modeling is a very important one and determines the thereafter written computer model. As a result it is essential to use a systematic and conceptual method, which allows flexibility to the extent to adjust it to your requirements and offers transparency to all partners in the process of conceptual modeling in order to improve the common system understanding and subsequently to improve the quality of the model. The proposed guideline is independently on the one hand, from the specific industry and on the other hand, also from the computer model used for the simulation. It can be applied to a varying complexity of SC and also to different levels of detail, which are represented by dimensions and levels.

REFERENCES

APICS Supply Chain Council, 2016. *APICS*. URL `http://www.apics.org/pe-home`.

Beckmann H., 2012. *Prozessorientiertes Supply Chain Engineering: Strategien, Konzepte und Methoden zur modellbasierten Gestaltung*. Springer Gabler, Wiesbaden.

ControlRisks, 2016. URL `https://www.controlrisks.com/`.

Crockford N., 1986. *An introduction to risk management*. Woodhead-Faulkner, Cambridge.

Deleris L. and Erhun F., 2005. *Risk management in supply networks using Monte-Carlo simulation. Proceedings of the Winter Simulation Conference, 2005*, 1643–1649.

Furian N.; O'Sullivan M.; Walker C.; Vössner S.; and Neubacher D., 2015. *A conceptual modeling framework for discrete event simulation using hierarchical control structures. Simulation Modelling Practice and Theory*, 56, no. 0, 82–96.

Harland C.; Brenchley R.; and Walker H., 2003. *Risk in supply networks. Journal of Purchasing and Supply Management*, 9, no. 2, 51–62.

Hillier F.S. and Lieberman G.J., 2005. *Introduction to operations research*. McGraw-Hill, New York, NY, 8. ed. ed.

Huan S.H.; Sheoran S.K.; and Wang G., 2004. *A review and analysis of supply chain operations reference (SCOR) model. Supply Chain Management: An International Journal*, 9, no. 1, 23–29.

Lambert D.M. and Cooper M.C., 2000. *Issues in Supply Chain Management. Industrial Marketing Management*, 29, no. 1, 65–83.

Lambert D.M.; García-Dastugue S.J.; and Croxton K.L., 2005. *AN EVALUATION OF PROCESS-ORIENTED SUPPLY CHAIN MANAGEMENT FRAMEWORKS. Journal of Business Logistics*, 26, no. 1, 25–51.

Law A.M., 2015. *Simulation modeling and analysis*. McGraw-Hill Education, New York, NY, 5. ed., in ed.

Lummus R.R.; Krumwiede D.W.; and Vokurka R.J., 2001. *The relationship of logistics to supply chain management: developing a common industry definition. Industrial Management & Data Systems*, 101, no. 8, 426–432.

Manuj I. and Mentzer J.T., 2008a. *GLOBAL SUPPLY CHAIN RISK MANAGEMENT. Journal of Business Logistics*, 29, no. 1, 133–155. ISSN 07353766.

Manuj I. and Mentzer J.T., 2008b. *Global supply chain risk management strategies. International Journal of Physical Distribution & Logistics Management*, 38, no. 3, 192–223.

Robinson S., 2008. *Conceptual modelling for simulation Part II: a framework for conceptual modelling. Journal of the Operational Research Society*, 59, 291–304.

Robinson S., 2011. *Choosing the Right Model: Conceptual Modeling for Simulation. Proceedings of the 2011 Winter Simulation Conference (WSC)*, 1423–1435.

Serdarasan S., 2013. *A Review of Supply Chain Complexity Drivers. Comput Ind Eng*, 66, no. 3, 533–540.

Stewart G., 1997. *Supply-chain operations reference model (SCOR): the first cross-industry framework for integrated supply-chain management. Logistics Information Management*, 10, no. 2, 62–67.

Tako A.A. and Robinson S., 2012. *The application of discrete event simulation and system dynamics in the logistics and supply chain context. Decision Support Systems*, 52, no. 4, 802–815.

Tang O. and Nurmaya Musa S., 2011. *Identifying risk issues and research advancements in supply chain risk management. International Journal of Production Economics*, 133, no. 1, 25–34.

Thierry C.; Bel G.; and Thomas A., 2010. *The Role of Modeling and Simulation in Supply Chain Management 1. SCS M&S Magazine*, 4, no. October, 1–8.

Tummala R. and Schoenherr T., 2011. *Assessing and managing risks using the Supply Chain Risk Management Process (SCRMP). Supply Chain Management: An International Journal*, 16, no. 6, 474–483.

Vester F., 2015. *Die Kunst vernetzt zu denken: Ideen und Werkzeuge für einen neuen Umgang mit Komplexität ; ein Bericht an den Club of Rome ; [der neue Bericht an den Club of Rome]*. No. 33077 in dtv Wissen. Dt. Taschenbuch-Verl, München, 10. auflag ed.

INVENTORY MANAGEMENT OPTIMIZATION

THE CONCEPT OF SEMI-VARIANCE AS A TOOL FOR SAFETY INVENTORY DECISIONS IN CASE OF UNCERTAIN DEMAND

Katrien Ramaekers[1]
Galina Merkuryeva[2]
Gerrit K. Janssens[1]

[1]Research Group Logistics, Faculty of Business Economics
Universiteit Hasselt – campus Diepenbeek
Agoralaan 1, B-3590 Diepenbeek, Belgium

[2]Department of Modelling & Simulation
Riga Technical University
Kalku street 1, LV-1658 Riga, Latvia
e-mail: {katrien.ramaekers, gerrit.janssens}@uhasselt.be, galina.merkurjeva@rtu.lv

KEYWORDS
Inventory management, safety inventory, semi-variance

ABSTRACT

An inventory system containing uncertainty, e.g. in demand or in lead-time requires to determine a safety inventory for re-ordering. The decision models need a probability distribution of the demand during lead-time. In the literature on inventory control, mostly a Normal distribution for describing this demand is assumed. Based on the knowledge of the variance and on the distribution assumption, the safety inventory is calculated, given a prescribed customer service level. However, the functional form of the probability distribution in practice might look different from the shape of a Normal distribution and by this wrong decisions are made which result in high costs or low service level. It is investigated here whether the use of semi-variance is more robust to a deviation in the shape from the Normal distribution. The determination of the safety inventory is worked out for a triangular distribution as an illustrative example.

INTRODUCTION

Operational inventory management deals with decisions on how much to order of a certain product and when to order. The former decision relates to a quantity, which is either a fixed quantity or a quantity based on the current inventory level (for example, to fill up the inventory to a certain level). The latter decision is based on the policy order based on either a time basis (for example, at the end of each week) or based on the inventory level (for example, if the current inventory level drops below a certain threshold, an order is placed).
The time between ordering and delivery of goods is called the lead-time. The lead-time may be fixed or may be uncertain (stochastic). Also the demand during lead-time (DDLT) mostly is stochastic. To avoid getting out-of-stock, companies hold extra inventory, which means more inventory than the expected value of demand during the lead-time. This extra inventory is called buffer inventory or safety inventory.
The level of safety inventory depends on the risk that the company is willing to take of an out-of-stock event. If the inventory on hand is not sufficient two situations may be observed: the demand results in lost sales or the demand results in backorders. The company decides on the level of safety inventory based on the probability of an out-of-stock event or on the expected number of units short while being out-of-stock. For this decision, the probability distribution of the DDLT needs to be known. Most textbooks and software assume that the DDLT follows a Normal distribution. An estimate of the mean and variance allows for the determination of the safety inventory, given the risk that the company would like to accept. But in reality the DDLT does not always have the characteristics of a normal distribution: the distribution is not always unimodal and it is not always symmetric. Several studies have shown that the shape of the demand distribution during lead time has an important influence on this decision and might lead to either wrong decisions in terms of service levels or will lead to higher costs than expected.

LITERATURE REVIEW

In the case of unknown but observable demand process, a commonly used approach is to employ replenishment formulas that are derived assuming a completely specified demand distribution, and to substitute statistical estimates for the demand distribution parameters. Limited historical data can be used to estimate the parameters of the demand distribution (Jacobs and Wagner, 1989). Their findings show that when demand variability is large, exponentially smoothed estimators can substantially outperform sample means and sample variances. When demand variability is relatively small, the cost of demand uncertainty is negligible, and the choice of statistical estimators is not critical.

Unfortunately, the demand distribution is rarely known and the lead-times are frequently random. Inventory managers are fortunate if they know the first two moments of these random variables. To address these issues, Ehrhardt

(1979) proposes the Power Approximation (PA) method. If the functional form of the DDLT is not (fully) known, the common assumption of making use of the normal distribution might be very harmful. Naddor (1978) finds that false assumptions about the distribution may lead to higher cost in the case of extreme distributions, but that, with realistic distributions, only the first and second moments are essential. On the other side, Bartezzaghi et al. (1999) show a significant impact of the shape of the demand distribution on the service level, based on a large set of experiments. Their analysis shows that the shape of the distribution is a primary factor in the determination of inventories. Also Lau and Zaki (1982) note that mean and variance are not sufficient for safety stock calculation, but also skewness and kurtosis should be accounted for. Furthermore, Käki et al. (2013) show the impact of the demand distribution shape on replenishment, based on experiments with qualitative shape characteristics (normal, positively skewed, negatively skewed, and bimodal).

Sometimes, it is even an unreasonable assumption that the demand obeys a known distribution. In such a case, for 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), and in Ramaekers and Janssens (2012), 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).

PROBLEM FORMULATION

The intention of this study is to investigate whether an alternative of estimating the variability of demand during lead-time would lead to less risk, compared to the use of the normal distribution, regarding errors in service level or in cost in case the DDLT follows an asymmetric or a multimodal distribution. The core idea is to use a different measure of variability, called the 'semi-variance' (to be explained in detail further on) as a basis for determining the safety inventory.

When mentioning about risk, one might compare the risk in inventory risks with those in investment in the financial world. 'Variance' is called a full domain risk measure while they can be contrasted with partial domain measures, which provide information for some distribution over some part of its domain. In the financial domain special relevance is given to the 'downside risk', focusing on return falling below some critical level (Grootveld and Hallerbach, 1999). In inventory management the interest goes into the 'upside risk', focusing on what happens above the critical level which is the inventory level. Just like the financial world focuses on downside risk measures based on mean-lower partial moments, our idea is to focus on mean-higher partial moments.

One approach to this 'upside risk' specifies the risk in terms of probability-weighted of deviations above a target. One example is the semi-variance, which has been introduced to Markowitz (1959, chapter 9) and which measures the variability below the mean in the financial case but above the mean in the inventory management case. The semi-variance is a special case of the more general lower partial moments (see Harlow and Rao (1989).

The Concept of Semi-Variance

The DDLT is expressed as a random variable *X* which follows a probability distribution with density *f(x)*. The cumulative distribution of *X* is written as *F(x) = Prob{ X <= x}*. The expected value of the DDLT is written as:

$$\mu_X = \int_{-\infty}^{+\infty} x \, dF(x) \qquad (1)$$

and its variance as:

$$\sigma_X^2 = \int_{-\infty}^{+\infty} (x - \mu_X)^2 dF(x) \qquad (2)$$

The variance may be split up in two parts defined as:

$$\sigma_X^2 = \sigma_X^{2-} + \sigma_X^{2+} \qquad (3)$$

with

$$\sigma_X^{2-} = \int_{-\infty}^{\mu_X} (x - \mu_X)^2 dF(x) \qquad (4)$$

and

$$\sigma_X^{2+} = \int_{\mu_X}^{+\infty} (x - \mu_X)^2 dF(x) \qquad (5)$$

The formula (5) is called the *(positive) semi-variance* which also might be rewritten as

$$\sigma_X^{2+} = E[max(0, X - \mu_X)^2] \qquad (6)$$

Formula (6) mostly will be estimated by means of the sample semi-variance, defined as

$$s_X^{2+} = \sum_{i=1}^{n} \frac{\left(max(x_i - \bar{x}, 0)\right)^2}{n} \qquad (7)$$

where

$$\bar{x} = \sum_{i=1}^{n} \frac{x_i}{n} \qquad (8)$$

The mean is estimated in the classical sense as the average value of the demand during a number of time periods. Once the mean is estimated the semi-variance of the DDLT can be estimated.

The reorder point *R* can be expressed as

$$R = \mu_X + k \, \sigma_X^+ \qquad (9)$$

with $k > 0$ a safety factor. Note that $\sigma_X^+ \leq \sigma_X$. In case $\sigma_X^+ = \sigma_X$, it means $\sigma_X = 0$. This means that the safety factor k should be larger than in the case the variance is used. For the Normal distribution (and in fact also for any symmetric distribution), $\sigma_X^+ = \sigma_X^- = \frac{\sigma_X}{\sqrt{2}}$. In this study it is assumed that the right-hand side of the distribution corresponds to the tail of a normal distribution. Therefore the value of *k* in formula (9) can be determined by $k = k_{Normal} * \sqrt{2}$, where k_{Normal} is the safety factor as obtained from the normal distribution. In practice the reorder point *R* is calculated as:

$$R = \bar{x} + k_{Normal} * \sqrt{2} * s_X^+ \qquad (10)$$

SOLUTION METHOD

An illustration is required to illustrate the reason why this change in risk measure is useful to investigate. The best way is to illustrate by means of a distribution from which the quantiles can be obtained analytically, for example the triangular distribution.

Let a variable Z be distributed as a triangular distribution with support on $[a,b]$ with mode m. Its density can be written as:

$$f(z) = \frac{2(z-a)}{(b-a)(m-a)} \; if \; a \le z \le m$$
$$= \frac{2(b-z)}{(b-a)(b-m)} \; if \; m < z \le b$$
$$= 0 \; otherwise \qquad (11)$$

Let the mean of a triangular distribution be written as:

$$\mu = \frac{a+m+b}{3} \qquad (12)$$

The right-side semi-variance is written as:

$$\sigma_X^{2+} = \int_{\mu}^{b} (x-\mu)^2 f(x)dx \qquad (13)$$

Two cases need to be considered (c = 1,2). The case c=1 relates to the condition $\mu < m$, while the case c=2 relates to the condition $\mu \ge m$. Let this integral be written as the sum of three parts I_{1c+}, I_{2c+} and I_{3c+}, where

$$I_{1c+} = \int_{\mu}^{b} x^2 f(x)dx \qquad (14)$$

$$I_{2c+} = -2\mu \int_{\mu}^{b} x f(x)dx \qquad (15)$$

$$I_{3c+} = \mu^2 \int_{\mu}^{b} f(x)dx \qquad (16)$$

First consider the case where $\mu < m$:

$$I_{11+} = \frac{2}{(b-a)(m-a)} \int_{\mu}^{m} (x-a)x^2 dx + \frac{2}{(b-a)(b-m)} \int_{m}^{b} (b-x)x^2 dx$$

which after some calculations leads to

$$I_{11+} = \frac{2}{(b-a)(m-a)} \left[\frac{m^4-\mu^4}{4} - a\left(\frac{m^3-\mu^3}{3}\right) \right] + \frac{2}{(b-a)(b-m)} \left[b\frac{(b^3-m^3)}{3} - \left(\frac{b^4-m^4}{4}\right) \right] \qquad (17)$$

$$I_{21+} = \frac{-4\mu}{(b-a)(m-a)} \int_{\mu}^{m} (x-a)xdx + \frac{-4\mu}{(b-a)(b-m)} \int_{m}^{b} (b-x)xdx$$

which after some calculations leads to

$$I_{21+} = \frac{-4\mu}{(b-a)(m-a)} \left[\frac{m^3-\mu^3}{3} - a\left(\frac{m^2-\mu^2}{2}\right) \right] + \frac{-4\mu}{(b-a)(b-m)} \left[b\frac{(b^2-m^2)}{2} - \left(\frac{b^3-m^3}{3}\right) \right] \qquad (18)$$

$$I_{31+} = \frac{2\mu^2}{(b-a)(m-a)} \int_{\mu}^{m} (x-a)dx + \frac{2\mu^2}{(b-a)(b-m)} \int_{m}^{b} (b-x)dx$$

which after some calculations leads to

$$I_{31+} = \frac{2\mu^2}{(b-a)(m-a)} \left[\frac{(m-a)^2}{2} - \frac{(\mu-a)^2}{2} \right] - \frac{2\mu^2}{(b-a)(b-m)} \frac{(b-m)^2}{2} \qquad (19)$$

Second, consider the case where $\mu \ge m$:

$$I_{12+} = \frac{2}{(b-a)(b-m)} \int_{\mu}^{b} (b-x)x^2 dx$$

which after some calculations leads to

$$I_{12+} = \frac{2}{(b-a)(b-m)} \left[b\left(\frac{b^3-\mu^3}{4}\right) - \left(\frac{b^4-\mu^4}{4}\right) \right] \qquad (20)$$

$$I_{22+} = \frac{-4\mu}{(b-a)(b-m)} \int_{\mu}^{b} (b-x)xdx$$

which after some calculations leads to

$$I_{22+} = \frac{-4\mu}{(b-a)(b-m)} \left[b\left(\frac{b^2-\mu^2}{2}\right) - \left(\frac{b^3-\mu^3}{3}\right) \right] \qquad (21)$$

$$I_{32+} = \frac{2\mu^2}{(b-a)(b-m)} \int_{\mu}^{b} (b-x)dx$$

which after some calculations leads to

$$I_{32+} = \frac{2\mu^2}{(b-a)(b-m)} \left[b(b-\mu) - \frac{b^2-\mu^2}{2} \right] \qquad (22)$$

Note that the negative semi-variance can be calculated in a similar way. But, as this part is not relevant for our purposes, this development is not included.

NUMERICAL ILLUSTRATION

In the numerical illustration the quality of both the safety stock based on variance and semi-variance is evaluated versus the exact value obtained from the triangular distribution. The triangular distribution has been chosen for this numerical example because the inverse cumulative distribution function can be expressed in an analytical way.

The inverse cumulative distribution function $F^{-1}(y)$ can be written as:

$$F^{-1}(y) = a + \sqrt{y(m-a)(b-a)} \quad if\ 0 < y < \frac{m-a}{b-a} \qquad (23)$$

$$F^{-1}(y) = b - \sqrt{(1-y)(b-m)(b-a)} \quad if\ \frac{m-a}{b-a} \leq y \leq 1 \qquad (24)$$

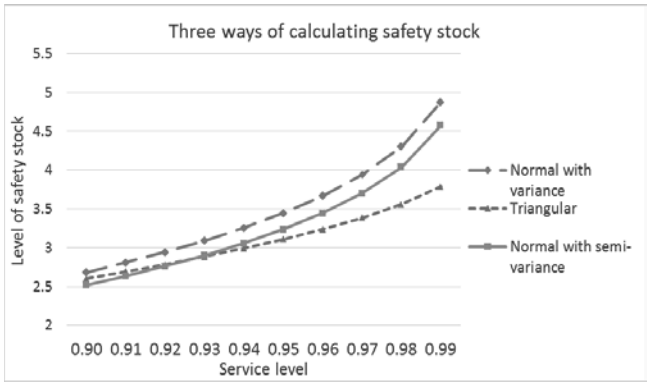


Figure 1: Level of safety stock in function of service level for the Triangular distribution (a=5, m=12, b=15)

A comparison of three ways of calculating the level of safety stock for a skewed distribution (μ < m) is shown in Figure 1. It can be seen that the approach with the semi-variance has a lower level of safety stock than the approach based on the variance and is closer to the exact curve of the triangular distribution. Figure 2 shows the same comparison for a skewed distribution (μ > m). The approach with the semi-variance has a higher level of safety stock than the approach based on the variance. The level is closer to the exact curve of the triangular distribution for nearly the whole spectrum of service levels. Only with the very high service levels the approach based on the variance is better. This phenomenon needs further investigation.

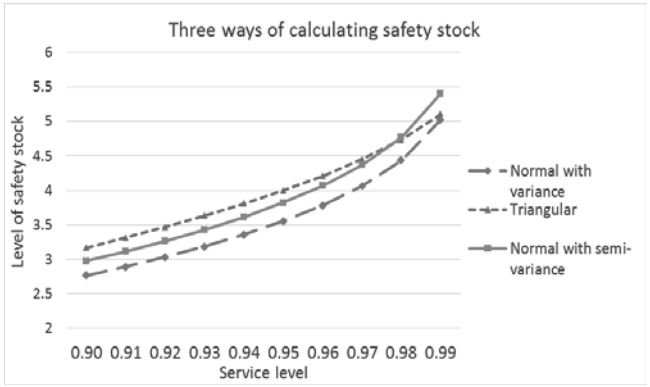


Figure 2: Level of safety stock in function of service level for the Triangular distribution (a=5, m=7, b=15)

Both figures 1 and 2 start from the assumption that the parameters of the distribution are known. In case the parameters need to be estimated by a small sample, the conclusions might be different. This phenomenon is illustrated in Figure 3. The related triangular distribution has parameters a=5, m=12 and b=15 (same as in Figure 1), which theoretically leads to mean = 10.667, variance = 4.389 and positive semi-variance = 2.095. Samples of size 20 are taken from this distribution and the levels of safety stock are calculated based on sample mean, sample variance and sample positive semi-variance. From one sample, Figure 3 is created. The sample has mean = 11.521, variance = 3.860 and positive semi-variance = 1.399.

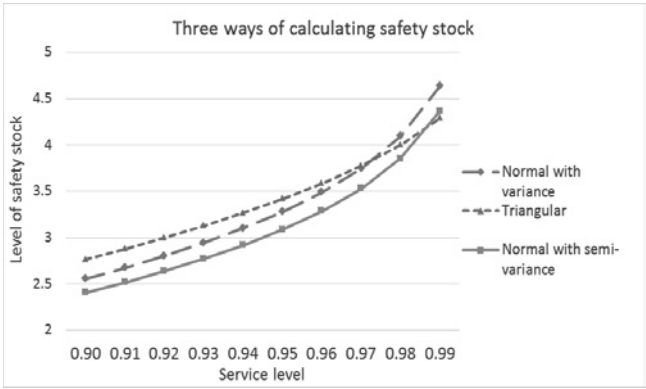


Figure 3: Level of safety stock in function of service level for a sample from the Triangular distribution (a=5, m=12, b=15)

The levels of safety stock based on the variance are higher than those based on the semi-variance. But the levels based on the variance are closer to the exact curve of the triangular distribution for the lower service levels, while those based on the semi-variance are closer for the higher service levels.

CONCLUSIONS

The performance of the method of determining safety stock levels, based on the variance and on a safety factor based on the Normal distribution, depends on the shape of the real but many times unknown demand distribution during lead-time. This fact inspires to use another way of determining the safety stock. This research paper has introduced the new concept of 'semi-variance' in this operational context. The paper illustrates this innovation by applying and evaluating the new concept by means of a case where the demand distribution during lead-time is known to be a triangular distribution with known parameters. Preliminary results show that the new concept can improve on the higher mentioned problem but not in a consistent manner. Further investigation is required to investigate the conditions in which it is a valid alternative. It looks interesting to check conditions like smaller/larger coefficients of variation, left/right skewness, bimodality or some combinations of these conditions. A suggestion might be to use the Compound Poisson distribution as variation of its parameters leads to several types of the conditions mentioned, as found in Ramaekers and Janssens (2007). The choice of this type of distribution forms a solid base for designing valid experiments in future research.

ACKNOWLEDGEMENTS

This work is supported by the Interuniversity Attraction Poles Programme initiated by the Belgian Science Policy Office (research project COMEX, Combinatorial Optimization: Metaheuristics & Exact Methods). The help of dr. Hanne Pollaris (UHasselt) for producing the figures is greatly appreciated.

REFERENCES

Bartezzaghi, E., R. Verganti and G. Zotteri. 1999. "Measuring the impact of asymmetric demand distributions on inventories." *International Journal of Production Economics*, 60-61, 395-404.

Chen, W.; J. Li and X. Jin. 2016. "The replenishment policy of agri-products with stochastic demand in integrated agricultural supply chains." *Expert Systems with Applications*, 48, 55-66.

Ehrhardt, R. 1979. "The power approximation for computing (s,S) policies." *Management Science,* 25, 777-786.

Grootveld, H. and W. Hallerbach. 1999. "Variance vs downside risk: is there really so much difference ?" *European Journal of Opererational Research*, 114, 304-319.

Harlow, W.V. and R.K.S. Rao. 1989. "Asset pricing in a generalised mean lower partial moment framework: theory and evidence" *Journal of Financial and Quantitative Analysis*, 3, 285-311.

Jacobs, R. and H. Wagner. 1989. "Reducing inventory system by using robust demand estimators." *Management Science,* 35, 771-784.

Janssens, G.K, and K. Ramaekers. 2011. "A linear programming formulation for an inventory management decision problem with a service constraint." *Expert Systems with Applications*, 38, 7929-7934.

Käki, A.; A. Salo and S. Talluri. 2013. "Impact of the shape of demand distribution in decision models for operations management." *Computers in Industry*, 64, 765-775.

Lau, H.-S., and A. Zaki. 1982. "The sensitivity of inventory decisions to the shape of lead time-demand distribution." *IIE Transactions*, 14, 265-271.

Markowitz, H.M. 1959. *Portfolio Selection: Efficient Diversification of Investments,* Wiley, New York.

Naddor, E. 1978. "Sensitivity to distributions in inventory systems." *Management Science*, 24, 1769-1772.

Ramaekers, K. and G.K. Janssens. 2007. "On the choice of a demand distribution for inventory management models." *European Journal of Industrial Engineering*, 2, No. 4, 479-491.

Ramaekers, K. and G.K. Janssens. 2012. "Service-oriented decisions on inventory levels in the case of incomplete demand information." *Logistics Research*, 5, Nos.1-2, 33-46.

MASS CUSTOMIZATION DYNAMICS SIMULATION FOR FASHION AND APPAREL MARKET

Jocelyn Bellemare
Department of Management and Technology, School of Business and Management, ESG
University of Quebec in Montreal (UQAM)
C.P. 8888, succ. Centre-ville, Montréal (Québec) H3C 3P8
Canada
E-mail: bellemare.jocelyn@uqam.ca

KEYWORDS
Fashion and Apparel industry 4.0, Mass Customization Program, Sticky information, Fitting Clothing, Products Configurator, Rapid manufacturing

ABSTRACT

This paper examines the potential of clothing mass customization configuration within the platform to identify the possibility to implement Industry 4.0 in the apparel industry. Even if, some manufacturers have managed this approach successfully, some of them have poorly mastered the concept. The increase in purchase returns for personalized and customized clothes both in stores and on the Web creates headaches for retailers because it affects their brand image. The first problem is related to the manufacturing aspects with measurements, adaptation of patterns and flexibility in methods and manufacturing deadlines. The second is the lack of knowledge and experience from the manufacturers to use properly the configuration systems. It has become increasingly important to understand how to create an approach for configurator implementation for clothing personalization and mass customization program. For producers to make the most of mass customization they need to better understand what can be done in terms of clothing personalization and mass customization capabilities. We discuss custom clothing for men in conjunction with the effects stemming from the evolution of mass production practices. This led us to explore from different angles the problems related to the automation of standard sizes and integration of "fits" done in traditional ways as well as computerized ways with respect to product adaptation. In this paper, we also analyze the mass customization concept and propose technological and operational approaches aimed at initiating useful discussions to better understand these issues.

INTRODUCTION

Past research has demonstrated the importance of understanding the mass customization of clothing within the context of trade globalization, which has led to ever more ferocious competition in the apparel industry. Moreover, as apparel products now seem to have an ever shorter life cycle, a phenomenon which is exacerbated by the introduction and implementation of new business models, businesses' commercial strategies must face mounting pressure. This situation forces the apparel industry players to revise their organizational strategies in order to survive in this highly competitive market. Organizations must reinvent themselves and find new ways to satisfy their customers. In order to grow, to maintain the current level of employment and possibly increase it, garment producers will need to develop new manufacturing strategies by orienting local production towards a flexible, quick-response system that allows for the production of various types of orders (small quantities, short deadlines, skilled labor, etc.). Thus, it will become essential for businesses to implement new strategies that correspond to the reality of current markets, in order to keep up with the rhythm of short cycle production. Businesses need to focus on flexibility, adaptability and agility (Pine, 1993).

FASHION AND APPAREL INDUSTRY 4.0

Fashion Apparel Industrie 4.0 called a ''smart apparel factory'' is the current trend of automation and data exchange in apparel manufacturing technologies. The combination of several major innovations in digital technology it includes the Internet of things, cloud computing, and cyber-physical systems communicate and cooperate with each other in real time used by participants of the value chain driving a new shift of change across the economy, with major implications for fashion market – including RFID, sophisticated sensors, digital printing and fabrication, 3-D product development and more. The consequences for fashion industry leaders are clear: more than ever before, they need to refocus on a few truly distinguishing core capabilities to create sustainable value in the future. In our view, these five priority of key success factors for Fashion Apparel Industry 4.0 : (1) customer excellence focus (The voice of the customer) and brand performance profile, (2) seamlessness in the omnichannel user experience integration, (3) renewed focus on physical retail, (4) operational excellence and innovation, (5) process Integration and Traceability. To remain strong and competitive, a company has to demonstrate its capacity to adapt in terms of creativity, production, quality, timing, and price.

MASS CUSTOMIZATION APPROACH

Reviewing the writings on this subject tells us that paradoxically, at a time where the global key word in most industries is standardization, the focus in the apparel industry

is on ‘‘uniqueness.’’ With the recent surge in the use of new media and telecommunication, consumers are more and more demanding and informed. They are no longer satisfied with standardized products that force them to make compromises. The Internet influences customers’ buying habits by creating needs that have to be satisfied instantaneously. In the clothing industry, these expectations not only imply having to constantly provide consumers with new options in terms of styles and colors, but also to allow them to find an affordable well-fitting clothing item and make it available to them almost as rapidly as if it was a standard-sized product. In order to meet these expectations, clothing companies must now propose custom-made products. Brands that offer personalized products (mass customization) are taking over both traditional and online stores. This is made possible by identifying the key points of body measurement necessary to produce well-adjusted, well-fitting garments. However, being able to take these measurements effectively and efficiently is crucial. Although efficient and affordable technologies are available to provide a body scan, few businesses are able to meet the requirements of custom-made products for the following reasons: lack of reliability of the measures provided by the body scan, problems related to the transmission of a large quantity of data to ‘’potential manufacturers, interface issues between the data generated by the body scan software and that used by pattern making, cutting and assembly.
Many apparel businesses are currently researching technological ways to produce, adjust, sell, and deliver, in a systematic and automatized fashion, personalized and made-to-measure products. Nevertheless, mass customization somehow remains misunderstood or is rarely used by actors in the clothing industry mainly because of the widely variable measurements, of the problems in adapting patterns, of the need for flexibility of manufacturing delays and methods. Many authors have produced research on mass customization; however, few of them have sought to identify the problems related to sizing and to so-called ‘hidden data’ coming from the customers (ease allowance, fullness, etc.).

Our objectives to develop a configurator for cothing mass customization, using computerized information systems, that could be used to analyze and decode measurement data coming from peripheral devices in order to identify as precisely as possible the necessary information to produce a well-fitting garment.

Hence, we need to identify the fundamental variables and data that are necessary to produce custom-made clothing. Parsimony in fundamental variables (length, circumference, density and textile matter behavior) will allow to significantly diminish the amount of data to analyze and send out in order to create an ‘intelligent’ pattern.

In this research, we hypothesize that the amount of non-essential data for pattern automation can be reduced by 65% (typical measurements), from the current situation. We shall not only try to reduce the quantity of data, but also to determine the minimal number of measures needed (minimum cardinalities). Moreover, using anthropometric measurements, like density, will enable us to identify key referential points which are essential to ensure proper fit. These referential points, when combined with data related to textile textures and behaviors, will allow for personalized pattern grading. To accomplish our primary objective we therefore must reduce the amount of data while increasing the quality of automated clothing patterns, thereby allowing for the production of well-adjusted custom-made clothing that meet customers’ needs and expectations

LITERATURE

The goal of mass customization is to efficiently provide customers with what they want, when they want it, at an affordable price. Inala (2007) contends that mass customization has become a competitive strategy for businesses that want to offer personalized products. A mass customizer must first identify the idiosyncratic needs of its customers, specifically, those product attributes along which customer needs to diverge the most (Piller and Blazek 2014). When clothing was made to measure, each garment was cut and assembled for individual customers (Istook 2002). As a result, it provided a personalized fit (Workman, 1991). This type of production is what Pine (1993) referred to as personalized and handcrafted production. Likewise, in order to be able to meet the demands of mass customization, all of a manufacturer’s operations have to be based, according to Zipkin (2001), on flexible processes that allow it to respond rapidly to customers’ requests. More often than not, mass customization consists in, for example, assembling basic items according to specific orders.

Mass customization therefore becomes a crucial development solution for businesses specialized in garment manufacturing and distribution (Pine, 1993). In fact, the demand for mass customization of clothing is only growing stronger. It has become possible thanks to the contribution of new technologies. Custom-made clothing requires a very thorough understanding of the expectations and specificities of each individual (Peterson, 2008). According to Pine (1993), the success of mass customization rests mainly on a successful integration of the value chain. In some respects, businesses must accomplish a feat by performing well on two axes that are generally on opposite sides of the spectrum in most businesses: maintaining short supply lead times while offering custom-made products that correspond to clients’ specifications.

There are mass markets for some customized products — the emergence of mass-customized apparel demonstrates that (Zipkin, 2001). The main problem of mass customization is related to the preparation of products according to customers’ requirements. Moon et Al (2014) states that because of their lack of knowledge and experience, consumers do not know what they really want. It is thus important to simplify their request by offering them some guidance. Doing so not only requires knowing a customer’s measurements and style, but also obtaining information that he never reveals: what literature refers to as “sticky information”. The term “sticky information” is defined by Von Hippel (1994) as information hidden by a customer that provides, in certain cases, a company with a key competitive advantage and offers significant opportunities for innovation. For example, consumers know their needs and tastes better than manufacturers. It is therefore difficult for a

manufacturer to obtain information that is either confidential or perceived to be so irrelevant that consumers reveal them sporadically, at best. This unknown data, like ease allowance, fit, proportion and the like, are essential to the production of custom-made garments. According to Ashdown (2013), they are at the source of most purchase returns occurring in stores.

CUSTOMER SATISFACTION

Customer satisfaction can kill: The increase in purchase returns for clothes both in stores and on the Web creates headaches for retailers because it affects their brand image, customer engagement and customer loyalty. People claim that it is very difficult to find clothes that fit them perfectly and they find that sizes vary from store to store. As a result, it appears important for stores to know their clientele and to offer clothes that fit customers adequately in order to increase their volume of sales per customer. Thus, some problems associated with mass customization must be corrected by the clothing industry, they are: the templates (blocks) used to create basic patterns are not adequate; the size standards and measurement charts have become obsolete; the sizing per territory/population rapidly changes; and, some of the information hidden by the customer must be decoded by manufacturers.

Faust and Carrier (2009) contends that errors in measurements still prevail in the clothing industry. Even if a customer is given a sizing chart, it is still difficult for him to take accurate measurements on his own. Ashdown (2013) has identified a few simple problems that might be encountered. For instance, when measuring waist circumference, it is necessary to stand straight in a natural position and to hold the tape measure parallel to the ground. A slight imbalance could result in errors of up to half an inch on the final garment. The main problem occurs when measuring the waist girth. Moreover, Park and Stoel (2002) mention that data transmission errors taking place during the data transfer process create problems at the time of order. As for the 3D body scan technology, it sends more than 300 000 data items during a sample body scan (Ashdown, 2007) which increases the complexity of selecting valid data in order to obtain reliable information.

Both methods (manual or 3D body scanner) of measurement have their strengths and weaknesses. For some authors, the biggest strength of the manual measurement is its ability to identify incoherent measurements while its most important weaknesses are the labor costs and the imprecision caused by human error when transcribing data (Fan et al., 2004). On the other hand, the strengths of the 3D body scanning are the speed and the low cost (nowadays) while its main weakness are in the measurement inconsistencies due to movement (Istook et al., 2011), the lack of accuracy when compared with manual measurements (Liu et al, 2014), and the difficulty to obtain correct measurements at feet position, for example (McKinnon & Istook, 2002). Accurate body measurements can be difficult to obtain with 3D body scanning due to factors such as posture, landmark indications, instrument position and orientation, pressure and tension exerted (Fan et al., 2004).

Not so long ago, body scanning was still at the stage of acceptance and maturation in the industry; the benefits of automation were not clearly visible (McKinnon & Istook, 2002). Female consumers who have been scanned generally react well to the results, yet women from specific socio-demographic groups are less comfortable with the idea of being body scanned (Loker et al., 2004).

As Whitestone & Robinette (1997) write, 3D body scanning is now an accepted tool in the apparel industry. As time goes by, the non-contact body measuring technologies generate more and more interest and applications in the apparel industry. It can be put to numerous applications: anthropometric measuring surveys, development of three-dimensional apparel, computer aided design (CAD), virtual garment environments and animation, mass-customization, etc. (Jones et al., 1995; Hardaker & Fozzard, 1998;, A., 2001; Koontz & Gibson, 2002; Xu et al., 2002; Ulrich et al., 2003; Bachvarov et al., 2014).

Ashdown (2013) indicates that computer systems need to accurately generate the information coming from both the pattern-making software and from the body scan. Issues arise when size charts and fit levels for different body types are not clearly established from the start. The key to success lies in the development, the architecture and the support of computer systems used to generate data based on individual body dimensions for pattern-making software, which need to be adapted individually. Despite the fact that all these approaches aim to produce apparel as accurately as possible, it appears that the great number of constraints makes it difficult to find a compromise between performance, accuracy and technicality during the production process.

CONFIGURATOR & CLOTHING PRODUCT DESIGN

Here, configuration processes play a crucial role to manage this task by providing customers support and navigation in co-designing their individual product or service. There is nothing simple about mass customization and it is not a simple strategy to undertake organizationally; it is not even a simple concept to comprehend (Hart, 1994). Today's market heterogeneity, increasing variety, steadily declining product life cycles, decreasing customer loyalty, and the escalating price competition in many branches of industry are the main motivators for firms going into mass customization (Pine 1993).

Configuration is an essential aspect of mass customization because it creates the possibilities to guide customers as they are making choices. Haug et al (2012) contends that the primary objective of a configurator is to facilitate the decision-making process of customers using a Web-based interface. Product configuration systems play an important role in supporting the mass customization paradigm, as it helps to determine the degree of personalization that a business will offer. Thus, the role of the configurator is to create a link between consumers and manufacturers (Inala, 2007). Mass customization does not equate to an increase in costs. According to Piller and Blazek (2014), using a configurator could significantly reduce costs since its Web-based technology diminishes the time required to take orders

and the application of toolkits for customer co-design may be the most used approach to help customers navigate choice in a mass customization system. In the current context, businesses use catalogs and manual production methods. Catalogs provide a predefined and limited number of combinations for a product without necessarily fulfilling all of a customer's specific needs (Quin and Yang, 2009). Manual configuration, on the other hand, essentially relies on the human expertise and necessitates competent and highly skilled workers (Rogoll and Piller, 2004). However, a lack of expertise eventually requires investments in terms of time and efforts; moreover, it forces employees to keep up to date with frequent technical changes and improvements. As a result, the configuration of a product to meet a customer's requirements can become a complex task which gets more demanding as the number of components and options increases. When the configuration requires numerous variations, the possibility of making errors also rises which can result in production delays. The repetition of subsequent steps may be required which can be costly. Thus, Ashdown (2007) contends, mass customization creates various technical challenges that need to be overcome before mass customized garments can be produced. The technological risks associated with a configurator project are essentially related to the development of a system that can share and process data and parameters (the parameter configurator) originating from various sources such as: the data entry tools (e.g. the Body scan), the basic garment patterns, the marker-making software, the automatic cutting table and the administrative and financial data. In short, none of the existing technological systems seem to provide a solution for mass customization in the apparel industry. Rogoll and Piller (2004) indicate that the optimal product configurator needs to create an interface between different programming languages and function entirely independently. Incidentally, these criteria add to the level of uncertainty associated with this type of installation. A product configurator must be used along with a high-performance technological platform so as to allow for interaction between customer and manufacturer as the product is designed. This creates an interface between the customer and the supplier which provides opportunities for value co-creation in both the apparel and fashion industries.

The most important mass-customization prerequisite is the understanding that mass customization itself is a highly customized strategy and you cannot imitate someone else's successful mass-customization strategy (Hart 1994). If prime producers want to make the most of this prospect, they will need to better understand what can be done in terms of clothing personalization and mass customization so as to formulate an appropriate strategy on how to use their measurement configurator. This research project will provide tools for fashion industry businesses that will allow them to gain a competitive edge through custom-made and short lead time projects. The opportunities created by the absence of such a service or system needs to be used by businesses in this industry to reposition themselves on the garment and apparel markets, both locally and internationally. This research offers great possibilities in terms of innovation and could constitute an outstanding opportunity for several actors in the fashion and clothing industry. Even though the local garment industry and that of emerging countries face each other on an uneven playing field, the local industry possesses a technological environment that could give it a significant advantage.

INDUSTRIAL DYNAMICS SIMULATION

The first stage of this research project is a preliminary study of the fundamental variables and data needed to produce a custom-made garment. This first step will allow for the production of a study which is itself an integral part of a larger research project. The proposed approach will aim, in part, to identify the fundamental variables and data essential to the fabrication of custom-made apparel. After this results have been submitted, the data obtained will be analyzed which will allow for the creation of a product parameter configurator. Moreover, in the near future, we will assess the modalities of implementation of this technology and its progressive use in the fashion and garment industries. The preliminary phase of this research project will take place in a manufacturing environment specialized in men's fashion.

At first, we will study the mass production and custom-made environments that exist in this industry. Next, we shall analyze three pants models provided by manufacturers specialized in athletic wear, sportswear and workwear. Each pattern will be analyzed and dissected in order to assess the fitting and grading methods used in relation to size and type of textile. From this first study, we will formulate a hypothesis on the fundamental variables needed to produce a garment using mass customization. In order to validate the fundamental variables that will enable us to create our configurator, we will conduct a study of the process involved in body measurement (length, circumference, density and stature) using both a body scan and manual measurements. We collected complete measurement data for 60 male subjects, aged 18 to 69.

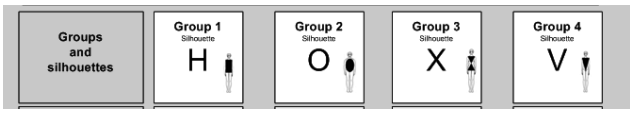


Fig. 1: Analysis of shapes processes and methods

In a same group of 60 men, 12 men will be recruited to allow us to model the variables and data linked to a production model that is part of a real rapid manufacturing process. So as to facilitate research based on individual shape groupings, we will use figure types represented by the letters H-O-X and V to categorize different types of silhouettes and redefining silhouettes from Rasband and Leichty (2006). Four morphotype-groups will be made up of men (of different stature) wearing a size 40 jacket and trousers of sizes 32 to 38. This innovative method significantly improves the recurring problem in the industry regarding classification systems of normalisation. It will then be possible to validate the data through our configurator and produce garments using rapid prototyping. A thorough examination of the clothing items produced will be carried out during the fitting phases in order to analyze their "fit". This will allow us to determine which variables appear to be problematic. Mass customization offers a new business model and growth opportunities for small manufacturing businesses and clothing companies. Indeed, from mass or

large volume production, businesses in this industry will be able to profit from this value-added advantage. According to Zipkin (2001), this type of production will be possible on a large scale because new technologies will become more easily accessible. This project originated from the idea of creating the "optimal" product configurator which would have the capacity to efficiently translate customers' desires and associate them with their anthropometric and anthropomorphic characteristics.

It appears obvious, following our measurements and interview activities involving 60 male individuals in the integral part of a larger research project, that the single pattern with respect to standard sizes is inadequate to meet the needs of the population. Personalization therefore is deemed to have a great future within the apparel industry. Based on measurements systematically made on several pairs of pants, it appears that manufacturers have a major issue with respect to consistency in their productions, not to mention that the underlying patterns are far from perfect.

Based on this data, three different pairs of pants of the same quality, brand and manufacturer will feel different on an individual. Independent of our approach to mass customization, we are attempting to solve this problem through this research because it is useless to try and find the perfect pattern for a given individual if the manufactured trousers do not conform to the pattern. In addition, following our measurement activities and meetings with master tailors, it appears that the measurements are taken manually or in an automated fashion by body scan will not suffice to guarantee a minimum fit criteria when it comes to custom trousers for given customers. Other types of data are determined essential and more important than measurements.

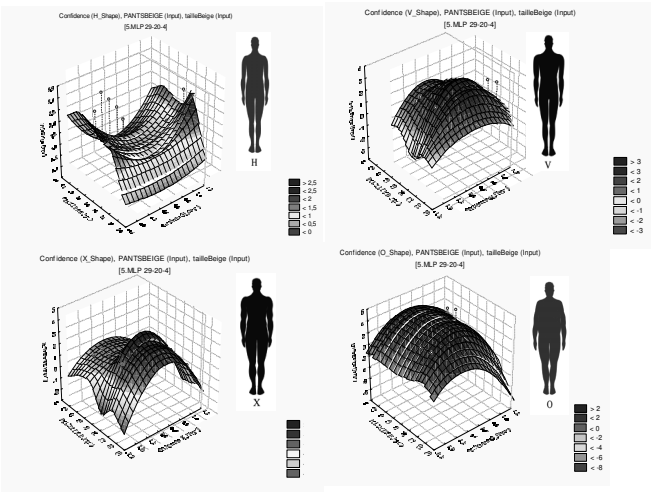


Fig. 2: Result of action and assessment of technological simulation tools used in the industry (cluster *fit*)

This is confirmed (figure 1) following our meetings with North American pattern makers and manufacturers. When analyzing the process of creating patterns for different sizes from a master pattern, it appears that it will be very difficult to create a specific pattern for every customer. This type of fit patterns automation is deemed neither feasible nor necessary.

Analysis results of the 60 individuals sampled show that we can identify approximately 12 customer profiles for every size. It would therefore be needed to design from known data, not one but 12 patterns for each size to accommodate the entire population with pants that would fit just as well as custom-made ones.

The issue with mass customization would then rest on rapidly identifying the customer profile from the 12 standard profiles, and produce a pair of pants from the corresponding pattern. Tests made with respect to neural network aspects show that it is possible to automatically classify all individuals with relatively fewer sizes than what is conventionally used (only 65% of typical sizes), however adding information from body analyzer/weight data providing fat and bone mass data (in the form of lean body mass statistics), body water percentage and data concerning fit perception. On the whole, these results allow us to envision the logistical aspects within an installation that would use mass customization methods. The methods also become different for patternmakers since instead of creating one pattern for each type (i.e: master pattern for size 32), 12 patterns for each silhouette type would then be created. Then, the same extraction methods for grade units using master grade units would yield the 12 patterns for each grade. Thus, by obtaining a body scan through Kinect Xbox-3D and data from a short survey/questionnaire, the manufacturer will then automatically obtain the silhouette data of target customers.

Currently, tests with a configurator, confirm the validity of our variables and the future potential for rapid prototyping by a mass individual production and assure a well-fitting garment via an online request. This method can be applied for professional, commercial, technical and mass consumer apparel. Through this work, it is also seen that it would be beneficial to label ready-to-wear trousers with silhouette-type information that best displays the style. This would no doubt allow the customer to filter more quickly through non-desired pairs or models. This project offers numerous innovative possibilities and could provide a major opportunity for those implicated in the apparel industry.

CONCLUSION

Digital capabilities are vital to move forward with Industry 4.0. Apparel industry businesses must be proactive, adopt, and adapt to new mindsets and management tools to take full advantage of information technologies. To successfully implement mass customization, it is of the utmost importance that they emphasize analysis, decision-making, performance evaluation, and added value. Indeed, flexibility is a must as the market increasingly expects it. Mass customization offers much potential for extending brand awareness, acquiring new markets and generating profits. Customers will be at the center of the changes to value chains, products and services. In the end, manufacturers and retailers will need to own relationships with the end customers who drive demand or at least integrate with platforms that allow them to access the end customers efficiently.

REFERENCES

Ashdown, S.P (2013), Creation of ready-made clothing : the development and future of sizing systems. Chapter 2. Faust, M. E. Carrier & S. Designing apparel for consumers: the impact of body size and shape, Cambridge, UK: Woodhead Publishing, (1): 17-31

Ashdown, S.P. (2007), Cambridge Sizing in clothing : developing effective sizing systems for ready-to-wear clothing, Woodhead Publishing in association with The Textile Institute, Boca Raton : CRC Press: 384.

Bachvarov, A., Maleshkov, S., Chotrov, D. (2014), ''Extending Configuration and Validation of Customized Products by Implicit Features in Virtual Reality Environments''. Proceedings of the 7th World Conference on Mass Customization, Personalization, and Co-Creation (MCPC 2014), Aalborg, Denmark, February 4th - 7th, Lecture Notes in Production Engineering 2014, pp. 189-199.

Fan, W.J. & Hunter, L., (2004), Clothing Appearance and Fit: Science and Technology, Woodland, Publishing Ltd., Cambridge.

Faust, M.-E., & Carrier, S. (2009), A proposal for a new size label to assist consumers in finding well-fitting women's clothing, especially pants: An analysis of size USA female data and women's ready-to-wear pants for North american companies. Textile Research Journal, 79 (16): 1446-1458.

Hardaker, C.H.M. & Fozzard, G.J.W. (1998), "Towards the virtual garment: three-dimensional computer environments for garment design", International Journal of Clothing Science and Technology, 10 (2): 114-12.

Hart, C.W.L. (1994), Mass customization: conceptual underpinnings, opportunities and limits, International Journal of Service Operations, 6 (1): 36-46.

Haug, A., Hvam, L., & Mortensen, N.H. (2012). "Definition and evaluation of product configurator development strategies". Computers in Industry 63(5), pp. 471-481.

Inala, K. (2007), Assessing product configurator capabilities for successful mass customization, University of Kentucky, theses.

Istook, C., Newcomb, E., & Lim, H (2011). 3D technologies for apparel and textile design. In J. Hu (Ed.), Computer Technology for Textiles and Apparel . Cambridge, UK : Woodhead Publishing.

Istook, C.L. (2002), Enabling mass customization: computer-driven alteration methods, International Journal of Clothing Science and Technology. 22 (1): 16-24.

Jones, PRM., Li, P., Brooke-W, K., & West, G. (1995), Format for human body modelling from 3-D body scanning, International Journal of Clothing Science and Technology, 7 (1): 7-16.

Koontz, M.L., & Gibson, I.E. (2002), Mixed Reality Merchandising: Bricks, Clicks and -- Mix. Journal of Fashion Marketing & Management, 6 (4): 381-395.

Liu, Z., Li, J., Chen, G., Lu, G. (2014), Predicting detailed body sizes by feature parameters. International Journal of Clothing Science and Technology, 26(2):118-130.

Loker, S., Ashdown, S. P., Cowie, L., & Schoenfelder, K. A. (2004), Consumer interest in commercial applications of body scan data. Journal of Textile and Apparel, Technology and Management, 4 (1): 1-13.

McKinnon, L. & Istook, C. (2002), "Body Scanning: the effects of subject respiration and foot positioning on the data integrity of scanned measurements", Journal of Fashion Marketing and Management, 6 (2): 103-121.

Moon, H & Lee, H-H. (2014), "Consumers' preference fit and ability to express preferences in the use of online mass customization", Journal of Research in Interactive Marketing, 8(2): 124-143.

Park, J. H. & Stoel, L. (2002), Apparel shopping on the Internet: Information availability on US apparel merchand Web sites. Journal of Fashion Marketing and Management, 6 (2):158-176.

Peterson, J. (2008), Mass customisation finds favour (clothing industry). Knitting International, 114 (1360): 36-37.

Piller, F., Blazek, D. (2014), Core Capabilities of Sustainable Mass Customization, Knowledge based Configuration, (chap.9) : 107-119.

Pine, B.J. (1993), Mass Customization: New Frontier in Business Competition, Harvard Business School Press, Boston, MA.

Qin, S., & Yang, L. (2009), Fuzzy optimisation modelling for apparel fit from body scanning data mining, Proceedings of the Sixth International Conference on Fuzzy Systems and Knowledge Discovery, Tianjin, China, (7) : 255-259.

Rogoll, T. & Piller, F. (2004), Product configuration from the customer's perspective: a comparison of configuration systems in the apparel industry, Proceedings of International Conference on Economic, Technical and Organisational Aspects of Product Configuration Systems, Lyngby (June): 28-29.

Ulrich, P. V., Anderson-Connell, L. J. & Wu, W. (2003), Consumer co-design of apparel for mass customization, Journal of Fashion Marketing & Management, 7 (4): 398-412.

Von Hippel, E. (1994), "Sticky Information" and the Locus of Problem Solving: Implications for Innovation. Management Science, 40 (4): 429-439.

Whitestone, J. J., & Robinette, K. M. (1997), Fitting to maximize performance of HMD systems, In J. Melzer & K. Moffit (Eds.), Head-Mounted Displays: Designing for the User, New York: McGraw-Hill, 206-215.

Workman, J.E. (1991), Body Measurement Specifications for Fit Models as a Factor in Clothing Size Variation, Clothing and Textiles Research Journal, Vol. 10 (1): 31-36.

Xu, B., Huang, Y., Yu, W., & Chen, T. (2002), Three-dimensional body scanning system for apparel mass- customization, Optic Engineering, 41 (7): 1475-1479.

Zipkin, P. (2001), The Limits of Mass Customization, Sloan Management Review, 42: 81-87.

SIMULATION-OPTIMIZATION: A SIMPLE APPROACH COMBINING METAHEURISTICS AND METAMODELS

Luiz Ricardo Pinto
Júlia Cobucci Morais
Gabriela Martins Nunes
João Flavio de Freitas Almeida
Department of Production Engineering
Federal University of Minas Gerais
Av. Antônio Carlos, 6627- Pampulha
Belo Horizonte, MG, 31.270-901, BRAZIL

KEYWORDS
Simulation-Optimization, Metamodeling, Simulated Annealing.

ABSTRACT

The use of simulation-optimization has increased in recent years. This technique is useful when the objective function and/or any constraints could not be assessed analytically. Beginners in simulation analysis frequently use commercial simulation packages, which include optimizers, to make this analysis. However, many of them see those optimizers as a black-box that searches for an optimum. Simulation-optimization techniques may be time-consuming. In this paper, we propose a simple combination of a metaheuristic and metamodel to show how easy is building simple algorithms that are able to perform this analysis in a competitive time. We have proposed a simple methodology where a simulation model, a metamodel and a metaheuristic work together in a simple loop to obtain solutions for an inventory problem. We also compare the solutions against those obtained by a commercial package.

INTRODUCTION

The ordinary simulation-optimization methodology uses a loop where a simulation model assesses the inputs proposed by an optimization model, which may use a variety of heuristics. As we see in Figure 1, each input vector θ_i feeds the simulation model which generate a set of performance measures, i.e., an output vector ω_i, which is analyzed by the optimization model that propose a new input vector θ_{i+1}. This loop remains until a predefined stop condition occurs. This methodology could be time-consuming, mainly in commercial simulation-optimization packages because they do not take the advantage of knowing, in advance, specific features of the system.

The basic idea proposed in this study is using a metamodel, which is an analytical model that represents the simulation model, to make a preliminary assessment of inputs proposed by heuristics, in order to simulate only promising configurations.

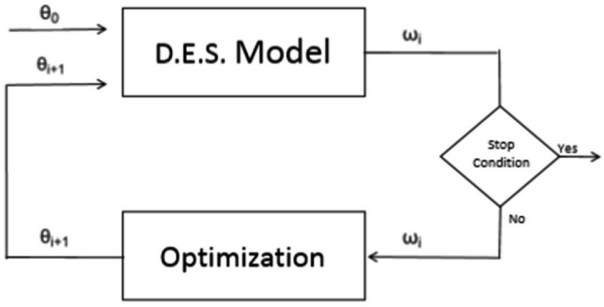


Figure 1: Ordinary simulation-optimization loop.

The addition of metamodel is shown in Figure 2.

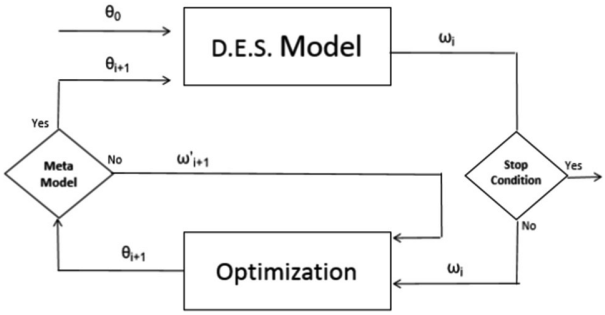


Figure 2: Simulation-optimization loop with metamodel.

The metamodel makes a preliminary evaluation of the input vector θ_{i+1} which is generated by the optimization model and if the response is not feasible or it is not considered good in a specific aspect, the output vector produced by metamodel $\omega'i+1$ is sent to optimization model to generate a new input vector. This process saves the run time that would be spent for that scenario in the simulation model. This is supposed to reduce the computing time, since the assessment made by metamodel is deterministic and therefore, much faster than evaluation through simulation model.

To illustrate the use of the proposed method against the use of ordinary simulation-optimization techniques, we used an

inventory system adapted from one proposed by Law (2007) and a similar model proposed by Biles et al. (2007). The preliminary results obtained in this study show the applicability of the proposed methodology and the good performance regarding time-consumption.

SIMULATION-OPTIMIZATION, METAMODELING AND DESIGN OF EXPERIMENTS

There are many packages to solve simulation-optimization problems, that are provided together commercial simulation software (Fu, 2002). Although all these packages seem like efficient, they may be time-consuming for some problems because they do not use specific features of the system which is simulated, i.e., the heuristics routines embedded into these optimizers are quite generic and the system spend much time "to learn" about the simulation model.

In this study, we have used a simple heuristic and metamodeling, just to illustrate the proposed methodology. We do not intend to make a very efficient heuristic algorithm neither a complex metamodeling in this study. The objective is showing how we can build a competitive alternative to perform simulation-optimization without the obligation to do something novel or even with a high degree of complexity. We use a well-known heuristic, Simulated Annealing (Kirkpatrick et al. 1983), and a simple polynomial regression to build the metamodel. In the same way, we use a modified Latin Hypercube Design (LHD) to plan experiments to generate data to build the metamodel.

Metamodels can be seen as simulation model surrogates, i.e., a model of a simulation model. These models, or metamodels, try representing the relationship between inputs and outputs of a simulation model. Barton (2015) presents a very useful tutorial on metamodeling, where he explains the basics, purposes, process and types of metamodels. However, the debate on this subject comes from decades ago. Barton himself addressed this theme twenty-five years ago (Barton, 1992).

Li et al. (2010) made an assessment of five types of metamodels commonly used in simulation studies. They also proposed an incorporation of metamodels in decision support systems to improve the performance of them. Can and Heavey (2012) also performed a comparative analysis of two types of metamodels and they studied three applications, one of them uses an inventory system close on the system we have used here. Other studies used kriging metamodel to study the same inventory system (Biles et al., 2007; Zakerifar et al., 2011). Chen and Li (2014) studied designs of experiment for metamodels and they also used the inventory system.

To build metamodels that provide similar responses to those provided by simulation models, we need good samples of the system behavior. There are several techniques to do this, and one frequently used to design experiments in simulation studies is Latin Hypercube Design (LHD) proposed by McKay et al. (1979). LHD is spacing filling, and the arrangement of samples provided by this type of design has many advantages. An overview of design for simulation experiments is given by Kleijnen (2005), where he explains the use of LHD for build kriging metamodels. In this study we did not use exactly LHD, but a design of experiments based on LHD instead.

Amaran et al. (2016) present a wide review of algorithms and applications of simulation-optimization. In many cases, the optimization model is driven by heuristics and there are many options for choosing. We have chosen Simulated Annealing, proposed by Kirkpatrick et al. (1983).

THE PROPOSED INVENTORY SYSTEM

The problem presented here is a version of the models proposed by Law (2007) and Biles et al. (2007). The inventory system refers to a single product, which is stored in a single warehouse. In this model, we include many choices of replenishment and lead time policies, not existing in the previous models.

There are 10 replenishment policies *r*, each one with a cost per order plus a cost per unit ordered as we show in Table 1. We must choose just one policy to manage the system. With exception of option number one, all others have a minimum payment, regardless the number of units ordered. Therefore, the choice of the replenishment policy can be critical in terms of costs.

Table 1: Replenishment policies.

ID (r)	Cost per order ($)	Unitary Cost ($/unidade)	Minimum Payment ($)
1	100	3	-
2	92	3	500
3	84	3	550
4	76	3	600
5	68	3	650
6	60	3	700
7	52	3	750
8	44	3	800
9	36	3	850
10	10	3	900

The lead time is also dependent on the policy chosen. Shorter lead times imply in additional overhead costs. Table 2 shows the options available for the lead time policy *t*.

Table 2: Lead time policies.

ID (t)	Time in days	Additional Fixed Cost ($)
1	UNIF(5.0,5.5)	-
2	UNIF(4.5,5.0)	10
3	UNIF(4.0,4.5)	20
4	UNIF(3.5,4.0)	30
5	UNIF(3.0,3.5)	40
6	UNIF(2.5,3.0)	50
7	UNIF(2.0,2.5)	60
8	UNIF(1.5,2.0)	70
9	UNIF(1.0,1.5)	80
10	UNIF(0.5,1.0)	90

Similar to the replenishment policy, we must choose just one option. Except for option number one, all others add an extra overhead cost.

The inventory is checked at the end of each week, and if it is below the replenishment point, we should trigger a replenishment order. The quantity of items requested must complete the entire warehouse level ***S***, and also supply the missing items of the outstanding orders that have come up until that time, without surplus units.

Therefore, the objective of the problem is selecting the size of the warehouse *S* and the replenishment point *s*, the replenishment (*r*) and lead time (*t*) policies to minimize the total cost, i.e., the sum of holding, shortage and order costs. In addition, we have constraints for average holding and shortage costs, which averages should not exceed \$100/day and \$30/day, respectively. The size of the warehouse *S* can range from 50 to 250 slots. Similarly, the replenishment point *s* can range from 50 to 250 units. However, for making the analysis easier, instead of using *S*, we will make our decision in terms of the difference $d = S - s$, i.e., we use as input vector *d, s, r, t*.

The demand is supposed to be 10 orders/day, following a Poisson distribution. The average number of units per order is 2.5, following the empiric distribution, where the number of units is: i) 1 unit – 17%; ii) 2 units – 33%; iii) 3 units – 17%; and iv) 4 units – 33%.

SYSTEM MODELING

We developed an entire simulation-optimization model covering all phases, from the simulation model, design of experiments (DoE), metamodeling, and simulation-optimization algorithm.

Firstly, we built the simulation model of inventory system described in section 3 in Arena. In this model, the input vector θ has four parameters: i) replenishment point *s*; ii) minimum lot size *d*; iii) identification of order policy *r*; and iv) identification of lead time policy *t*. The output vector ω has three performance variables: i) holding cost *Cth*; ii) shortage cost *Cts* and iii) total cost *Ctt*. All maintenance costs are on \$/day basis.

Secondly, we perform a design of experiments (DoE) in order to have a database for building the metamodel. This DoE was based on Latin Hypercube Design (LHD) for proposing the samples. In this design we propose 100 experiments, each one with a distinct input vector θ. For each vector θ_i, we run a simulation with 10 replications of 1 year long, to estimate the output vectors ω_i.

With all pairs θ_i and ω_i we built three metamodels to represent the holding, shortage and total costs (C*th*, *Cts* e *Ctt*). All metamodels used second order polynomial regression of the form given by (1).

$$f(\theta) = \sum_i \beta_i p_i(\theta) + \varepsilon \qquad (1)$$

Where $p_i(\theta)$ is a product of power functions, shown in Table 3.

Table 3: Functions used on metamodeling.

i	$p_i(\theta)$	*i*	$p_i(\theta)$	*i*	$p_i(\theta)$	*i*	$p_i(\theta)$
0	*1*	*5*	*s.d*	*10*	*r.t*	*15*	*s.d.r*
1	*s*	*6*	*s.r*	*11*	s^2	*16*	*s.r.t*
2	*d*	*7*	*s.t*	*12*	d^2	*17*	*s.d.t*
3	*r*	*8*	*d.r*	*13*	r^2	*18*	*d.r.t*
4	*t*	*9*	*d.t*	*14*	t^2	*19*	*s.d.r.t*

Figure 3 shows comparisons between results from metamodels and simulations: (a), (b) and (c) show 100 experiments used to build metamodels and (d), (e) and (f) show other 50 new random experiments, distinct from those ones.

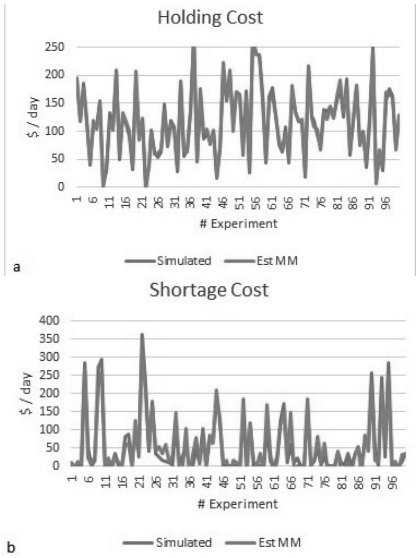


Figure 3 (a-b): Results from Simulation X Metamodels.

As we can see, the fitting is excellent for holding and shortage costs and a little bit worst for the total cost. Probably, we could improve this fitting using another type of metamodel, but the fitting seems appropriate for the purpose of this study.

Finally, the formulation of the optimization problem is given by (2), (3) and (4).

$$\underset{\theta \in \Theta}{Min}\ f(\theta) = E\left[Ctt(\theta, \varphi)\right] \qquad (2)$$

where θ is the vector of input variables, i.e., the values of *s*, *d*, *r* and *t*; $f(\theta)$ is the total cost on \$/day basis; $Ctt(\theta, \varphi)$ is the total cost of θ in replication φ ; and $E\left[Ctt(\theta, \varphi)\right]$ is the expected value of $Ctt(\theta, \varphi)$. The constraints are:

$$E[Cth(\theta, \varphi)] \leq MaxH \qquad (3)$$

$$E[Cts(\theta, \varphi)] \leq MaxS \qquad (4)$$

Where $E[Cth(\theta,\varphi)]$ and $E[Cts(\theta,\varphi)]$ are the expected values for holding and shortage costs, *MaxH* and *MaxS* are maximum values for them, respectively.

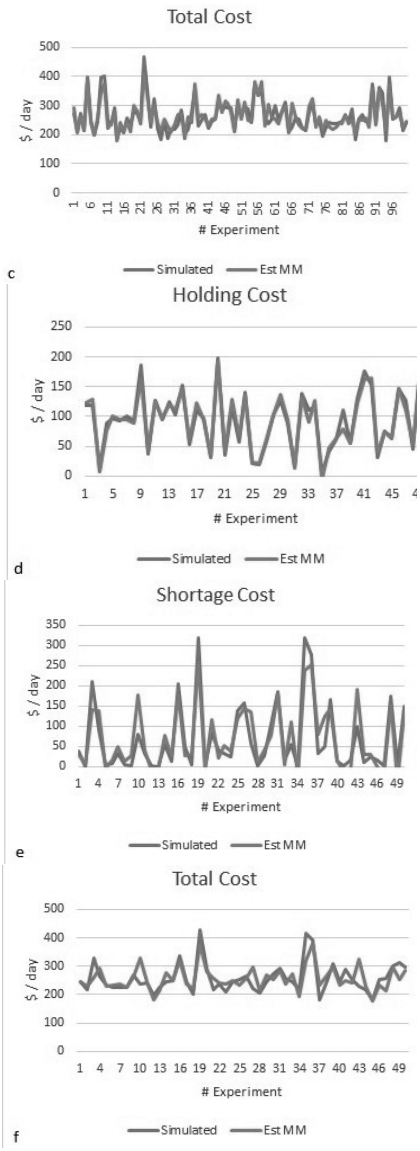


Figure 3 (c-f): Results from Simulation X Metamodels.

To get a solution, we use a simulation-optimization model, where the optimization model suggests candidate solutions θ_i, and the simulation model estimates the performance vector ω_i of each one, i.e., the values of *Ctt*, *Cth* and *Cts*.

The optimization model is not an optimization model in a formal sense, but a metaheuristic, which is search algorithm shown on Figure 4. The algorithm is a modified version of Metropolis Algorithm or simulated annealing (Kirkpatrick et al, 1983). The algorithm is based on a technique named annealing in metallurgy, where the initial temperature is *TIni* and it decreases in discrete steps, using a factor α, until a final temperature *TFin* or until occurs a maximum number of perturbations (*LPer*) without changing the current best solution.

Input data:
Set predefined data for *TIni, IPer, α, LPer, ProbMM, MaxCth, MaxCts*
Initialization:
  Read initial input vector θ_0 (s_0, d_0, r_0, t_0)
  Set $\theta = \theta_0$ and $\theta^* = \theta_0$
  Run simulation model for θ to estimate $\omega(\theta)$
  Set $Ctt^* = Ctt(\theta)$
  Set $T = TIni$ and $nc = 0$
Temperature loop:
repeat
  Set $i = 1$
  Perturbation loop:
  repeat
    Set $i = i + 1$
    Set $nc = nc + 1$
    Set $\theta_{try} = \theta$
    repeat
      Perturbation of θ_{try}
      Estimate ω_{try} $[Ctt(\theta_{try}), Cth(\theta_{try})$ and $Cts(\theta_{try})]$ via *Metamodel*
    until $(Ctt(\theta_{try}) < Ctt(\theta)$ and $Cth(\theta_{try}) < MaxCth$ and $Cts(\theta_{try}) < Maxcts)$ or with probability *ProbMM*
    Run simulation model for θ_{try} and estimate ω_{try}
    If $(Ctt(\theta_{try}) < Ctt(\theta)$ and $Cth < MaxCth$ and $Cts < Maxcts)$ or with probability $prob = exp\{-[Ctt(\theta try) - Ctt(\theta)]/T\}$
    then
      Set $\theta = \theta_{try}$ and $\omega(\theta) = \omega(\theta_{try})$
    end if
    If $(Ctt(\theta_{try}) < Ctt^*$ and $Cth(\theta_{try}) < MaxCth$ and $Cts(\theta_{try}) < Maxcts)$
    then
      Set $\theta^* = \theta_{try}$ and $\omega(\theta^*) = \omega(\theta_{try})$
      Set $i = i + 1$
      Set $nc = 0$
    end if
  until $i >= IPer$
  Set $T = \alpha.T$
until $T < TFin$ or $nc >= LPer$

Figure 4: Heuristic algorithm.

Basically, we start with an initial input vector θ_0, and the algorithm proposes a sequence of vectors, where each vector (θ_i) is a perturbation of the previous one (θ_{i-1}). For each vector θ_i, the algorithm assesses the performance vector, ω_i. The main heuristic algorithm is shown in Figure 4 and the perturbation algorithm on Figure 5.

Throughout the search, for each level of temperature, the algorithm proceeds a predefined number of perturbations (*IPer*) in the input data vector θ. Each perturbation changes vector θ to θ_{try}, which is assessed via metamodel to check its feasibility and its performance ω_{try} (costs). If the evaluation using metamodel is feasible and gets good performance, the vector θ_{try} is send to simulation model to confirm and/or refine the try input vector. Even θ_{try} leads a worst solution, the algorithm can sent it to simulation model with a static probability *ProbMM*.

```
Input data:
Set predefined data to Δs and Δd
Read the previous input vector θi-1 and check if it improved the solution
If θi-1 improved the solution
then make a local search in order to try improving ω
    If θi-1 changed s
    then
        If θi-1 increase s
        then continue increasing s
            s ← s + Δs
        else continue decreasing s
            s ← s – Δs
    end if
    If θi-1 changed d
    then
        If θi-1 increase d
        then continue increasing s
            d ← d + Δs
        else continue decreasing d
            d ← d – Δs
    end if
else make a random change in 2 elements of θi-1 (s or d and r or t)
Set θtry = θi-1
```

Figure 5: Perturbation algorithm.

Table 4 shows the initialization values of all parameters of the algorithm.

Table 4: Initializations parameters of the search algorithm.

Parameter	Description	Value
TIni	Initial value of the temperature parameter	180
IPer	Number of perturbations in each level of temperature	100
α	Factor used to decrease the temperature in each step	0.95
TFin	Final temperature	0.001
LPer	Maximum number of perturbations without changing the solution	2000
ProbMM	Probability to metamodel accept worst solution	0.1
MaxCth	Maximum value for the holding cost	100
MaxCts	Maximum value for the shortage cost	30

After the evaluation of θ_{try} by simulation model, if the output vector ω_{try} is really feasible and gets a better solution than current input vector θ, θ_{try} replaces θ. Even θ_{try} leads the worst solution, θ_{try} can replace θ with a probability given by (5):

$$prob = e^{\frac{-[Ctt(\theta_{try}) - Ctt(\theta)]}{T}} \qquad (5)$$

As we can see in (5), the probability *prob* has a direct relation with temperature *T*, i.e., in the beginning of the process (high temperatures) changes to worst solution to escape of local optima is more likely than in the end of the process (low temperatures). Similarly, the probability *prob* has inverse relation with the difference between $Cct(\theta)$ and $Cct(\theta_{try})$.

Furthermore, if θ_{try} leads a solution better than θ_{best}, θ_{try} replaces θ_{best}. The loop remains until the temperature reaches a minimum value (*TFin*) or until n iterations occur without changes in the best solution.

ANALYSIS AND RESULTS

We use five random experiments in order to assess the methodology. Despite randomness in values for *s*, *d*, *r*, and *t*, we have made the choice attempting combining a high, low and intermediate value for each one. We compare the results using a commercial simulation-optimization package (OptQuest for Arena), a single simulated annealing algorithm and a combination of simulated annealing and metamodeling as we proposed in this study. The performance of the algorithms and the quality of results depends on the stop conditions. In this case, we used two conditions: the value of final temperature (*TFin*) and the maximum number of iterations without changes in the best solution (*LPer*). The former we remain static and we run experiments varying the last condition. As we increase *LPer*, the result is better but the time-consumption is, obviously, high. All results are shown in Table 5.

Table 5: Results from simulation-optimization.

Measure	OPT	SA 500	SA 2000	SA MM 1000	SA MM 2000	SA MM NL
Time (min)	18.79 ± 2.84	7.21 ± 0.10	25.79 ± 9.70	2.20 ± 0.60	3.60 ± 1.42	190.57
***Ctt* ($/day)**	176.39 ± 0.2	177.15 ± 0.0	176.35 ± 0.1	178.32 ± 0.9	177.14 ± 0.4	176.32
Iterations SA	N/A	908 ± 2	3259 ± 1128	271 ± 77	445 ± 180	24055
Iterations MM	N/A	0	0	2196 ± 572	3463 ± 1566	240014
Tot iterations	872 ± 149	908 ± 2	3259 ± 1128	2467 ± 646	3908 ± 1746	264014
Best scenarios						
s	189	172	215	228	198	185
d	89	91	63	53	82	93
S	278	263	278	281	280	278
r	3	3	3	3	3	3
t	1	2	1	1	1	1

The first column in Table 5 refer to performance measure, and the following refer to the results via OptQuest, Simulated Annealing with *LPer* = 500 and *LPe*r = 2000, Simulated Annealing and Metamodel with *LPe*r = 1000 and *LPer* = 2000. The values are 95% of confidence interval.

In the last column of Table 5 we include a simulation where we did not set a limit for iterations. Therefore, the only stop condition was the algorithm reaching *TFin*. Furthermore, we reduce the temperature in slight steps using α = 0.95 and increase *TIni* to 180. These changes are supposed to make the search more refined and finding better results. In spite of the long time spent for algorithm simulating this scenario, the

objective was getting a benchmark to check how good are the other results.

As we can see, we have obtained similar results in terms of maintenance costs. Regard time-consumption the proposed methodology make much more iterations using just a fraction of the time of the others.

Different combinations of *s* and *d* lead a similar results in terms of costs, but as we can see, despite this variation in *s* and *d*, the value of *S* remains almost constant, around 278 units. The supposed reason for this is that, for some combinations of *s* and *d*, such that $S \approx 280$, there is a compensation in the sum of all costs. However, it is not true for any combination. For instance, $s = 50$, $d = 230$ $r = 3$ and $t = 1$ leads a very high *Ctt* = $ 408.21/day.

CONCLUSIONS

Combining metamodeling and heuristics into search algorithms as we proposed in this paper is viable and it leads competitive results in terms of the output quality. Regarding the time-consumption, the proposed methodology was more efficient than the others options we have used in this study. We showed that the development of algorithms to proceeds a simulation-optimization analysis may be quite simple and easy to implement and they could be seen as an alternative to the use of commercial packages.

REFERENCES

Amaran, S., N. V. Sahinidis, B Sharda and S. J. Bury. 2016. "Simulation optimization: a review of algorithms and applications". Annals of Operations Research 240 (1): 351–380.

Barton, R. R. 1992. "Metamodels for Simulation input-output relations". In Proceedings of the 1992 Winter Simulation Conference edited by J. J. Swain, D. Goldsman, R. C. Crain, and J. R. Wilson. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Barton, R. R. 2015. "Tutorial: Simulation Metamodeling". In Proceedings of the 2015 Winter Simulation Conference edited by L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, and M. D. Rossetti, eds. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Biles, W. E., J. P. C. Kleijnen, W. C. M. van Beers and I. van Nieuwenhuyse. 2007. "Kriging Metamodeling in constrained simulation optimization: an explorative study". In Proceedings of the 2007 Winter Simulation Conference edited by S. G. Henderson, B. Biller, M.-H. Hsieh, J. Shortle, J. D. Tew, and R. R. Barton, eds. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Can, B., C. Heavey. 2012. "A comparison of genetic programming and artificial neural networks in metamodeling of discrete-event simulation models". Computers & Operations Research 39: 424–436.

Chen E. J., M. Li. 2014. "Design of experiments for interpolation-based metamodels". Simulation Modelling Practice and Theory 44: 14–25.

Fu M. C. 2002. "Optimization for Simulation: Theory vs. Practice", INFORMS Journal on Computing 14 (3): 192–215.

Law, A. M. 2007. Simulation modeling and analysis, fourth edition, McGraw-Hill, Boston, MA. p. 48-58.

Li, Y.F., S.H. Ng, M. Xie, T.N. Goh. 2010. "A systematic comparison of metamodeling techniques for simulation Optimization in Decision Support Systems". Applied Soft Computing 10: 1257–1273.

Kirkpatrick, S., C. D. Gelatt Jr, and M. P. Vecchi. 1983. "Optimization by Simulated Annealing". Science 220 (4598): 671-680.

Kleijnen, J. P.C. 2005. "An overview of the design and analysis of simulation experiments for sensitivity analysis", European Journal of Operational Research 164: 287–300.

Mckay, M. D., R. J. Beckman and W. J. Conover. 1979. "A comparison of three methods for selecting values of input variables in the analysis of output from a computer code". Technometrics 21 (2): 239-245.

Zakerifar, M., W. E. Biles, G. W. Evans. 2011. "Kriging metamodeling in multiple-objective simulation optimization". Simulation: Transactions of the Society for Modeling and Simulation International 87 (10): 843–856.

AUTHOR BIOGRAPHIES

LUIZ RICARDO PINTO is Associate Professor at Universidade Federal de Minas Gerais, Brazil. His research interests are modeling and simulation analysis and applications. He holds a Ph.D. in Production Engineering / Operational Research from Universidade Federal do Rio de Janeiro (COPPE/UFRJ, 1999), Brazil. He was on sabbatical at Brunel University, UK (2011/2012). His email address is luiz@dep.ufmg.br.

JÚLIA COBUCCI MORAIS is undergraduate in Production Engineering at Universidade Federal de Minas Gerais, Brazil. Her research interests are operational research, simulation analysis and applications. Her email address is juliacobucci14@gmail.com.

GABRIELA MARTINS NUNES is undergraduate in Production Engineering at Universidade Federal de Minas Gerais, Brazil. Her research interests are operational research, simulation analysis and logistics. Her email address is gabrielamprod@gmail.com.

JOÃO FLÁVIO DE FREITAS ALMEIDA is Adjunct Professor at Universidade Federal de Minas Gerais, Brazil. His research interests are optimization and simulation modeling and applications. He holds a Ph.D. in Operational Research from Universidade Federal de Minas Gerais (DEP/UFMG, 2015), Brazil. His email address is joao.flavio@dep.ufmg.br.

LOGISTICS SIMULATION

SIMULATION OF LOGISTICS FOR CONSTRUCTION MANAGEMENT

Nikolaus Furian
Dietmar Neubacher
Siegfried Vössner
Graz University of Technology
Department of Engineering and Business Informatics
Email: nikolaus.furian@tugraz.at

Philip Santner
Michael O'Sullivan
Cameron Walker
University of Auckland
Department of Engineering Science
Email: michael.osullivan@aucklanduni.ac.nz

KEYWORDS
Discrete Event Simulation, Supply Chain Management, Construction Management

ABSTRACT

The construction sector is faced with increasing competition. Thus, efficiency and agility get more and more important. The first step in optimizing a system is to get a better understanding of the activities and their interconnections. One possibility in order to gather this kind of information is to build a simulation model. This paper describes the procedure of developing an entire simulation study of construction site supply chain logistics. A high degree of uncertainty is in the nature of every construction project. Permanent unpredictable changes of environmental factors (such as staff absenteeism, weather or breakdowns) complicate the planning phase tremendously. Simulation has been proven to be a suitable tool to address the challenges arising from uncertainty. The investigated construction site laid a new storm water pipe system in Auckland's central business district in New Zealand. The aims of the study were to simulate the logistics processes of the construction project (pipe delivery and soil removal) as well as to analyze its behavior and to work out potential improvements.

INTRODUCTION

In this paper we present a simulation study to improve the logistics for a construction project with the aim to build a new underground rail line in downtown Auckland, New Zealand. During the course of this project storm water pipes have to be replaced. The simulation model outlined in this paper focuses on the supply chain of pipe delivery and soil removal for this particular replacement. Thereby, an optimal number of trucks, as well as an optimal truck ordering policy are identified to ensure delays of the critical pipe jacking process are minimized. The work presented is based on the thesis of one of the authors (Santner 2017).

In recent years the research effort on construction site simulation has increased massively. The objective of those studies is to develop meaningful simulations in order to identify critical processes. This leads to the development of models that encompass multiple infrastructure configurations with different subprocesses, which is contrary to the simplification ethos which is central to simulation. A framework addressing this challenge was published by Voigtmann and Bargstdt (2010). For a general overview on the utilization of simulation in construction site management the reader is referred to AbouRizk (2010). Further, the different layers of construction supply chain from the source to the end customer can be summarized by: raw materials/component suppliers; labor market; and equipment manufacturer Cox et al. (2006). Of particular importance for this paper are the materials supply chains.

The paper is outlined as follows: the next section provides a description of the conceptual model of the simulation study and outlines the data acquisition process; the computer simulation, including scenarios and results, is described in the subsequent section; the paper concludes with final remarks and suggestions for further research.

THE CONCEPTUAL MODEL

In this section the conceptual model of the simulation is outlined. Conceptual modeling plays a crucial role within every simulation study as the content and purpose of the model is defined and documented. The conceptual model of the problem at hand was developed using the HCCM (Hierarchical Conceptual Modeling Framework), see Furian et al. (2015), and consists of the following building blocks: a general problem description; definition of objectives of the study; input and output factors; and the content of the model. Each building block is outlined in detail in the following sections.

Problem Description

This section provides an informal description of the construction site's logistic problem at hand. Therefore, first the process of pipe jacking is described briefly, followed by an illustration of the entire system.

Pipe Jacking
Pipe-jacking is a trench-less method for the installation of underground pipes. The underground area consists of three main parts; the shaft with hydraulic jacks and the pipe loading area, the inserted pipes with the soil transportation system (e.g. train), the mechanical excavation machine and the tunnel boring machine (TBM) at the front end of the tunnel. The main storage area for pipes and excavated soil is usually located on the surface.
Every pipe jacking cycle starts with the loading process of the pipe. A crane places the pipes from the surface down to the underground loading area. There the pipe gets connected with pipes that have already been inserted. Subsequently, hydraulic jacks exert force on the pipe and push the entire tunnel system into the ground. The excavation machine is operating during this process and produces excavated soil. After the pipe is fully inserted, the excavated soil is moved through the tunnel to the underground loading area. A crane conveys the tray with soil (muck skip) to the storage area where it gets emptied. This cyclic process is repeated until the desired tunnel length is reached.

Logistics Process
The New Zealand organization Auckland Transport is improving the public transport situation by building an underground rail line within the central business district of Auckland. As part of this project, storm water pipes have to be replaced by the pipe jacking method. The construction site is located in the central business district of Auckland. Thus, the site experiences high traffic density and volume as well as restricted space for storage, loading, and waiting areas.
The material supply chain problem consists of two main activities, pipe supply and soil removal. Pipe trucks deliver pipes from the supplier to the construction site. Simultaneously, soil needs to be removed. Dump trucks transport soil from the construction site to different disposal sites. Depending on the advancement of the tunneling process, the required number of daily trucks and trips may vary.
Figure 1 displays the layout of the construction site. It is structured into four main areas: pipe jacking area; storage area (pipe storage and the muck pit); loading/unloading area (where an excavator loads a truck from the muck skip or the gantry crane unloads pipes from a truck); and waiting area.

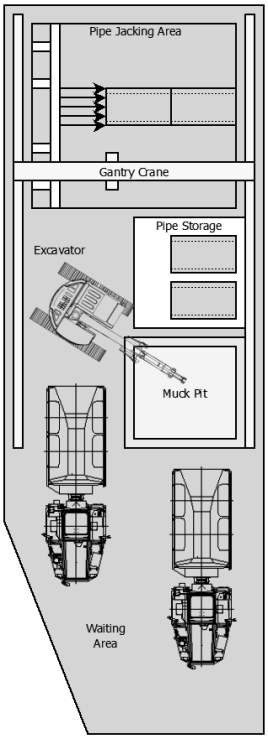


Figure 1: Layout Construction Site

Simulation Objectives

The objectives of the simulation study can be classified into three types: organizational objectives include the design and evaluation of current processes in order to avoid any interruption of the pipe jacking process due to surface problems (transport and logistics); general objectives require the simulation study to be carried out within 4 months, including a visualization model of the system; last, the modeling objectives state that different truck ordering policies for pipe delivery and soil removal are investigated.

Defining Inputs and Output Responses

Experimental factors are used to simulate different scenarios. Factors used included: the capacity of the muck pit; target value for pipes, i.e. the storage threshold that triggers a pipe order; the number of pipe and dump trucks; and travel times of trucks.

Outputs are measures that enable the assessment of whether the objectives of the simulation study have bee satisfied. Outputs collected during experiments include: the mean waiting time for trucks; the total time of delay caused by surface problems; and the maximum soil level.

Model Content

The next step within the process of conceptual modeling is to define the content of the model. Therefore, the structure and entities of the model, the individual and the system behavior have to be defined and documented. Furthermore, simplifications and assumptions have to be reported.

Model Structure

The entities included in the model are: trucks (dump trucks and pipe trucks); the unloading area; the pipe jacking area (representing the behavior under the surface); and traffic (representing delays for tracks based on traffic data). The gantry crane has been deliberately excluded and is represented within the behavior of the unloading area. Further, neither dumping grounds for soil nor the pipe supplier have been explicitly included but are represented by corresponding loading and unloading activities.

Individual Behavior

The individual behavior of the model is best described by the processes and activities of different entities within the model. Included activities can be summarized by: waiting activities; soil loading (at the construction site) and unloading (at the dumping ground(s)); pipe loading (at the pipe supplier) and unloading (at the construction site) and unloading (at dump ground); driving of trucks; and the pipe jacking itself which represents all activities below the surface. An overview of all processes in the model is given by figure 2.

System Behavior

The model control is a main feature of the HCCM framework. Control units can be a set of strategies or rules. Their task is to control the behavior of the defined entities. In the real world situation three areas were identified, at which control mechanisms were required: the PipeJacking Control Unit (deciding whether the pipe jacking activity has to be stopped due to problems with pipe delivery or soil removal), the Truck Access Control Unit (regulating the access of trucks), and the Truck Order Control Unit (defining ordering policies).

Each control unit is described by a set of rules. For instance, at the construction site there is a specifically defined policy, how trucks are granted access. Pipe deliveries have priority one, this means other deliveries have to wait until the pipe unloading activity is finished. One exception is the case in which the capacity of the muck pit has reached its maximum. In order to avoid delays of the pipe jacking process, the soil needs to be removed immediately.

Simplifications and Assumptions

Due to the nature of any real world problem, it is not possible to gather all information and data of a system. Some of the major simplifications made are: no traffic breakdowns of trucks; pipe jacking and unloading do not influence each other (exclusion of gantry crane); trucks can be ordered any time from the same operator; no delays at dumping ground and pipe supplier; no refueling of trucks necessary; pipe storage below the surface is not considered.

Data Acquisition and Validation

One of the biggest challenges of this simulation study was not only gathering information and data from the system, but rather to interpret and integrate it into the computer simulation. Unfortunately, some data was not available, because some processes had not been investigated yet by the company. The most critical steps to gather sufficient and accurate data were the collection of travel time data and data on breakdowns of the pipe jacking machinery.

Travel Times

Due to the heavy city traffic around the construction site, the trucks are faced with highly volatile traffic. Thus travel times vary significantly, depending on the weekday and time of the day. After the first round of interviews with the construction site's managers, it became clear that the necessary information was not available. The idea to track the trucks via GPS (Global Positioning System) was quickly discarded because of technical effort required. Hence, a program was developed, which collects data from Google Maps Directions API. With this tool it was possible to gather quarter hourly predictive travel times, based on historical data.

Breakdowns

Another main challange in the data acquisition was the integration of breakdown rates for the pipe jacking machinery in the model. The breakdown rate was given as hours per 12-hour shift for each component at the site. With this kind of data it was not possible to determine when and how long certain breakdowns occurred.They can be structured into six main breakdowns; Mechanical Fault (M), Electrical Fault (E), Water System (W), Surveyor (S), Ventilation System (V) and Others (O).

The challenge was to find a way to integrate the breakdown data into the simulation software. The chosen approach was to split up the breakdown data into two sets; long breakdowns ($>$ 12 hours) and short breakdowns ($<$ 12 hours). A re-sampling approach, called bootstrapping, was used in the simulation to generate long and short breakdowns that are based on real world data.

COMPUTER SIMULATION

In order to carry out experiments, the conceptual model was implemented with the JaamSim simulation software package (JaamSim 2016). JaamSim provides tools to not only simulate scenarios but also visualize simulation runs, which contributed to the verification and validation of the model with stakeholders of the construction company.

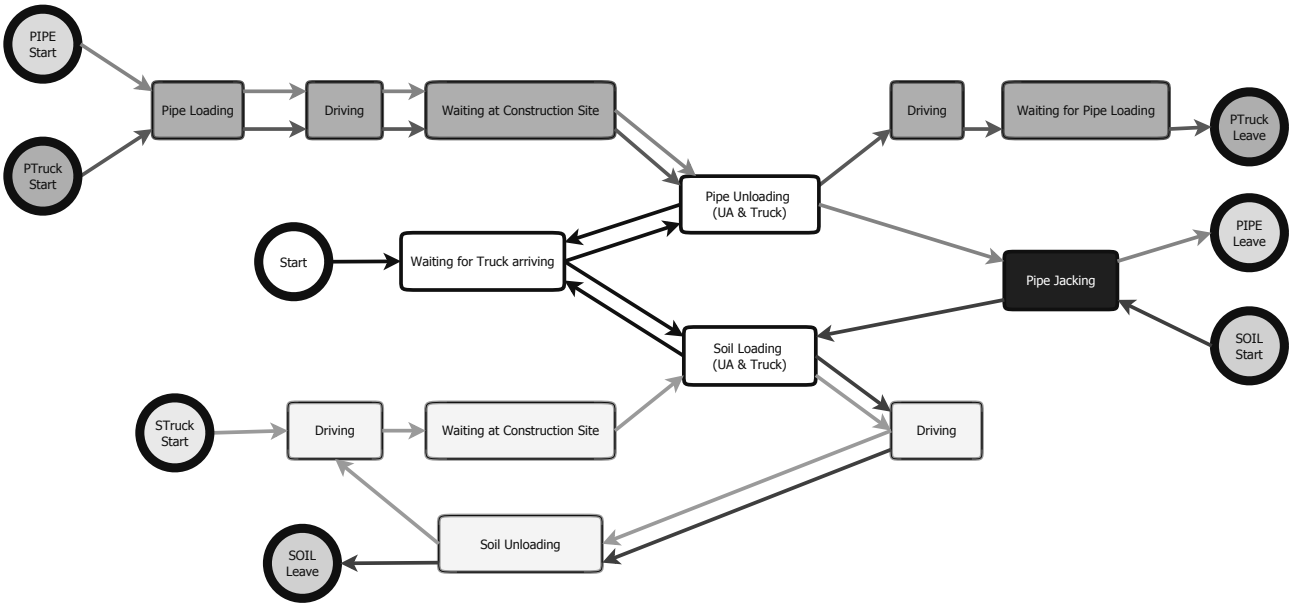


Figure 2: Entity Flow

Scenarios

One of the tasks was to define and simulate different possible order policies for the pipe and dump trucks, in order to see which one delivered the best results. The project team decided to integrate three possible scenarios: 5pm order policy; 7am order policy; and continuous order policy. For the first two policies the number of trips for the following or current day are defined at corresponding times depending on soil and pipe levels. While the 5pm policy neglects possible overnight production of soil and consumption of pipes it allows the truck company to plan trips in advance. The continuous order policy is similar to the JIT (Just in Time) strategy. A trip of a truck is ordered every time one truck load of soil ($7m^2$) is produced or one load of pipes (2 pipes) is consumed.

Results

This section presents the results and evaluation of the study. Each scenario was run 20 times in order to get statistically significant results. Around the construction site space is limited; the waiting area only provides space for two trucks. A third truck in the queue would have to look for a parking area near the construction site or would have to park at the roadside. Thus, cases with queue length of three or more must be avoided. Closer inspection of the construction site showed that the number of dump trucks for each day varies between zero and four. During the entire project, only one pipe truck was used to deliver pipes. However, in the simulation a second available pipe truck was added in order to see the impact of varying truck configurations to the system. For this purpose all possible configuration of available trucks, with $x_d \in \{1, 2, 3, 4\}$ dump trucks and $x_p \in \{1, 2\}$ pipe trucks, were simulated.
As there is restricted space within the construction site, only one truck is able to get loaded or unloaded at a time; arriving trucks have to wait in the waiting area in the meantime. This time is lost and causes delays to the supply chain. Hence, one question is which pipe and dump truck configuration causes the least total waiting time. Figure 3 shows the average total waiting time for trucks.

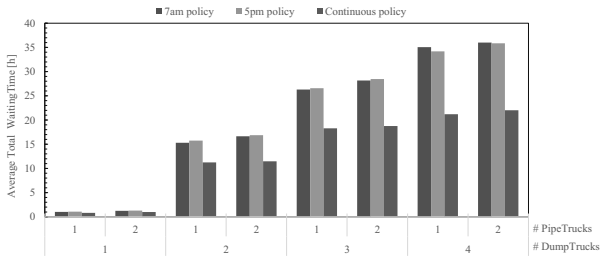


Figure 3: Average Total Waiting Time

The results for this scenario show, as one might expect, that the average total waiting time for trucks is lowest with a low number of total trucks. However, further simulation analyses have to be made in order to show whether one dump truck is able to remove soil from the construction site fast enough without causing any delay to the pipe jacking process.

Another conclusion can be drawn. The continuous order policy displays significantly better results in all cases. This can be explained by a more homogeneous distribution of truck arrivals over the day. With 7am and 5pm policy a bottleneck situation may occur especially in the morning which causes a longer than average waiting time and inconvenient queuing situations.
The objective of the next scenario was to show the behavior of the soil level of the muck pit with varying dump trucks and different order policies. The pipe jacking process needs to be stopped, if the capacity of the muck pit is reached. Thus, the focus of this simulation scenario is on cases where the soil level reaches a maximum (100m2).
Figure 4 displays the average maximum soil level of each scenario. The graph shows that the 7am order policy delivered better results than ordering the previous day. The lowest maximum soil level could be reached with the continuous order policy. However, it also shows that the maximum soil level does not change significantly with varying number of available dump trucks. In other words, an increasing the number of available dump trucks does not improve the situation at the muck pit.

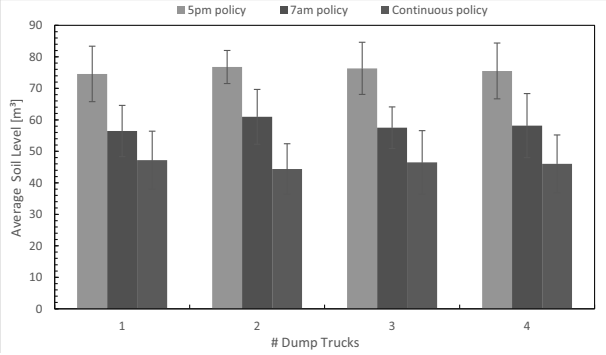


Figure 4: Average Maximum Soil Level

The next question to be answered by the simulation was the pipe delivery process. The key factor for assessing the pipe supply chain was the total pipe jacking delay time caused by a low pipe storage level. The analysis of the delay times is displayed in figure 5. Obviously, there are no problems caused by the pipe supply chain in the continuous and 7am order policy. Only the policy with a truck order at 5pm causes delays to the system. This result can be traced back to the fact that, with 5pm order policies, over-nigh production is not taken into consideration. More pipe trucks only marginally improve the situation.

CONCLUSION AND FURTHER RESEARCH

In this paper we presented the conceptual model of the construction logistics of a pipe-jacking construction site

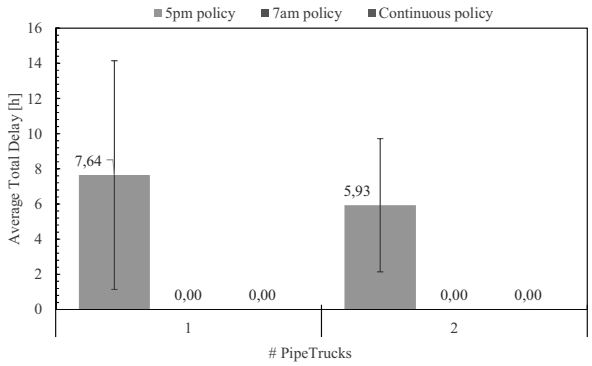


Figure 5: Delays of the pipe jacking process, caused by low pipe level

in downtown Auckland, New Zealand. Based on a conceptual model following the HCCM framework a computer simulation was built. Three different truck ordering policies were simulated and evaluated. The best results were achieved with the continuous order policy. However, this conclusion does not necessarily lead to a rejection of the 7am and 5pm order policy. The continuous order policy displays a very irregular order interval. This requires a very flexible truck driver (internal) or transport company (external). The next stage would be to coordinate the possible order policies with the transport companies. It needs to be evaluated whether irregular orders over one day (continuous order policy) are more favorable compared to fixed defined trips which may cause a higher truck waiting time especially in the morning (7am and 5pm order policy). The integration of truck roster (morning and evening shifts) in the simulation model was left for further research.
Further, it was concluded that the number of pipe trucks and dump trucks do not have significant impact on the performance of the construction supply chain.

ACKNOWLEDGMENTS

A major part of the paper is based on the master thesis Santner (2017).

REFERENCES

AbouRizk S., 2010. *Role of Simulation in Construction Engineering and Management. Journal of Construction Engineering and Management-Asce*, 136, no. 10, 1140–1153.

Cox A.; Ireland P.; and Townsend M., 2006. *Managing in Construction Supply Chains and Markets*. Report, Thomas Telford.

Furian N.; O'Sullivan M.; Walker C.; Vossner S.; and Neubacher D., 2015. *A conceptual modeling frame-*

work for discrete event simulation using hierarchical control structures. Simul Model Pract Theory, 56, 82–96.

JaamSim D.T., 2016. *JaamSim: Discrete-Event Simulation Software.*

Santner P., 2017. *Simulation of Logistics for Construction Management.* Master's thesis, University of Technology Graz.

Voigtmann J. and Bargstdt H., 2010. *Simulation von Logistikstrategien im Bauwesen. KIT Scientific Publishing 2010.*

COOPERATIVE DECISION MAKING MODELING IN TRANSPORTATION LOGISTICS DISPATCHING SYSTEM

Anton Ivaschenko
Samara National Research University
34 Moskovskoye shosse,
443086 Samara, Russia
E-mail: anton.ivashenko@gmail.com

Ilya Syusin
Magenta Technology
17c Curzon Street,
W1J 5HU, London, UK
E-mail: ilya.syusin@gmail.com

Pavel Sitnikov
ITMO University
14, lit. A, Birzhevaya liniya,
199034, Saint Petersburg,
Russia
E-mail: sitnikov@o-code.ru

KEYWORDS

Intermediary intelligent services, Smart logistics, Information technologies; Software platforms; Business processes modeling.

ABSTRACT

Modern business processes of dispatching in transportation logistics are primarily based on IT operational platforms with a dynamic real-time scheduling that enables you to enhance revenues whilst reducing operational costs. Considering the specific business requirements its implementation to practice can differ for various problem domains. Therefore there should be proposed an extension for business processes modeling notation that provides analysts an opportunity to formalize cooperative negotiation as a sub process of decision making in transportation logistics. In this paper there is proposed a new concept for Transportation Intermediary Service Platform (TISP) modeling. The examples of successful TISP implementation in practice are given for the distribution of products/services among fixed number of consumers and pickup and delivery service for unscheduled customers.

INTRODUCTION

In order to increase competitive performance most of modern companies in transportation logistics aim at combination of a wide variety of integrated services for their customers. Such an approach allows attracting new customers and providing transportation and courier services at reasonable price. For example, taxi companies grant courier services, and big transportation companies integrate both warehousing and distribution services.

At the same time for a number of transportation companies, it remains beneficial to stay small and flexible. This strategy helps them to reduce extra charges and adapt for a particular customer individually. An increasing group of customers prefers to get direct access to actual executors targeting high transparence and reliability of services.

The combination of these trends requires new approaches for smart business processes modeling based on modern information technologies. Modern trends in this area aim at virtualization of multiple transportation logistics service providers and customers' interaction in integrated information space, provided by specialized software solutions available on Internet. In this paper, we propose a new concept of Transportation Intermediary Services Platform (TISP) used to develop intelligent software solutions for transportation logistics.

Intelligent services platform is implemented as a software solution that provides virtual decision making points for smart transportation enterprises that can use them to compete and cooperate in integrated information space. The examples that illustrate some applications of TISP in real problem domain are implemented on the basis of Maxoptra platform, powered by Magenta Technology.

STATE OF THE ART

The concept of TISP is based on the modern principles of distributed simulation and decision-making support powered by multi-agent technology (Wooldridge 2002). The virtual world of negotiating service providers and customers should be treated as a complex network of continuously running and co-evolving intelligent agents. Such solutions are based on holons paradigm and bio-inspired approach (Leitao 2009), which 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) (Gorodetskii 2012).

Traditionally intermediary layers of Internet services have been considered from technical prospective (Hickson, 2008, Machiraju 2002) exploring various architectures of an overlay for federated service management, or web services management network. This concept relies on a network of communicating service intermediaries, each such intermediary being a proxy positioned between the service and the outside world. At the same time the benefits of Internet services implementation are studied as part of the e-Government paradigm (Layfield 2014).

As for the human beings represented by actors or agents, intermediary services should consider a combination of human and time factors. Interaction of customers and service providers powered by intermediary services generate and can be characterized by a big number of events that form Big Data and require modern technologies for its analysis (Bessis 2014).

Business processes modeling notation that supports such interaction should be flexible and dependent on unique customer requirements. This makes it reasonable to implement 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).

The idea of ISP was motivated by recent developments in transportation industry (Ivaschenko 2014). This proposed approach is close to 5PL (Fifth Party Logistics) concept, which is based on implementation of a number of services for customers and transportation companies provided by a specially designed software platform. 5PL platform is open for new transportation companies and even drivers and helps them negotiate with customers in integrated information space. 5PL provider owns no transportation resources itself but makes available a special service able to link suppliers and buyers. This service is based on the IT infrastructure, which plays the general role in 5PL business.

Customer representatives, transport managers, shippers, carriers, and even drivers become users of a certain IT platform. The purpose of this platform is to allocate incoming orders to appropriate resources, consolidate them improving consolidation and reducing idle time and generate efficient schedules for drivers and vehicles. Such service becomes attractive for small transportation companies and allows outsourcing dispatching functions for large logistics operators.

Still to ensure high efficiency of 5PL and ISP solutions both for customers and executors in terms of time and costs there is a request to implement modern technologies of business processes management based on decentralized architectures, distributed intelligence and multi-agent technology. This happens because of the increasing number of decision makers, high uncertainty and dynamics of changes, and flexibility of decision-making logic.

TISP CONCEPT SOLUTION VISION

The basic feature of the stated problem is a necessity to formalize the influence of human and time factors over the process of decision making. Both customers and transportation service providers possess independent behavior and cannot be managed by direct instructions. From the other side, in case the TISP provides durable solutions the users will trust it and wait for a certain period giving the system an opportunity to generate and compare separate options and analyze the influence over the whole network.

TISP solution is presented at Fig. 1. It introduces optimization functionality: the system starts helping its users to find the best combination of services according to their requirements. TISP has two basic features that are provided for customers and service providers: decision-making support (formalized as a part of business process using decision-making points, DMP) and allocation (scheduling) that is a regular part of transportation management software. DMP is a virtual platform for negotiating where multiple service providers can cooperate (by integrating their services) or compete to develop the best possible set of service options for a specific customer.
Therefore, by means of DMP the Intelligent services platform provide virtual decision-making points for smart transportation enterprises that can use them to compete and cooperate in integrated information space. TISP solution can be implemented using multi-agent technology.

Agents are introduced to represent the customers and service providers in the integrated information space and can be triggered both for simulation of customers activity and for representing the real customers in the process of searching for the integrated services. The architecture of multi-agent scene is quite simple: there are introduced three types of agents for Customers, Service providers and Services.

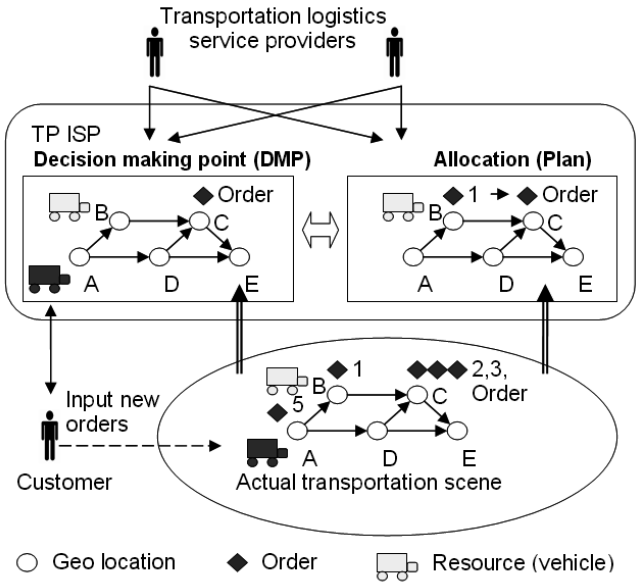


Figure 1: TISP Solution.

Customers and Service providers can interact according to their objectives and constraints and establish the links of cooperation according to which the Services are transmitted. The Customer agent tries to reduce the time and costs of the integrated service, and the Service Provider agent tries to increase its utilization delivering as many services as possible. The process of search can be presented as a sequence of local contracts between the customers and service providers.

DISTRIBUTION OF PRODUCTS/SERVICES AMONG FIXED NUMBER OF CONSUMERS

The proposed TISP solution was used to solve a number of actual problems in the domain of transportation logistics. It can be illustrated by two examples: automated distribution of products and services among the fixed number of

consumers that act in the role of customers and automated pick-up and delivery services.

The main challenge of the first problem is to deliver each item for relevant consumers (groups of consumers are predefined usually) and perform it in the most efficient way, which means to select the most appropriate route (transportation economy) and consider time windows of availability for each consumers (allows visit every route point once and make consumers informed and satisfied).

Groups of consumers are usually predefined and loyal, which means that they use the distribution service regularly. The example of products/services distribution process of products/services for 2 consumers is shown in Fig. 2.

The main steps of this process are:

1) Receive orders from all consumers. This is a very important step, because it allows creating the whole schedule before starting to execute it and minimize unexpected changes.
2) Create a schedule and trip-ticket. The result of this step is a real schedule and instructions that can be executed by employees.
3) Send trip-ticket to employees. This step executes an important communication function, because lets actors (employees) be informed about their tasks and consumers demands.
4) Start product/service delivery. This is a physical starting of products/service delivery.
5) Deliver product/service to the consumer 1. Product/service is delivered to the consumer and interaction between product/service provider and client can be started.
6) Accept product/service. The main result of this action – consumer has to be satisfied with the delivery service (product/service should be delivered in time and with acceptable quality).
7) Repeat steps 5, 6 for each consumer.
8) Finish product/service delivery. This is a physical finishing of products/service delivery.

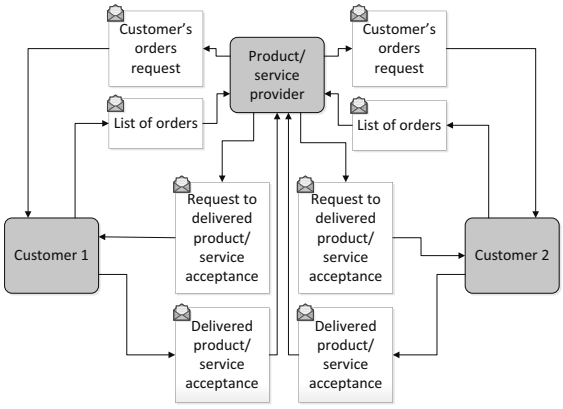


Figure 2: Products/Services Distribution Process for 2 Consumers.

This approach provides reaction to external events, such as "orders was modified" or "consumer is absent". It can be implemented using the direct interactions between product/service provider and consumer.
However, the problem is to make the proper decision as a reaction to unexpected event. This decision has to satisfy both product/service provider and customer.

Obviously, third-party actor should make decision. This actor needs the whole information about the process, demands and abilities of product/service provider and consumer. To provide such an opportunity we propose to introduce TISP as a third-party actor. TISP can perform specific actions (Decision-making point, DMP). DMP is a moment, when process can be modified according to different events and make decisions. In other parts, the process can be performed as defined, step by step.

The modified products/services distribution process is shown in Fig. 3.

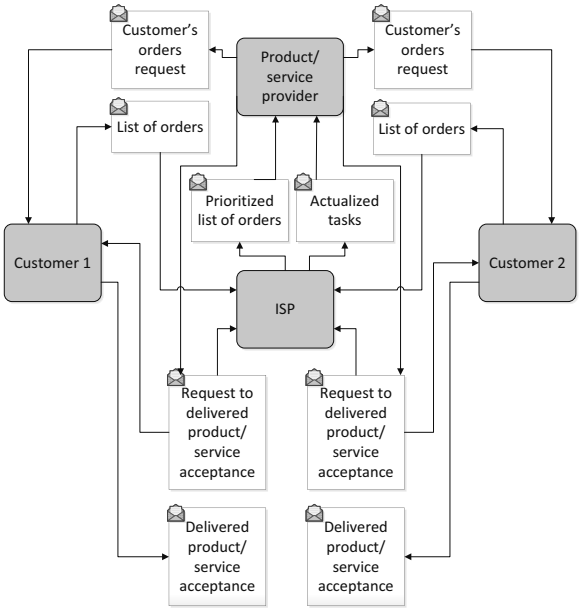


Figure 3: Modified Products/Services Distribution Process for 2 Consumers (Using TISP).

The description of the modified process is:

1) Receive orders from all consumers. This is a very important step, because it allows creating the whole schedule before starting to execute it and minimize unexpected changes.
2) **DMP: Prioritize Consumer's requests according to described criteria [3]-[5]. Define the most important and send them to product/service provider. This DMP allows making a decision "which orders will be processed" and creates tasks for delivery team.**
3) Create a schedule and trip-ticket. The result of this step is a real schedule and instructions, which can be executed by employees.

4) Send trip-ticket to employees. This step executes an important communication function, because lets actors (employees) be informed about their tasks and consumers demands.

5) Start product/service delivery. This is a physical starting of products/service delivery.

6) DMP: Receive responses from consumers (this action also includes receiving and processing all external events). Rearrange tasks in trip-ticket before deliver to the customer 1. As result we the new sub-process may be started (due to the new more important orders performing).

7) Deliver product/service to the consumer 1. Product/service is delivered to the consumer and interaction between product/service provider and client can be started.

8) Accept product/service. The main result of this action – consumer has to be satisfied with the delivery service (product/service should be delivered in time and with acceptable quality).

9) Repeat steps 6, 7, 8 for each consumer.

10) Finish product/service delivery. This is a physical finishing of products/service delivery.

PICKUP AND DELIVERY SERVICE FOR UNSCHEDULED CUSTOMERS

The challenge of this problem domain is to deliver product or service to a new customer that uses delivery service only once. Such customers require high service level: for instance, reminder messages via sms or call, very accurate delivery time etc. The main criterions of efficiency for this type of delivering are friendly communication with consumers and accurate delivery (on time, using customer's preferences).

Provided pickup and delivery service has to satisfy customer enough to use this service again and advise it to other customers. Fig. 4 shows the pickup and delivery process for 2 customers.

The steps of the pickup and delivery process:

1) One of the customers creates a new order. Usually this event is unexpected and product/service provider has to react appropriately.

2) Product/service provider checks order details, coordinate them (if necessary) and notices customer about successful order registration.

3) Finally customer accepts order details with product/service provider.

4) Product/service provider schedules delivery for customer's 1 order. Resource for order delivery is reserved.

5) Product/service provider may receive a new order from another customer (customer 2). In this case the new order must be also registered and scheduled; resource for delivery has to be reserved. But it may cause to rescheduling of customer's 1 order.

6) Reschedule delivery for customer 1. Cause to new iteration of negotiations with customer.

7) Notice customer 1 about the order change.

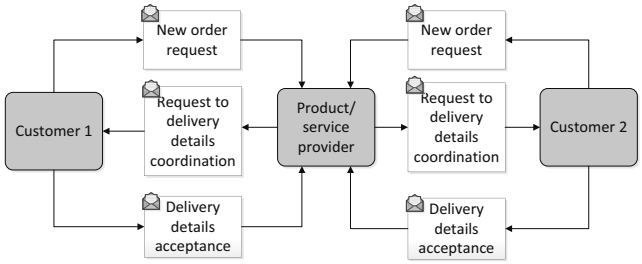


Figure 4: Pickup and Delivery Process for 2 Customers.

The problem is to handle new orders that are usually unexpected. Each new order may affect other orders, which are already scheduled, and reduce the quality level of delivery. Implementation of TISP allows increasing visibility of the process and reducing the number of iterations of negotiations between product/service providers and customers. The modified pickup and delivery process using TISP is shown in Fig. 5.

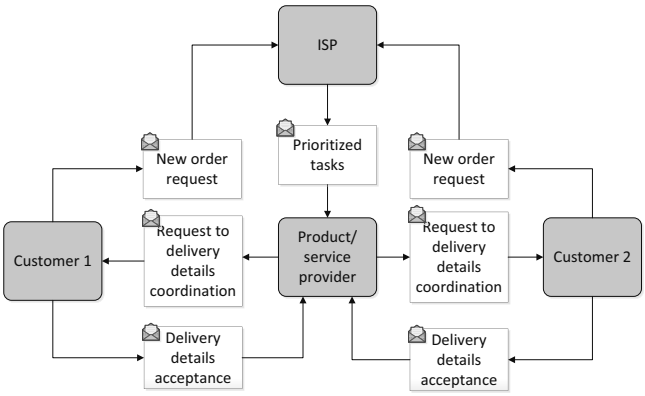


Figure 5: Modified Pickup and Delivery Process for 2 Customers (Using TISP).

The description of the modified process is:

1) One of the customers creates a new order. Usually this event is unexpected and product/service provider has to react appropriately.

2) Product/service provider checks order details, coordinate them (if necessary) and notice customer about registrations its order.

3) Finally customer accepts order details with product/service provider.

4) Product/service provider schedules delivery for customer's 1 order. Resource for order delivery is reserved.

5) Product/service provider may receive a new order from another customer (customer 2).

6) DMP: ISP analyses all new orders, prioritizes them according to criteria and creates a list of prioritized (important) orders. Orders from this list have to be performed first of all. Some orders may be exclude from the list, because of lack of resources for delivery.

7) Register new orders according to prioritized orders list.

8) Reschedule delivery for customer customers according to prioritized orders list. Cause to new iteration of negotiations with customer.

9) Notice customer 1 about the order change.

To illustrate implementation of TISP concept in practice there can be presented a Maxoptra platform powered by Magenta Technology. The system is distributed as SaaS-solution and provides monthly subscription. This is convenient for product/service providers, which are the main users of the system

Maxoptra has friendly and clear user interface, which allows to (see Fig. 6 – 7):

- monitor incoming orders and view order details;
- consider customer's particularity;
- schedule orders to the most convenient drivers;
- track driver's activity in the real time using GPS;
- notify dispatcher and drivers about the incoming orders.

Maxoptra is a useful tool for communications between product/service providers and customers, which allows to reduce interaction costs because of applying intermediary modeling concepts.

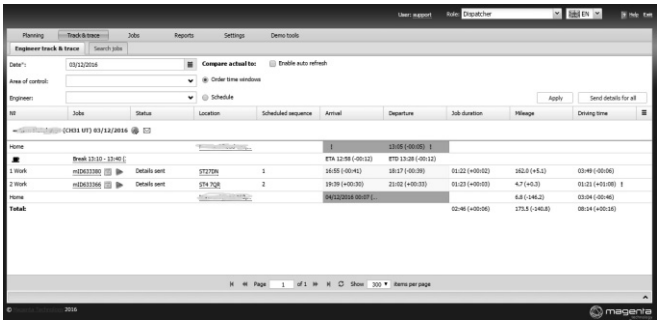


Figure 6: Maxoptra Tracking Screen.

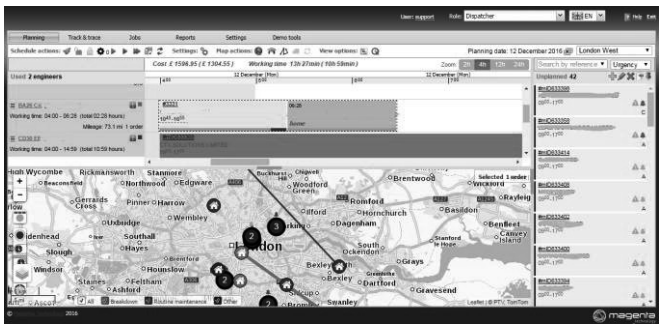


Figure 7: Maxoptra Dispatching.

CONCLUSION

The main problem of product/service providers in modern world is to make communication clear and accurate. On the one hand, provider needs to produce service or sell product with the highest price; on the other hand, customer wants to get service or good with the lowest price and the best quality.

Modeling the Transportation Intermediary Services Platform (TISP) allows developing IT solutions for smart transportation enterprises that can use them to compete and cooperate in integrated information space. The examples of successful TISP applications in practice and their implementation on the basis of Maxoptra engine illustrate the benefits of the proposed approach and can be recommended for various areas of transportation logistics.

REFERENCES

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

Fleischmann, A., Subject-oriented modeling and execution of multi-agent business processes / A. Fleischmann, U. Kannengiesser, W. Schmidt and C. Stary // Proc. IEEE/WIC/ACM International Conferences on Web Intelligence (WI) and Intelligent Agent Technology (IAT), USA, Atlanta, Georgia, 2013. – pp. 138 – 145

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

Hickson, A.; B. Wirth, and G. Morales. 2008. "Supply chain intermediaries study", University of Manitoba Transport Institute, 56 p.

Ivaschenko, A. 2014. "Multi-agent solution for business processes management of 5PL transportation provider". *Lecture Notes in Business Information Processing*, Vol. 170. 110-120.

Ivaschenko, A.; I. Syusin, and A. Fedosov. 2014. "Agent-based management for intermediary service provider". *Proc. of the 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT),* Warsaw, Poland, 428-432.

Layfield, C. 2014. "Opportunities in e-Government utilizing cloud computing: An EU perspective". *International Journal on Information Technologies and Security*, No. 2 (vol. 6), 3-26.

Leitao, P. 2009. "Holonic rationale and self-organization on design of complex evolvable systems", *HoloMAS 2009, LNAI 5696*, Springer-Verlag Berlin Heidelberg. 1-12.

Machiraju, V.; A. Sahai, and A. Moorsel. 2002. “Web services management network: an overlay network for federated service management”. Hewlett-Packard Laboratories, Hpl hp techreports, 234 p.

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

Comparison of Cost Performance of Fixed and Flexible Collection Strategy in Return Logistics Network

Di Zhang and Uwe Clausen
Fakultät Maschinenbau / Institut für Transportlogistik
Leonhard-Euler-Str. 2
D-44227 Dortmund, Germany
Email:{zhang@itl.tu-dortmund.de}

KEYWORDS

Returnable transport packaging, closed-loop network, green logistics, flexible collection strategy, operational simulation

ABSTRACT

To improve the collection efficiency in return logistics networks, the authors of this paper evaluate the serviceability of fixed and flexible collection strategies from the perspective of resource utilization. Collection frequency and total transportation cost are selected for comparison. The results show that the flexible collection strategy is more affordable than the fixed strategy. In general, I-Mode is superior to S-Mode, and small size trucks show a disadvantage in term of cost savings in less-than-truckload scenarios.

INTRODUCTION

As logistical systems evolve and become better organized, they create new opportunities for using packaging innovations to add value and minimize cost across a supply chain (Twede and Clarke 2004). Packaging has several objectives such as physical protection, barrier protection, containment or agglomeration, convenience, security and so on. StopWaste Partnership and RPCC (2007) through a cost comparison model to illustrate how it makes financial sense for the food manufacturers and the seafood distributors to switch to reusable transport packaging. Once-off packaging gradually loses its competitiveness and the returnable packaging is getting more and more popular.

With the broad spread of the concept of sustainable development, returnable articles turn to be a promising research topic in the modern society. Within such context, a different way of reflecting environmental concerns in logistics decisions is to manage the returned product flow and/or integrated both forward and reverse flows in the supply chains. This topic has been studied in details in the literature in recent years (Sheu et al 2005, Srivastava 2008, Salema et al. 2010, Ramos et al 2014).

Most of the articles figure out the reusable packaging system has an extensive application trends in some industries, but few of them made in-depth analysis of the distinguishing characteristics of once-off packaging and reusable packaging, or thought about whether the operational strategies of general logistics network are still applicable to return logistics networks. Based on the previous studies, this paper simulates return logistics network based on tailor-made bidirectional flexible service strategy, specifically for various collection scenarios of reverse channel. For the ease of understanding, the authors transform the conceptual resource utilization into comparable vehicle utilization, total cost, turnover rate and other indicators. Finally, an example of efficiency evaluation is carried out, which uses the total collection cost and trucks utilization rate as the key indicator and try to prove the flexible service strategy is more useful for return logistics network.

LITERATURE REVIEW

Twede and Clarke (2004) investigate the supply chain relationships that facilitate a reusable packaging system. They use two case studies to illustrate how a well-managed supply chain facilitates reusable packaging. Vasiliauskas and Bazaras (2006) analyze the current state of the world container market, describe problems related to empty container trips. Yun et al. (2011) considered an inventory control problem of empty containers in an inland transportation system with a simple policy to reposition containers from other hubs. Dang et al. (2012) deal with the problem of positioning empty containers in a port area with multiple depots. Furio et al. (2013) define two mathematical models (based on two different container movement patterns, i.e. with and without street-turns) to optimize land empty container movements among shippers, consignees, terminals and depots, along with minimizing storage costs.

As the first attempt examines the elasticity of substitution between self-owned and leased containers, and in turn analyzes the selection behavior of containers for sipping lines, Wu and Lin (2015) indicate that holding self-owned containers presents a relatively stable pattern; daily costs of holding different types of containers determine the usage proportion. Some strategic measures are developed to harmonize the competition to improve the overall efficiency. Akyüz and Lee (2016) develop a decision tool for a liner

shipping company to deploy its fleet considering vessel speeds and to find routes for cargos with repositioning of empty containers and transit time constraints. Xie et al. (2017) study the empty container inventory sharing and coordination problem in intermodal transport composed of a liner firm and a rail firm.

Researches (McGinnis and Kohn 2002) address the roles of process strategy, market strategy and information strategy in achieving logistic effectiveness. Autry (2005) studies the impact of reverse logistics formalization by exploring the relationships between formalization and liberalized returns policies, reverse logistics capabilities, and performance. Chen and Bell (2009) examine cases when the quantity of returned product is a function of both the quantity sold and the price, and address the simultaneous determination of price and inventory replenishment when customers return product to the firm. Based on a set of partitioning formulation, Ramos et al. (2014) discuss tactical and operational planning decisions of reverse logistics systems while considering economic, environmental and social objectives together.

Li et al. (2017) investigate the coordination strategies among different parties in a three-echelon reverse supply chain based on complete information. With the rapid development of e-commerce technologies, Batarfi et al. (2017) study forward and reverse integrated supply chain system: a refund return policy, two selling strategies, a single-channel (a retail channel) strategy and a dual-channel (a retail channel and an online channel) strategy are discussed. Hosoda and Disney (2017) investigate the dynamics of a closed-loop supply chain with first-order auto-regressive (AR(1)) demand and return processes, then establish an optimal linear policy in our CLSC setting to minimize inventory costs. Their CLSC model is a periodic review backlog system with constant lead times facing stochastic demand and return processes.

The work presented in this paper differs from the related literature in the following distinct ways. First, to the best of the authors' knowledge, no study has investigated the resource-utilization issues in return logistics networks, or considers the utilization elements as the key efficiency indicator of evaluation system. Second, no study has examined the effect of adopting a flexible service strategy on the facility location, inventory decision and the cost performance in the return logistics network (RLN).

COLLECTION STRATEGIES SIMULATION

Collection strategy is an important but easily overlooked subarea in reverse logistics (RL) and closed-loop supply chain (CLSC) researches. Zhang (2016) discussed a fixed collection strategy and compared the waiting time of returnable transport items (RTI) and the average turnover rate in 4 scenarios which have different economic collection quantity (ECQ) standard. Based on the concept of resource utilization and fixed collection strategy, an improved flexible collection strategy is proposed and discussed by Zhang (2017). The flexible strategy shows an advantage in reducing collection waiting time and improving the service efficiency.

Five steps before simulation

The goal of the simulation is to compare the cost performance of fixed and flexible strategies. Transport frequency and total transport cost in reverse channel are chosen as comparative indicators. Five steps are processed before simulation.

Step 1: Demand data processing.

Before simulation, the demands in bi-direction channels are tracked and aggregated. The RTIs received by the end customer determine the total amount of returnable RTIs in the area.

Step 2: Divide the service area.

After the demand processing, the up- and downstream customers are assigned to different depots based on distance limitation. The locations of alternatives depot are provided by the logistics department of investigation companies.

Step 3: Filter the best layout in service area

When the service areas are delineated, the initial layout is calculated by the ***warehouse*** module of CPLEX software to determine the location of depot (only two can be opened in four alternative depots, two customer sets: one for 34 forward customers and another for 65 reverse customers. Transport duration between depots and customers, initial inventory of RTIs, service capacity of each depot). In order to make the results closer to reality, location model considers both forward and reverse demand and summarize bi-directional transport time aim to minimize the weighted transport time within the planning area.

Step 4: Collection conditions preset

After the network layout is determined, the used returnable articles could be collected when the returnable quantity meet the economic collection quantity (ECQ) baseline. The usage of different types of truck will also impact collection efficiency. Simulation conditions are preset as below.

Table.1 indicates 8 simulation scenarios. I-IV are set to a single ECQ baseline (Fixed strategy) respectively and V-VIII are set for multiple ECQ sets with different truck arrangements (Flex strategy). For transportation, 80 and 120 are conventional capacity options, 100 is a new one. The transport costs of three types of trucks are 800/1000/12000 RMB without the distance factor (more detailed and complex calculation function will be executed in whole network simulation). All three can't load containers more than their capacity limitation. The letter f after the slash represents full load and i means some capacity is wasted.

Table. 1 collection scenarios in reverse channel

	I	II	III	IV
Fixed ECQ	60	80	100	120
	V	**VI**	**VII**	**VIII**
Flex ECQ	M1	M2	M3	M4
Unit for RTI is a Piece, unit for waiting time is a Day Three type of trucks: 80, 100, 120 *load capacity*				

Nr.	ECQ	Applicable scope			
		120+	120-101	100-81	80-60
M1	60/80/100/120	120/f	120/i	80/f	80/i
M2	60/80/100/120	120/f	120/i	100/i	80/i
M3	60/80/100/120	120/f	100/f	100/i	80/i
M4	60/80/100/120	120/f	100/f	80/f	80/i

Step 5: Set different scenes.

Fig.1 indicates all variable condition sets in the network simulation. Reverse collection methods are shown in column 3 and 4: S and I denote the single and the integrated collection mode, Fix and Fle represent the different collection strategies: either the fixed or flexible ECQ standard is used in collection activities. These 12 scenarios are just examples of flexible service strategy simulations. When the competitive environment or industry trends have changed, new simulation scenes could be modified based on the changes of demands requirements. All simulation will be carried out with Python 2,7,11.

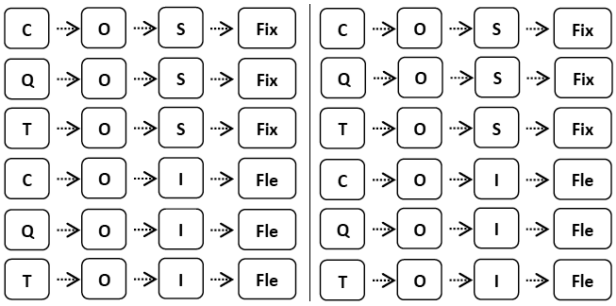


Fig.1. Simulation Scenarios of Flexible Strategy for Return Logistics Network

As the time-oriented indicators are already examined in Zhang (2016 and 2017), collection costs in different scenarios are calculated using equation (1):

$$TC = \sum_1^n (T_{fre} * tc_i) \quad (1)$$

$$tc_i = \min(T_{qua}|V_{FTL}, T_{qua}|V_{LTL}) \quad (2)$$

Equation (2) implies that the transportation cost could be compared and selected between FTL (full truck loading) and LTL (less-than-load) alternative mode according to T_{qua}(transport quantities).

The five steps before simulation prepare all parameters and set simulation conditions. The details can be adjusted based on practical collection operation.

Simulation results analysis

The truck utilization frequency in four fixed strategy scenarios (E60/E80/E100/E120) and one flexible strategy scenario (FlexiECQ) are compared in Fig.2. E60 scene accounts for nearly 30% of the total frequency of truck usage in 5 scenes, reaching the peak (about 34%) at Nr.59 customer. E80 scene takes about 20% before Nr.52 customer and reduces to 15% for remaining customer (except for Nr.61). The proportion of T-100 and T-120 size trucks used in E100 and E120 scenarios are stabilized at 15% respectively. In FlexiECQ scene, all three types of trucks are assigned tasks according to the amount of returnable packaging: T-80 takes 20%-25% of the total frequency of truck usage; T-100 and T-120 are occasionally used for a few customers. It is important to note that in FlexiECQ scene, the T-80 size truck is more commonly used for smaller size customers (nearly 25%) than large size customers(10%-20%).

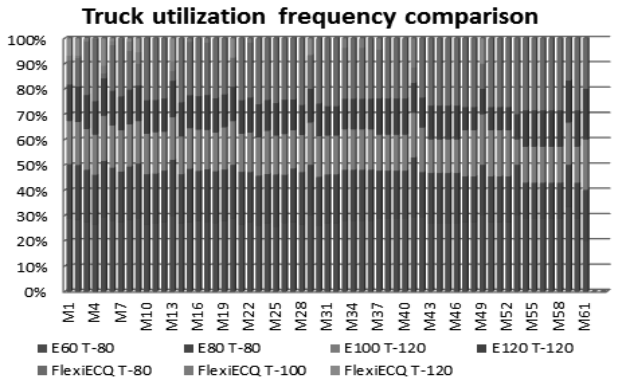


Fig.2. Truck Utilization Frequency Comparison in S Mode

As the value of total collection quantities in different scenarios are similar (Zhang 2017), the larger the loading capacity, the less the number of shipments. Ignoring the time element, Fig.3 shows the total collection cost in a variety of scenarios. The cost curve for E60 has the highest value for most customers; FlexiECQ is centered; and there is a clear gap between the values of E80/100/120 and E60 before customer Nr.31. Ignoring the time element, large-capacity truck cost less than small-capacity trucks. FlexiECQ shows the cost advantage compared with E60.

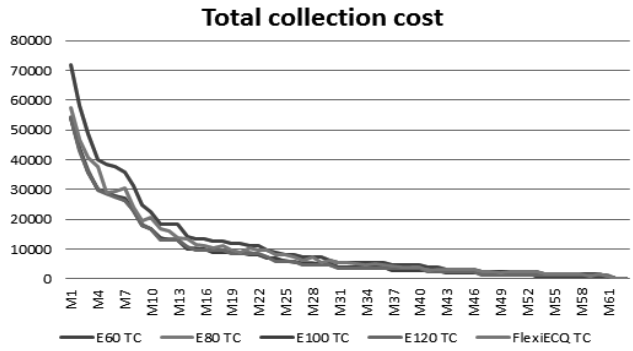


Fig.3. Total Collection Cost Comparison in S Mode

Beside separate collection mode, customers who are located within a certain distance or share a logistics transport line can be arranged/gathered using integrated collection (I) mode for the collection tasks. The detailed results of 11 customer groups are shown in Fig.4. Similar to the separately (S) mode, Fixed-E60 scene has the highest transportation frequency (28%-30%); the proportion of Fixed-80/100/120 are gradually reduced (20%/15%/10%). In most case, the FlexiECQ occupies more than 20% of the total truck utilization frequency. As the only exception, QD group has the maximum value of usage ratio of Fixed-60/80/100, which leads to the result that FlexiECQ has only accounted for about 15% of the total frequency. The reason for this phenomenon is that the customer completes the intensive production activities in a certain period. It may also be caused by seasonal changes/ raw material price fluctuations or some others reasons.

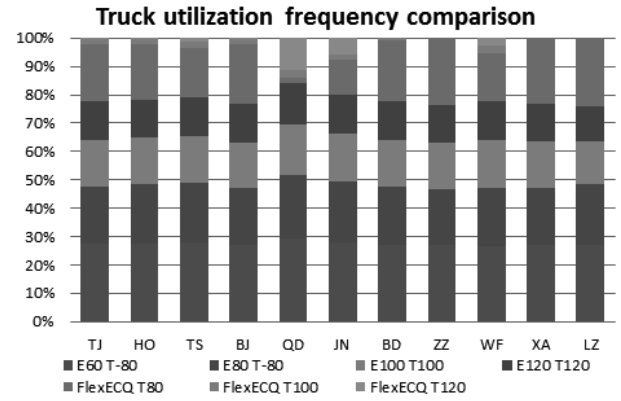


Fig.4. Truck Utilization Frequency Comparison in I Mode

Fig.5 compares the total collection cost in I mode. Under the premise that the total collection quantity is approximate, low truck capacity is accompanied by high transport frequency. Total cost in I mode follows an obvious rule: the value of Fixed-E60 and FlexiECQ are always higher than that of Fixed-E80/100/120, and the difference is gradually narrowing. Between Fixed-E60 and FlexiECQ, the Fixed-E60 costs more. Only in QD group is the CQ of FlexiECQ close to Fixed-80/100/120.

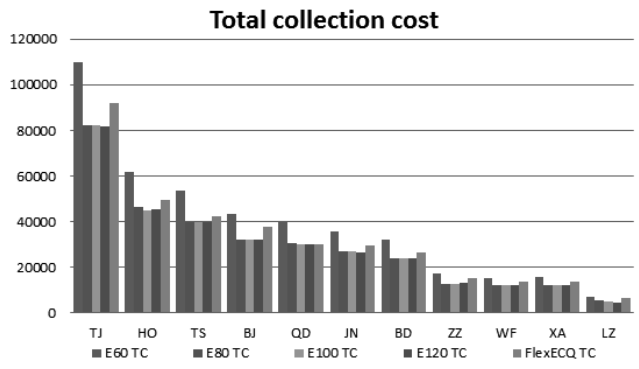


Fig.5. Total Collection Cost Comparison in I Mode

Fig.6 compares the truck utilization under truck schedule I-IV in FlexiECQ scene. Tasks are assigned to three types of trucks based on different criteria. It is obvious that T-80 type truck is more widely used. In TJ, truck T-100 is more popular than T-120 in schedule II and III; in DQ, the T-120 is more practical than T-80 and T-100. The integrated collection group is ranked in descending order according to the total amount of demand; the frequency of QD is even lower than the two groups behind him.

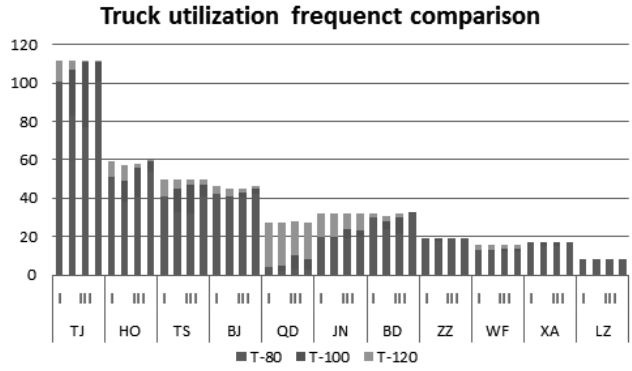


Fig.6. Truck Utilization Frequency Comparison in I Mode under Truck Schedule I-IV

Total collection cost of schedule I-IV is shown in Fig.7. Truck T-100 and T-120 are more frequently used in scenario II and III. It leads to a slight increase in the costs. For example, the usage frequency of TJ in schedule II and III is similar to I and IV, but the cost of II and III is obviously higher than I and IV, in which the truck T-100 is used more frequently. Group QD has the same situation. The truck utilization frequency of integrated group is much lower than the two customer group behind it, but the total collection cost in Fig.14 is closer to the latter two. The frequency of BD is almost equal with JN. Due to the more frequent use of truck T-80, the total cost of BD is lower than that of JN. It can be concluded that under schedule I-IV, more frequent use of small capacity trucks can reduce the total cost.

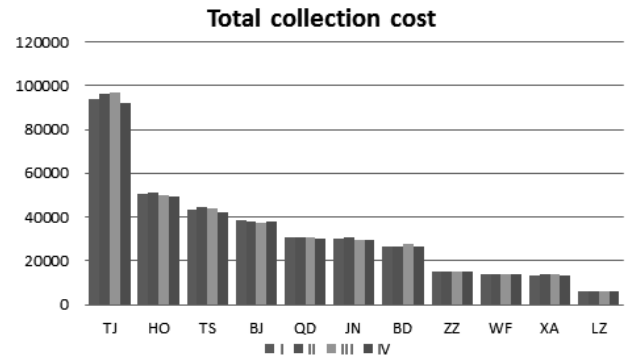


Fig.7. Total Collection Cost Comparison in I Mode under Truck Schedule I-IV

CONCLUSION

This paper compares and evaluates the economic indicators of fixed and flexible collection strategy in return logistics network. CPLEX and Python software have been used to solve collection efficiency problem. Simulation results show that the flexible service could significantly improve the network efficiency, shorten idle waiting time and save cost for the participants.

This conclusion can be regarded as a supplement to the previous findings: no matter if it is in S or I collection mode, the flexible collection strategy is better than the fixed. In general, I mode is superior to S mode, and small capacity trucks in different truck schedules have shown a disadvantage in term of cost savings.

Further researches could be focused on the whole network operation and the effect of different operation strategies on total profit.

REFERENCES

Autry, C.W., Formalization of reverse logistics programs: A strategy for managing liberalized returns, *Industrial marketing management*, 2005, Vol.34, 749-757

Akyüz, M.Hakan. and Lee, Chung-Yee, Servie type assignment and container routing with transit time constraints and empty container repositioning for liner shipping service networks, *Transportation Research Part B: Methodological*, 2016, Vol. 88, 46-71

Batarfi, R., Jaber, M.Y. and Alijazzar, S.M., A profit maximization for a reverse logistics dual-channel supply chain with a return policy, *Computers and industrial engineering*, 2017, Vol.106, 58-82

Bromiley, P. and Rau, D., Operations management and the resource based view: another view, *Journal of operations management*, 2016, Vol. 41, 95-106

Chen, J. and Bell, P.C., The impact of customer returns on pricing and order decisions, *European journal of operational research*, 2009, Vo.195, 280-295

Dekker, R., Bloemhof, J., Mallidis, I., Operations research for green logistics - an overview of aspects, issues, contributions and challenges. *European journal of operational research*, 2012, Vol.219, 671-679

Furio, S., Andres. C., Adenso-Diaz, Belarmino., Lozano. Sebastian, Optimization of empty container movements using street-turn: Application to Valencia hinterland, *Computers and Industrial Engineering*, 2013, Vol. 66, 909-917

Helfat, C.E. and Peteraf, M.A., The dynamic resource-based view: capability lifecycles, *Strategic management journal*, 2003, Vol.24, 997-1010

Hitt, M.A., Carnes, C.M. and Xu, K., A current view of resource based theory in operations management: a response to Bromiley and Rau, *Journal of operations management*, 2016, Vol.41, 107-109

Hosoda, T, and Disney, S. M., A unified theory of the dynamics of closedloop supply chains, *European Journal of Operational Research*, 2017, In Press.

Li, J., Wang, Z., Jiang, B. and Kim, Teabok., Coordination strategies in a three-echelon reverse supply chain for economic and social benefit, *Applied mathematical modelling,* 2017, Volume 49, 599-611

McGinnis, M.A. and Kohn, J.W., Logistics strategy-revisited, *Journal of business logistics*, 2002, Vol.23,1-18

Priem, R.L. and Butler, J.E., Is the resource-based 'view' a useful perspective for stratefic management research? *The academy of management review*, 2001, vol. 26, 22-40

Ramos, T.R.P., Gomes. M.I., Barbose-Povoa, A.P., Planning a sustainable reverse logistics system: Balancing costs with environmental and social concerns, *Omega*, 2014, Vol.48, 60-74

Rave Shankar, V.R., Analysis of interactions among the barriers of reverse logistics, *Technological forecasting and social change*, 2005, Vol.72, 1011-1029

Salema, MIG., Barbosa-Povoa, AP., Novais, AQ., Simultaneous design and planning of supply chains with reverse folws: a generic modelling framework, *European journal of operational research*, 2010, Vol.203, 336-349

Sheu, J.B., Chou, Y.H., Hu, C.C., An integrated logistics operational model for green-supply chain management. *Transportation research part E: Logistics and transportation review*, 2005, Vol.41, 287-313

Srivastava, S.K., Network design for reverse logistics. *Omega-International journal of management science*, 2008, Vol.36, 535-548

Twede, D abd Clark, R., Supply chain issues in reusable packaging, *Journal of Marketing Channels*, 2004, 7-26

Wu, W-M. and Lin, T-H., Selection behavior of the global container shipping industry for carrier-owned and leased containers, *Transport Policy*, 2015, Vol 37 , 11-19

Vasiliauskas, A.V. and Bazaras, D., Analysis of problems with containers as intermodal loading unit, *Transport and Telecommunication*, 2006, Vol.7, 232-239

Xie, YY., Liang, XY., Ma, LJ. And Yan, HM., Empty container management and coordination in intermodal transport, *European Journal of Operational Research*, 2017, Vol 257, 223-232

Zhang, D. and Clausen, U., Time-oridnted collection schedule simulation in closed-loop RTP network with multiple economic collection quantities and multiple production cycle, *30th European Simulation and Modelling Conference - ESM' 2016,* 440-447

Zhang, D and Clausen, U, Improving the resource utilization of reverse channel in return logistics network via flexible collection strategy, *XXII International Conference on Material Handling, Constructions and Logistics - MHCL 2017*

WEB REFERENCES

StopWaste Partnership and RPCC, A cost comparison model for reusable transport packaging, www.usereusables.com, 2007

AUTHOR BIOGRAPHIES

Di Zhang studied logistics network simulation and optimization at the TU Dortmund, Germany. Since 2012 she is PhD student in the institute of transport logistics. Her main research interests are sustainable logistics networks and their application for practical industries.

Uwe Clausen is the managing director of the Institute of Transport Logistics at the TU Dortmund University and director of the Fraunhofer-Institute for Material Flow and Logistics (IML). He is amongst others chairman of EffizienzCluster LogistikRuhr and a member of the Scientific Advisory Board of German Logistics Association (BVL).

EVALUATION IN TRANSPORT PLANNING: A COMPARISION BETWEEN DATA ENVELOPMENT ANALYSIS AND MULTI CRITERIA DECISION MAKING METHODS

Giuseppe Musolino
Corrado Rindone
Antonino Vitetta
DIIES
Università degli studi Mediterranea di Reggio Calabria
Feo di Vito, Reggio Calabria,
Italy
E-mail: corrado.rindone@unirc.it

KEYWORDS
Transportation, Decision-making, Discrete simulation, Model evaluation, DEA.

ABSTRACT

The success of transportation planning is affected, among other elements, by interaction among different stakeholders including decision makers, users and operators. Several evaluation methods are proposed in literature. They analyse effects of transportation planning, design and implementation processes.
In this paper an integrated evaluation process to support transportation planning is proposed. The proposed approach is a first tentative to incorporate in an integrated process and in a whole procedure transport simulation models, Multi Criteria Decision Making (MCDM) method and Data Envelopment Analysis (DEA). This, in to order to increase the stakeholder perception about effects of his/her choices towards some planned objectives. In particular, DEA and MCDM are experimented, as an integrated evaluation method inside the transportation planning process, and compared.

1. INTRODUCTION

This paper provides a research contribution in the wide topic concerning transportation planning process in a context where several actors (stakeholders) decide according to their own criteria. In this area, the specific problem of evaluation methods inside the planning process is studied. Planned actions produce effects, evaluated by stakeholders from different points of views. For this reason, in some cases conflict of interests could arise and they constitute a limitation in the follow-up phases of planning activities (design, realization).
(a) Transportation planning can be supported by simulation models and evaluation methods to measure and to compare a priori effects deriving from planning decisions (Cascetta 2009).
(b) Among evaluation methods, Multiple Criteria Decision Making (MCDM) supports decision makers in complex decision situations involving multiple actors with a plurality of points of view and objectives. These methodologies are applied for selecting non-dominated or the best scenario in relation to a set of defined criteria (Macharis et al. 2010).
(c) Among evaluation methods, Data Envelopment Analysis (DEA) is another interesting approach (Bouyssou D., 1999). DEA is a non-parametric method to measure relative efficiency of an homogeneous set of Decision Making Units (DMUs) and their use of multiple inputs to produce multiple outputs (Cooper et al., 2000).
More details of the state of the art about these methodologies are reported in section 2.
In the authors' knowledge, the following research questions are still open in transportation planning involving multi actors:
(a) integration of simulation models and MCDM methodologies inside the transport planning process, in order to prevent conflicts among multi stakeholders;
(b, c) use of DEA and MCDM methodologies as a multi criteria tool to support transportation planning.
The novelties of this paper consist on the proposition of:

- a process that integrates transport simulation models, MCDM and DEA methodologies in order to support transportation planning process characterized by multi actors and constrained resources (section 3.1);
- a procedure that implements the above process to obtain best performance scenarios (section 3.2).

After this introduction, the paper has four sections. Section 2 presents a state of the art concerning transportation planning process and DEA as a tool of MCDM. In section 3 a process and a procedure that integrate simulation models, MCDM and DEA are proposed. Section 4 illustrate the main elements of a prototypal application. The last section reports conclusive comments and future research perspectives.

2. STATE OF THE ART

(a) Transportation planning process drives decisions (i.e. realization – or not – of transport infrastructures and/or services) in order to reach objectives (i.e. sustainability) (López and Monzón, 2010; Praticò et al., 2015; Russo et al., 2016). However, objectives' importance could be not the same in relation to needs and values of different classes of stakeholders (Nik-Bakht and El-diraby, 2016). Public Engagement has been developed in recent years (OECD,

2009), in order to identify and to incorporate these different needs and values into the decision-making process. Different categories of stakeholders may be identified including institutions/authorities, transport operators (of infrastructures and services), transport services users, local communities and media (Bickerstaff et al. 2002; Cascetta and Pagliara, 2013). Different tools have been developed in order to facilitate the stakeholders engagement into the decision-making process,. In some cases, transport simulation models are essential in order to support ex-ante evaluations (Cascetta et al. 2015; Caggiani et al., 2017).

(b) Outputs of simulation models are used in MCDM analyses to evaluate indicators in order to achieve a trade-off between all competing objectives among the set of the non-dominated alternatives (Cascetta, 2009). In fact, decisions and its effects are intrinsically related to a plurality of points of view, which can be defined as criteria (Figueira et al. 2005). MCDM methodologies generate a large number of non-dominated solutions, respecting imposed constraints, which contain the best solution for each criterion, belonging to a frontier (Russo and Vitetta, 2006). One problem of MCDM methodologies is to identify and quantify weights to assign to the objectives and to related criteria (Opricovic, S., & Tzeng, 2002). Weights can be defined exogenous or they can be derived from external analysis (Schenkerman, 1991). One method that can overcome the weight definition problem is Data Envelopment Analysis (DEA).

(c) DEA was originally conceived to evaluate technical efficiency comparing homogeneous production units. DEA is generally used to compare a finite set of Decision Making Units (DMUs) represented by the quantities of inputs which they consume and the quantities of outputs which they produce (Charnes, et al. 1978).

Although results of DEA and MDCM have been compared in the literature, according to the authors' knowledge, the two approaches have not been experimented yet as an integrated evaluation method to support transport planning.

3. THE INTEGRATED PROCESS AND THE PROCEDURE PROPOSED

3.1 The integrated process

Transportation planning is an integrated process aimed to define a planned scenario that reaches objectives according to constrained resources (for instance financial budget). The planned scenario comprehends a sub set of measures (or interventions) to be implemented in the transportation system. A single intervention consumes a predefined quantity of resources; therefore, different planned scenarios can be defined with the same financial budget. For this reason, the decision-maker assumes a decision about which combination of measures (or which scenario) has to implement. The planned scenario results from a trade-off among stakeholders' conflicting interests.

The proposed integrated process is sketched in Fig. 1.

Objectives, *measures* and constrained *budget* are defined. Each stakeholder is considered as a DMU that in a *survey* attributes a relative importance to the objectives. At the same time, starting from the budget and from the predefined set of measures, each DMU selects a sub set of measures (scenario) in order to reach, from his/her point of view, the most important planned objectives.

A system of *simulation* models allows to quantify effects of each single scenario, explicitly selected by each DMU. Behavioral models simulate how the selected measures affect users travel choices in the dimensions of trip generation, distribution, mode and path (Cascetta, 2009). These models are based on random utility theory, which assume that transport users choose the alternative having the maximum perceived utility. The perceived utility may be modelled by means of probabilistic variables. Different model specifications may be used for the estimation of choice probability (percentage) of each alternative for each 'what if' scenario. Quantitative indicators of simulated effects can be compared with one another in order to indentify the best trade-off among conflicting interests. *Comparisons* among DMUs are supported by *DEA* and *MCDM*.

At the end of the integrated process, *best performance scenarios* result.

In order to implement this process the following procedure is proposed.

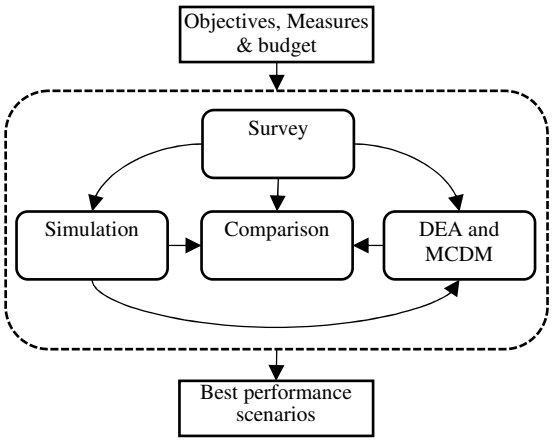


Figure1: The integrated process

3.2 Procedure

Objectives, measures and constrained budget are inputs of the procedure. The constrained budget does not allow implementation of all measures. Therefore, it is necessary a selection among measures, that represents a transport scenario.

In order to reach a trade-off among conflicting interests, the following procedure is proposed:

- first step, *survey* to collect preferences of stakeholders (DMUs) about transport scenarios (measures, resources and exogenous weights); this step is finalized to explicitate DMUs choices about objectives and resources allocation among a predefined set of interventions;
- second step, *simulation* by means of transport model to estimate effects of selected transport scenarios calculating indicators for each scenarios in relation to planning objectives; models are applied to estimate objective indicators related to the simulated effects of scenarios selected by interviewed DMUs;
- third step, *DEA* and *MCDM* to obtain the frontiers;
- fourth step, *comparison* between distances from DEA and MCDM frontiers.

The third and four steps are focused.
In the third step, DEA is developed, starting from the set of selected DMUs, to obtain an efficient frontier that contains best performance scenarios.
The production process represented by DEA in the proposed procedure is the following: each DMU uses a set of DEA inputs (costs for the scenario realization) to produce a set of DEA outputs (scenario indicators). The two vectors, associated to each scenario, are processed in a so-called output-oriented DEA. At the end of this phase the efficient frontier is obtained (DEA frontier). It constitutes the term of reference to compare all DMU alternative scenarios.
In the same step, MCDM is developed using the same data processed in DEA, to obtain a frontier that could not contain best performance scenarios. DMUs, inputs and outputs processed in DEA can be used as alternatives, costs and benefits in MCDM, respectively. Then, these data are elaborated to obtain a frontier adopting MCDM method (MCDM frontier). For instance, it is possible to consider a frontier that contains only the ideal scenario represented by the best values of indicators.

In the four step, the DEA frontier is compared to MCDM frontier measuring:
- the DEA distance from each point representing DMU to the DEA frontier;
- the MCDM distance from each point by means of MCDM frontier.

Similarities between the two distances indicate the possibility to adopt DEA as a tool in multicriteria evaluation of transport scenarios.

4. PROTOTYPAL APPLICATION

The application concerns the transportation planning process of the Città Metropolitana of Reggio Calabria in the south of the Italy (study area). The study area, with a population of about 275,000 inhabitants, comprehends three main urbanized poles: Reggio Calabria is the largest municipality and the most populated; the Tyrrhenian area has more infrastructural facilities (highways, railways, …); the Ionian area is under equipped in terms of quantity and quality of transport infrastructures and services (Figure 2).
At this stage of the research, the main elements of the application are studied. In the following, the integrated process (section 4.1) and the procedure (section 4.2) are specified for the study context.

4.1 The integrated process

In a long term perspective, the transport system of the Città Metropolitana of Reggio Calabria has to pursue sustainability objectives in each of its components:
- social sustainability, including impacts of transport system related to opportunities of people to have access to the urban functions/services (e.g. accessibility);
- economic sustainability, including impacts of transport system, related to cost efficiency (e.g. travel time);
- environmental sustainability, including external impacts of transport system, related to pollution (e.g. air, water, …).

4.2 Procedure

4.2.1 Survey

A set of stakeholders has been selected among academia, public administration and private operators. In the prototypal application, 18 stakeholders have been selected and interviewed, which are included in the sample. The survey is aimed to know stakeholder opinions in relation to interventions priorities on transport system to achieve sustainability objectives (economic, social and environmental).

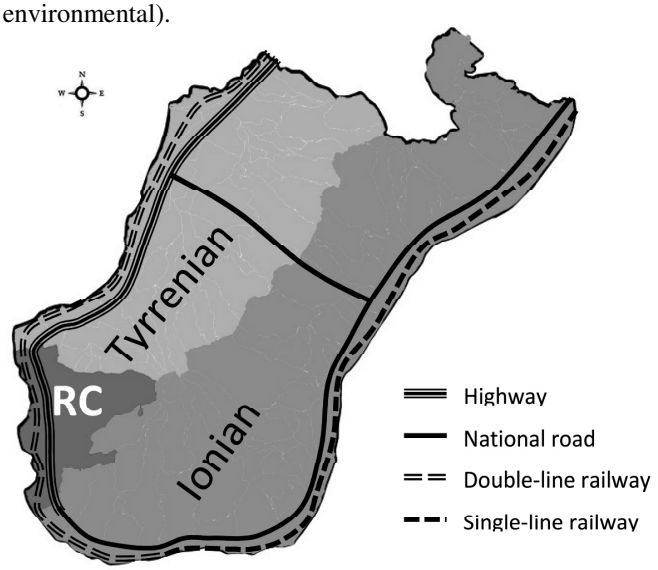


Figure 2: The study area

The questionnaire is subdivided into two sections. The first section concerns the set of the three planned objectives, vector **o** (o_1, economic sustainability; o_2, social sustainability; o_3, environmental sustainability). Each stakeholder should assign a relative weight (vector **u**) to each sustainability component, defined as a percentage from 0 to 100% (Table 1).

Table 1: Planned objectives for the study context

Set of objectives, **o**	Exogenous weights, **u**
economic sustainability, o_1	$u_1 \in [0 – 100\%]$
social sustainability, o_2	$u_2 \in [0 – 100\%]$
environmental sustainability, o_3	$u_3 \in [0 – 100\%$
$u_1 + u_2 + u_3$	100%

The second section concerns the set of planned interventions for transport system (infrastructures and services) of the study area. The interventions (vector **i**) are aggregated into three homogeneous sub sets (vector **M**):
- material infrastructures, M_1, including interventions concerning physical asset of the transport system;
- immaterial infrastructures and management, M_2, including interventions concerning technologies, system organization and governance;
- equipment, M_3, including interventions concerning vehicles and rolling stock.

For each intervention (i_i), a maximum amount of virtual monetary budget (b_i, vector **b**) is predefined (Table 2).
For simplicity sake, the index relative to the DMU is omitted in the following.
The cost c_i for realization of each intervention i_i, belonging to the homogeneous sub set M_j, consumes a predefined quantity of the available virtual monetary budget for the

subset (b_j). This quantity is expressed in terms of a percentage of b_j consumed by the cost c_i:

$$\alpha_{ij} = 100 \bullet c_i / b_j \qquad \forall j$$

In each subset, there is an intervention that has a cost equal to b_j quantity and then it has the value α_{ij} equal to 100.
Each stakeholder of the sample is considered as a DMU that:

- assigns a percentage at each objective (**u**);
- selects one or more interventions, that composes a scenario (**i**), and assigns to each intervention a percentage α_{ij}, respecting the budget constraint for each sub set M_j ($\sum_{i \in Mj} c_i \leq b_j$ or dividing by b_j and multiplying by100):

$$\sum_{i \in Mj} \alpha_{ij} \leq 100 \qquad \forall j$$

The percentage α_{ij} assigned by a DMU to each intervention varies from the minimum value of 0 (the DMU does not select the intervention i_i), to the maximum value of 100 (the DMU recommend the intervention i_i). The above constraints ensure that the total amount of virtual budget consumed to each selected scenario do not overcome the total constrained budget-in each subset.

Table 2: Planned interventions for the study area

Homogeneous subset **M**	Interventions	**i**	Virtual monetary budget (**b**)	Cost declared in the scenario (**c**)
Material infrastructures M_1	Roads in Ionian territory	i_1	60	Depending on the declared value of each DMU
	Sea-mountain Roads	i_2	15	
	Ionian Railways	i_3	50	
	Tyrrhenian Railways	i_4	100	
Immaterial infrastructures and management M_2	Road speed control	i_5	75	
	Information system for mobility	i_6	45	
	Integrated fares in public transport	i_7	60	
	Urban mobility restrictions	i_8	100	
Equipment M_3	Railway rolling stock acquisition	i_9	100	
	Vehicle acquisition for sharing mobility	i_{10}	15	
	Buses acquisition	i_{11}	30	
	Light urban railway	i_{12}	100	

The frequency of the budget assigned by the stakeholders to the interventions is reported in Figure 3. For example, the 65% of the respondents allocates to Ionian railway (i_3) a percentage α_{ij} comprised in the interval 75-100%.

4.2.2 Simulation

Each scenario selected by each DMU (**i**) constitutes an input for transport simulation models.
A discrete mode choice model is applied adopting a classifical form and literature parameters (Cascetta, 2009)**.**
The selected scenario corresponds to estimated values of a set of variables, which feed the mode choice model. The model application provides an estimation of mode choice percentage associated to the selected scenario.
Simulated effects are measured by several indicators and criteria. These values are used as inputs of three indicators, in order to calculate, in relation to the current scenarios, the variations of (vector **p**):

- economic sustainability (p_1), expressed in terms of total travel times (Russo and Musolino, 2012);
- social sustainability (p_2), expressed in terms of variation of Expected Maximum Perceived Utility (Ben Akiva and Lerman, 1985; Cascetta, 2009);
- environmental sustainability (p_3), expressed in terms of quantity of virtual budget allocated to public transport interventions.

The indicators are specified considering that an increasing values of the above indicators corresponds to an increase of sustainability goal.

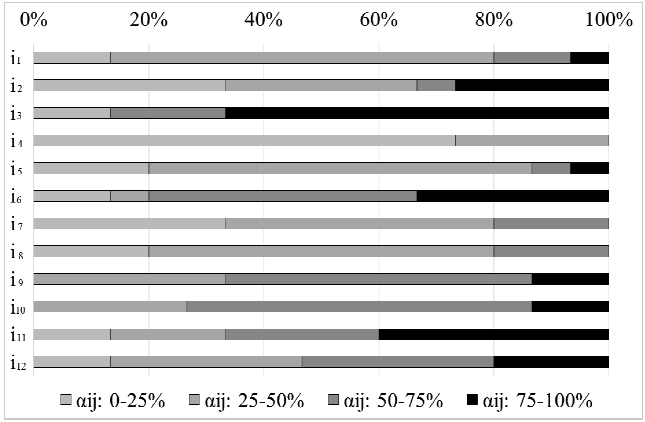


Figure 3: Frequency of declared α_{ij}

4.2.3 DEA and MCDM

Effects of selected transport scenarios (**i**) are compared by means of DEA and MCDM.
In relation to DEA, each selected scenario is represented by:

- one DEA input corresponding to the budget points assigned by each DMU (**b**);
- three DEA outputs corresponding to (**p**):
  - economic sustainability indicator (p_1) (output 1);
  - social sustainability indicator (p_2) (output 2);
  - environmental indicator (p_3) (output 3).

Assuming that the value **b** is the same for all DMUs, each scenario can be represented by a point in a 3D space that has the values of the three outputs as coordinates.
The frontier is obtained applying an output-oriented DEA.
In relation to MCDM, each alternative scenario is represented by the same indicators used in DEA (costs measured by means of budget points and benefits measured by means of sustainability indicators). The MCDM frontier corresponds to the ideal scenario represented by a point in the 3D space that has the coordinate equal to the maximum values of each indicators.

4.2.4 Comparisons

In this phase two distances are compared:

- the DEA distance of each point representing DMU to the output DEA frontier ($\mathbf{d}^{DEA}$);
- the MCDM distance measured from each point representing DMU to an ideal point representing the maximum values of objectives ($\mathbf{d}^{MDCM}$).

In order to compare homogenous entities, the two distances are standardised:

$$d_s^{MDCM/DEA} = (d^{MDCM/DEA} - d_{min}^{MDCM/DEA}) / (d_{max}^{MDCM/DEA} - d_{min}^{MDCM/DEA})$$

In Figure 4 a comparison between standardised DEA and MCDM distances is represented. Each point in the cartesian diagram has the two standardised distances as coordinates. Ideal points where DEA and MDCM output have full concordance lie on the segment connecting the minimum distances point (0, 0) and the maximum distances point (1, 1). The above segment is subdived into four quartiles, according four classes of distances. Many observed DMUs have a standardised distance falling into the first two quartiles. This means that two methods are acting similarly in terms of assessment of the effects, generated by the decisions of the individual DMU.

5. CONCLUSIONS AND RESEARCH PERSPECTIVES

One of the main critical elements in transportation planning process concerns the ex-ante evaluation of effects generated by the planned interventions/policies. These have to be evaluated in relation to the expectations of the potentional involved stakeholders that would like to achieve their objectives according to their interests. In this context, ex-ante evaluations of objectives can be simulated applying quantitative transport models. In this paper, a DEA is experimented as a tool for MCDM in order to support transportation planning.

An integrated process implemented by a procedure is proposed. A set of stakeholders is invited to express their objectives' importance and to indicate how to distribute a given budget of financial resources in order to realise a predefined set of transport interventions.

The first obtained results show that DEA and MCDM above specific conditions give similar results in a multicriteria evaluation context but not the same.

This paper represents a first step of a more general research, that will follow two directions in the next future. The first direction concerns the comparison between DEA and MCDM in order to evaluate the field of applicability of each method and the possible synergy. The second direction concerns the application of the proposed methodology to measure coherence of each stakeholder and to support stakeholder engagement.

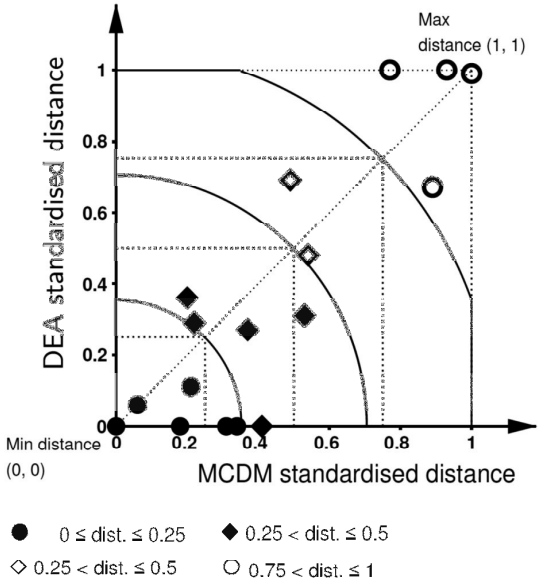


Figure 4: Comparison between DEA and MCDM distances

REFERENCES

Ben-Akiva M., Lerman S.R. (1985). *Discrete Choice Analysis. Theory and Application to Travel Demand*. The MIT Press, London, England

Bickerstaff, K., Tolley, R., and Walzer, G. (2002). "Transport planning and participation: the rhetoric and 25 realities of public involvement." *Journal of Transport Geography*, 10, pp. 61-73

Bouyssou D. (1999) "Using DEA as a tool for MCDM: some remarks", *Journal of the Operational Research Society* 50, pp. 974-978

Caggiani L., Camporeale R., Ottomanelli M. (2017). "Facing equity in transportation Network Design Problem: A flexible constraints based model". *Transport Policy* 55, pp. 9–17

Cascetta E., Cartenì A., Pagliara F., Montanino M. (2015). A new look at planning and designing transportation systems: A decision-making model based on cognitive rationality, stakeholder engagement and quantitative methods. *Transport Policy*, 38, pp. 27–39

Cascetta, E., (2009). Transportation Systems Analysis: Models and Applications 4rd ed., Springer, USA

Cascetta, E., and Pagliara, F., (2013). "Public engagement for planning and designing transportation systems." Procedia - Social and Behavioral Sciences, 87, pp. 103–116.

Cooper W.W. and Seiford L.M., Tone K. (2000). *Data Envelopment Analysis.* Kluwer Academic Publisher, Boston, USA.

Charnes, A., Cooper, W.W., Rhodes, E., (1978). Measuring the efficiency of decision making units. European Journal of Operational Research 2 pp. 429–444.

Figueira, J., Greco, S. & Ehrgott, M. (2005). Multiple Criteria Decision Analysis: State of the Art Surveys. Springer, New York.

López E., Monzón A. (2010) "Integration of Sustainability Issues in Strategic Transportation Planning: A Multi-criteria Model for the Assessment of Transport Infrastructure Plans". *Computer-Aided Civil and Infrastructure Engineering*, **25** (6), pp. 440–451

Macharis C., De Witte A., Turcksin L. (2010). The Multi-Actor Multi-Criteria Analysis (MAMCA) application in the Flemish long-term decision making process on mobility and logistics. Transport Policy 17 pp. 303–311

Nik-Bakht M., El-diraby T. E. (2016) "Communities of Interest–Interest of Communities: Social and Semantic Analysis of Communities in Infrastructure Discussion Networks" *Computer-Aided Civil and Infrastructure Engineering* 31, pp. 34–49

OECD. (Organisation for Economic Co-operation and Development). (2009). "Focus on Citizens: Public Engagement for Better Policy and Services." <http://www.oecd.org/gov/regulatory-policy/focusoncitizenspublicengagementforbetterpolicyandservices.htm> (June, 2015)

Opricovic, S., & Tzeng, G. H. (2002). "Multicriteria Planning of Post-Earthquake Sustainable Reconstruction" *Computer-Aided Civil and Infrastructure Engineering* 17, pp**.** 34–49

Praticò, F.G., Vaiana, R., Iuele, T. (2015), "Macrotexture modeling and experimental validation for pavement surface treatments", *Construction and Building Materials* 95, pp. 658–666.

Russo, F. & Vitetta, A. (2006) "A Topological Method to Choose Optimal Solutions after Solving the Multi-criteria Urban Road Network Design Problem". *Transportation* **33**(347). doi:10.1007/s11116-005-3507-7

Russo, F. & Musolino, G. (2012). "A unifying modelling framework to simulate the Spatial Economic Transport Interaction process at urban and national scales". *Journal of Transport Geography* 24, pp. 189–197

Russo, F. Rindone, C., Panuccio P. (2016). "European plans for the smart city: from theories and rules to logistics test case", *European Planning Studies* 24 (9), pp. 1709-1726

Schenkerman S. (1991). Use and abuse of weights in multiple objective decision support models. Decision Sciences. 22 (2), pp 369–378.

AUTHOR BIOGRAPHIES

GIUSEPPE MUSOLINO is assistant professor at Department DIIES of Università Mediterranea di Reggio Calabria, Italy. His research interests include land use-transport interaction models, microscopic traffic flow models, international maritime freight transport analysis. He published papers in international journals such as Journal of Transport Geography, International Journal of Shipping and Transport Logistics and Safety Science, among the others.

CORRADO RINDONE is technical responsible at Department DIIES of Università Mediterranea di Reggio Calabria, Italy. In the field of Transportation Engineering he has following research interests: transportation planning processes and products; parametric and non parametric evaluation methods in transportation; safety in transportation.

ANTONINO VITETTA is associated professor at the at Department DIIES of Università Mediterranea di Reggio Calabria, Italy. In the area of Transport System at urban and extraurban scales has studied: different aspects connected with the analysis of the mobility demand and supply; the interaction between the strategies of the decision making and the users and design involving with road urban transport.

TRAFFIC SIMULATION

Evaluation of car-following-models at controlled intersections

Laura Bieker-Walz
Michael Behrisch
Marek Junghans
Kay Gimm
German Aerospace Center (DLR)
email: laura.bieker@dlr.de

KEYWORDS
Microscopic simulation, Traffic research, Open source framework

ABSTRACT

Traffic simulations can help to investigate new traffic and transportation management solutions for overcoming problems like traffic jams, accidents or environmental pollution. For this a valid simulation model is needed. This paper provides an overview of the open source traffic simulation framework SUMO (Simulation of Urban MObility) and evaluates the implemented car-following models at controlled intersections with regard to vehicle positions and speeds. Particularly intersections can be bottlenecks for high traffic volumes and have a higher risk for accidents. Therefore this study focuses on traffic behavior at urban intersections.

Introduction

The increasing vehicular mobility has been offering many advantages for the population in urban areas, e.g. more comfort and flexibility. On the other hand, the rising amount of vehicles also lead to traffic problems like traffic jams, environmental pollution and accidents. To reduce these problems, particularly traffic and transportation management focuses on intelligent traffic management strategies.
Due to the complexity of managing mobility in urban areas, it would be very time-consuming, expensive and to some extent dangerous to test traffic management strategies in real world or test fields without theoretical evaluations before. Consequently, theoretical methods are necessary to analyze the benefits of traffic and transportation management strategies; and simulation frameworks can be one opportunity. The microscopic traffic mobility framework SUMO (Simulation of Urban MObility) is a time-discrete and open source simulation tool enabling such evaluations (Krajzewicz et al. 2012). SUMO can simulate fast and easily the traffic mobility of a large traffic network and provides many useful tools to evaluate the simulated data. An example simulation in SUMO can be seen in Figure 1.

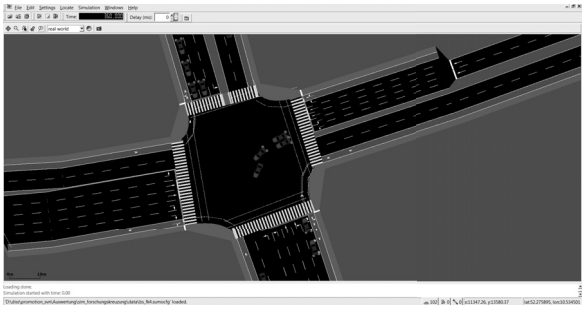


Figure 1: SUMO Simulation of the Research Intersection in Brunswick

Many microscopic traffic simulation tools are based on car-following models, which are well studied in research (Brackstone and McDonald 1999). They focus on the idea that the speed of a vehicle highly depends on the speed of the leading vehicle. Usually, the traffic behavior at intersections is often neglected in these models. But for traffic efficiency and safety the vehicle interaction at intersections have a high influence and are therefore in the focus of this research.
The paper is structured as followed: first a short introduction of traffic simulations and a description of some of the most common car-following models (Krauss, IDM and Wiedemann) will be given. Next, a controlled intersection in Brunswick (Germany) and its real world traffic data will be described. Afterwards, the simulation tool SUMO and the used simulation scenario will be presented. Finally, the simulation results and concluding remarks will be stated.

Traffic Simulation Models

For modeling traffic a large variety of different simulation models are available. These models can be divided mainly into three different types (Krauss 1998):

1. Macroscopic: average vehicle dynamics are simulated, e.g. traffic density.

2. Microscopic: vehicle dynamics are modeled for every single vehicle individually

3. Mesoscopic: a mixture of macroscopic and microscopic model, for instance vehicle queues

For the simulation of vehicle interaction a microscopic model is necessary. Vehicle dynamics are normally described as a function of the velocity and the position of each vehicle. A common process to describe these dynamics is to apply car-following and lane change models. This research concentrates on car-following only. All described models are implemented in the most recent version of SUMO.

Car-following models

The basic idea of the car-following theory is that the change in velocity v of a vehicle i depends on the velocity of the leading vehicle $i+1$ as well as the position difference (gap) and static parameters like the sensitivity or reaction time τ. (Krauss 1998)

$$\frac{dv_i(t)}{dt} = f(v_{i+1}(t), x_{i+1}(t) - x_i(t), \tau, ...) \tag{1}$$

Krauss model
The default car-following model of SUMO is the Krauss model (Krauss et al. 1997, Krauss 1998). In traffic simulation each vehicle can have two different motion types: free motion and interacting motion. In free motion, no leading vehicle limits the speed of the following vehicle. Therefore, its speed is bounded to its maximum (depending of the speed limit and the drivers desired speed):

$$v \leq v_{\max} \tag{2}$$

In case two vehicles interact with each other, both vehicles always try not to collide with each other. In this case at least one of the drivers reduces its speed that is not higher than the maximum safe velocity v_{safe}:

$$v \leq v_{safe} \tag{3}$$

The model is collision free, which means that no vehicle is driving faster than a safe speed v_{safe}. The safe velocity will be computed every time step using the following equation (Krajzewicz et al. (2002)):

$$v_{safe}(t) = v_l(t) + \frac{g(t) - v_l(t)\tau}{\frac{v}{b(v)} + \tau} \tag{4}$$

t : time step
$v_l(t)$: velocity of the leading vehicle in t
$g(t)$: gap between vehicle and leading vehicle i in t
τ : reaction time of the driver (usually 1 second)
b : deceleration function

In real life the acceleration of a vehicle depends on its physical ability and other effects like air resistance and others. To prevent that vehicles in the simulation are driving faster that it is possible in reality the desired speed v_{des} is calculated. The desired speed of each vehicle v_{des} is the minimum speed of the safe speed v_{safe}, the current speed plus the maximum acceleration and the maximum speedKrajzewicz et al. (2002):

$$v_{des}(t) = \min[v_{safe}(t), v(t) + a, v_{\max}] \tag{5}$$

Due to the imperfection of the human drivers, a random error is subtracted from the desired speed v_{des} Krajzewicz et al. (2002):

$$v(t) = \max[0, rand[v_{des}(t) - \epsilon a, v_{des}(t)]] \tag{6}$$

Intelligent Driver Model - IDM
The Intelligent Driver Model (IDM) is based on the Optimal Velocity Model (OVM) (Treiber and Kesting 2010):

$$\dot{v} = \frac{(v_{opt}(s) - v)}{\tau} \tag{7}$$

s : current gap to the leading vehicle
$v_{opt(s)}$: optimal velocity depends on gap s
τ : Time to adapt to the new speed
The OVM does not take the speed of the leading vehicle into account. It reacts only to the distance to the leading vehicle. Additionally, the OVM is very sensitive to accidents. Therefore the IDM was modeled, with the following acceleration equation (Treiber et al. 2000, Treiber 2017):

$$\dot{v} = a\left[1 - \left(\frac{v}{v_0}\right)^{\delta} - \left(\frac{s*(v,\Delta v)}{s}\right)^2\right] \tag{8}$$

v : current velocity
v_0 : desired velocity
s^* : desired gap

The desired gap s^* is calculated as follows:

$$s^*(v, \Delta v) = s_0 + \max\left[0, \left(vT + \frac{v\Delta v}{2\sqrt{ab}}\right)\right] \tag{9}$$

Every vehicle can have other values for the parameter of the model:
T : the time headway (between 0.8 -2 seconds)
s_0 : is the minimum gap (default: 2 meters)
a : acceleration (between 1-2 m/s^2)
b : deceleration (between 1-2 m/s^2)

Wiedemann model
The commercial traffic simulation Vissim uses the Wiedemann Model (Menneni et al. 2009). The Wiedemann model is a psycho-physical spacing model. If a faster vehicle is approaching a slower leading vehicle it will start to decelerate until it reaches its individual

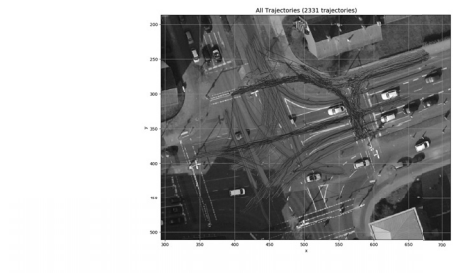


Figure 2: Real world picture of the intersection in Brunswick (Lines are representing detected trajectories red: vehicles, blue: bikes, pink: pedestrians)

threshold. The threshold is a function of speed difference and spacing. Human drivers are not able to perceive small speed differences and to keep their speed very accurate. Therefore, the vehicle will accelerate again if another threshold is reached (Fellendorf 1994).

Real World Traffic

The purpose of this research is to evaluate existing simulation models and compare the simulation results with data from human drivers. Therefore, one hour trajectory data (space-time-curves) from 23. January 2017 of human drivers recorded at a highly frequented junction (>20,000 traffic participants per day) in Braunschweig, Germany, intersecting two main roads was used. The Research Intersection is part of the test field AIM (Application Platform for Intelligent Mobility) and serves as a field instrument for detection and assessment of traffic behavior at complex urban intersections (Knake-Langhorst and Gimm 2016). Its infrastructure is equipped with several mono-cameras and multirange radar sensors to detect and track traffic participants in the inner part of the intersection. These trajectories provide static information about vehicle types (cars, motorcycles, trucks/vans, bicycles and pedestrians), vehicle sizes and time-variant information about their kinematic states (position, speed, acceleration) and headings when moving through the intersection. For details see (Knake-Langhorst and Gimm 2016) and (Schnieder and Lemmer 2012).
The trajectories' positions were mapped on the intersection as shown in Figure 2. The best view on the intersection is provided by the camera in the East (right street in Figure 2). Therefore, only trajectories from the East to the main street in the North were used for evaluation.

Simulation

SUMO is an open traffic simulation framework which is developed since 2001. A large amount of additional tools e.g. for routing, evaluations and emission calculation are

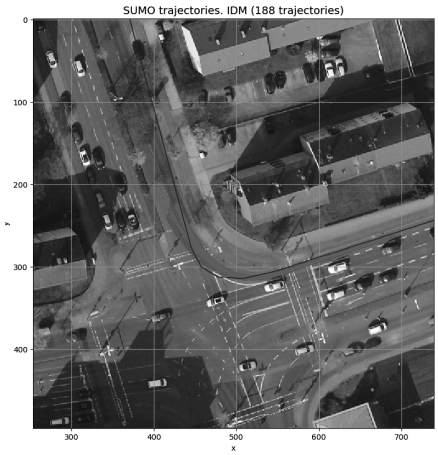


Figure 3: Real world trajectories (red) and simulated trajectories (blue)

available within SUMO. Furthermore, SUMO supports intermodal traffic systems including the use of public transport and the simulation of pedestrians. The source code of SUMO is freely available and can be extended by the users with their own algorithms and models. The user can only interact with SUMO via an interface called TraCI. Different traffic networks can be imported for example OpenStreetMap, VISUM, VISSIM and NavTeq. SUMO has been used in several international research project e.g. COLOMBO (Leich et al. 2016), Amitran and VABENE (Flötteröd and Bieker 2012).
SUMO version 0.31 was used to simulate the intersection in Brunswick with the three presented car-following models, see Figure 1. In the configuration file of SUMO can be stated which car-following model should be used. The simulated data was exported as floating car data (FCD) in XML format. The data evaluation was done in Python and the diagrams created with matplotlib (Hunter 2007).

Results

For the evaluation in this study the real world traffic trajectories of the intersection were compared with the simulated trajectories. In Figure 3 the trajectories mapped on the intersection are shown. While the real world trajectories are varying in the lateral movement of the lane; the simulated vehicles are always keeping the middle of the lane. An extension in SUMO allows vehicle to move on sub-lanes but this model is normally only used for overtaking, especially from bicycles, motorcycles or in case of building rescue lanes and will not influence lateral positioning on a free road. The trajectory data could be used to extend the sub-lane model to are more realistic lateral movement behavior, but are neglected in this study.
Figure 4 shows the approaching behavior of real world

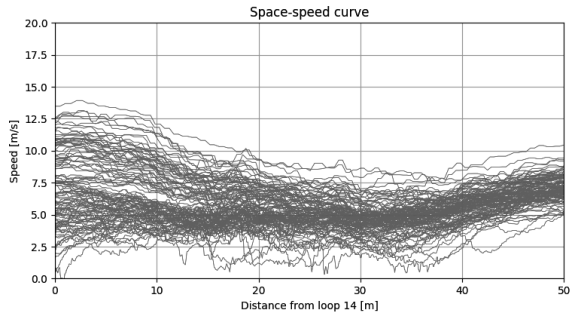


Figure 4: Real world traffic trajectories approaching a green traffic light

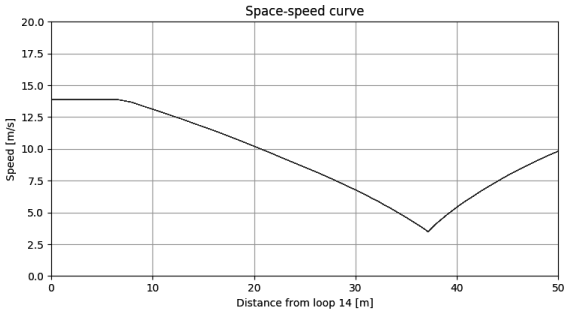


Figure 5: Simulated trajectories with Krauss model approaching a green traffic light

vehicles at a green traffic light. All trajectories which are passing the intersection without stop are considered in this study. It can be seen that the speed trajectories are oscillating a lot for human drivers. There is not one typical trajectory how the vehicles are driving over the intersection. In further research it would be interesting to see whether it is possible to cluster the trajectories to have a set of typical trajectories.

To simulate an ideal traffic case, only vehicles with could pass the intersect via green have been evaluated here. Compared to the real world trajectories, simulated trajectories appear to have less variation in all car following models (see Figures 5and 6). In Figure 5 the simulation results of the Krauss model are displayed. All vehicles have the same speed curves, which is due to the fact that drivers are keeping their desired speed until they approach the traffic light and are reducing their speed because of the pedestrian crossing. The IDM has a slightly larger oscillation of the speed curves, see Figure 6. The simulation results of the Wiedemann model are looking almost the same as the Krauss model and are therefore not illustrated here.

Furthermore, the simulation was feed with real world traffic demand to produce are a more realistic traffic behavior at the intersection. To compare the real world trajectories with the simulated trajectories the average speeds for every trajectory over the intersection were calculated and displayed in a box plot in Figure 7. While

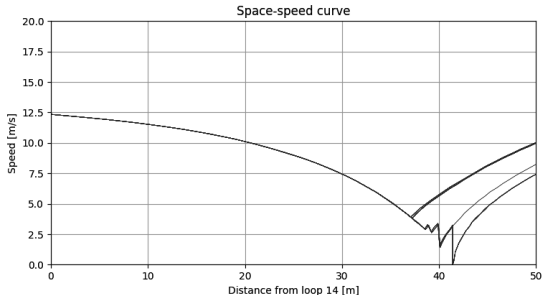


Figure 6: Simulated trajectories with IDM approaching a green traffic light

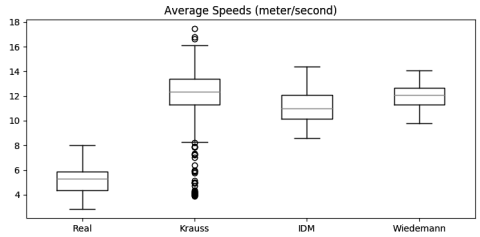


Figure 7: Comparison of the average speed for the vehicles in the area of the intersection

the average speed in the simulation reach almost the maximum of the allowed speed of 13.8 meters per second, the human drivers have to lower their speed significantly. One reason for this is that the driver have to decelerate and take care that no other traffic participant are in their blind spot. In the simulation the drivers always know where all other traffic participants are in each time step and therefore they do not have to break if no traffic participant is conflicting with its route.

The results how much time the vehicles needed to pass the intersection are similar and can be seen in Figure 8. The results of the IDM are closer to the results of the real world trajectories than the Krauss and the Wiedemann model. In average the vehicles need about 6 seconds to cross the intersection with the IDM and in real life while the Krass and Wiedemann model are underestimating the passing time by one second. The time difference might have not a high influence on the route of a single vehicle in the simulation but also might sum up for large simulation scenarios and routes.

Conclusions & future work

In this study the results of different car following models were compared with real world trajectories. Furthermore, this paper presented a short introduction to the traffic simulation framework SUMO. The SUMO framework includes a lot of tools for preparing a realistic traffic simulation and for the evaluation of the

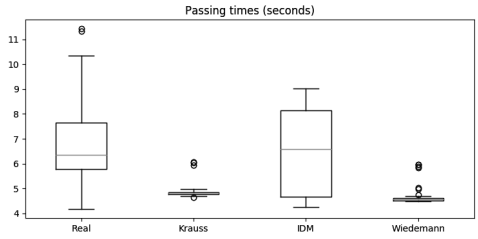


Figure 8: Comparison of the passing times for the vehicles in the area of the intersection

simulation results. In addition, SUMO provides different car-following models, which can be used for traffic evaluations. The described models are calibrated in other studies and are working well on highways and urban roads. But still some aspects like accident behavior are not modeled yet, every driver is always keeping a distance which is safe so that an accident should not happen.

Real world vehicle data can be used to modify the simulation models to have a more realistic driving behavior and therefore better simulation results. Our work can help to calibrate car following models to fit to the real world trajectories. Additionally, further studies can use this work e.g. for estimating the correct traffic light phases for induction loops or to develop a probabilistic car-following model.

Furthermore, traffic camera data could also provide live data for a SUMO simulation. Then the live simulation scenario could improve the traffic data results e.g. broken trajectories could be fixed by simulated trajectories or a trajectory forecast can be given by the simulation.

REFERENCES

Brackstone M. and McDonald M., 1999. *Car-following: a historical review. Transportation Research Part F: Traffic Psychology and Behaviour*, 2, no. 4, 181 – 196. ISSN 1369-8478. doi:https://doi.org/10.1016/S1369-8478(00)00005-X. URL `http://www.sciencedirect.com/science/article/pii/S136984780000005X`.

Fellendorf M., 1994. *VISSIM: A microscopic simulation tool to evaluate actuated signal control including bus priority*. In *64th Institute of Transportation Engineers Annual Meeting*. Springer, 1–9.

Flötteröd Y.P. and Bieker L., 2012. *Demand-oriented traffic management for incidents and disasters*. In *Second International Conference on Evacuation Modeling and Management (ICEM 2012)*.

Hunter J.D., 2007. *Matplotlib: A 2D graphics environment. Computing In Science & Engineering*, 9, no. 3, 90–95. doi:10.1109/MCSE.2007.55.

Knake-Langhorst S. and Gimm K., 2016. *AIM Research Intersection: Instrument for traffic detection and behavior assessment for a complex urban intersection. Journal of large-scale research facilities*, 2, no. A65. doi:http://dx.doi.org/10.17815/jlsrf-2-122.

Krajzewicz D.; Erdmann J.; Behrisch M.; and Bieker L., 2012. *Recent Development and Applications of SUMO - Simulation of Urban MObility. International Journal On Advances in Systems and Measurements*, 5, no. 3&4, 128–138.

Krajzewicz D.; Hertkorn G.; Rössel C.; and Wagner P., 2002. *SUMO (Simulation of Urban MObility) - an open-source traffic simulation*. In A. Al-Akaidi (Ed.), *4th Middle East Symposium on Simulation and Modelling*. 183–187. URL `http://elib.dlr.de/6661/`. LIDO-Berichtsjahr=2004,.

Krauss S., 1998. *Microscopic Modeling of Traffic Flow: Investigation of Collision Free Vehicle Dynamics*. Ph.D. thesis, Universität zu Köln.

Krauss S.; Wagner P.; and Gawron C., 1997. *Metastable states in a microscopic model of traffic flow. Phys Rev E*, 55, 5597–5602. doi:10.1103/PhysRevE.55.5597. URL `http://link.aps.org/doi/10.1103/PhysRevE.55.5597`.

Leich A.; Niebel W.; Bieker L.; Blokpoel R.; Caselli F.; Härri J.; Junghans M.; Saul H.; and Stützle T., 2016. *COLOMBO Deliverable 7.6: Project Final Report*. Tech. rep.

Menneni S.; Sun C.; and Vortisch P., 2009. *Integrated microscopic and macroscopic calibration for psychophysical car-following models*. In *Transportation Research Board 88th Annual Meeting*. 09-2773.

Schnieder L. and Lemmer K., 2012. *Anwendungsplattform Intelligente Mobilität-eine Plattform für die verkehrswissenschaftliche Forschung und die Entwicklung intelligenter Mobilitätsdienste. Internationales Verkehrswesen*, 64, no. 4, 62–63.

Treiber M., 2017. *Longitudinal Traffic model: The IDM*. URL `http://traffic-simulation.de/IDM.html`.

Treiber M.; Hennecke A.; and Helbing D., 2000. *Congested traffic states in empirical observations and microscopic simulations. pre*, 62, 1805–1824. doi:10.1103/PhysRevE.62.1805.

Treiber M. and Kesting A., 2010. *Verkehrsdynamik und -simulation: Daten, Modelle und Anwendungen der Verkehrsflussdynamik*. Springer-Lehrbuch. Springer Berlin Heidelberg. ISBN 9783642052279. URL `https://books.google.de/books?id=c7M8mQEACAAJ`.

A STOCHASTIC DRIVER DISTRACTION MODEL FOR MICROSCOPIC TRAFFIC SIMULATIONS

Manuel Lindorfer
Christian Backfrieder
Gerald Ostermayer
FH Upper Austria
Softwarepark 11, 4232 Hagenberg
Austria

Christoph F. Mecklenbräuker
Institute of Telecommunications
TU Wien
Gußhausstraße 25-25a, 1040 Vienna
Austria

KEYWORDS
Driver Distraction, Human Factors, Microscopic Traffic Simulation, Computer Modeling

ABSTRACT

In this paper we propose a novel approach towards integrating driver distractions with microscopic traffic simulations for the purpose of evaluating the impact of distracted driving on traffic flow in large-scale simulator studies. We use stochastic processes to model the frequency and duration of distractive tasks and, by doing so, generate distraction profiles for every individual vehicle in the simulation. We validate our model with the aid of an empirical data set and show that it is capable of reproducing the statistical properties derived from these data to a satisfying extent by performing dynamic simulations using the microscopic traffic simulator TraffSim.

INTRODUCTION

Distracted driving and the magnitude of its impact on road safety has prompted considerable attention in recent years. Throughout the years, various studies have been conducted in order to derive insights on these impacts and to understand the factors contributing to driver distraction, reporting that driver distraction and inattention is responsible for 25%-30% of police-reported crashes (Stutts et al. 2003). Other studies showed that distracted driving contributes to almost 80% of all crashes and 65% of all near-crashes (Dingus et al. 2006). Moreover, it was found that distracted driving is the leading cause of vehicle collisions in the United States (Hendricks et al. 2001; Young et al. 2007).

Driver distraction can be defined as "*the diversion of attention away from activities crucial for safe driving to a competing task*" (Lee et al. 2008), and generally results in a deterioration of driving performance. This includes, among others, an increase of the driver's reaction time or impacts on vision and steering behavior (Cooper et al. 2009; Dingus et al. 2016). Despite their ascertained influences on road safety, driver distractions and their large-scale effects on driving behavior and traffic flow in general have hardly been investigated in existing studies, whether it be naturalistic or laboratory experiments, statistical or simulation approaches (Wang et al. 1996; Redelmeier and Tibshirani 1997; Wilson et al. 2003; Stutts et al. 2003; LeBlanc 2006; Young et al. 2007; Owens et al. 2011; McKeever et al. 2013). The same holds true for the area of traffic simulation, which has become more and more popular in recent years, not least because of the emergence of Intelligent Transportation Systems (ITS). Although the increasing demand for the realistic simulation of vehicular traffic lead to various attempts to integrate human behavior with simulation models (Bando et al. 1998; Ahmed 1999; Treiber et al. 2006; Andersen and Sauer 2007; Jin et al. 2011), driver distractions and their evident impacts have largely been disregarded so far.

In order to bridge the gap between the vast number of studies available in literature and distraction simulation, we propose a computational model allowing for the integration of driver distractions with microscopic traffic simulators. The model is capable of simulating an arbitrary number of distraction types simultaneously and uses stochastic processes to generate distraction profiles for every single simulated vehicle, indicating whether the vehicle is distracted at a given point in time or not. The generation process is guided by a set of parameters specifying, among others, the frequency and duration of a distractive task or the drivers' exposure to potential distracting events. By taking these characteristics into consideration, we are able to simulate scenarios with a large number of vehicles while at the same time maintaining the statistical properties obtained from a naturalistic driving study (Stutts et al. 2003) with a comparatively small number of participants.

The remainder of this paper is organized as follows. The next section provides a review on driver distraction studies and notable efforts to integrate distracted driving with (traffic) simulations. Afterwards, the proposed distraction model is outlined in more detail, including its exemplary integration with the microscopic traffic simulator TraffSim (Backfrieder et al. 2015). The model is validated in the forthcoming section using empirical data gathered from a naturalistic driving study. Finally, the last section concludes the paper and gives an outlook on planned future work.

RELATED WORK

Distracted driving is an extensively studied, yet controversial topic in the scientific community. Hereinafter, we outline several studies which have been conducted in that respective field in recent years. Subsequently, we review some notable approaches towards the integration of driver distractions with traffic simulation models.

Studies on Distracted Driving

During the last decade, studies on driver distraction received a great deal of attention, not least since driver preoccupation with electronic devices while driving is becoming increasingly common (Regan 2004). Throughout the years,

numerous approaches towards studying the causes and consequences of distracted driving have entrenched in literature. These approaches can in principle be categorized into four groups.

The first category includes naturalistic experiments and field tests which record the drivers' behavior using cameras or kinematic sensors in a real driving scenario (Stutts et al. 2003; Young et al. 2007; Cooper et al. 2009; Wilson and Stimpson 2010; Fitch et al. 2013). Although these experiments allow to capture driving situations in a realistic environment, they are in most instances limited in terms of costs, time and the number of participants.

The same holds true for laboratory or driving simulator experiments, which are less risky compared to naturalistic studies, and which are among the most common methods to investigate driver distraction (Strayer et al. 2006; Beede and Kass 2006; Young et al. 2007; Cooper et al. 2009; Stavrinos et al. 2013).

The third group comprises so-called population-based studies which make use of existing databases (e.g. the Fatality Analysis Reporting System (National Highway Traffic Safety Administration 2015)) in order to derive relations between current crash trends and the drivers' preoccupation by secondary tasks (Redelmeier and Tibshirani 1997; Wilson and Stimpson 2010). The significance of the latter approach, however, strongly depends on the quality of the gathered data and the completeness of the underlying database.

To overcome the limitations of the first three approaches, researchers in the field have put considerable efforts into the development of cognitive models for the purpose of distraction prediction (Salvucci 2009; 2013). The main benefit of such computational modeling approaches is that they are comparatively fast, easier to set up and, most importantly, repeatable. Conversely, the models used in such experiments require extensive validation in order to ensure that their predictions correspond to real-world measures.

What a majority of these studies and models has in common is that they focus either on (i) distraction scenarios with just one or a few vehicles, or (ii) scenarios where drivers are exposed to a particular type of distraction. However, neither the impacts of driver distraction on traffic flow considering a large number of vehicles nor the simultaneous distraction of a significant number of drivers by different types of distractive activities has been investigated to date, as far as the authors know.

Distracted Driving in Traffic Simulations

Traffic simulation constitutes a useful method to investigate scenarios which are either too costly or too complex to be studied in a real-world setup or by other analytical methods. Throughout the years, numerous simulation frameworks have been developed by researchers in the field, designed for the simulation of vehicular traffic at different levels of abstraction (Fellendorf 1994; Behrisch et al. 2011). Thereby, microscopic traffic simulators provide the highest level of detail, as the movements of every single vehicle and its characteristics are modeled individually. Starting already in the 1950s, a multiplicity of behavioral models has been developed for the purpose of describing vehicle interactions at the microscopic level (Pipes 1953; Newell 1961; Gipps 1981; Treiber et al. 2000). Whilst most of these models aim to simulate driving under ideal conditions, the increasing demand for the realistic simulation of vehicular traffic lead to various attempts to integrate human factors such as a delayed reaction, anticipation capabilities or estimation errors with microscopic simulation models (Bando et al. 1998; Ahmed 1999; Treiber et al. 2006; Andersen and Sauer 2007; Jin et al. 2011).

Despite the indisputable outcomes of the abovementioned studies, driver distractions have largely remained disregarded in the scope of microscopic traffic simulations. A notable effort in this context has been made by (Yang and Peng 2009), who developed an error-able driver model capable of simulating both nominal and devious driving behavior. Driver distractions are viewed as recurring events which affect the driver in terms of increased workload and degraded control. Given a proper parametrization, the model is able to reproduce accident behavior that corresponds to field test results. Whilst the model is capable of simulating the simultaneous distraction of multiple drivers, it does not distinguish between different types of driver distraction, which, however, is inevitable when studying the large-scale impacts of distracted driving on traffic flow or road safety.

More recently, (Nourzad et al. 2014) proposed an integrative approach towards investigating the large-scale effects of driver distractions. They use a cognitive model to derive distraction profiles for different types of driver distraction, indicating whether or not a driver is distracted at a certain point in time. Subsequently, these profiles are integrated into the microscopic traffic simulator VISSIM (Fellendorf 1994). Although it can be considered as a first initial step for evaluating large-scale impacts of distracted driving with the aid of traffic simulations, their work also has its limitations. On the one hand, characteristics such as the number of occurrences or the drivers' exposure to specific types of distractive activities are not considered in the distraction profile generation process. On the other hand, their model was validated with the aid of simulation models rather than empirical data.

In order to overcome mentioned limitations, we introduce a computational model for the purpose of investigating the large-scale effects of distracted driving on vehicular traffic flow. The model is derived from empirical data gathered from a naturalistic driving study (Stutts et al. 2003) and uses stochastic processes to generate distraction profiles for individual vehicles. In contrast to (Nourzad et al. 2014), we consider global characteristics such as the frequency or duration of distracting events as well as the drivers' exposure to certain types of distraction. Moreover, we show that our model is capable of reproducing the statistical properties of the underlying naturalistic experiment for a large number of vehicles to a satisfying extent by means of dynamic traffic simulations.

DISTRACTION MODEL

In the following, the proposed distraction model is elaborated in more detail. Likewise, we demonstrate how the model can be integrated with traffic simulation frameworks with the aid of the microscopic simulator TraffSim (Backfrieder et al. 2015).

Reference Study

As already mentioned, we used the findings of a naturalistic driving study as a starting point for the development of our distraction model. The study conducted by (Stutts et al. 2003) involved the collection of unobtrusive video data from 70 volunteer drivers over a period of three hours of driving, with the aim of evaluating the driver's exposure to overall 17 different types of distraction (e.g. eating or drinking, talking on cell phone, reaching). Apart from providing insights on the causes and consequences of distracted driving, the study outcomes also comprise an extensive analysis of the recorded data, yielding a set of statistical parameters for each individual type of distraction:

- Exposure, i.e. the percentage of drivers engaging in the distractive activity
- Number of occurrences in the entire observation period
- Total duration, i.e. the cumulative duration among all drivers engaging in the distractive activity
- Average, minimum, maximum and standard deviation values for the duration of the distracting event

Beyond doubt, the work by (Stutts et al. 2003) is a valuable and still one of the most comprehensive contributions to the field of driver distraction research. Nevertheless, their study does not provide any particularized insights on abovementioned parameters. More precisely, their data is limited to aggregated or averaged descriptive quantities, while the actual samples have not been published to date.

Model Layout

In the following, we delineate the functional principle of the proposed distraction model with the aid of the microscopic traffic simulator TraffSim (Backfrieder et al. 2015). Figure 1 depicts the fundamental structure of the model, which can, loosely speaking, be separated into two phases. As one might expect, the *offline* phase relates to a set of preprocessing actions being executed before the actual simulation takes place, whilst the *online* part comprises all components which are involved in the distraction profile generation process in the course of an ongoing simulation. Hereinafter, we outline the model's core components in more detail.

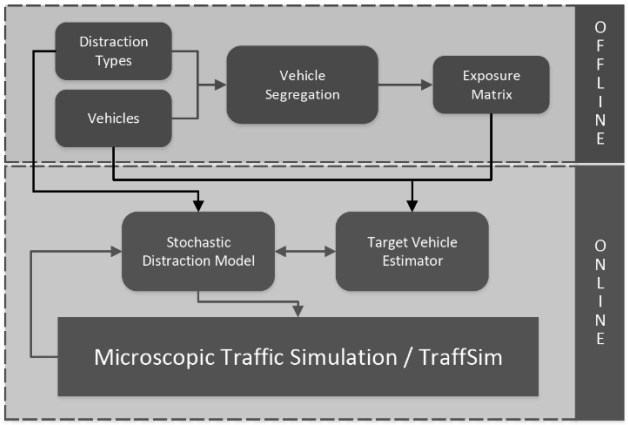


Figure 1: Functional Principle of the proposed Two-Phase Distraction Model

Vehicle Segregation

In a preliminary step, the model makes use of input data comprising the specification of all relevant distraction types as well as a set of vehicles used in the downstream simulation in order to generate a so-called *exposure matrix*. This matrix is basically an allocation of vehicles to one or multiple types of distracting activity. The generation of this matrix is thereby heavily influenced by the parametrization of the distraction types used as input to the model, which principally rests on the parameters derived in (Stutts et al. 2003). In course of the *vehicle segregation* process a uniformly distributed random variable $X \sim U(0,1)$ is compared with the exposure parameter e_d of a distraction type (cf. Equation (1)) in order to determine whether or not a certain vehicle is likely to engage in this particular type of distractive task. This process is repeated for each distraction type and every single vehicle, yielding a complete allocation matrix of each distraction type to a specific subset of vehicles.

$$\text{exposure} = \begin{cases} 1, & \text{if } X < e_d \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

Target Vehicle Estimator

The exposure matrix assembled in the first step is used as input for the *Target Vehicle Estimator* (TVE), which is triggered repeatedly during an ongoing simulation. In principle, this process is used to determine which of the simulated vehicles should be distracted by a given type of distracting activity at a certain point in time.
In order to find the respective target vehicle, a set of vehicles being worth considering for a given distraction type is composed in the first instance. Therefore, all vehicles which can be viewed as distractible are selected from the basic set of vehicles. To be more specific, this is the case for all vehicles that (i) have a corresponding allocation to the respective type of distracting task in the exposure matrix and (ii) are not distracted by any other kind of distractive activity at that point in time. Finally, the target vehicle is selected randomly from the remaining set of potentially distractible vehicles.
Although our assumption that a driver cannot be exposed to more than one type of distraction at the same time constitutes a simplification compared to reality, we consider this circumstance more as a parametrization rather than a modeling issue. Whilst the simultaneous distraction by different types of activities clearly makes a difference in a real driving scenario, the consequences associated therewith can be expressed simply by a proper parametrization in a simulation environment.

Stochastic Distraction Model

The *Stochastic Distraction Model* (SDM) is the core component in the proposed model and is directly coupled with the traffic simulation framework. Similar to other components in a microscopic simulation environment (e.g. crash detector, vehicle updater), the SDM is updated continuously in fixed time intervals in order to drive on the simulation. The main purpose of the SDM is to model both the frequency as well as the duration of distracting events, which are, in further consequence, assigned to vehicles determined by the TVE.

Frequency of Distractive Events

Apart from measuring the drivers' exposure to a specific type of distractive task, (Stutts et al. 2003) also evaluated how often that particular type of distraction emerged over the study period. A commonly used approach to model random, mutually independent occurrences of events are Poisson processes. Such processes are frequently used to model the arrivals of customers in queuing theory (Tijms 2003) or packets in wireless networks (Haenggi et al. 2009), for example. The inter-arrival times in Poisson processes, i.e. the time difference between two consecutive events, are independently exponentially distributed random variables with a rate parameter λ, representing the average number of arrivals per time unit (Taylor and Karlin 2014). We consider Poisson processes to be a suitable method do describe the arrivals of distractive events, as they occur (i) with a known average rate and (ii) independently of the time since their last occurrence.

Taking above deliberations into consideration, we model the time difference between two consecutive occurrences of a particular type of distractive task using an inverse transform sampling approach as delineated by Equation (2),

$$\Delta t = \frac{-\log(1 - X)}{\lambda} \qquad (2)$$

where Δt is the time until the next occurrence, X is a uniformly distributed random variable $X \sim U(0,1)$ and λ is the rate parameter of the exponential distribution. In our case, λ represents the average number of occurrences in a time interval of one second, and can be derived from the findings in (Stutts et al. 2003). In that manner, the first time of arrival for every single distraction type is evaluated at the start of a simulation. The process is repeated every time a new arrival is triggered. Subsequently, the corresponding distraction is parametrized using a random duration (see below) and assigned to a target vehicle obtained from the TVE.

Duration of Distractive Tasks

In recent years, considerable attention has been paid to investigating the amount of time drivers are distracted by different types of activities. Whilst most studies available in literature focus on a single type of distraction, (Stutts et al. 2003) provide insights on the average duration of 17 distinctive types of distracting tasks obtained from 70 volunteer drivers. Their findings comprise a statistical analysis of these data, including minimum, maximum and standard deviation values for the observed durations. However, they do not provide any detailed information relating to the actual statistical distribution of the ascertained values.

Being limited to the first two central moments of the empirical distribution, namely mean and variance, we use the method of moments (Bowman and Shenton 2004) to find a suitable estimation for the probability distribution of the duration of distractive activities. Apart from the fact that we do not have any detailed knowledge about the empirical samples, which is required for more sophisticated types of distribution fitting such as e.g. maximum likelihood estimation, the method of moments has the advantage of being efficient in terms of calculation effort.

More specifically, we model the duration of distracting events using the log-normal distribution. A similar approach was pursued by (Yang and Peng 2009), who found this to be the best fit based on empirical data gathered from the road-departure crash-warning system field operational test (LeBlanc 2006). The distribution parameters μ and σ, that are the mean and standard deviation of the natural logarithm of a log-normally distributed random variable, respectively, are thereby strongly dependent on the type of distraction being modeled, and are estimated as outlined by Equation (3),

$$\mu = \log\left(\frac{m^2}{\sqrt{v + m^2}}\right), \qquad \sigma = \sqrt{log\left(\frac{v}{m^2} + 1\right)} \qquad (3)$$

where m and v correspond to the mean and variance of the empirical data set, respectively. This calculation is performed every time the upstream Poisson process triggers an arrival of a distractive event. Subsequently, the target vehicle provided by the TVE is set to the distracted state for the estimated amount of time.

At this point we want to emphasize that in the scope of this work the method of moments has also been applied for other distributions which are entirely characterized by the first two central moments (e.g. the Gamma distribution), however, the estimates derived thereof were partially outside of the distributions' parameter space and, thus, not reliable.

MODEL VALIDATION

In this section we validate our model extensively with the aid of empirical data gathered in the naturalistic driving study presented in (Stutts et al. 2003). Subsequently, we show that the proposed model allows to reproduce the study results to a satisfying extent using different simulation setups.

Simulation Setup

For validation purposes we have chosen a scenario consisting of an identical setup as in our reference study. Thus, we simulated three hours of driving for 70 subject vehicles being exposed to altogether 17 types of potentially distracting activities. Each distraction type is parametrized in accordance to the empirical observations. For a full list of model parameters for the individual distraction types we refer to the findings in (Stutts et al. 2003). Moreover, we investigated the very same scenario considering a significantly higher number of vehicles. All simulations carried out in the scope of this work have been performed using the microscopic simulation framework TraffSim (Backfrieder et al. 2015).

Simulation Results

We have simulated the abovementioned scenario under consideration of $n \in \{70, 200, 1000\}$ vehicles, whereby each scenario was simulated 1000 times in order to generate statistically reliable results. Hereinafter, we confront the simulation results with the reference data in (Stutts et al. 2003) under various aspects. A summary of these findings including the average relative error for the parameters of interest can be found Table 1.

Table 1: Relative Error [%] between Simulated Quantities and Reference Data

Parameter	δx (n=70)	δx (n=200)	δx (n=1000)
Exposure	3,217	2,642	2,294
Frequency	1,579	1,125	4,503
$\Sigma\, t$	2,073	1,339	4,793
t	1,397	0,642	0,359
sd[t]	7,678	4,491	3,158

Exposure
The engagement of drivers in potentially distracting activities is an important figure when studying the large-scale effects of distracted driving. By randomly allocating vehicles to one or multiple of predefined types of distraction, the TVE ensures that not more drivers than expected are exposed to a particular distraction type. In that manner, we were able to reproduce the empirically determined values with a relative error of just 3.2% for the comparative scenario with n=70 vehicles. Increasing the number of vehicles to 200 and 1000 improves the situation even further to 2.6% and 2.3%, respectively. This can simply be attributed to the fact that with a growing number of vehicles the probability that the TVE selects a particular vehicle more often declines similarly.

Frequency of Occurrence
The number of occurrences is modeled using stochastic Poisson processes which determine the inter-arrival time between two consecutive distractive events. For the baseline scenario (n=70) this approach turned out to perform in a satisfactory manner, yielding a relative error in the range of 0,03% and 3,3%, or 1.58% on average. For the two remaining scenarios with an increased number of vehicles our findings are not of high significance, as the simulated figures can only be confronted with extrapolated and, thus, estimated values under consideration of the size of the study population.

Total Duration
The total duration denotes the aggregated amount of time of all drivers which engaged in a particular distracting activity throughout the observation period. For the comparative scenario with n=70 vehicles we were able to reproduce the observed values with a relative error of approximately 2%. Similar to the number of occurrences we are not able to confront this absolute quantity with a comparable observed value for the two remaining scenarios in a trustworthy manner, simply due to the limited size of the study population.

Average Duration and Standard Deviation
Another important aspect when studying the impacts of distracted driving is the amount of time drivers tend to engage in a distracting task. We model the latter using log-normally distributed random variables, whereby the distribution parameters are derived from the reference study using a method of moments estimation. Despite the simplicity of our approach and the limited knowledge about the actual distribution of the empirical values we were able to emulate comparatively appropriate values for both the average duration and its standard deviation. More precisely, the relative error for the average duration compared to the actual observed values is as little as 1.39%, 0.64% and 0.36% for the scenarios with 70, 200 and 1000 vehicles, respectively. With regard to the standard deviation, that error turns out to be somewhat larger, i.e. 7.68%, 4.49% and 3.15% for all three scenarios.

Minimum and Maximum Duration
As one might have noticed, the minimum and maximum duration of a certain type of distraction has not been taking into account in the model development process. This is simply due to the fact that the model was designed in such a way as to reproduce the average statistical quantities of an empirical data set rather than the corresponding lower and upper bounds. Notwithstanding, we were able to observe that the duration of distractive events lies within these limits in more than 97% of all cases for all three scenarios.

CONCLUSIONS AND FUTURE WORK

In this paper we have presented an integrative approach towards modeling driver distractions in the scope of microscopic traffic simulations. The proposed model allows for the simulation of multiple types of distractive activities simultaneously and makes use of stochastic processes to create distraction profiles for every single simulated vehicle. We validated the model with the aid of empirical data collected in a naturalistic driving study (Stutts et al. 2003) by means of dynamic simulations using the microscopic traffic simulator TraffSim (Backfrieder et al. 2015). The results reveal that the proposed model is able to reproduce the statistical characteristics derived from these data to a satisfying extent, with a relative error in the range of just a few percent for all investigated parameters. This is not only an indicator that the proposed model constitutes a proper starting point for studying the large-scale impacts of distracted driving in a simulation environment, but also moves simulation results closer to reality.
Given the proposed model as a starting point, future work could contain investigating the repercussions of different types of driver distraction on traffic efficiency (e.g. travel times, fuel consumption) and traffic flow dynamics (e.g. collective stability, accident frequency) on a larger scale. This, in turn, requires the development of new traffic models which allow to model the effects of distracted driving (e.g. reduced speed, increased reaction time) at the microscopic level. Moreover, future work could also contain to find an even more appropriate statistical distribution for modeling the duration of distractive tasks in compliance with data gathered from naturalistic or driving simulator experiments.

ACKNOWLEDGEMENT

The authors greatly acknowledge the support by the Austrian Research Promotion Agency (FFG) in the scope of the program “Industrienahe Dissertationen”.

REFERENCES

Ahmed K., 1999. “*Modeling Drivers Acceleration and Lane Changing Behavior.”* Ph.D. thesis, Massachusetts Institute of Technology.

Andersen G.J. and Sauer C.W., 2007. “Optical Information for Car-Following: The Driving by Visual Angle (DVA) Model.” *Human Factors* 49, No. 5 , 878-896.

Backfrieder C.; Ostermayer G.; and Mecklenbräuker C.F., 2015. "TraffSim – A Traffic Simulator for Investigations of Congestion Minimization through Dynamic Vehicle Rerouting." *International Journal of Simulation, Systems, Science and Technology* 15, No. 4, 38-47.

Bando M.; Hasebe K.; Nakanishi K.; and Nakayama A., 1998. "Analysis of Optimal Velocity Model with Explicit Delay." *Physical Review* E 58, No. 5429, 1035-1042.

Beede K.E. and Kass S.J., 2006. "Engrossed in Conversation: The Impact of Cell Phones on Simulated Driving Performance." *Accident Analysis and Prevention* 38, No. 2, 415-421.

Behrisch M.; Bieker L.; Erdman J.; and Krajzewicz D., 2011. „SUMO – Simulation of Urban Mobility – An Overview." In *Proceedings of the 3rd International Conference on Advances in System Simulation*, Barcelona, Spain, 55-60.

Bowman K.O. and Shenton L., 2006. "Estimation: Method of Moments." *Encyclopedia of Statistical Sciences* 3, 2092-2098.

Cooper J.M.; Vladisavljevic I.; Medeiros-Ward N.; Martin P.T.; and Strayer D.L., 2009. "An Investigation of Driver Distraction Near the Tipping Point of Traffic Flow Stability." *Human Factors* 51, No. 2, 261-268.

Dingus T.; Klauer S.; Neale V.; Peterson A.; Lee S.; Sudweeks J.; Perez M.; Hankey J.; Ramsey D.; Gupta S.; Bucher C.; Doerzaph Z.; Jarmeland J.; and Knipling R., 2006. "*The 100-Case Naturalistic Driving Study, Phase II – Results of the 100-Car Field Experiment*." Tech. rep., National Highway Traffic Safety Administration, Washington, D.C., USA.

Dingus T.; Guo F.; Lee S.; Antin J.F.; Perez M.; Buchanan-King M.; and Hankey J., 2016. "Driver Crash Risk Factors and Prevalence Evaluation using Naturalistic Driving Data." *Proceedings of the National Academy of Sciences* 113, No. 10, 2636-2641.

Fellendorf M., 1994. "VISSIM: A Microscopic Simulation Tool to Evaluate Actuated Signal Control Including Bus Priority." In *Proceedings of the 64th ITE Annual Meeting*, Dallas, TX, USA, 1-9.

Fitch G.M.; Soccolish S.A.; Guo F.; McClafferty J.; Fang Y.; Olson R.L.; Perez M.; Hanowski R.J.; Hankey J.; and Dingus T., 2013. "*The Impact of Hand-Held and Hand-Free Cell Phone Use on Driving Performance and Safety-Critical Event Risk*." Tech. rep., National Highway Traffic Safety Administration, Washington, D.C., USA.

Gipps P., 1981. "A Behavioural Car-Following Model for Computer Simulation." *Transportation Research Part B: Methodological* 15, No. 2, 105-111.

Haenggi M.; Andrews J.G.; Baccelli F.; Dousse O.; and Franceschetti M., 2009. "Stochastic Geometry and Random Graphs for the Analysis and Design of Wireless Networks." *IEEE Journal on Selected Areas in Communications* 27, No. 7, 1029-1046.

Hendricks D.; Fell J.; and Freedman M., 2001. "*The Relative Frequency of Unsafe Driving Acts in Serious Traffic Crashes: Summary Technical Report*." Tech. rep., National Highway Traffic Safety Administration, Washington, D.C., USA.

Jin S.; Wang D.H.; Huang Z.Y.; and Tao P.F., 2011. "Visual Angle Model for Car-Following Theory." *Phyisca A: Statistical Mechanics and Its Applications* 390, No. 11, 1931-1940.

LeBlanc D., 2006. "*Road Departure Crash Warning System Fiel Operational Test: Methodology and Results. Volume 1: Technical Report*." Tech. rep., University of Michigan, Ann Arbor, MI, USA.

Lee J.; Young K.; and Regan M., 2008. "*Driver Distraction: Theory, Effects and Mitigation*.", CRC Press, Florida, USA.

McKeever J.D.; Schultheis M.T.; Padmanaban V.; and Blasco A., 2013. "Driver Performance While Texting: Even a Little is Too Much." *Traffic Injury Prevention* 14, No. 2, 132-137.

Newell G.F., 1961. "Nonlinear Effects in the Dynamics of Car Following." *Operations Research* 9, No. 2, 209-229.

National Highway Traffic Safety Administration, 2015. "*Traffic Safety Facts 2015: A Compilation of Motor Vehicle Crash Data from the Fatality Analysis Reporting System and the General Estimates System*." Tech. rep., National Highway Traffic Safety Administration, Washington, D.C., USA.

Nourzad S.H.H.; Salvucci D.D.; and Pradhan A., 2014. "Computational Modeling of Driver Distraction by Integrating Cognitive and Agent-Based Traffic Simulation Models." *Computing in Civil and Building Engineering* 2014, 1885-1892.

Owens J.M.; McLaughlin S.B.; and Sudweeks J., 2011. "Driver Performance While Text Messaging Using Handheld and In-Vehicle Systems." *Accident Analysis and Prevention* 43, No. 3, 939-947.

Pipes L.A., 1953. "An Operational Analysis of Traffic Dynamics." *Journal of Applied Physics* 24, No. 3, 274-281.

Redelmeier D.A. and Tibshirani R.J., 1997. "Association Between Cellular-Telephone Calls and Motor Vehicle Collisions." *New England Journal of Medicine* 336, No. 7, 453-458.

Regan M.A., 2004. "New Technologies in Cars: Human Factors and Safety Issues." *Ergonomics Australia* 8, No. 3, 6-15.

Salvucci D.D., 2009. "Rapid Prototyping and Evaluation of In-Vehicle Interfaces." *ACM Transactions on Computer-Human Interaction (TOCHI)* 16, No. 2, 9.

Salvucci D.D., 2013. "Distraction Beyond the Driver: Predicting the Effects of In-Vehicle Interaction on Surrounding Traffic." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, New York, USA, 3131-3134.

Stavrinos D.; Jones J.L.; Garner A.A.; Griffin R.; Franklin C.A.; Ball D.; Welburn S.C.; Ball K.K.; Sisiopiku V.P.; and Fine P.R., 2013. "Impact of Distracted Driving on Safety and Traffic Flow." *Accident Analysis and Prevention* 61, 2013, 63-70.

Strayer D.L.; Drews F.A.; and Crouch D.J., 2006. "A Comparison of the Cell Phone Driver and the Drunk Driver." *Human Factors* 48, No. 2, 381-391.

Stutts J.; Feaganes J.; Rodgman E.; Hamlett C.; Meadows T.; Reinfurt D.; Fish K.; Mercadante M.; and Staplin L., 2003. "*Distractions in Everyday Driving*", Tech. rep., AAA Foundation for Traffic Safety, Washington, D.C., USA.

Taylor H.M. and Karlin S., 2014. "*An Introduction to Stochastic Modeling*." Academic Press.

Tijms H., 2003. "A First Course in Stochastic Models." John Wiley and Sons.

Treiber M.; Hennecke A.; and Helbing D., 2000. "Congested Traffic States in Empirical Observations and Microscopic Simulations." *Physical Review E* 62, No. 2, 1805-1824.

Treiber M.; Kesting A.; and Helbing D., 2006. "Delays, Inaccuracies and Anticipation in Microscopic Traffic Models." *Physica A: Statistical Mechanics and its Applications* 30, No. 1, 71-88.

Wang J.S.; Knipling R.R.; and Goodman M.J., 1996. "The Role of Driver Inattention in Crashes: New Statistics from the 1995 Crashworthiness Data System." *40th Annual Proceedings of the Association for Advancement of Automotive Medicine* 377, 392.

Wilson A.; Fang M.; Wiggins S.; and Cooper P., 2003. "Collision and Violation Involvement of Drivers who use Cellular Telephones." *Traffic Injury Prevention* 4, No. 1, 45-52.

Wilson A. and Stimpson J.P., 2003. "Trends in Fatalities from Distracted Driving in the United States, 1999 to 2008." *American Journal of Public Health* 100, No. 11, 2213-2219.

Yang S. and Peng H., 2009. "Development of an Errorable Car-Following Driver Model.", *Vehicle System Dynamics* 48, No. 6, 751-773.

Young K.; Regan M.; and Hammer M., 2007. "Driver Distraction: A Review of the Literature." *Distracted Driving* 2007, 379-405.

UTILISATION OF COMPUTER SIMULATION FOR DYNAMIC CALCULATION OF TRAIN DELAYS

Jan Fikejz and Josef Brožek
Department of Software Technologies
Faculty of Electrical Engineering and Informatics
University of Pardubice
Studentská 95, CZ-532 10 Pardubice
Czech Republic
E-mail: Jan.Fikejz@upce.cz, Josef.Brozek@upce.cz

KEYWORDS

Railway infrastructure models, train positioning, simulation, train delay, web services

ABSTRACT

This article deals about utilisation of computer simulation for dynamic calculation of train delays within the rail network model and satellite navigation. Attention is firstly aimed on the description of the trains position in the designed railway network model. Further attention is focused on the design of the dynamic calculation of train delays with utilization of reduced track profile and using computer simulation.

INTRODUCTION

Precise dynamic calculation of train delay is not exactly trivial and includes a wide variety of aspects affecting the resulting journey time. Such are both, technical parameters of (i) a train set (e.g. train acceleration or overall train set weight) and (ii) railway infrastructure (character of the line, velocity limits), and (ii) external influences like weather (e.g. temperature, weather conditions, visibility). Delay calculation is always dependant on the current train location in the railway network in given time, thus current information about location is an inseparable part of dynamic calculation of delay. In this case, it is not necessary to demand such accuracy as for safety systems to determine location and thus it is possible to use satellite navigation (GNSS – Global Navigation Satellite System) for finding the current location of a train.

POSSIBLE TYPES OF LOCALIZATION

Generally, localisation is prone to a wide range of approaches on how to identify the position of trains on a track. Put simply, localisation may be divided into the following three groups:

- Localization without the use of GNSS,
- GNSS using localization,
- GNSS-based, involving further support systems.

Trains localization without the use of GNSS

This type of trains localization often requires complementing the rail network infrastructure with additional construction elements, which entails higher costs of the actual implementation. On the other hand, this type of localization shows a high accuracy and reliability and is often used in the railway signalling technology. Essentially, it relates to the system of:

- ETCS (Ghazel 2104; Lieskovský and Myslivec 2010),
- Automatic train control (Chudacek and Lochman 1998; Lieskovský 2004),
- Track circuits (Dorazil 2008),
- RFID.

Trains localization using GNSS

When using GNSS for various application levels, it is necessary to take an indicated position error into consideration. Indicated position error is generally based on the nature of the satellite navigation. If we use systems that operate with the position information on an informative level only, we can tolerate a certain error; however, such inaccuracy is unacceptable in the railway signalling technology. However, various additional systems can be implemented to eliminate the error (completely or at least partially), thus making the position of the tracked object more accurate. The following systems can be listed in this group:

- EGNOS (Senesi 2012),
- Differential GPS (O'Connor 1997).

GNSS based localization involving additional support systems

As mentioned above, precise localization of trains using GNSS, especially for the needs of signalling technology, is a priori impossible. Nevertheless, the position of a rail vehicle can be determined significantly more precisely with the use of additional systems. This concern especially

solutions using inertial systems (Standlmann 2006), but also less known systems such as those based on GNSS and contactless eddy current measurement (Becker and Poliak 2008).

RAILWAY NETWORK MODEL

Undirected graph, as defined graph theory, is a natural candidate for a railway network model. Based on an analysis of data provided by the company SŽDC-TUDC (consisting of service regulations, passports, and codebooks), sets of algorithms were subsequently created, with which it was possible to generate a three-layer model of the rail network (Fikejz and Kavička, 2011). Roughly speaking, the track can be divided into individual so called supertracks, which consist of definition supra-sections (TDNU), where each supra-section contains track definition sections (TUDU) with mileposts (in hectometres). Basic aspects of the description of the rail network are collectively shown in Figure 1.

Mileposts (in hectometres) are shown in Figure with the distance in kilometres and are graphically represented using gray points. TUDU is recorded using a six-digit code (163105, 163106, 16307, 173202) and are graphically represented using solid lines (red, black, orange, brown). Individual supra-sections (CLS 007, CLS008, REG023) are shown in light blue and supertracks (B421021 1 and B421021 1A) are shown in dashed lines. A place significant in terms of transportation (branch line) is symbolized by a green square.

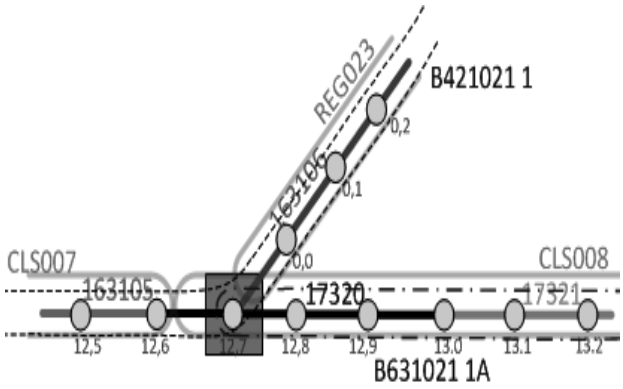


Figure 1: Basic aspects of the description of the rail network

The algorithm of railway network model (Fikejz and Kavička, 2011; Fikejz and Řezanina 2014.) was implemented directly on the database level using PL/SQL language. However, the algorithm had to be adjusted and generalized several times since there are various nonstandard conditions in the data, such as jumps in the mileposts (nonlinear growth of the kilometre succession between the mileposts) or change of an increasing kilometre sequence into a decreasing one and vice versa. The final model includes three data layers:

- **Data-Micro**, consisting of vertices and edges,
- **Data-Mezo**, include mezo-vertices and mezo-edges
- **Data-Macro**, containing super-vertices and super-edges.

Figure 2 presents the overall concept of a complete three-layer railway network model.

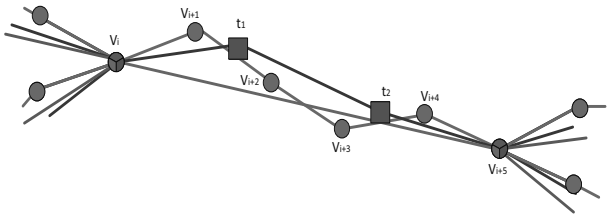


Figure 2: Illustration overall concept of a three-layer module

LOCALIZATION

The idea of trains localisation access to tracks is based on the correct pairing up of GPS information on position, provided by communication terminals, with the nearest vertex or edge of the graph. The discovered vertex/hectometre post disposes not only of a multi-dimensional key in the form of a GPS coordinate, it is also linked, through definition sections, to further information concerning the railway network infrastructure.

View of the situation that the model of railway infrastructure is stored in the database Oracle we can use the native database functions and operators. The `SDO_NN` (*nearest neighbor*) operator was selected in view of realising this unique trains localisation approach. The aforementioned operator searches for a geometric object that is closest to the object entered (like a point, for example). In other words, it is possible to find the nearest vertex, or more precisely edge in a model, from the current position trains, Figure 3.

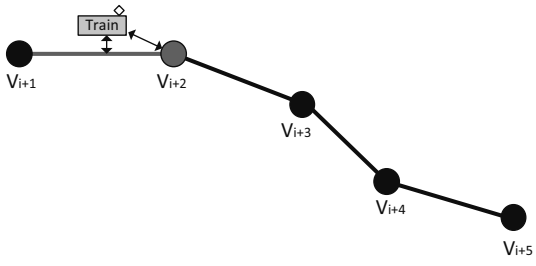


Figure 3: Main concept of localization

The actual detection of the current position of the trains can be divided into the following steps:

1. **Finding the nearest vertex and edge of the grap**h – from the current position of the trains given the three-layer railway network model

2. **Assessment of the relevancy of incoming GPS information from the communication terminal** – verification whether the current position is not burdened by a disproportionate error (like, for example, that the distance of the trains from the nearest vertex/edge is a mere few meters or tens of metres, or that the trains is still assigned to the same super-edge, provided that it should still be located on it)
3. **Calculation of the exact position of the trains on the edge of the mod**el – using perpendicular projection of the point (current trains position) onto the line

The trains position data are collected from the communication terminals. These communication terminals sent position information to the central database every 30 second.

REDUCED TRACK PROFILE

Railway infrastructure is rather varied and contains many areas affecting the train dynamic (Bandžuch 2006). For common traction calculations, it is possible to substitute the real track profile by a reduced set of substitute gradients, so called line resistance, which includes:

- gradient resistance
- curve resistance
- tunnel resistance

Curve resistance $\boldsymbol{sl_c}$ is substituted by fictive incline, for which the curve radius R<300 m for secondary and regional lines is defined by:

$$sl_c = \frac{500}{R-30}\ [‰] \qquad (1)$$

Reduced incline $\boldsymbol{sl_r}$ is then defined by:

where:

- $\boldsymbol{sl_1}$**to** $\boldsymbol{sl_k}$ -actual gradient in per mille (incline „+“, decline „-“)
- $\boldsymbol{sl_{c1}}$ **to** $\boldsymbol{sl_{cm}}$ -fictive gradient, substituting set of curves
- $\boldsymbol{l_1}$**to** $\boldsymbol{l_k}$ **-**length of gradients s_1 to s_k in meters
- $\boldsymbol{l_{c1}}$**to** $\boldsymbol{l_{cm}}$ -curve lengths 1 to m

$$sl_r = \frac{sl_1 l_1 + sl_2 l_2 + sl_3 l_3 + \cdots + sl_k l_k + sl_{c1} l_{c1} + \cdots + sl_{cm} l_{cm}}{l_1 + l_2 + \cdots + l_k}\ [‰] \qquad (2)$$

While condition
$(l_1 + l_2 + \cdots + l_k) \leq 2{,}5\ (l_{c1} + l_{c2} + \cdots + l_{cm})$ must be valid

An example of a reduced track profile is illustrated on the following Figure 5, in which arrows show the intended direction of movement of the train. The data in tables then depict:

- **kilometric position** of a change in profile $[km]$
- **the direction** where:
  - „+“ expresses change in direction
  - „−“ expresses change opposite the direction
- **track resistance** [‰] where:
  - positive value expresses track incline
  - negative value expresses track decline

TRAIN ACCELERATION

Train acceleration is influenced mainly by

- technical parameters of a locomotive (acceleration)
- overall weight of the train set

Using a special software tool for simulation train dynamics developed at University of Pardubice (Diviš and Kavička 2015), a set of measuring simulations focused on acceleration and breaking deceleration of trains was conducted for selected trains for the following track resistances:

- 0,5%
- 1,0 %
- 1,5 %
- 2,0 %

Assessment of data from individual measurings has shown that data can be approximated by a linear equation for other calculations.
If we consider a train set with the following parameters:

- Acceleration: 0,6497725 m/s^2
- Length: 96 m
- Weight: 234 t

then individual measured values of acceleration can be reflected in a graph on Figure 6.

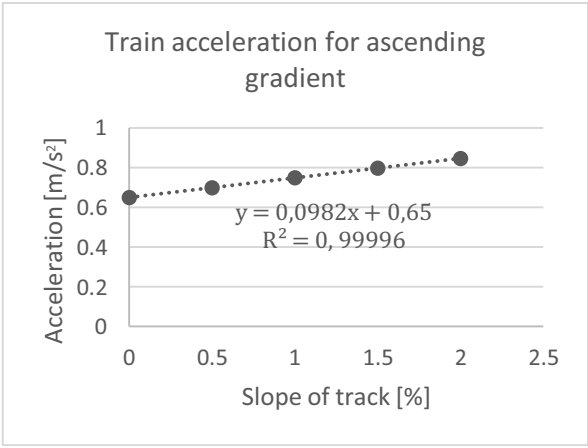


Figure 4: Example of train acceleration

From the graph above, it is clear that measured acceleration for incline can be approximated by a linear equation

$$y = 0{,}0982x + 0{,}65 \qquad (3)$$

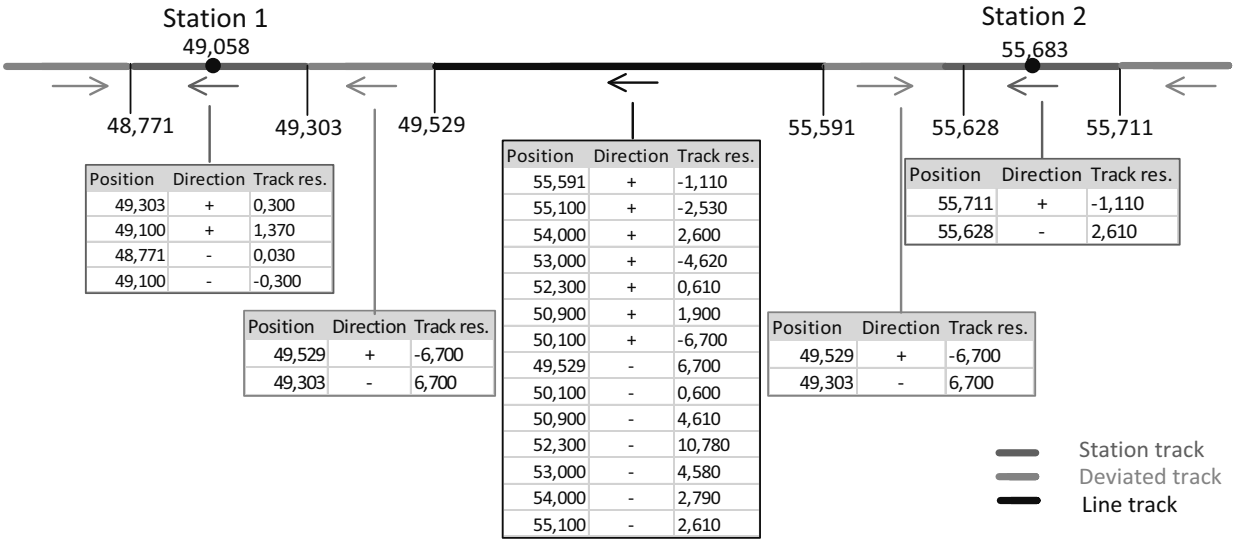
for reliability $R^2 = 0{,}99996$.

When data from measured acceleration for a declining track are applied, then through approximation we can achieve linear equation

$$y = -0{,}1044x + 0{,}6511 \qquad (4)$$

The algorithm itself is divided into 5 parts.

1. iterative calculation calculates the time $t_{acceleration}$ and path $s_{acceleration}$ needed to reach the required velocity v_{max}



Position	Direction	Track res.
49,303	+	0,300
49,100	+	1,370
48,771	-	0,030
49,100	-	-0,300

Position	Direction	Track res.
49,529	+	-6,700
49,303	-	6,700

Position	Direction	Track res.
55,591	+	-1,110
55,100	+	-2,530
54,000	+	2,600
53,000	+	-4,620
52,300	+	0,610
50,900	+	1,900
50,100	+	-6,700
49,529	-	6,700
50,100	-	0,600
50,900	-	4,610
52,300	-	10,780
53,000	-	4,580
54,000	-	2,790
55,100	-	2,610

Position	Direction	Track res.
49,529	+	-6,700
49,303	-	6,700

Position	Direction	Track res.
55,711	+	-1,110
55,628	-	2,610

Figure 5 : Example of reduced track profile

for reliability $R^2 = 0{,}99974$.

CALCULATING DELAY

For dynamic calculation of delay, it is possible to base the calculation on a simplified model, in which the overall journey time between two stations is given by adding their three parts (Figure 7):

- acceleration period to set velocity
- journey period in set velocity
- breaking period from set velocity to zero velocity

Figure 6: Concept of calculation

$$t_{total} = t_{accel.} + t_{\text{journey}} + t_{decel.} \qquad (5)$$

while it is supposed that acceleration and breaking are affected by track profile, however, the train has sufficient performance power to keep required velocity on route between the two track segments.

Calculation algorithm

Delay calculation algorithm uses general formulae for evenly accelerated linear movement, i.e. relations based on:

$$v = v_0 + at \implies t = \frac{v - v_0}{a} \qquad (6)$$

$$s = v_0 t + \frac{1}{2}at^2 \implies v^2 = 2as + {v_0}^2 \qquad (7)$$

2. iterative calculation calculates the time $t_{deceleration}$ and path $s_{deceleration}$ needed to reach zero velocity from velocity v_{max}
3. path of train movement in velocity v_{max} is calculated based on the difference between the overall path s_{total} and path for acceleration $s_{acceleration}$ and breaking $s_{deceleration}$, i.e.

$$s_{\text{journey}} = s_{total} - s_{accel.} - s_{decel.} \qquad (8)$$

4. the time t_{journey} of train movement in set velocity v_{max} is calculated
5. overall time t_{total} is calculated from formula 5

So, if we consider maximal velocity of a train on a regional line v_{max} (for example $65\frac{km}{h}$) and track according to figure 9, then the first part of the algorithm (for calculating $t_{acceleration}$ and $s_{acceleration}$) will conduct individual iterative calculations in points of gradient change on the track sl_i [‰] defined by kilometric location d_i [km]. If the current velocity of the train v_i is higher or equal to the required velocity v_{max}, then the iterative calculation is terminated and subsequently the exact time t_{last} and path s_{last}, when the train reached the required velocity v_{max} are calculated. For starting the calculation, a corresponding railway station $d_{station}$, kilometric location of which is given, is considered.

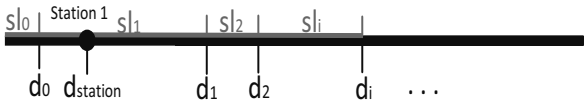


Figure 7: Concept of the acceleration algorithm

Concept of the algorithm for train acceleration to the required velocity v_{max}

1. $i = 1; v_0 = 0$
2. **if** $i = 1$
   then $s_i = d_i - d_{station}$
   else $s_i = d_i - d_{i-1}$
3. **if** $sl_i > 0$
   then $a_i = 0{,}0982 sl_{i-1} + 0{,}65$ (eq. 3)
   else $a_i = -0{,}1044 sl_{i-1} + 0{,}6511$(eq. 4)
4. $v_i = \sqrt{v_{i-1} + 2a_i + s_i}$
5. **if** $v_i < v_{max}$
   then $i = i + 1$ and go to step 2
6. $s_{last} = \frac{{v_{max}}^2 - {v_i}^2}{2a_i}$
7. $t_{last} = \frac{v_{max} - v_i}{a}$
8. $t_{toVmax} = \sum_{k=1}^{i-1} t_k + t_{last}$
9. $s_{toVmax} = \sum_{k=1}^{i-1} s_k + s_{last}$

Analogously, the breaking time $t_{deceleration}$ and the path $s_{deceleration}$ needed to stop the train at the station are calculated accordingly. The relation 8 determines the length of the line on which the train moves at a constant velocity and consequently from the relation

$$t = \frac{v}{s} \tag{9}$$

the journey length t_{journey} is calculated. Overall journey time t_{total} is then given by the sum of the partial times per the relation 5.
From the times between the individual railway stations t_{total}. and the time needed for the train to be serviced in the station $t_{service}$, it is then possible to calculate the total time of the journey to the required station.

For dynamic calculation of the delay of a moving train on the track, it is possible to calculate the time of journey/delay to the next railway station or to the selected station on the track from the knowledge of its current position (from the railway network model).

WEB SERVICES

For the dynamic calculation of train delays, web services were designed. These web services provide a basic set of information about:

- position of train
- distance form/to nearest train station
- travel time form/to nearest train station considering to actual position and track profile
- delay of train considering timetabling of trains

The main advantage of this approach is the hiding of the application logic of the localization mechanism from the final application. In the JAVA environment, we can use two different approaches. Their main difference is in the internal request processing, and in their architecture.

Design and implementation of web services
For the approach the set of the localization methods were prepared, which performed selected localization tasks. It possible to use these methods, which according to the train number, is able to find out the position of selected railway vehicle on the railway network and calculate the travel time and delay to the next station. Using the JSON protocol, this position data is then returned to the client in the final application. The concept of use of a web service is shown in Figure 9.

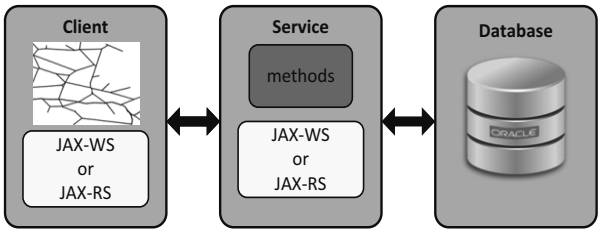


Figure 8: Concept of communication

SIMULATION OF DELAYED TRAINS

The selected train vehicles are equipped with communication terminals, which broadcast data including current GPS coordinates of the trains. When the vehicle is in motion, this communication terminal sends information about its position every 30 seconds.

Designed simulation model contains the core of discrete simulation utilizing standard calendar of process messages, which were, during the simulation, executed based on their time stamp Complete calculations for the dynamic calculation of the delay were subsequently implemented in the *infraRail* software tool.

For the simulation of variants of delayed trains are used:

- Real historical data (emulation of operation),
- Generated data.

The running application captures the current position of the train on the track with a set of information related to its position including the calculation of the current arrival delay to the next railway station, Figure 10.

CONCLUSION

The focus was on the proposal for dynamic calculation of train delays using the rail network model and satellite navigation. A multi-layered model of the railway network was designed reflecting the non-oriented graph. In addition, the algorithm was used to identify the position of trains in the railway network. This algorithm includes the search of the previous or next railway station. The article was also focused on description of the reduced track profile which was used for designs of the algorithm for dynamic calculation of train delays. Proposed algorithms have been implemented in the InfraRail software tool. Discrete simulation was used to test other variants of delayed trains, both based on historical and generated data.

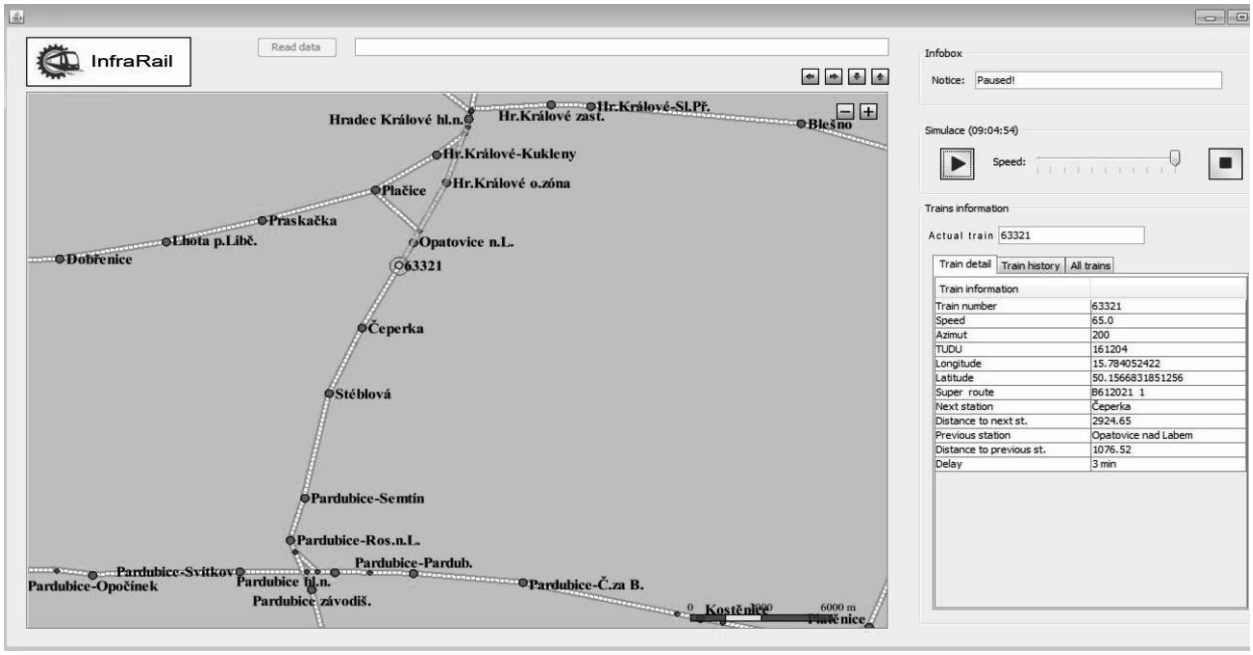


Figure 9: Running application

ACKNOWLEDGEMENTS

This work has been supported by the project "SGS_2017 Models of infrastructure and operation of land transport systems" (financed by the University of Pardubice).

REFERENCES

Becker, U. and J. Poliak. DemoOrt repositions trains with satellite. In: EURAILmag Business & Technology. 18. BLUE LINE & Bro, France, 2008,s. 216-219.

Chudaček, V. and L. Lochman. Vlakový zabezpečovací systém ERTMS/ETCS. In: Vědeckotechnicky sborník ČD, č. 5/1998

Dorazil, P. Základní vlastnosti kolejových obvodů bez izolovaných styků. Pardubice, 2008. Bachelor thesis. University of Pardubice. Supervisor: Milan Kunhart.

Fikejz, J. and A. Kavička. Modelling and simulation of train positioning within the railway network. In: KLUMPP, Matthias. ESM'2012. The European simulation and modelling conference. Ostende: EUROSIS - ETI, 2012, s. 366 -376. ISBN 978-9077381-73-1.

Fikejz, J. and A. Kavička. Utilisation of computer simulation for testing additional support for dispatching rail traffic. In: European Simulation and Modelling Conference, 2011. Ostende: EUROSIS - ETI, 2011. p. 225-231. ISBN 978-90-77381-66-3.

Fikejz, J. and E. Řezanina, Utilization of computer simulation for detection non-standard situations within the new data layer of railway network model. In: The 26th European Modeling & Simulation Symposium. Bordeaux, 2014 s. 371-377, ISBN 978-88-97999-32-4

Ghazel, M. Formalizing and subset of ERTMS/ETCS specifications for verification purposes. In:Transportation Research Part C: Emerging Technologies. Elsevier Limited, 2014, pp. 60-75 ISSN: 0968-090X

Kothuri, R. et al. Pro Oracle Spatial for Oracle database 11g. New York, NY: Distributed to the book trade worldwideby Springer-Verlag New York, c2007, xxxiv, 787 p. ISBN 15-905-9899-7.

Lieskovský, A. and I. Myslivec. ETCS a AVV - poprvé společně. In: EuroŽel, Žilina, 2010

Lieskovský, A. Automatické vedení vlaků Českých drah. In: Automatizace. Praha: Automatizace, 2004, roč. 10. ISSN 0005-125x.

O'connor, M. L. Carrier-phase differential gps for automatic control of land vehicles, In: Dissertation Abstracts International, Volume: 59-06, Section: B, page: 2876.; 158 p. 1997, Stanford University, ISBN: 9780591909272

Senesi, F. Satellite application for train control systems, In: The Test Site in Sardinia, Journal of Rail Transport Planning and Managemt. Elsevier BV, 2012, s. 73-78, ISSN:2210-9706

Bandžuch, Ľubomír. Modernizácia elektrifikovanej trate v rámci V. koridoru v úseku Košice – Poprad. Žilina, 2006. Diplomová práce. ŽILINSKÁ UNIVERZITA V ŽILINE. Vedoucí práce Doc. Ing. Gabriela Lanáková, PhD.

Diviš, R. and A, Kavička. Design and development of a mesoscopic simulator specialized in investigating capacities of railway nodes. In: The 27th European Modeling and Simulation Symposium (EMSS 2015): 12th International Multidisciplinary Modeling and Simulation Multiconference (I3M 2015). 1. Rende, Italy: CAL-TEK S.r.l, 2015, s. 52-57. ISBN 978151081376

HYBRID OPTIMIZING MODELS FOR PLANNING CHARGING INFRASTRUCTURES

Hubert Büchter
Fraunhofer-Institute for Material Flow and Logistics IML
Joseph-von-Fraunhofer-Str. 2-4
44227 Dortmund
Germany
E-mail: hubert.buechter@iml.fraunhofer.de

Sebastian Naumann
ifak - Institut f. Automation und Kommunikation e.V.
Werner-Heisenberg-Str. 1
39106 Magdeburg
Germany
E-mail: sebastian.naumann@ifak.eu

KEYWORDS

Optimization, Linear Programming, Combined Simulation, Transportation, Linear Model

ABSTRACT

Electric buses in public transport are increasingly being put into operation. Although batteries with large capacities are available, a net of charging stations can avoid peaks on the power grid. A properly designed charging infrastructure is a contribution to keep the balance between energy offer and energy demand. Increasing the number of charging stations, each with a small amount of power, enable in many cases the usage of existing transformer stations. This avoids the installation of new transformer stations and it reduces the whole investment.
The main requirements of a charging infrastructure are minimal costs for installation and for operation as well as minimal power peaks on the power grid. The solution introduced in this paper is a hybrid optimizer based on an analytical and a simulative approach. Each method has its own power and specific advantages. The optimization procedure sequentially executes both algorithms in a loop until a stop criterion is met. For optimization and for simulation well-known algorithms are applied and the problem specific models are developed systematically. The result is a planning tool for bus companies, energy providers and urban planners.

INTRODUCTION

The operation of electric vehicles in public transport is a contribution to the reduction of environmental impacts. Running a fleet of buses requires a well-designed charging infrastructure. This becomes more important if the requirements for the electricity network are taken into account. Big batteries with an overnight charging or small batteries with a rapid charging concept produce unwanted power peaks. A load-balanced grid requires many charging stations to spread the load in space and in time.
Many publications about the planning problem for charging infrastructures deal with non-public urban transport, e.g. (Frade et al. 2011; Kuchshaus et al. 2012; Chen et al. 2013). An overview of optimal planning for charging stations is in (Zheng et al.). That work deals with stochastic models with the focus on the grid and it does not meet the specific characteristic of public transportation, which is timetable driven. The complexity of the planning process for charging infrastructure is analyzed in (Lam 2014) and it shows that it is NP complete. Due to this result, any planning algorithm of practical relevance is an approximation. (Olsen and Kliewer 2016) extend a vehicle-scheduling problem by charging models with different battery characteristics. They find that in an often used operation range a linear approximation of the battery characteristic is sufficient. (Buechter and Naumann 2016a) and (Buechter and Naumann 2016b) used this approach in an optimizer that is an early predecessor of the here introduced approach. This optimizer focused also on battery swapping and line charging but did not consider the power grid in any way. The current work described in this paper fills this gap.
(Klemmt et al. 2009) worked on a similar optimizer for scheduling problems.
The focus of this paper is on public transportation with electrically driven buses. For all buses, there are special utilized bus stops, which are equipped with charging devices. All these devices need power lines to at least one transformer station. The task is to select the bus stops, which must be upgraded to charging stations.
Figure 1 shows an example with three transformer stations, six bus stops and two busses. All transformer stations, which can supply energy to one or more potential charging stations are connected by a question mark labeled line to the corresponding bus stop.

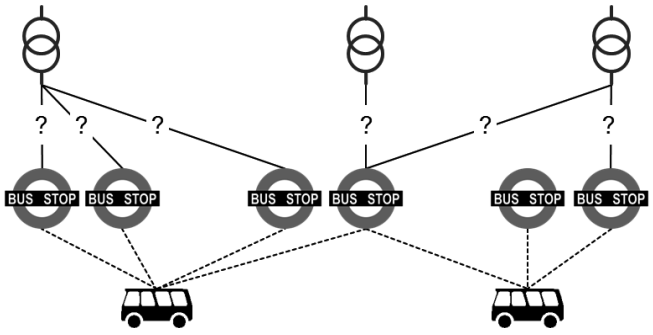


Figure 1: A Tiny Planning Problem

The planning algorithm calculates a list of all connections, which are required to keep all buses in operation. Figure 2 shows a possible result for this example. Only two transformers are required and two bus stops are upgraded to charging stations.

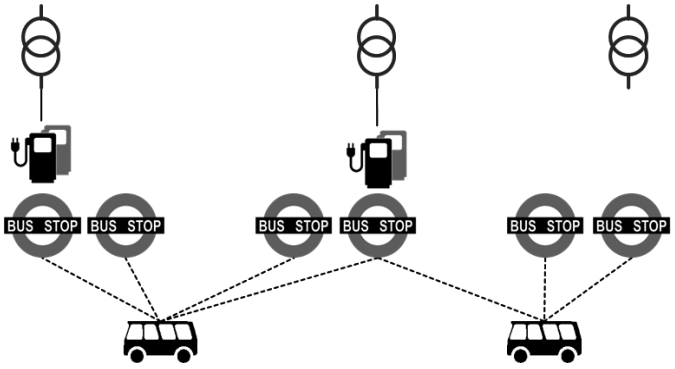


Figure 2: Solution for the Tiny Planning Problem

The planning process minimizes the costs for investment. The method is based on four different models, which are applied according to given limits and the desired results.
This paper is structured as follows: The first section introduces a linear system model. On this base, the following sections describe four different optimizing models. The next sections deal with the simulator and the iteration loop. Finally, we present some results and give an outlook to further work.

SYSTEM MODEL

The *optimization model* as well as the *simulation model* require a *system model*. The notation is in terms of linear algebra with mostly one row for each bus. The relevant physical objects for the system model are *buses*, *ways*, *bus stops* and *transformer stations*:

$$\left.\begin{array}{rll} \text{Buses} & i \in \mathbb{B}, & \mathbb{B} = \{1, \dots, l\} \\ \text{Ways} & j \in \mathbb{W}, & \mathbb{W} = \{1, \dots, m\} \\ \text{Bus Stops} & k \in \mathbb{S}, & \mathbb{S} = \{1, \dots, n\} \\ \text{Charging Stations} & & \mathbb{C} \subseteq \mathbb{S} \\ \text{Transformer Stations} & p \in \mathbb{T}, & \mathbb{T} = \{1, \dots, w\} \end{array}\right\} \quad (1)$$

If a bus stop will be equipped with a charging device, it becomes a *charging station* $\mathbb{C} \subseteq \mathbb{S}$. The estimation of $\mathbb{C}$ is not part of the system model but of the optimization model, which is introduced in the next section. The system model describes the behavior of charge of the batteries in a linear range. We assume that the power flow is constant over given periods. In this case, all energies are equal to the product of the time duration of power flow by the mean value of power:

$$\boldsymbol{e} = \int_{t_1}^{t_2} \boldsymbol{p}(\boldsymbol{t})\boldsymbol{dt} \quad \text{or} \quad \boldsymbol{e} = (\boldsymbol{t_2} - \boldsymbol{t_1}) \cdot \overline{\boldsymbol{p}} \quad (2)$$

We assume this linear battery characteristic and treat all physical values as mean values. All buses are equipped with permanently installed batteries.
Buses drive on ways and each way connects two bus stops. Details of the ways, such as junctions and traffic lights are not of interest. The *average drive time* for each way and for each bus is given in a matrix $\in \mathbb{R}^{+|\mathbb{B}|\times|\mathbb{W}|}$. The *average stop time* at each bus stop and for each bus is given in a matrix $\in \mathbb{R}^{+|\mathbb{B}|\times|\mathbb{S}|}$. The *traction energy consumption* for each way and for each bus is given in a matrix $\boldsymbol{Q} \in \mathbb{R}_0^{+|\mathbb{B}|\times|\mathbb{W}|}$. Additional to the traction energy there is a *base load* $\boldsymbol{pb_{base}}$ for e.g. air condition, lightning and heating as a power request for each bus.
The frequency with which the buses travel the individual routes is taken from the *schedule matrix for ways* $\boldsymbol{SW} \in \mathbb{N}_0^{|\mathbb{B}|\times|\mathbb{W}|}$. The frequency with which the buses stop at bus stops is taken from the *schedule matrix for stops* $\boldsymbol{SS} \in \mathbb{N}_0^{|\mathbb{B}|\times|\mathbb{S}|}$. Table 1 shows these Data structures in an overview. The operator $\circ$ is the Hadamard operator for a multiplying matrices or vectors element by element.

Table 1: Summary of Data Structures for Drive Frequencies and Stop Frequencies and for Drive Times and Stop Times

	frequency	average time	traction energy flow	base energy flow
on ways	$\boldsymbol{SW}$	$\boldsymbol{TD}$	$\boldsymbol{Q} \circ \boldsymbol{SW}$	$\boldsymbol{diag(pb_{base})} \cdot (\boldsymbol{SW} \circ \boldsymbol{TD})$
at bus stops	$\boldsymbol{SS}$	$\boldsymbol{TS}$	$\boldsymbol{0}$	$\boldsymbol{diag(pb_{base})} \cdot (\boldsymbol{SS} \circ \boldsymbol{TS})$

The *average operation time* $\overline{\boldsymbol{ot}}$ depends on these values:

$$\overline{\boldsymbol{ot}} = (\boldsymbol{SS} \circ \boldsymbol{TS}) \cdot \boldsymbol{1}_{|\mathbb{S}|} + (\boldsymbol{SW} \circ \boldsymbol{TD}) \cdot \boldsymbol{1}_{|\mathbb{W}|} \quad (3)$$

All buses operate cyclic over time. Each bus has its own *operation time* $\boldsymbol{ot}$, which is less equal than the *cycle time* $\boldsymbol{tc}$ for the entire system. Figure 3 shows an accumulated timing diagram.

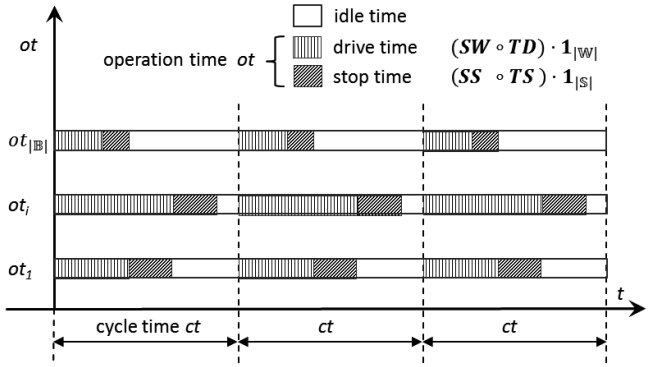


Figure 3: Accumulated timing diagram for cyclic operation. In practice drive times and stop times are interlaced

At the beginning of a cycle, the batteries have an *initial charge* $\boldsymbol{e_{ini}}$ and to the end, they must still have a *residual charge* of $\boldsymbol{e_{trm}}$ that is often smaller than the initial charge. This energy difference $\Delta\boldsymbol{e} = \boldsymbol{e_{trm}} - \boldsymbol{e_{ini}}$ can fed at a depot during the operating pause, but this is not considered here. For a complete cycle, the buses need an amount of energy that consists of the traction energy and the base energy:

$$e_i^- = \sum_{j=1}^{|\mathbb{W}|} sw_{ij} \cdot q_{ij} + pb_{base_i} \cdot ot_i - \Delta e_i \quad \text{for Bus } i \quad (4)$$

$$\boldsymbol{e}^- = diag(\boldsymbol{SW} \cdot \boldsymbol{Q}^T) + \boldsymbol{pb_{base}} \circ \boldsymbol{ot} - \Delta\boldsymbol{e} \quad \text{for all Buses} \quad (5)$$

Negative values of $\boldsymbol{e}^-$ are possible when the entire energy supply from the batteries can be provided without additional any charging of the batteries. Apart from a charging station in the depot, no charging infrastructure is required in this

case. All other cases which are in the focus of this paper require a charging infrastructure to provide an energy offer of $\boldsymbol{e^+}$. Figure 4 shows the energy level for one bus over the time.

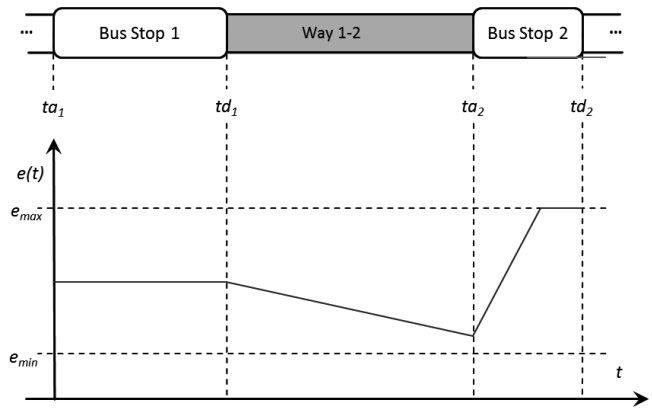


Figure 4: Energy for One Bus Arriving at Bus Stop 1 at ta_1 Starting at t_{d_1} and Driving over Way 1-2 to Bus Stop 2

The energy balance must grant while the energy offer will be as small in general in order to keep the costs low.

$$e^+ \geq e^- \tag{6}$$
$$\boldsymbol{e^+} \rightarrow \min \tag{7}$$

The charging controllers can use the energy offer full or even in part.

All charging stations must be connected to at least one transformer station that delivers the power. The total costs for the power lines are to minimize, while the power limits for the charging devices, and for the transformer stations must be observed.

A first approach for given power limits $\boldsymbol{pt_{max}}$ for the transformers, $\boldsymbol{ps_{max}}$ for the charging devices at the bus stops and the mean value for the operation time $\overline{ot}$ over all buses form necessary some conditions:

$$\left.\begin{matrix}\overline{ot} \cdot \sum_{p\in\mathbb{T}} pt_{max_p} \\ \overline{ot} \cdot \sum_{k\in\mathbb{S}} ps_{max_k}\end{matrix}\right\} \geq \sum_{i\in\mathbb{B}} e_i^+ \geq \sum_{i\in\mathbb{B}} e_i^- \tag{8}$$

The first term describes the energy offer from all transformers; the second term does this for all bus stops, which could be equipped with charging devices. It is up to the planning process to decide which of the feasible power lines are necessary and how to partition the transformer power to different charging stations.

Since we work only with mean values and we have no respect of the power and energy flows over the time, we can calculate with two assumptions: The *best case* and the worst case. In the best case, there are no time overlapping charges, which get their power from the same transformer. Then the maximal available power at a charging station k is the minimum of the charger limit ps_{max_k} and the sum of power limits pt_{max_p} of all connected transformers.

In the *worst case*, we assume that for a given transformer all connected charging stations are occupied by buses, which must be charged simultaneously. Then the available power of this transformer station must be divided to all connected consumers. Optimizer models for both cases are available.

OPTIMIZATION MODELS

This section shows various development stages for the optimization model. Starting with the simplest model, which only estimates the bus stops further models are introduced systematically. Each new model is an extension of its predeceasing model and additional aspects are considered. This further confines the solution space but the number of constraints and the number of variables increase. All optimization models use the introduced system model and the general form is

$$\boldsymbol{A} \cdot \boldsymbol{x} \geq \boldsymbol{b} \quad \text{Constraints} \tag{9}$$
$$\boldsymbol{c}^T \cdot \boldsymbol{x} \rightarrow \min \quad \text{Objective Function} \tag{10}$$

The objective function reflects in general the economic costs. However, it is also possible to consider other aspects like urban planning factors. Model extensions are possible in two ways, adding *new constraints*:

$$\begin{pmatrix}\boldsymbol{A_{11}}\\ \boldsymbol{A_{21}}\end{pmatrix} \cdot \boldsymbol{x} \geq \begin{pmatrix}\boldsymbol{b_1}\\ \boldsymbol{b_2}\end{pmatrix} \tag{11}$$

or introducing *new variables*, which requires an extended objective function:

$$\begin{pmatrix}\boldsymbol{A_{11}} & \boldsymbol{0}\\ \boldsymbol{A_{21}} & \boldsymbol{A_{22}}\end{pmatrix} \cdot \begin{pmatrix}\boldsymbol{x_1}\\ \boldsymbol{x_2}\end{pmatrix} \geq \begin{pmatrix}\boldsymbol{b_1}\\ \boldsymbol{b_2}\end{pmatrix} \tag{12}$$

$$\begin{pmatrix}\boldsymbol{c_1}\\ \boldsymbol{c_2}\end{pmatrix}^T \cdot \begin{pmatrix}\boldsymbol{x_1}\\ \boldsymbol{x_2}\end{pmatrix} \rightarrow \min \tag{13}$$

We start with the simplest possible model, which fulfills the energy balance between the buses and the charging stations only. There are no transformers so far and the charging power is the minimum of the power limit $\boldsymbol{pb}_{max}$ given by the buses and the power limit of the charging devices $\boldsymbol{ps}_{max}$.

The second model additionally introduces transformer stations and the connections to the chargers. The amount of power flow between the transformers and the charging stations are out of the scope of the second model.

The third model allows power limits given by the transformers. If some buses charge simultaneously from charges, which get their energy from the same transformer, then the maximal allowed power could be overdrawn temporarily.

If this is not a sufficient solution, the model must be extended to the fourth model. In this case the power flow is modeled by introduction of additional variables and additional constraints.

All these models build on each other in ascending order. The run time and/or memory requirements increase with the model number. A good strategy is the appliance of the different models in ascending order to get first results in a short time. If the constraints in a certain model do not holt, the procedure can be stopped and the cause can be found easily.

Table 2 gives an overview about the input and output data of the different models.

Table 2: Optimizer Models with Input Data and Output Data

Model	Input	Output
1	Traction energy Base power Bus schedule Max. charging power Max. power offer Costs for charging devices	List of charging points $\mathbb{C}$ Invest for charging devices Power flow from charging devices Power flow to buses
2	***Additional***: Costs for power lines	***Additional***: List of used transformers Invest for power lines
3	***Additional***: Max. transformer power	***Additional***: Best case power flow from transformers
4	***same as model 3***	***Additional***: Worst case power flow from transformers

Buses that stop at regular bus stops can be charged there if the stop is equipped with a charging device. There are three restrictions, which limit the energy transfer from the charger into the batteries. Table 3 shows them by name, graphical symbol and mathematical representation.

Table 3: Power Limits for Buses, Charging Stations and Transformers

Bus	Charging Station	Transfomer
		
$\boldsymbol{pb_{max}}$	$\boldsymbol{ps_{max}}$	$\boldsymbol{pt_{max}}$

Batteries have limited charging power. This limit can be individual for each bus ($\boldsymbol{pb_{max}}$). The charging devices have a maximal output power. This limit can be individual for each charging station ($\boldsymbol{ps_{max}}$). Each transformer has a power limit ($\boldsymbol{pt_{max}}$).
All batteries must be operated within specified limits of $[e_{min}, e_{max}]$ (see Figure 4). Therefore, the maximal charge $\Delta e = e_{max} - e_{min}$ is only possible at a specific bus stop if a bus arrives there with a charge of e_{min} .

Model 1: Selecting bus stops for charging stations

For this very first model, we only consider power limits for buses and for chargers as well as energy boundaries for buses. Within these limits, we chose the smallest amount of energy $e_{i,k}$ that is transferred at stop k to bus:

$$e_{i,k} = min\left(\Delta e_i\,, pb_i \cdot ts_{i,k}\,, ts_{i,k} \cdot ps_i\right) \tag{14}$$

For all buses and for all stops given by the stop schedule matrix $\boldsymbol{SS}$ the energy transfer is equal to the matrix $\boldsymbol{A}$ for the optimizer.

$$\boldsymbol{A} = min\left(\boldsymbol{\Delta e}, diag(\boldsymbol{pb_{max}}) \cdot \boldsymbol{TS}, \boldsymbol{TS} \cdot diag(\boldsymbol{ps_{max}})\right) \circ \boldsymbol{SS} \tag{15}$$

The solution vector $\boldsymbol{x_{ds}} \in \{0,1\}^{|\mathbb{S}|}$ selects those bus stops, which become charging ability. The goal is that no bus has to quit its service due to lack of energy. $\boldsymbol{A} \cdot \boldsymbol{x_{ds}}$ denotes the maximal possible energy transfer to each bus during a whole cycle. Finally the energy balance is:

$$\underbrace{\boldsymbol{A}}_{\boldsymbol{A_1}} \cdot \underbrace{\boldsymbol{x_{ds}}}_{\boldsymbol{x_1}} \geq \underbrace{\boldsymbol{e^-}}_{\boldsymbol{b_1}} \tag{16}$$

The indices of the subscripts mark this model as model one. The objective function assigns costs to each bus stop, which will be equipped with a charging device:

$$\underbrace{\boldsymbol{c}}_{\boldsymbol{c_1}} \cdot \underbrace{\boldsymbol{x_{ds}}}_{\boldsymbol{x_1}} \rightarrow \min \tag{17}$$

Model 2: Selecting transformer stations

The next step is to select transformer stations for supplying power to the charging devices. Model one delivers the charging stations, which need power from transformer stations. The assumption for the model two is that each transformer station can supply at least as much power as all connected chargers require. This model does not consider the power limits of the transformers what Model three does.
We introduce additional constraints for the connections and a new set of decision variables $\boldsymbol{x_{dts}} \in \{0,1\}^{|\mathbb{S}|\cdot|\mathbb{T}|}$, which indicate if a connection is chosen or not chosen. Model 1 becomes extended as follows:

$$\underbrace{\begin{pmatrix} \boldsymbol{A_1} & \boldsymbol{0} \\ -\boldsymbol{I}_{|\mathbb{S}|} & \boldsymbol{A_{22}} \end{pmatrix}}_{\boldsymbol{A_2}} \cdot \underbrace{\begin{pmatrix} \boldsymbol{x_{ds}} \\ \boldsymbol{x_{dts}} \end{pmatrix}}_{\boldsymbol{x_2}} \geq \underbrace{\begin{pmatrix} \boldsymbol{b_1} \\ \boldsymbol{0} \end{pmatrix}}_{\boldsymbol{b_2}} \tag{18}$$

$$\boldsymbol{A_{22}} \in \{0,1\}^{|\mathbb{S}|\times|\mathbb{S}|\cdot|\mathbb{T}|} \tag{19}$$

$$\boldsymbol{A_{22}}_{k,p} = \begin{cases} 1 \text{ if a connection between} \\ \quad \text{a charging station } k \text{ and} \\ \quad \text{a transformer station } p \text{ is possible} \\ 0 \text{ otherwise} \end{cases}$$

The objective function takes account of the additional costs for installing the power lines. The cost matrix $\boldsymbol{CTS} \in \mathbb{R}_0^{+|\mathbb{S}|\times|\mathbb{T}|}$ contains these values, where cts_{kp} represents the cost of a line from transformer station p to bus stop k. For the optimization model, $\boldsymbol{CTS}$ is transformed into a vector by line wise read:

$$\boldsymbol{cts}_i = \boldsymbol{CTS}_{(\mathrm{p}-1)\,|\mathbb{T}|+k} \tag{20}$$

The cost function of Model 1is extended to this vector:

$$\boldsymbol{c_2} = \begin{pmatrix} \boldsymbol{c_1} \\ \boldsymbol{cts} \end{pmatrix} \qquad \underbrace{\begin{pmatrix} \boldsymbol{c_1} \\ \boldsymbol{cts} \end{pmatrix}^T}_{\boldsymbol{c_2}} \cdot \underbrace{\begin{pmatrix} \boldsymbol{x_{ds}} \\ \boldsymbol{x_{dts}} \end{pmatrix}}_{\boldsymbol{x_2}} \rightarrow \min \tag{21}$$

Each selected connection contributes to objective value.

Model 3: Considering transformer limits

Model 2 makes it possible to determine the necessary connections between the charging stations and the transformer stations. However, only the installation costs but not possible power restrictions were taken into account so far. Model 3 ensures that even in the case of a maximal transformer power is less than the maximum charging power of the buses; these can always be supplied with sufficient energy. It is assumed that the buses, which simultaneously receive their energy from the same transformer, do not overload these, even if they are charged with maximum charging power. This condition is only taken into account in Model 4.

$$\underbrace{\begin{pmatrix} \boldsymbol{A_1} & \boldsymbol{0} \\ -\boldsymbol{I}_{|\mathbb{S}|} & \boldsymbol{A_{22}} \\ \boldsymbol{0} & \boldsymbol{A_{32}} \end{pmatrix}}_{\boldsymbol{A_3}} \cdot \underbrace{\begin{pmatrix} \boldsymbol{x_{ds}} \\ \boldsymbol{x_{dts}} \end{pmatrix}}_{\boldsymbol{x_3}} \geq \underbrace{\begin{pmatrix} \boldsymbol{b_1} \\ \boldsymbol{0} \\ \boldsymbol{b_1} \end{pmatrix}}_{\boldsymbol{b_3}} \tag{22}$$

$$\boldsymbol{A_{32}} \in \mathbb{R}_0^{+|\mathbb{S}|\times|\mathbb{S}|\cdot|\mathbb{T}|} \tag{23}$$

$$\boldsymbol{A_{32}} = \boldsymbol{SS} \circ \boldsymbol{TS} \cdot \begin{pmatrix} \boldsymbol{pt_{max}}^{\mathrm{T}} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{pt_{max}}^{\mathrm{T}} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{pt_{max}}^{\mathrm{T}} \end{pmatrix} \tag{24}$$

The model contains no additional variables and the objective vector remains unchanged:

$$\boldsymbol{c_3} = \boldsymbol{c_2} \tag{25}$$

The additional constraints reduce the solution space. In case of under dimensioned transformer power it is possible, that Model 2 is solvable but Model 3 is unsolvable. For this reason, a successive application of the different models makes sense.

Model 4: Supporting simultaneous charging

The previous model only made the decision, between which transformer stations and which bus stops a powerline has to be installed. Model 4 calculates the amount of power flow over these connection lines. In this model, the sum of the charging power delivered from a transformer must not exceed its limit even if all connected charger operate simultaneously. This pessimistic assumption generally leads to higher costs, but also to a more robust charging infrastructure.
This model covers up the last constrain block from Model 3. The power flow between a transformer and an charging station is expressed by $\boldsymbol{x_p} \in \mathbb{R}_0^{+|\mathbb{S}|\cdot|\mathbb{T}|}$. The model becomes:

$$\underbrace{\begin{pmatrix} \boldsymbol{A_1} & \boldsymbol{0} & \boldsymbol{0} \\ -\boldsymbol{I}_{|\mathbb{S}|} & \boldsymbol{A_{22}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{A_{42}} & \boldsymbol{A_{4a,3}} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{A_{4b,3}} \end{pmatrix}}_{\boldsymbol{A_4}} \cdot \underbrace{\begin{pmatrix} \boldsymbol{x_{ds}} \\ \boldsymbol{x_{dts}} \\ \boldsymbol{x_p} \end{pmatrix}}_{\boldsymbol{x_4}} \geq \underbrace{\begin{pmatrix} \boldsymbol{b_1} \\ \boldsymbol{0} \\ \boldsymbol{0} \\ \boldsymbol{b_1} \end{pmatrix}}_{\boldsymbol{b_4}} \tag{26}$$

The first block of the introduced additional constraints ensure that sum of the outgoing power flow from the transformers do not exceed their maximum.

$$\boldsymbol{A_{42}} = \begin{pmatrix} \boldsymbol{pt_{max}}^{\boldsymbol{T}} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{pt_{max}}^{\boldsymbol{T}} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{pt_{max}}^{\boldsymbol{T}} \end{pmatrix} \tag{27}$$

$$\boldsymbol{A_{4a,3}} = -\begin{pmatrix} 11\ldots1 & 00\ldots0 & \cdots & 00\ldots0 \\ 00\ldots0 & 11\ldots1 & \cdots & 00\ldots0 \\ \vdots & \vdots & \ddots & \vdots \\ 00\ldots0 & 00\ldots0 & \cdots & 11\ldots1 \end{pmatrix} \tag{28}$$

The following block of constraints ensures that for all buses their energy demand is provided at least.

$$\boldsymbol{A_{4b,3}} = \boldsymbol{SS} \circ \boldsymbol{TS} \cdot \begin{pmatrix} 10\ldots0 & 10\ldots0 & \cdots & 10\ldots0 \\ 01\ldots0 & 01\ldots0 & \cdots & 01\ldots0 \\ \vdots & \vdots & \ddots & \vdots \\ 00\ldots1 & 00\ldots1 & \cdots & 00\ldots1 \end{pmatrix} \tag{29}$$

The model contains additional variables, which do not influence the objective values. Finally, the objective vector becomes:

$$\boldsymbol{c_4} = \begin{pmatrix} \boldsymbol{c_3} \\ \boldsymbol{0} \end{pmatrix} \tag{30}$$

The additional variables x_p do not contribute to the objective value.

SIMULATION MODEL

The hybrid planning process consists of two steps. The first one calculates all necessary average values from the input data. The optimizer works with these values and produces a plan, which bus stops must be equipped with charging devices and which transformer station supplies the needed power. These results are the input for the simulation process.
The simulation checks the feasibility of the solution. If at least one bus runs out of energy, the optimization is executed again but with increased energy demands. If charging times at different charging stations overlap more than a given acceptable limit, the objective function gives penalty to the effected charging stations. Figure 5 shows this in a flow diagram.

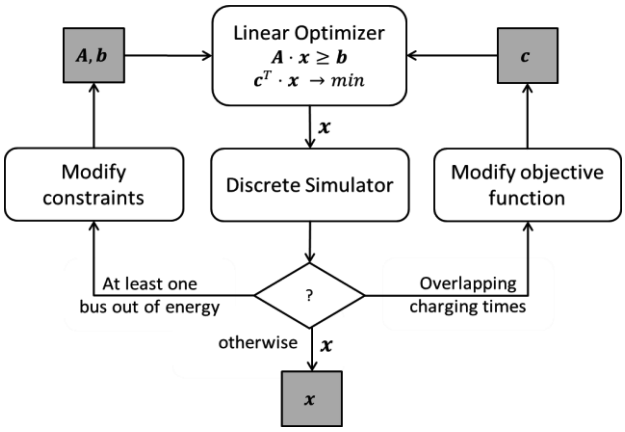


Figure 5: Control flow of the optimization process

The simulation runs on the linear model in order to get short simulation times. For the final run of the simulator, a more sophisticated model can be used to reflect the nonlinear behavior of the battery charge. This needs more time for the simulation process and it is out of the scope of this paper.
The simulator works event driven and processes the data from complete bus schedules. The events are departures from bus stops and arrivals at bus stops. The event processor does not generate events and there is no inner loop back into the event queue. Since any synchronization between buses for passenger transfers or other reasons are considered in the timetable and bus schedule design. For this reason, an independent processing of all buses becomes possible. This is the key for simulating all buses concurrently.
The output matrices $\boldsymbol{SS'}$ and $\boldsymbol{SW'}$ contain the number of visits of all stops and of all ways by the different buses. If all buses complete their runs successfully the output matrices $\boldsymbol{SS'}$ and $\boldsymbol{SW'}$ are equal to the given schedule matrices $\boldsymbol{SS}$ and $\boldsymbol{SW}$**.** Otherwise, hints for modifying the infrastructure can be derived from the output matrices.

RESULTS AND CONCLUSIONS

The introduced procedure was implemented in Octave and the basic algorithms for optimizing and for simulation are external functions. These are the most time consuming components and an interpreter like Octave runs inherently rather slow. The code for the optimizer is the GNU Linear Programming Kit (GLPK). The simulator was especially coded in C++ for this application. The memory complexity is shown in table 5 and the number auf constraints and variables are shown in table 4.

Table 4: Memory Requirements of the Optimization Models

Model	Number of Constraints	Number of Variables
1	$\lvert\mathbb{B}\rvert$	$\lvert\mathbb{S}\rvert$
2	$\lvert\mathbb{B}\rvert + \lvert\mathbb{S}\rvert$	$\lvert\mathbb{T}\rvert \cdot (\lvert\mathbb{S}\rvert + 1)$
3	$\lvert\mathbb{B}\rvert + 2\lvert\mathbb{S}\rvert$	$\lvert\mathbb{T}\rvert \cdot (\lvert\mathbb{S}\rvert + 1)$
4	$2\lvert\mathbb{B}\rvert + 3\lvert\mathbb{S}\rvert$	$\lvert\mathbb{T}\rvert \cdot (2\lvert\mathbb{S}\rvert + 1)$

Table 5: Memory Complexity of the Optimization Models

Model	Memory Complexity
1	$\mathcal{O}(\lvert\mathbb{B}\rvert \cdot \lvert\mathbb{S}\rvert)$
2	$\mathcal{O}(\lvert\mathbb{S}\rvert \cdot \lvert\mathbb{T}\rvert \cdot (\lvert\mathbb{B}\rvert + \lvert\mathbb{S}\rvert))$
3	$\mathcal{O}(\lvert\mathbb{S}\rvert \cdot \lvert\mathbb{T}\rvert \cdot (\lvert\mathbb{B}\rvert + 2\lvert\mathbb{S}\rvert))$
4	$\mathcal{O}(\lvert\mathbb{S}\rvert \cdot \lvert\mathbb{T}\rvert \cdot (2\lvert\mathbb{B}\rvert + 3\lvert\mathbb{S}\rvert))$

In order to reduce the model size, we eliminate all variables that do not contribute to the objective value. After this first step of *model reducing*, some constraints could become obsolete. In a second step, we eliminate them and loop back to the first step until no reduction is possible. The reduced models need significant less memory and the optimization runs faster without any influence to the results.

Quantitative results for the complexity improvements by model reduction are not available yet and are left to further work. For first tests, artificial data form a model generator were sufficient. Next results came from investigations with three bus operator companies in East Germany and one in Poland. However, in these first studies the power grid was not integrated in the corresponding models. Real life bus systems for two German cities (Magdeburg and Braunschweig) are being modeled currently. The energy suppliers provide the data for the position and power capability of their transformer stations, but it is not easy to determine the costs of connecting the potential charging stations to the transformers. Finally, some costs became available and some other were estimated so far. First results are available and are to be verified. Currently we work on the medium sized model for the city of Braunschweig, Germany. The model consists of 1153 bus stops, 1294 ways and 65 buses.

REFERENCES

Buechter, H. and Naumann, S.. 2016a. “Optimizing Charging Infrastructure for Electrically Driven Buses.” In: European Simulation and Modelling Conference ESM 2016, pages 384–386.

Buechter, H. and Naumann, S. 2016b. “A Hybrid Planning Method for Charging Infrastructure for Electrically Driven Buses in Public Transportation.” In: Grzegorz Sierpiński (Editor) Intelligent transport systems and travel behavior. 13th Scientific and Technical Conference "Transport Systems. Theory and Practice 2016" Katowice, Poland, September 19-21, 2016 : selected papers., Cham: Springer (Advances in intelligent systems and computing, Volume 505), pages 175-185.

Chen, T. D., Kockelman, K. M., Moby Khan, et al. 2013. “The electric vehicle charging station location problem: A parking-based assignment method for Seattle.” In: Transportation Research Board 92nd Annual Meeting, volume 340, 2013, pages 13-1254

Frade, I.; Ribeiro, A.; Goncalves, G.; Antunes, A.P. 2011. “Optimal location of charging stations for electric vehicles in a neighborhood in Lisbon, Portugal.” In: Transportation Research Record, Journal of the Transportation Research Board, 2011, pages 91-98.

Klemmt, A.; S. Horn; G. Weigert; K.J. Wolter. 2009. "Simulation-based optimization vs. mathematical programming: A hybrid approach for optimizing scheduling problems." In: Robotics and Computer-Integrated Manufacturing (25), pages 917–925.

Kuchshaus, V.; Bioly, S.; Klumpp, M. 2012. "Deducting E-Mobility Loading Station Locations from City Parking Data." In: 17th International Working Seminar on Production Economics, Conference Proceedings, Innsbruck 20.-24.02.2012, Bd. 2: Innsbruck, pages. 81–92.

Lam, A. Y. S.; Leung, Y.-W.; Chu, X. 2014. "Electric vehicle charging station placement: Formulation, complexity, and solutions." In IEEE Transactions on Smart Grid (5), pages 2846-2856.

Olsen, N., Kliewer, N. 2016, "Electric Vehicle Scheduling - A study on charging modeling for electric vehicles.", In: Operations Research Proceedings, 2016, pages 82–102.

Zheng, Y.; Dong, Z. Y.; Meng, K.; Zhao J.H.; Qiu, J., 2014. “Electric Vehicle Battery Charging/Swap Stations in Distribution Systems: Comparison Study and Optimal Planning” In: IEEE Transactions on Power Systems, Vol. 29, No. 1, January 2014, pages 221-229.

SENSOR NETWORK SIMULATION

SIMULATION OF OPTIMIZED NONLINEAR FREQUENCY MODULATION IN PULSE COMPRESSION RADAR

Jiří Roleček
Pavel Bezoušek
Karel Juryca
Department of Electrical Engineering
University of Pardubice
Studentská 95, Pardubice, 532 10
Czech Republic
E-mail: Jiri.Rolecek@upce.cz

KEYWORDS
Radar, Pulse Compression, LFM, NLFM, Doppler effect, signal distortion, Optimization.

ABSTRACT

In this paper, a radar pulse, which is a nonlinear frequency modulated signal, is mathematically modeled and analyzed. The impact of the Doppler effect (due to the target motion), and impact of signal distortion in radar components on pulse compression efficiency, is studied. The aim is to optimize both the signal waveform and the compression algorithm. Two mitigation methods of parameter deterioration are considered, namely: a pre-distortion and a pulse suppression filter response adaptation. In the case when only a linear and time invariant distortion is considered, both equalization methods lead to the same results. Unfortunately, the Doppler effect is not a time invariant procedure, and the high power amplification in a radar transmitter is basically a non-linear operation. That is why both methods should be used simultaneously in balance.

INTRODUCTION

Essential radar characteristics are the maximum detection range of small targets and the range resolution of two closely spaced targets. While the maximum range is related to the energy of the transmitted pulse, the range resolution depends on the signal bandwidth B (Richards at al.). Thus, surveillance radars with high maximum ranges and moderate range resolutions should transmit pulses of a high time-bandwidth product $B\tau_p$. The necessary signal bandwidth is achieved by an appropriate intra-pulse modulation.
The received signal, reflected off some object is amplified and filtered in a standard heterodyne receiver, and then it is correlated with the transmitted signal in the matched filter, where a time shifted replica of the transmitted signal is generated. This procedure is called the pulse compression. The main parameters of the compressed signal are the main-lobe width and the side-lobe to the main-lobe ratio (side-lobe suppression). The first parameter is closely related to the range (time) resolution of two targets. If these targets are not sufficiently separated, then a main-lobe width cannot be resolved. The side-lobes cover the area in the extent of the transmitted pulse length to both sides around the main lobe. A good side-lobe suppression factor enables the radar to also detect weak reflections close to the strong ones.
The Doppler effect deteriorates the signal compression parameters, broadening the main lobe and lowering the side-lobe suppression. The Doppler effect impact depends highly on the type of the intra-pulse modulation used. It is well known, that selection of the linear frequency modulation (LFM) conserves the main lobe width, even at long pulses with high and moderate bandwidths. Unfortunately, it also shows high side-lobes due to its rectangular-shaped modulation spectrum. The nonlinear frequency modulation (NLFM) could effectively suppress the side-lobes, with only a minor widening of the main lobe. But the NLFM reaction to the Doppler effect critically depends on the actual FM response. That is why the modulation optimization, with respect to the Doppler effect, is a part of this research. We are also interested in the effect of a signal distortion in radar components on the mentioned pulse compression characteristics. The possibility of minimization of this effect by pre-distortion and by adaptation of the matched filter response were analyzed using signal and processing algorithm models.

SIGNAL MODEL DESCRIPTION

The radar phase or frequency modulated pulsed signal can be described as follows:

$$s(t) = A(t).e^{j2\pi f_0 t} e^{j\varphi(t)}$$

$$A(t) = A_0 \quad \text{for} \left| t - kT_{\text{op}} \right| \le \frac{\tau}{2}$$

and

$$A(t) = 0 \quad \text{for} \left| t - kT_{\text{op}} \right| \rangle \frac{\tau}{2}$$

Where f_0 is the carrier frequency
A_0 is the pulse amplitude
τ is the pulse length
T_{op} is the pulse repetition period
k is an integer
$\varphi(t)$ is the modulated phase.

In practice the signal is generated and processed in the form of its complex envelope. It means without the carrier

term $e^{j2\pi f_0 t}$. The signal path in a radar starts in a Direct digital synthesizer (DDS), which generates the signal complex envelope according to the prescribed $\varphi(t)$ response. The instantaneous phase $\varphi(t)$ is related to the modulated frequency $f(t)$ by:

$$\varphi(t) = 2\pi \int_{-\tau/2}^{t} f(\gamma) d\gamma$$

Particularly, we analyzed the signal for an S-band ground-based surveillance radar with f_0 = 2.8 GHz, τ = 80 μs and T_{op} = 1.05 ms. The radar signal bandwidth is 1.25 MHz.

The signal power spectrum density PSD $S(f)$ is closely related to the $f(t)$ response. Using the stationary phase assumption (Vizitiu et al. 2014) the following relation between $f(t)$ and $S(f)$ could be found:

$$\int_{-\frac{T_{op}}{2}}^{t} |s(\theta)|^2 d\theta = \int_{-\infty}^{f(t)} S(\gamma) d\gamma$$

A typical $f(t)$ response and the corresponding phase $\varphi(t)$ are shown in figures 1 and 2 (the blue curve). In figure 1 we can see two precipitous regions of $f(t)$ at the pulse edges, responsible for the compressed pulse side-lobe suppression. In this particular case, the signal PSD has the form of the Taylor window (figure 3).

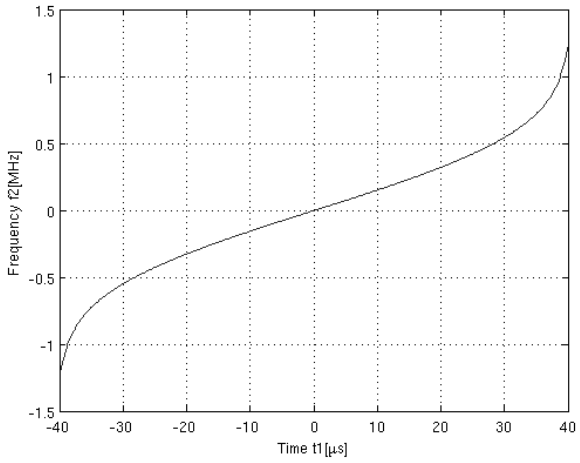


Figure 1: Frequency Modulation

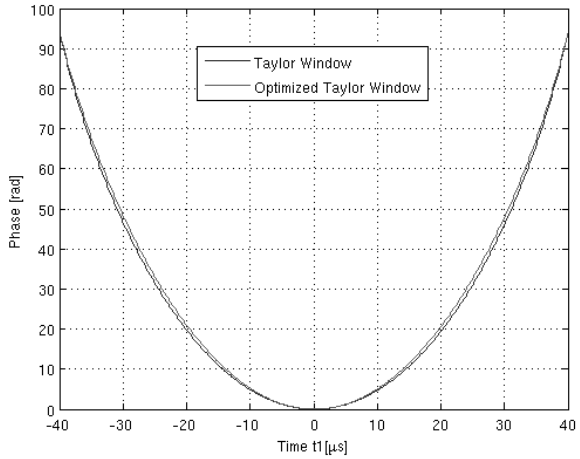


Figure 2: Phase Shift

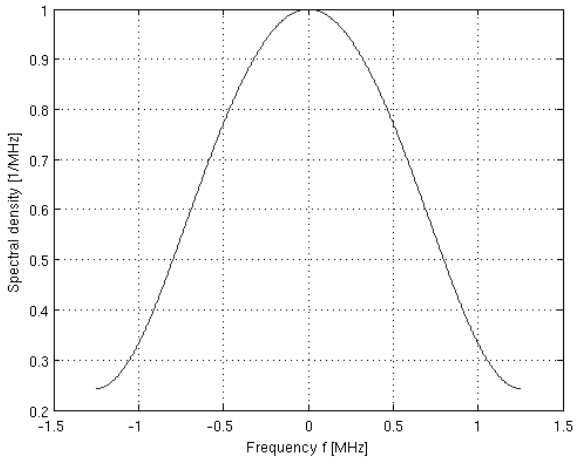


Figure 3: The Taylor window used as a PSD

The signal bandwidth B is an equivalent width of the PSD curve i.e. a rectangle of the same area width.

After reception and pre-processing of the reflected signal $s_r(t)$ and its transformation to complex envelope, a matched filter is applied, where the received signal complex envelope is correlated with a replica of the transmitted signal complex envelope. If the signal distortion is negligible, the filter output has the form of the delayed autocorrelation function (ACF) of the transmitted signal complex envelope. Below the "complex envelope" specification is dropped using the term signal only.

An autocorrelation function of the signal with the frequency modulation shown in figure 1, is demonstrated in figure 4. Since the signal ACF is generally an Inverse Fourier Transform of the signal PSD, the ACF main-lobe width is a reciprocal to the signal bandwidth B. Its side-lobes levels depend on the form of the signal PSD, particularly on its edge "roll-off" quality. Thus, standard window functions may be used in the role of a signal PSD to achieve the desired ACF. Unfortunately, if the Doppler effect and a signal distortion are taken into account, the situation becomes more complicated.

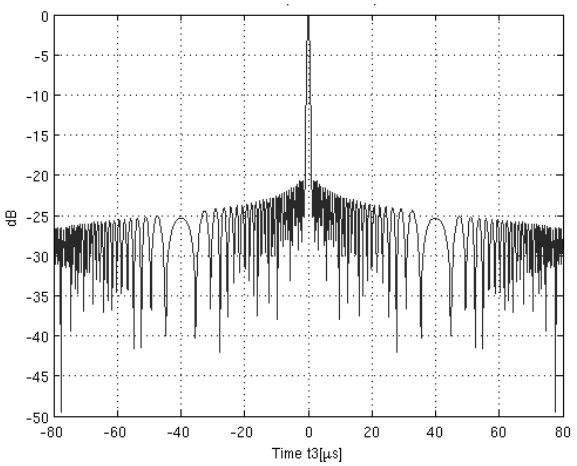


Figure 4: Autocorrelation function of a signal with a Taylor window PSD.

DOPLLER EFFECT SIMULATION

Surveillance radars are used to detect and track primarily moving targets. Then the reflected signal frequency is shifted by the Doppler shift f_d (Richards et al. 2014):

$$f_d = -\frac{2v_r}{c} f_0$$

where v_r is the radial component of the target to radar velocity, and c is the speed of light. If the target approaches the radar, the relative velocity v_r is negative. It is worth noting, that in a primary radar application the Doppler shift is twice greater than in communication, because the signal is affected twice: at its impact and at the reflection of the target.

In our complex signal model it could be simply assumed by a multiplication by a factor of $\exp(j2\pi f_d t)$:

$$s_r(t) = s(t) \cdot e^{j2\pi f_d t}$$

Due to a large variety of possible target movements, the Doppler shifts of the both polarities with the same probability should be assumed. From this consideration we conclude, that the signal frequency modulation $f(t)$ should be either symmetrical or anti-symmetrical around the pulse center. But the symmetrical alternative is quite impractical, since it needs four regions with a steep $f(t)$ response instead of two regions in the case of an anti-symmetrical response (figure 1).

In figure 5, the result of pulse compression simulation is shown with a received signal, reflected off the target with a relative radial velocity of 600 m/s. The magnified mainlobe shows, that in our setup, the result is Doppler´s effect tolerant, i.e. sidelobe level and mainlobe width aren´t substantially influenced.

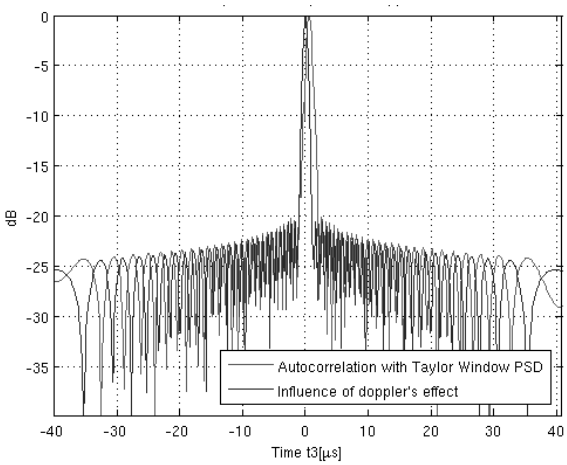


Figure 5: Influence of Doppler´s effect

LINEAR DISTORTION SIMULATION

Similarly important, is to integrate a signal distortion due to the radar component imperfections into the simulation process. The distortion could be either linear or non-linear. The non-linear one, mainly the AM-PM conversion, is highly specific for a particular device and it will be evaluated experimentally. In this research phase, only the linear distortion was incorporated in the form of a FIR distortion filter with the following transfer function:

$$H(f) = \frac{1}{1 + j2\pi \dfrac{f - f_0}{B_d}}$$

where B_d is the distortion filter bandwidth and f is the signal frequency. In figure 6, there is shown the transfer function dependence on the relative frequency $f_1 = f - f_0$ in the range of $\pm B$ (B is a radar signal bandwidth). In this case, the distortion filter bandwidth was chosen to be B_d = 10 MHz.

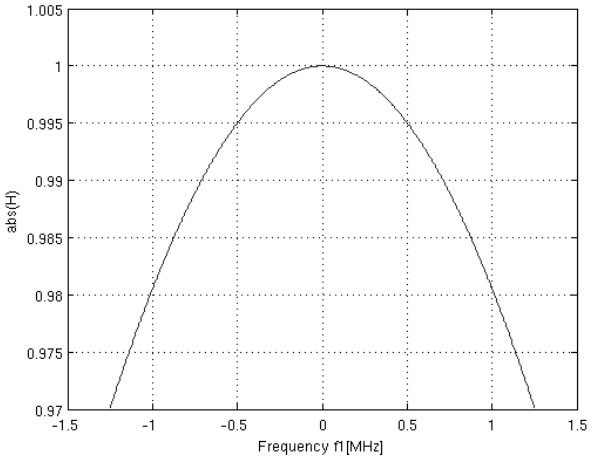


Figure 6: Distortion filter transfer function

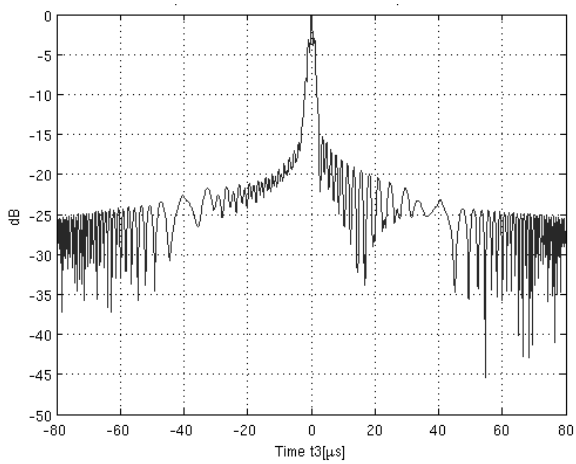


Figure 7: Pulse compression results on signal distorted by the radar system. Signal modulation using the Taylor window.

REDUCTION OF THE UNDESIRABLE IMPACTS

There are basically two ways, how to reduce the impacts of the Doppler effect and signal distortion on the pulse compression quality. They are: the signal pre-distortion in front of a high power amplifier and the signal equalization at reception. It is obvious, that the pre-distortion could be performed in the signal generator, simultaneously with its modulation. On the other hand, the signal processing at receiver, from its input to the pulse compression filter, is assumed to be linear and time invariant. As a result, the first

reduction method consists of a modification of the signal phase modulation in the signal generator. The second method comprises of an adaptation of the pulse compression filter response.

It was shown, that the ACF side-lobe reduction of the signal, without the Doppler shift, can be effectively performed using a window function with a corresponding phase modulation. For instance, we found the Taylor window to be adequate in respect to the side-lobe suppression and the main-lobe width widening. But, if the Doppler effect is considered, slight modification should be applied. We optimized the phase modulation $\varphi(t)$ by adding a correction term $\varphi_C(t)$ to the Taylor window modulation response $\varphi_T(t)$:

$$\varphi(t) = \varphi_T(t) + \varphi_C(t)$$

where the correction term $\varphi_C(t)$ is an even polynomial of the time t:

$$\varphi_C(t) = \sum_{n=1}^{N} a_n \left(\frac{4t^2}{\tau^2} \right)^n ,$$

where the coefficients a_n are subjects of the optimization, focused on side-lobe minimization simultaneously with the main-lobe width conservation. The same procedure could be applied to optimize signal pre-distortion in order to suppress a signal non-linear distortion in the high power amplifier of the radar transmitter.
The modification methods of the compression filter response could be divided into two groups: the lossless method and the lossy one. There is a simple, quite universal method, leading to the side-lobe level reduction. This method consists of attenuation of the filter edge coefficients. The coefficients of the matched filter are equal to the complex conjugates of the transmitted signal samples, which, in our case, have all the same amplitudes. Then this method is very efficient in side-lobe reduction, but it substantially widens the main-lobe and lowers its amplitude. So this method should only be used with much caution. The true lossless methods are possible only in the case of a time invariant linear distortion of the signal. Then the best solution is to change the filter response to be matched directly to the distorted signal.

RESULTS

To show the impact of the signal distortion on the pulse compression, and to demonstrate the proposed pre-distortion procedure, we used the described linear signal distortion model. The effect of the linear distortion on the pulse compression is shown in figure 7. Only the Taylor window phase response $\varphi_T(t)$ is used in the modulation here. When we use the optimized correction term in the modulating phase response $\varphi_C(t)$, the results are much better, as could be seen in figure 8. This is true even if the difference between the Taylor window $\varphi_T(t)$ and the optimized term is relatively small (fig. 3). During optimization, the N^{th} order correction polynomial was set in a broad range. Finally, the order as small as $N = 5$ was found to be sufficient.

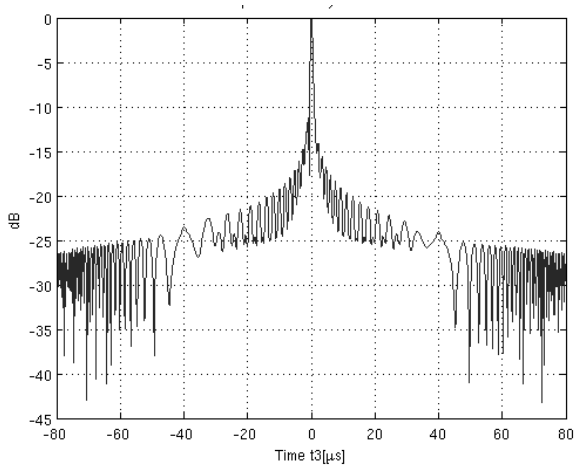


Figure 8: Pulse compression result on signal distorted by the radar system. Modulation using the Taylor window and optimized pre-distortion.

We can see, that using pre-distortion only, a partial improvement in the pulse compression is achieved. In the case of linear distortion, much better results could be reached by adapting the compression filter coefficients (see figure 9). However, in combination with the Doppler effect, the advantage of this method significantly falls.

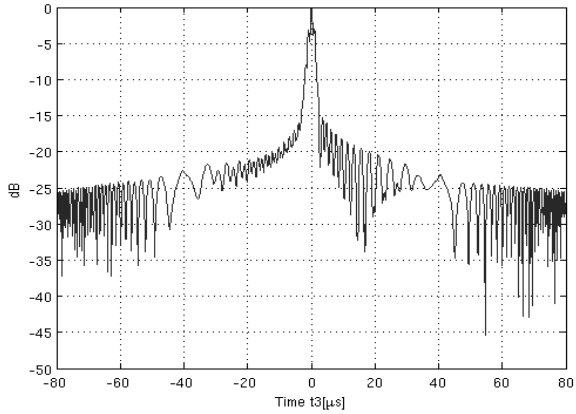


Figure 9. Pulse compression result of the distorted signal using the method of adapted compression filter coefficients

CONCLUSIONS

This part of the research shows, that there are more approaches to reduce effects of signal distortion and the Doppler effect on the pulse compression in a pulse radar system with nonlinear intra-pulse frequency modulation. Measurements on a real radar are prepared, in order to make our model more robust and adapted to real situations. This work will lead to improvements, that will be integrated in new radar systems.

REFERENCES

Boukeffa S., Jiang Y., Jiang T. 2011. "Sidelobe reduction with Nonlinear Frequency Modulated Waveforms". *IEEE 7th International Colloquium on Signal Processing and its Applications*. Penang Malaysia. ISBN 978-1-61284-413-8

Kurdzo J. M., Cheong B. L., Palmer R. D., Zhang G. 2014. "Optimized NLFM Pulse Compression Waveforms for High-Sensitivity Radar Observations". *International Radar Conference*. Lille France. ISBN 978-1-4799-4195-7

Richards M. A., Scheer J., Holm W. A., Melvin W. L. 2014. "*Principles of modern radar"*. Raleigh, NC: SciTech Pub., 2014. ISBN 978-1-891121-52-4.

Vizitiu I. C., Enache F., Popescu F. 2014. "Sidelobe Reduction in Pulse-Compression Radar Using the Stationary Phase Technique: an Extended Comparative Study." *International Conference on Optimization of Electrical and Electronic Equipment*. ISBN 978-1-4799-5183-3

BIOGRAPHIES

JIŘÍ ROLEČEK was born in Pardubice, Czech Republic in 1982 and went to the Brno University of Technology, where he studied biomedical engineering and cybernetics and obtained his degree in 2006. Since 2011 he is with the University of Pardubice at Faculty of Electrical Engineering as a PhD student.

PAVEL BEZOUŠEK was born in Ostrava, Czechoslovakia in 1943. He received his MS degree from the Czech Technical University in Prague in 1966 and his PhD degree from the same university in 1980. He was with the Radio Research Institute of the Tesla Pardubice from 1966 till 1994, where he was engaged in microwave circuits and systems design. Since then he is with the University of Pardubice, now at the Faculty of Electrical Engineering and Informatics. Presently he is engaged in radar systems design.

An Open Architecture Framework for the Electronic Warfare Modeling & Simulation

Sang Yeong Choi*
Hyun Seo Kang, Hyoung Jun Kwon, Sug Joon Yoon**

* School of Defense Science, Hansung University, Myongji University,
South Korea, E-mail: metayoung@gmail.com
** Department of Mechanical and Aerospace Engineering, Sejong University

KEYWORDS
Open Architecture, Electronic Warfare, Modeling and Simulation, Reusability, Interoperability

ABSTRACT

This paper presents an open architecture framework for the electronic warfare modeling & simulation (EW M&S), called "OAFEw". OAFEw is intended to be used for the development of an EW simulation model in a reusable and interoperable way. The basic idea of OAFEw is that EW M&S components can be easily re-assembled to suit the user's needs with its building blocks defined in the common conceptual reference model and rules governing all stages through a EW M&S life cycle. We showed its usefulness with an illustration of some OAFEw implementation for a simple scenario.

INTRODUCTION

EW is any action involving the use of the electromagnetic spectrum or directed energy to control the spectrum, or impede adversary assaults via the spectrum. EW includes three major subdivisions: electronic warfare support (ES), electronic protection (EP), and electronic attack (EA) (Schleher 1999). ES involves searching for threat signals for the purpose of immediate threat recognition, identifying and locating to help planning and conduct of counter operations. EA involves measures taken to defeat or neutralize threat electronic assets by means of jamming, chaff, flare, and so on. EP comprises countermeasures to enemy EA. In the EW, the threat is a kind of electro-signals emitted from a search or track radar. The target or the EW system may detect the threat signal and take a counter measure such as jamming and flare against the threat emitter. Thus the EW engagement happens between an array of emitting threats and receiving EW systems.

EW modeling and simulation (M&S) is to capture the progression of actions or processes that EW assets undergo during an EW engagement, and to reproduce the EW reality in the simulated computer environment. During the research and development of EW equipment, the EW M&S supports a variety of engineering activities such as concept verification, trade-off analysis, design synthesis, testing and evaluation. Particularly in the simulation based design enabling cost reduction, time reduction, and performance enhancement, the EW M&S becomes its essential tool. However, the lack of M&S reusability, scalability, and interoperability has been degrading its usefulness.

In this paper, we propose an open architecture framework for electronic warfare modeling & simulation, called "OAFEw". OAFEw is intended to foster reusability and interoperability of EW M&S in the simulation based design of EW equipment. The basic idea of OAFEw is that "EW M&S SW components are to be re-assembled to suit the user's needs with its building blocks defined in the common conceptual reference model and rules governing all stages through an EW M&S life cycle." This paper presents the OAFEw for the EW M&S.

In the next Section, we will review the relevant research works. Then Section 3 describes the OAFEw in details. In Section 4, we will show its usefulness by showing an illustrative example of OAFEw implementation with a simple EW scenario. Finally we will have conclusions in Section 5.

RELATED LITERATURE REVIEW

Defense Modeling and Simulation

Defense modeling and simulation (DM&S) has a wide area of its application such as training, analysis, and weapon system development. Thus DM&S was recognized as important defense assets. That led to the US DoD-wide objective of finding ways to support reuse and interoperability of defense simulations. The high level architecture (HLA) (IEEE 2000) is one of the products to establish common high level simulation architecture to facilitate the interoperability of all types of models and simulations among themselves and with C4I systems, as well as to facilitate the reuse of M&S components. The Simulation Interoperability Standards Organization (SISO) focuses on facilitating simulation interoperability across government and non-government applications worldwide, and published base object model (BOM) template specification as a SISO standard (IEEE 2005). BOM provides a component framework for facilitating reuse and composability. The BOM concept is based on the assumption that piece-parts of models, simulations, and federations can be extracted and reused as modeling building-blocks or components. The interplay within a simulation or federation can be captured and characterized in the form of reusable patterns. These patterns of interplay are sequences of events between simulation elements. The representation of the pattern of interplay is captured in the BOM.

For the integration of different simulations, the federation execution and development process (FEDEP) (IEEE 204) and distributed simulation engineering and execution process (DSEEP) was developed (IEEE 2011). The DSEEP is a more

generalized process. The DSEEP is intended as a high-level process framework into which the lower-level systems engineering practices native to any distributed simulation user can be easily integrated. Further, for the implementation independent formulation of the simulation battlespace. a conceptual modeling language(CML) was introduced (Clark). The CML is primarily graphical depiction of a process with links to the knowledge acquisition/knowledge engineering artifacts providing domain details. The OOS CML uses a color-coded abstract model based on Coad's graphical language (Coad 1999).

EW Simulation

Gallagher et al (Gallagher 1988) proposed a bi-path process for the development of the generic EW simulation. The first path was to create the engineering models, and the second path was to take the models under a formal software development process, prepare them for software implementation in a full scale simulation. This method is a classical multi-disciplined team approach requiring development activity from the outset to the end, with modular design. We call it "from-the-outset" approach. The from-the-outset approach definitely lack of reusability. Gupta et al. (Gupta 2012) studied about the electronic warfare simulation based on service oriented architecture. The electronic warfare simulation was structured to SOA and achieved the effect of dynamic sharing and reusability.

Discussions

As reviewed in the literature review, in the defense modeling and simulation communities, much effort has been made to acquire M&S with the reusability and interoperability. Defense modeling and simulation communities made a lot of contributions by establishment of standards such as high level architecture (HLA), base object model (BOM) conceptual modeling language (CML), federation development and execution process(FEDEP), defense simulation engineering and execution (DSEEP), etc. They help foster the reusability and interoperability of M&S, but cannot provide the exact total solution for a specific domain. The reason is because its ultimate achievement eventually depends on the specific domain. Thus there is still an ever present need to reach the reusability and interoperability goal, which has been left to M&S developers of a specific domain. In this regards, our contribution is one of the embodiment of such efforts to the EW modeling and simulation domain.

OPEN ARCHITECTURE FRAMEWORK FOR EW MODELING AND SIMULATION: OAFEw

Basic Concept

An open architecture framework of electronic warfare modeling & simulation (called "OAFEw") is intended to foster easy composition of the EW simulation model, and to be used for the life cycle management of EW simulation components on the basis of common conceptual reference model so that it can achieve the reuse and interoperability of EW M&S. The basic concept of OAFEw is: Firstly, we assemble M&S components according to a EW scenario on the basis of the EW conceptual reference model. Secondly, we use common templates for the component specifications that are embodiment of elements in the EW conceptual reference model. Thirdly, we keep the common rules for component assembling, its implementation, reuse, and for governing all stages through an EW M&S life cycle.
OAFEw comprises a conceptual reference model, an interplay use case model, an interplay component library, and a simulation model, which are shown in Figure 1.

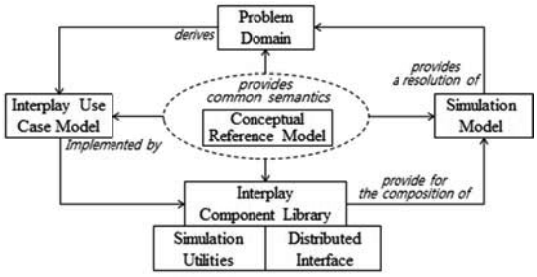


Figure 1: Constituents of OAFEw

The conceptual reference model is a conceptual model of an EW real world in the problem domain. It is noted that a conceptual model represents the real world to be implemented in the simulation model. The conceptual reference model provides common semantics to be referenced by other OAFEw constituents.
The interplay use case model is to model use cases within the context of EW simulation domain. The interplay use case model is associated with a EW scenario with which a certain problem should be resolved. The scenario thus can be formulated using each of interplay use cases.
The interplay component library is a collection of EW simulation components implemented according to specifications of the interplay use cases. The components in the interplay component library must have both internal interfaces to the simulation utilities such as a simulation engine, and external interfaces to the high level architecture runtime infrastructure (HLA/RTI) to be compliant with the HLA of distributed simulation. Thus there are two types of components in the interplay component library. These are simulation utility components and distributed interface components. Simulation utility components are for simulation execution and collection of simulation history data. Distributed interface components are for the interoperability of the components residing on the other federate on the EW federation.
The simulation model is to mimic an EW real world related with a problem domain. The simulation model is composed of components in the interplay component library. The components are reused for the simulation composition associated with an experiment scenario in the problem domain.

OAFEw Governance

OAFEw governance is to govern the life cycle of OAFEw constituents: the interplay use case model, the conceptual reference model, the interplay component library, the simulation model. The governing rules are specification rules, life cycle management rules, and verification rules.

Specification Rules

Specification rules constrain the specification method of the OAFEw constituents such as the conceptual reference model, the interplay use case model, and the distributed interface model. Figure 2 shows the graphical depiction of their relationships constraining each other.

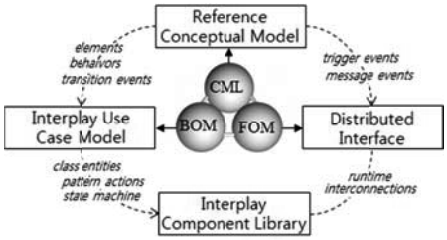


Figure 2: Specification methods

Firstly, the conceptual reference model provides OAFEw constituents with a common dictionary which explains what is going on in an EW real world. The conceptual reference model should be described using CML. The CML elements are shown schematically in Figure 3. The CML describe both static and dynamic properties of EW entities in a way that an Ew element generates an event, then the event again activates the element behavior and alters the state of the element. Every element belongs to a category and has characteristics. Piece and game space inherit from the element, and they belong to a battle space. For the sake of the fuller modeling power of the OAFEw constituents, we extend notion of the events to three types which are partly adopted from BOM: transition event, message event, trigger event. The transition event is an event that changes the inner state of an element. Either the message event or the trigger event is an event sending a external element. The message event designates the receiver, but the trigger event does not designate the receiver.

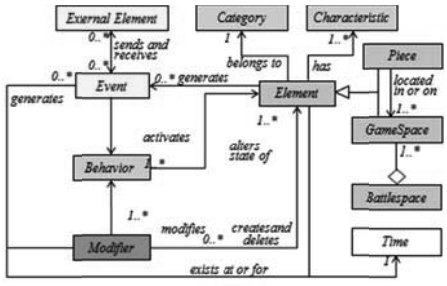


Figure 3: Conceptual modeling language

Secondly, the OAFEw use cases in the interplay use case model have to be captured so that the series of use cases may compose an experimental scenario for the resolution of problems. The interplay is the same notion as in the BOM standard, which is a sequence of pattern actions involving one or more conceptual entities appeared in the EW conceptual reference model. The pattern action is a single step in a pattern of interplay that may results in a state change of a conceptual entity. Thus, each interplay use cases should be specified using BOM standard. Further, the specification of interplay use cases have to use the same dictionary of conceptual entities, behaviors, events in the EW conceptual reference model, because they should share the common semantics with each other. The use case specifications are then to be implemented as reusable components in the interplay component library. The reusable components provide the simulation model with reusable building blocks.

Finally, the distributed interface should be specified with the HLA object model template (OMT) in order to be compliant with HLA. The object and interaction in the OMT have to be defined using either the message event or the trigger event in the conceptual model in order to share common semantics throughout all of the OAFEw constituents.

Life Cycle Management Rules

The OAFEw interplay components undergo two stages in the EW simulation life cycle: development stages, reuse stages. Figure 4 shows the two stages. The development stages are shown on the left hand side of Figure 4, and the reuse stages are on the right hand side. The development stages get started with the definition of EW operational scenario, and developing the interplay use case model which is a specification of interplay in the scenario. Then the interplay components are implemented. The way of implementation is not necessary to state in the life cycle management rules, because there may be many methods for their implementations. However, associated activities have to receive verification, validation and accreditation (VV&A). The VV&A will be addressed in the 3rd rule of Section 3.2.3 in more details. The verified components are then managed at the repository, waiting for their reuse. At the reuse stages on the right hand side of Figure 4, we firstly identify the problem to be resolved with the simulation, and define the problem domain. We proceed to select associated components to be assembled for the simulation model associated with the problem domain, and execute the simulation model. Throughout its execution, we get outcomes such as measure of performance or measure of effectiveness to give a resolution to the relevant problem.

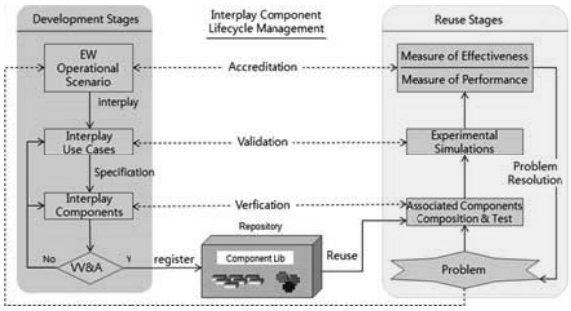


Figure 4: Life cycle management rules

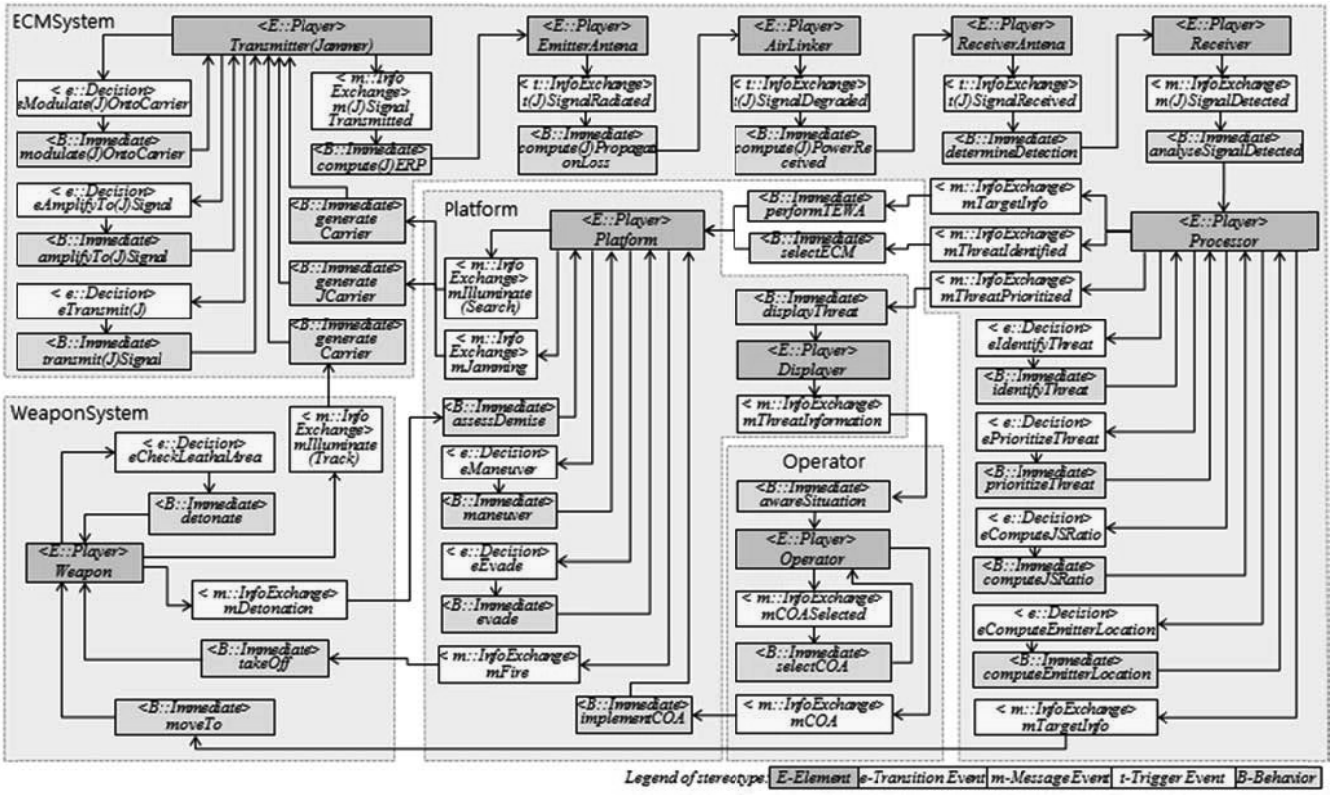


Figure 5: EW reference conceptual model

VV&A Rules

Verification, validation and accreditation (VV&A) are important to reach at the right resolution of the problem. Verification is the process to assure that interplay components are implemented in the right way as specified in the interplay use case model, additionally to assure that interplay use cases models are specified in the right way. On the other hand, validation is to assure the fidelity of the interplay use case specification enough to be used for the intended purpose. Finally accreditation is the official certification of the user authority. VV&A should be performed in the life cycle of the interplay components as shown in Figure 4.

Further, the conceptual entities in the conceptual reference model should be verified prior to the specification of the interplay use cases. The interplay use cases should be validated to ensure the right depiction of a real-life EW engagement scenario.

AN ILLUSTRATION OF SOME IMPLEMENTATION OF OAFEw AND ITS USEFULNESS

As aforementioned, an open architecture framework for electronic warfare modeling & simulation (OAFEw) consists of a conceptual reference model, an interplay use case model, an interplay component library, and a simulation model. In order to show the usefulness of OAFEw and its validity, we implemented OAFEw version 1.0. By the suffix version 1.0 of the OAFEw, we mean that OAFEw is still under going for its elaboration.

EW Conceptual Reference Model Version 1.0

The first implementation of OAFEw is its conceptual reference model. By now version 1.0 was implemented. Basically, OAFEw conceptual reference model version 1.0 includes the static and dynamic terms of the EW simulation entities which may provide a EW simulation modeler with a common dictionary sharing common semantics between OAFEw constituents. Figure 5 shows the conceptual reference model depicted using CML explained in Section 3. The OAFEw conceptual reference model consists of 4 components: *ECMSystem*, *WeaponSystem*, *Platform*, and *Operator*. *ECMSystem* comprises *Transmitter*, *EmitterAntena*, *AirLinker*, *ReceiverAntena*, *Receiver*, and *Processor*. *Transmitter* or *Receiver* may share a common antenna, so EmitterAntena or ReceiverAntena may be the same. *AirLinker* is an EW environment degrading the signal transmitted. Each of *ECMSystem* components have events and their behaviors invoked by the events as shown on the top and the right hand side of Figure 5. *WeaponSystem* is a threat in the EW simulation. *WeaponSystem* is equipped with an illuminator containing Transmitter and Antenna. an example of *WeaponSystem* may be a surface to missile. Platform is the platform of a weapon system. They may be also equipped with either *WeaponSystem* or *ECMSystem*, or both. An aircraft may for example become *Plaform* containing *ECMSystem*. The aircraft is a target in a classical EW simulation. The aircraft detects the threat signal of a missile or the missile launcher, and then jams the receivers of them in the EW simulation.

An Illustrative Example of Some Implementation

The interplay use cases was implemented with the EW interplay components with C++ and stored in the EW interplay component repository. The EW interplay components become building blocks of a EW simulation. It allows users to compose the interplay components according to the engagement scenario of a wide variety of EW threats, targets, and electronic countermeasure equipment without having to reprogram of the simulation software at the outset. Then we had the EW simulation model configuration for our illustrative example, which is shown in Figure 6.

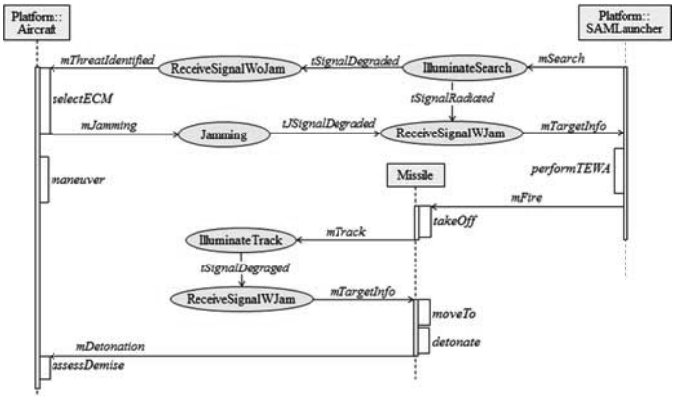


Figure 6: An example of EW simulation configuration

Figure 7 shows a snap shot of its execution. It is a constant time-step driven simulation of the missile threat versus the aircraft target engagement. During a run of the engagement simulation, all positions of the entities are stepped through time and RF signal propagation is simulated. This demonstration shows a one-on-one engagement where the left one is an air threat equipped with an emitter, and the other is a ground launcher target equipped with a jammer.

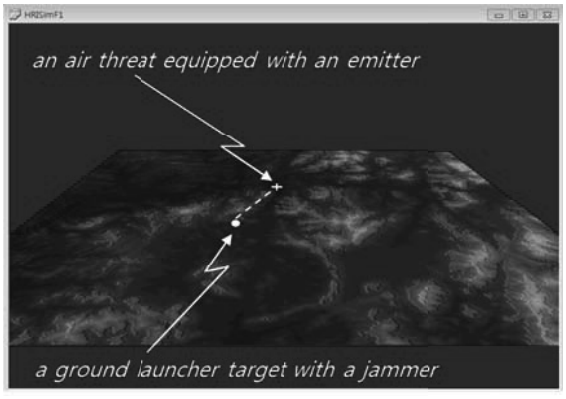


Figure 7: Snap shot of the EW simulation execution

CONCLUSIONS

In this paper, we explained an open architecture framework for an electronic warfare modeling & simulation (called "OAFEw") for the development of an EW simulation in a reusable and interoperable way. By implementing OAFEw version 1.0, we have seen its usefulness. However, its fuller application is still on going, which is left for us as further works.

ACKNOWLEDGEMENT

This study was supported by the Research Program funded by the Agency for defense development in South Korea.

REFERENCES

Schleher, D. Curtis. 1999. Electronic warfare in the information age, Norwood, MA: Artech House Publishers, 624 p.

IEEE standard for modeling and simulation (M&S) High Level Architecture (HLA)-Framework and rules. 2000. IEEE Std 1516-2000, IEEE, USA.

IEEE standard for modeling and simulation (M&S) high level architecture (HLA)-Federate interface specification. 2000. IEEE Std 1516.1-2000, IEEE, USA.

IEEE standard for modeling and simulation (M&S) high level architecture (HLA)-Object model template (OMT) specification. 2000. IEEE Std 1516.2-2000, IEEE, USA.

IEEE: Base Object Model (BOM) Template Specification. 2005. IEEE Std 003-2006.

IEEE: IEEE Recommended Practice for High Level Architecture (HLA) Federation Development and Execution Process (FEDEP). 2003. IEEE Std 1516.3-2003.

IEEE: IEEE Recommended Practice for Distributed Simulation Engineering and Execution Process (DSEEP). 2011. IEEE Std 1730-2010.

Clark R. Karr, Conceptual Modeling in OneSAF Objective System, Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC).

Coad, P., Lefebvre, E., and DeLuca, J. 1999, ″"Java Modeling In Color With UML: Enterprise Components and Process″", Book and CD-ROM edition, Prentice Hall Professional Technical Reference (PTR), June 15.

Ray E. Gallagher, et al, EW-GEMS: System concepts and development process of an electronic warfare simulation, Proceedings of the 1988 Winter Simulation Conference, pp.783-790.

S.K. Gupta, G. Siva Prasad, and J. Nanda Kishore, Electronic Warfare Simulation-based on Service Oriented Architecture, Defense Science Journal, Vol. 62, No. 4, July 2012, pp. 219-222, DOI: 10.14429/dsj.62.929, 2012, DESIDOC.

AUTHOR BIOGRAPHY

SANG YEONG CHOI was born in Busan, Korea and went to Korean Military Academy, Korean National Defense University, Cranfield University of UK, where he studied defense science and obtained Ph.D. in 1989. He is an adjunct professor in School of Defense Science, Hansung University and MyongJi University. Previously, he was a professor at Korean National Defense University.

ON THE EFFECTS OF THE VARIATIONS IN NETWORK CHARACTERISTICS IN CYBER PHYSICAL SYSTEMS

Géza Szabó, Sándor Rácz
Ericsson Research, Budapest, Hungary
Email: {geza.szabo,sandor.racz}@ericsson.com

József Pető
Budapest University of Technology and Economics, Hungary
Email: pjoejoejoe@gmail.com

Rafael Roque Aschoff
Pernambuco Federal Institute of Education, Science, and Technology, Brazil
Email: rafael.aschoff@gprt.ufpe.br

KEYWORDS
network characteristics, cyber-physics, Gazebo

ABSTRACT

The popular robotic simulator, Gazebo, lacks the feature of simulating the effects of control latency that would make it a fully-fledged cyber-physical system (CPS) simulator. The CPS that we address to measure is a robotic arm (UR5) controlled remotely with velocity commands. The main goal is to measure Quality of Control (QoC) related KPIs during various network conditions in a simulated environment. We propose a Gazebo plugin (OurPlugin 2017) to make the above measurement feasible by making Gazebo capable to delay internal control and status messages and also to interface with external network simulators to derive even more advanced network effects. Our preliminary evaluation shows that there is certainly an effect on the behavior of the robotic arm with the introduced network latency in line with our expectations, but a more detailed further study is needed.

INTRODUCTION

A cyber-physical system (CPS) is a mechanism controlled or monitored by computer-based algorithms, tightly integrated with the internet and its users. Unlike more traditional embedded systems, a full-fledged CPS is typically designed as a network of interacting elements with physical input and output instead of as standalone devices. For tasks that require more resources than are locally available, one common mechanism is that nodes utilize the network connectivity to link the sensor or actuator part of the CPS with either a server or a cloud environment, enabling complex processing tasks that are impossible under local resource constraints. Among the wide diversity of tasks that CPS is applied we focus on robot control in this paper.

Currently, one of the main focus of cloud based robotics is to speed up the processing of input data collected from many sensors with big data computation. Another approach is to collect various knowledge bases in centralized locations e.g., possible grasping poses of various 3D objects.

Another aspect of cloud robotics is the way in which the robot control related functionality is moved into the cloud. The simplest way is to run the original robot specific task in a cloud without significant changes in it. For example, in a Virtual Machine (VM), in a container, or in a virtualized Programmable Logic Controller (PLC). Another way is to update, modify or rewrite the code of robot related tasks to utilize existing services or APIs of the cloud. The third way is to extend the cloud platform itself with new features that make robot control more efficient. These new robot-aware cloud features can be explicitly used by robot related tasks (i.e. new robot-aware services or APIs offered by cloud) or can be transparent solutions (e.g., improving the service provided by the cloud to meet the requirement of the robot control).

Designing cyber-physical systems is challenging because of a) the vast network and information technology environment connected with physical elements involves multiple domains such as controls, communication, analog and digital physics, and logic and b) the interaction with the physical world varies widely based on time and situation. To ease the design of CPS, robot simulators have been used by robotics experts. A well-designed simulator makes it possible to rapidly test algorithms, design robots, perform regression testing, and train AI system using realistic scenarios.

There are various alternatives, sets of tools that make it possible to put together a CPS simulation environment, but it is very difficult, needs a lot of interfacing with various tools and impractical. The requirements of a widely applicable CPS are the following:

- Should be modular in terms of interfacing with the CPS
- Should be modular in terms of interfacing with network simulator, realization environment
- Should be able to cooperate with widely applied environments

We chose Gazebo as our target robot simulation environment that we intend to extend with new functionalities to make it capable of being applied as a CPS. Gazebo (Gazebo 2017) offers the ability to accurately and efficiently simulate populations of robots in complex indoor and outdoor environments. It has a robust physics engine, high-quality graphics, and convenient programmatic and graphical interfaces. Gazebo is free and widely used among robotic experts.
The main challenge with the design principle of Gazebo is that the control of actuators is deployed and run practically locally to the actuators. In this case, there is no need to consider the effects of a non-ideal link between the actuator and the controller. Considering the CPS context, as controllers are moved away from actuators, it becomes natural and even necessary to analyze the effects of the network link between them.
Gazebo has a plugin system that we target to use to provide us an interface to our modular network simulation environment. The goal of this paper is to show the design principles of the network plugin and provide the research community with a tool for further research in CPS.

THE MEASUREMENT SETUP THAT WE GO FOR

The CPS that we address to measure is a robotic arm (UR5 (UR5 2017)) controlled remotely with velocity commands. The main goal is to measure Quality of Control (QoC) e.g., cumulated PID error during trajectory execution, cumulated difference in joint space between the executed and calculated trajectories, etc. related KPIs during various network conditions in this setup.
Figure 1 shows the use case with real hardware that we target to simulate in Gazebo. The left side of the figure (Hardware) shows the same data elements described in (roscontrol 2017), whereas the right side of the picture (Realization) uses the same colors for the boxes to describe a specific realization. In the specific case, the UR5 can be accessed via TCP/IP ports 50001 to send command messages and port 50003 to read the robot status messages. The trajectories are computed by MoveIt (MoveIt 2017). MoveIt sends trajectories to the controller manager which starts a velocity controller (yellow), a specific type of `ros_control`. The `ur_modern_driver` (Andersen 2015) implements the hardware resource interface layer by simply copying the velocity control packets to the proper TCP sockets. A middle node can be deployed between the robot driver and the robot (green) that can alter the network characteristics.
A trivia approach to setup the above architecture in a simulation environment is provided by Universal Robots. Universal Robots simulator software (URSim 2017) is a java software package that makes it possible to create and run programs on a simulated robot, with

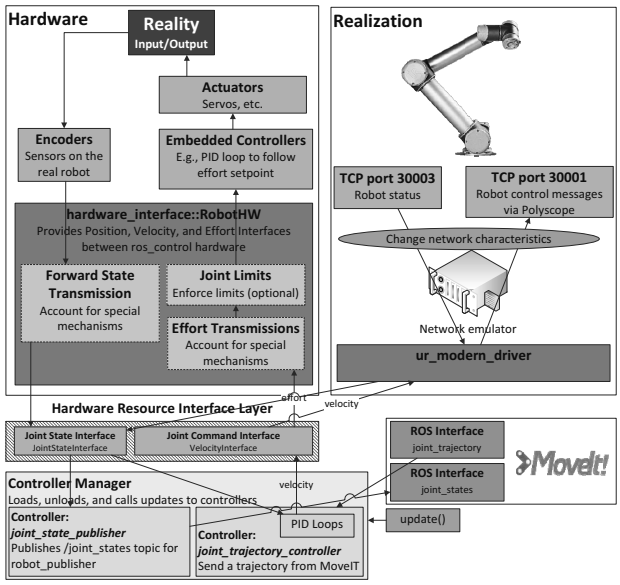


Figure 1: Target architecture to be realized with simulator

some limitations. The limitation of this solution is that it is capable to simulate only one robot. There is no chance to integrate the robot in complex environments as you can configure with Gazebo e.g., interacting with other mechanical elements in the workspace, check collisions with the environment, etc.

MOTIVATION AND RELATED WORK

COMPETITIONS

A frontier method to push research groups to their limits is to organize competitions. DARPA, a research group in the U.S. Department of Defense, announced the DARPA Robotics Challenge with a US $2 million dollar prize for the team that could produce a first responder robot performing a set of tasks required in an emergency situation. During the DARPA Trials of December 2013, a restrictive device was inserted into the control computers of each competing team and the computer that formed the 'brain' of the robot. The intent of the network degradation was to roughly simulate the kind of less than perfect communications that might exist during those kinds of emergency or disaster situations in which these robots would be deployed. The restrictive device –, a Mini Maxwell network emulator from InterWorking Labs – alternated between a 'good' mode and a 'bad' mode of network communication, every sixty seconds. 'Good' minutes permitted communications at a rate of 1 Mbps (in either direction) and a base delay of 50 ms (in each direction.) 'Bad' minutes permitted communications at a rate of 100 Kbps (in either direction) and a base delay of 500 ms (in each direction.) At the end of each minute, a transition occurred from bad-to-

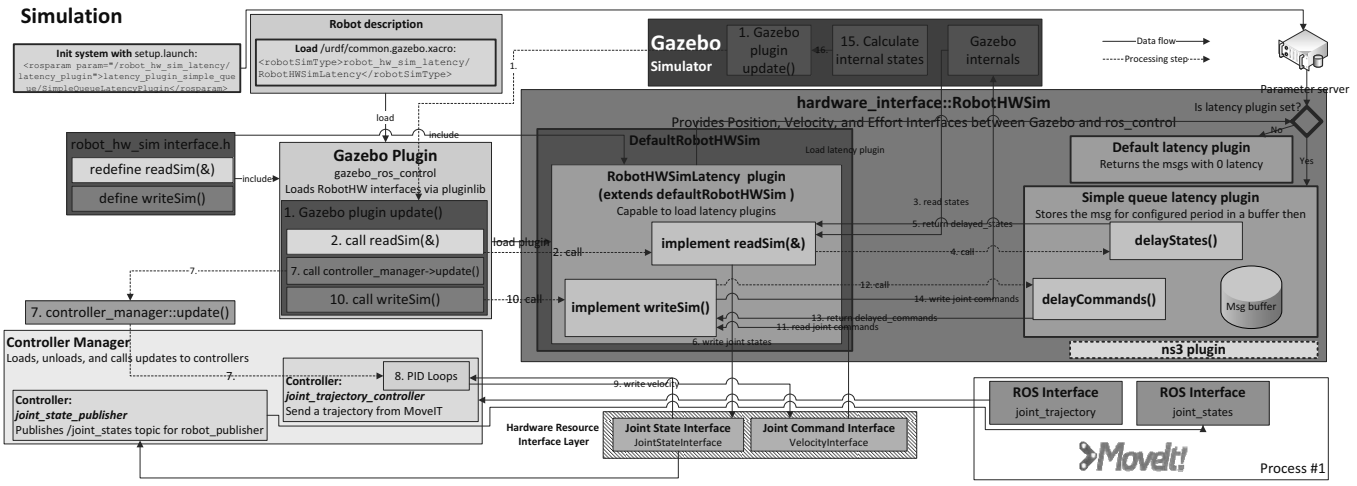


Figure 2: Gazebo architecture

good or good-to-bad. A side effect of these transitions was packet-reordering.

The impact of network degradation on the teams was larger than expected. Informal feedback suggested that several teams did not realize that rate limitation induces network congestion or the ramifications of that congestion. Some teams appeared to have been surprised by the behavior of the network protocol stacks, particularly TCP stacks, in the operating systems underneath their code. (DARPA 2017) The above experiences would have been probably less striking to the teams if they were able to test the network characteristics changes in a simulation environment.

A recent competition Agile Robotics for Industrial Automation Competition (ARIAC)(ARIAC 2017) targets industrial related applications. ARIAC is a simulation-based competition is designed to promote agility in industrial robot systems by utilizing the latest advances in artificial intelligence and robot planning. There is no tricky network environment in the ARIAC competition. The industry relies on robust low-delay protocols. That is why it is an interesting aspect to see what happens when those links and protocols are exchanged. For instance, what are the possible performance improvements or degradation when the control or sensors data processing in an industrial scenario are moved further away from the actuators and how different protocols would fare under various network characteristics?

WHY GAZEBO?

In both of the above competitions, Gazebo provided the simulation infrastructure. In a more structured study about the level of how wide-spread the various simulator tools were done in (Ivaldi et al. 2014). It showed that Gazebo emerges as the best choice among the open-source projects.

Authors of (omnetros 2017) describes some early experiments in linking the OMNET++ simulation framework with the ROS middleware for interacting with robot simulators in order to get within the OMNET++ simulation a robot's position which is accurately simulated by an external simulator based on ROS. The motivation is to use well-tested and realistic robot simulators for handling all the robot navigation tasks (obstacle avoidance, navigation towards goals, velocity, etc.) and to only get the robot's position in OMNET++ for interacting with the deployed sensors. Our goal is the other way around, thus to introduce the effects of the network simulator into the robot simulator.

The roadmap of Gazebo development shows that version 9.0 arriving at 2018-01-25 will have support to integrate network simulation (ns-3 or EMANE). Further information regarding if this feature will be like (omnetros 2017) or the one we propose in this paper is not available yet.

PROPOSED METHOD TO SIMULATE THE EFFECTS OF NETWORK CHARACTERISTICS

In ROS, topics (rostopics 2017) are named buses over which nodes exchange messages. ROS currently supports TCP/IP-based and UDP-based message transport. ROS nodes are standalone executables running with individual process IDs in the operating system. One practical way to introduce latency in current ROS deployment is via defining network namespaces among nodes. For a certain namespace, custom delay, jitter, drop characteristics can be defined with `tc` like in (nwnstc 2014). The main issue is that there is a MoveIt node as an individual process, but the whole joint controller-actuator control loop is realized within Gazebo as one other process. The only topic based communication happens between the MoveIt and the monolith Gazebo process. So this kind of solution cannot be applied to our problem.

We have to dig deeper in the architecture of Gazebo and realize the CPS system within. To keep the architecture modular, we decided to implement the proposed method as a Gazebo plugin. While the setup most of these plugins fits well in the current Gazebo architecture and can be done via configuration files, there are still patches needed to be applied on core functional elements of the Gazebo code.
Figure 2 shows the architecture of the proposed method. The coloring of the figure follows the way in (roscontrol 2017). Green represents new added plugins, modules, functionalities. The working of the system is the following. As a first step –, a launch file that triggers the whole simulation to run – setups a parameter on the ROS parameter server. This parameter defines the specific latency plugin that will be loaded.
The launch file initiates the Gazebo simulation. Gazebo loads the `gazebo_ros_control` plugin (left most blue box) that main purpose is to interface with the ROS controller manager. This module needed a small tweak. The original code passed the address of the messages from the controller manager to the simulation, performed the actions (read status, calculate commands) triggered by the `update()` function in a sequential manner. There was no modification of these input variables during the calculations in the original code. In our system, the messages are copied and stored to make it possible to perform further actions on the messages.
Gazebo loads configuration files from the `common.gazebo.xacro` file in which it is specified that our custom `RobotHWSimLatency` plugin should be loaded instead of the `DefaultRobotHWSim` plugin. Our `RobotHWSimLatency` plugin is the extension of the `DefaultRobotHWSim` plugin with modified read and write functions and with the task to load a custom latency plugin. The latency plugin to be loaded is the one that was setup by the parameter server. The current options include a) the default latency plugin that practically returns the messages with no introduced latency and b) the simple queue latency plugin. This later has a configurable size of the queue to store the messages in them. In each simulation tick (100Hz), the messages are shifted one position forward in the queue and when they reach the end of the queue they are provided to Gazebo as the currently valid message. In the same way, an interface plugin to cooperate with external network simulators like ns3 (ns3 2017) can be also implemented here. We described the detailed call sequence of the plugin system in details on (OurPlugin 2017).

EVALUATION

We evaluated our proposed method on various Key Performance Indicators (KPIs). The most straightforward evaluation is the visual inspection of the robotic arm movement. For this purpose, we loaded the robot model into rviz and used a ros package to visualize the markers on the way the robotic arm passed through. Figure 3 is a screenshot from rviz which shows the visualized trajectories. The bottom left corner of the picture is the starting point of the robotic arm. It passes through the waypoints one-by-one from number 1 to 5. The black lines are the trajectories, while the lines with various colors show the effect of introducing latency into the system. The cyan color shows the reference scenario with 0 latency. In all other cases, we introduced latency in the system in both the command writing and status reading direction and rerun the trajectory planning and execution scenario. The upper right corner of the picture shows a magnified part around the trajectories. The trajectories were planned with the `RRTConnectkConfigDefault` planner.

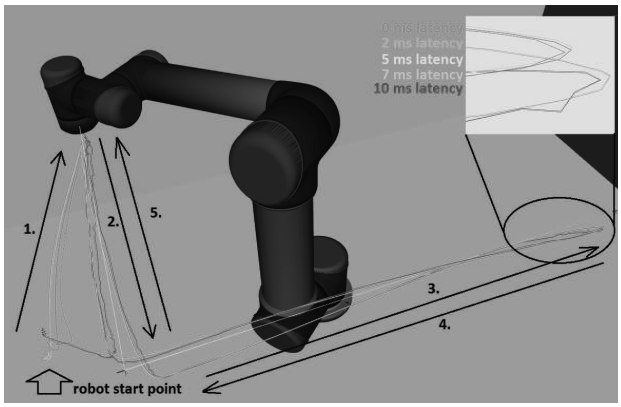


Figure 3: The visualized trajectories

The visualized trajectories show the expected behavior of the system. Increasing the latency increases the deviance from the original trajectories. It should be noted that the planned trajectories are straight in Cartesian-space. To move along these trajectories the robotic arm needs complex movements in the joint-space, thus even the movement in a straight line causes deviation from the reference trajectory. In the other way around, if the planned trajectories were straight in the joint-space, we would see a movement in circles by the robotic arm, but the effect of the latency was more negligible.
Figure 4 shows the velocity commands sent to the robot in the function of time. Analyzing the velocity commands in such details reveals that comparing the different scenarios are not straightforward for several reasons. One is that the planning is non-deterministic, and a slight difference during the initialization of the gazebo environment ends up with some different planned trajectory. The execution of the trajectories depends on the environment status as well, and it is never the same. Joint 4 shows the expected effect on the velocity commands levels as well, thus the induced latency causes increased velocity command deviation compared to the reference scenario. It is also a clear observation that around 10 ms latency, the system starts to get unstable.

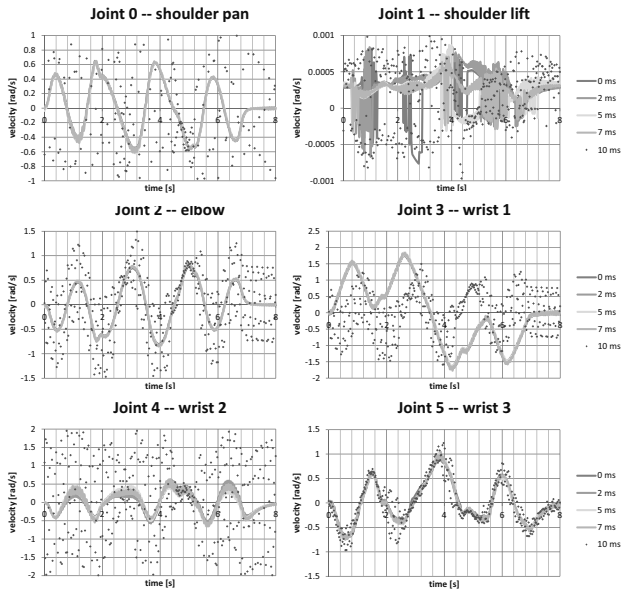


Figure 4: The velocity commands sent to the robot

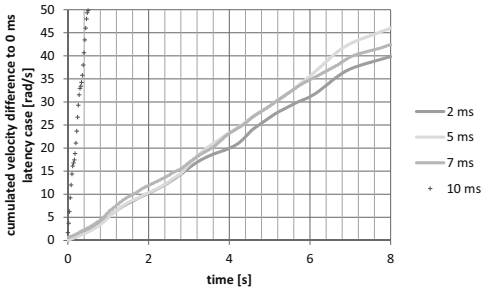


Figure 5: The cumulated difference of the velocity commands comparing to the reference scenario

This is likely due to the various updating frequency parameters that Gazebo employs to run the simulation. It needs definitely further work to make it clear how the introduced latency affects other characteristics or behaviors, such as the robot commanding frequency, whole physical simulation steps, internal message timings.
Figure 5 shows the cumulated difference of the velocity commands comparing to the reference scenario. The 2 ms latency scenario is the closest to the reference as it is expected. In the first 3 sec of the trajectory execution the 5 ms scenario is closer to the reference than the 7 ms scenario, but around 6 sec, the 5 ms scenario collects so much error that shows bigger deviation than the 7 sec scenario. The 10 sec scenario has another magnitude of error, and thus cut off the diagram after the first second.

CONCLUSION AND FURTHER WORK

In this paper, we proposed a plugin (OurPlugin 2017) to extend the capabilities of the current Gazebo robotic simulator and turn it into a CPS system. The realization of the proposed method is a plugin to Gazebo. The plugin fits into the modular design of Gazebo. As of the interface is available, it eases to test various network effects on the robot control. Based on our preliminary evaluations it does affect the QoC KPIs of the robot control.
The evaluation showed behavior which is expected and reasonable, but also cases which show that the whole system needs fine-tuning. We plan to evaluate the working mechanism of the system with the help of the ROS, gazebo and research communities. We plan to do more extensive measurements with the tool. We plan to interface it with various radio network simulators and see the effects of the radio on the QoC KPIs. In a similar way, we plan to investigate the how the system behaves when taking into account not only the network links characteristics but also the protocols for message exchanging. We also plan to compare the level of similarity of the simulation to real robot HW controlled in a real radio network. We are taking part in the ARIAC competition and we plan to evaluate if the tool can provide any advantage for us in any of the use cases of the competition.

REFERENCES

Andersen T.T., 2015. *Optimizing the Universal Robots ROS driver*. Tech. rep., Technical University of Denmark, Department of Electrical Engineering.

ARIAC, 2017. *Agile Robotics for Industrial Automation Competition (ARIAC)*. URL `http://gazebosim.org/ariac`.

DARPA, 2017. *Network Emulation at the DARPA Robotics Challenge*. URL `https://iwl.com/white-papers/network-emulation-at-the-darpa-robotics-challenge/`.

Gazebo, 2017. *Gazebo*. URL `http://gazebosim.org/`.

Ivaldi S.; Peters J.; Padois V.; and Nori F., 2014. *Tools for simulating humanoid robot dynamics: A survey based on user feedback*. In *2014 IEEE-RAS International Conference on Humanoid Robots*. 842–849.

MoveIt, 2017. *MoveIt*. URL `http://moveit.ros.org/`.

ns3, 2017. *ns-3*. URL `https://www.nsnam.org/`.

nwnstc, 2014. *Network Namespaces and Traffic Control*. URL `http://gigawhitlocks.com/2014/08/18/network-namespaces.html`.

omnetros, 2017. *How to link OMNET++/Castalia with ROS*. URL `http://cpham.perso.univ-pau.fr/WSN-MODEL/castalia-ros.html`.

OurPlugin, 2017. *Gazebo latency plugin*. URL `https://github.com/Ericsson/robot_hw_sim_latency`.

roscontrol, 2017. *Data flow of ros_control and Gazebo*. URL `http://gazebosim.org/tutorials/?tut=ros_control`.

rostopics, 2017. *ROS Topics*. URL `http://wiki.ros.org/Topics`.

UR5, 2017. *UR5*. URL `www.universal-robots.com/ur5-robots`.

URSim, 2017. *URSim*. URL `https://www.universal-robots.com/download/?option=28545#section16632`.

IDENTIFYING THE OPTIMAL TRANSMISSION RANGE IN DEPTH-BASED ROUTING FOR UWSN

Mohsin Jafri, Simonetta Balsamo and Andrea Marin
DAIS, Università Ca' Foscari Venezia, Italy
email: mohsin.jafri@unive.it, balsamo@unive.it, marin@unive.it

KEYWORDS
Depth-Based Routing, optimal transmission range, energy efficiency, underwater acoustic networks.

ABSTRACT

Routing in Underwater Wireless Sensor Networks (UWSNs) is a challenging problem because of the intrinsic characteristics of this class of wireless networks (long propagation delay, mobility of nodes, etc.) and because of the performance indices that must be taken into account, such as the network throughput, the packet delivery ratio and the energy cost. In particular, routing algorithms must grant a low energy cost in order to maximise the lifetime of the network's nodes. In this study we focus on a popular routing protocol for UWSNs, namely the Depth-Based Routing (DBR). Specifically, we study the impact of the transmission range of the nodes on the network performance indices, with particular attention to its energy efficiency. The study is based on an extensive set of simulations performed in AquaSim-NG using a library that has been developed with the aim of providing an accurate estimation of the nodes' energy consumption. The main outcome of our work is showing the relation between transmission range providing the optimal DBR energy efficiency and the density of the nodes in a UWSN.

INTRODUCTION

Routing protocols in UWSNs aim at providing high network connectivity, low energy consumption and low packet delay by capitalizing the intrinsic characteristics of acoustic communication. From the functional point of view, routing protocols in UWSNs have to transmit the sensed data collected by the underwater nodes to some sink nodes on the surface which will eventually transmit them to the base-station to be processed. Underwater nodes are usually equipped with batteries which are difficult to replace or recharge and for this reason energy preservation must be a key-factor in the design of routing algorithms and a simple flooding strategy turns to be highly inefficient. In the literature, several strategies have been proposed in order to introduce new routing protocols or to optimize previously proposed ones including specific node deployment strategies, localization schemes and transmission range selection.

Depth-Based Routing (DBR) (Yan et al. 2008) is a localization-free routing scheme and only relies on depth information of nodes in order to transfer data from the source to the sink node. When a node transmits a packet all of its neighbors can receive it due to the broadcast nature of the considered acoustic transmission, however only the low-depth neighbors are eligible for forwarding. When a packet is sent, the protocol aims at selecting the neighbor which is nearest to the surface as forwarder so that the number of hops is reduced and, as a consequence, the end-to-end delay and the energy consumption are also reduced.

The actual deployment of a UWSN using DBR must face some design problems concerning the identification of the optimal configuration parameters for the protocol such as the constants for the computation of the holding time and the depth threshold value. However, it should be clear that also the configuration of the physical layer parameters affects the performance of the network. Specifically, the node transmission range strongly influences the energy cost of the protocol. High transmission ranges consume more energy and increase the probability of interferences but allows DBR to cover longer distances with one hop. This consideration suggests that there must exist optimal values for the transmission ranges (see (Zorzi and Pupolin 1995) for an analytical model of terrestrial networks addressing this problem).

Harris et.al (Harris III and Zorzi 2007) propose a simulation model to compute an accurate transmission power required to meet the SNR threshold of $20dB$ at the receiver for various intermediate distances among the nodes. They also devise a model for an acoustic channel and provide its comprehensive implementation in NS2 by employing passive sonar equation. We use this work for modelling the correct transmission events in our simulation model. In the literature of underwater networks, aspects of physical layer have been taken into account for improving the performance of routing and MAC protocols. To this aim, efficient localization strategies, optimal transmission range selection and design of operational modes of acoustic modems showed to be helpful in increasing the network lifetime, improving the robustness of its connectivity and decreasing the end-to-end delay.

Porto et.al (Porto and Stojanovic 2007) propose an extended form of Distance-Aware Collision Avoidance Protocol (DACAP) by augmenting it with optimized transmission power and range selection for sensor nodes. The fine control of these parameters leads to an improvement of the energy efficiency while the network connectivity is preserved. Similarly to the outcome we have in this paper, the authors find out that the selection of the optimal transmission range in DACAP depends on the network density. However, in contrast with DACAP, it is not necessary true for DBR that the optimal transmission range is the minimum radius that ensures the network connectivity as it will be evident from our experiments. Gao et.al (Gao et al. 2012) provide an analytical model for the evaluation of the network power consumption. Based on this model, they propose a method for obtaining the optimal transmission range for a randomly deployed network.
Although all of these papers aim at specifying the optimal transmission range for the combination of some MAC layers and routing protocol, still to the best of our knowledge there is no work considering DBR optimal transmission ranges by taking into account the detailed implementation of the network (e.g., busy terminal problems and so on). To cover this gap, we adopt our implementation (Jafri 2017) of DBR in AquaSim-NG (Martin et al. 2015) which is a NS3 based simulator and its libraries have been designed with a more efficient and detailed simulation framework for UWSNs.

Contributions In this paper we address the problem of estimating the optimal transmission range for DBR based UWSNs by resorting to a detailed simulation model that takes into account a broad set of relevant aspects of actual network deployments. To this aim, we extended the DBR implementation of AquaSim-NG (Martin et al. 2015) in order to include an accurate modelling of nodes' energy consumption taking into account the operational modes of the modems. The simulator is open access and can be downloaded from the official repository of AquaSim-NG (Jafri 2017). We emphasized the cross-layer interactions between the physical and the routing layer. Finally, our simulation model is able to tackle the problem of the busy terminal which is well-known to be important for the estimation of the network energy efficiency (Yan et al. 2008). We have considered several scenarios and we have experimentally derived a relation between the optimal transmission range and the node density.

DBR AND ITS SIMULATION MODEL

In this section we briefly recall DBR and present the main features of our simulation model. We take a bottom up approach based on the layer partition of the protocol stack. Particular attention will be devoted to the analysis of the power consumption and the loss probability at the physical layer.

Modelling the power consumption at the physical layer

At the physical layer, the transmission power consumption of an acoustic signal in UWSNs is computed by using the passive sonar equation presented in (Harris III and Zorzi 2007) which gives the Signal to Noise Ratio (SNR) at the receiver based on some parameters among which a major role is played by the transmission power and the Attenuation-Noise (AN) factor. This last factor is computed according to the well-known Thorp's formula (Harris III and Zorzi 2007):

$$10\log_{10}\alpha(f) = \begin{cases} 0.11f^2/(1+f^2) + 44f^2/(4100+f^2) \\ +2.75*10^{-4}f^2 + 0.003 & \text{if } f \geq 0.4\text{kHz} \\ \\ 0.002 + 0.11(f/(1+f)) + 0.011f & \text{if } f < 0.4\text{kHz} \end{cases}$$

where f is the frequency measures in kHz and $\alpha(f)$ is measured in dB/km.
Total attenuation $A(l,f)$ is computed by combining the total absorption loss $\alpha(f)$ and the spreading loss:

$$10logA(l,f) = k*10log(l) + l*10log(\alpha(f)), \quad (1)$$

where k is the spreading coefficient. Following (Harris III and Zorzi 2007), we compute the total attenuation in $dbre\mu Pa$ which is the standard unit used to compute the signal loss in acoustic communications. The first term of Equation (1) models the spreading loss and the second the attenuation loss.
Algorithm (1) accurately predicts the required transmission power considering various distances between the communicating nodes. By targeting specific SNR at the receiver, the passive sonar equation gives the required transmission power which majorly increases with the distance (see, e.g., (Harris III and Zorzi 2007)).

DBR network layer and its simulation model

In DBR, nodes use pressure-based sensors to estimate their depth and rely on this information to transmit the packets to the on-surface sink. As DBR is a controlled-flooding based scheme the correct setting of its parameters, namely the *depth threshold* and the *holding time*, plays a pivotal role for obtaining high performance with a low energy consumption. Intuitively, the forwarder selection is based on the packet scheduled sending time which is decided on the basis of computation of the holding time. The packet holding time is proportional to the depth difference between the sender and the candidate forwarder and hence it favors the nodes that allow the packets to cover longer distances towards the

Algorithm 1 Computation of transmission power consumption

```
AN[i] ← Attenuation Noise factor for ith frequency
of signal bandwidth
k ← Spreading coefficient
d ← Euclidean distance between nodes
Thorp(f[i]) ← attenuation loss for ith frequency
of signal bandwidth
Noise(f[i]) ← noise loss for ith frequency of signal
bandwidth
Pr ← SNR threshold of receiver
Pt ← Transmission power required to successfully
transmit signal
Num_freq ← Number of frequencies in the
bandwidth of signal
DI ← Directivity Index

for i ← 0 to Num_freq do
    AN[i] ← − (k ∗ 10 ∗ log10(d) + d ∗ Thorp(f[i]) +
    DI + log10(Noise(f[i])));
    if AN[i] > AN[max_index] then
        max_index ← i
    end if
end for
Pt = Pr − AN[max_index];
return Pt;
```

sinks. The depth threshold is used to prevent nodes with low depth difference to become candidate forwarders. During the holding time duration, nodes discard the enqueued packet upon finding its transmission from a lower depth neighbor. DBR targets lowest depth neighbor of sender as an optimal packet forwarder which is also helpful in suppressing transmissions of other eligible neighbors of sender node.

According to (Yan et al. 2008) in DBR the holding time is obtained as follows:

$$f_{\mathrm{DBR}}(d) = \left(\frac{2\tau}{\delta}\right) * (T - d),$$

where T is the maximal transmission range of a node, τ is the maximum propagation delay of one hop, i.e., $\tau = T/v_0$ (where v_0 is the sound propagation speed in the water), d is the depth difference between the sender and the receiver and δ is a scaling factor of the holding times which is chosen in order to achieve the optimal performance of the network and to minimize the hidden terminal problem. The analysis of the impact of these configuration parameters on the network performance has been done in (Yan et al. 2008). Nevertheless, in this paper we focus on the impact of a configuration parameter at the physical layer, namely the transmission range, on the network performance expressed in terms of the expected packet delivery ratio and the energy cost.

PROBLEM STATEMENT

When we deploy an UWSN using DBR routing protocol, the setting of the network layer parameters, i.e., the holding time and the depth threshold, is helpful to minimize the energy consumption but may be not sufficient. In fact, the selection of an optimum transmission range at the physical layer may drastically reduce the network energy cost (and hence its lifetime) while maintaining a reasonably high packet delivery ratio. Transmission range plays a pivotal role in determining the energy consumption and the packet delivery ratio in a UWSN implementing DBR. Let us focus on the energy cost defined as the expected energy required to successfully send a packet to the sink node. Short transmission ranges cause problems in the network connectivity and hence frequently require packet retransmissions that cause a high energy consumption. On the other hand very long transmission ranges require more energy per packet and cause the increase of the number of redundant transmissions caused by hidden terminals. In this work, we seek the optimal value of the transmission range given a certain node density that results in a low energy consumptions and maintains a reasonable high packet delivery ratio. Moreover, an appropriate choice of the transmission range reduces the busy terminal problem (Zhu et al. 2014) by limiting the burden on more stressed nodes from the network traffic.

SIMULATION EXPERIMENTS

In this section we address the problem of identifying the optimal transmission range of sensor nodes with respect to the energy cost of the network by resorting to the simulation model. We study UWSNs with various numbers of nodes deployed in a fixed space of $500m \times 500m \times 500m$ according to a uniform random distribution. The number of nodes varies from 100 to 800 and hence we recreate the scenarios that are similar to those that have been previously studied for other purposes in (Yan et al. 2008). The depth-threshold is 1/4 of the maximum transmission range, and the mobility pattern is a random walk. For MAC layer, we implement Broadcast MAC protocol (Mirza et al. 2009) which efficiently supports the functioning of flooding-based routing protocols. The source node is placed in the bottom of the network. Multiple on-surface sinks have been deployed and the source node transmits a single packet after every two seconds. Table 1 summarizes the experiment setting.

In order to identify the optimal transmission range, we compute the following performance indices: (i) Energy cost of network defined as the expected energy required to successfully deliver a packet measured in Joule per packet, (ii) Packet delivery ratio and (iii) Total number of transmissions of network. For each measurement we performed 20 independent experiments and build the

Parameter	Value
Network size	500m ×500m ×500m
Deployment	Random uniform
Initial energy of nodes	500J
Packet size	64 Bytes
Node mobility speed	2 m/s
Receiving power consumption	0.1 W
Idle power consumption	1 mW
Mobility pattern	Random walk
δ	Transmission range
f	$3kHz$

Table 1: Simulation Parameters

confidence intervals at 95% whose width is always below 7% of the measured value.

Impact of transmission range on the energy cost of network, packet delivery ratio and total number of transmissions

In this experiment we study the network energy cost as function of the transmission range of the sensor. Figure 1 shows the results of our experiments, i.e., the estimates of the energy cost of the network as a function of the transmission range for networks with 500 to 800 nodes. We observe that for very low transmission ranges the cost of retransmissions due to broken routes becomes prohibitive from the point of view of the energy consumed by the networks, whereas as the transmission range increases we have both to face the problem of the higher cost for the transmission of the single packet and the explosion of the number of retransmissions due to the hidden terminal problem and the consequent increased number of collisions.

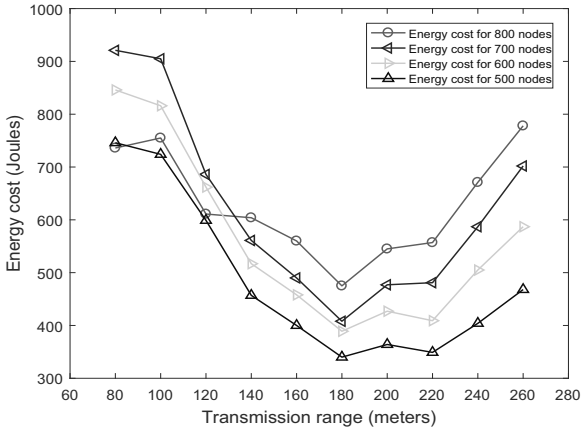


Figure 1: Energy cost of the network as a function of the transmission range.

We can also observe that as the density of the nodes increases, the cost for redundant transmissions and the consequent collisions become dominant in increasing the energy cost of the network even in its optimal working point. For the four considered network densities we have an optimal transmission range of approximatively 180 meters. We will see later on that above a certain density of nodes the optimal transmission range tend to stabilize to this value under the assumptions of Table 1.

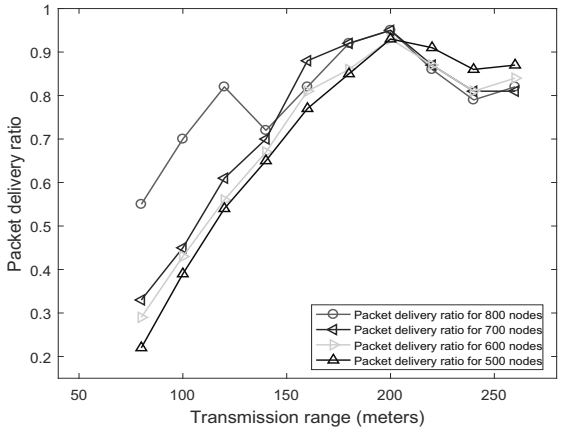


Figure 2: Packet delivery ratio with different node densities.

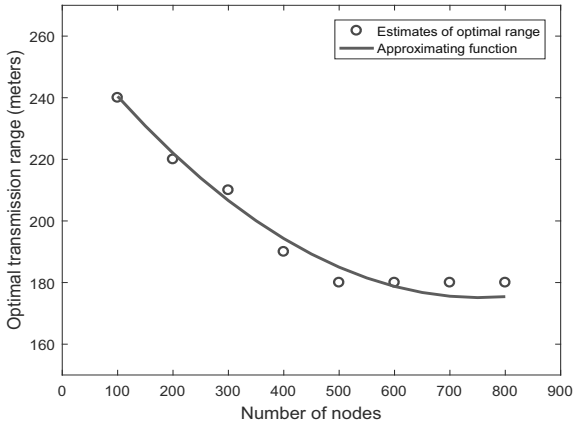


Figure 3: Optimal transmission range for different node densities with minimum energy cost.

Consulting Figure 2, the packet delivery ratio quickly increases with the sharp increase in the transmission range thanks to availability of multiple paths between source node and the sinks. However, after reaching at the maximum point, it declines due to the redundant transmissions and problems caused by the busy terminals. Interestingly, the transmission range associated with the optimal packet delivery ratio is coherent with the value which optimize the energy cost.
It is also worth of notice that as observed in (Zhu et al. 2014) there is a strong correlation between high packet

delivery ratio and reduction of the busy terminal problem. It is worthwhile of notice that the packet delivery ratios decrease after reaching the maximum but appear to become more stable. Also for what concerns the optimal packet delivery ratio, the experiments suggest that the networks with density of 500 nodes outperform those with higher densities in case of transmission ranges longer than 200 and this may suggest that finding the optimal densities could be an interesting problem for future works.

Optimal transmission range as function of the node density

In order to experimentally study the connection between the optimal transmission range and the network node density we have run a large set of simulations for each given density and identified the optimal value for the energy cost. This has been done by assuming the convexity of the function $E_c = f(r)$, where E_c is the energy cost as function of the transmission radius r. Then we have proceeded by using a bisection method. Figure 3 shows the optimal transmission range for various numbers of deployed nodes. We observe that for networks with a number of nodes higher than 500 the optimal transmission range stabilizes at approximatively 180 meters. As observed in Section , this value optimizes both the network energy cost and its packet delivery ratio. As the number of deployed nodes decreases, the optimal transmission range increases to 240 meters associated with 100 nodes as number of intermediate forwarders decreases causing the decrease in total energy consumption of network.
According to our experiments if ρ is the network node density expressed in expected number of nodes for km^3, we can say that the optimal transmission range r^* for DBR decreases with higher ρ as:

$$r^* \propto \rho^{1/6}.$$

In Figure 3 we plot the function $745/\rho^{1/6}$ and we can see that it provides a good approximation of the estimates of the optimal range. We observe that this result is quite different from the empiric law proposed in (Porto and Stojanovic 2007) for DACAP where the optimal transmission range was found to decrease with β as $1/\sqrt{\beta}$ where β is the 2-dimensional node density.

CONCLUSION

In this work we have studied the impact of the configuration of the nodes' physical layer parameters on the performance of DBR routing protocol. In order to reach our goal, a new simulator based on AquaSim-NG has been developed that with respect to its predecessors provides an accurate modelling of the modem operational modes, the cross-layer interactions required by this protocol and the busy terminal problem. The simulator can be downloaded at the official repository of AquaSim-NG (Jafri 2017). Specifically, we have addressed the problem of determining the optimal transmission range providing the lowest energy cost given the network density. To this aim we first studied the behavior of the energy cost as a function of the transmission range for networks with given node densities and empirically verified that this optimal value exists. Then, we have looked for this optimum value for different node densities. Finally, we studied the relation between the network density and the optimal transmission range. As expected, we found that sparse networks require higher optimal transmission ranges, but that this values tends to decrease slowly with denser networks.

REFERENCES

Gao M.; Foh C.H.; and Cai J., 2012. *On the selection of transmission range in underwater acoustic sensor networks. Sensors*, 12, no. 4, 4715–4729.

Harris III A.F. and Zorzi M., 2007. *Modeling the underwater acoustic channel in ns2.* In *Proceedings of the 2nd international conference on Performance evaluation methodologies and tools.* ICST, 18–26.

Jafri M., 2017. *AquaSim Next Generation : Libraries, DBR implementation by Mohsin Jafri.* `https://github.com/rmartin5/aqua-sim-ng/blob/master/model/aqua-sim-routing-ddbr.cc`. Accessed: 2017-04-04.

Martin R.; Zhu Y.; Pu L.; Dou F.; Peng Z.; Cui J.H.; and Rajasekaran S., 2015. *Aqua-Sim Next Generation: A NS-3 Based Simulator for Underwater Sensor Networks.* In *Proceedings of the 10th International Conference on Underwater Networks & Systems.* ACM, 18–22.

Mirza D.; Lu F.; and Schurgers C., 2009. *Efficient broadcast MAC for underwater networks. Procedings of ACM WUWNet, Berkeley, CA, USA.*

Porto A. and Stojanovic M., 2007. *Optimizing the transmission range in an underwater acoustic network.* In *OCEANS 2007.* IEEE, 1–5.

Yan H.; Shi Z.J.; and Cui J.H., 2008. *DBR: Depth-Based Routing for underwater sensor networks.* In *Adhoc and Sensor Networks*, Springer. 72–86.

Zhu Y.; Cui J.H.; Peng Z.; and Zhou Z., 2014. *Busy Terminal Problem and Implications for MAC Protocols in Underwater Acoustic Networks.* In *Proceedings of the International Conference on Underwater Networks & Systems.* ACM, 1–11.

Zorzi M. and Pupolin S., 1995. *Optimum transmission ranges in multihop packet radio networks in the presence of fading. IEEE Transactions on Communications*, 43, no. 7, 2201–2205.

ELECTRONICS SIMULATION

DYNAMIC SWITCHING OF PROCESSOR SIMULATION MODEL ACCURACY

Johannes Kohl, Dietmar Fey
Department of Computer Science,
Chair of Computer Architecture
Friedrich-Alexander-University Erlangen-Nurnberg (FAU), Germany
{johannes.j.kohl, dietmar.fey}@fau.de

Jürgen Bäsig
Technische Hochschule Nürnberg
Department Electrical Engineering,
Precision Engineering,
Information Technology
Nürnberg, Germany
juergen.baesig@th-nuernberg.de

KEYWORDS
PLC, processor architecture, design space exploration, processor verification

ABSTRACT

Instruction Set Simulators (ISS) provide faster simulation speed and lower implementation effort compared to Cycle Accurate Simulators (CAS). This high simulation speed is gained by a less detailed model description losing most parts of the architecture specific behavior. In system level simulations a high resolution cycle accurate view of architecture and program behavior is oftentimes only required for a small part of a simulated code. To reach the point of interest with the correct program context the simulation of the entire preceding code is mandatory. By starting the simulation in ISS and switching to the CAS once the marked area is reached the simulation is running faster and the important part is simulated with the desired precision. In this paper we discuss the synchronization issues regarding the dynamic switching between ISS and CAS and state the loss of accuracy of a switched simulation in comparison with full cycle accurate simulation.

INTRODUCTION

In system level simulations of System on Chip (SoC) designs and network simulations as well as in the development of Application Specific Instruction-set Processors (ASIPs) instruction set simulators are essential. Their high simulation speed, executing millions of instructions per second, and the short implementation time as well as their fast adaptability to instruction set changes make them best suited in both cases. These advantages come with an inaccuracy in cycle count as well as a lack of details concerning architecture behavior. If at some point during instruction set simulation the processor model's cycle accurate behavior with a detailed view on the architecture internals is of interest a ISS is not sufficient. Especially if the required high resolution view should equal a full time CAS simulation.
In principle, there are two different ways to estimate the cycle count in question. The first one is the analytical, or mathematical way where an architecture is parameterized and described with mathematical equations. These methods are based on architecture properties as well as on information about the executed code. This can be applied to data extracted from the ISS with little to not traceable impact on the simulation speed. The second one summarizes the average execution time of each instruction during the ISS simulation. However cache effects as well as branch prediction behavior are usually based on average performance values and therefore increase the inaccuracy especially if the considered program part behaves atypically. In case of architecture exploration these two approaches do not deliver the desired details about the actual processor's internal behavior.
A solution combining both cycle count accuracy and a detailed architecture behavior view is the simulation of the considered program part in a CAS environment. At the point in time the ISS reaches the program area of interest the program context is extracted and loaded into the CAS model. After the program section has been executed on the CAS model the context is transferred back to the ISS environment continuing the simulation. In system level simulations all components interacting with the processor model during the cycle accurate phase must be part of the CAS environment. The cycle accurate simulation changes the simulation environment, but it sill leads to inaccuracies. One major reason for this is the unknown cache content after simulator switching. Additionally the change of pipeline depth leads to a pipeline bubble that causes different behavior at the switch point. That effect is further increased if the cycle accurate model supports instruction reordering with multiple dispatches per cycle.
In this paper we introduce the idea of dynamic simulator switching based on a hybrid processor model containing both a cycle accurate as well as an instruction set simulator. The two cores are able to exchange the current program context including registers, caches, branch prediction unit and memory to ensure the correct continuation of program execution. For all components in question a basic behavior model for the ISS that is extended with cycle accurate behavior for the CAS simulation is required. Both simulators are accessing is the same context data. Additionally the core contains a switch point detection logic automatically changing the active core based on predefined switch points. In Fig. 1 the main components of a hybrid processor simulation are shown. Due to interface sharing a switching of cores has no influence on the communication with external components attached to the hybrid core.
This paper is divided into five sections. In this first section

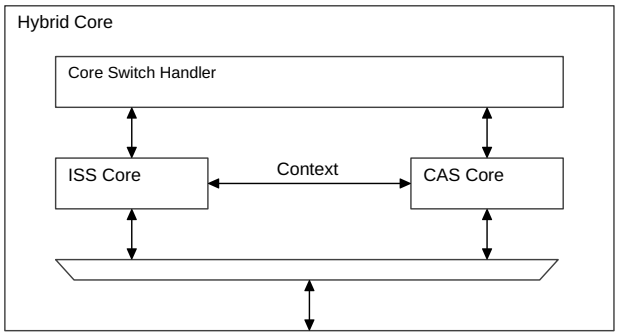


Figure 1: Structure and main components of a hybrid processor model with dynamic accuracy switching.

the idea of dynamic simulation accuracy switching with a hybrid processor model core is discussed. In the next section existing processor simulation and cycle count estimation techniques are evaluated. Afterwards the dynamic core switching and the inaccuracy sources in such a system are explained in detail. The fourth section demonstrates simulation accuracy results based on a hybrid PLC processor simulation model compared to ISS and mathematic cycle count estimations. Finally an overview about future work is given.

Related Work

Dynamically changing the simulation environment during a program simulation is a technique widely applied in development of software. Hybrid simulators like HySim (Kraemer et al. 2007, Jovic et al. 2012) or the hybrid instruction set simulation framework introduced in (Qiu et al. 2011) switch between an ISS and a virtual machine. Parts of the application are directly executed on the host machine in native form. The goal of these hybrid simulators is the gain of simulation speed in software development for Application Specific Instruction Set Processors. In processor architecture development the same technique can be applied to gain a more profound view on a design without losing too much simulation speed. Therefore suitable instruction set simulators and cycle accurate simulators are evaluated in the following.

Several kinds of instruction set simulators are used in academic research, ranging from event based modeling (Luckham and Vera 1995) and timed automata (Wang et al. 2011) to advanced simulation models like Sniper (Carlson et al. 2011), a x86 multi core simulator which tackle the Multi-Core scalability problem. Additionally there are Architecture Description Language (ADL) based ISS, like HARMLESS (Kassem et al. 2009; 2008) and ArchC (Rigo et al. 2004) that loom into the area of CAS. The strength of these simulators is a fast architecture description and generation. However their cycle accuracy is limited. On the side of cycle accuracy fully featured CAS like LISA (Hoffmann et al. 2001), with its just-in-time compiler (Braun et al. 2004), and equally gem5 (Binkert et al. 2011) are present. These simulators provide high accuracy simulation but are limited to a number of predefined architectures that are not suited for special purpose processor simulation.

Another technique is cycle count approximation based on

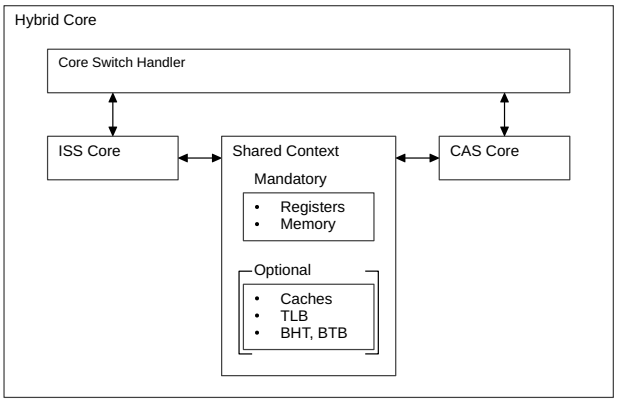


Figure 2: Hybrid core with mandatory as well as optional shared context elements.

mathematical and statistical models. A faster equation based cycle count estimation with interval modelling is presented in (Eeckhout 2010). This method requires the amount of instructions executed, processor architecture properties like dispatch width, cache misses and prediction performance values.

For a hybrid core design with resource sharing it is necessary that both simulators are implemented within the same environment. During a core's inactivity phase any influence on the simulation speed of the running core must be kept to a minimum. Due to the mentioned requirements a hybrid core including an ArchC based ISS and a SystemC based CAS resulting in an overall SystemC base system is a suitable combination. In addition to that SystemC supports state of the art Electronic Design Automation (EDA) simulators and enables switching between ISS and hardware level simulation.

Dynamic Simulator Core Switching

The primary focus of a dynamic hybrid processor simulation is high speed, low detailed simulation up to a switch point and from there on a cycle accurate view on the architecture. If both simulators support the same instruction set switching between them can occur at each point during program execution as long as the program context is transfered correctly. A switch point logic defines the initially active core and controls switching based on program counters. Each time the currently active core changes the program counter it is checked against the switch point list. A simulator change is performed if switch point and switching condition occur at the same time. During a core switch the current processor context is transfered to the other core. All information absolutely necessary to ensure a correct execution of the program must be part of the context. This includes the processor register values and the data and instruction memory content. To speed up the switching between cores we propose resource sharing between simulators. For instance a shared memory both cores operate on makes time intensive copying of content from one core to another obsolete. A more detailed visualization of a hybrid core with necessary and optional shared context elements is shown in Fig. 2.

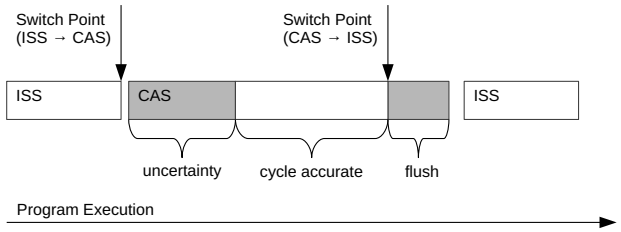


Figure 3: Schematic illustration of a uncertainty gap occurring between ISS and CAS simulation switching.

Switching from ISS to CAS

Switching from an ISS without any architecture modelling to a CAS model with pipelining, caches or branch prediction requires additional steps to ensure correct program continuation. Low detailed ISS can not provide all context information required to initialize the CAS model into the same state it would have within a cycle accurate only simulation at the switching point. This lack of initialization leads to a gap of uncertainty that depends on the architectural differences between the two cores and the extent of context transferred. In figure Fig. 3 a change from ISS to CAS with uncertainty gap is visualized. The gap is overcome at the point when the hybrid core's cycle accurate behavior matches a full time CAS simulation behavior. In case of short programs in CAS path this point might never be reached. Additionally different cache and branch prediction behavior occurring late within the program section cause an insurmountable uncertainty gap.

Another source of inaccuracy is pipelining. The correct initialization of a multiple stage pipeline design in CAS requires that all this information is available in the ISS simulator. But a pure ISS simulator does not provide any pipeline information. In case the ISS supports pipeline modelling the provided pipeline information must be transformed to the CAS model. If the two pipeline models differ in the number of stages the transformation is not practical. Switching in this direction always leads to an empty pipeline and an inaccuracy comparable to a branch misprediction.

Cache Content Inconsistent cache content directly affects correct program execution and must therefore be avoided at all times. If the CAS model contains cache behavior modelling the caches must be brought to a consistent state before continuing the simulation in CAS. This is required because changes made to data and instruction areas during ISS are directly written to the main memory without updating the CAS model's cache. This applies to all memory hierarchy levels between the core and the last level memory that has already been shared.

The easiest way to ensure cache consistency is to clear all instruction set and data cache content before starting the CAS. This leads to a maximal cache miss count within the cycle accurate simulated program part and enlarges the uncertainty gap further. It adds no additional functionality to the ISS it even slows down the switching of simulators a little. Moreover clearing caches leads to a deterministic behavior in CAS program execution because each first access to a memory area leads to a cache miss and makes the inaccuracy predictable.

A second possibility is to make the cache consistent with the memory by updating all stored cache lines to their new values. Depending on the differences between the last program part simulated in CAS and the current program part the result of this method can differ widely. Cache behavior can reach from full cache miss if CAS has recently executed a totally different program area to full cache hit scenarios. This leads to unpredictable accuracy dependent on preceding switch points and is therefore not practicable.

Thirdly a cache model can be added to the ISS simulation tracking memory operations and update the cache content accordingly. At switching point the cache context is simply copied to the CAS cache model This requires identical cache configurations in both models or at best a shared cache resource that doesn't even require the content copying step. Cache models in ISS offer a high cache behavior accuracy for CAS simulations but do not eliminate differences in cache content completely. One source for deviations are out-of-order architectures with multiple in-flight load operations, where cache content can be changed before before the responsible load operation commits to the permanent register map. Additionally, the extra effort to model cache behavior in ISS slows down the simulation. Shared caches can be applied to multilevel memory hierarchies and only increase the modelling and simulation effort within the ISS.

Branch Prediction Incorrect Branch History Table (BHT) and Branch Target Buffer (BTB) content are corrected within the core and increase execution time but do not affect correct program execution. Therefore the prediction unit context updating is optional and due to the single implementation without content copies in main memory limited to content clearing and ISS model context transferring. Leaving the prediction unit state unchanged is possible but depends on the last program part executed in CAS and leads to unpredictable behavior and is therefore not explored.

Analog to caches clearing the BHT results in deterministic behavior and does not affect ISS simulation speed. In most processors the BHT uses a two bit counter for branch behavior prediction that results in a one out of four chance to initialize the BHT content correctly. BTB clearing on the other side can influence the uncertainty gap stronger because target program counter is usually 32 bit wide. Applying the branch prediction models to the ISS and transferring it to or sharing it with the CAS model can decrease the diversities in program execution even further but also increases ISS simulation time.

Switching from CAS to ISS

Switching simulators in this direction does not imply an uncertainty gap but can lead to inconsistent memory content. One important point is the content of write back caches that result in inconsistent memory content if not written back. Flushing a write back cache and transferring all dirty cache

lines to the main memory is mandatory. This increases the time required for switching from CAS to ISS but has no effect on CAS accuracy. If the cache sharing method described above is applied this is no longer necessary because writing back dirty cache lines is performed by the ISS cache model. Another issue are memory write back buffers and currently active write operations on the data bus. Both must be finished before switching to the ISS core to avoid main memory corruptions. Furthermore all currently active read operations on the bus system must be cleared out to avoid open bus transactions at the next CAS activation.

Case study with a Hybrid PLC Processor Simulation

As foundation of the hybrid simulation core SystemC is chosen due to its flexibility, simulation speed and because it is widely used in instruction set and cycle accurate ADLs. A PLC processor model is used to demonstrate the dynamic core switching techniques described above. The PLC architecture is pipelined with multi-issue execution and contains first level caches for data and instructions. Moreover it uses a BHT for branch behavior prediction and a BTB for branch target prediction. The instruction set is compatible with the Siemens Statement List (STL) and contains 140 instruction types adding up to 1,400 instruction variations. The core disassembles instructions into an internal micro code representation with an average ratio of 1.6 micro instructions per macro instruction. The cycle accurate SystemC based model of a PLC processor is designed for a parallel research project and applied here. The ISS part is covered by the already existing ArchC PLC processor model introduced in (Kohl et al. 2015).

Results

In the following section the impact on simulation speed and hybrid simulation accuracy is presented. A set of four test programs is used to evaluate the hybrid PLC core simulator. The first application implements a control loop which contains a discrete PID-Controller and a 1st-order lag element (PT1 element). The second program is a quick sort algorithm sorting 5,000 data sets. A fast Fourier transformation of an input data stream is used as third application. To cover the bit logic instruction field of the PLC processor a CRC calculation based on logic operations is used as a fourth program.

Hybrid Simulation Execution Time

To show the impact of the hybrid simulation approach on execution time the required time for different programs and CAS rates is measured. The CAS program share is increased in three steps from 0 % to 0.5 % of the executed instructions. A zero percent CAS rate equals a full ISS simulation and the 100 % a full CAS. For better diversification three different starting switch points are selected for each program and applied to all CAS rates. The measurement takes the shared resources cache and branch prediction units and the additional computation effort within the ISS into account. As both BHT and BTB show little impact these two shared resources are

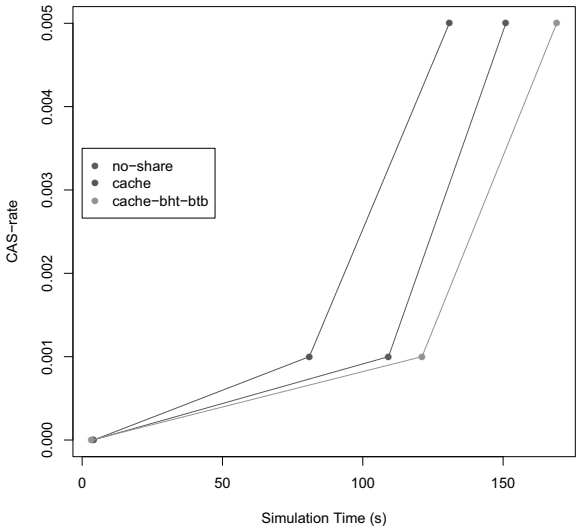


Figure 4: Hybrid core simulation time increase with increasing CAS simulation part ratio.

combined in the simulation tagged ”iss-cache-branch”. The average behavior of all programs and switch point behaviors is shown in Fig. 4. The higher the CAS rate gets the slower the simulation becomes. Additionally the experiment shows that the effort to maintain both cache and the branch prediction units within the ISS adds the highest amount of execution time to the ISS model.

Moreover the added switching point detection and processor core switching functionality has an impact on simulation time. Both simulators perform the switch point detection on each instruction execution. The dominating effort in switch point detection is the search for the current program counter within the list of active switch points. With increasing list size the detection effort increases. To isolate the detection effort a list of switch points with program counters that never occur within the executed program is used. This eliminates the impact of the core switching process that occurs after switch point hit and reduces the overhead to the detection. Due to the long simulation time of CAS the impact is not visible until after more than 1.000.000 switch points are added to the list. In case of ISS simulation an effect is measurable if more than 100 switch points are added to the list. In case switch points that occur during program execution are added to the CAS the time required for switch point checking has no measurable effect on simulation time. The impact of core switching during simulation is measured by increasing the amount of core switches. For a correct comparison the resulting simulation time is compared to simulations with the same CAS rate but only a single switch point. The switching frequency is measured in switches per executed instructions. The experiment shows that as long as the switching rate is less than 0.01 switches per instruction the impact on simulation speed is not measurable. Moreover applications requiring such a high switch rate over the entire simulation run are rare.

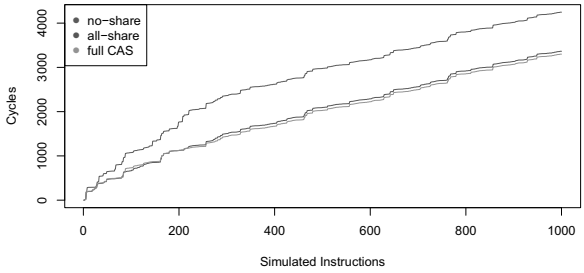


Figure 5: Impact of cache and branch prediction sharing on absolute simulation accuracy

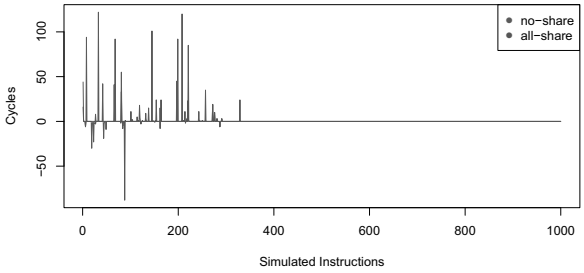


Figure 6: Instruction timing comparison of a program part executed in hybrid simulation CAS to a full CAS

Simulation Accuracy

An important factor of dynamic core switching is the cycle accuracy reached in hybrid simulation compared to a full CAS simulation. One aspect of cycle accuracy is the absolute cycle count of a switched program section. To take the impact of cache and branch prediction unit sharing between the two simulators into account additional experiments are performed. The absolute difference between the pure CAS and the hybrid simulation is measured after each instruction executed. In all test cases the relative difference of absolute cycle count decreases with increasing CAS simulation parts. Without the sharing of cache and branch prediction resources the investigated programs show an average inaccuracy of 3.5 % after 10.000 simulated CAS cycles. This value decreases to 0.25 % if all resources are shared between the ISS and CAS simulators. Moreover the impact of the uncertainty gap after the switch decreases with increasing CAS simulation. An example of the absolute difference in clock cycles per instruction for one switch point within the PID program is shown in figure 5.
Measuring differences in absolute cycle count does not indicate how accurate the hybrid simulation is in terms of architectural behavior. To determine a hybrid simulations accuracy in showing an architecture's internal behavior a more profound evaluation is required. By measuring the amount of cycles between each instruction and comparing them to those of a full time CAS model, a better assessment is possible. After reaching complete identical behavior in terms of equal cycle count between all instructions an identical architecture behavior is assumed. This point in simulation marks the end of the uncertainty gap and depends on the executed program code. In the plot presented in Fig. 6 the difference in cycles between instruction is plotted over the amount of executed instructions for different resource sharing models. This diagram also shows the PID program and the same switch point as the previous one. In this example the simulation without resource sharing overcomes the uncertainty gap after 329 instructions. The simulation with full resource sharing on the other hand reaches this point after 291 instructions. In this case the resource sharing decreases the uncertainty gap by 11.5 %. The average for all investigated programs and switch points reaches a 15.4 % shorter uncertainty gap.

Conclusion

In this paper a hybrid simulation technique switching between instruction set simulation and cycle accurate simulation is presented. The simulation environment supports shared resources for fast context switch between simulators and allowed switching at any point in simulation. The crucial points of dynamically switching between simulators with different levels of density of architecture information are discussed. Especially the impact of processor components with strong influence on simulation accuracy including cache and branch prediction are assessed. By sharing these components between the two simulators and keeping them up to date during the ISS execution, the simulation speed of the ISS is increased only slightly but the accuracy of the program parts executed on the CAS is increased significantly.
The simulation technique is applied to a PLC processor simulation with an ISS model based on ArchC and a CAS model written in SystemC. Simulations show that the hybrid simulation technique can be applied for fast simulation up to a point of interest. From there on the high accuracy CAS simulator takes over.

ACKNOWLEDGEMENT

The Authors gratefully acknowledge the support by profichip GmbH and the Bayrische Staatsministerium für Wirtschaft Infrastruktur und Technologie (StMWIVT), in the context of the R&D program IuK Bayern under Grant No. IUK-1308-0009.

REFERENCES

Binkert N.; Beckmann B.; Black G.; Reinhardt S.K.; Saidi A.; Basu A.; Hestness J.; Hower D.R.; Krishna T.; Sardashti S.; Sen R.; Sewell K.; Shoaib M.; Vaish N.; Hill M.D.; and Wood D.A., 2011. *The Gem5 Simulator. SIGARCH Comput Archit News*, 39, no. 2, 1–7.

Braun G.; Nohl A.; Hoffmann A.; Schliebusch O.; Leupers R.; and Meyr H., 2004. *A universal technique for fast and flexible instruction-set architecture simulation. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 23, no. 12, 1625–1639.

Carlson T.E.; Heirmant W.; and Eeckhout L., 2011. *Sniper: Exploring the level of abstraction for scalable and accu-*

rate parallel multi-core simulation. In *2011 International Conference for High Performance Computing, Networking, Storage and Analysis (SC)*. 1–12.

Eeckhout L., 2010. *Computer architecture performance evaluation methods*. Morgan & Claypool Publishers, San Rafael, Calif. ISBN 978-1608454679.

Hoffmann A.; Schliebusch O.; Nohl A.; Braun G.; Wahlen O.; and Meyr H., 2001. *A methodology for the design of application specific instruction set processors (ASIP) using the machine description language LISA*. In *Computer Aided Design, 2001. ICCAD 2001. IEEE/ACM International Conference on*. 625–630.

Jovic J.; Yakoushkin S.; Murillo L.; Eusse J.; Leupers R.; and Ascheid G., 2012. *Hybrid simulation for extensible processor cores*. In *2012 Design, Automation Test in Europe Conference Exhibition (DATE)*. 288–291.

Kassem R.; Briday M.; Bechennec J.L.; Trinquet Y.; and Savaton G., 2008. *Simulator generation using an automaton based pipeline model for timing analysis*. In *International Multiconference on Computer Science and Information Technology, 2008. IMCSIT 2008*. 657–664.

Kassem R.; Briday M.; Béchennec J.L.; Trinquet Y.; and Savaton G., 2009. *Instruction Set Simulator Generation Using HARMLESS, a New Hardware Architecture Description Language*. In *Proceedings of the 2Nd International Conference on Simulation Tools and Techniques*. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), Simutools '09, 24:1–24:9.

Kohl J.; Fey D.; and Bäsig J., 2015. *Using ArchC to model a Hard-PLC processor*. In *Industrial Instrumentation and Control (ICIC), 2015 International Conference on*. 507–511.

Kraemer S.; Gao L.; Weinstock J.; Leupers R.; Ascheid G.; and Meyr H., 2007. *HySim: A fast simulation framework for embedded software development*. In *Hardware/Software Codesign and System Synthesis (CODES+ISSS), 2007 5th IEEE/ACM/IFIP International Conference on*. 75–80.

Luckham D. and Vera J., 1995. *An event-based architecture definition language*. *IEEE Transactions on Software Engineering*, 21, no. 9, 717–734.

Qiu J.; Gao X.; Jiang Y.; and Xiao X., 2011. *An ultra-fast hybrid simulation framework for ASIP*. In *Electronics, Circuits and Systems (ICECS), 2011 18th IEEE International Conference on*. 711–714.

Rigo S.; Araujo G.; Bartholomeu M.; and Azevedo R., 2004. *ArchC: a systemC-based architecture description language*. In *16th Symposium on Computer Architecture and High Performance Computing, SBAC-PAD*. 66–73.

Wang R.; Song X.; Zhu J.; and Gu M., 2011. *Formal modeling and synthesis of programmable logic controllers*. *Computers in Industry*, 62, no. 1, 23–31.

A NEW METHOD TO TRANSFORM PETRI NETS TO DIGITAL CIRCUITS USING INPUT-DRIVEN REACHABILITY GRAPHS

Christoph Brandau and Dietmar Tutsch
School of Electrical, Information and Media Engineering
University of Wuppertal, Germany
e-mail:{brandau, tutsch}@uni-wuppertal.de

KEYWORDS
Petri net, Computer-aided analysis, Model analysis, Model design, Digital Circuit Petri Nets, Discrete Simulation, Input-driven reachability graph, Transformation, Digital circuit

ABSTRACT

This paper introduces the transformation process from digital circuit Petri nets (DCPN) to digital circuits. It describes the complete process from modeling a Petri net to the generation of VHDL code. First, a short introduction to DCPN is given. Then the transformation process will be discussed and several strategies are shown.
This paper introduces a new way to generate sequential circuits from Petri nets. For this purpose, the input-driven reachability graphs will be introduced. Combinational logic will be transformed by the use of truth tables and optimization like Quine-McCluskey or Karnaugh maps. All strategies are implemented in the tool Logical Petri Net (LPN). The LPN is described in (Brandau et al., 2016a).

INTRODUCTION

Nowadays, associated with the continuous increase of the complexity of digital circuits, new approaches, and tools supporting their design must be introduced. In this paper, we offer the description of the transformation process from Petri nets to digital circuits. These nets will be transformed into a hardware description language like VHDL. To reach this goal, we extend the standard Petri nets with additional elements.
The motivation is to design a digital system using an accessible graphical description. Petri nets provide many methods to analyze the modeled nets which can be used creating safety relevant circuits. Among others Markov chains are used for the analysis.
There have already been several approaches to do so. In (Gomes et al., 2014), standard Petri nets are extended to IOPT-nets with input and output elements. In (Bukowiec et al., 2014), (Yi et al., 2014), (Wiśniewski et al., 2017) the behavior of Petri nets is described by every single element of the Petri net. They will use transition oriented or transition/place oriented conditionals to describe the Petri net in a hardware description language. In addition, removing and adding token in the net will be performed depending on the transitions connected to a place. Every part of the net will be translated individually. Additionally, colored Petri nets are used in these papers. In contrary, our approach is based on the extension of standard Petri nets. Each of these approaches has advantages and disadvantages, so we decided to define an own Petri net type. These so-called Digital Circuit Petri Nets (DCPN) are described in (Brandau et al., 2016b).

Regarding a description of digital circuits the standard Petri net elements are insufficient and the standard Petri net definition needs to be extended including input and output places, subplaces, netconnectors, subtransitions and inhibitor arcs. Fig. 1 shows a net with these extensions.

MODELING

In this section, the Petri net model is presented. At first, the elements of the DCPN are briefly described. Then, the modeling possibilities are discussed. The standard elements like places, transitions, and arcs (Murata, 1989) will be extended with the following new elements.

Input and Output Places

To represent digital circuits with Petri nets, input and outputs must be defined, to get access to the circuit from outside. For this purpose, two new types of places are added. These are, firstly, the input places, which are represented as black circles with a green filling. On the other hand, these are the output places also presented as black circles, but with a yellow filling.

Later on, the input places will be filled with any possible marking representing the digital input pattern and the result at the output places will be shown for the analysis of the Petri nets. The input and output places are required to generate a port list giving the connections to the surrounding system components.

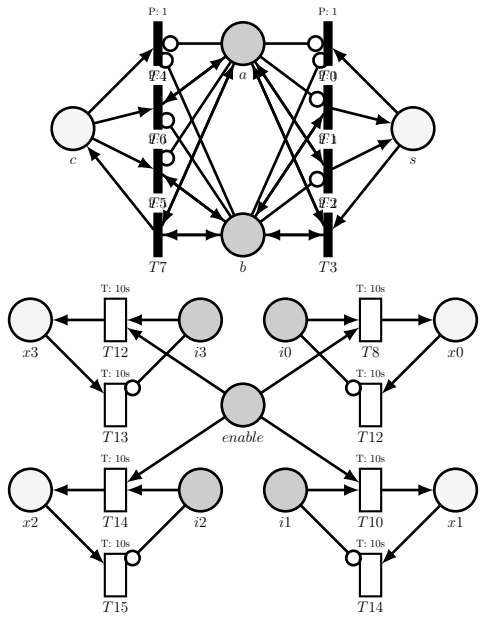


Figure 1: Digital Circuit Petri Net with input places (green/dark gray) and output places (yellow/light gray) and inhibitor arcs. Furthermore, the elements from standard Petri nets are also included. The upper left subnet represents a half adder as a modelled DCPN. The rest of the net is a four bit memory with an enable input to store the given values.

Subnets

For readability of Petri nets, a division into subnets is desirable. Regarding a connection of multiple items within a net, it is necessary, that subnets provide multiple in- and output arcs. This is also indispensable for the presentation of a wiring from outside. In the hardware description, it is common to model system components individually and combine them to an entire system. For this purpose, it is useful to transfer Petri nets to a similar structure.

For this paper, the previously existing subnet types are not sufficient because multiple inputs and outputs to and from the subnet are needed, so two separate types have been introduced. These are the subplaces and the subtransitions. In order to prepare the components of a subnet available to the outside, a new element is introduced. It is called netconnector. These components fulfill a different task, depending on the type of the subnet.

Subplaces are used to represent a subnet whose arcs lead from or to transitions and subtransitions. Subplaces replace places in case that this part of the net should be more detailed. Therefore, the basic behavior is equivalent to the place type of the Petri net. This means that arcs must not connect any subplaces or places with each other.

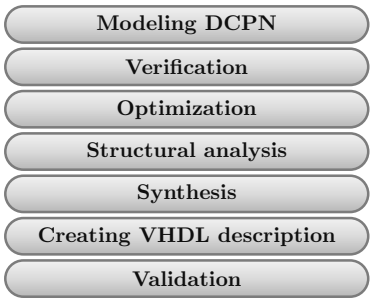


Figure 2: The transformation process steps to transform a DCPN into a digital circuit.

The symbol in the Petri net corresponds to an unfilled circle surrounded by a second unfilled circle. For each input and output arc of the subplace, a netconnector within the subnet is created and will be represented in the form of a wider unfilled arrow. This element is given the name of the element, from which the arc leads into or out of the subplace.

As a further kind of subnet, Petri nets have been expanded with subtransitions. Their behavior is similar to that of a transition. Arcs between subtransitions and transitions are therefore not allowed. Only arcs between subtransitions and subplaces or places are allowed. The presentation is an unfilled rectangle surrounded by a second unfilled rectangle. In subtransitions the places located outside and connecting to the subnet are represented as closed unfilled semi circles. The subnets are based on the macroplaces which are presented in (Karatkevich, 2008).

TRANSFORMATION

Seven steps will be executed for the transformation of a designed DCPN into a digital circuit. They are described in this section and can be seen in Fig. 2. For every step, several strategies exist which will be explained in this section. The transformation of combinational logic uses the reachability graph or the reduced reachability graph that is introduced in (Karatkevich, 2007). A detailed transformation of a combinational DCPN is shown in Brandau et al. (2016a).

The transformation process can take place in two different ways. First, the transformation of the entire DCPN can be performed by transforming each subnet of the DCPN by starting with the innermost subnets. This is because the overlying nets in the hierarchy use the circuits of the internal nets. On the other hand, a breakup of the hierarchy can be performed to produce an overall circuit from the entire DCPN. To resolve the subnets to the main net, they will be reintegrated into the main net.

The modeling of the proposed system takes place as a DCPN. Here, the properties from the previous sections are used. This is followed by the verification of the

Table 1: Strategies for the verification of a DCPN.

S1	Existing input places and output places
S2	Existing transition
S3	Arcs at input places and output places
S4	Transitions without arcs
S5	Places without arcs
S6	Subnets without arcs
S7	Connection
S8	Strong connection
S9	Static conflicts

model to detect erroneous or contradictory characteristics of the DCPN and to identify these faults in the net. The further used strategies to identify potential improvements in the net are shown in Table 1.
The first step is to check if input places and output places exist. Input places can be missing, but output places are relevant for the circuit creation. Next step is to check if there are transitions in the net. The circuit will have no functionallity if there is no transition in the net. Then the net is checked for arcs from or to the input and output places. If there are no arcs to these elements, these elements have no impact on the circuit. The same applies to transitions, subplaces, and subtransitions without arcs.
The next strategies are the connection and strong connection. In this step, the DCPN will be divided into several nets if possible. For this purpose, the net will be searched for connected subnets by starting at the output places. All connected nets are considered as new Petri nets and will be transformed separately in the next steps. At the VHDL creating process, they will be put together to one circuit. As the last step of the verification, the static conflicts of transitions will be searched and saved for later used strategies.
The next transformation step is the optimization of the Petri net. Table 2 shows the used strategies. Elements without an impact on the behavior of the circuit will be removed in this step. Furthermore, this step occurs in the detection and grouping of redundant elements. These elements can be parallel transitions or transitions in row. Given net symmetries can be used to reduce the size of the net.

Table 2: Strategies to optimize the DCPN

S10	Remove elements without arc
S11	Combine parallel transitions
S12	Combine sequential transitions
S13	Remove places without input arc
S14	Remove not connected subnets
S15	Reduce net symmetries

The next step is a general structural analysis, in which the breakdown by circuit type must be carried out. These types are combinational logic and sequential logic. All strategies are shown in Table 3. Faulty modeling can also appear in this step because a comprehensive analysis of the system takes place. This step is the important and most sophisticated step in this chain because the transformation starts and will be finished within the next step. Also, a lot of reachability graphs and other calculations are done in this step. First, the net type must be determined. The net type is divided into three types:

timeless net contains only timeless transitions

timed net contains timeless and timed transitions

strong timed net contains only timed transitions

Table 3: Strategies to classify the found subnets to sequential or combinational circuits.

S16	Determine net type
S17	Create reachability graphs
S18	Find cycles
S19	Termination
S20	Timeless termination
S21	Defined end state
S22	Defined timed state
S23	State free
S24	Determine circuit type

Strategy 17 calculates reachability graphs for every input marking. The input markings are all possible capabilities of a token in the input places. An input place can hold either no or a single token. For n inputs, this means that there are 2^n possibilities. Then each reachability graph will be checked for cycles and a general termination of the graph.
Every net has to be checked whether it terminates timeless. Otherwise, the transformation can not be carried out. Also, timeless nets must have a defined final state in all reachability graphs. All other net types must have a defined timed state. This means that only one state can be reached at a time. All active timeless transitions must be fired until only timed transitions are active.
The last thing which would be checked before a statement of the circuit type can be given is the state free strategy. For this purpose, all input combinations of the net are compared with all others. If the same results are always present at the outputs, regardless of the sequence, then the net is state free. If the net terminates timeless has a defined final state and is state free, then this net is combinational logic. All other nets are sequential logic.
The structural analysis will be followed by the synthesis of the net to the target architecture. The strategies are shown in Table 4. This transformation step will

be divided into two parts. The first is the synthesis of combinational logic with the strategies 25 to 28. First, the truth table will be created which can be used for the VHDL-description. Otherwise, a conjunctive normal form can be calculated from the truth table. Also, the algorithms from Quine-McCluskey, Espresso or Karnaugh maps can be used to determine the equations. Further information about this transformation strategies is shown in (Brandau et al., 2016b).
The other part is the synthesis of sequential logic with the strategies 29 to 34. First, the sequential type of the circuit must be determined. This type can be synchronous or asynchronous. After this step, the circuit clock can be calculated or given by the designer of the DCPN. For calculation, the greatest common divisor of all switching times will be used. If the designer sets the clock to a fixed value, the switching time of every transition has to be verified. After this, the input-driven reachability graph (IRG) will be calculated. The creation of this graph will be discussed in the next section. Also, the reducing of timeless states and the expansion of timed transitions will be described in the next section.

Table 4: Strategies for the synthesis of the given DCPN. The strategies 25-28 are for combinational logic and the strategies 29-34 are for sequential logic.

S25	Create truth table
S26	Calculate conjunctive normal form
S27	Set optimization targets
S28	Optimize equations
S29	Determine sequential circuit type
S30	Calculate clock
S31	Calculate wrong timed transitions
S32	Create input-driven reachability graph
S33	Reduce timeless states in IRG
S34	Expand IRG for timed transitions

The transformation of DCPN in the hardware description language VHDL affiliates in which the interface of the complete net is created. Furthermore, the description of the behavior or structure is created depending on the detected circuit type. Table 5 shows all used strategies.
First, the divided nets from strategy 7 must be merged. Also the input places and output places which were removed in strategy 3 must be added to the net. Next, the names of all places have to be checked if the used signs are allowed. Forbidden names are changed. Then the entity for the main DCPN and all subplaces and subtransitions will be created.
The architecture description is the next step. For this, the created truth tables, equations, and input-driven reachability graphs will be used. For the description of sequential circuits a Moore machine with three processes is used. An example of a transformation from an

Table 5: Strategies to create the entity and architecture in VHDL. The strategies depends on the determined circuit type.

S35	Merge subnets
S36	Add input and output places
S37	Check element names
S38	Create interface for main net
S39	Create interface subplace
S40	Create interface subtransition
S41	Write truth table
S42	Write equation
S43	Write sequential state transitions
S44	Write synchronous state transitions
S45	Write asynchronous state transitions
S46	Write assorted follow state
S47	Write follow state
S48	Write outputs

DCPN with two detected subnets is described later in this paper. The example Petri net is shown in Fig. 1.
After the transformation is completed, a validation of the circuit can be started. It will be achieved by a comparison of the simulation results between the DCPN model and the resulting digital circuit. The simulation of the DCPN is carried out by an event-based simulation because every firing transition is one event. A test bench for the created circuits will be generated for the simulation. Table 6 shows the used strategies.

Table 6: Strategies to validate the resulting digital circuit.

S49	Create testbench
S50	Evaluate simulation results

Fig. 3 shows all strategies for the transformation process. The red labeled strategies are possible termination points of the process. The yellow marked strategies are points where warnings can be obtained. These warnings are for sequential logic if there given clock did not match the switching time of the transitions.

INPUT-DRIVEN REACHABILITY GRAPH

As described in the last section, the input-driven reachability graph is used for generating sequential circuits. The graph consists of a set of nodes N. They represent the internal state of the Petri net. A node n_i contains a marking M_i of the given DCPN. The marking of the input places is not saved in the node. Furthermore, a node contains a set of directed arcs A_i that pointing to its successors.
A directed arc a_j has a set of conditions C_j. The arc also has a reference to the target node n_j to which the

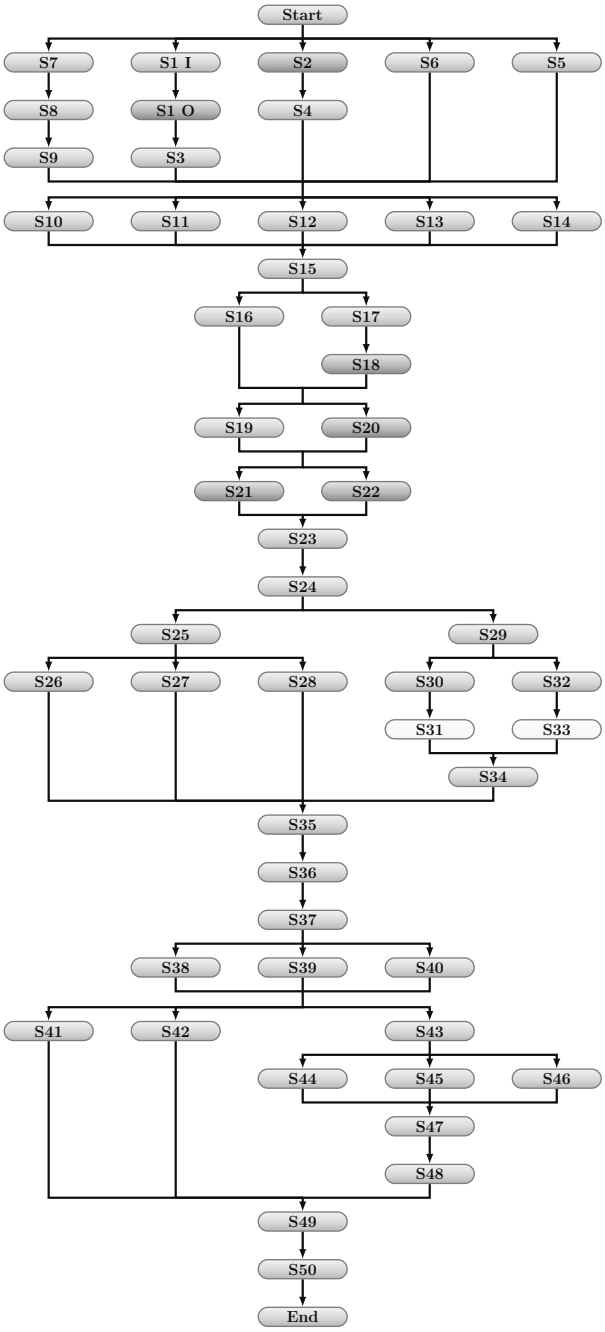


Figure 3: This is the transformation process with all used strategies and branches. The red marked strategies are possible termination strategies. The yellow marked strategies can generate warnings.

arc points. The arc always starts in its associated node n_i. A condition c_k contains the information of the input marking M_{in_k} and the switching time $time(t_j)$ of the transition t_j. Timeless transitions are assumed with a switching time of zero. The set C_j can have several conditions with the same input marking, but the transition must then be a different one. This can occur if several transitions can fire at the same time. All conditions that transfer the state of the net from a node n_i to n_j are summarized in one arc. From this, the definition for a node n_i, an arc a_j, and a condition c_k are obtained:

$$\begin{aligned} n_i &= \{M_i, A_i\} \\ a_j &= \{C_j, n_j\} \\ c_k &= \{M_{in_k}, time(t_k)\} \end{aligned} \tag{1}$$

The pseudo-code presented in Algorithm 1 is used to create the input-driven reachability graph. The initial marking M_0 and the Petri net PN are the required inputs for the algorithm. The starting node n_0 is generated from the initial marking M_0. This node n_0 is added to the set of nodes N and to the set of nodes N_{temp} to be checked. After that, the algorithm runs while a node has to be checked and is present in N_{temp}. For this purpose, a node n is selected from the set N_{temp} and all input combinations are simulated for this node to find new nodes that are not element of N. The tested node n will be removed from the set N_{temp}.

Algorithm 1 Input driven reachability graph

```
function createIRG(M_0, PN)
    Create n_0 from M_0
    N ← n_0
    N_temp ← n_0
    while N_temp ≠ ∅ do
        n ∈ N_temp
        N_temp ← N_temp \ n
        Fill PN with marking from n
        for all m ∈ M_input do
            for all t = {t | t ∈ T_all ∧ m|t⟩} do
                m' = m -t-> m'
                Create n_new from m'
                if n_new ∈ N then
                    n_exist = {n_i | n_i ∈ N ∧ n_i = n_new}
                    addArc( m, t, n, n_exist )
                else
                    addArc( m, t, n, n_new )
                    N ← N ∪ n_new
                    N_temp ← N_temp ∪ n_new
```

For n, all possible combinations of markings M_{input} will be placed in the input places. For each of these combinations, the transitions t will be checked, if they are active. Each active transition fires, from which the marking m' is produced. From this, the node n_{new} is generated, by

filtering out the inputs from m'. If the node already exists in N, then the equivalent node n_{exist} is taken from N and the method addArc will be called with the marking m, the transition t, the actually checked node n and the existing node n_{exist}. If the node does not yet exist, then it will be added to both sets N and N_{temp}. Then also the method addArc will be called with the same elements. The only change is to use n_{new} instead of n_{exist}.
Algorithm 2 describes the procedure for adding new arcs to the input-driven reachability graph. As a first step, the condition c is generated from the marking m and the transition t. From t only the switching time is used. From this the arc a is generated, which additionally receives the target node n_{new}. Now all arcs in A will be checked in node n, whether they have the same target node n_{new}. If such an arc a_i already exists, then the condition c is added to this arc. Otherwise, the set A of the element n is extended by the arc a.

Algorithm 2 Add arc to IRG

```
function ADDARC(m, t, n, n_new)
    Create c from m and t
    Create a from c and n_new
    for all a_i ∈ A ∈ n do
        if n_i ∈ a_i = n_new then
            Add c to a_i return
    A ← A ∪ a, A ∈ n
```

The generated graph can be used, if all transitions have the same switching time and this is also the clock or all transitions are immediate ones. Otherwise, the graphs have to be further adapted with the following strategies. First, all timeless conditions have to be check, if they can be transformed. This means that all arcs of the IRG are checked for timeless transitions as conditions. If a further arc with the same condition starts at the target node, then these conditions are combined into a signle one.
After this step, all timed conditions will be checked, if they have the same switching time t_{time} as the clock clk of the circuit. The node will be divided into $\frac{t_{time}}{clk}$ nodes to represent the valid switching time in the created digital circuit if the switching time is larger than clk. After this step, the synthesis for sequential circuits is done and the hardware description can be created.

EXAMPLE TRANSFORMATION

The transformation for the given DCPN in Fig. 1 will divide the net into two subnets. The upper left subnet is a combinational circuit and represents a half adder. The calculated equations for this net are:

$$s = (\neg a \wedge b) \vee (a \wedge \neg b)$$
$$c = (a \wedge b)$$

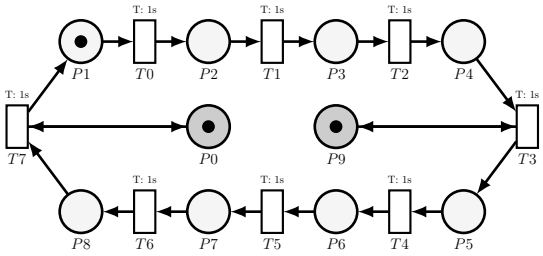


Figure 4: The modeled DCPN represents a ring counter, which switches every clock event to the next step. The only exceptions are the transitions T_3 and T_7, where the input place P_0 or P_1 must have a token.

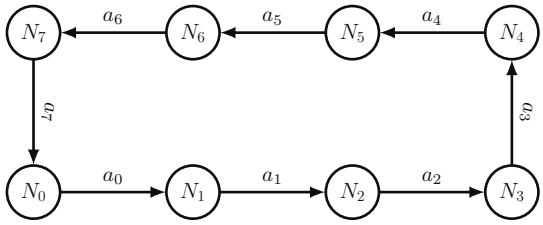


Figure 5: This figure illustrates the input-driven reachability graph for the given net in Fig. 4. The nodes are discribed in Tab. 7 and the conditions are in Tab. 8.

These equations are implemented in VHDL and the transformation of this partial net is completed. Further informations to this transformation can be obtained by reading (Brandau et al., 2016b).
As next step, the second subnet will be transformed. First, the structural analysis detected this net as a sequential net. The clock will be calculated to $10ns$ because every transition has the same switching time. Next, the input-driven reachability graph will be created. This graph is without any optimization very bad readable. The optimizations are one of our next steps in research. So we show another example for an input-driven reachability graph in this paper.

Table 7: All reachable nodes from the IRG shown in Fig. 5.

name	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8
N_0	1	0	0	0	0	0	0	0
N_1	0	1	0	0	0	0	0	0
N_2	0	0	1	0	0	0	0	0
N_3	0	0	0	1	0	0	0	0
N_4	0	0	0	0	1	0	0	0
N_5	0	0	0	0	0	1	0	0
N_6	0	0	0	0	0	0	1	0
N_7	0	0	0	0	0	0	0	1

The DCPN is shown in Fig. 4 and the resulting IRG is shown in Fig. 5. This graph will be transformed

into VHDL as a three process state machine. The three processes are used for the calculation of the following state, the calculation from the outputs and the state memory. The arc conditions are given in Tab. 8 and the node states are in Tab. 7.

Table 8: All arcs with thier conditions from the input-driven reachability graph in Fig. 5. Every condition is discribed as a tuple of token in the input places. The switching time is the same in every transition and is dropped in this table.

arc	**conditions** (P_0, P_9)			
a_0	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
a_1	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
a_2	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
a_3	$(0,1)$	$(1,1)$		
a_4	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
a_5	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
a_6	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
a_7	$(1,0)$	$(1,1)$		

CONCLUSION AND FUTURE WORKS

This paper presented the transformation process to generate digital circuits from DCPN. This process is divided into the steps modeling, verification, optimization, structural analysis, synthesis, creating VHDL description and validation. Several strategies of this steps have been shown. This process can be realized with our tool Logical Petri Net.

One focus of this paper is based on the input-driven reachability graphs. They were presented for the first time in this paper. These IRG serve to extract sequential logic from Petri nets. They work like Medwedyev machines. These graphs work for the transformation into sequential circuits, but they still do not have much optimization. This optimization is one of our next goals. Our current research is aimed at extending the tool LPN to establish more strategies in the tool. The shown strategies are implemented, but there is not much optimization in the structural analysis. So this can consume much time for larger nets.

In addition, we want to develop a parallel simulation. We distribute the creation of the reachability graphs on multiple processors. The number of the created graphs increases by 2^n to the number of inputs n for the analysis. For large DCPN even the creation of an input-driven reachability graph can also be divided into parts to simulate it on several processors.

REFERENCES

Brandau C.; Potthoff N.; and Tutsch D., 2016a. *Logical PetriNet - A Modeling Tool to Describe and Transform Petri Nets into Digital Circuits*. In *ESM'2016 - 30th European Simulation and Modeling Conference*. Eurosis-ETI.

Brandau C.; Potthoff N.; Tutsch D.; and Lepich T., 2016b. *Digital Circuit Petri Nets: A new Petri Net Type to Describe and Transform Digital Circuits for Product Safety Engineering*. In *The 6th IEEE International Conference on Consumer Electronics - Berlin*. IEEE, IEEE, Berlin, Deutschland.

Bukowiec A.; Tkacz J.; Adamski M.; and Wisniewski R., 2014. *Dual synthesis of Petri net based dependable logic controllers for safety critical systems*. In *Human System Interactions (HSI), 2014 7th International Conference on*. 243–248.

Gomes L.; Moutinho F.; Pereira F.; Ribeiro J.; Costa A.; and Barros J.P., 2014. *Extending input-output place-transition Petri nets for distributed controller systems development*. In *Mechatronics and Control (ICMC), 2014 International Conference*. 1099–1104.

Karatkevich A., 2007. *Dynamic Analysis of Petri Net-Based Discrete Systems*, Springer Berlin Heidelberg, chap. Reduced Reachability Graphs. ISBN 978-3-540-71560-3, 27–62.

Karatkevich A., 2008. *On macroplaces in Petri nets*. In *Design Test Symposium (EWDTS), 2008 East-West*. 418–422.

Murata T., 1989. *Petri nets: Properties, analysis and applications. Proceedings of the IEEE*, 77, no. 4, 541–580. ISSN 0018-9219.

Wiśniewski R.; Karatkevich A.; Adamski M.; Costa A.; and Gomes L., 2017. *Prototyping of Concurrent Control Systems With Application of Petri Nets and Comparability Graphs. IEEE Transactions on Control Systems Technology*, , no. 99, 1–12.

Yi Z.; Mu X.; Zhao P.; and Yi Y., 2014. *Software-hardware Interaction Analysis Based on Petri Net*. In *Control and Decision Conference (2014 CCDC)*. 2815–2820.

FLUID SYSTEMS SIMULATION

MONTE-CARLO SIMULATION OF DAILY PRECIPITATION AND RIVER FLOW CONDITIONAL SPATIO-TEMPORAL FIELDS

Nina A. Kargapolova
Institute of Computational Mathematics and Mathematical Geophysics
Pr. Lavrent'eva 6
630090 Novosibirsk
Russia

Novosibirsk State University
Pirogov St. 1
630090 Novosibirsk
Russia

E-mail: nkargapolova@gmail.com

KEYWORDS

Stochastic Simulation, Conditional Field, Multiplicative Model, River Flow, Precipitation.

ABSTRACT

In this paper, a stochastic model of river flow and daily precipitation conditional fields is considered. This model is based on real data from weather stations situated in Novosibirsk region and gauging stations on Berd' river. Spatial heterogeneity of precipitation distributions is taken into account. Several examples that illustrate dependence of river flow on different precipitation regimes are given.

INTRODUCTION

There are several factors that influence on a river flow. Among them groundwater inflow, precipitation and snowmelt in the river basin and artificial drainage are usually considered as key factors defining a river flow (Komlev 2002). It is obvious that it is not always possible to carry out a full-scale experiment and to study on its basis influence of each specific factor on the river flow. By contrast, numerical simulation may help to solve this problem. In this paper precipitation influence on the river flow is studied. One of the approaches to numerical study of such influence is a dynamical-probabilistic approach. In this approach precipitation in a river basin are considered as a random field and entry of water into the river from its water system is simulated on a basis of a deterministic dynamical hydrological model. In this paper, another approach is suggested: both precipitation and river flow are considered as a joint random process. Corresponding real data analysis, model description, simulation algorithms and results of numerical experiments are given in this paper.

REAL DATA ANALYSIS & MAIN ASSUMPTIONS

The model of daily precipitation and river flow conditional random fields is based on real data from 56 weather stations situated in Novosibirsk region (Russia). Data about daily precipitation (mm) was collected since 1969 to 1983. Two weather stations (Maslyanino and Stariy Iskitim) are also gauging stations on Berd' river where in the same years average daily river flow (m^3/s) was measured. Since the river is usually covered with ice from the first days of November till the end of April, river flow was measured only from May to October. Figure 1 represents a map of Novosibirsk region with marked weather and gauging stations and Berd' river catchment basin. Simulation of a precipitation field was done on gauging stations and in nodes of a rectangular grid, covering the catchment basin, with a grid step $\Delta x = \Delta y = 3\ km$ in South-North and West-East directions. Total number of points where simulation of precipitation was done is equal to $N = 1158$ (1156 grid nods and 2 gauging stations). It should be noted that only 2 weather stations lie within the simulation area. In this paper, a model in which river flow is simulated in a single point (on one of the gauging stations) is considered.

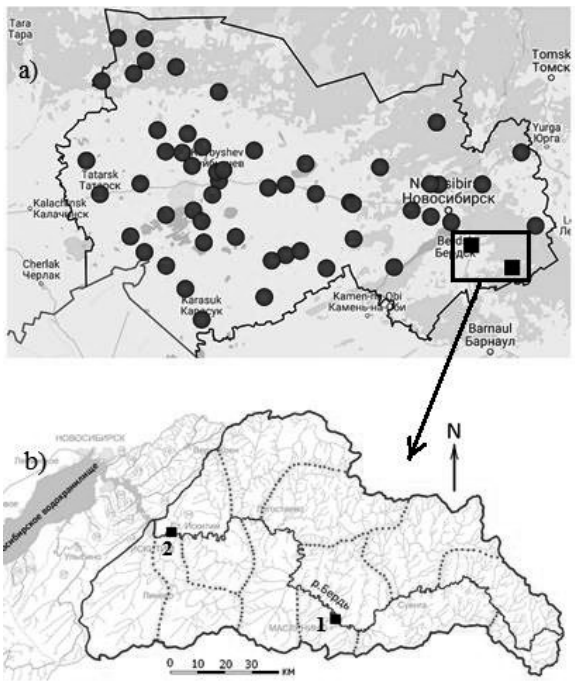


Figure 1: a) – a Map of Novosibirsk Region, b) – Berd' River Catchment Basin. Circles – Weather Stations, Squares – Weather and Gauging Stations. 1 – Maslyanino, 2 – Stariy Iskitim

Precipitation

For precipitation spatial field simulation, a modification of an approach presented in (Marchenko and Ogorodnikov 1991; Ogorodnikov and Sereseva 2015) was used. It is supposed that precipitation χ_i in a point number $i, i = 1,\ldots, N$ may be presented in a form

$$\chi_i = \omega_i \xi_i,$$

where ω_i is a value of a precipitation indicator random field ω (if $\omega_i = 1$ a day is considered as a wet day (there are non-zero precipitation) in a point i, if $\omega_i = 0$ it's a dry day) and ξ_i is amount of precipitation in a point i under a condition of precipitation presence. Random fields ω and ξ are independent on each other. These fields are supposed to be homogeneous in sense of their correlation structure. Since real data allow to estimate correlation coefficients of a precipitation field only between weather stations and give no information about correlation coefficients between nods, it was necessary to approximate real data based correlation coefficients with correlation functions of continuous arguments and calculate correlation matrixes C^{ω}, C^{ξ} of fields ω and ξ using these functions. Approximating functions have a form

$$c\left(x_i, y_i, x_j, y_j\right) = c\left(x_j - x_i, y_j - y_i\right) =$$
$$= c(x, y) = \exp\left(-\left[ax^2 + bxy + by^2\right]^d\right),$$

where $\left(x_i, y_i\right), \left(x_j, y_j\right)$ are Cartesian coordinates of points number i and j. Parameters of the approximating functions are given in Table 1. Elements of the correlation matrixes C^{ω}, C^{ξ} are denoted as $c_{ij}^{\omega}, c_{ij}^{\xi}$ $(i, j = 1,\ldots, N)$ respectively.

Table 1. Parameters of the Approximating Correlation Functions. July

Field	*a*	*b*	*c*	*d*
ω	0.0021	-0.0026	0.0010	0.2130
ξ	0.0032	-0.0090	0.0095	0.3010

In (Ogorodnikov and Sereseva 2015) it was supposed that fields ω and ξ are homogeneous in sense of their one-dimensional distributions, but in (Kargapolova and Ogorodnikov 2017) it was shown that these fields are heterogeneous. Assume that types of marginal distributions do not depend on i, but distribution parameters may vary from point to point. Marginal distribution of the indicator field ω in each point i is defined with a probability $p_i = P(\omega_i = 1)$. Gamma-distribution with density

$$f_i(x) = x^{k_i - 1} \exp\left(-\frac{x}{t_i}\right) \Big/ \Gamma(k_i) t_i^{k_i}$$

is assumed to be a one-dimensional distribution of the field ξ in point i. Parameters $p_i, k_i, t_i, (i = 1,\ldots, N)$ were determined on weather stations for each month on a basis of real data using maximum likelihood method and following IDW-interpolation from weather stations to grid nods. It should be noted that usage of gamma-distribution as a one-dimensional distribution of precipitation amount and IDW-interpolation for distribution parameters calculation is a common approach in Statistical Meteorology (see, for example (Hartkamp et al. 1999; Richardson and Wright 1984; Sluiter 2009)).

For simulation of a spatio-temporal precipitation field it is also necessary to define a time-correlation structure of fields ω and ξ. Considered in this paper model is constructed under an assumption that both fields ω and ξ are stationary and their temporal correlation functions don't depend on spatial coordinates:

$$r^{\omega}\left(x_i, y_i, x_j, y_j, t_1, t_2\right) = r^{\omega}\left(t_1, t_2\right) = r^{\omega}\left(t_1 - t_2\right) = r^{\omega}(h),$$
$$r^{\xi}\left(x_i, y_i, x_j, y_j, t_1, t_2\right) = r^{\xi}\left(t_1, t_2\right) = r^{\xi}\left(t_1 - t_2\right) = r^{\xi}(h).$$

Values of $r^{\omega}(h), r^{\xi}(h)$ $(h = 0, 1, 2, \ldots)$ may be defined as corresponding sample correlation coefficients. It appeared that sample correlation coefficients of fields ω and ξ are statistically non-significant for $h \geq 3$ and for $h \geq 2$ respectively. This means that temporal correlation matrices R^{ω}, R^{ξ} of fields ω and ξ may be considered as Toeplitz band matrices. Spatio-temporal correlation matrixes of fields ω and ξ are defined as direct products $R^{\omega} \otimes C^{\omega}, R^{\xi} \otimes C^{\xi}$ (Ermakov and Mikhailov 1982).

River Flow

Real data analysis and later simulation show that there are 2 functions that give equal in quality approximation of sample marginal distribution function of a river flow on a gauging station. The first one is a piecewise-linear approximation of a sample distribution function combined with an exponential tail. The second one is a mixture of 2 (for Maslyanino) and 3 (for Stariy Iskitim) Gamma-distributions. In this paper, mixture of Gamma-distributions was used.

Since in this paper river flow is simulated only on one gauging station (either Maslyanino or Stariy Iskitim) there is no need to define spatial correlation structure of river flow. Values of a temporal correlation function (under the assumption that river flow is a stationary process) can be assumed to be equivalent to its sample values, but since sample size is relatively small, statistical errors of temporal correlation coefficients estimations are significant. It was decided to use a function

$$corr(h) = \frac{1}{A + B} e^{-\alpha h} \left(B \cos \beta h + A\right)$$

as a temporal correlation function. This function approximates well sample correlation values.

There is a rather tricky question: "How to define correlation coefficients between precipitation in the grid nods and river flow on a gauging station?" An approach that was proposed

in (Kargapolova and Ogorodnikov 2017; Shlychkov et al. 2015) was used. Let $t_i^{\omega}(\Delta t)$ be a correlation coefficient between precipitation indicator in a nod number i and river flow on a chosen gauging station Δt days later. We suppose that $t_i^{\omega}(\Delta t)$ may be presented in a form

$$t_i^{\omega}(\Delta t) = r_{\omega s}(\Delta t) \cdot c_{ik}^{\omega},$$

where $r_{\omega s}(\Delta t)$ is a correlation coefficient between precipitation indicator and river flow on a gauging station Δt days later, c_{ik}^{ω} is a defined above spatial correlation coefficient of precipitation indicator in a nod number i and on a gauging station ($k = N-1$ or $k = N$ depending on a gauging station under consideration). Correlation coefficients $t_i^{\xi}(\Delta t)$ between precipitation amount and river flow are defined in an analogous manner.

SIMULATION ALGORITHM

Even if it is necessary to simulate only a spatial joint random field of precipitation indicators, precipitation amount and river flow, its joint correlation matrix

$$\begin{pmatrix} & & & & & t_1^{\omega}(0) \\ & C^{\omega} & & \Theta & & \vdots \\ & & & & & t_N^{\omega}(0) \\ & & & & & t_1^{\xi}(0) \\ & \Theta & & C^{\xi} & & \vdots \\ & & & & & t_N^{\xi}(0) \\ t_1^{\omega}(0) & \dots & t_N^{\omega}(0) & t_1^{\xi}(0) & \dots & t_N^{\xi}(0) \quad 1 \end{pmatrix}$$

is a $(2N+1)\times(2N+1)$ square matrix. Here Θ is a $N \times N$ zero matrix. Correlation matrix's dimension turns simulation of a joint field using Choletsky or spectral decomposition of the correlation matrix to be time and RAM consuming operation. Moreover, when so high-order matrixes are used for simulation, huge computational errors accumulate that make simulation algorithms instable. Simulation of a spatio-temporal field in this situation seems to be a practically unsolvable problem. To reduce matrix dimension and to take into account dependence of a river flow on precipitation, method of conditional distributions may be used (Ogorodnikov and Prigarin 1996; Sobol 1973). Let us formulate a simulation algorithm and then discuss each step of the algorithms in detail.

Algorithm.

Step 1. Using threshold transformation of a Gaussian process, a field ω of precipitation indicators with correlation matrix $R^{\omega} \otimes C^{\omega}$ is simulated.

Step 2. Independently on ω a field ξ of precipitation amount with correlation matrix $R^{\xi} \otimes C^{\xi}$ is simulated using inverse distribution function method.

Step 3. Fields ω and ξ are pointwise multiplied to form a field χ of precipitation.

Step 4. When field χ is simulated, a single point conditional field ψ of river flow is simulated.

In the first step of the algorithm simulation of the field ω in each point i (both for spatio and spatio-temporal models) is done using a threshold transformation of a stationary Gaussian process η^{ω}:

$$\omega_i = \begin{cases} 1, \eta_i^{\omega} \le c_i, \\ 0, \eta_i^{\omega} > c_i, \end{cases}$$

where threshold values c_i are defined from equations

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{c_i} \exp\left(-t^2/2\right) dt = p_i,$$

and correlation matrix $C^{\eta^{\omega}}$ of the process η^{ω} is defined by C^{ω}, p_i (Kargapolova 2017).

In the second step, the field ξ is simulated on a basis of a well-known inverse distribution function method (see, for example, Ogorodnikov and Prigarin 1996; Prigarin 2005). In the context of this method an intermediary Gaussian process η^{ξ} with correlation matrix $C^{\eta^{\xi}}$ (that is defined by described above C^{ξ} and $f_i(x)$) is simulated and then transformed into the field ξ using equality

$$\xi_i = F_i^{-1}\left(\Phi\left(\eta_i^{\xi}\right)\right),$$

where $F_i(x)$ is a corresponding to $f_i(x)$ distribution function and $\Phi(x)$ is a Gaussian distribution function. It may happen that numerically calculated matrix $C^{\eta^{\xi}}$ of a Gaussian process η^{ξ} is non-positively defined. In such case a regularization procedure must be applied to $C^{\eta^{\xi}}$ before simulation of the process η^{ξ} (Ogorodnikov and Prigarin 1996).

In the last step, a conditional single point random field ψ dependent on ω and ξ is simulated. Simulation formulas for simulation of conditional fields are given, for example, in (Ogorodnikov 2013).

NUMERICAL EXPERIMENTS

Before any of numerical experiments are done, a model must be verified. For model verification characteristics that are not model input parameters are usually used. Figure 2 show an example of such characteristic. On basis of real and simulated data a probability of the event "precipitation on weather stations Maslyanino and Stariy Iskitim is greater than l mm simultaneously" was estimated. Hereafter 10^6 simulated trajectories were used for estimations. Simulated data sample estimations of this probability $p(l) = P(\chi_{N-1} > l, \chi_N > l)$ differ from corresponding real

data sample estimations, but since statistical error of real data estimations is relatively huge, reproducing of $p(l)$ by the model may be considered as satisfying. Verification of the model also showed that model reproduces well such characteristics as probabilities $P(\psi < h \mid \chi_{N-1} < m, \chi_N < l)$, $P(\psi > h \mid \chi_{N-1} > l)$, etc. Naturally, for model verification only data from weather and gauging stations is used.

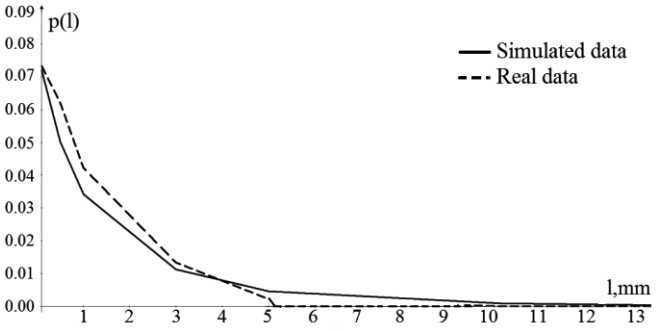


Figure 2: Probability $p(l) = P(\chi_{N-1} > l, \chi_N > l)$

Since model passed verification, it may be used for study of a river flow as a function of precipitation in the river basin. One of characteristics that was studied is a behavior of the river flow in Berd's lower reach in case of precipitation absence in the river basin. Table 2 shows probability that river flow ψ in Stariy Iskitim is greater than $E\psi = 14.79\ m^3/s$ in case of nd consequent dry days on nn % of grid nods. Here $E\psi = 14.79\ m^3/s$ is an average river flow in Stariy Iskitim, estimated on a basis of real data. This characteristic shows not only dependence of river flow on precipitation, but also indirectly shows relationship between river flow and other influencing factors.

Table 2. Probability That River Flow in Stariy Iskitim is Greater Than Its Average Value in Case of nd Consequent Dry Days on nn % of Nods. July

		nn %			
		30	50	70	100
nd, days	1	0.32	0.28	0.21	0.15
	3	0.29	0.13	0.09	0.08
	5	0.15	0.07	0.04	0.02
	7	0.09	0.05	0.01	0.01

Simulated trajectories of a conditional joint random field may also be used for study of the river flow in case of heavy rainfall. Table 3 shows several values of a conditional distribution function $CF(h)$ of the river flow in Maslyanino under a condition that precipitation in all grid nods within a 40 km radius of Maslyanino is greater than 12 *mm.* For comparison, real data based values of a river flow non-conditional distribution function $F_\psi(h)$ in Maslyanino are also given in the Table 3.

Table 3. Conditional Distribution Function of the River Flow in Maslyanino Under a Condition That Precipitation in All Grid Nods Within a 40 km Radius of Maslyanino is Greater Than 12 *mm.* July

h, m^3/s	5	6	7	10	12	13
$CF(h)$	0.64	0.69	0.75	0.94	0.99	1.00
$F_\psi(h)$	0.22	0.47	0.65	0.88	0.90	0.92

CONCLUSION

Considered in this paper model is a parametric one. This makes possible application of the model for study of a river flow in different geographical areas varying model parameters depending on a region under consideration. After proper model setting and thorough check it is planned to use this model in a cross-disciplinary project which is focused on environmental changes in the Ob River lower reach in case of artificial river diversion and currently appearing permafrost thaw.

ACKNOWLEDGEMENTS

This work was supported by the Russian Foundation for Basis Research (grants No 15-01-01458-a, 16-01-00145-a, 16-31-00123-mol-a, 16-31-00038-mol-a) and the President of the Russian Federation grant (No MK-659.2017.1).

REFERENCES

Ermakov, S.M. and G.A. Mikhailov. 1982. *Statistical simulation*. Moscow: Nauka. [Ермаков, С.М. и Г.А. Михайлов. 1982. Статистическое моделирование. Москва: Наука.]

Hartkamp, A.D.; K. De Beurs; A. Stein and J.W. White.1999. "*Interpolation Techniques for Climate Variables*". NRG-GIS Series 99-01. Mexico, D.F.: CIMMYT.

Kargapolova, N.A. 2017. "Stochastic Simulation of Non-stationary Meteorological Time-series. Daily Precipitation Indicators, Maximum and Minimum Air Temperature Simulation Using Latent and Transformed Gaussian Processes". In *Proceedings of the 7th International Conference on Simulation and Modeling Methodologies, Technologies and Applications* (Madrid, Spain, July 26-28).

Kargapolova, N.A. and V.A. Ogorodnikov. 2017. "Conditional stochastic model of daily precipitation and river flow joint spatial field". In *Proceedings of the International Workshop "Applied Methods of Statistical Analysis. Nonparametric Methods in Cybernetics and System Analysis - AMSA'2017* (Krasnoyarsk, Russia, Sept. 18-22).

Komlev, A.M. 2002. *Regularities of formation and methods of river flow estimation*. Perm: Perm University Publishing. [Комлев, А.М. 2002. *Закономерности формирования и методы расчетов речного стока*. Пермь: Изд-во Перм. ун-та.]

Marchenko, A.S. and V.A. Ogorodnikov. 1991. Probabilistic models of dry and rainy days sequences. Preprint No 933. Novosibirsk: CC SB AS USSR. [Марченко, А.С. и В.А. Огородников. 1991. "Вероятностные модели последовательности сухих и дождливых суток". Препринт № 933. Новосибирск: ВЦ СО АН СССР.]

Ogorodnikov, V.A. 2013. *Numerical modelling of discrete random processes and fields.* Novosibirsk: NSU.

Ogorodnikov, V.A. and S.M. Prigarin. 1996. *Numerical modelling of random processes and fields: algorithms and applications*. The Netherlands, Utrecht: VSP.

Ogorodnikov, V.A. and O.V. Sereseva. 2015. "Multiplicative numerical stochastic model of daily sums of liquid precipitation fields and its use for estimating statistical characteristics of extreme precipitation regimes". *Atmospheric and Oceanic Optics,* Vol. 28, No 4, p. 328-335.

Prigarin, S.M. 2005. *Methods for the numerical simulation of random processes and fields*. Novosibirsk: ICMaMG SB RAS. [Пригарин, С.М. 2005. *Методы численного моделирования случайных процессов и полей*. Новосибирск: ИВМиМГ СО РАН.]

Richardson, C. W. and D.A. Wright. 1984. *WGEN: A Model for Generating Daily Weather Variables.* U. S. Department of Agriculture, Agricultural Research Service, ARS-8.

Shlychkov, V.A.; V.A. Ogorodnikov and O.V. Sereseva. 2015. " Joint numerical stochastic model of the daily river flow time series space-time fields of daily sums of liquid precipitation". In *Proceedings of the International Conference "Interexpo Geo-Siberia"* (Novosibirsk, Russia, April 20-22), Vol. 4, p. 150-154. [Шлычков, В.А.; В.А. Огородников и О.В. Сересева. 2015. "Совместная численная стохастическая модель временных рядов суточного стока реки и пространственно-временных полей суточных сумм жидких осадков". В *Трудах международной конференции Интерэкспо Гео-Сибирь* (Новосибирск, Россия, 20-22 апреля), Т.4, 150-154.]

Sluiter, R. 2009. "Interpolation methods for climate data". Literature review. The Netherlands, De Bilt: KNMI.

Sobol, I.M. 1973. Numerical Monte Carlo methods. Moscow: Nauka. [Соболь, И.М. 1973. *Численные методы Монте-Карло*. Москва: Наука.]

AUTHOR BIOGRAPHY

NINA KARGAPOLOVA was born in Novosibirsk, Russia and went to the Novosibirsk State University (NSU), where she studied Mathematics and obtained her Bachelor and Master degrees in 2008 and 2010. After 3 years of postgraduate training she got a Doctor degree in Mathematics (area of competence – Computational Mathematics). During the last 4 years she has been working as a researcher at the Institute of Computational Mathematics and Mathematical Geophysics (Russian Academy of Science) and at the same time she is Associate Professor at NSU.

FORECAST OF SELECTED QUALITY INDICATORS OF WASTEWATER FLOWING TO THE TREATMENT PLANT USING SELECTED BLACK BOX METHODS

Bartosz Szeląg[*], Krzysztof Barbusiński[**], Agnieszka Operacz[**], Jan Studziński[***]
[*] Kielce University of Technology, 25-314 Kielce, Poland
bszelag@tu.kielce.pl
[**] Silesian University of Technology, 44 – 100 Gliwice, Poland
krzysztof.barbusiński@polsl.pl
[***] Systems Research Institute, Polish Academy of Sciences, 01-447 Warsaw, Poland
studzins@ibspan.waw.pl

KEYWORDS
Support vector machines, boosted tree, MARS, black–box, wastewater modelling.

ABSTRACT

In this study some black-box models are employed to model various indicators of the wastewater quality (characterized by biochemical oxygen demand, chemical oxygen demand, the content of total and ammoniacal nitrogen, suspended solids and total phosphorus) at the inflow to a sewage treatment plant using as the predictors the sewage inflow and temperature as well as the indicators measured in the previous time steps. Considering all tested methods one can see that in most cases the best predictions have been obtained by the method of multivariate adaptive regression splines. Moreover the models based only on inflow and temperature inputs have got a noticeably lower errors than those once conditioned by indicators.

INTRODUCTION

It is a difficult task to operate a wastewater treatment plant because that involves maintaining many technological processes on an appropriate level in order to get a satisfactory degree of contaminant reduction. However, due to the stochastic character of the wastewater inflow and quality, affected by rainfall depth, by amount of water produced, by season of the year and contaminant load, the sewage inflow is characterised by dramatic and abrupt changes. Because of that inflowing contaminant load and wastewater amount lead to disturbances in the facility operation and may result in wrong decisions taken by the plant operator responsible for the wastewater treatment. The practice shows that it can be costly and time-consuming to restore the proper operation of the facility and therefore it is important to forecast both the quantity and quality of the wastewater inflowing treatment plant. Modelling those occurances is important because that creates the possibility to identify abnormal events due to random factors (e.g. intensive rainfall). That offers time to the plant staff to prepare the facility for the events by selecting appropriate settings in crucial plant units.

The bioreactor parameters can be determined using physical models describing the kinetics of biochemical changes in nitrogen and carbon compounds in individual bioreactor units (Harremoës et al. 1993, Wichern et al. 2001, Mannina et al. 2016). Such models make possible to forecast operation of separate (storm water, urban) and combined sewage systems. Those models are based on systems of differential equations describing processes occurring in a plant (Leandro and Martins 2016, Krebs et al. 2013). The models need to be calibrated with high resolution. Due to a large number of parameters that must be defined and to interactions between them a lot of calibration problems are then to solve (Weijers and Vanrolleghem 1997, Belia et al. 2009, Szeląg et al. 2016). The applicability of physical models is limited because of their uncertainties.
Therefore some black-box methods can be employed to model the quantity and quality of wastewater inflowing a treatment plant (Häck and Kö hne 1996, Hamed at al. 2004). In those methods at the training stage the model structure is produced being a basis for defining dependent model variables (flow to the treatment plant, content of biogenic compounds). A literature review with respect to black-box methods was carried out for the study. In Dellana and West (2009) ARIMA (autoregressive integrated moving average) models and artificial neural networks (ANN) have been successfully applied to simulate the concentration of suspended solids, total nitrogen and phosphorus in the wastewater inflow. El–Din and Smith (2002) demonstrated the ARIMA model applicability for prediction of chemical oxygen demand and suspended solids in stormwater. Raha (2007) used ANN to model wastewater quality (biogenic compounds) but his simulation results showed much discrepancy between measured and modelled values. Verma et al. (2013) employed a number of models, including Support Vector Machines (SVM), Random Forests (RF), Multivariate Adaptive Regression Spline (MARS) and k-Nearest Neighbour (k-NN) methods to forecast the concentration of total suspended solids. Small differences between simulated and measured values have been then found. The analyses conducted by Minsoo et al. (2016) confirmed the possibility of using k-NN method to model wastewater quality indicators such as chemical oxygen demand, total nitrogen and phosphorus. Artificial neural

networks and boosted trees (BT) method were successfully applied to determine total nitrogen by Hosseini (2011).
The models for predicting wastewater quality indicators presented in the literature were developed exclusively on the basis of values of a given parameter (i.e. of the contents of total nitrogen, carbon, phosphorus and suspended solids in the wastewater inflow) measured on previous days at regular times intervals. Obtaining such results is possible by means of continuous online monitoring. If monitoring is not available then determination of individual indicators describing the content of carbon, nitrogen and phosphorus in the wastewater must be done at a laboratory. However, regarding potential failures of measurement devices or technical problems related to determination of selected indicators of wastewater quality, a possibility of obtaining successive inputs data for the models is limited. Therefore it is reasonable to develop statistical models to calculate wastewater quality indicators at the sewage inflow on the basis of parameters that can measured in a relatively easy way without specialist equipment, or without necessity to perform their time-consuming determinations, e.g. BOD_5.
In this study three black-box methods were used to model the wastewater quality indicators. The study analyses the possibility of predicting selected contaminant indicators based on the sewage inflow and temperature measurements, and the possibility of predicting quality indicators based on their measurements.

THE OBJECT OF INVESTIGATION AND THE MEASUREMENTS DATA

The investigations concerned the urban wastewater treatment plant located in the commune of Sitkówka–Nowiny. The plant receives sanitary wastewater from the separate wastewater system mostly of Kielce city and of Sitkówka–Nowiny commune. The design capacity of the treatment plant is 72.000 m^3/d, and it is capable of serving a population equivalent (P.E.) of 275.000.
Wastewater delivered to the treatment plant is mechanically pre-treated using step screens and aerated grit chambers, with separate grease traps. Then wastewater is pumped to four primary settling tanks from which primary sludge is delivered to the secondary sludge pumping station. Leaving the mechanical unit of the plant, wastewater is conveyed, via a flume, into biological unit. Bioreactor, with separate denitrification and nitrification tanks, constitutes the core component of the treatment plant. In the bioreactor, contaminants are finally removed from wastewater. In the initial denitrification tank, nitrogen compounds are partially removed from wastewater. Afterwards, wastewater is directed to dephosphatation tanks for the removal of phosphorus compounds. Then wastewater together with activated sludge is transferred to four secondary settling tanks, from where after clarification it flows to the receiving water body, i.e. the river Bobrza.
An online monitoring conducted by the waterworks at the treatment plant since 2012 provides measurements of parameters describing inflowing wastewater quantity and temperature. The parameters concerning inflowing wastewater quality (contents of carbon, nitrogen, and phosphorus compounds and of suspended solids) are determined at the plant laboratory.

METHODOLOGY

In this study the calculations have been performed for two cases. The first case concerns the analysis of the ability to model the wastewater quality indicators such as chemical and biochemical oxygen demand (COD, BOD_5), total and ammoniacal nitrogen content (TN, NH_3), total suspended solids (TSS), and total phosphorus (TP) only on the basis of the results of their recent measurements. In the treatment plant the measurements of the wastewater quality indicators are conducted at various time intervals (from 1 to 5 days), therefore in this study the models were developed to simulate the indicators with a variable forward time step. In the models the time interval between successive measurements of the selected input data provided an additional explanatory variable.
The second case concerns analyzing the possibility to predict the values of contaminant indicators in wastewater on the basis of measurements of the flow and temperature of wastewater inflowing the plant. Those data and relevant assumptions provided the basis for developing statistical models to calculate COD, BOD_5, TSS, TN, TP and NH_3 using the MARS, BT and SVM methods.
To make the modelling process successful and to properly assess the models obtained, the data were partitioned into the training and testing sets (75% and 25%). Prior to the start of the modelling, input and output data normalization was performed by means of the transformation (Rutkowski 2006):

$$\overline{A}_i = \frac{A_i - \min A}{\max A - \min A} \tag{1}$$

where: $\overline{A}_i$ - normalized value of i-th element from set A containing the measurements, A_i – measured value of i-th element from set A, max$A$ – maximum value of all elements in A, min$A$ - minimum value of of all elements in set A.
MARS method is one of the numerous data mining tools used for solving regression problems (Zhang & Goh 2016, Gutiérrez et al. 2009). The method is an extension of the classic approach to determine predictors in regression models. In MARS method the variation ranges of input data are divided into three intervals, in which the analysed variables may have a different effect on the analysed phenomenon. Interval limits are established on the basis of threshold value (t); if the value of the variable analysed is below or above t, then it will have a different weight or sign. To differentiate between the variable values being smaller or greater than threshold value t_i, the following basic functions are used:

$$h(X) = \alpha_i \cdot (\max(0, X - t)) \tag{2}$$

where: $h(X)$ – vector of basic functions for individual variables (x_i), for which the following condition is satisfied:

$$x_i - t_i = \begin{cases} x_i - t_i; & for \quad x_i > t_i \\ 0; & for \quad x_i \le t_i \end{cases} \qquad (3)$$

In MARS method, regression dependence has a form of a spline function, which is a linear combination of the product of basic functions and individual weights, and which can be written as follows:

$$f(X) = \alpha_0 + \sum_{m=1}^{M} \alpha_m \cdot h_m(X) \qquad (4)$$

where: $X=[x_1, x_2, \ldots, x_i]$ – vector of input data, α_m – values of weights, h_m – basic functions.

To estimate model parameters, a special algorithm was designed to search the space of observations in order to specify threshold values (nodes). Using selected nodes, basic functions are created, which together with appropriate weights constitute a foundation for the description of the analysed phenomenon. It should be noted that so far that method has not been used to model the operation of a wastewater treatment plant.

The algorithm included in MARS is based on recursive partitioning of the features space and it comprises two alternately occurring stages, which go on until stopping criterion is reached. It constitutes the value of generalised cross validation (Friedman 1991). At the first stage the model complexity is increased by the addition of basic functions until the maximum number, declared by the user, is reached. Then the deletion procedure (the so-called pruning) is triggered, which eliminates less important basic functions, and thus independent variables, from the model.

Boosted trees (BT) constitute an implementation of the stochastic gradient boosting method applied to classification and regression problems (Friedman 2002). The method concept involves the generation of decision trees. Each subsequent tree is used to identify cases misclassified by previous trees. Computations show that for some estimation and prediction issues, the forecasts obtained by growing boosted trees are much closer to actual values of variables modelled than the solutions produced by single regression trees.

Support Vector Machines (SVM) cover a group of methods developed by Vapnik (1998), firstly exclusively for classification purposes, which expanded over time to include regression issues (SVR). For that reason the dependence between the model output and input variables can be non-linear. As a result, in this method a non-linear transformation of *N*–dimensional space to *K*–dimensional features space of larger size is applied. In this study the support vector regression method with a radial kernel function was applied to predict wastewater quality. The radial kernel function was used to minimise the functional of the following form:

$$\sum_{i=1}^{m} \frac{c}{m} |y_i - f(x_i)|_{\varepsilon} + \frac{1}{2} \cdot \|f\|_k^2 \qquad (5)$$

where: $|y_i - f(x_i)|_\varepsilon = \max\{0, |y_i - f(x_i) - \varepsilon\}$, ε – permissible error, $\|f\|$ - norm f in Hilbert space, c – a constant selected by the user, depending on value ε (Burges 2000), m – size of the training set, $f(x_i)$ – value of function $f(x)$ at point x_i, described by equation:

$$f(x) = \sum_{i=1}^{N_{sv}} \left(\alpha_i - \alpha_i^{'}\right) \cdot K(x, x_i) + w_0 \qquad (6)$$

in which: w_0 – deviation, N_{sv} – number of support vectors, dependent on c and ε; α_i, $\alpha_i^{'}$ –Lagrange multipliers, $K(x,x_i)$ – kernel function with radial basic functions (Burges 2000).

Criteria of model assessment

To evaluate predictive abilities of mathematical models for forecasting biochemical and chemical oxygen demands, suspended solids, total and ammonia nitrogen, and total phosphorus concentrations, the following formulas were used:

- mean absolute error (MAE):

$$MAE = \frac{1}{n} \cdot \sum_{i=1}^{n} \left| y_{i,obs} - y_{i,pred} \right| \qquad (7)$$

- mean absolute percentage error (MAPE):

$$MAPE = \frac{1}{n} \cdot \sum_{i=1}^{n} \left| \frac{y_{i,obs} - y_{i,pred}}{y_{i,obs}} \right| \cdot 100\% \qquad (8)$$

where: $y_{i,obs}, y_{obl}$ – measured and calculated values of variables modelled, n – data set size.

RESULTS

On the basis of the measurements of the quality and quantity of wastewater flowing to the treatment plant, ranges of measurement variations were established (Table 1). The data in Table 1 show that the values of chemical and biochemical oxygen demands, of suspended solid, of total and ammonia nitrogen, and of total phosphorus concentrations varied largely. Consequently, wastewater quality indicators concerned, which are inputs data for models depicting the kinetics of changes in carbon, nitrogen and phosphorus compounds in the bioreactor, differ much either. Therefore it is necessary to forecast those parameters in order to model bioreactor operation.

Table 1. Ranges of variation of values of parameters describing quantity and quality of raw wastewater

Variable	Min	Max
Q, m^3/d	32564	86592
T, °C	10.6	20.9
BOD_5, mg/dm^3	109,0	557,0
COD, mg/dm^3	331.0	1050.0
SS, mg/dm^3	126.0	572.0
NH_4, mg/dm^3	24.4	65.9
TN, mg/dm^3	39.9	124.1
TP, mg/dm^3	3.1	12.6

Among the methods considered in the study, only in MARS method the estimation algorithm allows to delete independent variables that have a negligible effect on the dependent variable. Consequently this method was employed firstly to simulate selected quality indicators of wastewater inflowing the plant. Using the independent variables established by MARS method, simulations of concentrations of biogenic compounds and suspended solids were performed using also other methods: SVM and BT. To do it the values of wastewater quality parameters were used that were obtained from the measurements. Finally, the forecasts of quality indicators (BOD_5, COD, SS, TN, NH_4 and TP) of wastewater inflowing the treatment facility were prepared based on the wastewater temperature and flow measurements.

Table 2. Independent variables by forecasting wastewater quality indicators (BOD_5, COD, SS, TN, NH_3, TP)

Indicators	Variables	
	Quantitative	Qualitative
COD	Q(t-1), Q(t-2), T(t-1), T(t-2)	COD(t-1)
BOD_5	Q(t-1), Q(t-2), Q(t-3), T(t-1)	BOD_5(t-1)
SS	Q(t-1), Q(t-2), Q(t-3)	SS(t-1)
TN	Q(t-1), Q(t-2), T(t-1)	TN(t-1)
NH_4	Q(t-1), T(t-1), T(t-2), T(t-3)	NH_3(t-1)
TP	Q(t-1), Q(t-2), Q(t-3), T(t-1)	TP(t-1)

Table 2 presents determined independent variables providing a basis for predictions of wastewater quality indicators. Tables 3 and 4 give values of matching errors (MAE, MAPE) by BOD_5, COD, SS, TN, NH_4 and TP simulations against the measurements data. The simulation values were produced using statistical models, after five-fold crossvalidation. As to the parameters of the methods used by the modelling they were as follows: in MARS method the number of basic functions varied from 5 to 10, in SVM method *c* values ranged from 9 to 11, and in BT method the best models were obtained after 16 ÷ 50 iterations.

Table 3. Fitting errors of wastewater indicator models against the measurements data by using sewage temperature and flow as predictors; a) MARS and SVM models; b) BT models

a)

Variable	MARS				SVM			
	training		test		training		test	
BOD_5	32.51	13.54	35.11	14.12	39.30	15.46	39.67	16.42
COD	95.94	14.34	104.54	14.53	116.05	15.16	118	16.44
SS	45.38	15.28	45.81	15.53	48.66	17.47	51.67	18.33
TN	4.77	6.65	5.24	6.74	5.72	7.49	5.83	7.51
NH_4	3.79	7.81	3.81	7.90	4.25	8.03	4.25	8.82
TP	0.82	12.37	0.90	12.43	0.90	13.00	0.96	13.48

b)

Variable	BT			
	training		test	
BOD_5	MAE	MAPE	MAE	MAPE
COD	36.59	15.02	38.62	16.04
SS	108.42	14.59	114.1	15.58
TN	44.63	14.48	44.97	15.35
NH_4	5.28	6.88	5.39	7.01
TP	3.67	8.26	4.04	8.49

Data in Table 2 indicate that to predict the concentrations of biogenic compounds and of suspended solids, it is sufficient to know values of BOD_5, COD, SS, TN, NH_4, TP from the previous measurements. In case of calculating chemical and biochemical oxygen demands and the concentrations of suspension, of total nitrogen and of phosphorous, the values of daily inflow and wastewater temperature provide the explanatory variables.

Table 4. Fitting errors of wastewater indicator models against the measurements data by using quality indicators as predictors; a) MARS and SVM models; b) BT models

a)

Variable	MARS				SVM			
	training		test		training		test	
BOD_5	32.51	13.54	35.11	14.12	39.30	15.46	39.67	16.42
COD	95.94	14.34	104.54	14.53	116.05	15.16	118	16.44
SS	45.38	15.28	45.81	15.53	48.66	17.47	51.67	18.33
TN	4.77	6.65	5.24	6.74	5.72	7.49	5.83	7.51
NH_4	3.79	7.81	3.81	7.90	4.25	8.03	4.25	8.82
TP	0.82	12.37	0.90	12.43	0.90	13.00	0.96	13.48

b)

Variable	BT			
	training		test	
BOD_5	MAE	MAPE	MAE	MAPE
COD	36.59	15.02	38.62	16.04
SS	108.42	14.59	114.1	15.58
TN	44.63	14.48	44.97	15.35
NH_4	5.28	6.88	5.39	7.01
TP	3.67	8.26	4.04	8.49

On the basis of the data presented in Tables 3 and 4 one can conclude that the lowest values of the prediction errors for different concentrations were obtained by the models based on the temperature and daily wastewater inflow into the plant. In case of modelling BOD_5 based on Q and T (Table 2), the smallest errors were obtained using MARS method (MAE=35.11 mg/ dm^3 and MAPE=14.53%), while the highest errors were obtained by SVM method (MAE=39.67 mg/ dm^3 and MAPE=16.43%). Regarding the model to forecast BOD_5 based on BOD_5(t-1) input, the obtained prediction errors are similar to those once got by SVM and MARS methods (MAE=41.98 mg/dm^3, MAPE=17.47% and MAE=40.83 mg/dm^3, MAPE=17.03%, respectively). The model to predict COD based on Q and T manifested the

lowest errors in case of MARS method and the highest ones when SVM method was applied.

It can be seen from Table 3 that errors for suspended solids prediction based on Q(t-1), Q(t-2) and Q(t-3) measurements are comparable by MARS method (MAE=45.81 mg/dm^3 and MAPE=15.53%) and BT method (MAE=44.97 mg/dm^3 and MAPE=15.35%). The model for SS prediction obtained by SVM method shows the worst predictive abilities (MAE=51.67 mg/dm^3 and MAPE=18.33%). The model for SS prediction based on SS(t-1) inputs showed similar error values (Table 4) by all methods applied (MARS, SVM, BT) as MAE and MAPE values varied only slightly, ranging 54.93÷56.67 mg/dm^3 and 19.27 ÷ 20.10 %, respectively.

Analysing the calculated errors by the models predicting the total nitrogen (Table 3) based on Q and T inputs, one can see that MARS method (MAE=5.24 mg/dm^3 and MAPE=6.74%) produced slightly smaller error values than the two other methods. In the models for TN forecast, developed by means of BT and SVM methods and based on Q(t-1), Q(t-2) and T(t-1) inputs, the values of errors are MAE=5.39 mg/dm^3, MAPE=7.01% and MAE=5.83 mg/dm^3, MAPE=7.51%, respectively. By modelling TN based on TN(t-1) inputs, the determined MAE and MAPE errors are greater compared with the models based on Q and T. The respective error values are 7.05 mg/dm^3 and 9.63% by MARS method, 7.02 mg/dm^3 and 9.49% by SVM method and 7.14 mg/dm^3 and 9.67% by BT method.

In the models predicting ammonia nitrogen NH_4 based on Q and T (Table 2), the lowest error values were found for MARS method (MAE=3.81 mg/dm^3 and MAPE=7.90%). Slightly worse NH_4 simulation results were received with SVM method (MAE=4.25 mg/dm^3 and MAPE=8.82%) and BT method (MAE=4.04 mg/dm^3 and MAPE=8.49%). Regarding the models for NH_4 based on NH_4(t-1) inputs, the errors achieved (Table 4) are greater than these once got by MARS method (MAE=4.84 mg/dm^3 and MAPE=10.57%), by SVM method (MAE=4.97 mg/dm^3 and MAPE=10.57%) and by BT method (MAE=4.91 mg/dm^3 and MAPE=10.63%).

From Tables 3 and 4 results also that the smallest prediction errors for total phosphorus (TP) models based on Q and T (Table 2) were obtained with MARS method (MAE=0.90 mg/dm^3 and MAPE=12.43%). As to other methods applied, the fitting errors calculated were slightly greater by SVM method (MAE=0.96 mg/dm^3 and MAPE=13.48%) as well as by BT method (MAE=0.92 mg/dm^3 and MAPE=12.90%). By TP models based on TP(t-1) inputs, the errors received (Table 4) are higher compared with the models applying Q and T inputs (Table 2).

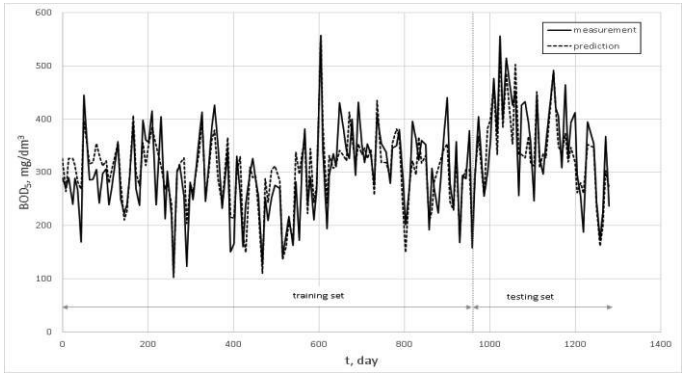


Figure 1. Comparison of the measurements and computations of BOD_5 while using MARS method.

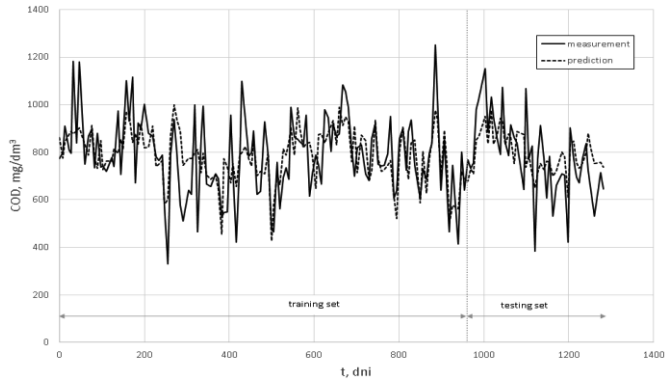


Figure 2. Comparison of the measurements and computations of COD while using MARS method.

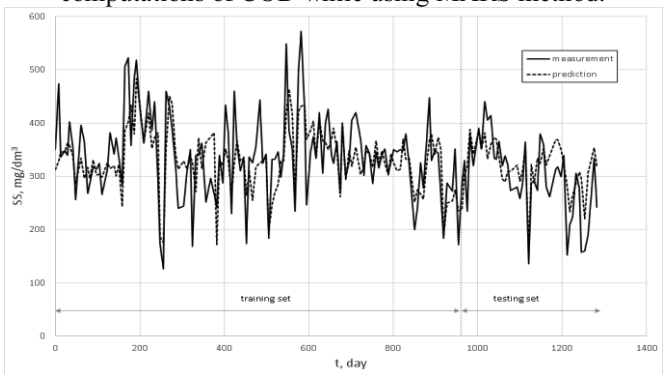


Figure 3. Comparison of the measurements and computations of SS while using MARS method.

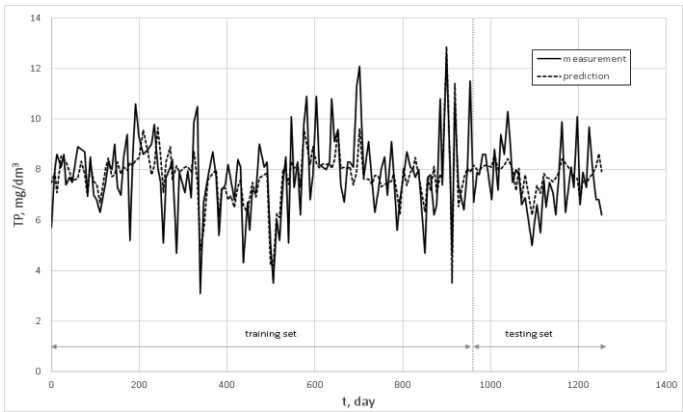


Figure 4. Comparison of the measurements and computations of TP while using MARS method.

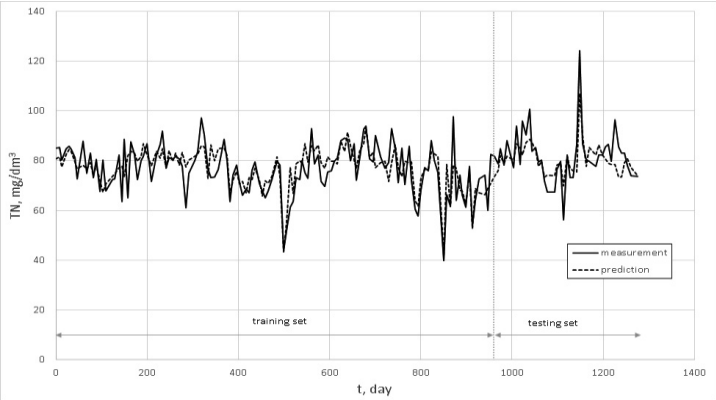


Figure 5. Comparison of the measurements and computations of TN while using MARS method.

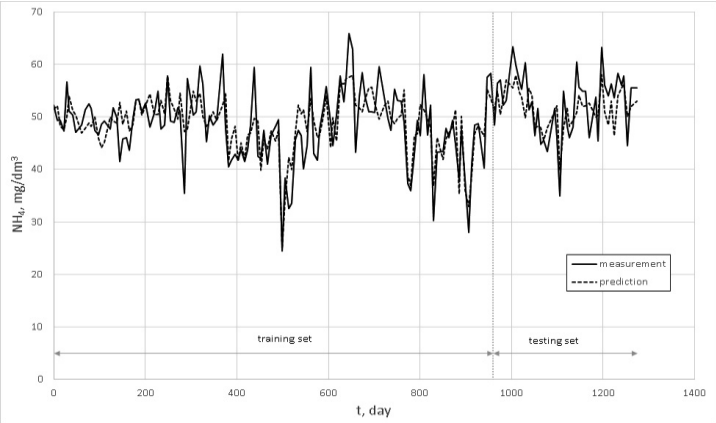


Figure 6. Comparison of the measurements and computations of NH_4 while using MARS method.

Figures 1 ÷ 6 present in a graphic form values of individual biogenic compounds and of suspended solids, measured and determined in simulations using as the predictors sewage temperature and daily sewage inflow to the treatment plant.

CONCLUSIONS

Analyses conducted for the study showed that for forecasting the values of wastewater quality indicators, including biochemical and chemical oxygen demands and the concentrations of total and ammonia nitrogen and of total phosphorus, the following time series methods of boosted trees, of support vectors and of multivariate adaptive regression splines can be used. Using those methods the best matching of simulation results against the measurements was found by MARS method, what is confirmed by the calculated values of absolute and relative errors of modelling.

Additionally, the computations occurred showed that for forecasting the concentrations of biogenic compounds and of suspended solids, it is possible to rely exclusively on the measurements of daily wastewater inflow and of its temperature.

There is to note that the prediction of wastewater quality indicators on the base of only wastewater inflow and temperature measurements is simple and easy and such approach does not require tedious and time-consuming laboratory tests of those parameters that are commonly executed while operating a treatment plant.

Such approach makes also possible to forecast simply the concentrations of biogenic compounds in wastewater in cases of failures of monitoring systems installed in the plants. The settings of crucial units of the treatment plant bioreactor shall be selected based on simulation results of quality and quantity of inflowing wastewater. If such simulations can be done easy by means of the methods mentioned then the plant operator gets a convenient tool ensuring a successful reduction of contaminants in the wastewater treated.

REFERENCES

Belia E., Amerlinck Y., Benedetti L., Sin G., Johnson B., Vanrolleghem P. A., Gernaey K.V., Gillot S., Neumann M. B., Rieger L., Shaw A. & Villez K. 2009. "Wastewater treatment modelling: Dealing with uncertainties." *Water Science and Technology*, 60(8), 1929–1941.

Burges C. 2000. "A tutorial on support vector machines for pattern recognition." In: *Knowledge discovery and data mining,* U. Fayyad (eds.), Kluwer, pp. 1 – 43.

Dellana S.A. & West D. 2009. "Predictive modelling for wastewater applications: Linear and nonlinear approaches." *Environmental Modelling and Software*, 24, 96-106.

Friedman J. 1991. "Multivariate Adaptive Regression Splines." Annals of Statistics, 19, 1-141.

Friedman J. H. 2002. "Stochastic gradient boosted." *Computational Statistics and Data Analysis*, 38(4), 367 – 378.

Gutiérrez G., Schnabel Á., S. & Contador J. F. L. 2009. "Using and comparing two nonparametric methods (CART and MARS) to model the potential distribution of gullies." *Ecological Modelling*, 220(24), 3630-3637.

Hamed M., Khalafallah M.G. & Hassanein E. A. 2004. "Prediction of wastewater treatment plant performance using artificial neural network." *Environmental Modeling and Software*, 19, 919–928.

Harremoës P., Capodaglio A. G., Hellstrom B. G., Henze M., Jensen K. N., Lynggaaard-Jensen A., Otterpohl R. & Soeborg H. 1993. "Wastewater treatment plants under transient loading-performance, modeling and control." *Water Science and Technology*, 27(12), 77 - 115.

Häck M. & Kö hne M. 1996. "Estimation of wastewater process parameters using neural networks." *Water Science and Technology*, 33(1), 101-115.

Krebs G., Kokkonen T., Valtanen M., Koivusalo H. & Setälä H. 2013. "A high resolution application of a stormwater management model (SWMM) using genetic parameter optimization." *Urban Water Journal*, 10(6), 394 – 410.

Leandro J. & Martins R. 2016. "A methodology for linking 2D overland flow models with the sewer network model SWMM 5.1 based on dynamic link libraries." *Water Science and Technology*, 73(12), 3017 – 3026.

Mannina G., Cosenza A. & Viviani G. 2016. "Sensitivity and uncertainty analysis of an integrated membrane bioreactor model." *Desalination and Water Treatment*, 57(21), 9531 – 9548.

Minsoo K., Yejin K., Hyosoo K., Wenhua P. & Changwon K. 2016. "Evaluation of the k – nearest neighbour method for forecasting the influent characteristics of wastewater treatment plant." *Frontiers of Environmental Science & Engineering*, 10(2), 299-310.

Rutkowski L. 2006. "Artificial intelligence methods and techniques." *(in Polish).* Warszawa, PWN.

Szeląg B., Kiczko A. & Dąbek L. 2016 “Analiza wrażliwości i niepewności w modelach hydrodynamicznych na przykładzie zlewni zurbanizowanej (Sensitivity and uncertainty analysis of hydrodynamic models in the urban catchments – case study).” *Ochrona Środowiska*. 38(3), pp. 15 – 21.

Weijers S.R. & Vanrolleghem P.A. 1997. “A procedure for selecting best identifiable parameters in calibrating activated sludge model no.1 to full-scale plant data.” *Water Science and Technology*, 36(5), 69–79.

Wichern M., Obenaus F., Wulf P. & Rosenwinkel K.H. 2001. “Modelling of full-scale wastewater treatment plants with different treatment processes using the Activated Sludge Model no. 3.” *Water Science and Technology*, 44(1), 49 – 56.

Vapnik V. 1998. “Statistical Learning Theory.” John Wiley and Sons. New York.

Verma A., Wei X. & Kusiak A. 2013. “Predicting the total suspended solids in wastewater: A data-mining approach.” *Engineering Applications of Artificial Intelligence*, 26(4), 1366–1372.

Zhang W. & Goh A. T.C. 2016. “Multivariate adaptive regression splines and neural network models for prediction of pile drivability.” *Geoscience Frontiers*, 7(1), 45 – 52.

Raha D. 2007. “Exploring Artificial Neural Networks (ANN) Modelling for a Biological Nutrient Removal (BNR) sewage treatment Plant (STP) to Forecast Effluent Suspended Solids.” *Indian Chemical Engineering*, Vol. 49, no. 3.

El-Din A.G., Smith D.W. 2002. “Modelling approach for high flow rate in wastewater treatment operation.” *Journal of Environmental Engineering and Science*, Vol. 1, No.4, pp. 275-291.

ENERGY FORECASTING AND OPTIMIZATION

NEURAL NETWORKS MODELS AND ELECTRICITY DEMAND FORECASTING: AN ECONOMETRIC APPROACH

Francis Bismans
BETA-University of Lorraine, Nancy, France
COEF, Nelson Mandela University, South Africa
E-mail: Francis.Bismans@univ-lorraine.fr

Igor Litvine
COEF, Nelson Mandela University, Port Elizabeth, South Africa
E-mail: Igor.Litvine@mandela.ac.za

KEYWORDS
Neural networks (ANN), electricity consumption, forecasting, econometric models, recessions.

ABSTRACT

This paper deals with a so-called feedforward neural network model which we consider from a statistical and econometric viewpoint. It was shown how this model can be estimated by maximum likelihood. Finally, we apply the ANN methodology to model demand for electricity in South Africa. The comparison of forecasts based on a linear and ANN model respectively shows the usefulness of the latter.

INTRODUCTION

Artificial neural networks (ANN) subsume a set of models which have been developed in the cognitive sciences to understand the functioning of human brain. These models were originated in the publications by (McCulloch and Pitts 1943) and in the study of perceptron by (Rosenblatt 1958). However, at that time the capabilities of computing were very limited, so these early models were too simple to explain the complexities of the actual operation of the brain. Consequently, with growth of computing powers, more complex ANN structures and network learning methods were designed, peculiarly in the investigations of (Rumelhart et al. 1986) and (McClelland et al. 1986).

In this study we are interested in applying econometric approach to ANN models. From this viewpoint, the inspiring and path breaking contribution is that of (Kuan and White 1994). More recently, (Kuan 2008) gave a review of the matter from an econometric perspective. In brief, ANN models for an econometrician constitute a specific set of non-linear models and "learning" is understood as an estimation of model parameters.

Applications of econometric ANN models are been numerous in the field of market finance. For a presentation of main results, see e.g. (Franses and van Dijk 2000, chapter 5). Forecasting, especially macroeconomic, was also an area to prospect well, as evidenced by the studies of (Swanson and White 1997, McMenamin (1997), Zhang, G. et al. 1998, Rech 2002, White 2006, Medeiros et al. 2006, and Ozdemir et al. 2010), a short list in the vast literature on the subject.

Fundamentally, the aim of this paper is to develop an application of neural networks focused on the electricity production in South Africa**.**

THE ANN (*k*,*q*) MODEL

From an econometric perspective, the "single hidden-layer feedforward" model will be written as follows:

$$y_t = \boldsymbol{\alpha}'\mathbf{z}_t + \sum_{j=1}^{q} \beta_j G(\boldsymbol{\gamma}'_j \mathbf{z}_t) + \varepsilon_t, \quad t = 1,\cdots,T, \quad (1)$$

where y_t is the dependent variable (= output), $\mathbf{z}_t = (1, y_{t-1},\cdots, y_{t-p}, x_{1t},\cdots, x_{kt})'$ is the vector of the explaining variables, including the constant and the delayed values of y_t, $\boldsymbol{\gamma}$ and $\boldsymbol{\alpha}$ are $(p+k+1)$ vectors of parameters and the *G*s are the activation functions. Consequently, the relationships between y_t and $\mathbf{x}_t$ are possibly nonlinear. Moreover, as usual, ε_t is a Gaussian white noise with null mathematical expectation and constant variance.

The activation functions are usually restricted to those which have values between zero and one. From this viewpoint, the logistic function

$$G(x) = 1/(1+e^{-x}), \quad x \in \mathbb{R}, \quad (2)$$

is the leading choice. However, many other choices are possible such as smooth cumulative distribution functions, sine and cosine functions, hyperbolic tangent functions, etc. – see (Kuan, 2008). Furthermore, (McMenamin, 1997) has even proposed to use the π-based activation function, say π^x.

In total, eq. (1) and the relevant specific activation functions constitute the ANN (*k*, *q*) model. Fundamentally, it belongs to the class of nonlinear models and subsumes many other models, well known in the econometric literature – see e.g. (Franses and van Dijk, 2000) – such as the switching-regression model or the Smooth Transition Autoregressive (STAR) model. Better, as noted by (Kock and Teräsvirta, 2011), ANN-model is also a so-called "universal approximator".

The problem will be drastically simplified if ANN (*k*, 1) model is considered:

$$y_t = \boldsymbol{\alpha}'\mathbf{z}_t + \beta G(\boldsymbol{\gamma}'\mathbf{z}_t) + \varepsilon_t, \quad t = 1,\cdots,T, \qquad (3)$$

where $\varepsilon_t \sim N(0,\sigma^2)$ and *G* is the logistic function (2)

THE DATA

The series to be "explained" and to be used for prediction, denoted Elec, is that of the monthly electricity production in South Africa for the period 2002-M1 to 2010-M6. It furnishes the production in thousands of Mwh.
The explicative variables are three in number:

- The consumer prices of services index, denoted CPI, which gives a picture of the prices variation for all urban areas (2012-M12 = 100);
- The total volume of manufactured production, denoted Prod, which constitutes an index with 2010 = 100;
- A binary variable, Rec, which takes the value one during the recessions and the value zero during the expansion phases of the economy. This dummy series was constructed based on the dating of the South African business cycle by (Bismans and Majetti 2012) and (Bismans and Le Roux 2013). The dating of cyclical turning points is obtained by means of the BBQ algorithm. (BBQ is an acronym for Bry-Boschan Quarterly.) Of course, the final series resulting from the algorithm implementation must be monthly.

Therefore, all the series are thus monthly. Precise also that these series – except Rec – have been downloaded from the data base of the South African Reserve Bank.

ESTIMATION OF THE ECONOMETRIC MODEL

The search for an adequate model follows the general-to-specific (GETS) methodology vindicated by David Hendry in several publications. (See for a recent and path-breaking reference, (Hendry and Doornik 2014), especially chapter 1). The essence of this procedure is to begin the discovery process by specifying a general unconstrained model (GUM) with many variables and many lags. It is necessary to reduce the initial model by using significance tests on the coefficients and also the models selection criteria (Akaike, Bayesian Information Criterion and Hannan-Quinn). The elimination of irrelevant variables is carried out sequentially by applying at each stage the criterion of meaning and information.
Following this approach, Table 1 presents the final results of the linear process.

Table 1. The final linear model

	Coefficient	Std. Error	p-value
Elec (-2)	0.754	0.0875	0.000
Elec (-4)	-0.453	0.0834	0.000
Rec	-1608,9	494.1	0.002
Rec (-1)	1654	524.8	0.002
Rec(-6)	1568	297.9	0.000
Prod (-1)	128,9	14.38	0.000

Some comments are due at this point.

Firstly, given the corresponding p-values, all the variables in the final model are significantly different from zero at one percent. Secondly, the variable "Rec" is peculiarly significant. Logically, the entry in recession lowers immediately the electricity output, but after one month and especially six months, the relationship becomes positive. Thirdly, the manufactured production contributes positively, but with only one lag, to the electricity production.
Given Eq. (3), it remains to estimate the part

$$\beta G(\boldsymbol{\gamma}'\mathbf{z}_t) + \varepsilon_t,$$

where $\beta = 1$, G is the logistic Cumulative Distribution Function (CDF) and $\mathbf{z}_t$ is the vector compounded by the variables Elec(-2), Elec(-4), Rec, Rec(-1), Rec(-6) and Prod(-1).
Now, Table 2 shows the ML estimators of the gamma parameters.

Table 2. The ANN-model (non-linear part)

Elec (-2)	Elec (-4)	Rec	Rec (-1)	Rec (-6)	Prod (-1)
150	-151.7	-0.002	-0.0002	0.009	0.04

As a general rule the estimated coefficients must not be interpreted in the same way as we do for a linear model. They are only used to get a better prediction of the electricity consumption.

FORECASTS

The out-of-sample forecasts are presented in the following table for a horizon of six months.

Table 3. Dynamic forecasts

Horizon	Actual values	Linear model	ANN-model
2010-01	20124.4	18900.0	18748.4
2010-02	18861.7	18124.1	17972.6
2010-03	20914.6	17818.5	18590.3
2010-04	19844.7	17782.6	18187.1
2010-05	21149.2	17741.2	19675.9
2010-06	21352.6	18284.8	19660.4

The comparison of point predictions depicts that globally, the ANN-model is a better tool for forecasting in comparison to the linear model.
To refine the analysis, two additional evaluation indicators are computed: the Root Mean Square Error (RMSE) and the Mean Absolute Percentage Error (MAPE). Finally, the Theil's coefficient is equally implemented.

Table 4. Some evaluation statistics

	Linear model	ANN-model
RMSE	2480.2	1605.9
MAPE	10.945	7.526
Theil's U	1.97	1.24

All the statistics deliver the same information: the predictive performance of the neural network model is superior to that of its linear counterparty. However, it is possible to push the interpretation further by considering the coefficient of Theil. Indeed, Theil's *U* takes a value of one with the naïve model. Values lesser than 1 indicate an improvement comparatively to this naïve benchmark and values higher than the unity translate a deterioration of the forecasts, always with respect to this benchmark.
From this viewpoint, both used models - linear and feedforward neural network - do less well than a simple random walk, for which the best prediction of a variable in *t* is the observed value of this variable during the immediately preceding period.
Explaining this apparent paradox is not difficult: it follows from this that the strong seasonality in the series was not taken into consideration. Nevertheless, one must recall that the only objective of this contribution was to compare linear and ANN models from a forecasting perspective.

CONCLUSIONS

Undoubtedly, the study has proved empirically that the ANN-model demonstrated superior predictive properties than the linear one. However, two limitations of the canonical ANN benchmark structure will be exceeded in the future: on one side, considering one node in the hidden layer is a simplification, at best temporary, to be abandoned; on the other side, the seasonality should be modeled explicitly. It's our prospective way!

REFERENCES

Bismans, F. and R. Majetti. 2012. "Dating the South African Business Cycle", *Journal for Development and Leadership*, 1, pp. 1-12.

Bismans, F. and P. Le Roux. 2013. "Dating the Business Cycle in South Africa by Using a Markov-switching Model", *Studies in Economics and Econometrics*, 37, 25-39.

Fine, T.L. 1999. *Feedforward Neural Network Methodology*, New York: Springer-Verlag.

Franses, P.H and D. van Dijk. 2000. *Non-linear Time Series Models in Empirical Finance*, Cambridge: Cambridge University Press.

Hendry, D.F. and J.A. Doornik. 2014. *Empirical Model Discovery and Theory Evaluation*. *Automatic Selection Methods in Econometrics*, Cambridge (MA)-London: The MIT Press.

Kock, A.B. and T. Teräsvirta. 2011. "Forecasting with Nonlinear Time Series Models", *The Oxford Handbook of Economic Forecasting*, Eds M.P. Clements, and D. Hendry, Oxford-New York: Oxford University Press, 61-87.

Kuan, C-M. 2008. Artificial Neural Networks, *The New Palgrave Dictionary of Economics*, 2nd edition, Eds Durlauf, S.N., and L.E. Blume, London-New York, Palgrave Macmillan.

Kuan, C-M. and H. White. 1994. "Artificial Neuronal Networks: An Econometric Perspective", *Econometric Reviews*, 13, 1-91.

McClelland, J.L.; D.E. Rummelhart; and the PDP Research Group. 1986. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol. 2, Cambridge, MA: MIT Press.

McCulloch, W.S. and W. Pitts. 1943. "A Logical Calculus of the Ideas Immanent in Nervous Activity", *Bulletin of Mathematical Biophysics*, 5, 115-133.

McMenamin, J.S. 1997. "What not Pi? A Primer on Neural Networks for Forecasting", *The Journal of Business Forecasting Methods & Systems*, 16 (3), 1-20.

Medeiros, M.C.; T. Teräsvirta; and G. Rech (2006), "Building Neural Networks Models for Time Series: A Statistical Approach", *Journal of Forecasting*, 25, 49-75.

Ozdemir, O.; A. Aslanargun; and S. Asma. 2010. "ANN Forecasting Models for ISE National-100 Index", *Journal of Modern Applied Statistical Methods*, 9, 579-583.

Rech, G. 2002. "Forecasting with Artificial Neural Network Models", Working Paper Series in Economics and Finance, n° 491, Stockholm School of Economics.

Rosenblatt, F. 1958. "The Perceptron: A Probabilistic Model for Information Storage and Organisation in the Brain", *Psychological Reviews*, 62, 386-408.

Rummelhart, D.E.; G.E. Hinton; and the PDP Research Group. 1986. *Parallel Processing: Explorations in the Microstructure of Cognition*, Vol. 1, Cambridge, MA: MIT Press.

Swanson, N.R. and H. White. 1997. "A Model Selection Approach to Real-Time Macroeconomic Forecasting Using Linear Models and Artificial Neural Networks", *Review of Economics and Statistics*, 79, 540-550.

White, H. 2006. "Approximate Nonlinear Forecasting Methods", in *Handbook of Economic Forecasting*, vol.1, Eds Elliot, G., C.W.J. Granger and A. Timmerman, Amsterdam: Elsevier, 459-512.

Zhang, G.; B.E. Patuwo; and M.Y. Hu. 1998. "Forecasting with Artificial Neural Networks: The State of the Art", *International Journal of Forecasting*, 14, 35-62.

AUTHORS BIOGRAPHIES

FRANCIS BISMANS obtains his PhD in economics from the University of Liège, Belgium. Today, he is professor in economics and statistics at the University of Lorraine, France. He published in 2016 the books *Business Cycles in the Run of History,* Springer, New York-Heidelberg, and in French, *Probabilités et statistique inférentielle*, Ellipses, Paris .He is also Research Associate at the Nelson Mandela University, Port Elizabeth, South Africa.
E-mail: Francis.Bismans@univ-lorraine.fr

IGOR LITVINE has obtained a PhD in mathematical statistics from Kiev National Taras Shevchenko University, Ukraine. He also obtained a PhD in economics from the University of Lorraine, France. He is now professor at the Nelson Mandela University and also the director of the Center of Expertise in Forecasting (COEF) in this university. He has numerous publications in the fields of statistics and computer science.
E-mail: Igor.Litvine@mandela.ac.za

OPTIMISATION OF COMPRESSED AIR SYSTEM'S ENERGY USAGE THROUGH DISCRETE EVENT SIMULATION: COMPRESSOR PERFORMANCE

Robbie Mulvany
Alan Arokiam
Abdelhafid Belaidi
Faculty of Engineering & Science
University of Greenwich
Chatham Maritime
Kent
UK

John Ladbrook
Michael Higgins
Dunton Technical Centre
Ford Motor Company
Essex
UK

KEYWORDS
Compressed Air, Discrete Event Simulation, Energy Optimisation, EnergyBlocks, AirBlocks

ABSTRACT
Compressed air systems (CAS) utilised in manufacturing processes require significant energy input for operation. The estimated cost of producing compressed air is considered high with little transparency available when assessing its value in manufacturing. There is currently poor awareness of the performance of CAS in relation to its equipment utilisation and energy optimisation.
This paper presents a modified approach to the EnergyBlocks methodology for representation and simplification of compressed airflow profiles in discrete event simulations (DES). The presented AirBlocks methodology significantly reduces the aggregate data required to represent the dynamic and interdependent nature of CAS. Combining the AirBlocks approach with manufacturing throughput productivity simulations allow a productivity oriented compressed air demand profile to be developed. This offers the capacity to estimate periods of sustained peak, average and minimum air demand, incidents of production stoppages due to air starvation and, identify waste and saving potential in the system. This paper includes an industrial case study where the AirBlocks approach was used in evaluating the performance of an existing CAS. Through simulation - poor compressor utilisation and regular incidents of air starvation were identified as symptoms of insufficient CAS volumetric capacity and an oversized compressor system in an automotive engine manufacturing plant.

INTRODUCTION

It is estimated that global primary energy demand will increase by 40% between 2007 and 2030 with industrial demand projected to grow most rapidly (OECD 2009). During this period industry is expected to account for 20% of the total world energy demand with electrical energy making the largest single contribution. The electrical energy required to meet the demands of CAS in industry accounted for between 10% to 30% of overall industrial energy consumption (Radgen and Blaustein 2001). In 2001 it was estimated that this energy demand required 80TWh of electricity and produced 55 million tonnes of CO2 in the European Union (EU-15) alone (Radgen and Blaustein 2001). Considering between 20-25% of input electrical energy is delivered as usable compressed air energy (Kreith 2000) and 60% or less of air consumed actually make a direct contribution to the product or service which it was intended (Foss 2002), CAS are one of the most expensive in terms of energy utilisation. Figure 1 shows the cost of energy delivery (US$/gigajoule) in comparison to natural gas, steam and electricity.

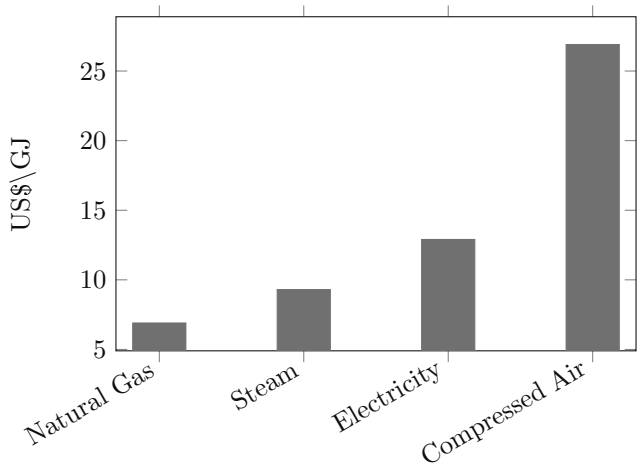


Figure 1: Cost of energy delivery (Yuan et al. 2006)

The energy cost in operating a CAS is undoubtedly high. A compressors average lifespan is 13 and 16 years for 10-110kW and 110-300kW compressors respectively (Radgen and Blaustein 2001). During this time the energy cost will reach up to 78% of the total set up and running cost of the system (Saidur et al. 2010). Figure 2 shows

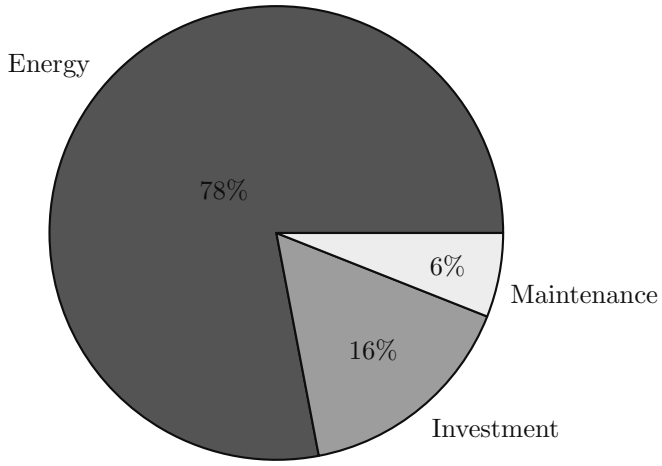


Figure 2: Life time cost of a compressed air system (Saidur et al. 2010)

the proportions of cost attributed to each key factor in procuring and operating a CAS.

CAS energy consumption can be further exacerbated through poor practices, poor understanding and the prioritisation of production reliability over system efficiency and energy value. It is reported in literature that compressed air has the perception of being free to the user. Reasoning proposed by McKane and Medaris (2003) suggests that within the manufacturing industry there is a common disconnection between the known presence of compressed air in the distribution network and the associated electrical energy cost required to create this presence. Furthermore, Saidur et al. (2010) states; the only time that issues such as leakage and filter replacement get attention is when air and pressure loss begin to interfere with regular production.

This viewpoint was given additional insight from interviews with 19 European enterprises which found system reliability to be the most important performance criterion as the cost of a CAS breakdown inevitably leads to lost production. Cost was rated as the least important performance criterion for CAS (Radgen and Blaustein 2001). However the emphasis which was placed on reliability was not realised in practice as a report by US Department of Environment (US DOE 2001) found 35% of interviewed end users experienced unscheduled shutdowns with 60% of these shutdowns lasting for 2 days or more.

It is widely recognised in literature that there are vast improvements to be gained in CAS efficiency with much attention given to technical approaches for improving pneumatic component performance, compressor control strategies, systems maintenance, procurement, system and plant design (Saidur et al. 2010). It is however alleged that the implementation of technical measures designed to improve energy efficiencies in the industrial environment are very low (Radgen & Blaustein 2001), a point which has been reiterated by Galitsky and Worrell (2003), Rohdin and Thollander (2006), O'Driscoll and O'Donnell (2012) and Fleiter et al. (2012). Reasons for this lack of uptake are primarily seen as organisational, where a combination of cost accountability and awareness of potential savings are invisible to key decision makers. Failure to highlighting specific performance indicators required to promote energy efficiency in a CAS are largely due to the amalgamation of all electrical energy consumption costs as a single general overhead (Marshall 2013). Furthermore, difficulties in implementing improvements to CAS efficiency can arise from complex management structure within an organisation. Such structures can see responsibility and prioritisation diluted where potential measures for system improvements must pass through different departments with differing functions, e.g. finance, maintenance, procurement (Radgen and Blaustein 2001).

It is estimated that over 50% of industrial plant air systems have potential for large energy saving projects with relatively low project costs US DOE (2001). The perceived lack of focus on implementation measures within the literature, coupled with poor organisational accountability in industry - support the case for the need of a comprehensive data driven approach to improving compressed air systems energy efficiency which targets the decision making process.

The aim of this paper is to present a modified approach to the EnergyBlocks methodology to be used for the refinement and representation of compressed airflow profiles in DES. The presented AirBlocks methodology creates an environment in which a CAS performance can be evaluated with regard to its total manufacturing production demand to identify potential for improving energy utilisation. Evaluation can be carried out on multiple levels of aggregation including individual machine components, single machines, multi-machine operations and a complete manufacturing line. The paper briefly reviews recent and seminal literary contributions which highlight the notable developments in modelling and simulation of energy flows in manufacturing systems.

MODELLING & SIMULATION

The implementation of CAS energy efficiency measures in the manufacturing industry have remained low despite continuous developments on the topic. Current decision mechanisms are largely failing to recognise the substantial technical and economic potential for energy saving. The biggest influencing factor in decision making is that of financial savings. Such decisions must be informed with detailed knowledge of the systems current performance (Talbott 1992). But due to the contextually specific nature of most existing data/information it is difficult to relate it to individual situations faced by CAS users.

Information tools are highlighted by Radgen and Blaustein (2001) as a platform to transcend organisa-

tional structural barriers where responsibility for differing aspects of CAS are spread through company departments and levels. Simulation is a modelling tool used to understand system performance and identify potential for improvements (Tako and Robinson 2010). Where an investigation seeks to understand system performance over time and identify potential improvements, simulation can offer the following advantages (Panneerselvam 2006);

- Experimentation times can be compressed.
- Performance can be studied under multiple scenarios.
- Success or failure on a simulated system has no adverse effects on production.

Drawbacks to simulation modelling are the time and cost required to create and verify effective and system comparable models (Wilson et al. 2015).
Thollander et al. (2009) combined energy auditing with production optimisation and simulation as a means to inform strategic investment in a Swedish foundry. Generating performance data which directly correlated to production for assumed future energy cost variations gave decision makers the relevant insight to system performance in the appropriate context to make informed investment decisions.
Maxwell and Rivera (2003) propose the use of a dynamic system simulation to investigate the effects of air pressure on the performance of CAS. The systems approach taken by Maxwell and Rivera (2003) addresses a core element of CAS the interdependent and dynamic relationship between supply and demand. However, the deterministic approach in modelling compressed air demand gives a narrow view of the effects of production variability. This approach does not assess the effect of uncertainty in production such as random occurrences of peak demand and production interruptions. Furthermore the low resolution offered by a fixed 10 second time interval and short (24hr) production period of the demand profile fails to consider the longer term variability in production output and possible instantaneous occurrences of peak air demand.
The integration of production planning and energy performance analysis using DES is a topic which has recently been the focus of some attention. A study by Berglund et al. (2011) highlights the broad separation of plant information systems and production data in manufacturing. The author proposes an integrated approach to handling production and facility energy consumption which evaluates the combined impact of process energy from manufacturing operations and resources, facility energy and building services using DES (Figure 3). While the full range of energy systems proposed by Berglund et al. (2011) is beyond the scope of this study it does highlight the necessity for any developments with respect to DES of CAS to be considered in the context of a more comprehensive approach for integrating energy flow into a production planning environment.
Herrmann et al. (2011) states that dynamic interactions of processes and auxiliary equipment must be considered when planning and controlling manufacturing systems. The author highlights the estimation of time based energy consumption to be of major importance if factory systems are to be considered as a collective. While no single comericially available software yet support such an analysis the author presents three general simulation paradigms which are most commonly pursued for energy oriented manufacturing system simulation (Figure 4).
Herrmann et al. (2011) proposed a generic simulation environment which integrates elements of Paradigm B

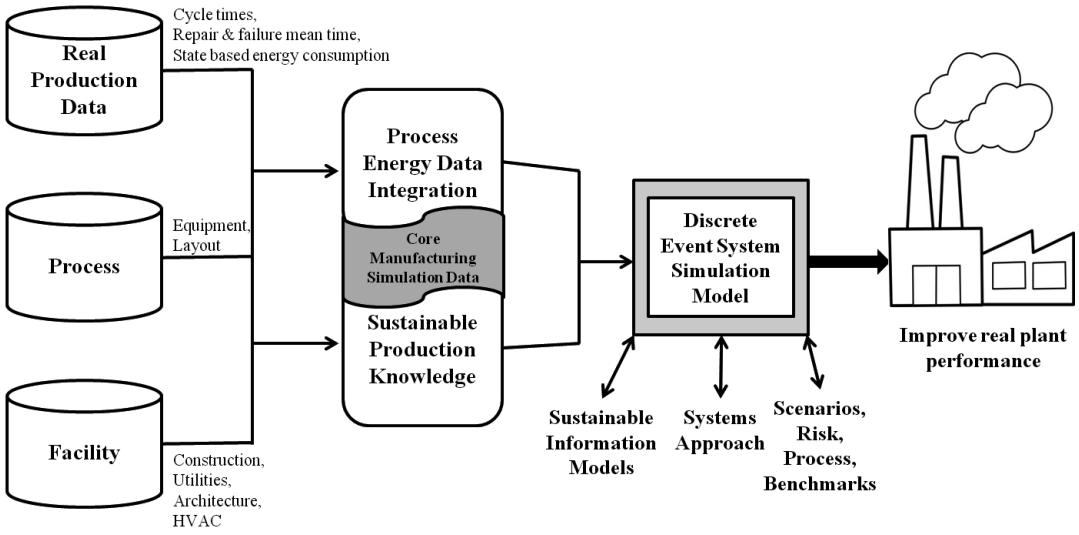


Figure 3: Integration of production & facility energy consumption (Berglund et al. 2011)

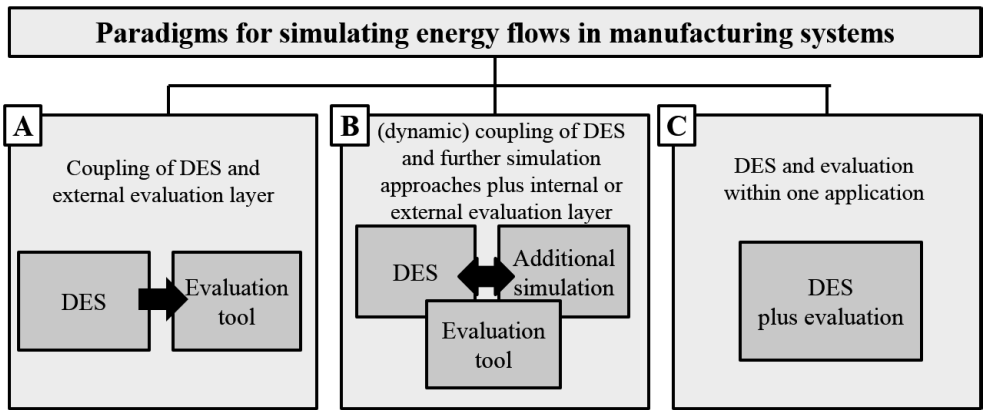


Figure 4: Energy flow simulation paradigms for manufacturing (Herrmann et al. 2011)

and C. This approach offers a platform for a comprehensive analysis of energy system performance with respect to production requirements but is not without drawbacks. The combination of discrete event and discrete time simulation implicitly contains fixed time interval, which if too short will unnecessarily increase demand on computational resources. Subsequent computations may occur where the system remains unchanged offering no benefit, conversely if the time intervals are too large, important events may lack sufficient detail and remain overlooked (Carter and Price 2000).
Discrete event and discrete time simulations have extensive fundamental differences in discretization algorithms, reference locality, sequencing / scheduling and data structures (Nutaro 2007). Modellers would still require a understanding of the differences in simulation type, methods and nature of the model (stochastic & deterministic) (Tako and Robinson 2010).

PROPOSED SOLUTION

Approach

The conservation of mass approach is the adopted method for this study due to its conceptual simplicity and suitability to a fixed volume system analysis with negligible heat transfer. The conservation of mass is described as the time rate of accumulation of mass within a control volume equals the difference between the total rates of mass flow in and out across the boundary (Bejan et al. 1996). As this study is assessing the relation of demand, system capacity (receiver & distribution network volume combined) and supply in an open system the control volume is identified as the CAS capacity with supply and demand dictating the flow of mass across the systems boundaries (Cengel and Boles 2013). The fundamental equations that govern this approach are;

$$\frac{dm_{cv}}{dt} = \sum_{i}(\dot{m}_i) - \sum_{o}(\dot{m}_o) \tag{1}$$

where:
$\dot{m}_i$ = input mass flow rate
$\dot{m}_o$ = output mass flow rate
m_{cv} = the total mass contained within the control volume at time t,
To approximate the appropriate system capacity and any occurrences of starvation the mass within the total systems volume at the upper pressure threshold P_1 and the lower pressure threshold P_2 must be determined. Assuming ideal gas behaviour and negligible changes in air temperature, the conservation of mass equation can be expressed as;

$$\frac{(dP_s)}{dt} = \frac{(\dot{m}_i - \dot{m}_o) * R_g * T}{V_{sv}} \tag{2}$$

where:
P_s = the system pressure
R_g = ideal gas constant for air
T = average temperature in the receiver
V_{sv} = combined receiver and network volume.
The required volume of the combined receiver and network volume at the upper pressure threshold and the lower pressure threshold must be equal, thus the volume is approximated as follows;

$$V_{sv} = \frac{m_{mr} * Rg * T}{(P_1 - P_2)} \tag{3}$$

where:
m_{mr} = the maximum mass reduction in a single simulation period.
The combined receiver and network volumes are considered as a single quantity as this allows a balance between network and receiver capacity to be determined during the design process. In order to determine the appropriate compressor output and system capacity using this model, a direct deterministic multivariate optimisation method is employed. This method is referred to as a univariate search where all but one variable is fixed allowing the local optimum value to be found. This variable is in turn fixed and another variable is optimised with the process repeated until there are no further improvements in the objective function (Smith 2016). The

resultant data set can be used to approximate the global optimum for the decision making process. When determining the local optimum value the interval halving method is employed as described by Singiresu (2009). The convergence criterion is - the minimum final interval within a tolerance range where there are no occurrences of starvation.

Methodology

The methodology for describing compressed air consumption of production equipment is a modification of the EnergyBlocks planning methodology presented by Weinert et al. (2011). The authors present an approach based on segmenting the power profiles of production equipment in their operating state (see Figure 5).

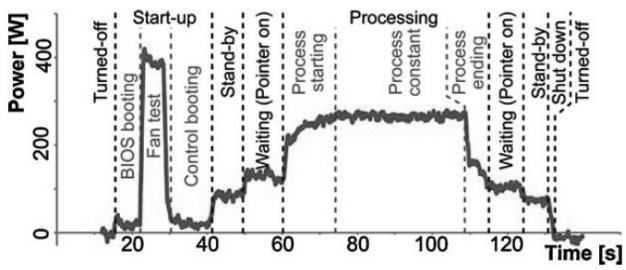


Figure 5: Energyblocks - partition of a power profile (Weinert et al. 2011)

With this approach a production process is modelled as a sequence of EnergyBlocks which represent the whole production process chain. Furthermore a database of EnergyBlocks allow any process chain to be modelled which contains similar process parameters and machine specifications. This approach is modified for the refinement of compressed air profiles for use in DES. Two primary justifications for modifying this process are; (1) a machines compressed air profile may constitute more than one superimposed processes for example, if a machine demonstrated a downstream leakage characteristic when no machining activities are occurring then the logical assumption can be made that the leakage is symptomatic of the machines performance. Thus, any additional air demand processes will result in the consolidation of both contributing factors. (2) The second justification is in reducing the number of interpreted events in the process by defining underlying behaviours, each contributing event can operate concurrently as opposed to sequentially. For the sake of distinction the modified approach is referred to as AirBlocks and a comparative of each approach is made between Figures 6 & 7 using a minimum quality lubrication (MQL) airflow profile for a machine tool.

The AirBlocks approach offers a significant reduction in the datapoints required to represent the same number of processes, the number of events required to represent the cycle is reduced by almost half offering improved computational efficiency in simulation. In addition, underlying contributers such as leakage are more accurately represented allowing greater analysis of any specific individual elements contribution to the production demand.

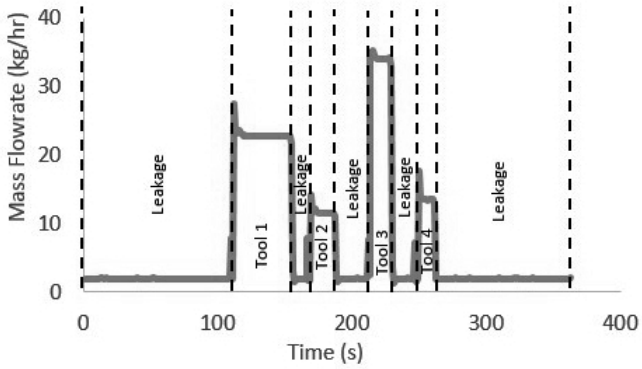


Figure 6: Energyblocks approach

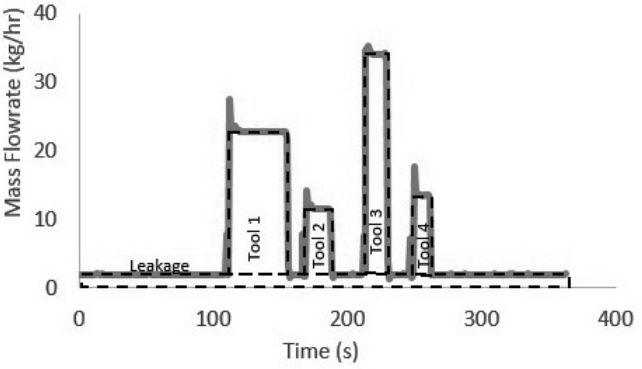


Figure 7: Airblocks approach

Simulation

The simulation tools used by the industrial partner for this study are Lanners commercial DES software Witness® linked to bespoke front end within Microsoft® Excel®, further details of which are available in Wilson et al. (2015) & Higgins (2013). This research aims to develop a universally applicable approach to simulating compressed air system performance with regard to its manufacturing production demands a database based method was used to generate and drive CAS performance simulations using DES throughput and productivity data and machine level compressed air demand profiles. This topic has seen development from Randell and Bolmsjo (2001) where a proof of concept for the integration of information between platforms was presented. Interfacing DES software and relational database management systems (RDBMS) is supported by many priority vendors with the most appropriate example presented by Waller (2012). A MySQL RDBMS was selected as the most applicable platform to interface with Witness. This study was limited to an integration of information only although further systems integration is planned.

The basic structure of the database simulator can be categorised as Paradigm B as shown in Figure 4. This structure offers greater functionality where a high level of detail is required when analysing air flows and dynamic interactions between subsystems. Although the level of modelling and simulation complexity is increased and differing model aspects are spread across multiple expert tools, it has been observed that the integration of simulation tools, data storage infrastructure and evaluation tools is necessary to efficiently manage and utilise data generated in a manufacturing environment. Studies carried out by Skoogh and Johansson (2008), Randell and Bolmsjo (2001), Sargent (2003), Weinert et al. (2011), Herrmann and Thiede (2009) and Lind et al. (2009) offer justification and insight to the integration of a database into production, supply chain and energy simulations. A drawback to the multiple platforms required in this simulation approach expressed by Herrmann et al., (2011) is reduced transferability. As this approach is reliant on a fundamental aspect of DES the event set - it could be argued that provided suitable input data can be generated from any alternative DES platforms, this approach is transferable. Furthermore, the selection of RDBMS can increase the platforms transferability if selected appropriately. The simulation structure is presented in Figure 8.

RESULTS & DISCUSSION

The simulation results offered insight into the performance of the existing compressor and capacity arrangement. By comparing simulation results to observed behaviours in the CAS an evaluation of the cause of the incidences of starvation was able to be made.

The existing compressed air supply consisted of approx. 100 m^3/hr (STP Standard Temperature & Pressure) compressors and a total network volumetric capacity of 1.5m^3. Compressor utilisation is approximately 65% (Figure 9) which corresponds with a value of 60% estimated from observation. Charting the existing volumetric capacity against the compressor output shows the initial capacity to be in close proximity to the threshold where occurrences of starvation are likely (Figure 10). Considering the simulation result is derived from a simulated compressor supply with 100% availability, it does not account for the impact a control strategy would have. Thus it is reasonable to assume that the actual threshold for starvation would in reality occur at a marginally increased volumetric capacity but utilisation would be expected to remain the same as it is a function of the compressors output only.

Allowing for the impact of a control strategy, this result was consistent with the behaviour observed in the CAS during operation thus it allowed an informed course of

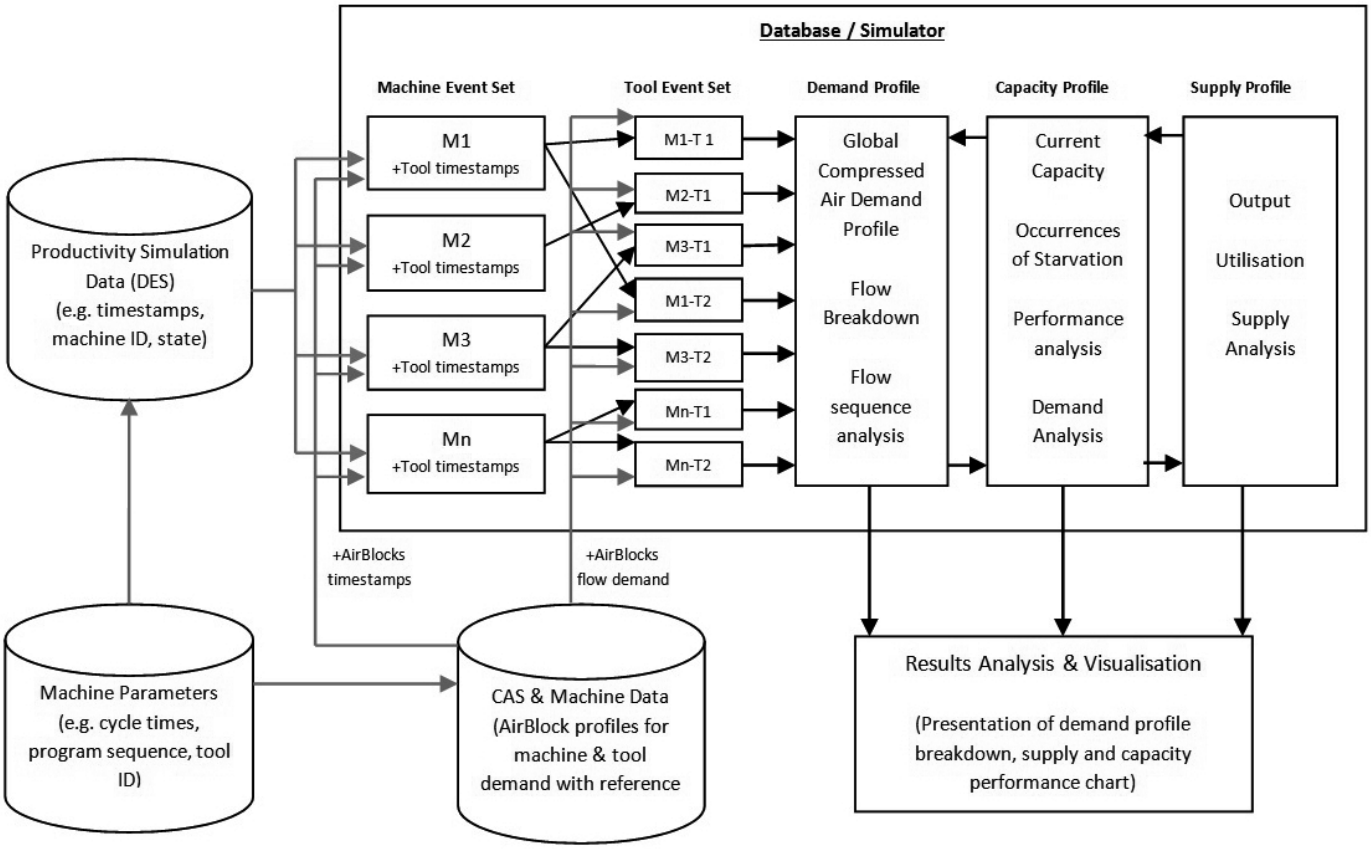


Figure 8: Simulation structure of production oriented CAS simulation

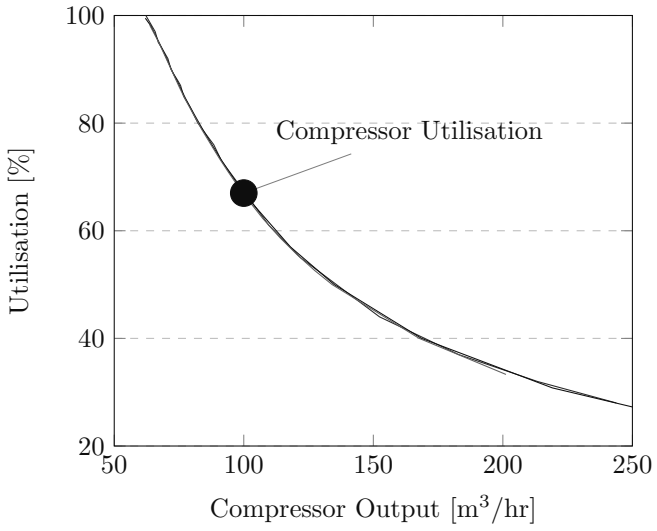


Figure 9: Compressor Utilisation

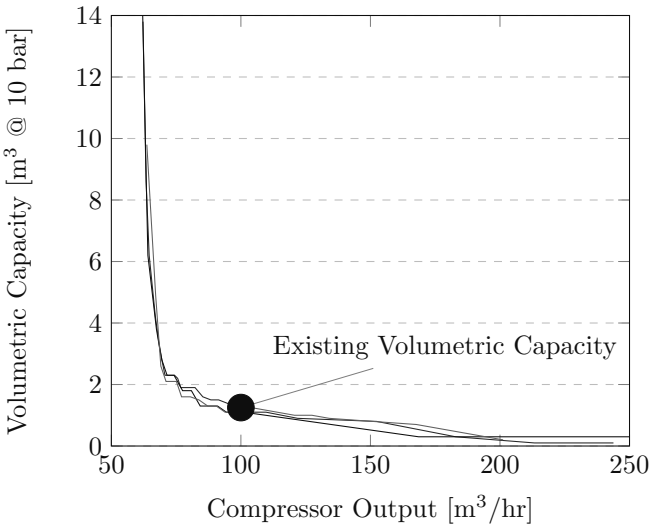


Figure 10: Network Volumetric Capacity / Compressor Output

action to be taken. To overcome the likelihood of further production interruptions due to starvation, the volumetric capacity was increased to 2.5 m^3.

Such insight was not previously available to the industrial partner. Its impact will have long-term positive consequences to the machine availability and therefore productivity seen within the manufacturing facility.

CONCLUSION & OUTLOOK

This paper presented the AirBlocks methodology for representation and simplification of compressed airflow profiles in DES. By combining the AirBlocks approach with manufacturing throughput productivity simulations, a compressed air demand profile was able to be produced which was representative of real world demand on CAS in discrete maunfacturing. This approach allowed the variablity in discrete manufacturing to be accounted for while the reduction in aggregate data enabled long periods of production simulation to be carried out.

The case study presented an application of the AirBlocks methodology as a means of evaluating a CAS by assessing the performance of an air compressor and receiver to a simulated air demand. Within the broader aim of the research study the simulation outcome was aimed to assist decision makers in the design, procurement and implementation process of CAS design and operation. The case study was successful in both its specific aim (compressor performance analysis) and within the broader aim of this study. However the approach is not without weakness. Currently no compressor control strategy was accounted for in the model which would invariably affect the simulation outcome in terms of the starvation threshold and required compressor output. If a control strategy activation range was considered the effect of a narrow activation and deactivation pressure range would increase the required compressor output to meet the demand and the capacity would also be expected to increase. Thus the current simulation results must be considered with a large safety factor.

The novelty aspects of the research are found in the following;

- The simulation method proposed offers a novel approach to simulating a CAS supply, capacity and demand relationships whilst retaining the discrete event simulation paradigm characteristics of event sets and sequential data processing.
- The use of a database simulator builds on previous research by applying it to a CAS data analysis task.
- The AirBlocks method of data simplification is a novel approach to transferring dynamic system data to discrete event data.

Further research in this field will address the impact of a compressor control strategy on a CAS supply system and explore energy reduction measures for CAS.

REFERENCES

Bejan A.; Tsatsaronis G.; and Moran M., 1996. *Thermal Design and Optimization*. John Wiley & Sons Inc.

Berglund J.; Michaloski J.; Leong S.; Shao G.; Riddick F.; Arinez J.; and Biller S., 2011. *Energy Efficiency Analysis for a Casting Production Systems. 2011 Winter Simulation Conference*, 1060–1071.

Carter M.W. and Price C.C., 2000. *Operations Research: A Practical Introduction*. CRC Press LLC, Boca Raton, Florida.

Cengel Y.A. and Boles M.A., 2013. *Thermodynamics: An Engineering Approach*, vol. 53. McGraw

Hill, 5th ed. ISBN 9788578110796. doi:10.1017/CBO9781107415324.004.

Fleiter T.; Hirzel S.; and Worrell E., 2012. *The characteristics of energy-efficiency measures a neglected dimension. Energy Policy*, 51, 502–513. ISSN 03014215. doi:10.1016/j.enpol.2012.08.054.

Foss R., 2002. *Managing Compressed Air Energy Part I: Demand Side Issues.* URL `http://www.maintenancetechnology.com/2002/09`.

Galitsky C. and Worrell E., 2003. *Energy Efficiency Improvement and Cost Saving Opportunities for the Vehicle Assembly Industry: A Guide for Energy and Plant Managers.* Tech. Rep. January, U.S. Environmental Protection Agency.

Herrmann C.; Thiede S.; Kara S.; and Hesselbach J., 2011. *Energy oriented simulation of manufacturing systems Concept and application. CIRP Annals - Manufacturing Technology*, 60, no. 1, 45–48. ISSN 00078506. doi:10.1016/j.cirp.2011.03.127.

Higgins M., 2013. *Fitness of Simulation within the Automotive Industry.* Ph.D. thesis, Cardiff University.

Kreith F. (Ed.), 2000. *The CRC Handbook of Thermal Engineering.* CRC Press LLC, Boca Raton, Florida.

Marshall R., 2013. *Using Kpi ' S for Peak Efficiency. Compressed Air Best Practices*, 38–42. URL `https://www.compressedairchallenge.org/library/articles/2013-07-CABP.pdf`.

Maxwell G. and Rivera P., 2003. *Dynamic Simulation of Compressed Air Systems.* In *2003 ACEEE Summer Study on Energy Efficiency in Industry, Conference Proceedings 3*. 146–156.

McKane A. and Medaris B., 2003. *The Compressed Air Challenge: Making a Difference for US Industry.* In *Energy Efficiency in Motor Driven Systems*, Springer - Verlag, New York. 34 – 40.

Nutaro J., 2007. *Discrete event simulation of continuous systems.* In *Handbook of Dynamic Systems Modeling*, ... 1–23.

O'Driscoll E. and O'Donnell G.E., 2012. *Industrial power and energy metering a state-of-the-art review. Journal of Cleaner Production*, 41, 53–64. ISSN 09596526. doi:10.1016/j.jclepro.2012.09.046.

OECD, 2009. *World Energy Outlook 2009.* Tech. Rep. 4, International Energy Agency. doi:10.1049/ep.1977.0180.

Panneerselvam R., 2006. *Operations Research.* PHI Learning Private Limited, New Delhi, 2nd ed.

Radgen P. and Blaustein E., 2001. *Compressed Air Systems in the European Union.* Tech. rep., ADEME, Fraunhofer ISI, DoE, ECE, Stuttgart.

Randell L. and Bolmsjo G., 2001. *Database driven factory simulation: a proof-of-concept demonstrator.* In B.A. Peters; J.S.Smith; D.J. Medeiros; and M.W. Rohrer (Eds.), *Proceeding of the 2001 Winter Simulation Conference*. vol. 2. ISBN 0-7803-7307-3. ISSN 02750708, 977–983. doi:10.1109/WSC.2001.977402.

Rohdin P. and Thollander P., 2006. *Barriers to and driving forces for energy efficiency in the non-energy intensive manufacturing industries in Sweden. Energy*, 31, no. 12, 1836–1844.

Saidur R.; Rahim N.; and Hasanuzzaman M., 2010. *A review on compressed-air energy use and energy savings. Renewable and Sustainable Energy Reviews*, 14, no. 4, 1135–1153. ISSN 13640321. doi:10.1016/j.rser.2009.11.013.

Singiresu S.R., 2009. *Engineering Optimization: Theory and Practice.* John Wiley & Sons Inc., 4th ed.

Smith R., 2016. *Chemical Process Design and Integration.* John Wiley & Sons Ltd, 2nd ed.

Tako A.a. and Robinson S., 2010. *Model development in discrete-event simulation and system dynamics: An empirical study of expert modellers. European Journal of Operational Research*, 207, no. 2, 784–794. ISSN 03772217. doi:10.1016/j.ejor.2010.05.011.

Talbott E.M., 1992. *Compressed Air Systems : A Guidebook on Energy and Cost Savings.* The Fairmont Press, Inc., Lilburn. CA, 2nd ed.

Thollander P.; Mardan N.; and Karlsson M., 2009. *Optimization as investment decision support in a Swedish medium-sized iron foundry A move beyond traditional energy auditing. Applied Energy*, 86, no. 4, 433–440. ISSN 03062619. doi:10.1016/j.apenergy.2008.08.012.

US DOE, 2001. *Assessment of the market for compressed air efficiency services.*

Waller A., 2012. *Proceedings of the 2012 Winter Simulation Conference.* In C. Laroque; J. Himmelspach; R. Pasupathy; O. Rose; ; and A. Uhrmacher (Eds.), *WITNESS Simulation Software.* IEEE. ISBN 9781467347815.

Weinert N.; Chiotellis S.; and Seliger G., 2011. *Methodology for planning and operating energy-efficient production systems. CIRP Annals - Manufacturing Technology*, 60, no. 1, 41–44. ISSN 00078506. doi:10.1016/j.cirp.2011.03.015.

Wilson J.; Arokiam A.; Belaidi H.; and Ladbrook J., 2015. *A simple energy usage toolkit from manufacturing simulation data. Journal of Cleaner Production*, 122, 266–276. ISSN 09596526. doi:10.1016/j.jclepro.2015.11.071.

Yuan C.Y.; Zhang T.; Rangarajan A.; Dornfeld D.; Ziemba B.; and Whitbeck R., 2006. *A Decision-Based Analysis of Compressed Air Usage Patterns in Automotive Manufacturing. Journal of Manufacturing Systems*, 25, no. 4, 293–300.

INVESTIGATION OF THE MODELLING EFFECTS ON THE STEAM GENERATOR'S BEHAVIOUR DURING THE EARLY STAGES OF A STATION BLACKOUT IN A CANDU 6 REACTOR

Roxana-Mihaela Nistor-Vlad
Daniel Dupleac
Ilie Prisecaru
Politehnica University of Bucharest
Splaiul Independenței 313, Sector 6
060042, Bucharest, Romania
roxanamihaelanistorvlad@gmail.com,
danieldu@cne.pub.ro, prisec@gmail.com

Chris Allison
Innovative Systems Software
3585 Briar Creek Ln
Ammon, ID 83406, USA
iss@srv.net

KEYWORDS

Nuclear engineering, model analysis, sensitivity analysis, deterministic, system analysis.

ABSTRACT

RELAP/SCDAPSIM is a system tool that allows the used to model reactor systems and analyze a wide range of transients. It was initially designed to analyze LWRs, but starting 2009 Politehnica University of Bucharest worked intensively to demonstrate that the code could also predict the behaviour of a CANDU reactor.

This paper shows the evolution of a Station Blackout accident in a CANDU 6 reactor using RELAP/SCDAPSIM, with variations in modelling the steam generator of a PHWR, and also listed the parameter that influence the dryout phenomena in the steam generators secondary side. The analysis performed has shown no major effects of the steam generators modelling on the parameters selected to be highlighted, and the progression of the accident has shown only minor differences in the events timing.

INTRODUCTION

A Station Blackout scenario considered in this analysis is a transient initiated by the loss of AC power (Class IV) coincident with the loss of all on site standby (Class III) and emergency electric power supplies.

The steam generators play an essential safety function during the SBO accident scenario, serving as the primary function of removing the heat from the primary system by transferring the heat from the primary coolant (D_2O) circulated through the "U" tubes to the secondary side containing H_2O as a coolant. After the reactor trip, the decay heat from the fission products consists of about 7% of full power heat production. The most common method of removing the decay heat is through the steam generators. The heat removal function will be ensured by maintaining a minimum water inventory in the steam generators (Chaplin 2016).

In this analysis the RELAP/SCDAPSIM code was used to develop the full plant model. The main purpose of this paper is to observe the impact of the steam generators modelling on the behaviour of the plant during the early stages of a SBO transient. For this purpose four different configurations of the steam generators were analyzed.

SBO ANALYSIS IN CANDU 6 REACTORS

RELAP/SCDAPSIM, designed to predict the behavior of reactor systems during normal and accident conditions, is being developed at Innovative Systems Software (ISS) as part of the international SCDAP Development and Training Program (SDTP). RELAP/SCDAPSIM uses the publically available SCDAP/RELAP5 (Siefken et al. 1996) models developed by the US Nuclear Regulatory Commission (NRC) in combination with proprietary (a) advanced programming and numerical methods, (b) user options, and (c) models developed by ISS and other STDP members (Innovative Systems Software 2017; Allison and Hohorst 2008).

RELAP/SCDAPSIM is designed to describe the overall reactor coolant system (RCS) thermal hydraulic response and core behavior under normal operating conditions or under design basis or severe accident conditions. The RELAP5 models calculate the overall RCS thermal hydraulic response, control system behavior, reactor kinetics, and the behavior of special reactor system components such as valves and pumps. The SCDAP models calculate the behavior of the core and vessel structures under normal and accident conditions. The SCDAP portion of the code includes user-selectable reactor component models for LWR fuel rods, Ag-In-Cd and B_4C control rods, BWR control blade/channel boxes, electrically heated fuel rod simulators, and general core and vessel structures. The models calculate the damage progression in the reactor core: heat-up, oxidation and meltdown of fuel rods and control rods, ballooning and rupture of fuel rod cladding, release of fission products from fuel rods and disintegration of fuel rods into porous debris and molten material. The SCDAP portion of the code also includes models to treat the later stages of a severe accident including debris and molten pool formation, debris/vessel interactions, and the structural failure (creep rupture) of vessel structures. The

latter models are automatically invoked by the code as the damage in the core and vessel progresses (Allison and Hohorst 2008) .

The full plant model (as shown in Figure 1) used in this analysis is a detailed model of the core (16 thermal hydraulic channels, each of them describing the behaviour of a group of specific fuel channels, based on the radial power distribution in the core); the fuel assemblies along with the pressure tube, annular CO_2 gas, and the calandria tube were modelled using SCDAP components.

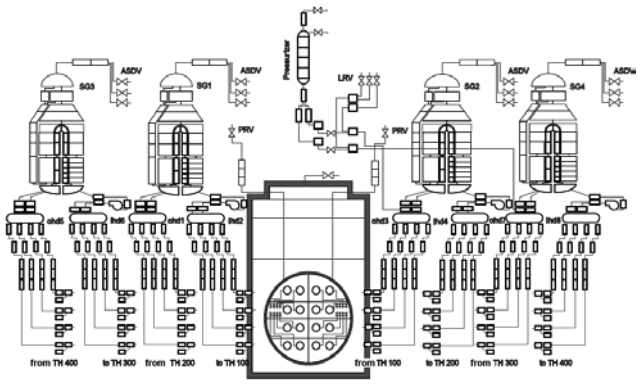


Figure 1: CANDU 6 full plant model

During the past 15 years, Politehnica University of Bucharest has developed a full plant model of a generic CANDU reactor, with slight variations up to the present, regarding the number of the thermal hydraulic fuel channels describing the core, fuel bundles modelling (using RELAP heat structures or SCDAP components), adding more and more features to better represent the behaviour of the plant during several accident conditions.

SG MODEL IN RELAP/SCDAPSIM

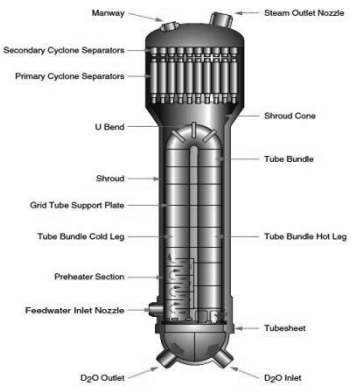


Figure 2: Steam generator in a CANDU reactor (Chaplin 2016)

Since the SBO scenario has been analyzed several times before (Dupleac et al. 2009; Dinca et al. 2015; Zhou and Novog 2017) employing RELAP/SCDAPSIM, this analysis will focus on the influence of different modelling configurations on the dryout of the steam generator, void fraction in the secondary side of the SGs, and also in the U tubes, natural circulation mass flow in the PHTS, pressure in the SGs during the early stages of a SBO in a CANDU 6 reactor.

The main components of the steam generators modelled in RELAP/SCDAPSIM are: the hot and cold legs of the "U" tubes, the primary coolant inlet and outlet, the preheater section, the cyclone separator, the steam dome, the downcomer, the riser, and the shrouds. Most of the steam generators components were modelled using pipe components, single volumes, a separator, and single junctions describing the connections between the components. The steam generator of a CANDU 6 reactor is shown in Figure 2.

Steam Generator Modelling - Case 1

First case which was considered is the one used in the publication reports on the results of an IAEA coordinated research project (CRP) on benchmarking severe accident computer codes for heavy water reactor applications (IAEA 2013).

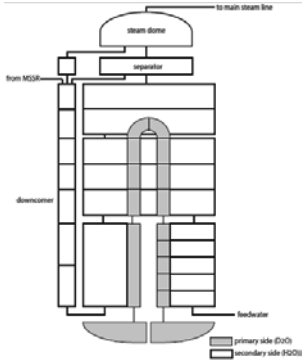


Figure 3: SG Base Case Nodalization Scheme

Figure 3 shows the nodalization scheme for the steam generator. The "U" tubes are discretized in 15 volumes, and separate volumes (single volumes) for the inlet and outlet plenum. The secondary side consists of the preheater, riser, separator, drum and downcomer volumes discretized in different number of nodes to properly model the steam generator performances. The feed water is injected in the preheater, a time-dependent volume gives the temperature and a time-dependent junction is used as a boundary condition for the feed water flow rate.

The pipe component is simply a series combination of single-volume and single-junction components. The single-junction is the basic hydrodynamic flow unit in RELAP5. The input data specifications describing the basic junction properties and conditions for the junctions associated with other types of components (pipes, branches, etc.) are identical to those described for the single-junction component. Pipe components offer input conveniences, since most characteristics of the volumes and junctions in a pipe are similar or change infrequently along the pipe, and input data requirements can be reduced accordingly.

Because of the sequential connection of the volumes, junctions are generated automatically rather than being individually described.

Steam Generator Modelling - Case 2

The second case starts from the previous model used in the first case analysis, with the variation of the number of volumes of the steam generator secondary side components (Figure 4) from 3 volumes of the riser to 9 volumes, and the preheater section from one volume to 5 volumes. The primary side of the steam generator was also divided in multiple nodes corresponding to the detailed nodding used on the secondary side.

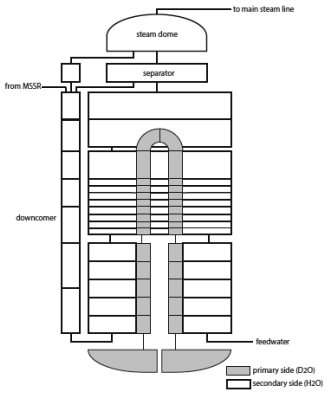


Figure 4: SG Detailed Nodding Scheme 1

Steam Generator Modelling - Case 3

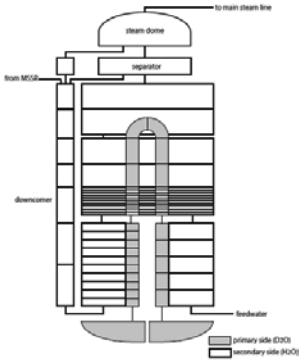


Figure 5: SG Detailed Nodding Scheme 2

The third case starts from the model used in the base case analysis, with the variation of the number of volumes used to describe the steam generator secondary side components (Figure 5), from 3 volumes of the riser to 12 volumes, and the preheater section from one volume to 10 volumes. The primary side of the steam generator has been modelled based on the nodalization used for the secondary side.

Steam Generator Modelling - Case 4

The fourth nodalization proposed for this study consists in a single volume describing the secondary side of the steam generator, the downcomer's nodding being maintained as it was in the previous proposed nodding schemes.

The single-volume component is the basic hydrodynamic cell unit in RELAP5. Note that the pipe component may be thought of simply as a series collection of single-volumes joined by single-junctions. The input data specifications describing the basic volume geometries and conditions for the other types of components (pipes, branches, etc.) are identical to those for the single volume component. The flow area, length, and volume of the cell must be input. These three parameters must be consistent or an input error results. Thus, it is recommended that one of these three quantities be input as zero, allowing the code to calculate its value consistent with the two nonzero entries. For complex geometries, the requirement that the area, length, and volume be consistent may require the modeler to accept a compromise on one or more of the input parameters. This situation arises when the modeler attempts to include a region with a varying flow area varies within a single hydrodynamic cell. A compromise is needed because the average flow area for the geometry may not adequately represent the flow path in the region. The input flow area determines the flow velocity, the input length affects the calculated frictional pressure drop, and the input volume contributes to the overall fluid system volume. An additional constraint is that the length input for a vertical cell must be enveloped by the elevation gain of the cell. The modeler should select the compromise that would least affect his particular problem.

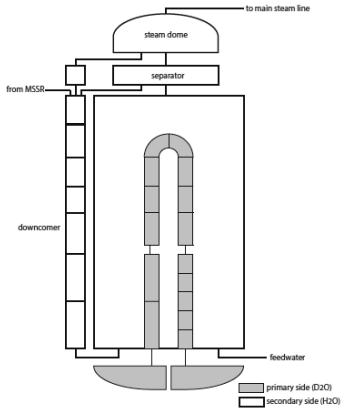


Figure 6: SG Simplified Nodding Scheme

The RELAP5 thermal-hydraulic model solves eight field equations for eight primary dependent variables. The primary dependent variables are pressure (P), phasic specific internal energies (U_g, U_f), vapor volume fraction (void fraction) (α_g), phasic velocities (v_g, v_f), noncondensable quality (X_n), and boron density (ρ_b). The independent variables are time (t) and distance (x). The secondary dependent variables used in the equations are phasic densities (ρ_g, ρ_f), phasic temperatures (T_g, T_f), saturation temperature (T_s), and noncondensable mass fraction in noncondensable gas phase (X_{ni}). The equations are described in the Models and Correlations Code Manual (NUREG/CR-5535/Rev 1-Vol IV)

SBO ANALYSIS AND RESULTS

The heat transfer from the PHTS to the steam generators causes the water boil-off and the increase of the pressure in steam generators secondary side. When the steam generators secondary pressure reaches the set point for the opening of MSSVs, the steam is discharged from the secondary side to the environment outside the containment. Following SBO primary pumps trip and the flow rate through reactor core decreases rapidly to the level of natural circulation. The SGs mass inventories and therefore the water level in the steam generators continuously decrease as a result of boil-off. When the SGs secondary side inventories are depleted, the SGs are no longer a heat sink to remove heat from PHTS and the natural circulation in PHTS is ceased.

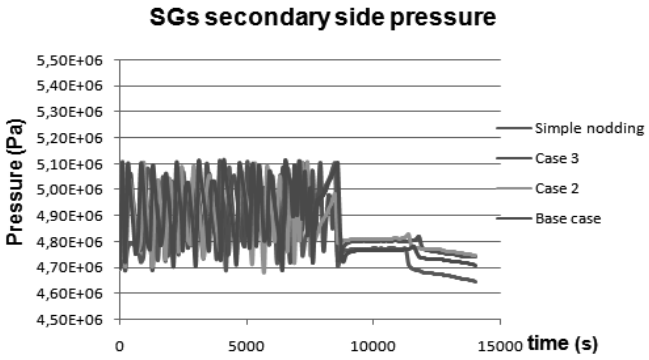


Figure 7: Pressure in SGs secondary side

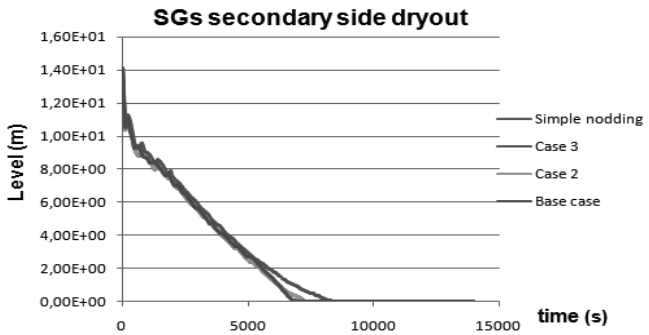


Figure 8: Dryout phenomena in SGs secondary side

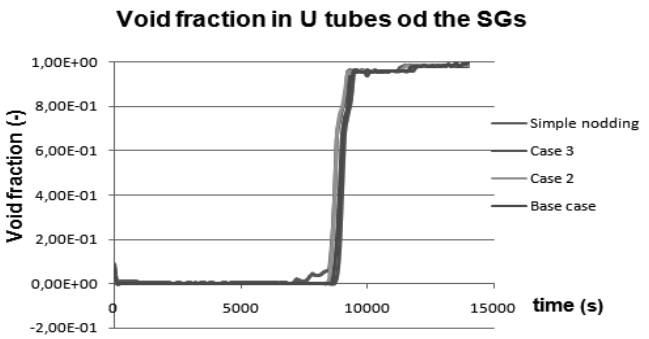


Figure 9: Void fraction in SGs primary side

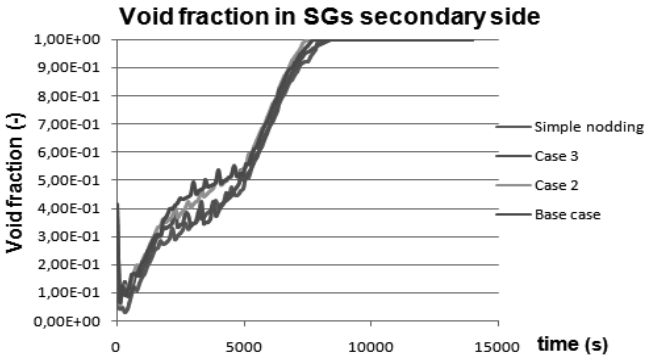


Figure 10: Void fraction in SGs secondary side

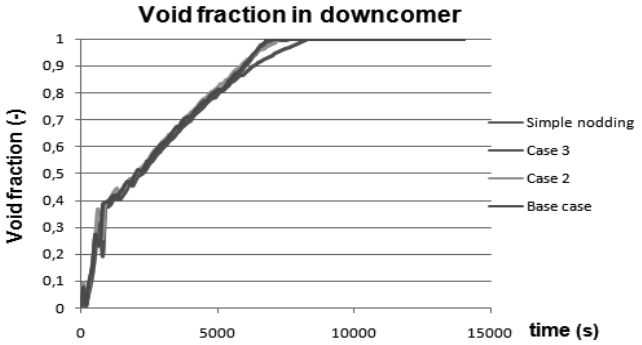


Figure 11: Void fraction in SGs downcomer

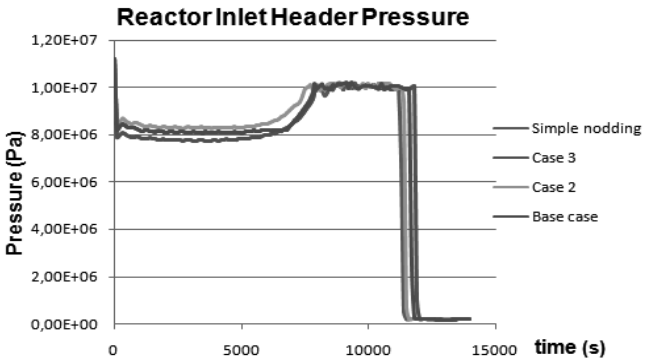


Figure 12: Pressure in RIH

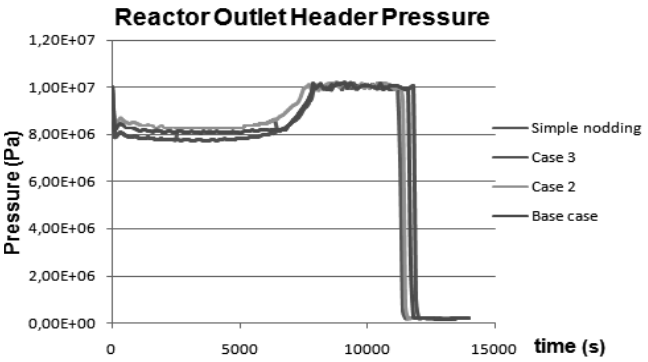


Figure 13: Pressure in ROH

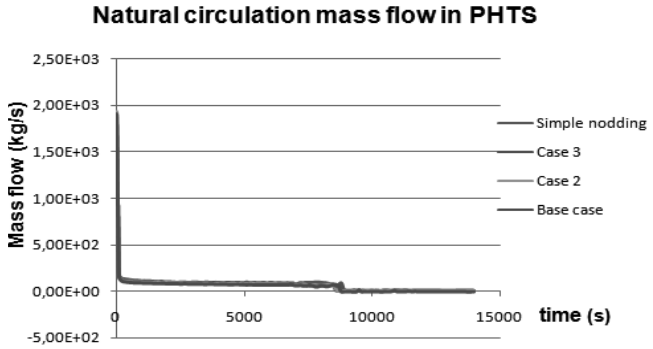


Figure 14: Natural circulation in PHTS

Table 1: Time Sequence Significant for SGs Behaviour

Event	Time (s)			
	Case 1	Case 2	Case 3	Case 4
Loss of Class III and Class IV power	0.0	0.0	0.0	0.0
Reactor trip (shutdown)	0.0	0.0	0.0	0.0
Secondary side of SGs is dry (at all SGs)	7700	8000	8400	7800
First time LRVs open	9000	8600	8800	9100
First fuel channel breaks	11700	11600	12000	11500

CONCLUSIONS

This modelling approach has intended to show the influence of the sensitivity studies through the variation in the number of axial volumes used to describe a major component of a CANDU reactor to some of the parameters of the plant during the early stages of a severe accident.

The simplest subdivision of a model into a set of control volumes or nodes is obtained by dividing the entire model into approximately equally-sized nodes. Appropriate node size is governed by several factors: numerical stability, run time, and spatial convergence. Numerical stability requires that the ratio of the node length to diameter be unity or greater. In practice, this ratio is much larger than one, but this "rule" provides a lower limit. Generally, nodes should be defined as large as possible without compromising spatial convergence of the results. That is because node size directly influences run time; the smaller the node, the smaller the maximum time step size to remain numerically stable. The material Courant limit dictates that the time step not exceed the node length divided by the maximum fluid velocity. Determining spatial convergence in the numerical results is a less straightforward process. However, suitable nodalization is problem dependent, and the user must exercise some judgment as to where in the model nodalization sensitivity studies are warranted.

The event timing showed no major differences in the evolution of the accident in a CANDU 6 reactor with different configurations of the steam generators, and the evolution of the parameters show similar trends with only slight differences in evolution, considered as being insignificant for the progression of the accident.

Advanced analyses should be made in order to determine the influence of using a single volume instead of pipe components in RELAP describing the SGs secondary side taking in count the flow direction and path in the opposite side of the preheater section.

REFERENCES

Chaplin, R. A. 2016. "The Essential CANDU - Nuclear Plant Systems." UNENE, Department of Engineering Physics, McMaster University Hamilton, Ontario, Canada

L. Siefken, E. Coryell, E. Harvego, and J. K. Hohorst 1996. SCDAP/RELAP5/MOD3.3 Code Manual,. s.l. : NUREG/CR-6150 – INEL-96/0422 – Revision 2., 1996

RELAP5&RELAP/SCDAPSIM Software. RELAP5 Thermal-Hydraulic Safety Analysis Software. [Online] Innovative Systems Software, 2017. [Cited: 31 August 2017.] www.relap.com

C. M. Allison, J. K Hohorst 2008. "Role of RELAP/SCDAPSIM in Nuclear Safety", *Proceedings of the TopSafe Conference,* Dubrovnik, Croatia, 2008

Dupleac, D., Mladin, M., Prisecaru, I. 2009. "Generic CANDU 6 plant severe accident analysis employing SCDAPSIMRELAP5 code.", Nuclear Engineering and Design Volume: 239 Issue: 10 Pages: 2093-2103

Dinca, E., Dupleac, D., Nistor-Vlad, R. M., Bonelli, A., Siefken, L. J., Allison, C. M., Hohorst, J. K. 2015. "Analysis of a SBO in a CANDU using RELAP/SCDAPSIM/MOD3.6", *Proceedings of the 7th International Conference on Modelling and Simulation in Nuclear Science and Engineering* (7ICMSNSE), Ottawa Marriott Hotel, Ottawa, Ontario, Canada, October 18-21, 2015

Zhou, F. and Novog, D. 2017. "RELAP5 Simulation of CANDU Station Blackout Accidents with/without Water Make-up to the Steam Generators.", Nuclear Engineering and Design, Volume 318, July 2017, Pages 35-53

Nuclear Safety Analysis Division, 2001. "RELAP5/MOD3.3 CODE MANUAL VOLUME IV. Models and Correlations.", NUREG/CR-5535/Rev 1-Vol IV. Idaho Falls, Idaho, December 2001

IAEA 2013. "TECDOC - 1727, Benchmarking Severe Accident Computer Codes for Heavy Water Reactor Applications." International Atomic Energy Agency, Vienna, 2013, 311 p. ISSN 1011–4289; no. 1727; ISBN 978–92–0–114413–3

AUTHOR BIOGRAPHY

ROXANA-MIHAELA NISTOR-VLAD went to the University Politehnica of Bucharest in 2009 and studied Power Engineering and obtained the bachelor degree in Power Engineering and Nuclear Technologies in 2013**.** During the master studies, also Nuclear Engineering, she worked for a Technical Support Organization for the Nuclear Safety Department as a junior nuclear engineer. After finishing her master's, in 2015, she started working as a Teaching Assistant at University Politehnica of Bucharest, working in parallel for her PhD thesis on Accident analysis in CANDU 6 reactors.

AEROSPACE SIMULATION

TIME MANAGEMENT OF HETEROGENEOUS DISTRIBUTED SIMULATION

Clément Michel
Janette Cardoso
Pierre Siron
ISAE-SUPAERO, University of Toulouse
10 avenue Édouard Belin
BP 54032 - 31055 Toulouse CEDEX 4, France
Email: {firstname.lastname}@isae-supaero.fr

KEYWORDS
Aerospace, Distributed Processors, Model design, Discrete simulation, Simulation interfaces, Cyber-Physical Systems, Heterogeneous Systems, HLA, Distributed Simulation

ABSTRACT

Cyber-physical systems (CPS), by their very nature, mix continuous and discrete behavior and are modeled by heterogeneous components. Formal analysis cannot always handle such complex systems and simulation is a necessary step. In particular, distributed simulation is very useful for the validation of CPS for two main reasons: either the CPS itself is distributed (e.g., a fleet of UAVs) or the CPS is too complex and/or has too much models (e.g., an aircraft). We discuss in this paper the impact of distributing the simulation of a system: which are the rules that must be applied to guarantee a correct behavior between the different simulators? If a centralized simulation already exists (using Discrete Event simulation), which hypothesis must be made for the Distributed Discrete Event simulation? The co-simulation framework used and discussed in this work is Ptolemy-HLA. It allows a Ptolemy model to be distributed using the high-level architecture (HLA) standard.

INTRODUCTION

The analysis of cyber-physical systems (CPS) is a complex task due to the heterogeneity of the parts involved, as they integrate different methodologies and tools.
Simulating different parts of a CPS requires different abstraction and tool supports, and the lack of interoperability between tools poses a major challenge. Because of the nature of the CPS, or because of its complexity, distribution can be necessary.
Distributing a simulation brings its own challenges, as it requires all the simulation elements to conform to a collection of rules in order for the elements to communicate between them.

In this work, the Ptolemy framework is used for modeling the CPS system, and HLA is the standard chosen for the distribution. The IEEE High-Level Architecture (HLA) is a standard for distributed discrete-event (DDE) simulation. CERTI is a HLA-compliant RTI (Run Time Infrastructure). Ptolemy II is a modeling and simulation tool for heterogeneous systems, well suited for modeling CPS since it provides different models of computation (MoC). In this paper, the simulation entities, called **federates** are Ptolemy models, but the approach presented in this paper can be easily used for other simulators. We focus on the time representation issues introduced in a distributed simulation, using HLA as a standard for the co-simulation and Ptolemy as a simulator component. Interoperability and reuse are important targets, so the coupling between simulators is an important issue in the design. The first step in a distributed simulation is to find the partition of a simulation model consisting of a set of sub-models. Which sub-models can (or must) be put together in a *simulator* belonging to the distributed simulation? How to guarantee that the distributed simulation will be valid (according to some criteria)?
The purpose of this work is twofold: Providing properties that ensure a correct time coordination between a federate and HLA, and studying the impact of the distribution in the Ptolemy-HLA framework in order to build models requiring timely input. By studying this specific coupling, we intend to find general problems as well their solutions concerning the time management of heterogeneous distributed simulations.
We will start by presenting the characteristics of both Ptolemy and HLA. Then, we will discuss the impact of an HLA-based distribution, before introducing elements needed for the Ptolemy-HLA simulation to be conservative.

PTOLEMY

Ptolemy II is a Java open-source simulation and modeling tool intended for experimenting with system design techniques, particularly those that involve combinations of different types of models (Ptolemaeus 2014). A Ptolemy simulation, called model, is com-

posed of a Director and software components called actors, that execute concurrently and communicate through messages (called events) carrying values, sent via interconnected ports. An actor that is executed is said to be fired.

The collection of rules that governs concurrent execution of the actors and the communication between them is called a Model of Computation (MoC). The MoC for each actor is implemented as a Director, a software component that dictates how actors should be fired. In this paper, we focus on the Discrete Event (DE) and Continuous MoCs.
Ptolemy uses a model of time known as the superdense time, represented by a tuple (t, n), where t is called the model time and n is called the microstep. The model time represents the time at which some events occurs, and the microstep represents the sequencing of events that occur at the same model time (Manna and Pnueli 1993, Ptolemaeus 2014). The events are ordered in the event queue first by t, then by n.
An event e is noted $e((t, n), v)$ with (t, n) the timestamp and v the value of the event. To ensure determinism, the order in which actors are fired when multiple events are queued for the same timestamp (t, n) is given by the actor rank, a topological sort that lists the actors in data-precedence order.
In the DE MoC, a model advances its logical time to the timestamp of the next event in the queue, and the actor to whom this event is destined for is executed. In the Continuous MoC, the model advances its logical time to a timestamp computed by the solver (Runge-Kutta 23 or Runge-Kutta 45), and all actors are fired at once.
Ptolemy provides a so-called `TimeRegulator` interface with a `proposeTime` method (Ptolemaeus 2014). This interface is implemented by attributes that wish to be consulted when a Director advances time. The Director will call the `proposeTime` method, passing it a proposed time to advance to, and the method will return either the same proposed time or a smaller time.

HLA

The High-Level Architecture (HLA) is a standard for distributed discrete-event simulation. In HLA terminology, the entire system to be simulated is represented by a *federation*, which is a collection of *federates* (simulation entities or simulators).
The execution of a federation uses a middleware, called the *Runtime Infrastructure* (RTI). The federation have a *Federation Object Model* (FOM), a file that contains the definition of data structures (called objects) exchanged between the federates.
In a federation, messages are exchanged between the various federates through the RTI. Those messages, called events, can be timestamped. Messages that are time-stamped are said TSO for Time Stamp Order.

Among the services defined by the HLA standard (IEEE-SA Standards Board 2010), we will focus on object management and on time management.

The object management service allows for federates to exchange messages and values. We focus here on two functions: *Update attribute value (UAV)*, that sends a message to the federation, and *Reflect Attribute Value (RAV)*, that delivers a message from the federation.
A federate is said *regulating and constrained* when it both sends and receives TSO messages. When all the federates are regulating and constrained, the federation is said to be *conservatively synchronized* (Kuhl et al. 2000). Regulating federates generate TSO messages that must occur no earlier than $h_c + lah$, with h_c being the current HLA time for the federate, and lah being its lookahead, a value that establishes a lower bound on the timestamps that can be sent.
While a Discrete Event simulation simply advances its time to t when wanting to, a Distributed Discrete Event simulation first asks for the permission to advance to h, and only advances to h when granted its authorization.
A time constrained federate requests to move its logical time forward by first asking the RTI to do so, either through a *Next Event Request (NER)* or a *Time Advance Request (TAR)*. The RTI replies to this request by sending all TSO messages up to h to the federate, then sending back a *Time Advance Grant (TAG)*. The TAG is written $TAG(h)$ and is an "authorization" for the federate to advance to the logical time h.

TAR
Consider a federate is at time h and asks for a time h_1 through a $TAR(h_1)$; the next time value is always h_1 granted by $TAG(h_1)$. So $TAR(h_1) \rightarrow TAG(h_1)$. The figure 1 illustrates the reception of a $RAV(h')$ messages with $h < h' \leq h_1$ during the advancing phase to h_1: the logical time is not updated to h' and is eventually granted to h_1.

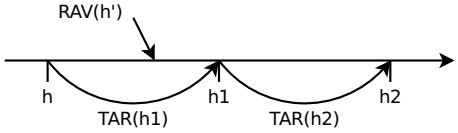


Figure 1: TAR Time Advance Policy Illustrated.

NER
Consider a federate is at time h and asks for a time h_1 through a $NER(h_1)$. If it receives a message $RAV(h')$ with $h < h' \leq h_1$, it advances its time to h':

$$NER(h_1) \rightarrow \begin{cases} TAG(h'), & \text{if } RAV(h') \\ TAG(h_1), & \text{if no } RAV(h') \end{cases}$$

Figure 2 illustrates this advance policy. If the federate is still willing to go to h_1, a new $NER(h_1)$ must be asked.

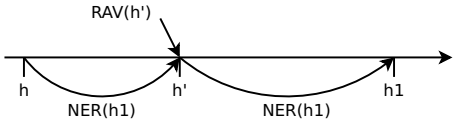


Figure 2: NER Time Advance Policy Illustrated.

IMPACT OF THE DISTRIBUTION AND COUPLING

To take part in a HLA federation, a simulator needs a HLA *coupling interface*, that handles the coupling between the simulator and HLA. Let us consider Figure 3 where two parts are depicted:

- Distribution: ruled by HLA standard that guarantees the time advancing is coherent for all federates (no simulator will advance to a time in the past, and TSO message are delivered in time-stamp order);
- Coupling: put in accordance the time advancing mechanism of the simulator with the HLA time advancing mechanism and so guarantee an ordered data exchange between the simulators in the federation.

The *object* management services (UAV, RAV) and the *time* management services (NER, TAR) are independent. However, the delivery of a RAV message is done during the time advancing phase.
Let be t be the simulation time and h the HLA time. The *distributed* simulation in Figure 3 is a Federation where Federate `f1` simulates a model `M1` and Federate `f2` simulates a model `M2`. Each federate has its own calendar queue, and events are sent/received through the RTI using HLA services UAV and RAV. The time is advanced using HLA services TAR (or NER).

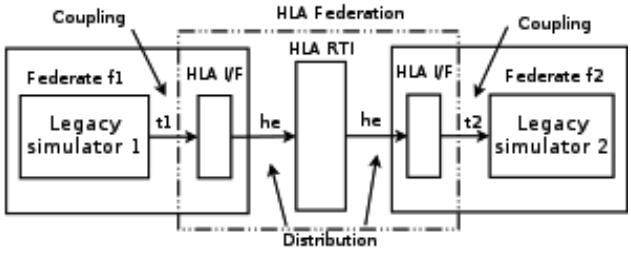


Figure 3: Distribution and coupling

Let be:

- t_1 the timestamp at which $f1$ wants to produce an event (to be sent through the RTI);
- h_e the timestamp of this corresponding event in the HLA service UAV sent through the RTI; it is also the timestamp of the callback RAV;
- t_2 the timestamp of an event put in `f2` calendar queue after the reception of the RAV from the RTI;
- h_c the current HLA time;
- lah the lookahead;
- TS the HLA time step, defined only when using TAR; $h_c + TS$ is called *next point in time*.

In such a distributed simulation, the event time stamp of a message can be altered twice, $t_1 \rightarrow h_e$ and $h_e \rightarrow t_2$, due to:

- Different time representations;
- the distribution itself and its rules, e.g. a federate cannot send an event earlier then h_c+ lookahead;
- the coupling *HLA* $\leftrightarrow$ *simulator* providing a coherent way for both time advancing mechanisms.

Indeed, a coupled distributed simulation does not come without a cost, and $t_2 > t_1$. In the next section, the Ptolemy-HLA coupling will be presented. Let us point out that in this paper we will focus on NER time management but TAR was also designed and implemented.

PTOLEMY-HLA

Ptolemy-HLA (Lasnier et al. 2013) aims to provide a distributed, conservative simulation. Both HLA and Ptolemy offer conservative time management options (Buck et al. 1994, Fujimoto 2003), then the interface binding them together must be conservative as well. In this section, we detail the different conditions needed for a Ptolemy-HLA conservative simulation.

Ptolemy/HLA Coupling Design

For a Ptolemy model to work as a HLA federate, the model's Director must be a DE Director. However, it can very well contain composite actors that themselves contain a Continuous Director each. Three Ptolemy components are added for Ptolemy and HLA to communicate:

- *HlaManager*: encompasses the high-level HLA rules and services such as time management and object management.
- *HlaSubscriber* actor: receives events from the HLA federation through RAVs (pictured on Figure 4b).
- *HlaPublisher* actor: sends Ptolemy events to the HLA federation through UAV service (pictured on Figure 4c).

Events circulating through the federation (even when not sent to another federate) must respect both HLA and Ptolemy time restrictions. In order to coordinate both Ptolemy time advancing and HLA time advancing, the `proposeTime` method was extended in order to allow Ptolemy to query the RTI for a time t. When a Ptolemy

federate wants to advance its time to the timestamp of the earliest event available, it first interrogates the RTI through the `proposeTime` algorithm, and wait for the RTI to grant it.
Algorithm 1 displays the `proposeTime` for the NER.

Algorithm 1 Extended `proposeTime` algorithm (NER)

```
1: NER(t_asked)
2: while TAG(h_received) not received do
3:     TICK()
4: end while
5: t_asked ← h_received
6: if RAV received then
7:     Schedule HlaSubscriber firing at t_asked
8: end if
9: return t_asked
```

The Ptolemy-HLA framework allows a federation (`f1`, `f2`) with any combination of federates time management: (NER,NER), (NER,TAR), (TAR, NER), (TAR,TAR). This is true also for a federation with more than two federates.
The events exchanged between the federates see their timestamp manipulated and shifted as depicted on Figure 3. In order to coordinate HLA and Ptolemy time, the coupling interface in the simulator has two time lines. In this section, we consider that both time lines have the same computer representation, so both can be directly compared.
In the sequel, let us present how the timestamp can change in the Ptolemy-HLA framework when sending and receiving an event according to the time management.

Sending a Ptolemy event through the RTI:
Let be t_{1_c} the current (Ptolemy) logical time of federate `f1` and h_{1_c} be the current HLA time. The federate wants to send an event $e(t_{1_c})$ (through the HlaPublisher actor) using a HLA UAV(h_1) service; timestamp h_1 depends on the federate's time management:

$$\text{NER: } h_1 = h_{1_c} + lah, \text{with } h_{1_c} = t_{1_c} \tag{1a}$$

$$\text{TAR: } h_1 = \begin{cases} h_{1_c} + lah, & \text{if } t_{1_c} < h_{1_c} + lah \\ t_{1_c}, & \text{otherwise} \end{cases} \tag{1b}$$

Receiving a RAV from the RTI:
Let be t_{2_c} the current (Ptolemy) logical time of federate `f2` and h_{2_c} be the current HLA time. The reception of the HLA RAV(h_1) service at h_{2_c} wakes up a HlaSubscriber actor with a Ptolemy event $e(t_2)$; the timestamp t_2 depends on the federate's time management:

$$\text{NER: } t_2 = h_1 \tag{2a}$$

$$\text{TAR: } t_2 = h_{2_c} + TS \tag{2b}$$

These rules are necessary in order to produce a conservative distributed simulation that respects the rules of both HLA and Ptolemy (i.e. no TSO message sent earlier than h_c+ *lookahead*, no event insertion in Ptolemy's past).
It can be seen from equations 1a, 1b, 2a and 2b that the difference $t_2 - t_{1_c}$ between the production of an event at time t_{1_c} in `f1` and its consumption at time t_2 at `f2` depends on the time management of each federate and its parameters:

$$t_2 - t_{1_c} = f(lah, (TS), t_{1_c}, h_{1_c}, h_{2_c}).$$

Let us consider again Figure 3, where both federates `f1` and `f2` use a NER time management with h_c current HLA time for both federates and a same lookahead *lah*. Notice that they can also have different lookaheads.
Let be t_{1_c} the current (Ptolemy) logical time and $e_{\text{f1}}(t_{1_c})$ the event that wakes up the HlaPublisher actor in `f1`. By using equation 1a, the corresponding UAV is sent with timestamp $h_1 = t_{1_c} + lah$ (regardless of `f2` time management). The RAV event with timestamp h_1 received by `f2` is queued at HlaSubscriber actor as a event e_{f2} with timestamp $t_2 = h_1$ given by equation 2a. Thus, the (NER,NER) configuration introduces in the distributed simulation a delay $\delta_{NER-NER} = t_2 - t_{1_c} = lah$.

Ptolemy/HLA coupling: implementation issues

Coupling Ptolemy with HLA requires to take into account that the time representation is different: CERTI RTI uses IEEE 754 `double` (with dynamic precision) and Ptolemy uses its own `Time` representation (fixed precision). These two time representations must be well converted.
Timestamps in Ptolemy are represented under the form $t = n * r$ with n an integer (called the *time value*) and r a Java *double* (called the *time resolution*). Thus, time values in Ptolemy can only adopt values that are multiple of r, e.g. r, $2r$, $3r$, etc...
Despite the time representation difference, one must guarantee that:

- Rule 1: No event should be inserted into the simulator's past — $t_{RAV} > t_{PtII}$
- Rule 2: No UAV should be sent in the past of the HLA federation — $h_{UAV} > h_{HLA}$

Non observance of these rules violates the causality of either the federation or the simulator, outputting wrong results.
At the earliest stages of Ptolemy-HLA, the question of the conversion between different (computer) representation was eluded, even if it was well-known that a function $\hat{f}$ such as $\hat{f}(h_{HLA}) = t_{PtII}$ and $\hat{f}^{-1}(t_{PtII}) = h_{HLA}$ was unobtainable.

In the following, the difference between HLA and Ptolemy (or any other simulator that does not use *double*) time representation is taken into account. In this work, we introduce two functions f and g such as:

$$\begin{cases} f(h_{HLA}) = t_{PtII} \\ g(t_{PtII}) = h'_{HLA} \end{cases}$$

Several properties are required from f and g in order to respect temporal coherence in TAR and NER cases, such as:

- f and g are monotonous strictly increasing functions
- $\forall h_1, h_2, h_3 \quad h_3 \geq h_2 > h_1 \Rightarrow h_3 \geq (g \circ f)(h_2) > h_1$
- $\forall t_1, t_2 \quad t_2 \geq t_1 \Rightarrow (f \circ g)(t_2) \geq t_1$.

We find these properties hard to fulfill from a mathematical point of view, even probably they cannot be fulfilled if f and g are not the identity function and this is not possible with heterogeneous systems. In such a context, we can find and use algorithmic solutions in order to use existing conversion functions and to solve the list of the identified problems.
Thus, each time that t and h must be compared in Algorithm 1, we introduce f and g, two monotonous, non-strictly increasing functions, with the hypothesis that algorithmic modifications will allow the functions to work in our context. Considering the f and g functions, Algorithm 1 is modified to Algorithm 2 in the following way:

- NER is called only if $g(t_{asked}) > h_{current}$: prevents Ptolemy from requesting a time advance if the requested time is imperceptible by the HLA time.
- $t_{asked} \leftarrow t_{current} + r$, if $f(h_{received}) \leq t_{current}$, r being Ptolemy's resolution: prevents HLA from granting a time advance imperceptible by the Ptolemy time.

Equations 1a, 1b, 2a and 2b are also modified in order to take f and g into account. For instance, 1a becomes $h_1 =, h_{1_c} + lah,$ with $h_{1_c} = g(t_{1_c})$.
Finally, we keep the association between the return of a function and its argument, i.e. when a value $t_1 = f(h_1)$ is received, the association between t_1 and h_1 is memorized, and we will output h_1 when evaluating $g(t_1)$.
Algorithm 2 ensures a framework where rules 1 and 2 are observed.
Because of this difference in HLA and Ptolemy time representation, the theoretical delay induced by the distribution described in the previous sub-section can be slightly altered by a value ϵ introduced by the conversion between the time representations.

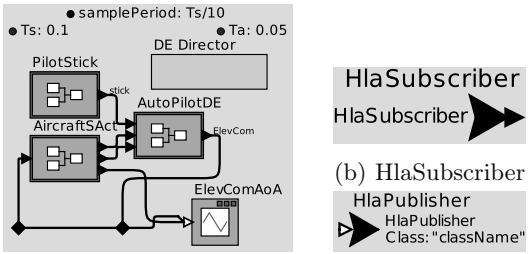
APPLICATION

A full case study of a longitudinal flight control (Lasnier et al. 2013) will be considered in this section.

Algorithm 2 Modified `ProposeTime` Algorithm (NER)

```
1:  if g(t_asked) > h_current then
2:      NER(g(t_asked))
3:      while TAG(h_received) not received do
4:          TICK()
5:      end while
6:      if RAV received then
7:          if f(h_received) > t_current then
8:              t_asked ← f(h_received)
9:          else
10:             t_asked ← t_current + r
11:         end if
12:         Schedule HlaSubscriber firing at t_asked
13:     end if
14: end if
15: return t_asked
```

Centralized simulation



(a) Centralized F-14. (b) HlaSubscriber (c) HlaPublisher

Figure 4: F-14

The centralized (hierarchical and heterogeneous) Ptolemy model, based on a Matlab model of an F-14 aircraft, is pictured on Figure 4a. `Aircraft` and `Stick` are composite actors modeled in the continuous MoC and the `AutoPilotDE` composite actor (controller) uses a DE MoC. As in a real aircraft, signals in the continuous domain need to be sampled in order to be used in the DE domain.
Let us highlight the model of the `AutoPilotDE` block. The block comports one output `ElevCom` that is the elevation command sent to the `Aircraft` block, and takes in three inputs:

- `stickS`: The current action on the stick.
- `alphaS`: The current vertical velocity of the F-14.
- `qS`: The current pitch rate of the F-14.

If only one of the inputs receives an event, the `AutoPilotDE` block will consider the others inputs as absent. Thus, receiving the events at three different timestamps would trigger three different rounds of computation that would output incorrect `elevCom` values. As a consequence, the events concerning variables `stickS`, `alphaS` and `qS` must be fed at the same timestamp (t, n) for the output to be correct. Thanks to the topological sort, we are guaranteed to have all these events

produced by the `PeriodicSampler` before actors within `AutoPilotDE` are fired, processing the three events at once.

Distributed simulation

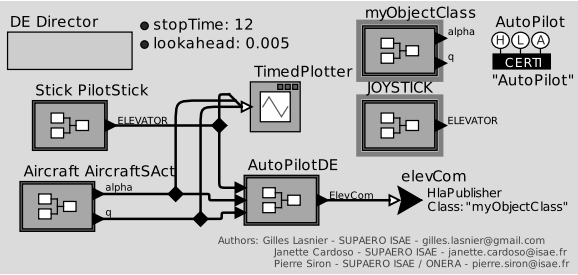


Figure 5: Federate `f14AutoPilotDE`.

A distributed version of the F-14 is obtained by creating a new federate model for each of the three actors in Figure 4a: `PilotStick`, `AutoPilotDE`, and `PilotStick`. For each actor, an input port is connected to an `HlaSubscriber` (pictured on Figure 4b) and an output port is connected to an `HlaPublisher` (pictured on Figure 4c). The `f14AutoPilotDE` federate, containing the `AutoPilotDE` block, is pictured on Figure 5. The `Stick PilotStick` and `Aircraft AircraftSAct` are composite actors containing the `HlaSubscriber` receiving events, respectively, from `f14PilotStick` and `f14Aircraft` federates, while the `elevCom` actor is an `HlaPublisher` sending events to the federation.

The three federates use the NER time advancing mechanism with a same lookahead lah.

Let us consider $t_{stick}, t_{alpha}, t_q$ the timestamps of the events arriving, respectively, at inputs `stickS`, `alphaS` (generated by the `f14Aircraft` federate) and `qS` (generated by the `f14PilotStick` federate).

According to equation 1a, the UAVs at `f14Aircraft` and `f14PilotStick` federate are sent with timestamp $h = g(t + lah)$, with h the current HLA time for the federate when the event is produced. The RAVs received at `f14AutoPilotDE` federate are then queued as events e with timestamps $t' = f(h)$ according to equation 2a.

For the `AutoPilotDE` block inside `f14AutoPilotDE` federate to work properly, events for all the inputs are required to be received at the same time. Thus, $t'_{stick} = t'_{alpha} = t'_q$. By expressing t' as a function of t, we obtain the following condition:

$$f(g(t_{stick}) + lah) = f(g(t_{alpha}) + lah) = f(g(t_q) + lah)$$

Since the three federates are Ptolemy federates, they have the same f, g functions. Moreover, as all the federates have the same lah, we can reduce the condition to

$$t_{stick} = t_{alpha} = t_q$$

This condition is satisfied as the `PeriodicSampler`s (actors sampling the output of continuous models in `Aircraft` and `Stick` blocks in Figure 4a) outputs all the events at the same time since they have the same period. Thus, the `AutoPilotDE` block receives and processes its inputs the same way than in the centralized model, albeit in a time-shifted manner. The bias lah introduced by the distribution does not, in this case, generate significant changes in the simulation behavior. So, the `AutoPilotDE` block maintains its expected behavior under some conditions on the other federates.

RELATED WORK

In (Deschamps et al. 2017), the partitioning of a simulation is discussed, addressing the simultaneity at the inputs of all components (of a distributed simulation). A formalism is proposed to prove the equality of delays by design and so the distributed simulation is valid. The case study of an f14 aircraft is presented in (Michel 2017) where, for a given partition, the *cyber* component is analyzed to guarantee that the data consumptions are simultaneous using different implementations: memory blocs, internal clock.

An aspect in distributed simulation is the multiplicity of times and the different representation of these time values. One must guarantee that the times advancing of the distributed simulation are done correctly (e.g., conservation simulation (Fujimoto 2003)). The time representation of each simulator, and the entity that allows for co-simulation – HLA or master algorithm in FMI – must be well dealt in the coupling between them. A detailed analysis of time representation in the FMI framework is done in (Cremona et al. 2017). In particular this paper discuss the choice of resolution to be used when the FMUs (components of a co-simulation) have different resolutions. The coordination of different times issue appears also in real cyber-physical systems (Shrivastava et al. 2016).

In (Nägele and Hooman 2017), a HLA/simulator coupling, called wrapper, is modeled using POOSL (Parallel Object-Oriented Specification Language). They use PoRTIco, a HLA compliant RTI and their federates are also regulating and constrained.

CONCLUSION AND PERSPECTIVES

Ptolemy-HLA framework allows to run valid distributed simulations for Event-Driven (NER) and Time-Stepped (TAR) advancing mechanisms. For a same federate model, the user can choose several parameters such as the time advancing mechanism (TAR or NER) and the lookahead. The framework implementation as well several demos are available at (The Ptolemy Project 2017). The coupling algorithm between Ptolemy and HLA in the case of NER mechanism was detailed in this paper. For highlighting a valid coupling, first we consider that

the (computer) representation of its timelines are the same, then we extended the algorithm for taking their representation difference into account. The TAR mechanism was also implemented in (The Ptolemy Project 2017). In both cases a time bias between the federates, expressed by equations 1a to 2b, is introduced but the algorithm guarantees an valid and efficient distribution. When considering different time representation, the bias can be slightly increased by a value ϵ introduced by the time conversion.

The `f14` Federation presented in this paper was obtained from the centralized model following the characteristics of a CPS: a federate for the cyber part (the controller), a federate for the plant (the aircraft), and a federate for the pilot stick that can be replaced by a real hardware in this federation. However, the distribution of a model needs some thought process about the model. In a centralized model some hypotheses are implicit, for instance the reception of all data generated by uphill actors. This constraint is ensured by the topological sort handled by Ptolemy that determines a deterministic execution order. This execution order can also be ensured in a distributed context, as well as the (logical) simultaneity of the inbound events of an actor, even considering the timestamp modification introduced by HLA. An analysis must be performed by the user to ensure that the chosen distribution is correct since a simulation can be distributed according to different mappings. For helping the user to make this analysis, the next step is the design of an extra "layer" that both helps the user to distribute a model and ensure that the centralized model is correctly distributed taking into account the bias introducing by the distribution and the coupling.

ACKNOWLEDGEMENT

This research was partly supported by the French Ministry of Defense through financial support of the Direction Générale de l'Armement.

REFERENCES

Buck J.T.; Ha S.; Lee E.A.; and Messerschmitt D.G., 1994. *Ptolemy: A Framework for Simulating and Prototyping Heterogeneous Systems. International Journal of Computer Simulation*, 4, 155–182.

Cremona F.; Lohstroh M.; Broman D.; Tripakis S.; and Lee E.A., 2017. *Hybrid Co-Simulation: It's About Time.* Tech. Rep. UCB/EECS-2017-6, University of California at Berkeley.

Deschamps H.; Siron P.; Cardoso J.; and Cappello G., 2017. *Toward a Formalism to Study the Scheduling of Cyber-Physical Systems Simulations.* In *2017 IEEE/ACM 21st International Symposium on Distributed Simulation and Real Time Applications (DS-RT) (DS-RT'17).* Rome, Italy.

Fujimoto R.M., 2003. *Parallel Simulation: Distributed Simulation Systems.* In *35th Conference on Winter Simulation.* Winter Simulation Conference, WSC '03. ISBN 978-0-7803-8132-2, 124–134.

IEEE-SA Standards Board, 2010. *IEEE Standard for Modeling and Simulation (M & S) High Level Architecture (HLA): Federate Interface Specification.* Institute of Electrical and Electronics Engineers, New York. ISBN 978-0-7381-6247-8.

Kuhl F.; Dahmann J.; and Weatherly R., 2000. *Creating Computer Simulation Systems: An Introduction to the High Level Architecture.* Prentice Hall PTR, Upper Saddle River, NJ. ISBN 978-0-13-022511-5.

Lasnier G.; Cardoso J.; Siron P.; Pagetti C.; and Derler P., 2013. *Distributed Simulation of Heterogeneous and Real-Time Systems.* In *IEEE/ACM 17th International Symposium on Distributed Simulation and Real Time Applications.* IEEE Computer Society, 55–62.

Manna Z. and Pnueli A., 1993. *Verifying Hybrid Systems.* In *Hybrid Systems*, Springer, Berlin, Heidelberg, Lecture Notes in Computer Science. ISBN 978-3-540-57318-0 978-3-540-48060-0, 4–35. doi:10.1007/3-540-57318-6_22.

Michel C., 2017. *Distributed Simulation of Cyber-Physical Systems.* Tech. rep., ESIEA, ISAE-SUPAERO.

Nägele T. and Hooman J., 2017. *Co-Simulation of Cyber-Physical Systems Using HLA.* In *2017 IEEE 7th Annual Computing and Communication Workshop and Conference (CCWC).* 1–6. doi:10.1109/CCWC.2017.7868401.

Ptolemaeus C. (Ed.), 2014. *System Design, Modeling, and Simulation Using Ptolemy II.* Ptolemy.org.

Shrivastava A.; Derler P.; Baboudr Y.S.L.; Stanton K.; Khayatian M.; Andrade H.A.; Weiss M.; Eidson J.; and Chandhoke S., 2016. *Time in Cyber-Physical Systems.* In *2016 International Conference on Hardware/Software Codesign and System Synthesis (CODES+ISSS).* 1–10.

The Ptolemy Project, 2017. *Ptolemy Project Home Page.* `https://ptolemy.eecs.berkeley.edu/`.

Real-Time Simulation of Large Aircraft Fuel Systems

Stephen Wright
Department of Engineering, Mathematics and Design,
University of West of England,
Frenchay Campus, Coldharbour Lane,
Bristol, UK

Alvery Grazebrook
Fuel Control and Indication,
Airbus Operations Ltd,
Pegasus House, Aerospace Avenue, Filton,
Bristol BS34 7PA, UK

KEYWORDS

Aerospace, Discrete simulation, Real-time, Model testing.

ABSTRACT

This paper presents a method for the real-time simulation of the fluid-mechanical components of the fuel storage and transfer systems of large civil aircraft. The simulation includes modelling of fuel flow, fuel properties and related measurement sensors, under all regions of the aircraft's flight envelope. The Fluid Network Model (FNM) has been successfully deployed in the core of several hardware-in-the-loop integration facilities for the testing of large transport aircraft avionics, including that of the Airbus A380. The principles, features, and limitations in pursuit of its real-time performance goal are presented, and qualitative test results are described to illustrate its application and utility.

INTRODUCTION

The fuel load of modern civil aircraft can account for over half of its take-off weight: thus sophisticated storage and distribution networks, in combination with management avionics, play an essential role in maintaining flight control. One of most celebrated examples in recent years is the Airbus A380; a full load around 250 tonnes of fuel is stored in multiple tanks throughout the airframe, shown in Figure 1. (Langton et. al. 2009)

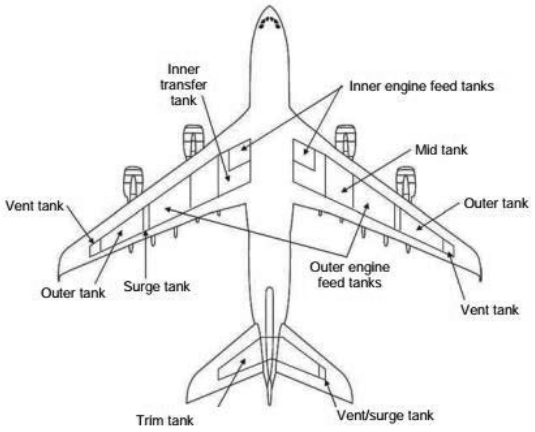


Figure 1: A380 Fuel Tank Layout

Fuel is stored and transferred between tanks by a complex network of pumps, pipes and valves, ensuring that required reliability standards are met under a vast range of flight conditions and failure scenarios. For example, the fuel system equipment within the wing of the A380 is shown in Figure 2. (Langton et. al. 2009)

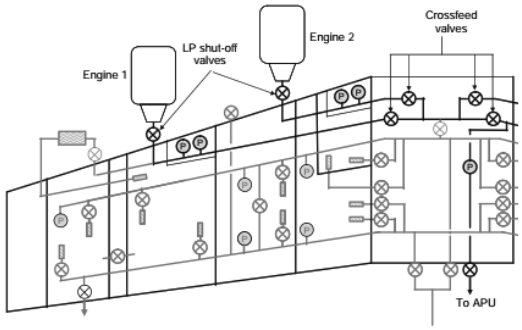


Figure 2: A380 Wing Fuel System Interconnect

As well as ensuring that fuel is continuously delivered to the engines, several other functions are required. These include the monitoring of fuel mass in each individual tank, maintenance of aircraft centre of gravity within the stability envelope of the airframe, monitoring that fuel temperature remains within acceptable limits, and diversion of fuel to heat exchangers to provide cooling of other systems. In order to achieve all these functions reliably and with minimal crew workload, fuel systems are controlled by complex avionics, which in turn interact with other systems and their avionics. Development and test of avionics hardware is invariably performed in simulated Hardware In Loop (HIL) environments prior to flight testing (Moir and Seabridge, 2011)

Several physical mechanisms are supported by the FNM, and each has its own fidelity requirements to support the avionics. The fidelity is assessed against observable behaviour. For example, the size of aircraft such as the A380 leads to significant changes in head pressures between tanks during pitch and roll manoeuvres; one tank is physically raised above another, sometimes to a height of several metres. This effect is exploited by control functions to provide emergency transfer capabilities following failures. In order to test this capability, the FNM must model fuel head-pressure driven flow, with sufficient accuracy to satisfy the expected flow-rates configured within the avionics' software. Even under normal operating conditions, in which fuel is transferred by pump pressure, pipe flow-rates vary due to head-pressure, and the effect must be accounted for.

Thus, the increase in both the scale and functionality of avionics in the last twenty years (Butz 2007) has created a need for more sophisticated HIL-hosted models such as the FNM. Conversely, the availability of high-level programming tools (Mathworks Simulink User's 2015), and low-cost, high-performance computing platforms (Mathworks Simulink Coder User's 2015) have made their development and deployment possible.

THEORY OF OPERATIONS

The FNM simulates the entire network of tanks, pipes, pumps and valves within the aircraft, and their associated sensors. These sensors consist of discrete position detectors to confirm that valves have achieved the fully open or fully shut position, discrete pressure sensors to confirm that activated pumps are delivering their expected pressures, discrete level sensors indicating the fluid surface being above a fixed height, and analogue probes whose capacitance varies with the quantity and properties of fuel covering them: these are used by the avionics to infer the quantity of fuel in a given tank.

A fuel system model must account for the masses of fuel in each tank during refuel, defuel, engine burn-down, and tank-to-tank transfer operations. The FNM enhances this basic capability with calculation of flow-rates for these operations, modelling of the density and dielectric properties of the fuel itself, and introduction of the relationship between fuel quantity and surface height for each tank. The model predicts the fuel flows and quantities in the tank/pipe network for all tank, pump, valve, pitch, and roll configurations, by the real-time solving of a topologically correct resistive network driven by pressure sources representing activated pumps and fuel head pressures. For example, the combined head pressure 'P' due to the fuel in a wing-tank at a given roll angle is described in Figure 3.

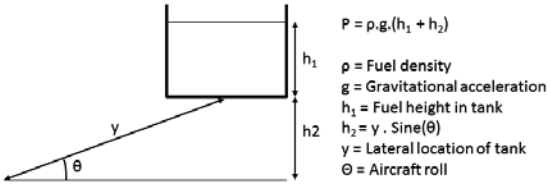


Figure 3: Head Pressure Calculation

This value may be calculated for any condition and then used to drive a pressure source within the model. Similarly, a pump may be modelled by a pressure source that drops to zero when it is deactivated, allowing rapid modification of pump characteristics without having to consider its effects on every network scenario. This approach requires a considerable investment in solver and model development, and is computationally intensive to achieve real-time performance.

Previous generations of FNM developed in the early 1990's adopted an approach based on interpolated look-up tables (LUT) of known flow rates based on a selected set of operational scenarios. This worked well for pumped transfers, but not so well for gravity driven flows. Current tools and execution platforms allow this more sophisticated approach to be taken, which yields clear benefits in terms of coverage across operating scenarios.

Modelling of the system as a topologically correct pressure-driven network gives the additional benefit that all nodes within the network are observable. This allows clear visualisation of the behaviour of the model and its configuration via a Graphical User Interface (GUI), as shown in Figure 4.

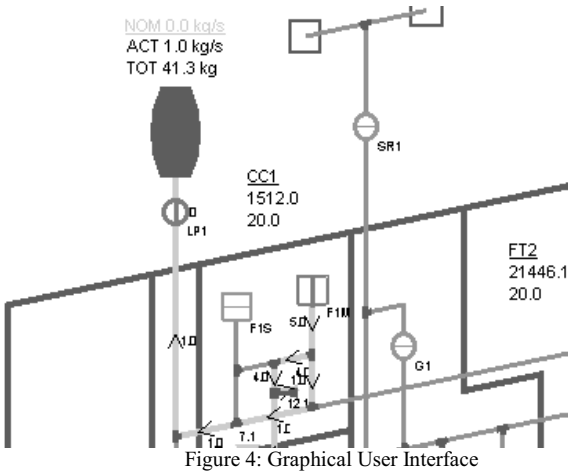


Figure 4: Graphical User Interface

Different grades of aviation fuel have different chemical properties that affect their density and dielectric properties with temperature, as described by the long-established Clausius-Mossotti Law (Van Rysselberghe 1931). Temperature expansion must be modelled in order to predict potential fuel overflows at given temperatures and masses, and dielectric properties must be modelled as the fuel quantity gauging system is based on capacitive effects, which must be rigorously tested. Mixing of these properties within tanks and pipe networks must also be modelled, to accommodate the mixing of different grades of fuel within the aircraft. This introduces a need to model the temperature evolution of the fuel during operation. A lumped parameter temperature model is used, relating fuel temperature to outside air temperature, and fuel mixing from different fuel sources. The parameters are set based on higher fidelity thermal models.

The aircraft fuel tanks are formed from the aircraft structure. Therefore they have complex shapes, and the shape changes based on structural load. The relationship between fuel quantity and surface height is very different from the linear function implied by the straight-sided tank in Figure 3. As the aircraft moves, the pitch/roll angles vary. This relationship is managed by a three-dimensional look-up table using tank pitch, roll and fuel volume to give surface height, in order to inform head pressure, overflow, and pump/valve starvation point calculations. This algorithm is extended with 3D co-ordinate data for each capacitance probe to derive the "wetted length" of each element, and thus deduce the capacitance value reported to the avionics.

The majority of avionics testing facilitated by the FNM is the management of degraded system conditions, and the model must therefore accommodate appropriate fault injection. Some of these conditions may be created by the simple overriding of FNM outputs: for example sensor or actuator failures. However, some faults yield more subtle behaviours, and must be modelled inside the FNM: examples include valve shaft breakage in which the device remains in the open or shut position regardless of the actuated position; another example is leakage of fuel into our out of tanks.

As stated, the FNM is intended for deployment within HIL facilities against actual avionics equipment, making deterministic real-time operation its primary requirement. The fluid-mechanical operation of the fuel system is slow by the standards of most aircraft sub-systems, and this is reflected in the avionics iteration rate of only 1 Hertz. This has led to a model iteration rate of 4 Hertz being specified. This low iteration rate requirement has made the ambitious pressure network solution viable.

TECHNICAL CHALLENGES

The FNM is an example of a multi-source, multi-sink resistive effort-flow network, shown in its most basic form in Figure 5.

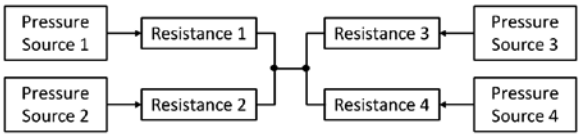


Figure 5: Basic Multi-source Multi-sink Network

The introduction of multiple (i.e. more than two) sources driving a shared bus presents particular mathematic challenges, and various techniques for solving this class of problem in real-time exist, particularly in the electrical domain (Dessaint et. al. 1999). However, the FNM's fuel system application presents several specific challenges. The scale of the model (particularly for the A380 application) is of moderate scale, as defined by the number of pressure sources, sinks, switched resistive elements and network junctions, each of these being in the order of 100. A basic network generally assumes an additional simplification of all pressure sources being ideal (i.e. the same pressure is delivered for all flows rates), and all resistive elements being linear (i.e. the flow through it is directly proportional to the pressure across it): neither of these assumptions is valid for the FNM.

Although head pressures may be considered ideal, the pressure/flow characteristic of the aircraft's fuel pumps is of the form shown in Figure 6. The figure shows that a steadily increasing pressure drop is experienced with increasing flow, corresponding to a notional internal resistance within the element. At much higher flows the pumping device itself approaches a maximum rate and pressure tails off more rapidly to zero.

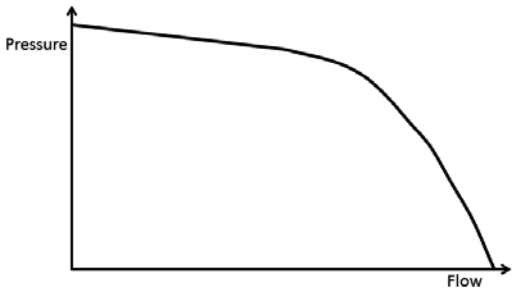


Figure 6: Typical Fuel Pump Pressure/Flow Characteristic

As stated, the characteristic of the resistive elements is also non-ideal: specifically the resistance rises roughly in proportion to flow (due to turbulence effects in the network), yielding a flow that is approximately proportional to the square root of the pressure across the element, as shown in Figure 7.

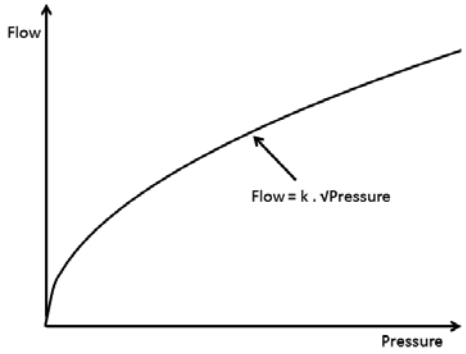


Figure 7: Typical Pipe Pressure/Flow Characteristic

The network also contains another, even more complex, class of non-linear component: Non-Return Valves (NRV). These are the fluid equivalent of electrical diodes, presenting a near-ideal conductor when pressurised in one direction, but an infinite resistance in the other. Modelling of infinite resistance is also required in the modelling of valves when in a shut position, yielding truly zero flow.

At a topological level, the fluid networks of some aircraft (for example the A380) also contain pipe loops, yielding recursive dependencies for many solver algorithms. Thus the basic form of network being modelled by the FNM is more fully illustrated in Figure 8.

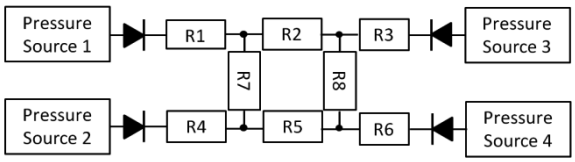


Figure 8: Evolved Multi-source Multi-sink Network

Simultaneous solving of the interaction of all these characteristics in deterministic real-time is the underlying goal of the FNM.

MODELLING APPROXIMATIONS

The FNM must provide sufficient fidelity to allow testing of aircraft avionics under all test scenarios and execute in a sufficiently timely fashion to support real-time operation.

Thus the FNM's design embodies the basic "art" of model design: the identification of appropriate approximations and abstractions that are acceptable for an intended purpose.

It is interesting to note that fuel designers also employ highly detailed simulations developed on the Flowmaster tool (Tu and Lin 2011). These use iterative algorithms to solve detailed models of all non-linear components within the network, which yield accurate flow predictions but take significant time to converge to a solution and may not converge to a solution at all. This clearly renders them inappropriate for real-time applications.

Although an essential part of the design of the fluid-mechanical system, modelling of transient effects generated by fuel in-flow and "surge" effects when stopping flow within the network is not required for avionics testing. Therefore the FNM does not attempt to model these effects, and only predicts fuel flow under steady-state conditions.

In order to allow linear solving techniques to be employed, the fuel pump characteristic shown in Figure 6 is modelled as a bi-modal linearized approximation of this curve, shown overlaid on the actual curve in Figure 9.

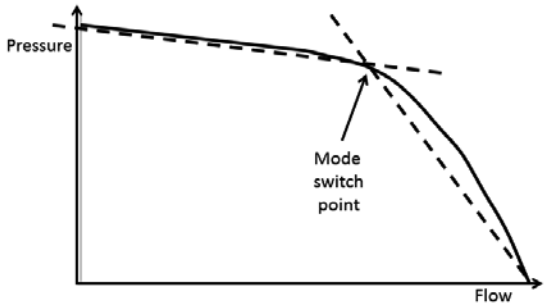


Figure 9: Approximated Fuel Pump Pressure/Flow Characteristic

Switching between the two modes is performed automatically within the model, based on the pump flow rate from the previous iteration of the model. Selection of the model parameters (especially the model switching point) must be made by analysis of supplied pump data, with lines manually fitted to plotted curves to give sufficient accuracy across the operational range. This approximation is valid, as for a given operational condition the pump flow is constant within one or other mode, and switching effects are transitory. The technique of resistance-switching based on delayed state is also used to implement a model of NRVs within the network. For this component, resistance is switched between zero and infinity based on the detected pressure on either side of the component. More broadly, for some solver implementations, state-delay is required to break the recursive dependencies inherent in the topological loops such as those shown in Figure 8. This approach is satisfactory in most network topologies.

The use of delayed state, shared across the elements of the model, is driven by the selection of fixed-step solver techniques, which is in turn is demanded by the requirement for deterministic real-time performance. This efficiently resolves the recursive dependencies inherent in solving these topologies and non-linear functions in real-time, but imposes distinct behaviour limitations. In the case of the bi-modal approximation used for pump modelling, switching between modes will yield an incorrect flow prediction for a single cycle: as such events are rare, this error is acceptable. A similar single-cycle error will occur in the case of an NRV switching from a forward-biased to reverse-biased condition, which will manifest as a reverse flow through the element for one cycle. The effect will also occur in the case of the resistive element, in which the resistance will converge to an exact solution over a number of model iterations. Again, as in practice the aircraft fuel system generally remains in steady state once a configuration has been selected by the controlling avionics, the associated error is acceptable.

IMPLEMENTATION

The FNM is implemented in Mathworks Simulink, and automatically translated to C via the Simulink Coder plugin tool. An example of the high-level Simulink input for the A380 FNM is shown in Figure 10.

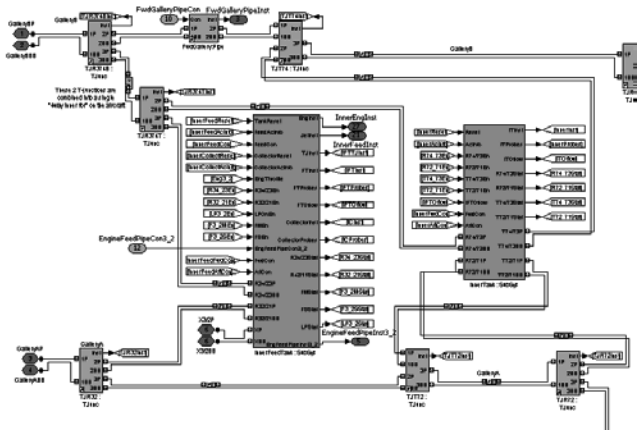


Figure 10: Fluid Network Model Simulink Fragment

A combination of the Simulink graphical input language and embedded C "S-functions" (Mathworks Developing S-Functions 2015) is employed. Appropriate use of the graphical language allows the rapid interconnection of lower-level and generic library blocks, and allows the graphical layout to mirror the topology of the actual network. The Simulink graphical language allows for very efficient dependency calculation, an essential process for the multiple dependency paths inherent in a large network. In complement, the use of C S-functions allows for optimisation of core, performance-limiting algorithms; by giving the developer more control over the final compiled C code that than would be available via the Simulink Coder. C S-functions are also used for inclusion of legacy code used to implement the tank volume/surface-height translation algorithms previously described.

Each aircraft-specific FNM version is constructed from a generic library of fuel system components, implementing components such as tanks, pumps, valves, pipes and interconnecting junctions. The library is shown in Figure 11.

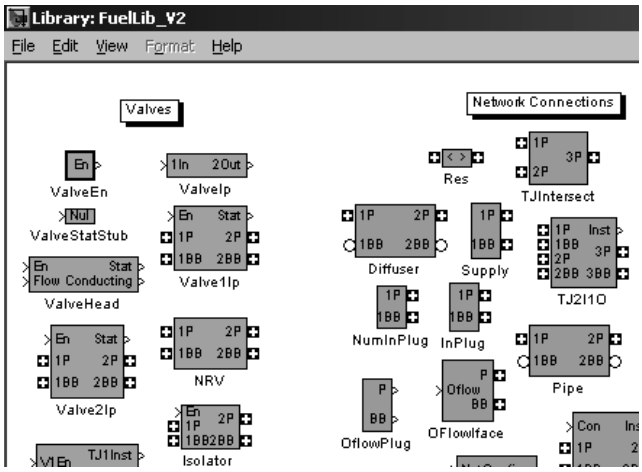


Figure 11: Fluid Network Model Generic Library

These generic components are configurable to a particular aircraft application, either at build-time via Simulink or C constant settings, or at run-time via model inputs. For example, network pipe resistance values remain constant for a given aircraft, and are therefore configured at build-time. Conversely, different pump pressure settings are sometimes desired by avionics test personnel, and are therefore made to be run-time configurable from an initial default setting.

The use of high-level coding methods and management of basic functionality via a generic library achieves two essential features of the model: re-use across multiple aircraft projects and implementation of additional functionality. For example, the FNM and its underlying techniques were initially developed for the Airbus A380 aircraft, and later effort was applied to make the method cross-compatible for the Airbus A400M program: this approach yielded considerable benefits in its rapid application to the Airbus A350 program.

Within the generic fuel system library, solving of network flows may be accomplished by a variety of solving techniques. The choice of solver selection is driven by such factors as performance, calculation accuracy, run-time configurability, and tool licencing costs. In practice, the real-time requirement of the FNM's avionics-test application, and the fixed-rate iteration methods that this implies, is the most constraining driver in this choice. For example, initial versions of the FNM employed the commercially available SimPowerSystems library (formerly "PowerSystem-Blockset") (Dessaint et. al. 1999). Subsequently, this was superseded by an alternative method developed by Wright (to be described in future publications), and deployed as a drop-in replacement within the fuel system library shown in Figure 11. Both of these deployed solutions share a common format, and adopt an electric paradigm for representing the generic effort-flow concept.

Thus, although the FNM is intended for real-time applications, it would be entirely feasible to introduce a more accurate iterative-solver technique for non-real-time applications.

The scale and complexity of an aircraft fuel system's operation makes clear GUI s such as that shown in Figure 4 an essential development tool in order to allow test engineers to visualise the current configuration and operation of the system under test. In order to maintain flexibility, the FNM supports a large array of model status outputs such as fuel quantities, properties and individual network segment flows, allowing loose coupling with any desired external interface. Thus, separate GUIs have been created using Microsoft C++ based libraries (shown in Figure 4), Java graphical libraries, and the Tk graphical library (Ousterhout and Jones, 2009).

The use of automatic and manually coded C to implement the FNM allows its deployment to a variety of computational platforms, and this portability is further enhanced by the provision of a simple signal-based Application Programming Interface (API) to allow insertion and extraction of signal data, and iteration of the model by a controlling program. Thus the model has been ported to various platforms.

MODEL VERIFICATION

The avionics testing depends in part on the model fidelity. It is necessary to verify the FNM model's function to predict network flows under a wide range of aircraft conditions. The most significant aircraft operations to verify are on-ground refuel and defuel, feeding fuel to the engines, and transfer of fuel between tanks during flight. In addition to supporting tests of operational scenarios, many test conditions are artificial, deliberately exercising combinations of elements within the model, rather than representing an in-service scenario.

For the verification process, the detailed non-real-time Flowmaster models described are taken as an oracle against which tests are performed. In the case of the development of new aircraft designs, empirical aircraft data is only available long after the testing has commenced. After flight test data is generated, it is used to validate the model-based oracle, giving an indirect comparison of FNM fidelity. This approach is taken as empirical aircraft data is frequently limited in scope and the comparison method must compensate for measurement and test condition errors. Typically, the target is for flow-rates within 20% of Flowmaster data for all scenarios; an error of less than 5% is achieved for most scenarios.

APPLICATIONS

The Airbus A380 was the first and arguably most celebrated application of the FNM, and the vast increase in fuel management avionics introduced for that aircraft

provided the motivation for its initial development. The utility and flexibility of the techniques led to it being re-used for the development of similar new HIL facilities for the A400M and A350 aircraft. The economies of scale given by the technique have also led to it being retrospectively fitted to the A330/340 and A320 HIL facilities, as part of the support equipment renewal process for these long-lived programs. Here, the FNM's methods supersede the previous interpolated LUT approach, implemented with obsolete development tools and execution platforms. The expanded scope of the FNM improves the capability for HIL testing during in-service investigations and system upgrades.

HIL testing of avionics is only one of many Model Based Engineering processes within the fuel system domain, one of them being the Airbus Fuel System Modelling Environment (AFSME). This tool is used for analysis of inter-tank flow, aircraft centre of gravity control, state chart specification of fuel management functions, simulation of equipment failure, simulation of bulk-fluid thermal effects and analysis of tank inerting using nitrogen-enriched air. Although intended for real-time applications, the fidelity of the FNM can be applied to AFSME in non-real-time studies. It offers more accurate simulations for certain types of analysis if used in place of the LUT-based models currently employed.

FUTURE DEVELOPMENTS

Scope for future development broadly falls into two categories: development and evaluation of alternative flow solver algorithms and tools, and improved overall model fidelity through introduction of improved fuel library components. Some of the possible areas of investigation are described here.

The two solver libraries currently demonstrated by the FNM have been selected or developed with the goal of maintaining strict deterministic real-time capability and supporting Simulink development. Relaxing this requirement would allow a range of alternative methods to be considered. For example, non-deterministic solvers may be sufficient for non-real-time or "soft" real-time performance (i.e. practically meeting time constraints but non guaranteeing to do so). For example, a range of physical modelling libraries based on the Modelica (Fritzson 2015) language exist, and should be evaluated.

Considerable improvements to the existing solver methods are also feasible. In the first instance, the performance improvements brought by Wright's upgraded solver easily permit iteration rates to be increased from the current 4 Hertz to 10 or even 100 Hertz: this could allow application in higher-rate physical models, e.g. hydraulics, as well as allowing faster-time behaviour in Fuel systems to be represented. As described, the introduction of state-delay terms to resolve the recursive dependencies inherent in topological loops and non-linear functions creates inaccuracies and potential instabilities, and methods are being investigated to statically resolve such functions deterministically within a single iterative loop. Some progress has been made in incorporating these techniques into the FNM's solver.

Improved solver performance also offers the possibility to improve model fidelity by added more model detail (and thus complexity). For example, current tank simulations consider tanks as a single unpartitioned volume, while they are in reality composed of multiple linked bays, which introduces errors by failing to capture stepped flows and their effect on coverage of fuel measurement probes. Expanding the model to individually model these bays would largely be a matter of engineering, rather than computer science.

REFERENCES

Butz H "The Airbus Approach to Open Integrated Modular Avionics (IMA): Technology, Methods, Processes and Future Road Map" *Signal*, 2007

Dessaint L, Al-Haddad K, Le-Huy H "A power system simulation tool based on Simulink" *IEEE Transactions on Industrial Electronics*, (Volume:46 , Issue: 6), 1999

Fritzson P "Principles of Object-Oriented Modeling and Simulation with Modelica 3.3: A Cyber-Physical Approach" Wiley, 2015

Langton R, Clark C, Hewitt M, Richards L "Aircraft Fuel Systems" Wiley, 2009

Mathworks "Simulink User's Guide" Mathworks Inc. 2015

Mathworks "Simulink Coder User's Guide" Mathworks Inc. 2015

Mathworks "Developing S-Functions" Mathworks Inc. 2015

Moir I, Seabridge A "Aircraft Systems: Mechanical, Electrical and Avionics Subsystems Integration, 3rd Edition" Wiley, 2011

Ousterhout JK, Jones K "Tcl and the Tk Toolkit" Addison-Wesley, 2009

Tu Y, Lin GP "Dynamic Simulation of Aircraft Environmental Control System Based on Flowmaster" *Journal of Aircraft, Vol. 48, No. 6,* 2011

Van Rysselberghe P "Remarks concerning the Clausius-Mossotti Law", *Journal of Physical Chemistry,* 1931

AN EVALUATION FRAMEWORK FOR UAV SURVEILLANCE APPLICATIONS

Michael Ettlinger
Bilge Sarp
Christopher-Eyk Hrabia
Sahin Albayrak
Technische Universitt Berlin, DAI-Lab
Ernst-Reuter-Platz 7
10587 Berlin, Germany
E-mail: christopher-eyk.hrabia@dai-labor.de

KEYWORDS
Autonomous Systems, Agent-based Simulation, Unmanned Aerial Vehicle, UAV Coordination, UAV Surveillance

ABSTRACT

Unmanned Aerial Vehicles (UAVs) are becoming more and more popular in many different fields. One of the fields is surveillance. UAVs can have access to wide range of area, where people usually struggle. Before multiple UAVs perform such a mission, it is desirable to simulate the system on an abstract level in a virtual environment. This allows to find the perfect configuration, before starting with an expensive implementation. This project covers a configurable simulation framework that enables the implementation of different Multi-UAV algorithms as well as different project settings. Thereby the system can be modified based on various technical and business needs. The measurements make it possible for developers or system administrators, to make decisions about the system, before it is deployed and tested with physical drones. The framework is evaluated with an example business case, where a system planner considers the amount of UAVs, that shall be implemented for a given surveillance scenario in the security context.

Introduction

Currently, drones are used not only for military missions, but also for many civilian applications. For instance they are used for delivering goods, observing an area, as well as advertising services and products. Today most drones are remote controlled by people. However, as drones will be used more in the future, manual human-control for each drone will become less feasible. Therefore the control of UAVs will be taken over more and more by algorithms enabling autonomous operation.
As UAV systems can become very expensive, many UAV applications are simulated before physical implementation. To simulate such a use case, many configuration settings and algorithms are needed, which can vary dependent on the environment. In this paper we present a testing framework for the UAV surveillance domain, to measure key performance indicators (KPI) and to make decisions about configurations or algorithms based on the results. In particular we are targeting the evaluation of higher level coordination and cooperation algorithms. In the example surveillance use case we analyse how many drones should be used in a given environment with a given configuration. The simulated drones have goal-directed adaptive behaviour and they can adapt to the changes in the crowd. Since drones should move independently and need to organize themselves, their movements are inspired by swarm algorithms, which provide agents a collective and self-determining behaviour. Boid flocking algorithm is used in the testing framework as an example algorithm and it provides a basis to demonstrate such a collective behaviour.
The simulation testing framework allows comparing different configurations for different use cases. It gives the opportunity to run various algorithms and see how the behaviour of the agents changes. The system is based on the Mesa framework (David Masad and Jacqueline Kazil 2015). Mesa enables agent-based modelling and a visualization in the web browser. Mesa is a flexible playground for software developers to test different techniques, to find out most appropriate algorithm and to compare configurations. The system allows developers to compare different algorithms and business people to test different settings for different requirements.
The remainder of this paper is structured as follows, in Section Preliminaries and Related Work studies that are related to the mission of the paper are described. Section Description of the Use-Case introduces the use-case scenario and implemented specific algorithms. In Section Description of the model components of the system are explained and in Section Evaluation results of an example use case are described. The final Section Conclusion summarises the paper and highlights future steps.

Preliminaries and Related Work

Our system is a sample test framework for agent-based model UAV applications and a considerable amount of literature is available in agent-based modelling and simulation field.

Cheng et al. (2012) present a decentralized task allocation model based on both task stimulus intensity and a respond-

ing threshold. Natural behaviour of insects has influenced developing the response threshold method as they can perform complex tasks with swarm intelligence. In this study a agent-based simulation environment (SWARM simulator) was used to perform these objectives. In SWARM simulator agents can freely correlate and communicate with each other and the environment via a schedule of discrete events and messages. The mission scenario is about searching for targets in an initially unspecified area and destroy them. Effective distribution of UAVs to obtain optimum coverage of area and attack the target is tried to achieve in simulation. UAVs' actions were determined based on rules and various swarm sizes and multiple targets were implemented in two sets of simulations.

Ghnemat et al. (2008) present swarm intelligence algorithms that are inspired from natural termite nest building and ant culturing algorithm to tackle dynamical and spatial organization emergence. Agent-based modelling techniques are used for modelling. REPAST(recursive porous agent simulation toolkit) and OpenMap as geographical information system (GIS) software is used for developing the simulation. REPAST is an open source agent-based modelling and simulation platform. Some features of the tool are, it is fully object-oriented, implemented in Java, and it supports Java, C#, C++, etc.

In the study of McCune and Madey (2013), the authors propose a new agent behaviour capable of partitioning a search area by solving an optimization problem for assigning swarms efficiently to a topology. Agent-based simulations were developed to test swarm solutions. Simulations were conducted using MASON that is an agent-based simulation framework from George Mason. The simulation determines whether or not a swarm of UAVs can clean a given region based on the parameters given.

Many agent-based modelling simulations have been implemented in previous studies using various frameworks (NetLogo, Repast, or MASON). There are also many more flight simulators e.g. Microsoft Flight Simulator, X-Plane, FlightGear, etc. that can be used to simulate UAV flight dynamics (Gimenes et al. 2008), but are not suitable to simulate many UAV on an abstract management level. Moreover, using these simulators require particular domain knowledge and high expertise. Additionally they are highly computationally demanding frameworks. This system makes use of the Mesa framework to create a simulation environment. Among other simulation environments Mesa provides less complex simulation experience and is less computationally demanding. Mesa is written in Python and is therefore OS independent which makes Mesa particularly different from other simulation frameworks. It allows to develop external modules and different agents. Moreover, Python enables interactive analysis of model output data, through the IPython Notebook (Pérez and Granger 2007) or similar tools. Another significant advantage of Mesa is that it is easy to set up and use. As it comes with built-in components such as agent schedulers and spatial grids, users can quickly create agent-based models. Mesa also allows user to create visualization using browser-based interface and it enables to use Javascript to make customization in front-end. Finally Mesa's data collector allows users to conveniently analyse and export results.

Description of the Use-Case

We introduce a surveillance use case to provide a basis for the model description in the following Section Description of the model. A group of UAVs is used to perform surveillance for provocative activity on a specified group of people in a demonstration area. UAVs organize themselves, continually checking for suspicious activity in the crowd. All measured suspicious activities are reported to other UAVs and the central control system. The UAVs share their knowledge peer-to-peer every time they are within communication range.

The demonstration area represents a part of a city which has a base station and several buildings. The UAVs share their observations with the base station every time they recharge their batteries. Buildings cannot be entered neither by crowd members nor by UAVs. The base stations are connected to the control centre which collects all observations.

The following subsections provide more details about the use case specific algorithm implementations and illustrate possible development scenarios using our framework.

Crowd Member Algorithm

In our example use case, people motion is implemented based on the flocking behaviour of birds. To simulate the movement of a crowd of people, a flocking algorithm based on *Boids* is implemented. *Boids algorithm* has three fundamental rules. The first rule is cohesion where objects keep staying near neighbouring object to ensure togetherness. The second rule of the Boids algorithm is alignment that provides objects to head in the same direction. The third rule is separation which avoids collision among objects. In our system, the original Boids' rules Rydell (2014) are slightly modified and the new algorithm is called Crowd Member Algorithm.

The first rule in Crowd Member Algorithm calculates the centre of mass in a given radius to move closer to a set of agents. The second rule determines the moving away direction (in a given radius) from other crowd member agents to avoid overcrowding and the third rule calculates the average velocity of crowd members to move with a set of agents.

Depending on the position of the agents, the weights of these rules differ. For instance if an agent is somewhere alone in the grid, then the value calculated by the first rule will be more important than the value from the other rules. Conversely, in another scenario, when an agent is in the middle of the crowd, then adjusting its velocity to the average velocity of the crowd will be more important to move along with its neighbour. Therefore the location of the agent influences the assessment of these rules.

After determining the priority of rules for each agent, then possible new locations are calculated. The algorithm attempts to move the agent to the first possible location, if there is no obstacle, then the agent moves to its new location in the

first try. If there is an obstacle, then the agent tries to move to the option with the second highest priority and it continues until it finds an available place. As this is an iterative process, everything above is recalculated in each simulation step.

UAV Algorithm

The UAV algorithm is a custom-developed swarm algorithm enabling a decentralised, robust and self-organised operation of many UAV. The main goal of the UAV algorithm is to visit every location on the grid as often as possible and to minimize the age of oldest visited area (map cell). Each UAV memorizes when each cell was visited the last time and based on this knowledge, UAVs go to the oldest cell, they know of.

The algorithm allows UAVs to share their knowledge among themselves. When two UAVs meet, they compare the information they have gathered and exchange the knowledge (the last visited time of cells). The idea here is that a UAV always intends to go to the cell that has not been visited for a long time, and if a UAV visits one cell, the other UAVs do not need to go to this cell, because it has recently been visited. After knowledge sharing, the other UAVs know about this activity and discard the recently-visited-cell from their target list.

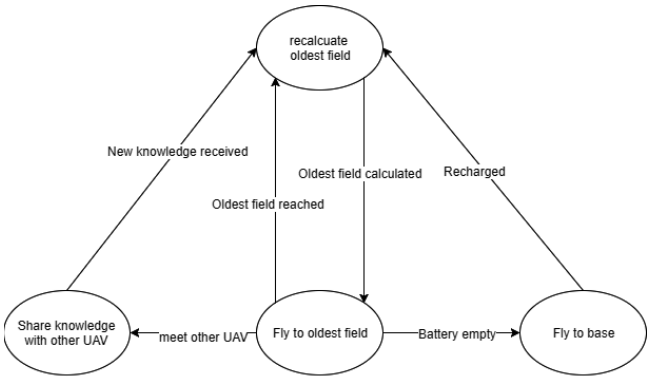


Figure 1: State diagram of an UAV agent

If two UAVs want to go to the same location, the one with highest priority goes to this cell, the other one that has a lower ID moves towards another direction. The priority is determined by the UAV ID.
Obstacles are blocking UAV movement. If there is an obstacle in the direction of UAV's target location, the UAV excludes the obstacle position from its possible steps array, and goes to the closest free cell in the target direction.
If a UAV battery level is lower than the defined threshold, it determines the base station's position as its target location, and directly moves to the base station. After recharging its battery, the UAV continues its usual surveillance activity. The possible UAV states and corresponding transitions are visualised in Figure 1.

Clustering Algorithm

In each simulation time step the UAVs make observations of the crowd. Each observation has identifying features (A picture of the crowd member could be a real life equivalent). As it is not possible in a real world application to get a perfect picture that identifies the crowd member correctly 100% of the time, this uncertainty is also implemented in the simulation. Every observation includes random noise. The magnitude of this noise can be configured in the configuration. The command centre uses all these identifying features with included noise of the observations to detect which observations portrait the same crowd member and groups them together. To group these observations, different clustering algorithms can be chosen. The implementations to choose from are perfect clustering, which groups all observation 100% correct regardless of errors, random error clustering, which is similar to perfect clustering, but randomly makes mistakes when grouping, and K-means clustering from the sklearn package.

Description of the model

The basic principle of the architecture of Mesa is modularity and it has a set of built-in components that can be combined and modified to create various models. The architecture consists of the modules Modelling, Analysis, and Visualization. Furthermore, the Mesa framework supports a grid as environment representation. UAVs, crowd members and obstacles are modelled as agents. Mesa allows to display and manage the implementation of different control algorithms. Even though Mesa is providing a suitable foundation for many agent-based simulation environments, it was necessary to develop certain extensions. Therefore the framework was forked and extended. To improve the visualization, a templating engine was created, which allows to customize the HTML. Bootstrap is used to display the components. For various ineffectiveness of Mesa functions, caches were created to increase overall performance. Mesa normally displays only mobile agents. To add obstacles like buildings, the visualization was also extended.

Modules

The package algorithms includes swarm algorithms and clustering algorithms of the model. After every step of every agent, data about the observations are saved in their own instance of those algorithms. The swarm algorithms are split up in *UavSwarmAlgorithm* and *CrowdMemberSwarmAlgorithm* classes. These algorithms are explained in detail in Section Description of the Use-Case.
The class *EnvironmentModel* represents the environment as well as all actors in the environment. The environment also holds an instance of the configuration, which allows to adjust various settings of the simulation. Height, Width and grid are all generated from a map, which is explained later. *LastError* as well as *lastCrowdDetectionRate* are statistical evaluations which can be displayed for the user of the system. Schedule

is an interface to the user of the system. It allows to navigate through the simulation steps in a timeline and to execute individual steps. The Data Collector allows statistical evaluation of the data.
CommandCenter is the interface to the user of the simulation. It holds all observations and the estimated grid. In a real world scenario, law enforcement can use this information to decide if they should act on the demonstration.
The system has two different grids, one representing all agents in the system and the other one representing the system perception, which can visualise either individual agents or the global system view.
The system can be configured using a configuration file, for instance with parameters such as width of map representation in the browser, crowd member detection range, threshold describing the upper bound of the error between new identifying features to still be considered equal, etc. The mapping system allows to describe the environment map in a file, which then is used to generate all the initial agents and obstacles. In the system config and test config, map, and algorithm files are interchangeable.
Three integration tests suits have been created to evaluate our system. The first one is for general test, the second suite is for test configurations, the third one was created specifically for the case discussed in Section Evaluation.

Agents

All agents of our system inherit from the Mesa-Agent class.
CrowdMemberAgent represents one person in the crowd. This agent walks around on the grid. The attribute *willingnessToMove* is a random factor representing how likely the crowd member is to move at any given point. Every crowd member has a *baseSuspiciousness* attribute which is a factor, which indicates how provocative a person is. When it is higher it acts more often suspiciously. UAVs make many observations to be able to make a guess about the suspiciousness of *CrowdMemberAgents*. *IdentifyingFeatures* is an array of 100 random variables. These random variables represent identifying features which could be a picture of a face in a real world example. At each observation of the crowd member the UAV gets representation of those features with added noise to simulate a real camera object detection system. The UAV can then use those features to categorize observations and group observations of the same person together. At any point in time the crowd member can either act suspicious or not. This feature is represented in the *currentSuspiciousness*.
ObstacleAgent represents an obstacle in the environment, like building or other concrete structures. *ObstacleAgents* are passive agents. They are represented in the grid, but do not interact with other agents.
UavAgent represents one UAV in the system. Throughout the run time, the UAVs fly around according to their swarm algorithm and observe all crowd members, which again act according to a swarm algorithm, in a specified range. At every step of the model, the method *detectCrowd* is called. In this call observations of all surrounding crowd members will be generated and saved.
The different state variables of the *CrowdMemberAgent*, *UavAgent* and the *EnvironmentModel* are described in the Tables 1, 2, and 3.

Evaluation

The testing framework allows to test a variety of KPIs in various environments. This evaluation describes an example business case in detail and then continues to give an overview of other KPIs, which can be tested with the system.

Comparing amount of UAVs in given setting

As an example application of the system, we created an environment, where the map, the general configuration and the algorithms stay the same. Only the amount of UAVs is evaluated with different settings in order to find an optimal amount of UAVs for the given use case. This test allows administrators of the system to test how well different amounts of UAVs perform in a given environment applying a certain algorithm. The assumption of this test is, that the first UAV has the highest performance. For each additional UAVs, the performance gain decreases. At the end of the test, an administrator should be able to decide, which amount of UAVs is in the budget, while also producing good results in surveillance. The amount of UAVs varies from 1 to 10. Each test case is executed 100 times and then the results mean is taken. Each test case encompasses 2000 simulation steps of the system.
In the following we explain available KPIs in our system and illustrate them with results of our example use case.

KPI: Has seen all Crowd members
The KPI *Has seen all Crowd members* shows the required number of steps to detect all Crowd members at least once.

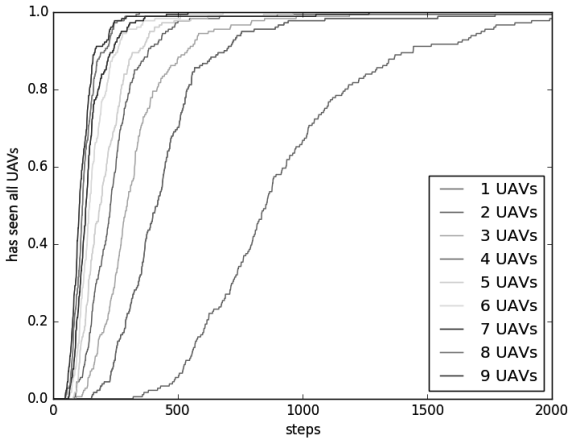


Figure 2: KPI: Has seen all Crowd members

As visible in Figure 2, 1 UAV usually takes 1000 steps to half the crowd members. By adding a second UAV, the per-

Variable	Description
baseSuspiciousness	Variable indicating how likely a crowd member acts in a suspicious way.
currentSuspiciousness	Observed suspiciousness level that is changing over time

Table 1: State Variables of the CrowdMemberAgent.

Variable	Description
map	n x m matrix: Obstacles in map
observations	Observations of an UAV that have been reported to the base station.

Table 2: State Variables of the Environment Model.

formance increases drastically to around 400 steps to see half of the crowd. Adding further UAVs only provides a marginal gain in performance. After having 6 UAVs, the additional benefit is negligible. Therefore the ideal number of UAVs to see all UAVs as soon as possible is between 2 and 6.

KPI: Occurrences per crowd member

One UAV normally takes at least 500 steps to detect all crowd members, therefore the line only increases from 0 at around step 500 as shown in Figure 3. The more UAVs are used, the sooner all crowd members are detected, and therefore the earlier the line increases from 0. After the initial detection, all graphs show an almost linear increase. Therefore it can be assumed that, the more UAVs are added to the system, the more observations are made. 6 UAVs produce around twice as much observations as 3 UAVs. As the increase is only almost linear, further information can be inferred. Considering max occurrences (Figure 4), we see that after 2000 steps, 9 UAVs produce around 1790 observations. 1 UAV produces around 230 occurrences. If it was fully linear, 9 UAVs should produce 2070 occurrences. In a qualitative analysis of the visualization, many cases can be seen where UAVs disturb the flight path of other UAVs. This is especially visible when multiple UAVs want to charge at the same time. Therefore it can be inferred that the more UAVs there are, the fuller the airspace is and the more UAVs are disturbed by other UAVs. Reducing the number of UAVs helps to prevent this effect.

KPI: Age of last seen oldest field

The KPI *Age of last seen oldest field* is the maximum visiting age of all known cells. This gives insight, in how long it usually takes for UAVs to detect every single cell once.

Figure 5 shows the age of the last observation for the oldest visited cell in the grid. The first time the graph diverts from the straight diagonal trend marks the point when the system has detected all locations. The height of the graph gives insight of how long it usually takes to see every cell again.

As visible in Figure 5, 1 UAV usually takes around 750 to 1000 steps to detect all fields. Adding a second one reduces this time to only around 400 steps. After having 3 UAVs, the duration reduction is only minimal. The height of the graph gives an insight into how long it usually takes to recheck every cell again. A straight line at age 300 indicates that every cell is seen at least every 300 steps of the system.

Every configuration shows a constant increase in age describing a cell detection rate decrease. The cell detection rate for 9 UAVs decreases after around 1500 steps. 1 UAV alone has a steady detection performance decrease throughout the first 2000 steps.

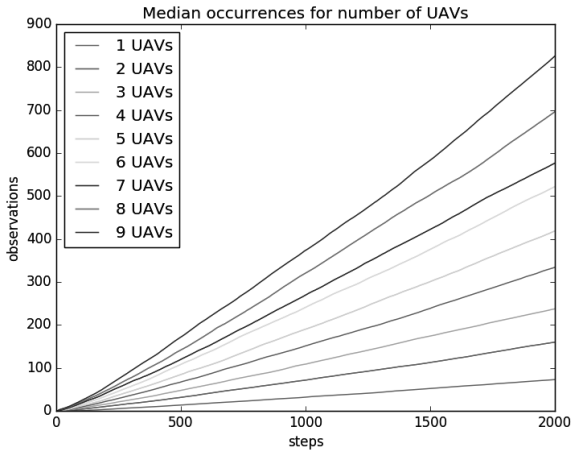


Figure 3: KPI: Median Occurrences per Crowd Member

This comes mostly from a limitation of the chosen grid framework. The chosen Grid implementation of Mesa allows only to have a two dimensional field of agents. Therefore crowd members can block UAVs, because UAVs cannot fly over crowd members. Without this limitation the graph should be approximately linear as crowd collisions would not be possible. This limitation is planned to be addressed in future.

Leaving out the decrease of performance, it can be seen that 2 to 5 UAVs produce the best cell observation performance. Less than 2 UAVs show a very bad performance. More than 5 UAVs only results in a marginal performance increase.

KPI: Suspiciousness accuracy

Suspiciousness accuracy is the main KPI of our use case. It combines all other KPIs and describes how well the system can detect which crowd members act peaceful and which are rioting. The suspiciousness is a factor that UAVs detect by watching crowd members. The suspiciousness accuracy gives insight into how well the observations reflect the actual base suspiciousness.

Variable	Description
observationAgeGrid	n x m matrix: time stamp where UAV thinks each cell has been seen the last time
penaltyGrid	n x m matrix: Information about obstacles positions. Also contains information where other recently met UAVs are flying to, to avoid two UAVs flying in the same direction.
observations	All observations that have been made by the UAV and observations that have been made by other UAVs that have shared their information.
State	Current UAV state, e.g. flying to the base station or executing surveillance operation.

Table 3: State Variables of the UavAgent.

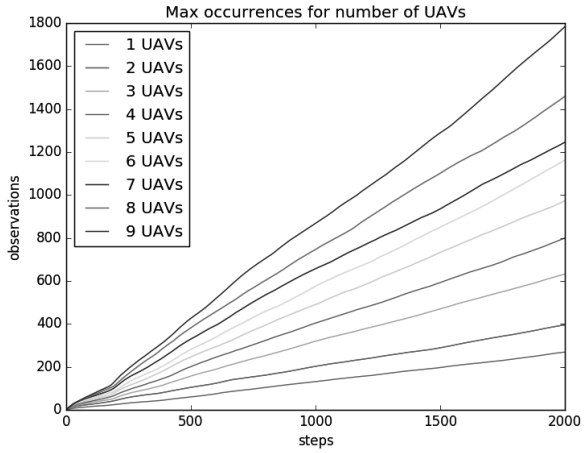


Figure 4: KPI: Max. Occurrences per Crowd Member

The Figure 6 shows that it takes 1 UAV around 1000 steps to reach 90% accuracy. Having 2 to 4 UAVs increases this accuracy a lot. After adding even more the accuracy does not increase much. Therefore the best amount of UAVs to detect suspiciousness lies between 2 and 4.

UAV count evaluation conclusion

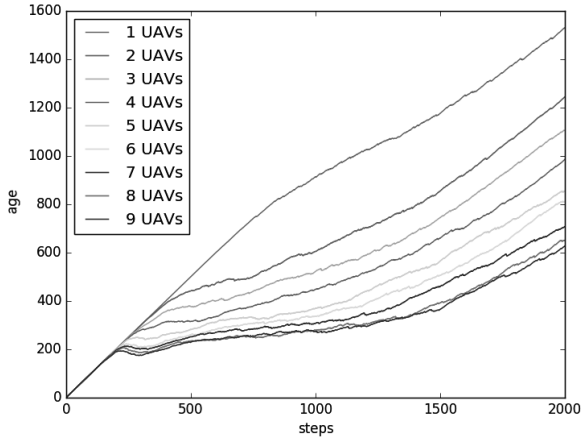


Figure 5: KPI: Age of oldest Field

All described KPIs, show that at least 2 UAVs are needed to have an acceptable prediction rate. Some KPIs show performance gains up until 4, 5 or 6 UAVs. More than 6 UAVs does not show strong performance gains in any KPI. As seen in Section KPI: Occurrences per crowd member fewer UAVs work more efficient, as they do not disturb other UAVs. Also fewer UAVs would result in lower cost. From the KPIs for the given configuration, map and algorithms 2 to 4 UAVs would produce the best results.

Performance test

The performance test allows to test the performance that is used for each simulation module. This allows to compare the simulation performance of crowd member step, UAV step, UAV observation and command centre calculation.

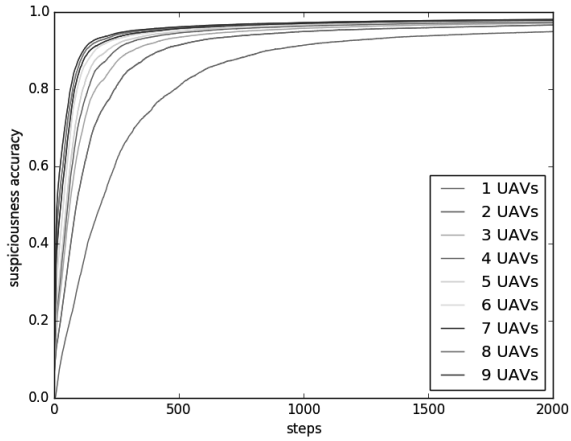


Figure 6: KPI: Suspiciousness Accuracy

The performance test in Figure 7 gives clear insight into the run time complexities of various parts of the simulation. The configuration, map and algorithm stayed the same as in Section Comparing amount of UAVs in given setting. As we received the best results with 2 to 4 UAVs in the former tests we have chosen the amount of 3 UAVs for the following tests. While the duration of crowd member movement, UAV movement are linear in run time, the other two values seem to be more problematic. Command centre calculation increases steadily, while the UAV observation duration increases quite rapidly. The first problem comes very likely from the fact that clustering of observations takes longer when more observations need to be taken into consideration. The latter

problem seems to originate from the same underlying issue. Inserting new observations takes longer, when there are already a lot of observations present. When UAVs meet and share their knowledge, this is also more resource intensive, when more observations are taken into account.

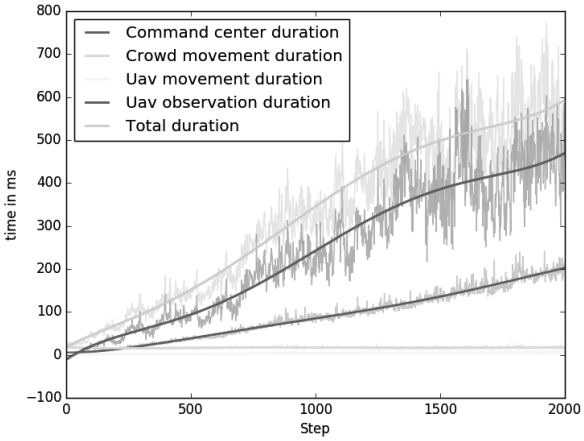


Figure 7: Execution Time Comparison of different Modules

An algorithm with a quadratic run time complexity can never move from simulation to a real world scenario, as it would be impossible to scale up in time. A solution could be an observation cache. Every few steps all observations that belong to the same crowd member could be merged into one observation with a weight factor. This would make it easier for the clustering algorithm to split all observations by crowd members. Such an enhancement reduces the overall run time complexity to linear, which then allows a deployment in a real world scenario. This finding illustrates how our framework can help to detect bottlenecks in a particular implementation.

Further test options

Our testing suite allows to adapt tests for a variety of variables. To compare simple settings, the test configuration of the *ConfigurableTestSuite* can be used. More complex tests can be added by creating an additional test suite, as demonstrated with *UavCountComparisonTestSuite* in Section Comparing amount of UAVs in given setting.

Conclusion

The purpose of our system is to create a higher level Multi-UAV simulation system that allows for the evaluation of different algorithms, configurations for UAV management. Our configurable simulation framework can be customized based on different technical and business needs. From a developer point of view, different coordination and self-organisation algorithms can be implemented and compared. Moreover, system planners can easily modify the configuration to apply different evaluation and statistic plans. Based on these results users can make decisions and plan further actions.
In our example evaluation about security surveillance we have demonstrated how the framework can be used to determine an optimal amount of UAV for a given setting. Moreover, this enabled us to illustrate how a performance bottleneck in the implementation can be detected.
Some processes (such as observing agents in a given radius) take a long time in the Mesa framework, which was used as foundation for our work, because Mesa does not make use of any caching strategy. Especially when displaying new simulation steps, the missing render cache decreases the performance. In order to solve this problem, a custom caching method was implemented which in the end significantly increased the overall performance of the system. Finally, as Mesa's web interface is not configurable enough, styling of the front end could only be done using complicated workarounds. In future work, we would like to address mentioned limitations, e.g. implement more advanced caching methods in order to get maximum efficiency from the Mesa framework and evaluate more coordination and self-organisation frameworks.

REFERENCES

Cheng H.; Page J.; and Olsen J., 2012. *Dynamic Mission Control for UAV Swarm via Task Stimulus Approach*. *American Journal of Intelligent Systems*, 2, no. 7, 177–183.

David Masad and Jacqueline Kazil, 2015. *Mesa: An Agent-Based Modeling Framework*. In Kathryn Huff and James Bergstra (Eds.), *Proceedings of the 14th Python in Science Conference*. 53 – 60.

Ghnemat R.; Bertelle C.; and Duchamp G.H., 2008. *Agent-based modeling using swarm intelligence in geographical information systems*. In *Innovations in Information Technology, 2008. IIT 2008. International Conference on*. IEEE, 69–73.

Gimenes R.; Silva D.C.; Reis L.P.; and Oliveira E., 2008. *Using Flight Simulation Environments with Agent-Controlled UAVs*. In *Autonomous Robot Systems and Competitions: Proceedings of the 8th Conference*.

McCune R.R. and Madey G.R., 2013. *Agent-based simulation of cooperative hunting with UAVs*. In *Proceedings of the Agent-Directed Simulation Symposium*. Society for Computer Simulation International, 8.

Pérez F. and Granger B.E., 2007. *IPython: a system for interactive scientific computing*. *Computing in Science & Engineering*, 9, no. 3.

Rydell A., 2014. *Multi-agent swarm control in virtual worlds*. Master's thesis, Linkping University, Department of Computer and Information Science, The Institute of Technology.

ENGINEERING SIMULATION

APPLYING THE MODEL-DRIVEN ARCHITECTURE APPROACH TO DYNAMIC STRUCTURE APPLICATIONS

Min ZHU, Clément FOUCHER, Vincent ALBERT, Alexandre NKETSA
LAAS-CNRS, Université de Toulouse, CNRS, UPS, Toulouse, France
Email: {Min.Zhu, Clement.Foucher, Vincent.Albert, Alexandre.Nketsa}@laas.fr

KEYWORDS
Model Driven Architecture, Reconfigurable Architectures, Meta-Model, Simulator

Abstract

Model-Driven Architecture (MDA) is a system engineering approach which consists in separating the model description from the execution platform. It allows building a model without detailed knowledge of the target platform, as well as retargeting the execution platform without changing the model itself.
We present a meta-model called Partial Reconfigurable DEVS (PRDEVS) that is able to represent dynamic structure changes of a model. We base our approach on the DEVS formalism, which is modular and hierarchical. Our description paradigm differs from the previous DEVS-based dynamic meta-models in that it explicitly deals with adding and removing components. This approach is closer to the general reconfigurable embedded system design methodology. Both a software and a FPGA-based hardware platform are considered as dynamic execution platforms.

INTRODUCTION

The Discrete Event System Specification (DEVS) formalism introduced by Zeigler et al. (2000) is a strong mathematical foundation for specifying hierarchical and modular models. The DEVS formalism allows to build discrete event systems and provides algorithms for simulation. DEVS models are made of atomic components, which define a behavior, and coupled components which can hold several other components and describe the way they are connected. As the initial DEVS formalism was not designed to handle structure changes, either in model composition or communication, various extensions were proposed to address these dynamic systems. However, our point of view on existing formalisms which allow to describe reconfigurable systems is that they are either too high level, making it difficult to apply to real systems, or too deeply linked to the execution platform.
The Model-Driven Architecture (MDA) approach by Object Management Group (2016), derived from Model-Driven Engineering (MDE), consists in separating the application model description from the execution platform. This brings various benefits, such as allowing the teams working on an application to be independent from the ones working on the platform, or enabling deploying an application built from a single model on various platforms. A complete MDA specification consists in a Platform-Independent Model (PIM), one or several Platform-Dependent Models (PDM), and sets of interfaces correspondence to allow building a Platform-Specific Model (PSM) by merging PIM with PDM, as depicted on Figure 1.

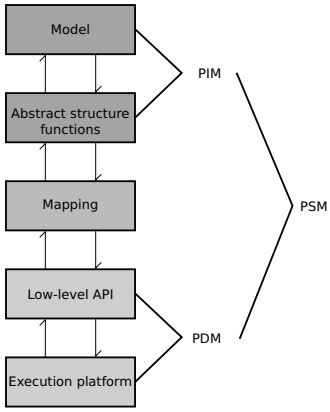


Figure 1: MDA Structure of PRDEVS

In this paper, we propose a DEVS-based PIM formalism able to describe applications whose structure can evolve dynamically. The aim of this formalism is to comply with hardware-reconfigurable architectures such as FPGAs, without however restricting the execution to these platforms. We thus take in consideration the generally observed approach used when building such systems, but remain sufficiently high level to target any platform supporting dynamic architecture change of the application, such as software-based implementation. We also introduce PDM candidates in order to show the compliance of the PIM with such architectures.
First, we present the existing work related to DEVS and its extensions. Then the background of DEVS formalism and reconfigurable hardware are presented. Afterwards, the PRDEVS model at PIM level is presented with its syntax and semantics. A software oriented implementation with its possible PSM is then presented. Finally, we conclude this article and present a view of what remains to be done.

RELATED WORK

Zeigler et al. (2000) initially introduced the DEVS formalism in the late 70's as a way to build models with a discrete-event approach using a mathematically defined formalism. DEVS was later extended with Parallel DEVS (PDEVS), and we now reference the initial DEVS formalism as Classic DEVS (CDEVS). In this article, the acronym DEVS thus refers to the general DEVS ecosystem rather than to the original CDEVS formalism.

DEVS is a hierarchical set of components of two kinds: atomic components define a behavior while coupled components gather and link other components, either atomic or coupled. The DEVS formalism is inherently static, and dynamic structure behavior can only be emulated, e.g. using a selector to enable or disable models over time. Several extensions have been proposed aiming at dynamically adapting the models structure during the simulation.

DSDEVS, defined in Barros (1997), is based on a 4-tuple network structure where atomic components can connect directly with other atomic components by a set of influencers I. The network executive χ is a specific component whose state represents the network structure. The γ function allows to obtain the network structure from the current state of χ. χ thus takes responsibility for all changes of model structure, meaning that components in the model can not take decision on structural adaptation. DSDE (Barros (1998)) is a parallel version of DSDEVS.

The principle of dynDEVS as described in Uhrmacher (2001) is that each atomic component has its own model transitions function ρ_α which controls its own structural transformation. At coupled component level, the equivalent function is the network model transition function ρ_N. However, dynDEVS assumes a static set of ports which is not adapted for most of the dynamic applications. Uhrmacher later developed $\rho-$DEVS, a dynDEVS variation supporting dynamic ports.

The formalisms mentioned above stay at theoretical level: they propose an abstract simulator and can be adapted to an actual execution environment.

There are recent works like RecDEVS which consider hardware models of computation (MoC) together with DEVS. RecDEVS (Madlener (2013)) proposes a model based on DEVS for final use on reconfigurable hardware like FPGA. In RecDEVS the system executive C_χ is in charge of the structure changes. However RecDEVS takes into account some hardware specificities from the beginning, like component communication relying on a bus structure with an address notion. This limits the model to a use on the target platform defined by RecDEVS. Thus, there is no separation between PIM and PDM. There are other limits on the meta-model itself, like the fact that a component deletion can only be triggered by the component itself. Eventually, there is no final implementation on FPGA, the workflow only goes to SystemC simulation.

BACKGROUND

Parallel DEVS (PDEVS)

PDEVS (Zeigler et al. (2000)) is the root formalism for DEVS extensions dealing with parallelism. It defines two kinds of components: atomic components, which are the base elements defining a behavior, and coupled components, which gather various other components and define their relationships. PDEVS allows different components to evolve simultaneously and provides resolution mechanisms to deal with conflicting simultaneous events.

Atomic Models

PDEVS defines an atomic component as an indivisible unit implementing a behavior. It can evolve in reaction to an external event (external transition), or when a timeout occurs (internal transition). The formal definition is as follows:

$$M =< X, Y, S, s_0, \delta_{ext}, \delta_{int}, \delta_{con}, \lambda, \tau > \text{ where}$$

X $= \{(p, v) \mid p \in InPorts, v \in X_p\}$ is the set of input ports and values, where
  $InPorts$ is the set of input ports
  X_p is the set of allowed input values for port p

Y $= \{(p, v) \mid p \in OutPorts, v \in Y_p\}$ is the set of output ports and values, where
  $OutPorts$ is the set of output ports
  Y_p is the set of possible output values for p

S is the set of sequential states

s_0 is the initial state of the component

δ_{ext} : $Q \times X \to S$ is the external state transition function, where
  $Q = \{(s, e) \mid s \in S, 0 \leq e \leq \tau(s)\}$ is the set of total states, with e the time elapsed since latest transition

δ_{int} : $S \to S$ is the internal state transition function

δ_{con} : $Q \times X \to S$ is the confluent transition function

λ : $S \to Y$ is the output function

τ : $S \to \mathbb{R}^+_{0,\infty}$ is the time advance function

When an event X occurs on an input port, the external transition function δ_{ext} is called which may result in a state change. The time advance function τ associates a time to each state which, when reached, triggers the output function λ then the internal transition function δ_{int}. Note that τ accepts both 0 and ∞ as values. When simultaneous external and internal events occur, the confluent function δ_{con} is called instead of δ_{int} or δ_{ext} to solve the conflict. δ_{con} can be as simple as calling δ_{int} or δ_{ext}, which is a way of prioritizing between these two functions, or can be a totally different function.

Coupled Models
A coupled component is a way of linking other components. Externally, it behaves like an atomic component and thus can be used in another coupled to form a hierarchical model. Its definition is:

$$N =< X, Y, D, \{M_d\}, EIC, EOC, IC > \text{ where}$$

X, Y	as defined for atomics
D	is the set of components names
$\{M_d\}$	is the set of components in this coupled, with $d \in D$
EIC	is the external input coupling function
EOC	is the external output coupling function
IC	is the internal coupling function

The three coupling functions directly link ports between them:

EIC	links $p_N \in InPorts_N$ to $p_d \in InPorts_d, d \in D$
EOC	links $p_d \in OutPorts_d, d \in D$ to $p_N \in OutPorts_N$
IC	links $p_a \in OutPorts_a,\ a \in D$ to $p_b \in InPorts_b$, $b \in D, a \neq b$

Parallel Dynamic Structure DEVS (DSDE)

DSDE (Barros (1998)) defines a specific component, χ, whose state encodes the structure of the network, i.e. the current network structure can be obtained at any time from χ state using the structure function γ. A transition of χ thus can represent a change of the network structure. The DSDE component is defined as:

$$DSDE_N =< X_N, Y_N, \chi, M_\chi > \text{ where}$$

N	is the network name
X_N, Y_N	$\equiv X, Y$ in PDEVS
χ	is the name of the dynamic network executive
M_χ	is the model of the executive χ

The model of the executive χ is an extended definition of an atomic model defined as:

$$M_\chi =< X_\chi, S_\chi, s_{0,\chi}, Y_\chi, \gamma, \Sigma^*, \delta_\chi, \lambda_\chi, \tau_\chi > \text{ where}$$

γ	$: S_\chi \to \Sigma^*$ is the structure function
Σ^*	is the set of network structures

According to this definition, the set of components is defined, but their state is not. Barros thus defines the new components states after a χ transition to be equal to the same components state before transition (plus time advance) if the component existed, or to be the initial state if the component didn't exist.
In this definition, χ is the only component allowed to change the network structure. Moreover, the connections between the components of the network are also defined by χ state, i.e. a simple change of connector without affecting the atomic components themselves must be treated as a χ transition.

RecDEVS

$$N_{Rec} =< X_{ext}, Y_{ext}, D, C_\chi > \text{ where}$$

$D:$	Set of all available DEVS components
$C_\chi:$	is the network executive which is a DEVS atomic

RecDEVS defines an unique identifier ID for each component. The creation of new RecDEVS components consists of a fixed sequence of messages as follows:

* if the component C^{ID}_{orig} wants to create a new component of type $d \in D$, it sends a message (C^{ID}_{orig},C_χ,(new d)) to the network executive.
* C_χ receives the message and performs an external transition δ_{ext}. This will create a new RecDEVS component C^{id}_d and add it to the list of instantiated components
* A confirmation message (C_χ, C^{ID}_{orig}, (confirm C^{id}_d)) with the address of the new component is then sent to the originator.
* Starting from the reception of the confirmation message, the originator can address the newly created component.

FPGA

A Field-Programmable Gate Array (FPGA) is an integrated circuit designed to be configured to form a complex digital circuit. The majority of FPGA architectures carry out combinatorial logic using Lookup Tables (LUTs) (Koch (2012)), associated to flip-flops to form sequential circuits. Several LUTs and flip-flops form a base reconfigurable resource (e.g. configurable logic block (CLB) in Xilinx technology) which is the smallest reconfigurable unit in a FPGA. Configuration of the underlying structure in the FPGA can be stored using various technologies. The Static RAM-based FPGA is a common FPGA architecture. There are also flash-based FPGAs. A bitstream is the configuration data to be loaded on board to implement the desired logic.
Partial Reconfiguration (PR) (Feist (2012)) is the ability to dynamically modify the architecture hosted on the FPGA by loading a partial bitstream (i.e. a configuration of a specific area of the FPGA) while the remaining logic continues to operate without interruption.

PRDEVS PIM: SYNTAX AND SEMANTICS

We propose a PIM syntax based on PDEVS and inspired by RecDEVS. The platform-independent model is, as stated from the name, a model which is built to represent an application, without necessary knowledge of the simulation environment or of the target platform that will run the application or the application simulation. Thus, our PIM syntax must be able to represent dynamic structure models, but it shouldn't assume anything about how the structure changes will actually be

applied. This is the first difference with RecDEVS, as the RecDEVS meta-model makes no distinction between PIM and PDM. The target architecture is assumed from the model definition, notably with the use of address notions within the models.

PRDEVS PIM Abstract Syntax

A PRDEVS is a model which contains all required information about components, structure, and allows for structural changes. Though we use PDEVS and RecDEVS as a base reference, we slightly rewrite some definitions to clarify specific points. A major difference with RecDEVS is that we do not use a specific component like C_χ which stores the network structure in its state: we directly manipulate the sets, adding and removing elements. The first motivation for this approach is to be closer to the current engineering approach to describe dynamic systems. This is slightly equivalent to the software notion of new/delete instructions for object manipulation. Moreover, this way of doing offers the ability to reach structure states which may not have been predicted when first designing the system, allowing for auto-adapting systems to be more flexible.

Concerning the D set, on one hand PDEVS defines D to be *the set of components names*, i.e. a list of all the names of the components *inside* the coupled. On the other hand, RecDEVS definition of set D is "a list of available component names", and this set is compared to a list of components *types*. They both define the set $\{M_d \mid d \in D\}$, which contains the components themselves. In our component sets-based description, we rather directly manipulate the components sets, and we think this description is redundant, as one can be obtained from the other. So we chose to merge these two sets, so that the D set directly contains the components themselves. Instead of storing the names of the components, we rather use the notions of *identifiers* and *types*.

The *identifier* follows the notion introduced by RecDEVS where an identifier $ID \in \mathbb{N}$ is attributed to each component. For the *type* notion, we can make the connection with object-oriented programming, where there can be various instances of a *class*, we call *objects*. Here, the notion of *type* is equivalent to *class*, i.e. it defines a component structure and initial state, but there can be various components ($\equiv$ *objects*) sharing the same type with a different state. The identifier is then used to differentiate the components. We use here a definition close to RecDEVS but formalize the notation: we use T as the list of defined types, i.e. a list of components types which can be instantiated.

A PRDEVS component then has an identifier, which is unique and dynamically defined, and a type, which can be shared.

Main PRDEVS Component

The dynamic structure ability relies on a library of available components, each being of a specific type, which can be added to the system. The library is defined as $L = \{C_t \mid t \in T\}$. Components in the library expose a null identifier, as it is defined on instantiation.
Components in use still have a type $t \in T$, but also an identifier id, and are noted C_t^{id}. The notation can be simplified to C^{id}. Unlike RecDEVS, we do not restrict $id \in \mathbb{N}$: although a PSM implementation will probably have to impose such a restriction, the PIM doesn't require so. We thus define an arbitrary set ID which contains the allowed identifiers.

$$PRDEVS =< L, C^{Top} > \text{ with}$$

L $= \{C_t \mid t \in T\}$, library of available components
C^{Top} a coupled containing the application structure

To have a clear and simple definition, we do not include the sets T and ID in the definition. Indeed, T can be retrieved from L using the definition $T = \{t \mid C_t \in L\}$, while ID can be any arbitrary set. We also define the set ID_{PRDEVS} which contains all the identifiers of the components in C^{Top}, regardless of the hierarchy.

Coupled Component

A coupled component will be very alike PDEVS definition:

$$N^{nid} =< X, Y, D, EIC, EOC, IC > \text{ with}$$

D $= \{C_t^{id} \mid t \in T, id \in ID\}$ the set of components contained in the coupled

All of the sets defined here can be linked to a specific component by displaying its identifier, e.g. X^{nid} refers to the set of inputs X of coupled component C^{nid}. A coupled component C^{nid} does not have a specific type, as the component structure can change during execution when a component is added to or removed from D^{nid}.
For convenience, we define a few additional sets:

ID^{nid} $= \{id \mid C^{id} \in D^{nid}\}$, the set of all identifiers of components in the coupled whose identifier is nid
D_N the set of all coupled in the PRDEVS
ID_N $= \{id \mid C^{id} \in D_N\}$, the set of coupled identifiers

Atomic Component

As our formalism deals with structure changes, the atomic components definition adds a structure change function to the PDEVS atomic formalism:

$$M_t^{mid} =< X, Y, S, s_0, \delta_{ext}, \delta_{int}, \delta_{con}, \lambda_{SC}, \lambda, \tau > \text{ with}$$

λ_{SC} $: S \to SC$ the structure function with
$SC = \{addComponent(), removeComponent(), addPort(), removePort(), addConnection(), removeConnection()\}$ the list of structure change functions.

The remark on sets identifiers applies to atomic components, e.g. S^{mid} is the set of states S of component C^{mid}.

As for coupled, we define the following sets:

D_M the set of all atomics in the PRDEVS
ID_M $= \{id \mid C^{id} \in D_M\}$, the set of atomics identifiers

PRDEVS Implementation

The L set can be represented as a list of available models, which are defined by the modeler or provided by a predefined library. The list must match a type name and a component. This way, when adding a component to a model, only its type must be provided, and the list is used to retrieve model information.

Components Common Characteristics

The coupled and atomic models share two properties: the Ports set and the identifier. Any number of ports can be present on a component.

The main difference with RecDEVS is that they use the identifier to manage communication between components on a message-passing paradigm. As we separate the implementation from the high-level model, we do not presuppose of a communication scheme in the PIM.

The ports of a component are implicitly defined by the couples (p, v) of X and Y sets. However, a specific definition can be derived to formally identify the ports as mathematical objects. This representation of ports as independent objects, i.e. not only as names belonging to a set, and whose allowed values are defined by the X and Y sets, is easier to manipulate.

The ports of the components must be uniquely identified, but only among a component. Thus, the name of the port on the component is sufficient to identify it uniquely, as long as the component itself has a unique identifier. Moreover, a port has a direction (input/output), and a set of available values. We then introduce a definition of what is a port:

$$P^{id}_{Name} =< Name, Dir, Type > \text{ with}$$

id the identifier of the component
$Name$
$\in InPorts^{id} \cup OutPorts^{id}$
$Direction$
$\in P_{dir} = \{in, out\}$
$Type$ $\in P_{type}$

The set of allowed types P_{type} can be defined as a set of allowed definition sets, and will most likely contain $\mathbb{N}$, $\mathbb{R}$, $\mathbb{B} = \{True, False\}$, etc.

Concerning the connection between two ports, it must define a source and a sink. The data type does not have to be recorded, but type match between source and sink should be checked when the connection is created.

Coupled Models

The coupled models are actually sets of sets: a set of components, and a set of connections between ports.

By using the tree representation for holding the structure, the set of components D is represented by the children of the node. Using the previous definition of ports, the X and Y sets can be better represented as sets of ports. Thus, X and Y will be merged into a list called "Ports" containing ports as defined previously. The remaining three sets EOC, EIC and IC are represented by a list of two elements: a source and a sink.

Atomic Models

The atomic models are leafs of the model tree, so they can be defined as in the abstract syntax.

Structure Change Functions Exposed by the Simulator

Unlike RecDEVS, we do not restrict which component can call a structural change function. Indeed, RecDEVS states that a component can only delete itself, not another component. The main justification provided is that it avoids accidentally deleting a component which is still in use. By only allowing self-deletion, the component can announce its own deletion to linked components before committing deletion. But this approach doesn't seem to be a good answer to this issue. Indeed, deleting a component which is still in use in an application may be a conception error. Restricting the remove call to the component itself does not solve the case where the component itself is badly defined, and forget announcing its deletion to some of related component. Imposing such a constraint do not avoid errors, so the restriction is irrelevant. We believe the application correctness is up to the modeler, and to avoid such errors, applications should be checked for correctness, e.g. using formal methods.

By allowing any component to call structure change functions, we let the modeler decide how to handle its structure changes: all components can be autonomous and directly trigger the functions, or there can be one or more components in the model which are in charge of the structure, only them being able to call these functions.

Most functions defined here could possibly fail in some circumstances, but we do not want to handle exceptions or errors in this early definition, so we assume their use is made with correct parameters. Error-checking will be part of future work.

Structure change functions have a different priority level than other messages in the simulator: the simulator has a list of pending SC tasks and the list is executed only at the end of an imminence cycle. As a first approach, we chose to execute all SC functions in zero time relatively to the simulator, i.e. the simulator is paused while the structural changes are carried on. In future work

however, we are planning to allow structural changes to be applied while the simulation is running in order to allow taking full advantage of hardware partial reconfiguration technology. This will require structural checks on the model such as making sure that a newly added component will not be required for simulation until it is fully operational. This can be carried on by separating the SC function call from its return, obtaining of the new identifier on a separate external transition of the component which called the add function.

• addComponent : $T \times ID_N \rightarrow ID$
Adds a new component into an existing coupled component. The new component type and the hosting coupled ID are provided as parameters. The identifier of the new component is returned. The abstract function $getAvailableId : \emptyset \rightarrow ID$ determines an available identifier in ID. Abstract function $getNewComponent : T \rightarrow D$ returns a new component from the library matching the given type, while $getExistingComponent : ID_{PRDEVS} \rightarrow D$ returns an existing component from its identifier, as displayed on Algorithm 1.

input : $t \in T$; $id_{host} \in ID_N$
output: $id \in ID$
Data: $newId \in ID$; $newC \in D_M$; $hostC \in D_N$

newId ← `getAvailableId()`
newC ← `getNewComponent(`t`)`
newC.id ← newId
hostC ← `getExistingComponent(`id_{host}`)`
hostC.D ← hostC.D ∪ {newC}
return newId

Algorithm 1: addComponent procedure

• removeComponent : $ID_{PRDEVS} \rightarrow \emptyset$ Removes the existing component, whose identifier is passed as a parameter, from the PRDEVS. The abstract function $getParentComponent : D \rightarrow D_N$ returns the parent component of a model, as displayed on Algorithm 2.

input: $id_{removed} \in ID_{PRDEVS}$
Data: $rem_C \in D, hostC \in D_N$

remC ← `getExistingComponent(`$id_{removed}$`)`
hostC ← `getParentComponent(`$remC$`)`
hostC.D ← hostC.D \ {remC}

Algorithm 2: removeComponent procedure

• addPort : $ID_N \times Name \times P_{type} \times P_{dir} \rightarrow \emptyset$
Adds a port with name $Name$ to a coupled component whose identifier is provided, with the associated *type* and *direction*, as displayed on Algorithm 3.

input : $id \in ID_N$; $p_n \in Name$; $p_t \in P_{type}$; $p_d \in P_{dir}$
Data: $new_{port} \in Port$; $hostC \in D_N$

$new_{port} \leftarrow (p_n, p_t, p_d)$
hostC ← `getExistingComponent(`id`)`
hostC.port ← hostC.port ∪ $\{new_{port}\}$

Algorithm 3: addPort procedure

• removePort : $ID_N \times Name \rightarrow \emptyset$
Removes the port name $Name$ from the component whose ID is provided. The abstract function $getPort : D_N \times Name \rightarrow Port$ gets an existing port from a coupled component, as displayed on Algorithm 4.

input : $id \in ID_N$;$p_n \in Name$
Data: $hostC \in D_N$; $removed_{port} \in Port$

hostC ← `getExistingComponent(`id`)`
$removed_{port}$ ← `getPort(`$hostC$,p_n`)`
hostC.port ← hostC.port \ $\{removed_{port}\}$

Algorithm 4: removePort procedure

• addConnection : $ID_N \times Name \times ID_N \times Name \rightarrow \emptyset$
Adds a connection between two ports. The ports are referred to using the combination of the component identifier and the port name. The $Z_{i,d}$ definition must be respected, i.e. the two components must be in the same coupled, or one of the two component must be a coupled and the other one a component inside the coupled, and one must be an input and the other an output. Moreover, the definition interval *type* must match between the two ports, as displayed on Algorithm 5.

input : $id_1 \in ID_{PRDEVS}$; $p_n1 \in Name$;
$id_2 \in ID_{PRDEVS}$; $p_n2 \in Name$
Data: $hostC1 \in D_N$; $hostC2 \in D_N$;
$connection \in PortConnection$

$hostC1$ ← `getParentComponent(`id_1`)`
$hostC2$ ← `getParentComponent(`id_2`)`
$connection \leftarrow \{(id_1, p_n1), (id_2, p_n2)\}$
if $hostC1.id$=$hostC2.id$ **then**
  $hostC1.IC = hostC1.IC \cup \{connection\}$
else if $hostC1.id$=id_2 **then**
  $hostC1.EOC = hostC1.EOC \cup \{connection\}$
else if $hostC2.id$=id_1 **then**
  $hostC2.EIC = hostC2.EIC \cup \{connection\}$

Algorithm 5: addConnection procedure

• removeConnection : $ID_N \times Name \times ID_N \times Name \rightarrow \emptyset$
Removes a connection between two ports using the same notation as the previous function, as displayed on Algorithm 6.

input : $id_1 \in ID_{PRDEVS}$;$id_2 \in ID_{PRDEVS}$;$p_n1 \in Name$;$p_n2 \in Name$
Data: $hostC1 \in D_N$; $hostC1 \in D_N$;
$connection \in PortConnection$

$hostC1$ ← `getParentComponent(`id_1`)`
$hostC2$ ← `getParentComponent(`id_2`)`
$connection \leftarrow \{(id_1, p_n1), (id_2, p_n2)\}$
if $hostC1.id$=$hostC2.id$ **then**
  $hostC1.IC = hostC1.IC \setminus \{connection\}$
else if $hostC1.id$=id_2 **then**
  $hostC1.EOC = hostC1.EOC \setminus \{connection\}$
else if $hostC2.id$=id_1 **then**
  $hostC2.EIC = hostC2.EIC \setminus \{connection\}$

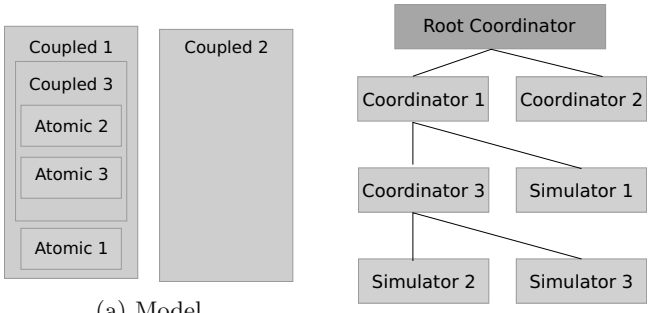
Algorithm 6: removeConnection procedure

AN EXAMPLE OF PRDEVS PDM AND PSM

While our final aim is to implement PRDEVS on reconfigurable hardware, we choose to first develop a software PDM using well-known tools. This is intended as a proof of concept to check that the simulator structure is reliable.

Software PDM Definition

We propose a PDM for software simulation that implements PRDEVS abstract syntax. This implementation is based on Zeigler's abstract simulator described in Zeigler et al. (2000). This simulator deals with a static hierarchy of components and we add the dynamic SC functions to treat SC-messages from λ_{SC}.
The simulator described by Zeigler uses a hierarchical tree of models, which has a coordinator object for each coupled model and a simulator object for each atomic model. To each of the simulator objects, a model object describing the structure of the represented atomic is associated. There is a single root coordinator which lead the hierarchical tree. The root coordinator contains a list of imminent models and their next event time. This list is updated at the beginning of each event step.
Under root coordinator are coordinators which can be the parents of coordinators or simulators which are the leaves of this hierarchical tree. A correspondence from model to simulation can be seen in Figure 2. We name this a one-to-one correspondence process. This is part of the initialization phase.



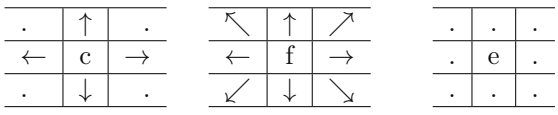
(a) Model (b) Controller

Figure 2: One-to-One Correspondence

As a first approach, a flat representation is chosen for our implementation of PRDEVS. That is to say, the levels of hierarchy are ignored as if all the atomics were directly instantiated in C^{TOP}. A list of available components L is held in the root coordinator. A simulator or a coordinator which correspond to a type can be created using the SC-functions.
There are four types of messages sent between simulators and/or coordinators during simulation: X-message, Y-message, *-message and SC-message which correspond respectively to δ_{ext}, λ, δ_{int} and λ_{SC} of atomic components. The first three messages are defined as Zeigler's: imminent models are chosen based on their minimal next event step and the imminent models triggers *-message. If the conditions to trigger λ on the current state are met, a Y-message is then integrated into the Y-message bag. The bag is sent to the target model and is received as a X-message at the end of each event cycle. The model which received X-message will move to upcoming state and wait for next cycle. SC-message, in some aspect similar to Y-message, will build a SC-message bag and the dynamic SC functions are executed at the end of each event cycle. The SC-message is then treated in zero-time compared to the simulation time.

Use Case Definition

We apply PRDEVS syntax by creating a PSM simulation implementing a game: Within a $size \times size$ grid co-exist three types of players: chicken, fox and egg. Each cell can hold only one player and players move under certain rules, as shown in Figure 3: chickens can move randomly around in four direction while foxes can move randomly around in all eight directions; eggs can not move. There is a rules model recording the position of all players and judging if each move is authorized.
Each round, all players are imminent and move simultaneously. If a chicken reaches another chicken, an egg will be laid randomly around and it stays at the same position. If a fox reaches a chicken, the chicken is eaten and its cell occupied by fox. The game ends when there are no chicken any more or if foxes are blocked by eggs.
Chickens and foxes components communicate using their ports. *valid* is an input port and *askAvailability* is an output port.
$P_{valid} =< valid, in, \{isChicken, isFree, else\} >$
$P_{askAvailability} =< askAvailability, out, (positionX \in size, positionY \in size) >$
Players calculate their destination themselves and verify with the rules component before moving.
The state machine for the Chicken model is as shown in Figure 4. The initial state of a chicken is $S1$. When the chicken component is imminent, it receives a *-message to execute the internal transition and randomly defines the desired destination. It moves to state $S2$. Then an output is sent to verify the availability of this position. It moves to state $S3$ and wait for an input. The Y-message arrives to the rules model which responds ac-



(a) Chicken (b) Fox (c) Egg

Figure 3: Moving Rules for Players

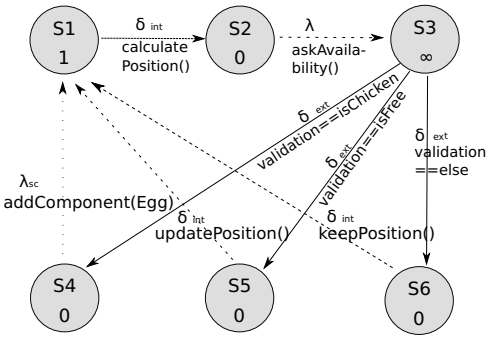


Figure 4: Internal View of the State Machine of Chicken

cording to the availability. Depending on this response, the chicken model moves to state $S4$, $S5$ or $S6$ and then the SC-function or the internal transition is called.
When the SC-function addComponent is triggered, the simulator stores the call. After simulation cycle is over, the root coordinator copies the Egg library object into the C^{Top} component. After what, the simulation cycle resumes.
The game starts with an initial numbers of players, an example is presented on figure 5a. After the simulation runs, we found the result as shown in figure 5b.

c	c	.	.
.	.	f	.
.	.	.	.
.	e	.	.

(a) Game Starts

.	.	.	.
f	e	.	.
.	.	.	.
.	e	.	.

(b) Game Ends

Figure 5: Example of game turns

Hardware PDM Overview

A hardware PDM is being formalized, relying on several reconfigurable areas and two buses: one control bus and one data bus. There is a correspondence table for each area and its address. Each reconfigurable area will have a type and certain limits since their hardware resources can be different.
Reconfigurable areas can store a component model. An API (Application Programming Interface) for the SC-functions can be implemented at software level in specific area containing a processor. The API is able to match the configuration of the FPGA, such as which type of reconfigurable areas are suitable for which type of component model.

CONCLUSION AND FUTURE WORK

In this article, we presented a system engineering approach using formal modeling which aims at supporting dynamic architectures. This approach leads to a separation between the model and the final application platform. We presented a PIM model which can adapt its structure dynamically, both component- and connection-related. With the formal SC-functions presented in this article, a PIM level of PRDEVS is defined. A possible PDM and PSM is introduced with application on software by a game. With this practical example of PRDEVS syntax, the feasibility of the dynamical formal model is verified. However in practice, the SC-functions are only applied at the end of each event cycle. One future work will be to integrate the SC-functions during the simulation without pausing the other components, in order to allow extending beyond simulation purposes and take advantage of PR.
Finally, the main objective will be the FPGA-based implementation of a use case defined at PRDEVS PIM level. This will require real-time handling, and partial reconfiguration scheduling to allow reconfigurations to be completed before using the component, where the structure change was carried on in zero time in simulation.

References

Barros F.J., 1997. *Modeling Formalisms for Dynamic Structure Systems. ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 7, no. 4, 501–515. ISSN 1049-3301. doi:10.1145/268403.268423.

Barros F.J., 1998. *Abstract simulators for the DSDE formalism.* In *1998 Winter Simulation Conference. Proceedings (Cat. No.98CH36274)*. IEEE, vol. 1, 407–412. doi:10.1109/WSC.1998.745015.

Feist T., 2012. *Vivado design suite. White Paper*, 5.

Koch D., 2012. *Partial Reconfiguration on FPGAs: Architectures, Tools and Applications*, vol. 153. Springer Science & Business Media.

Madlener F., 2013. *A Model of Computation for Reconfigurable Systems.* Ph.D. thesis, Technische Universität, Darmstadt.

Object Management Group, 2016. *MDA - The Architecture of Choice for a Changing World.* URL `http://www.omg.org/mda/`. [Online; accessed 19-January-2017].

Uhrmacher A.M., 2001. *Dynamic Structures in Modeling and Simulation: A Reflective Approach. ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 11, no. 2, 206–232. ISSN 1049-3301. doi:10.1145/384169.384173.

Zeigler B.P.; Praehofer H.; and Kim T.G., 2000. *Theory of modeling and simulation: integrating discrete event and continuous complex dynamic systems.* Academic press, Orlando, FL, USA, 2nd ed. ISBN 0127784551.

MODELLING AND OPTIMAL CONTROL WITH ENERGY REGENERATION OF A 6DOF MOTION PLATFORM WITH PERMANENT MAGNET LINEAR ACTUATORS

E. Thöndel
Department of Electric Drives and Traction
Czech Technical University
166 27, Prague,
Czech Republic
E-mail: thondee@fel.cvut.cz

KEYWORDS
Motion Platform, Linear Actuator, Permanent Magnet Linear Synchronous Motor, Optimal Control, Energy Regeneration.

ABSTRACT

The paper deals with the modelling, simulation and optimal control of parallel robotic structures (in particular those with six degrees of freedom) driven by permanent magnet linear electric motors using energy regeneration and during power outages. By virtue of their kinematic and dynamical properties, these robotic structures can be used for a variety of purposes, for instance in industrial automation or simulation technology, where motion platforms (especially hexapods) are used for motion cueing. Electric drive parallel mechanisms provide several significant benefits, in particular with regard to optimum energy distribution and regeneration options. The main research motivation is the prevailing need to develop a solution ensuring that the parallel mechanism can safely return after a power cut to the default or 'park' position in a fully controlled way (simulators, for instance, have to be capable of sliding back to the boarding position) using solely the energy accumulated in the system. This research goal requires the development of exact and fast simulation models providing insight into transient system behaviour and, subsequently, determining the optimum control strategy. Therefore, the paper will provide a detailed simulation model which can also be used in optimizing the structures of the mechanism and determining the optimum servo drive design.

INTRODUCTION

Parallel robotic structures play a crucial role in many applications. The author's technical interest is focused in particular on the area of simulation technology, where motion platforms with six degrees of freedom are used for the purposes of motion cueing. Nevertheless, the results of the research can be applied universally and employed in many other areas, such as industrial automation.
Motion platforms with six degrees of freedom were originally designed for use in aircraft simulators (Stewart 1965). However, the system soon found many uses in other industries, such as automated production or testing.
The platform is comprised of six linear actuators arranged in a parallel kinematic structure. Different platform types employ different linear actuators, with the oldest ones featuring hydraulic cylinders. Gradually, hydraulic solutions started to be replaced by electromechanical actuators (Thöndel 2011). However, another switchover is imminent, following the recent breakthroughs in the area of permanent magnet electric drives, as purely electric linear actuators now start to offer a fully-fledged alternative to hydraulic or electromechanical solutions. The main benefit of electric actuators is a much higher dynamic range, allowing systems based on this method to be used in a wider variety of applications.
Wind tunnel analyses and measurements are performed by all airborne research organisations. A new motion platform, intended for aerodynamic process measurements and determination of dynamic parameters, is currently being developed by the CIDAM (Centre for Intelligent Drives and Advanced Machine Control) competence centre with the support of the Technology Agency of the Czech Republic (TACR). With the new platform researchers will be able to execute certain limited manoeuvres and test the aircraft's control algorithms directly in a wind tunnel.
A schematic model of the mechanism described in this paper is provided in Figure 1.

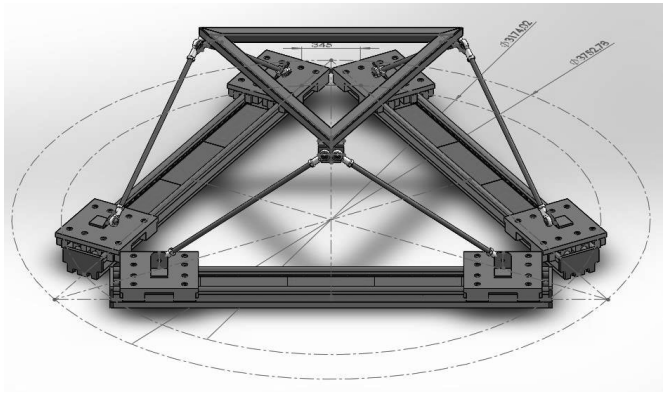


Fig. 1: Motion Platform with Six Degrees of Freedom Driven by Linear Actuators

Given the properties and requirements of permanent magnet linear motors a different linear actuator design was used in this project. Unlike hydraulic or electromechanical systems with flexible actuator arm lengths, the permanent magnet electric actuators used in this motion platform feature fixed-length arms. To reach the required position the arm's bottom joint slides along a rail fitted with permanent magnets.
A detailed analysis of the kinematic structure and the principle of linear motors are provided in (Thöndel 2014).

PERMANENT MAGNET SYNCHRONOUS LINEAR MOTOR

Permanent magnet synchronous motors (PMSMs) belong to the most recent motor generations, finding use especially in applications requiring accurate position and speed servo-control (such as industrial robots). Unlike other electric motors, PMSMs can be easily arranged in a linear shape. This feature will be leveraged in the development of the mathematical model.

The mathematical model of a linear motor can be derived from the properties of typical rotational motors – all one has to do is to "cut and unroll" the engine in one's mind (Miller et al. 2002). All properties of the mathematical model of a rotational motor can be recalculated to a linear form if the radius of the "unrolled" rotational motor is known. The radius can be written as:

$$2\pi r = 2p\tau \rightarrow r = \frac{\tau p}{\pi}, \tag{1}$$

where r is the desired radius, p the number of pole pairs and τ the pole pitch.

Rotational electric machines are typically analysed in a suitable coordinate system rotating synchronously with the selected quantity. In this way, the examined AC quantities can be transformed to the corresponding DC quantities. A system rotating with a synchronous speed seems to be best suited for synchronous machines. In technical literature, this transformation is known as the d-q transform, or Park's transformation.

The currents can be transformed to the d-q system by means of a gradual α-β transformation (Clarke transformation) to the stationary orthogonal system:

$$\begin{bmatrix} i_\alpha \\ i_\beta \end{bmatrix} = \frac{2}{3}\begin{bmatrix} 1 & \frac{-1}{2} & \frac{-1}{2} \\ 0 & \frac{\sqrt{3}}{2} & \frac{-\sqrt{3}}{2} \end{bmatrix}\begin{bmatrix} i_a \\ i_b \\ i_c \end{bmatrix} \tag{2}$$

and subsequently to the rotating d-q system:

$$\begin{bmatrix} i_d \\ i_q \end{bmatrix} = \begin{bmatrix} cos\vartheta_e & sin\vartheta_e \\ -sin\vartheta_e & cos\vartheta_e \end{bmatrix}\begin{bmatrix} i_\alpha \\ i_\beta \end{bmatrix} \tag{3}$$

The mathematical model has been determined under the following assumptions:

- The system is powered by a three-phase symmetrical power source with harmonic voltages.
- All phases have the same resistances and inductances.
- The magnetization characteristic is linear.
- Losses in iron are not considered.

The following expressions apply to the different stator winding phases:

$$\begin{aligned} u_a &= R_s i_a + \frac{d}{dt}\psi_a \\ u_b &= R_s i_b + \frac{d}{dt}\psi_b \end{aligned} \tag{4}$$

$$u_c = R_s i_c + \frac{d}{dt}\psi_c$$

where R_s is the stator winding resistance and ψ_a, ψ_b and ψ_c are the magnetic fluxes in the form:

$$\begin{aligned} \psi_a &= L_s i_a + \psi_M cos(\omega_e t) \\ \psi_b &= L_s i_b + \psi_M cos\left(\omega_e t + \frac{2}{3}\pi\right) \\ \psi_c &= L_s i_c + \psi_M cos\left(\omega_e t - \frac{2}{3}\pi\right) \end{aligned} \tag{5}$$

In the above equation, L_s is the stator winding inductance and ψ_M the rotor magnetic flux, rotating with respect to the stator with the speed ω_e.

The final equations of the electrical part of the motor can be obtained by means of a d-q transform of the above formulas:

$$\begin{aligned} u_d &= R_s i_d + L_s \frac{d}{dt} i_d - L_s \omega_e i_q \\ u_q &= R_s i_q + L_s \frac{d}{dt} i_q + L_s \omega_e i_d + \psi_M \omega_e \end{aligned} \tag{6}$$

The foregoing set of equations can be extended with the equation of the electromechanical moment, which can be derived using the law of conservation of energy, assuming mechanical losses and losses in iron are not considered.

$$\frac{3}{2}Re\{U_s I_s\} = M\omega_m + 3R_s I_s^2, \qquad \omega_m = \frac{\omega_e}{p} \tag{7}$$

$$M = \frac{3}{2}\psi_M p i_q \tag{8}$$

Now, the radius expression determined earlier can be used and the above model "unrolled" into linear form.

$$\begin{aligned} \omega_e &= p\omega_m = p\frac{v}{r} = \frac{\pi}{\tau}v \\ M &= Fr = \frac{\tau p}{\pi}F \end{aligned} \tag{9}$$

The linear motor equations can be obtained by substituting into the rotation motor equations (6) and (8):

$$\begin{aligned} u_d &= R_s i_d + L_s \frac{d}{dt} i_d - L_s \frac{\pi}{\tau} v i_q \\ u_q &= R_s i_q + L_s \frac{d}{dt} i_q + L_s \frac{\pi}{\tau} v i_d + \psi_M \frac{\pi}{\tau} v \\ F &= \frac{3}{2}\frac{\pi}{\tau}\psi_M i_q = K_F i_q \end{aligned} \tag{10}$$

where K_F is the force constant, v the motor's mechanical speed ($\dot{y}_j$) and F the acting force.

In the derived equations, the current component i_d generates a magnetic flux inverse to the magnetic flux generated by the permanent magnets. In practice, the value of this current component is maintained at zero by means of independent vector control. Assuming that $i_d = 0$, the foregoing equations can be further simplified without any impact on the model's suitability for the design and testing of the motion platform. Following this step, the resulting synchronous linear motor differential equation can be written as:

$$\frac{d}{dt}i_q = -\frac{R_s}{L_s}i_q + \frac{1}{L_s}u_q - \frac{2}{3}\frac{K_F}{L_s}v \qquad (11)$$
$$F = K_F i_q$$

MATHEMATICAL MODEL OF THE PARALLEL MECHANISM

The analytical expression of the dynamic behaviour of the parallel mechanism is highly complex, as the equation of motion cannot be obtained without forward kinematic transformation, an exceedingly complicated operation for this mechanism type.
The equation of motion has the following general form (Thöndel 2011):

$$m_{red}\ddot{y} + \frac{1}{2}\frac{\partial m_{red}}{\partial y}\dot{y}^2 = F - m_{red}g - B\dot{y} \qquad (12)$$

In this expression, F is the acting force of the linear motor according to expression (11), y the displacement of the linear motor, B the viscous friction coefficient, g the gravitational acceleration and m_{red} the reduced mass of the given actuator's load. Generally, this mass depends on the position of the platform and therefore cannot be determined without forward kinematic transformation.

$$m_{red} = f(y_1, \quad \dots, \quad y_6) \qquad (13)$$

A mathematical model was created in the MATLAB/ Simulink/SimMechanics environment for simulation purposes to provide numerical solutions of the forward kinematic transformation (Figure 2).

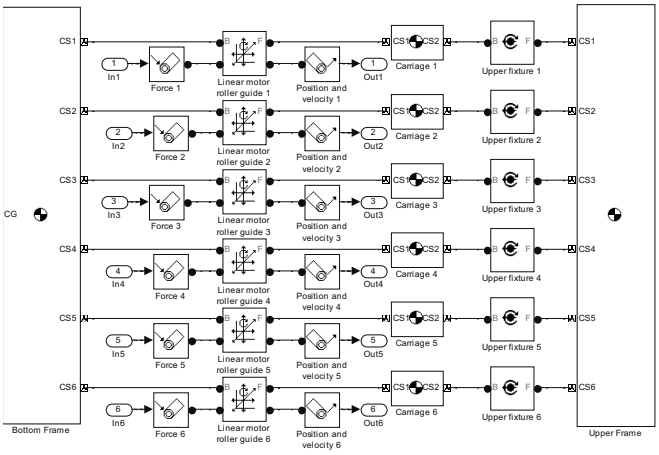


Fig. 2: Mathematical Model of the Mechanism in the MATLAB/Simulink/SimMechanics Environment.

In addition, the simulation model is planned to be used also for control algorithm design and testing. For this purpose, the model can be linearised, assuming that the total load mass m_{load} is divided approximately evenly between individual actuators. In such a scenario, the following expression holds true:

$$m_{red} = const = \frac{m_{load}}{6} \qquad (14)$$

Under these simplified conditions, the equation of motion for an actuator can be expressed in the following linear differential matrix form, which is suitable for control algorithm designs and analyses. The control algorithm must be robust enough with respect to the m_{red} variation expressed by the formula (13).

$$\frac{d}{dt}\underbrace{\begin{bmatrix} i_q \\ v \\ y \end{bmatrix}}_{x_j} = \underbrace{\begin{bmatrix} -\frac{R_s}{L_s} & -\frac{2}{3}\frac{K_F}{L_s} & 0 \\ \frac{K_F}{m_{red}} & -\frac{B}{m_{red}} & 0 \\ 0 & 1 & 0 \end{bmatrix}}_{A_j}\begin{bmatrix} i_q \\ v \\ y \end{bmatrix} + \underbrace{\begin{bmatrix} \frac{1}{L_s} & 0 \\ 0 & -g \\ 0 & 0 \end{bmatrix}}_{B_j}\underbrace{\begin{bmatrix} u_q \\ 1 \end{bmatrix}}_{u_j}$$

$$y_j = \underbrace{[0 \quad 0 \quad 1]}_{C_j^y}\begin{bmatrix} i_q \\ v \\ y \end{bmatrix} \qquad i_j = \underbrace{[1 \quad 0 \quad 0]}_{C_j^i}\begin{bmatrix} i_q \\ v \\ y \end{bmatrix}$$
$$j = \langle 1 \quad \dots \quad 6 \rangle \qquad (15)$$

CONTROL SYSTEM

In each state, the system contains a certain amount of energy. This can be divided into three different types:

- kinetic energy resulting from the movement of individual actuators,
- potential energy given by the positions of individual actuators, and thus the whole device, and
- electrical energy accumulated in the reactance of the servo-loops of the drives

These energy components can be mutually converted (not without losses, naturally) or the energy can be dissipated via frequency converter brake resistances or conducted back to the electrical grid. In the case of a power failure, the system cannot use additional energy. Therefore, the envisaged algorithm has to provide a control solution ensuring that the energy accumulated in the system is gradually dissipated via the system's frequency converter brake resistances so as to allow the platform to descend safely to the default park position via an emergency trajectory dynamically calculated and continuously updated during the 'landing' procedure based on current state of all actuators.
By joining the mathematical description of the individual actuators (15) into one model, we obtain one linear dynamic system with six inputs and six outputs:

$$\begin{aligned} A_s &= diag(A_1 \quad \dots \quad A_6) \\ B_s &= diag(B_{11} \quad \dots \quad B_{61}) \\ C_s^y &= diag(C_1^y \quad \dots \quad C_6^y) \\ C_s^i &= diag(C_1^i \quad \dots \quad C_6^i) \\ x &= [x_1^T \quad \dots \quad x_6^T]^T \\ u &= [u_{q1} \quad \dots \quad u_{q6}]^T \\ y &= [y_1 \quad \dots \quad y_6]^T \\ i &= [i_1 \quad \dots \quad i_6]^T \end{aligned} \qquad (16)$$

where B_{11} to B_{61} is always the first column of the respective matrix B_1 to B_6.
In the following design the implementation of the control on a digital computer (PLC) is intended, that is, a discrete version of the controlled system is required. By simple discretization

by the Euler method we obtain the corresponding matrix of the discrete system:

$$\begin{aligned} A_d &= I + A_s T_s \\ B_d &= B_s T_s \\ C_d^y &= C_s^y \\ C_d^i &= C_s^i \end{aligned} \tag{17}$$

where I is the unit matrix, and T_s is the sampling period.
In order to eliminate the steady-state error (caused by static gravitational force), it is advisable to add an integrator to the system (Wang 2009). Taking a differential operation on both sides of the system state equation and add additional state variables, we obtain an augmented system:

$$\begin{bmatrix} \Delta x(k+1) \\ y(k+1) \\ i(k+1) \\ u(k+1) \end{bmatrix} = \begin{bmatrix} A_d & 0 & 0 & 0 \\ C_d^y A_d & I & 0 & 0 \\ C_d^i A_d & 0 & I & 0 \\ 0 & 0 & 0 & I \end{bmatrix} \begin{bmatrix} \Delta x(k) \\ y(k) \\ i(k) \\ u(k) \end{bmatrix} + \begin{bmatrix} B_d \\ C_d^y B_d \\ C_d^i B_d \\ 1 \end{bmatrix} \Delta u(k) \tag{18}$$

By substituting (17) to (18) we get the final system, for which we design an optimal control.

$$\begin{aligned} \underbrace{\begin{bmatrix} \Delta x(k+1) \\ y(k+1) \\ i(k+1) \\ u(k+1) \end{bmatrix}}_{x_e(k+1)} &= \underbrace{\begin{bmatrix} I + A_s T_s & 0 & 0 & 0 \\ C_s^y + C_s^y A_s T_s & I & 0 & 0 \\ C_s^i + C_s^i A_s T_s & 0 & I & 0 \\ 0 & 0 & 0 & I \end{bmatrix}}_{A_e} \underbrace{\begin{bmatrix} \Delta x(k) \\ y(k) \\ i(k) \\ u(k) \end{bmatrix}}_{x_e(k)} \\ &+ \underbrace{\begin{bmatrix} B_s T_s \\ C_s^y B_s T_s \\ C_s^i B_s T_s \\ 1 \end{bmatrix}}_{B_e} \Delta u(k) \\ y(k) &= \underbrace{[0 \quad I \quad 0 \quad 0]}_{C_e^y} \begin{bmatrix} \Delta x(k) \\ y(k) \\ i(k) \\ u(k) \end{bmatrix} \\ i(k) &= \underbrace{[0 \quad 0 \quad I \quad 0]}_{C_e^i} \begin{bmatrix} \Delta x(k) \\ y(k) \\ i(k) \\ u(k) \end{bmatrix} \\ u(k) &= \underbrace{[0 \quad 0 \quad 0 \quad I]}_{C_e^u} \begin{bmatrix} \Delta x(k) \\ y(k) \\ i(k) \\ u(k) \end{bmatrix} \end{aligned} \tag{19}$$

The aim of the control is to minimize the control error and at the same time to optimize the energy flow through the system so that the system can be brought into the parking position even when the power supply is disconnected. Let us define a cost function that converts this task into an optimization problem.

$$J = (ref - Y)^T Q (ref - Y) + \Delta U^T R \Delta U + U^T S I_i \tag{20}$$

The first term minimizes the error between the predicted value Y and the set-point signal *ref*; the second term respects the change of the control voltage ΔU and the last term represents the product of the control voltage U and the current component I_i generating the applied force. The value of the product corresponds to the actual power, and because of the optimization, the positive value and the negative value is considered to be consumed and regenerated energy respectively. Matrices Q, R, and S are diagonal weight matrices that can be used to tune the performance of the control.
Based on the derived augmented state-space model, it is possible to predict the behavior of the output variable over a defined time horizon T_p.

$$Y = F_y x_e(k) + \Phi_y \Delta U \tag{21}$$

where

$$\begin{aligned} Y &= \begin{bmatrix} y(k+1)^T & \dots & y(k+1+T_p)^T \end{bmatrix}^T \\ \Delta U &= \begin{bmatrix} \Delta u(k) & \dots & \Delta u(k+T_p) \end{bmatrix}^T \\ F_y &= \begin{bmatrix} C_e^y A_e^0 \\ \vdots \\ C_e^y A_e^{T_p} \end{bmatrix}, \Phi_y = \begin{bmatrix} C_e^y A_e^0 B_e & 0 & 0 \\ \vdots & \ddots & 0 \\ C_e^y A_e^{T_p - 1} B_e & \dots & C_e^y A_e^0 B_e \end{bmatrix} \end{aligned}$$

Similarly, it is possible to define the predicted behavior of other variables in the cost function.

$$I_i = F_i x_e(k) + \Phi_i \Delta U \tag{22}$$

$$U = F_u x_e(k) + \Phi_u \Delta U \tag{23}$$

By substituting (21), (22) and (23) into (20) and subsequent modification, we get the final formula for the cost function.

$$\begin{aligned} J &= \Delta U^T \underbrace{\left(\Phi_y^T Q \Phi_y + R + \Phi_u^T S \Phi_i\right)}_{H} U^T \\ &+ \Delta U^T \underbrace{\left[\underbrace{\left(2\Phi_y^T Q F_y + \Phi_u^T S F_i + \Phi_i^T S F_u\right)}_{f_1} x_e(k) - \underbrace{2\Phi_y^T Q}_{f_2} ref \right]}_{f} \end{aligned} \tag{24}$$

By minimizing the cost function, we obtain the optimal control that takes into account not only the error between the output and the set-point but also the energy balance of the control with respect to the energy regeneration.
If we do not take into account the constraints of the control variable (voltage), it is possible to find the solution analytically by means of a derivative. In this case, the control law leads to a state feedback.

$$\begin{aligned} \Delta U &= H^{-1}\left(f_2 ref - f_1 x_e(k)\right) \\ u(k) &= u(k-1) + \Delta U(1 \dots 6) \end{aligned} \tag{25}$$

The control variable is limited by the value of the voltage u_{dc} on the DC link of the frequency inverter. The voltage can be directly measured and predicted in the case of emergency descent by solving the following differential equation:

$$u_{dc} \frac{du_{dc}}{dt} = -\frac{1}{C_{dc}} u^T i \tag{26}$$

where C_{dc} is the capacity of the DC link.

The minimum of cost function must be searched by numerical methods while respecting the control voltage constraints, which significantly increases computational demands.

SIMULATION RESULTS

The derived simulation model was implemented in the MATLAB/Simulink environment. A general overview of the parameters used in the simulation is provided in Table 1 below.

Table 1: Model Parameters Used During Simulation

Parameter	Value	Unit
R_s	5.7	Ω
L_s	40	mH
K_F	145.5	N/A
m_{load}	500	kg
B^*	0	kg.s^{-1}
C_{dc}	2.32	mF
T_p	50	steps
T_s	0.01	s

* Mechanical losses were not taken into consideration in the initial simulation.

A summary overview of the simulation results is provided in Figure 3 below, where a simulation result is shown comparing the response of the system on the emergency descent with $(S > 0)$ and without $(S = 0)$ consideration of energy regeneration in the control law.
It can be seen that in the case of disregarding the regenerated energy, voltage drops rapidly on the DC link, which quickly leads to the shutdown of the control system and the free fall of the mechanism. The optimal control with the use of energy regeneration finds a solution where the DC link voltage drop is minimal.
So far, it has not been possible to verify the simulation results with actual measurements, as the platform does not exist yet. The purpose of the mathematical and simulation model is to lay the foundations for the subsequent design and sizing of the mechanism and its control system.

CONCLUSION

The paper provides a description of the modelling and the optimal control of a 6DOF motion platform actuated by permanent magnet linear motors. It introduces a detailed mathematical model and presents the design of the optimal control with respect to the energy regeneration. The model and the control described here will be employed to design an actual mechanism to be later used for testing and in advanced vehicle and flight simulators.

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 simulation engineer and developer at Pragolet since 2009.

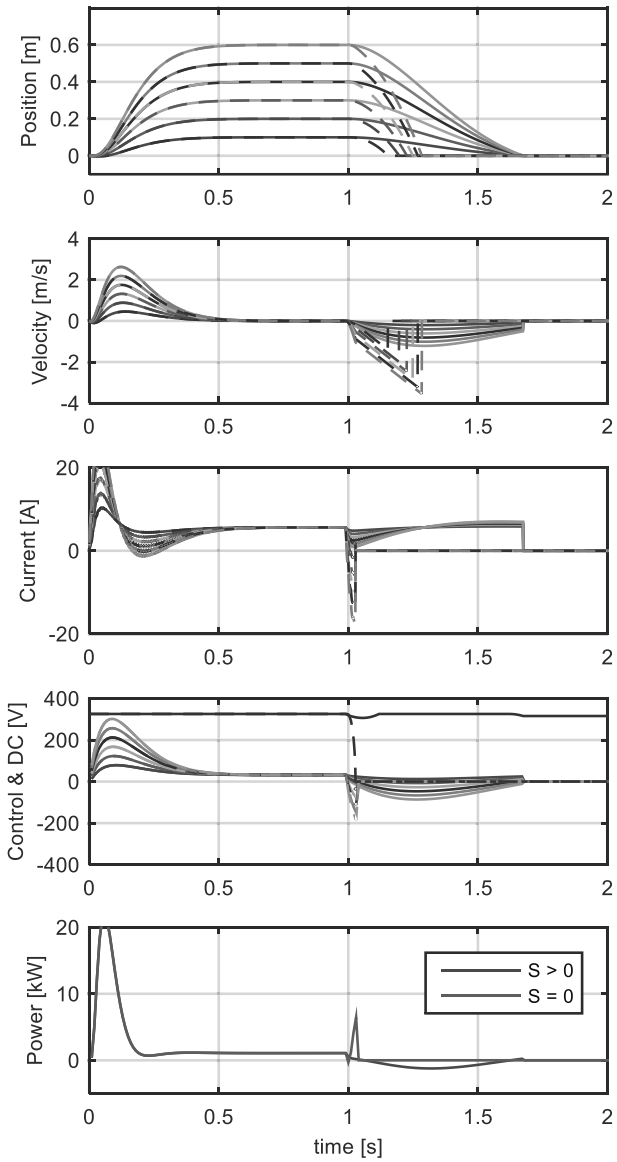


Fig. 3: Simulation Results – a power-cut is detected at the time 1 second.

REFERENCES

Stewart, D. 1965. "A Platform with Six Degrees of Freedom". UK Institution of Mechanical Engineers Proceedings, Vol 180.

Thöndel, E. 2011. "Design and Optimal Control of a Linear Electromechanical Actuator for Motion Platforms with Six Degrees of Freedom". Intelligent Automation and Systems Engineering, Springer, 65-77.

Thöndel, E. 2014. "Modelling and Simulation of a 6DOF Motion Platform with Permanent Magnet Linear Actuators for Testing in Wind Tunnel". In 28th European Simulation and Modelling Conference, Porto, Portugal.

Wang, L. 2009. "Model Predictive Control System Design and Implementation Using Matlab®". Springer.

Miller, CE; Zyl, AW and Landy CF. 2002. "Modelling a Permanent Magnet Linear Synchronous Motor for Control Purposes". IEEE Africon 2002.

HUMAN COMPUTER INTERFACES

HUMAN-COMPUTER INTERFACE FOR COMMUNICATION AND AUTOMATED ESTIMATION OF BASIC EMOTIONAL STATES

Svetla Radeva, Strahil Sokolov and Dimitar Radev
Department of Information Technologies
University of Telecommunications and Post
1st Acad. Stefan Mladenov Str., 1700 Sofia,
Bulgaria
E-mail: svetla_ktp@abv.bg

KEYWORDS
Man-machine interfaces, Neural networks, Image processing, Sampling, Estimation.

ABSTRACT

The considered Human-Computer Interface (HCI) use Electroencephalography (EEG) activity or other electrophysiological measures of brain functions as new non-muscular channels for control and communication with smart devices and smart mobile applications for disabled persons. The research aims developing of technology for communication with smart mobile applications, based on processing of recorded electrophysiological signals at execution of different mental tasks. The interface use depends on the interaction of two adaptive controllers: the user, who generates brain signals that encode intent and the computer system that translate these signals into commands that accomplish the user's intent for connection with smart mobile applications. Estimation of the basic emotional states is a step of building up such interface. The recorded brain signals with experimental setup for two basic emotional states after noise filtering are estimated on the base of clustering and classification using Convolutional Neural Network (CNN) and statistical features. Classification is performed with support vector machines (SVM). Since the human emotions are modelled as combinations from physiological elements such as arousal, valence, dominance, liking, etc., these quantities are the classifier's outputs. Features was selected on the base of principal component analysis for face emotion estimation. The best achieved correct classification performance for EEG is about 68.2%. Classifier combination is used to return the final score for the particular subject.

INTRODUCTION

Human-Computer Interface depends on the interaction of two adaptive controllers: the user, who generate brain signals that encode intent and the interface system, that translate these signals into commands that accomplish the user's intent (Radeva and Radev, 2015a). In this sense for use such interface both user and system must acquire and maintain to each other. The user encodes intent in signal features that the interface can measure. The HCI measures these features and translates them into device commands. This dependence, both initially and continually, on the adaptation of user to system and system to user is the fundamental principle of HCI, based on measurements of EEG. The human – computer interface (HCI) is based on the fundamentals of human-computer interaction which consists of the following:

- Human brain decides the instruction for delivering to thinking activity;
- This decision, from human-brain, is transfer to human peripheral(s) by nervous system;
- From human peripheral(s), this decision is transferred to computer peripheral;
- From computer peripheral the decision, which is now computer command is transferred to CPU (computer brain);
- Principal component analysis for estimation of basic emotional states;
- CPU executes the task.

The time taken by human brain to decide on the first step and CPU to execute the instruction on the last step is almost negligible, because only human brain and computer brain are the active part of this interaction. The rest steps are a medium which-just bridging a gap between human thinking process and CPU understanding process. If we can somehow bridge this gap via some automatic means, then a brain-computer interface will convert human brain thoughts and estimates basic emotional states directly into computer brain instructions or executing programs.
In recent years, multimodal approaches for human emotion estimation have emerged. Researches have tried to incorporate the EEG signals for emotional state analysis and this represents a challenging area of modern research. Recent advances in the area of EEG analysis are said to deliver promising emotion recognition results already. There still seem to be gaps in terms of stability and accuracy in those algorithms. This is what motivated us to research in the area to provide a framework for reliable EEG emotional state estimation with main application of human-computer interface (Radeva and Radev, 2015b).

TWO-DIMENSIONAL HUMAN EMOTION MODEL

The human emotion is a highly subjective phenomenon: it has been accepted by phsycologists that multiple dimensions or scales can be used to categorize emotions. The two–dimensional model of emotion, shown in Figure 1 is introduced in (Wang et al. 2011). The valence axis represents the quality of an emotion ranging from unpleasant to pleasant. The arousal axis refers to the quantitative activation level ranging from calm to excited state.

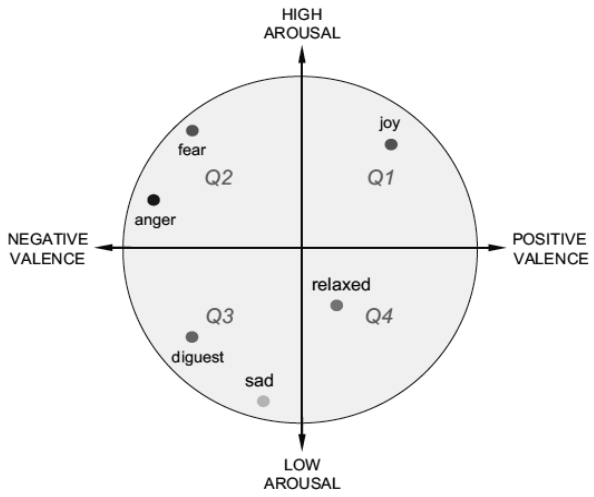


Figure 1: Two–dimensional human emotion model

This approach for recognition of EEG-based emotions uses time and frequency features. The time features are in fact some statistical quantities such as means and standard deviations of the raw signals and its first and second derivatives as well. The classified emotions are: joy, relax, sad and fear. The best accuracy is about 66% among three types of classifiers.

In (Li and Lu 2009) the authors use the EEG signal to classify two basic emotions: happiness and sadness. These emotions are evoked by showing subjects pictures that contain facial expressions of smile and cry. The authors propose a frequency band searching method to choose an optimal band into which the recorded EEG signal is filtered. They use Common Spatial Patterns (CSP) and linear Support Vector Machine (SVM) to classify these two emotions. To investigate the time resolution of classification, they explore two kinds of trials with lengths of 3*s* and 1*s*.

Classification accuracies of 93.5% and 93% are achieved on 10 subjects for 3 *s* and 1 *s* trials, respectively. Their experimental results indicate that the gamma band (roughly 10 Hz to 30 Hz) is suitable for EEG-based emotion classification.

In (Kothe et al. 2013) the achieved accuracy of the emotional valence is about 71%. The technique relies on changes in the power spectrum of short-time stationary oscillatory EEG processes within the standard EEG frequency bands. The classification stage is logistic regression with elastic-net regularization. The features are extracted from very limited set of electrodes and the dimensionality is further reduced with Principal Component Analysis (PCA). The performance is evaluated for arousal, valence and modality separately. As can be expected the arousal is with highest classification accuracy (over 90%).

In (Horlings et al. 2008) is given an approach for emotions recognition using brain activity. The following types of features are used: EEG frequency band power, cross-correlation between EEG band powers, peak frequency in alpha band and Hjorth parameters. The performances for classification in five classes is above 30%. These results are indicative for a significant variability of EEG-based features for emotion recognition among different subjects.

EEG PROCESSING

The experimental system includes 3D camera Panasonic HDC-Z0000, sender Spectrum DX9 DSMX, Sony GoPRO – GoPro HERO3, Nikon D902D smart TV Samsung UE-65HU8500 + LG60LA620S, ACER K11 Led projector, Linksys EA6900 AC1900 smart router, Pololu Zumo Shield, 8 core/32GB RAM/4TB HDD/3GB VGA computer for video processing that translate EEG signals into computer commands. This implemented system is 10-20 System - an international standard for EEG electrode placement locations on the human scalp. The standard defines a grid relative to physical landmarks on the head, such as the indentation between the nose and forehead (nasion), and the bump on the back of the head (inion) at the occipital protuberance.

For EEG processing was used the spherical spline method, where the human head is modelled as a sphere. The parameter for spline flexibility is set to its default value of 4. The process of analysis of recorded EEG data is connected with the spectral power of the signal in a set of following standard frequency bands: θ (theta - frequency range from 4 *Hz* to 7*Hz*), α (alpha - 8 *Hz* to 13 *Hz*), β (beta-low - 14 *Hz* to 29 *Hz*) and γ (gamma - 30 *Hz* to 45*Hz*). Among many methods for features selection, was chosen the Minimum Redundancy and Maximum Relevance (mRMR) criterion, presented in (Peng et al., 2005). The relevance RL of the set of selected features $F=\{f_1,f_2,..\}$ and target classes C was defined as:

$$RL=\frac{1}{|F|}\sum_{f_i\in F} I(f_i,C) \qquad (1)$$

where I denotes the mutual information. The redundancy RD of the features was defined as follows:

$$RD=\frac{1}{|F|^2}\sum_{f_i f_j\in F} I(f_i,f_j) \qquad (2)$$

For incremental search $\max[I(F,C)]$ is equivalent to $\max[RL(F,C)-RD(F)]$. Our second suggestion for features is mRMR selection from all possible sets of ratios $\frac{Act_{c,b}}{Act_{k,b}}$, $c\neq k$ and $\frac{Mob_{c,b}}{Mob_{k,b}}$, $c\neq k$, where *c* and *k* denote the EEG channel and b is the activity (θ, α or β).

FACIAL FEATURES EXTRACTION

In this investigation we join the rapid cascaded classifier with the accurate monolithic one within the two-level combined cascade of classifiers instead of using them independently. This is realized in order to achieve higher detection and lower false alarm rates.The proposed approach for face detection and validation is based on our previous research (Velchev et al. 2016). It utilizes the OpenCV face detection algorithm (Viola and Jones 2004) and a convolutional neural network. The two-level cascade of classifiers is called "combined" since it combines different types of classifiers, which have been proved in the course of time: the first level is represented by the Haar-like features' cascade of weak classifiers, which is responsible for the face-like objects

detection, and the second level is a CNN for the objects' verification, shown on Figure 2.

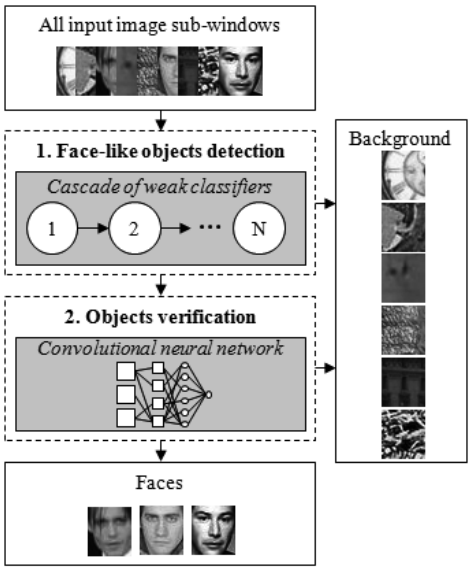


Figure 2: Combined cascade of neural network classifiers

In this phase the fairly fast face detector is also able to deliver faces in frontal pose. This depends on the training set of images for the CNN. In our approach this has proven to be useful since we are using short-length videos of the subjects' faces may have slight fluctuations off the frontal pose. For our experiments we have used not only our measurements, but as well the DEAP dataset from Database for Emotion Analysis using Physiological Signals (Koelstra et al. 2012). It consists of multimodal data physiological including EEG signals taken from 32 leads. Each subject participates with 40 trials for 60 s. The EEG signals were downsampled to 128 *Hz*, bandpass filtered (4 *Hz* to 45 *Hz*). The all data was averaged to the common reference. The performance is validated and evaluated using the k-fold technique, where testing part is extracted from the whole dataset. The rest of the dataset is used to train the classifier. This procedure repeats (10 times in our case) and the accuracy is calculated as an average of the accuracies in the iterations. The arousal and valence scores in dataset are given as fractional numbers ranging from 1 to 9. We have quantized these scores to 3, 5 and 7 levels and the testing was performed for each case. The calculated classification accuracies versus dimensionality of the feature vectors is seen on Figure 3.

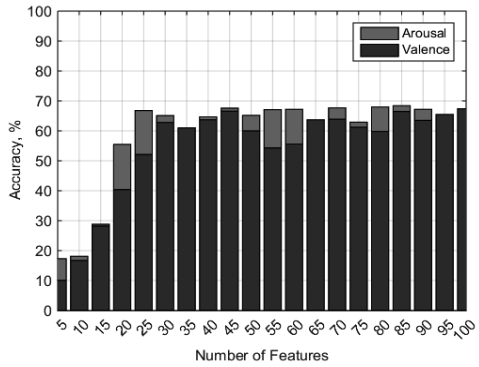


Figure 3: Classification accuracies for arousal and valence versus dimensionality of the feature vectors

Face emotion estimation runs in parallel and contributes to the improvement of the scores generated by the EEG analysis module. The proposed PCA-based face emotion estimation provides an additional improvement for the EEG signal analysis.

CONCLUSIONS

An approach for automated multimodal EEG and face-based estimation of human emotions was presented. The accuracy was investigated for different levels of EEG arousal and valence. For EEG, the maximal accuracy of 68.2% is achieved when the arousal and valence is classified in only three levels. The improvement of the classification is delivered via parallel face emotion analysis system based on PCA. The used classifiers are SVM and proposed score-level decision fusion. In our future work we will seek to implement future improvement of developed interface using Active Appearance Models as well as 3D facial emotion recognition as further improvement in order to achieve our objective and create HCI for people with motor disabilities.

ACKNOWLEDGEMENT

This paper is a part of a research project of the Scientific Research Fund at the Bulgarian Ministry of Education and Science DFNI I02/2014 "Human-computer interface for medical assisting systems for life improving of people with propelling problems". The authors are thankful for financial support.

REFERENCES

Horlings, R., Datcu, D. and Rothkrantz, L. J. M. 2008. "Emotion Recognition using Brain Activity," Proc. of CompSysTech'08, vol. II, 1–6.

Koelstra, S., Mühl, 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," IEEE Transactions on Affectiv Computing, vol. 3 (1), 18–31.

Kothe, C., Makeig, S., Soleymani, M. and Onton, J. 2013. "Emotion Recognition from EEG During Self-Paced Emotional Imagery," Proc of the Humaine Association Conf. on Affective Computing and Intelligent Interaction (ACII), 855–858.

Li, M. and Lu, B.-L.2009."Emotion classification based on gamma-band EEG", Engineering in Medicine and Biology Society, Proc. of. The Annual Int. Conf. EMBC 2009, 1223–1226.

Peng, H., Long, F. and Ding, C. 2005. "Feature Selection Based on Mutual Information: Criteria of Max-Dependency, Max-Relevance, and Min-Redundancy," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 27 No 8, 1226–1238.

Radeva, S. and Radev, D. 2015. "Brain - Computer Interface for Communication via Record Electrophysiological Signals". Proc. of the Second Int. Conf. on Digital Information Processing, Data Mining and Wireless communications, Dubai, UAE, 38–48.

Radeva, S. and Radev, D. 2015. "Signal Processing at Different Mental Tasks Execution for Brain-Computer Interface". Proceedings of the ESM'2015, Leicester, UK, 159 – 163.

Velchev, Y., Radev, D. and Radeva, S. 2016. "Features Extraction Based on Subspace Methods with Application to SSVEP BCI", International *Journal of Emerging Engineering Research and Technology* Vol. 4, No 1, (Jan), 52-58.

Viola, P., and Jones. M. J. 2004. "Robust real-time face detection." *Int. journal of computer vision*, vol. 57, No. 2 137-154.

Wang, X.-W. Nie, D.and Lu, B.-L. 2011. "EEG-Based Emotion Recognition Using Frequency Domain Features and Support Vector Machines," Proc. of the ICONIP'11, Part I, 734–743.

SIMULATION-BASED USER INTERFACES FOR DIGITAL TWINS: PRE-, IN-, OR POST-OPERATIONAL ANALYSIS AND EXPLORATION OF VIRTUAL TESTBEDS

Torben Cichon and Juergen Rossmann
Institute for Man-Machine-Interaction
RWTH Aachen University
Ahornstrasse 55, 52074 Aachen, Germany
E-mail: cichon@mmi.rwth-aachen.de

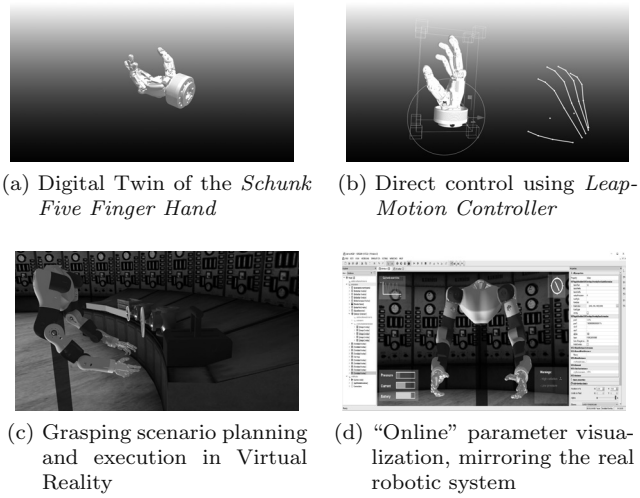
KEYWORDS
3D Simulation, Multi-Body Dynamics, Robotics, Teleoperation, User Interfaces, VR/AR/MR

ABSTRACT

There is a need for a more natural interaction with complex systems in terms of usability, safety, and collaboration potential. No matter if it's the development and operation of a teleoperated robot, supervision of autonomous actions, planning or optimization of industrial processes, an intuitive and direct control possibility by observing all essential information is the key for an optimal man-machine-interaction. In our case, such UI then comprise (a) an experimentable digital twin of the real asset, (b) intuitive control by means of new UI hardware, (c) the virtual exploration of the evaluated process, (d) the preparation and visualization of sensor data, and (e) the process evaluation before, during or after the execution of a task. This holistic approach puts VTBs in between the user and the real system during the development, the execution, and the evaluation phase. This should enable the user to intuitively and safely interact with complex systems utilizing 3D simulation.

MOTIVATION

New developments in Virtual Reality (VR) and everything accompanied led to a rising interest in this kind of technology regarding 3D software and Input/Output (I/O) hardware. But still, in research not only an appealing visualization is important but also the physical plausibility is fundamental. The applications of VR are momentarily mainly based on consumer electronics and computer games but seams to strive more and more into the scientific research. Additionally, the interaction, cooperation, or even collaboration of man and machine represent key characteristics of next generation robotics. Working without cages and sharing the same work space humans can safely interact with robots physically, in application fields like industry, ambient assisted living, or search and rescue to name a few. Mainly, this research tries to enhance the symbiosis of the human leading and the technology enabling on the one hand, and on the other hand decreasing risks of failure and enhancing safety of operator and system. Combining technology, knowledge, and research in these applications with the use of 3D simulation software and current VR hardware seams to be a promising way of reducing the complexity of the underlying system, enhancing the safety, and bridging the gap between developer, user, and the overall system and process.



(a) Digital Twin of the *Schunk Five Finger Hand*

(b) Direct control using *Leap-Motion Controller*

(c) Grasping scenario planning and execution in Virtual Reality

(d) "Online" parameter visualization, mirroring the real robotic system

Figure 1: From development to in-process monitoring: Using Virtual Testbeds and the *LeapMotion Controller* for intuitive manipulation tasks.

INTRODUCTION

We want to motivate our approach by exemplary show manipulation processes (cf. Fig. 1) with the help of such simulation-based user interfaces (UIs).
Intelligent and soft robotic systems are in the focus of today's research with regards to man-machine interaction and man-robot collaboration. Due to the dexterity and the usability of human-like tools this kind of man-robot cooperation uses hand-like end effectors, for example the *SCHUNK 5-Finger-Hand*, first introduced in Liu et al. (2008). Virtual Testbeds (VTBs) can then be used for

designing and testing the robotic hardware with a virtual model during its development phase (Fig. 1a). Additionally, a sophisticated rigid body simulation can be used to evaluate feasible grasping scenarios of the hand already in this early stage of developments. But with state of the art user interfaces for such systems it can be a time consuming, tedious process of applying the correct force/torque on the different motors of the phalanges. Thus, we use the *LeapMotion Controller*[1] as a **natural user interface**, which directly maps the joints of the human hand to the virtual robotic ones (Fig. 1b). Additionally, grasping is not limited to the hand itself, it is more an overall process of full-body movement interacting with the environment. Like picking up a coin from a table it is natural for humans to use multiple hands for swiping it to an edge or exploiting the environment to reduce the complexity of the task. Therefore, Virtual Testbeds can be used to push the manipulation process into its context, of for example grasping and turning a valve with a teleoperated mobile robotic system, like in a real scenario (see Fig. 1c).

Another central aspect of user interfaces is the visualization. Besides state of the art VR technology, like *Oculus Rift* and *HTC Vive* for VR, *Vuzix MR300* and *Google Glass* for AR, or *Microsoft Hololens* and *Acer Mixed Reality Headset* for MR, the pre-processing and visual preparation and presentation of external and internal information is of paramount importance. The user should directly and intuitively see all necessary information in his sight. Therefore, we want to use 3D simulation to generate highly customized user interfaces (like in Fig. 1d) to visualize for example the battery consumption or forces and torques of each joint motor, either of the real robotic system or its digital twin.

Thus, establishing Virtual Testbeds and their 3D simulation core as the **central interface** for the user to interact with complex systems is the main intention of this contribution. This comprises (a) new input devices for VTBs, (b) enhanced visual aids generated within the VTB, and (c) ensuring seamless connectivity to internal VTB modules and external hardware. Combining such new user interface constituents and then exploring VTBs is the pivotal use case of this paper. This exploration starts already in the design and development phase, where modeling becomes more natural. But it can also be used "online", mirroring a real system, or even afterwards for evaluating and optimizing processes. All in all, such new user interfaces should lead to a holistic system of a human user, the VTB as a mediator, and the real system.

RELATED WORK

In this chapter we will present the related work w.r.t the general use of digital twins and Virtual Testbeds, the interface between user and complex system, and finally the use of simulation technology as a mediator.

Virtual Testbeds Using Digital Twins

No matter in which application, simulation tools are almost used in every field of research. Mostly, these tools are focusing on individual aspects or specific application areas. A more holistic approach to 3D simulation, especially used in robotics, is provided by so-called Virtual Testbeds (VTBs), presented for example in Rossmann et al. (2013), where complex technical systems and their interaction with prospective working environments are first designed, programmed, controlled, and optimized in 3D simulation, before commissioning the real system.

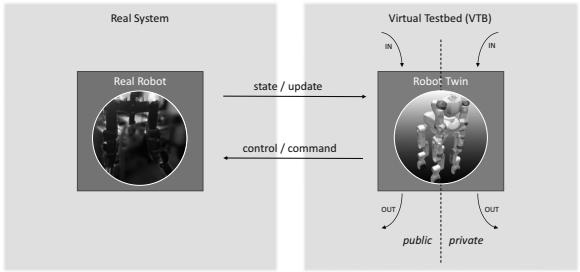


Figure 2: Digital Twin: Example of a digital twin including data exchange and privacy conditions

Within these VTBs we use **digital twins** of the real system. Exemplary shown in Fig. 2 we present the digital twin of a mobile robot (in its development stage), where we can use the virtual asset to test and develop the asset itself and also everything interacting with it. Additionally, the digital twin can be updated according to the current robot' state and could also control or send individual commands to its real counterpart. The privacy condition of the digital twin can be useful to separate externally retrievable and private data. But all in all, this whole system comes down to the simplified essential requirement of providing the **same in- and output** of the real asset and corresponding digital twin, as shown in Fig. 3.

Figure 3: Simplified equality of asset and digital twin

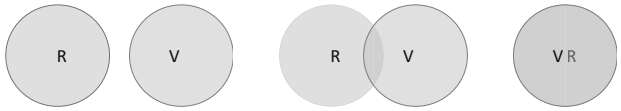
Interfacing Digital Twin and Real System

The User Interface (UI) is the layer where interactions between human users and the VTB occur. This inte-

[1] https://www.leapmotion.com

raction includes the **input** to and **output** from the system.

I/O Technologies
The input can be categorized in terms of different input device technologies, like "controller-based", "optical", or even "haptic" devices. Fong et al Fong and Thorpe (2001) already defined major interface types for (vehicle) teleoperation, where they named VR and haptic interfaces "Novel". Such a novel haptic feedback in 3D simulation is quite rare in current research and can mostly be found in rehabilitation applications. Regarding teleoperated robots, we already conducted this approach using a customized exoskeleton and VTBs in Cichon et al. (2016a).



(a) VR (b) AR (c) MR

Figure 4: Distinction and fusion of virtuality (V) and reality (R) based on data source

The output of the system in state-of-the-art 3D simulation tools is mostly done on standard monitors but can also be extended to "Virtual Reality" (VR), "Augmented Reality" (AR), and "Mixed Reality" (MR). These fields of visual output can be categorized by the associated hardware, primer purpose of use, or the amount of fusion from virtuality and reality like in Fig. 4.
Although it is not directly part if the user interface itself, to bridge the gap from user to real system, the interface between VTB and final hardware is of course also important. As the mostly used middleware in robotics we use the Robot Operating System (ROS). A first introduction about ROS is given in Quigley et al. (2009), whereas in Muratore et al. (2017) a real-time capable approach of mobile robotics is presented which can also be integrated in different 3D simulator software.

Using Simulation In-the-loop
Although the use of simulators in-the-loop or even as a mediator is quite limited, the general approach of using models for teleoperation is an ongoing research topic. Willaert et al. (2012) discussed the approach of model-mediated telemanipulation especially with the goal of stability assurance. Special focus is put on the model consistency (model adjustments versus discrete model jumps) where they underlined the power of using models for prediction. The approach of using so-called mental models for human robot interaction is motivated in Sheridan (2016) and brought to application in Cichon et al. (2016b) as an conceptual extension of VTBs towards simulation-based control and support.

UI Software
A general support operator setup during teleoperation of mobile robots can be found in Schwarz et al. (2017), where they also give a current view on the application site of mobile robotics utilizing ROS functionalities. Such visualization comprises different tools to visualize for example (a) live images, (b) 3D point clouds, (c) command line error log, (d) actuator diagnostics, (e) 2D height map, (f) network statistics, etc. An introduction to operator interfaces utilizing AR and VR is given in Roßmann et al. (2010), where special visualizations are applied in search-and-rescue applications.

UI Hardware
Using state-of-the-art technology as input devices for robotic control was already done in several fields of research. The evaluation and analysis of the devices themselves, comparing specifications, and their use in various fields of application are the main foci of research. Regarding the *LeapMotion* sensor Weichert et al. (2013) use an industrial robot to analyze the devices' accuracy and repeatability, whereas Guna et al. (2014) combine the technology with high-precision motion tracking to show the prospects and limits of *LeapMotion* as a professional tracking system. In the application field of ambient assisted living Bassily et al. (2014) use the *LeapMotion* for human-robot interaction for elderly or physically impaired people. Gromov et al. showed in Gromov et al. (2016) the use of a gesture recognition bracelet using IMU and EMG signals of the user's arm, namely *Thalmic Labs MYO*[1]. The *MYO* was used to control multiple mobile robots by speech, arm movement, and hand gestures to select, localize and communicate task requests and spatial information.

CONCEPT

Although the concept of simulation-based cognitive man-robot collaboration encompasses more, we will focus here on simulation-based UI and thus the in- and output of the VTB to develop, use, explore, and analyze digital twins in VTBs. Thus, we will present the four main compartments: Fusion of real and virtual data, natural and intuitive UI input, flexible and feasible UI output, as well as fast and simple data analysis throughout the whole life cycle of a system.
Combining real and virtual information can lead to a new approach on how to use and interact with complex systems. Diving into virtual worlds instead of just looking at (raw) data on a monitor can create a deeper sense of **immersion**. This immersion can then be enhanced by using the correct UI in terms of hardware and software.
Although keyboard and mouse are naturally not the most **intuitive input devices** on the market, long year

[1] `https://www.myo.com`

usage led to a widespread acceptance and intuition of the user and thus simple buttons and mouse movements can be the method of choice for intuitive UI. Going one step further, the most natural interface of a human operator is mostly his body and especially his hands. Thus, we want to include new UI hardware which support body/finger tracking and even force feedback.
Due to the human fixation on seeing and feeling things, audio- visual- and haptic- **feedback** are core elements of new user interfaces. Haptic feedback in Virtual Testbeds, applied on exoskeletal teleoperated mobile robots has already been presented in Cichon et al. (2016a), and will only be addressed briefly in the scope of this paper. Additional help by visual means will be addresses regarding rendering techniques, as well as feasible data representations. Although audio feedback is also helpful, we will address this (for now) just mediated by the real system directly and not via 3D audio simulation.
Within the scope of such new user interfaces we want to be able to **explore** the digital world. This comprises the exploration of the digital twin itself, but also in the direct sense of exploring environments with the help of VTBs. This can then be used prior to fabrication of the system by testing and optimization, as well as "online" during the use of the real system. Additionally, VTB functionalities, like logging presented in Atorf et al. (2015), can be used for a post-mission wrap-up and evaluation. Overall, generating such new user interfaces for VTBs enables us to predict possible outcomes and answer some "What if...?" questions ahead to their execution.

IMPLEMENTATION

The following section describes the implementation of the proposed concept. This covers the Input/Output hardware with an integration layer, the visualization of data, and the final setup of an ideal user interface for a given application utilizing real and virtual data.

Input/Output

We implemented a generic integration of different UI hardware to use the full spectrum of VTBs, utilizing all internal frameworks. A modular, object-oriented implementation scheme leads to a layer of input devices that can be abstracted from the simulation system, but easily connected with for example the Rigid Body Simulation (RBS), the Sensor Simulation (SS), the Rendering (R), or the ROS framework (see Fig. 5).
For now, we have implemented the *LeapMotion* sensor, *Thalmic Labs MYO* wristband, and even haptic input devices like the *Geomagic Touch X*. Due to the fact that all input devices are implemented following the same scheme we will present the **LeapMotion** sensor exemplary in the following in more detail.
From the user perspective the point of origin is one *LeapMotionExtension*, which can be added to any 3D model in a VTB. This Extension holds an object *LeapMotionDevice* in the I/O communication layer. The *LeapMotionDevice* can then establish the communication to one *LeapMotion* sensor, with the *LeapMotion* SDK and the device driver. It polls the current data (frames, hand exoskeleton, ...) from the sensor, which are directly relayed to the Extension where the data types are converted and written to the I/O-board. With this scheme all hardware dependencies are encapsulated in one object which can also be transferred to a separate thread for example.

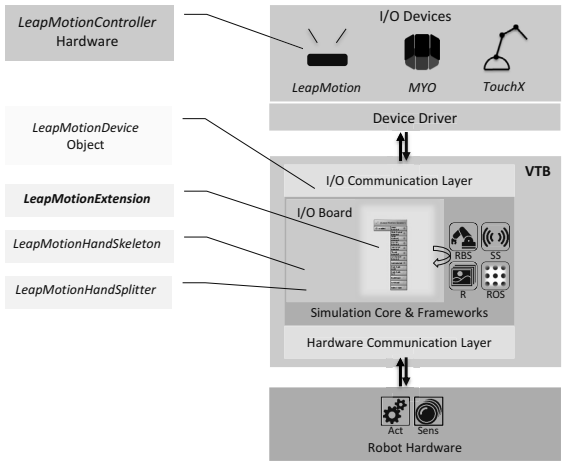


Figure 5: Implementation Scheme of a variety of I/O Devices, especially the *LeapMotion* Sensor

Using the I/O board of the VTB we can then access all data of the extension. Additionally, we can connect each property to another I/O board of the same type. To simplify these connections it is even possible to extend the I/O data types with a *LeapMotionHandSkeleton* which can be transmitted to a *LeapMotionHandSplitter* node handling the mapping. For example, we can use the hand frame of the *LeapMotion* to move the frame of the Schunk hand, and connect the phalanges position (via the splitter node) to rigid body based motors to move each finger according to the user's hand movement.
Additional information of the sensor, like raw data or an included gesture recognition, can then also easily be incorporated in the I/O framework. This leads to an infinite amount of possibilities to connect single or multiple gestures to all aspects of simulation. One example would be the direct control of the Schunk hand, only if the *MYO* wristband recognizes a "fist" gesture.

All in all, such "natural" UI hardware increases the so-called **embodiment** of the user. This embodiment is supported by visualizing the human body in VR (cp. Fig. 6), which is very important for an intuitive use.
As a further step, it is also possible to use the ROS inter-

face (presented in Cichon et al. (2016c)) to connect the VTB with ROS-compatible hardware. Thus, the UI hardware can directly be used to interact with real systems by connecting the I/O Board to the ROS framework. Consequently we can directly control ROS-capable hardware, like for example the real *Schunk Hand*, mediated via the VTB.

Data Visualization

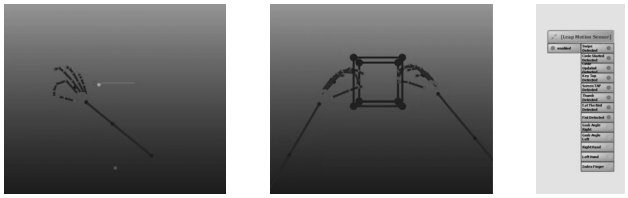
We extended the scope of the visualization framework and expanded its capabilities to support stereoscopic VR goggles. This led to the possibility to define customized head-up-displays for monitor or VR views and project costmaps or occupancy grids into the 3D scene. Additional immersion could be achieved by using a live ROS audio stream, directly transmitted to the user. With respect to the aforementioned state of the art in teleoperation and visualization and the GUI of Schwarz et al. (2017) we can incorporate parts of it within the VTB and also push selected information into the view of a first person operator with a stereoscopic headset.

Fusion of Reality and Virtuality

Utilizing the aforementioned Input/Output and visualizations it is now possible to design a customized user experience for a given application. This includes the choice of an adequate input hardware selection (which can also be a combination of multiple hardware devices), and a customized visual head-up-display.
Combining visualization and input devices with gesture recognition we can even move and place user interfaces in the 3D scene. One prominent example is using the 'bloom' gesture to show an interface which then can be used to edit some properties.
Due to the used implementation and the underlying VTB it is now also possible to choose which data source to use, virtual or real.

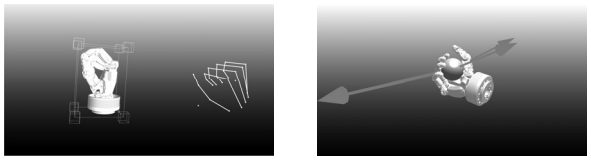
APPLICATIONS

Applications for using simulation-based user interfaces for digital twins range from single aspect to whole processes and systems in their development, assessment, or in use. Of course, the first aspect of using such new input devices is to model something in the 3D world. This comprises simple geometric modeling of boxes (see Fig. 6b) but also virtual interactions like buttons or sliders (see Fig. 6a).



(a) Interaction: Slider (b) Modeling (c) I/O Board

Figure 6: Leap Motion Extension used for modeling



(a) Movement and Grasping Types (b) Force Mangnitude and Contact Evaluation

Figure 7: Manipulation Force Analysis

The aforementioned manipulation and grasping scenario is one aspect of the CENTAURO project[1]. Thus, before an exoskeleton is manufactured or the real *Schunk Hand* is not at hand, it is possible to use the *LeapMotion* with the digital twin to evaluate grasping forces in simulation first (see Fig. 7).
The overall goal of the CENTAURO project is to develop a human centered teleoperated mobile robotic system for disaster scenarios. Besides a robust and dexterous robot, one main focus is on the UI to reduce the workload of the operator. Using the simulation-based UI, the operator should be able to inspect all necessary information as he needed visualized in a VR goggle. Additionally, he should be able to switch from the direct control of the real system to its digital twin in the VTB which uses the real sensor input to generate an environment model in the rigid body simulation. Switching into VR then allows to test possible actions safely first, before execution in reality.

Such new user interfaces can also be used in industrial application. One example is the monitoring of assembly processes. Here, the VTB can be used to visualize internal critical parameters prior to failure, or can be used in AR to show possible alternative processes.

CONCLUSION AND OUTLOOK

Taking everything into consideration, we have presented a holistic UI framework which is already in use for a set of UI hardware visualized on VR hardware. It comprises multiple input devices using the same implementation scheme and interface layer no matter if controller based (*MYO*), optical (*LeapMotion*), or even haptic (*Touch X*). It enables feasible data visualizations even for stereoscopic or dynamic user interfaces. Finally, a modular integration into the VTB allows access to all other frameworks incorporated into the 3D simulation system, independent of the application. Combined with a ROS hardware interface for a direct connection and feedback loop this is already in everyday use. The coinciding pos-

[1] `https://www.centauro-project.eu`

sibilities and the modularity of the overall implementation leads a more immersive and embodied experience interfacing complex systems. It becomes more natural to evaluate systems and explore their use and the virtual world itself in VTBs, enabling the user to analyze the system pre-, in-, and post operation.
Besides various other prospects, we already initiated developments of using these UIs and VTBs in AR or MR hardware in industrial applications, which seems to be very promising. Using the developed UI framework in the *Microsoft Hololens* or the *Vuzix M300* could lead to a good collaboration of industrial robots and humans, working side by side. Assembly processes, workplace optimizations, or worker guidance with the help of VTBs are just some examples of the possible applications.

ACKNOWLEDGEMENTS

This project has received funding from the European Unions Horizon 2020 research and innovation program under grant agreement No 644839.

REFERENCES

Atorf L.; Cichon T.; and Roßmann J., 2015. *Flexible Data Logging, Management, and Analysis of Simulation Results of Complex Systems for eRobotics Applications*. ESM.

Bassily D.; Georgoulas C.; Guettler J.; Linner T.; and Bock T., 2014. *Intuitive and adaptive robotic arm manipulation using the leap motion controller*. In *ISR/Robotik 2014; 41st International Symposium on Robotics; Proceedings of*. VDE, 1–7.

Cichon T.; Loconsole C.; Buongiorno D.; Solazzi M.; Schlette C.; Frisoli A.; and Roßmann J., 2016a. *Combining an exoskeleton with 3D simulation in-the-loop*. In *9th International Workshop on Human Friendly Robotics (HFR)*. 31–34.

Cichon T.; Priggemeyer M.; and Roßmann J., 2016b. *Simulation-based Control and Simulation-based Support in eRobotics Applications. Applied Mechanics & Materials*, 840.

Cichon T.; Schlette C.; and Roßmann J., 2016c. *Towards a 3D simulation-based operator interface for teleoperated robots in disaster scenarios*. In *Safety, Security, and Rescue Robotics (SSRR), 2016 IEEE International Symposium on*. IEEE, 264–269.

Fong T. and Thorpe C., 2001. *Vehicle teleoperation interfaces. Autonomous robots*, 11, no. 1, 9–18.

Gromov B.; Gambardella L.M.; and Di Caro G.A., 2016. *Wearable multi-modal interface for human multi-robot interaction*. In *Safety, Security, and Rescue Robotics (SSRR), 2016 IEEE International Symposium on*. IEEE, 240–245.

Guna J.; Jakus G.; Pogačnik M.; Tomažič S.; and Sodnik J., 2014. *An analysis of the precision and reliability of the leap motion sensor and its suitability for static and dynamic tracking. Sensors*, 14, no. 2, 3702–3720.

Liu H.; Wu K.; Meusel P.; Seitz N.; Hirzinger G.; Jin M.; Liu Y.; Fan S.; Lan T.; and Chen Z., 2008. *Multisensory five-finger dexterous hand: The DLR/HIT Hand II*. In *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*. IEEE, 3692–3697.

Muratore L.; Laurenzi A.; Hoffman E.M.; Rocchi A.; Caldwell D.G.; and Tsagarakis N.G., 2017. *XBotCore: A Real-Time Cross-Robot Software Platform*. In *IEEE International Conference on Robotic Computing, IRC17*.

Quigley M.; Conley K.; Gerkey B.; Faust J.; Foote T.; Leibs J.; Wheeler R.; and Ng A.Y., 2009. *ROS: an open-source Robot Operating System*. In *ICRA workshop on open source software*. vol. 3, 5.

Rossmann J.; Kaigom E.G.; Atorf L.; Rast M.; and Schlette C., 2013. *A Virtual Testbed for Human-Robot Interaction*. In *Computer Modelling and Simulation (UKSim), 2013 UKSim 15th International Conference on*. IEEE, 277–282.

Roßmann J.; Kupetz A.; and Wischnewski R., 2010. *An AR/VR based approach towards the intuitive control of mobile rescue robots. International Journal of Computer, Electrical, Automation, Control and Information Engineering*, 4, no. 11, 411–418.

Schwarz M.; Rodehutskors T.; Droeschel D.; Beul M.; Schreiber M.; Araslanov N.; Ivanov I.; Lenz C.; Razlaw J.; Schüller S.; et al., 2017. *NimbRo Rescue: Solving Disaster-response Tasks with the Mobile Manipulation Robot Momaro. Journal of Field Robotics*, 34, no. 2, 400–425.

Sheridan T.B., 2016. *Human–Robot Interaction Status and Challenges. Human Factors: The Journal of the Human Factors and Ergonomics Society*, 58, no. 4, 525–532.

Weichert F.; Bachmann D.; Rudak B.; and Fisseler D., 2013. *Analysis of the accuracy and robustness of the leap motion controller. Sensors*, 13, no. 5, 6380–6393.

Willaert B.; Van Brussel H.; and Niemeyer G., 2012. *Stability of model-mediated teleoperation: Discussion and experiments. Haptics: Perception, devices, mobility, and communication*, 625–636.

SURGERY ASSISTANT BASED ON AUGMENTED REALITY

Anton Ivaschenko
Samara National Research University
34 Moskovskoye shosse,
443086 Samara, Russia
E-mail: anton.ivashenko@gmail.com

Alexandr Kolsanov, Aikush Nazaryan
Samara State Medical University
89 Chapayevskaya st.,
443099 Samara, Russia
E-mail: avkolsanov@mail.ru

KEYWORDS

3D anatomy, augmented reality, image-guided surgery, simulation, visualization, medical data.

ABSTRACT

On the basis of recent advances in surgery modeling and augmented reality there were developed a new solution for surgery assistance in real time. The solution consists of 3 modules: preoperative planning, 3D imaging and surgery navigation. New simulation models and algorithms were introduced for surgery focused visualization and decision making support. The developments were successfully probated at clinics of Samara State Medical University for a number of medical cases. This paper describes the proposed solution and its implementation in practice.

INTRODUCTION

Augmented reality (AR) is one of the most challenging technology trends of simulation and modeling nowadays. At the same time the problem of its implementation in image-guided surgery remains open. Considering the existing capabilities of available equipment there was developed an original solution for AR surgery assistant based on anatomy focused visualization (Ivaschenko 2015, 2016, 2017).

Under this research there was developed a special hardware solution to capture the movements of surgery instruments and coordinate them with surgeon's focus by means of specifically designed headset. In addition to it there was implemented a software solution capable to simulate individual 3D surgery scenes based on reconstruction of CT / MRI images, support preoperative planning, and provide visualization of surgery scenes over the human body for surgery navigation and decision making support. Details of the proposed approach are presented below.

STATE OF THE ART

Image-guided surgery (IGS) is being successfully implemented in modern neurosurgery and helps to perform safer and less invasive procedures (Mezger 2013, Galloway 2015). During the IGS procedure the surgeon uses tracked surgical instruments in conjunction with preoperative or intraoperative images in order to indirectly guide the procedure. IGS systems present the patient's anatomy and the surgeon's precise movements related to the patient to computer monitors in the operating room. The source images are captured by cameras or electromagnetic fields.

IGS refer to the area of computer assisted surgery (CAS) that is based on application of various computer technologies for surgical planning and guiding or performing surgical interventions (Haaker 2005). CAS include the phases of medical imaging (using CT, MRI, X-rays, ultrasound, etc.), image analysis and processing, preoperative planning and surgical simulation, surgical navigation, and robotic surgery. Using the surgical navigation system the surgeon uses special instruments, which are tracked by the navigation system.

The basic challenge of image-guided surgery implementation in practice is the complexity of data visualization. Surgeon needs the convenient, adequate and real-time picture being presented over the area of surgery intervention scene. At the same time the presented image should not be overcomplicated with minor details attracting attention to the most essential points. Therefore there is a strong request for optimization of medical data visualization in image-guided surgery.

Major modern trends of medical data visualization are widely explored in (Holzinger 2014). It is noted that the main goal of medical data visualization is to combine several data sets to analyze multiple layers of a biological system at once. The system should interlink all related data sets (e.g., images, text, measured values, scans) and offer visual analytics to support experts. This approach supports the idea of maximum effective visualization of complex medical data for medical professionals instead of automatic decision-making.

Interactive visualization is strongly required in various software solutions that provide decision-making support. Interactive IT devices, including VR and AR goggles, tablets and panels allow tracking the user's activity, which can be used to understand the user's behavior aspects and patterns and adapt the interactive application logic of data presentation. Successful application of 3D web-based anatomical visualization tool using virtual reality (VR) technology is presented in (Said 2015).

AR has high perspectives of its implementation in surgery. Using modern solutions (Krevelen 2007, Navab 2004, Rusev 2014) there can be solved a number of problems specific for medical data visualization. The goal of user interface design

is to make the user's interaction as simple and efficient as possible, in terms of accomplishing user goals. Technological advances, exploding amounts of information, and user receptiveness are fueling AR rapid expansion from a novelty concept to potentially the default interface paradigm in coming years (Singh 2013).
At the same time AR faces the same core usability challenges as traditional interfaces, such as the potential for overloading users with too much information and making it difficult for them to determine a relevant action. However, AR exacerbates some of these problems because multiple types of augmentation are possible at once, and proactive apps run the risk of overwhelming users.

SOLUTION VISION

The proposed solution for AR surgery assistant is developed as a multifunctional complex that allows planning surgical operations based on preoperative MRI and CT examinations by building a 3D model of internal organs and tissues. The solution includes the systems of 3D imaging, preoperative planning, and surgery navigation. The solution is illustrated by Fig. 1.

Figure 1: The Concept of AG Surgery Assistant

Based on CT and MRI studies of the patient radiologist creates 3D reconstructions of the zones of surgery operational interest, and tissues the subject to destruction. Consequently the complex saves the time of the radiologist due to automatic segmentation of vessels, organs and neoplasm. There was developed an original segmentation technology described in (Nikonorov 2014). One the next stage the surgeon plans an operation based on the resulting 3D model. During this process he establishes anatomical landmarks, chooses the trajectory and safe limits of surgical access.

The navigation system allows displaying the 3D model using AR goggles. The individual anatomical model is projected over the patient's actual organ, together with the data on the surgery plan and the current clinical parameters of the patient. Tracking system includes originally designed headset, markers surgery instruments and a specialized tracking system (see Fig. 2). Original design allows capturing the movements of surgery instruments coordinated with the surgeon focus.

At the beginning the visualization system projects a personified anatomical 3D model onto the skin. Due to this, the surgeon determines and outlines the optimal entry points, reduces the surgical field and the volume of surgical access. During the operation, the surgical navigation system provides the surgeon with continuous monitoring of the access trajectory, the position of the surgical instrument and comparison with the surgical plan. Thus, it helps the surgeon consciously increase the radicalism and accuracy of surgical intervention.

The system targets the maximum safety and efficiency of complex surgical interventions with minimal damage to the patient's tissues. Therefore the system provides:

- assistance in the analysis and planning of a future operation in a visual 3D model based on a preoperative CT / MRI study;
- navigation during the operation with the help of the visualized model and the operation plan superimposed on the surgery field.

To solve the mentioned above challenges of AR implementation in practice the standard functionality was extended by the possibility of surgeon’s focus coordination based on intelligent analysis of the intervention process. Embedded software with intelligent decision-making support captures the user’s focus in the form of event chains and compares them with typical scenarios of surgery intervention.

IMPLEMENTATION

AR headset equipped with surgery instruments markers and tracking system is presented in Fig. 4. The designed headset allows presenting 3D models of human body parts and other medical data that is contextually required over the real surgery intervention field.

Figure 2: AG Headset Adopted for Surgery Assistant

The human body is represented by 3D models with realistic appearance and possibility to interact with surgery instruments.The inner parts of human body are simulated in the scene by soft body models and surgery instruments are

simulated by rigid bodies. These models are fashioned with the help of specifically designed shaders and algorithms that simulate jars, liquids and blood.

3D simulation of human body parts for real medical cases is based on automated reconstruction of KT and MRI images with post processing in manual mode. The resulting image can be used by surgeon for preliminary planning of surgery intervention and modeling of possible alternatives.

The system allows loading medical data in DICOM format from various sources and exporting images to JPEG and PNG.

Solution functionality includes:

- 3D reconstruction of typical and atypical anatomical structures with visualization of normal and pathological cases on the basis of X-ray studies (organs, vessels, ducts, affected parts);
- surgery intervention planning by setting the basic stages and visual anatomical information;
- visualization of personified topographic and anatomical data of the patient in augmented reality mode;
- accompanying surgical intervention with monitoring the location of surgical instruments;
- training of students and beginners in the video material recorded during the operation.

Fig. 3 – 4 present the results of human body parts 3D reconstruction for surgery planning and intervention decision making support.

Figure 3: 3D Surgery Scene Reconstruction and Simulation

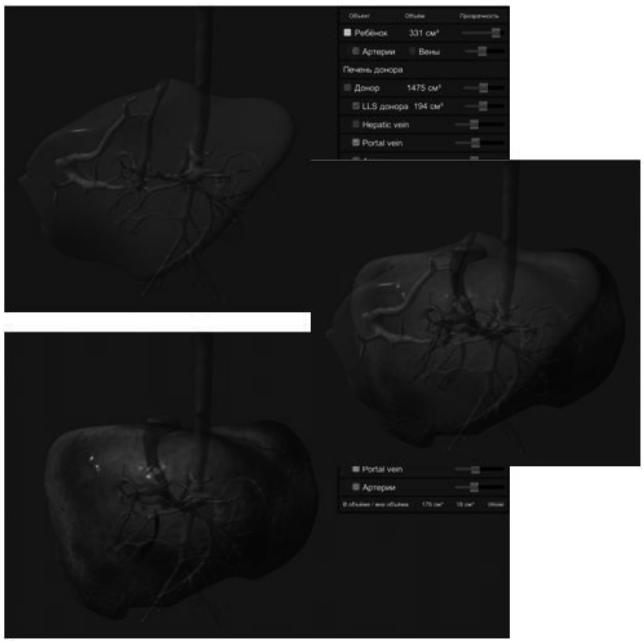


Figure 4: Planning the Liver Partial Transplantation

The given example illustrates selection of the proportion of the liver in a related transplantation.

CONCLUSION

The proposed AR surgery assistant is based on implementation of AR and provides preoperative planning, 3D imaging and surgery navigation. There were developed an original solution for hardware and software that provides focused visualization for surgery intervention.

The solution allows to increase the radicality and precision of surgical intervention, reduce the damage to the patient's tissues, the time of surgery and the amount of blood loss; obtain a weighted surgical plan with the optimal resection volume, entry points, access pathway; improve communication between doctors in interdisciplinary cases and ensure the conduct of training rehearsals procedures

REFERENCES

Galloway, RL Jr. 2015. “Introduction and historical perspectives on image-guided surgery”. *In Golby, AJ. Image-Guided Neurosurgery*. Amsterdam: Elsevier. 3-4.

Haaker, RG.; M. Stockheim, M. Kamp, G. Proff, J. Breitenfelder, A. Ottersbach. 2005. “Computer-assisted navigation increases precision of component placement in total knee arthroplasty”. *Clin Orthop Relat Res* 433, 152-9

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.

Ivaschenko, A.; A. Kolsanov, A. Nazaryan, A. Kuzmin. 2015. “3D surgery simulation software development kit”. *Proceedings of the European Simulation and Modeling Conference 2015 (ESM 2015)*, Leicester, UK, EUROSIS-ETI. 333-240

Ivaschenko, A.; N. Gorbachenko, A. Kolsanov, A. Nazaryan, A. Kuzmin. 2016. “3D scene modelling in human anatomy simulators”. *Proceedings of the European Simulation and*

Modeling Conference 2016 (ESM 2016), Spain, EUROSIS-ETI. 307-314.

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

Krevelen, R. 2007. "Augmented Reality: technologies, applications, and limitations". Vrije Universiteit Amsterdam, Department of Computer Science.

Mezger, U.; C. Jendrewski, M. Bartels. 2013. "Navigation in surgery". *Langenbecks Arch Surg.* 398: 501-14.

Navab, N. 2004. "Developing killer apps for industrial Augmented Reality". Technical University of Munich, *IEEE Computer Graphics and Applications* IEEE Computer Society.

Nikonorov, A.; P. Yakimov, Y. Yuzifovich, A. Kolsanov. 2014. "Semi-automatic liver segmentation using Tv-L1 denoising and region growing with constraints". *9th German-Russian Workshop on Image Understanding*, Koblenz, Germany. 1-4.

Rusev, I.; R. Ruisev, T. Vassilev. 2014. "An approach for implementing an intelligent user interface". *International Journal on Information Technologies and Security*, No. 4 (vol. 6), 43-50

Said, C.S.; K. Shamsudin, R. Mailok, R. Johan, H.F. Hanaif. 2015. "The Development and Evaluation of a 3D Visualization Tool in Anatomy Education" *EDUCATUM - Journal of Science, Mathematics and Technology.* Vol. 2. No. 2, 48-56.

Singh, M.; M.P. Singh. 2013. "Augmented Reality interfaces". *Natural Web Interfaces* IEEE Internet Computing. 66-70

BIOLOGICAL DATA SIMULATION

THE CONTROLLED DEVELOPMENT OF (MEDICAL) SELF-DIAGNOSIS SYSTEMS WITH SELF-ENFORCING NETWORKS

Christina Klüver and Jürgen Klüver
University of Duisburg-Essen
Computer Based Analysis of Social Complexity
45117 Essen, Germany
christina.kluever@uni-due.de; juergen.kluever@uni-due.de

KEYWORDS
*S*elf-Enforcing Network, self-organized learning, cue validity factor, neural networks, medical self-diagnosis, monitoring symptoms

ABSTRACT

The usage of the Internet to find medical information and the increasing number of according websites and apps has been discussed controversially over years. In this article a Self-Enforcing Network (SEN), which is a self-organized neural network, is proposed to develop medical self-diagnosis systems, which enables different applications according to the search behavior of users, but containing the knowledge of experts. Two different prototypes are described: the first one can be used as "system-checker", starting with symptoms inserted by users and SEN proposes possible diagnosis; the second one is based on an existing diagnosis, which often leads to further disorders. The last one can be used e.g. for "monitoring" symptoms. The systems can be easily enlarged according to frequent searches in the Internet.

INTRODUCTION

In last years the worldwide search of websites to find medical information and/or advices is increasing (e.g. Griggs, 2015; Zuccon et al., 2015; Dubowicz and Schulz, 2015; Mueller et al., 2017) and especially self-diagnosis can become a problem because for example "Dr Google" (Avery et al., 2012; Robertson et al., 2015) is often the first contact "partner" and can have an influence on the relationship to the consulted physician (Karnam and Raghavendra, 2017). Yet a general warning to use such websites because of psychological and physical problems resulting of the self-diagnosis (e.g. Aiken et al., 2012; Shen et al., 2015) is not enough because the search for information in the Internet has for a long time become a common practice.
It is no wonder that apps using a name like "'WebMD', 'Doctor Online', 'Virtual Doctor', 'Dr Android MD Diagnosis' and 'Pocket Doctor'" (Lupton and Jutel, 2015) have high visitor numbers, who do not know, which algorithms are used for the according solutions.
The development of suited diagnosis systems is extremely difficult because of a frequently large number of symptoms and according diseases, dependent laboratory outcomes etc. The techniques vary, but the tendency in last years is the use of Computational Intelligence methods:
In general, for different problems in medical application artificial neural networks are used, in particular supervised learning as well as self-organized learning types, as decision support for different kinds of diagnosis (for a short summary e.g. Amato et al., 2013; Pombo et al., 2015). To improve medical diagnosis systems different neural networks and support vector machines (SVM) are developed (Rodríguez et al., 2016; Vassis et al., 2015; Peker et al., 2016). Especially for online self-diagnosis or medical guidance a convolutional neural network is proposed (Yao et al., 2016), a fuzzy-ART as an "intelligent home disease pre-diagnosis system" (Kim, 2016), a long-short-term memory (LSTM) framework to construct the system for self-diagnosis "Android" (Liu et al., 2016), to name only few recent developments.
In Klüver (2016a) a Self-Enforcing Network (SEN) system was proposed inter alia for such purposes, which is reliable with respect to its operations and results. Using the different functionalities of SEN the developer of such a system used as symptom-checker can control the data basis and the results. If a diagnosis is not unambiguous, the symptoms belonging to a diagnosis must be checked if they are all taken into consideration.
In this article further developments are shown as the introduction of the *component-wise analysis* to get more insights about the differences between the data basis and the given symptoms by a user. This functionality is useful when taking into account that a person will have different or diffuse symptoms, which are not necessarily typical for only one disease. In addition the development of another SEN is presented, based on a selected diagnosis, which allows the *monitoring* of symptoms by a person to check if there are initial signs for typical further disorders.
First the basic logic of the network and its usage are described. The next section deals with the methodical approach how to construct a suited database, namely how to combine certain symptoms with the according diseases. Afterwards the operations of the SEN with several examples are demonstrated.
The main goal of this article is to show that the development of SENs as self-diagnostic systems enables a twofold usage, namely as "symptom-checker" and/or "symptoms monitoring". It is easy to methodically ensure the reliability and medical validity of such a diagnosis system, provided that cooperation with health professionals occurs.

THE SELF-ENFORCING NETWORK (SEN)

The SEN is a self-organized learning neural network developed by our Research Group "Computer Based Analysis of Social Complexity" (CoBASC). Here we give only a general description of the network (for details cf. e.g. Klüver and Klüver, 2013; Klüver, 2016 a and b).
To *train* the network, first a *semantical matrix* is needed, containing the objects, the corresponding attributes, and the affiliation degree of an attribute to an object. Depending on the analysis the data basis can be imported as a csv-file containing the real data, or, as in this case, a developer must carefully define the affiliation degree between attributes and objects.
For the construction of a diagnosis system the numerical values of the matrix are obtained by scaling techniques known for long in, e.g., quantitative social research (Freeman, 1989). One might call these methods as quantifications of qualitatively perceived or observed proportional relations (as shown in Table 1).

Main Components Of SEN

The operations of a SEN start by analyzing the values v_{sm} of the semantical matrix and by transforming the values of the semantical matrix into the weight matrix of the network. The weight matrix, hence, is generated from the semantical matrix and *not* at random. This is very important when analyzing real data to ensure that the learning algorithm does not arbitrarily change the values.
In consequence, the weight value w_{oa} between object ***o*** and the attribute ***a*** is the according semantical value v_{oa}:

$$w_{oa} = c * v_{oa} \qquad (1)$$

c is a constant usually defined as $0 \leq c \leq 1$. It has the same function as the well-known learning rate in standard neural networks.
The learning rule of a SEN that varies the values of the weight matrix is:

$$\mathrm{w}(t+1) = \mathrm{w}(t) + \Delta\mathrm{w}, \text{and} \qquad \Delta w = c * w_{oa} \qquad (2)$$

In addition a *cue validity factor* (cvf) is used, which is a measure how important certain attributes are for membership in a given category (Rosch and Mervis, 1975), meanwhile playing a role in different fields (e.g. Klüver and Klüver, 2007, Biel et al., 2016).
By using cvf-values it is possible to distinguish between the degrees of importance of an attribute for the analysis (Klüver, 2017). If the value of the cvf = 1 or higher, the attribute is most important; if cvf = 0 than the attribute is not considered for the training and in consequence not for clustering. The learning rule becomes

$$\Delta w = c * w_{oa} * cvf_a \qquad (3)$$

The *topology* of SEN is dependent on the specific problem (Klüver 2016 a and b); for the development of a diagnosis system a two-layer topology is best suited containing the attributes (symptoms) as input neurons and the according objects (diseases) as output neurons.
As in each neural network the dynamics of a SEN is generated by so-called *activation functions*. A user of a SEN can choose between different activation functions. In all cases a_j is the activation value of the receiving neuron j, a_i are the activation values of the sending neurons i, and w_{ij} as usual are the according weight values. For the diagnosis system the logarithmic-linear function (LLF), developed by us, has the best-suited result:

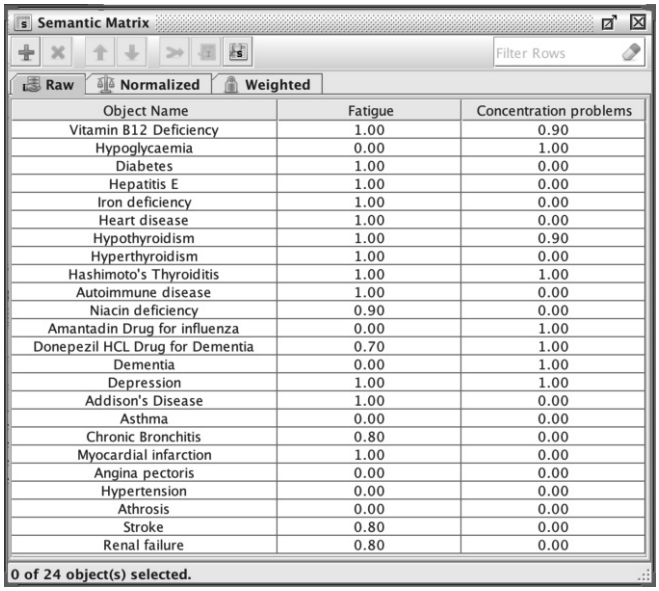
$$a_j = \sum \begin{cases} lg_3(a_i + 1) * w_{ij}, & if\ a_j \geq 0 \\ lg_3(|a_i + 1|) * -w_{ij}, & \text{else} \end{cases} \qquad (4)$$

One can interpret the use of the logarithm as a dampening factor that is "internal" to the function; the basis 3 of the logarithm was chosen simply because basis 2 would generate too small activation values and basis 4 too large values.
After the learning process is finished, a user can insert a so-called *input vector* containing the different attributes, in the context of this article meaning symptoms of a person.
The results of a SEN system are *visualized* in different ways to allow a fast interpretation: the visualization of the computed similarities according to the *highest* activated neuron (ranking), the *smallest* difference between the vectors (distance), and the "map visualization" (Klüver, 2016b), representing the *approximated similarity* between all objects.
If the map visualization of the trained network allows the detection of clusters, polygon features select the objects within one cluster for further analysis.

METHODICAL PROCEDURE

For the first prototype of a SEN based diagnosis system five general symptoms, namely 'fatigue', 'difficulty concentrating', 'listlessness', 'nervousness' and 'sweating', which are e.g. typical symptoms of students before exams, were used to find possible diseases (Klüver, 2016 a).[1]
To get the information about the symptoms, 10 medical and specialized websites, e.g. for endocrinology, are consulted for each disease and in addition the ICD-10 (International Statistical Classification of Diseases and Related Health Problems). The symptoms, which coincide on all websites for a specific disease are interpreted as essential and have a value of 1.0, less typical or rare symptoms have the value of 0.1. Accordingly a pain scale or intensity of different symptoms can be defined.
The first prototype containing 14 diagnosis and 41 symptoms is now enlarged with some of the most frequent chronic diseases in Europe described in Fehr et al. (2017), including in total 22 diseases, conditions, and the initial 2 drugs, which produce the described symptoms as side effects, and 52 symptoms (Fig. 1).

[1] Google showed for these symptoms 705.000 results. WebMD has following information: "There are 98 conditions associated with difficulty concentrating, fatigue, feeling faint and poor concentration." "There are 124 conditions associated with difficulty concentrating, disorientation, dizziness and fatigue." "There are 105 conditions associated with difficulty concentrating, fatigue, forgetfulness and memory problems." http://www.medmaster.net/medsearcherengin.html gives an overview about different medical web sites (accessed on January 2016 and August 2017).



Semantic Matrix

Filter Rows

Raw | Normalized | Weighted

Object Name	Fatigue	Concentration problems
Vitamin B12 Deficiency	1.00	0.90
Hypoglycaemia	0.00	1.00
Diabetes	1.00	0.00
Hepatitis E	1.00	0.00
Iron deficiency	1.00	0.00
Heart disease	1.00	0.00
Hypothyroidism	1.00	0.90
Hyperthyroidism	1.00	0.00
Hashimoto's Thyroiditis	1.00	1.00
Autoimmune disease	1.00	0.00
Niacin deficiency	0.90	0.00
Amantadin Drug for influenza	0.00	1.00
Donepezil HCL Drug for Dementia	0.70	1.00
Dementia	0.00	1.00
Depression	1.00	1.00
Addison's Disease	1.00	0.00
Asthma	0.00	0.00
Chronic Bronchitis	0.80	0.00
Myocardial infarction	1.00	0.00
Angina pectoris	0.00	0.00
Hypertension	0.00	0.00
Athrosis	0.00	0.00
Stroke	0.80	0.00
Renal failure	0.80	0.00

0 of 24 object(s) selected.

Figure 1: Excerpt of the Semantical Matrix

The values in the matrix, as mentioned above, mean the degree of affiliation of the symptoms with respect to the diseases. The possible data exchange between SEN and e.g. Excel through a csv-file as a tabular representation allows doctors a fast verification of the values. More important is the fact that professionals can discuss the values to find a good agreement about the degree of affiliation.

THE SYMPTOM-BASED SYSTEM

In this SEN system a cvf was used for the symptoms, which are characteristic for chiefly one specific disease. For example, “fatigue” on the one hand is a symptom occurring in many different diseases, “craving for salt food” on the other is typical for Addison’s Disease. Therefore each symptom characteristic for just one disease got a cvf enlarged value of 2.0, for example “increased sensitivity on cold”, “aching muscles”, “hair loss”, “unexpected pain”, and “craving for salt food”.

Because of the enlargement of the initial diagnosis system with additional diseases, as developers we had to check if the diagnoses have the typical according symptoms to enable an unambiguous diagnosis. This is especially important because all here included diagnosis have several common symptoms.

SEN starts the learning process of the diseases and the according symptoms as described in previous section; c (learning rate) is 0.1, the activation function is the linear-logarithmic one, and the learning process finishes after 3 iterations.

The patient case described in Klüver (2016 a) is taken again, where only several externally (by family members) *observed* symptoms were given into the system, namely: fatigue (1.0), muscle cramp (1.0), thirsty (1.0), hungry (1.0), lost weight (1.0), head ache (0.7), dry skin (1.0), aching muscles (1.0), aching joints (1.0).

The result is shown in Fig. 2 containing the computed ranking and distances:

Ranking

Vector: Symptoms person 3

+1.04 Diabetes
+0.90 Hepatitis E
+0.77 Addison's Disease
+0.51 Hypothyroidism
+0.35 Donepezil HCL Drug for Dementia
+0.32 Hyperthyroidism
+0.31 Hashimotos Thyroiditis
+0.21 Niacin deficiency
+0.16 Amantadin Drug for influenza
+0.13 Renal failure
+0.11 Athrosis
+0.09 Vitamin B12 Deficiency
+0.09 Iron deficiency
+0.09 Heart disease
+0.09 Autoimmune disease
+0.09 Depression
+0.07 Low blood sugar level Hypoglycaemia
+0.03 Chronic Bronchitis
+0.03 Myocardial infarction
+0.02 Hypertension
+0.02 Stroke
+0.00 Dementia
+0.00 Asthma
+0.00 Angina pectoris

Distance

Vector: Symptoms person 3

+0.47 Diabetes
+0.47 Hepatitis E
+0.83 Addison's Disease
+1.29 Hyperthyroidism
+1.41 Niacin deficiency
+1.46 Hypothyroidism
+1.47 Renal failure
+1.49 Hashimotos Thyroiditis
+1.51 Vitamin B12 Deficiency
+1.52 Autoimmune disease
+1.53 Heart disease
+1.54 Amantadin Drug for influenza
+1.56 Iron deficiency
+1.60 Athrosis
+1.61 Depression
+1.67 Low blood sugar level Hypoglycaemia
+1.69 Stroke
+1.71 Myocardial infarction
+1.71 Chronic Bronchitis
+1.75 Hypertension
+1.80 Angina pectoris
+1.81 Asthma
+1.88 Donepezil HCL Drug for Dementia
+2.15 Dementia

Figure 2: The proposed diagnosis by SEN

The result in Fig. 2 shows the bar charts and computed values for the two measured differences. On the left side the ranking corresponds to the highest activation values; on the right side the Euclidean distance is shown for the difference between the input vector of the user and each disease. It is important to note that the proposed diagnosis do not differ after enlarging the system.

Both computed results show that the first three diagnoses are identical. One sees in addition that the two computing methods shown on the right and left side are very suitable to check the validity of the system, i.e. to check the unambiguity of the results. A satisfactory diagnosis certainly means that it is no artifact, due to specific algorithms. To be sure, the comparison of the results of two different methods is not a decisive validity proof but a strong indicator that the result is more than a chance one. In this case the first two diagnoses are not only equal, but also near in the values; the third proposed one has a less activation value (ranking) and a grater value in the distance.

The final diagnosis by the responsible doctors in the hospital showed that SEN was insofar right, as *both* Diabetes and Hepatitis E were the responsible diseases. Even in this extremely difficult case SEN was able to propose a sound solution because the similarity of the computational results hints at just this possibility, namely the simultaneous incidence of more than one disease.

This example shows that in cases where the SEN system proposes more than one disease with nearly equal values users and the responsible doctors should take into regard the possibility that more than one disease might be the cause for the different symptoms.

Component-Wise Analysis Of The Results

The obtained results allow no insight about the differences in the trained and inserted vectors. One can ask why e.g. Diabetes and Hepatitis E are possible candidates in respect to the inserted symptoms by a user, since several of symptoms belonging to the both diseases are very different.
An additional tool was implemented in SEN, which enables the component-wise analysis of the inserted input vector and the trained network. In this case the colored visualization refers to the differences between the components of the vectors; only the symptoms inserted by a user are analyzed in comparison to all diseases. (Fig 2):

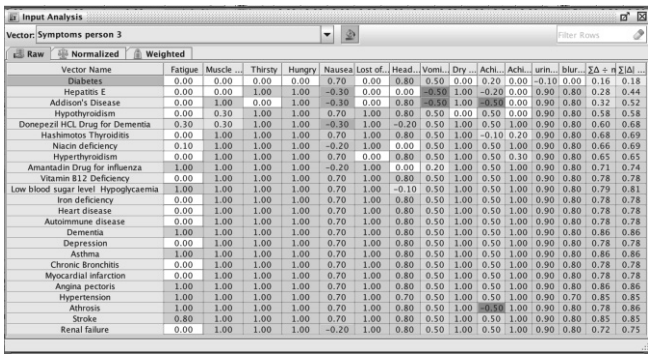


Figure 3: Result of the input analysis with SEN

If there exist no difference, the distance is of course 0 and the values are represented without labeling; the green color indicates that the input vector has a positive value for this attribute, the semantical matrix contains an according smaller value or 0 (represented by color intensity). The red color is used if there is an existing value in the semantical matrix, but a lower or no value in the input vector. In the last two columns the deviations are computed: the penultimate column displays the activation value of the according object vector; the last column consists of the absolute deviance between input vector and objects. In both cases the smallest difference is to the objects Diabetes (D) and Hepatitis E (H).
In Table 1 the analysis is shown consisting only of the decisive differences in the attributes:

Table 1: Differences in the attributes

Disease: / Symptoms	D	H	**User**
Thirsty	**1.0**	0.0	**1.0**
Hungry	**1.0**	0.0	**1.0**
Nausea	**0.0**	**1.0**	**0.7**
Head ache	0.0	**0.8**	**0.8**
Aching joints	0.0	0.0	**0.2**
Urinating more as usual	**1.0**	0.0	**0.9**
Blurring vision	**0.8**	0.0	**0.8**

The described new procedure in SEN allows a developer an automatic analysis of similarities and differences, the finding of unique or decisive attributes to increase the cvf, and at least the possibility to inform the user about missing symptoms to find an unambiguous diagnosis, and about the symptoms, which are important for a user, but not typical for the proposed diagnosis (Klüver, 2016 a).

„Center-Modus“ Visualisation For Users

For a layman it is frequently not very useful to show these different results because such tables and visualizations are on a first sight not easily or intuitively interpretable. That is why our research group additionally constructed the visualization option “center modus” (Klüver and Klüver 2013). In this modus the user sees a plane divided in concentric circles. The different diseases are placed by the algorithm at the periphery, the input of the user, i.e. the vector of his symptoms, is placed in the center. When starting SEN the objects are “drawn” to the center. In the end the user sees the diseases ordered according to their geometrical distances to the center. The nearer the symbol for a disease is to the center, i.e. to the user’s input, the more probable the disease is the cause for the symptoms. Fig. 4 shows the results of the diagnosis in the center modus:

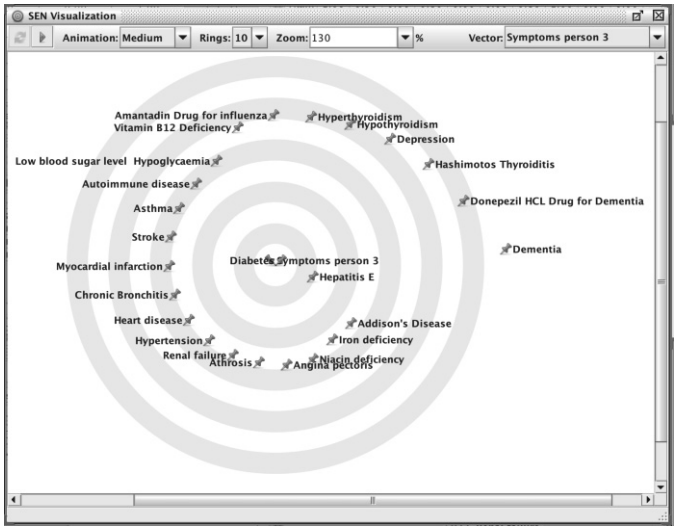


Figure 4: The center-modus visualisation of SEN

The result shows a user that Diabetes indeed is most probable, and also Hepatitis E. All the other diseases have a grater distance to the center and accordingly are less probable. This visualization has the advantage that a user also is informed about the diseases included in such a system.

Map-Visualisation For Object Selection

This visualization clusters the objects according to their approximated similarities. It enables e.g. to see, which diseases have commonalities in their symptoms (Fig. 5). In addition the diseases are placed near to a user (marked with a blue color in Fig. 5), which are more probable due to the inserted symptoms.
Further analysis is also possible through the selection of objects or clusters. Taking into account that diagnosed persons are informed about possible complications and/or are asking “Dr. Google”, “self-diagnosis” systems can be developed as “self-control” as proposed in the next section. The map visualization helps to select the disease assigned to a person, as shown in Fig. 5:

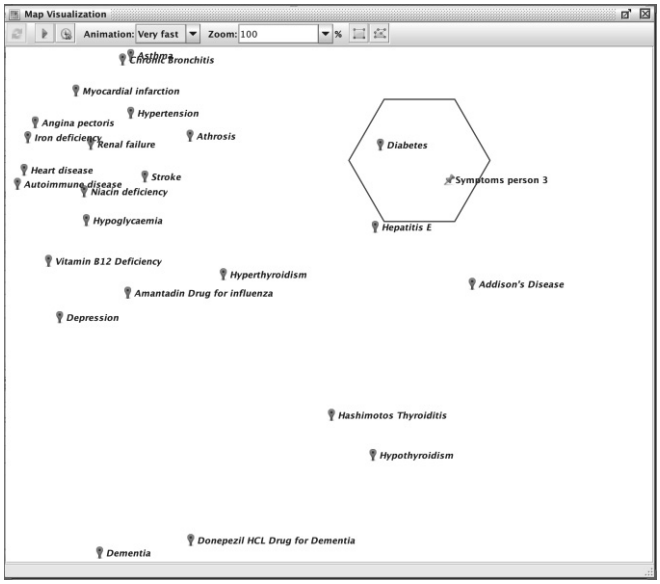


Figure 5: Selection of Diabetes for the development of a SEN-system

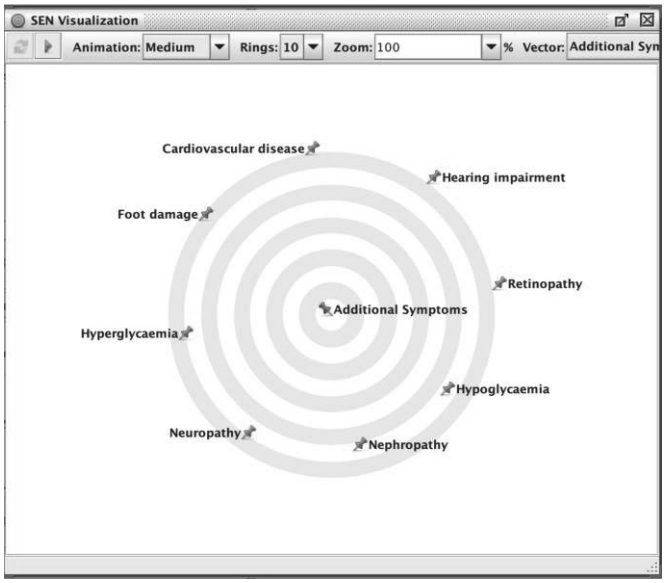


Figure 6: Result of SEN: Hypoglycaemia and Nephropathy are slight attracted to the center.

In this case the symptoms (attributes) are part of the next SEN-system, considering that they have an influence on the possible complications.

DEVELOPING OF A DIAGNOSIS-BASED SYSTEM

The preceding section described a diagnosis system based on symptoms, inserted by the users. It is possible to change this procedure by basing the system on a given diagnosis and showing the most probable additional complications.

The diagnosis of "Diabetes" signifies a confrontation of patients and family members with severe possible complications over years (Alanazi et al., 2017; Brock et al., 2016) as *Cardiovascular* disease, *nerve* damage (Neuropathy), *kidney* damage (Nephropathy), *eye* damage (Retinopathy), Hypoglycaemia, Hyperglycaemia, *foot* damage, *skin* conditions, and *hearing impairment*, to name only the main consequences. Tkachenko et al. (2017) emphasize that own researches with e.g. Google can have a positive effect on the "surveillance of Type 2 diabetes".

The problem is again that some symptoms are very general, as e.g. headache or tingling in feet and arms. With a diagnosis like Diabetes these symptoms can already have a meaning for the persons directly concerned. The usage of SEN could have the advantage that patients and family members can control (observed) symptoms if there is a tendency to a mentioned complication.

For the development of such a "Diabetes"-SEN-system 44 symptoms and laboratory outcomes are defined as attributes and 8 diseases or possible complications as objects.

According to a study, Diabetic kidney disease (DKD) affects approximately 20–40 % of individuals who have diabetes (Molitch and Hahr, 2017). The percentage is high enough to produce fears in the involved persons.

To check the system, at the beginning some general symptoms for Nephropathy are inserted into SEN as input vector: fatigue (1.0), concentration problems (1.0), sweating (1.0), nausea (0.1), urinating more than usual (1.0), tingling in feet or hands (1.0), itching (0.5) swelling of hands (1.0), pale skin (0.2), and headache (1.0). The result is shown in Fig. 6:

According to the symptoms, e.g. headache, pale, concentration problems, Hypoglycaemia is beside Nephropathy also attracted, but all complications seem not to be acute.

Adding additional symptoms or laboratory results into the system as increased nausea (0.5), protein in the urine (0.5), increasing of tingling und itching (1.0), loss of appetite (1.0) and an increased pale skin (0.6), the result shows that the complications of Nephropathy should be considered seriously, as shown in Fig. 7:

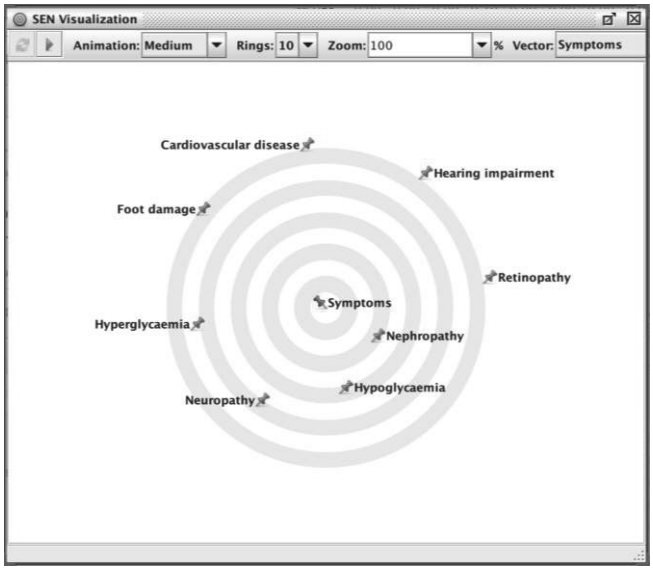


Figure 7: Nephropathy is near to the center

The result shows that not only the problem of Nephropathy (kidney damage) becomes more serious, but also that Hypoglycaemia and Neuropathy are also more attracted to the center and the symptoms should be carefully observed.

Again a SEN-system can inform a user, which symptoms are decisive for the changes, and give additional information about necessary medical examinations.

CONCLUSION AND FURTHER WORK

These prototypes already show the potential of a SEN for self-diagnosis systems. The network is easily to enlarge with symptoms and diseases and the semantical matrix allows transparency in respect to the used data. The two different computation methods and the different activation functions allow a control if the database must be corrected, or if the user must be asked for additional symptoms. The usage is quite simple and easy to learn because the user just has to insert his symptoms together with his assessment of the intensity of his symptoms. As in the SEN developed as orientation system for students (Klüver et al. 2015), the user will for this task get a list of possible symptoms, which he can activate.

An automatic extension of the database will be developed with additional algorithms to evaluate the system with a large database, and, most important, to develop an interactive user interface for asking about symptoms and giving advices for further analysis, for example blood tests. A connection to platforms for doctors to exchange experience or to decide, which values should be chosen for the semantical matrix, is also possible.

Finally our research group plans to develop several SEN systems for different medical disciplines. These systems will be able to "communicate", i.e. for example that one system looks to other systems if they contain the symptoms given by a user and which diseases are connected with them. Such "hybrid systems" should allow the automatic but controlled enlargement of the systems.

REFERENCES

Aiken, M.; G. Kirwa; M. Berry; and C.A. O'Boyle. 2012 "The Age of Cyberchondria." *Royal College of Surgeons in Ireland Student Medical Journal*, 5, 71-74.

Alanazi, W.A; M.B. Uddin; S. Fakhruddin; and K.E. Jackson. 2017. "Recurrent insulin-induced hypoglycemia induces AngII and COX2 leading to renal (pro)renin receptor expression and oxidative stress." *International Journal of Medicine*, 5(1), 71-78.

Amato, F.; A. López; E.M. Peña-Méndez; P. Vaňhara; A. Hampl; and J. Havel, 2013. "Artificial neural networks in medical diagnosis." *Journal of Applied Biomedicine*, Vol. 11, 47-58.

Avery, N.; J. Ghandi; and J. Keating. 2012. "The 'Dr Google' phenomenon—missed appendicitis." *NZMJ*, Vol. 125 Nr. 1367, 135-137

Benigeri M. and P. Pluye. 2003. "Shortcomings of health information on the internet." *Health promotion international*, 18(4), 381-386.

Biel, M.; B. Hammer and T. Villmann. 2016. "Prototype-based models in machine learning". Wiley *Interdisciplinary Reviews: Cognitive Science*, Vol. 7, Nr. 2, 92-111.

Brock, C.; B. Brock; A. Grave Pedersen; A. Mohr Drewes; N. Jessen; and A. D. Farmer. 2016. "Assessment of the cardiovascular and gastrointestinal autonomic complications of diabetes." *World J Diabetes*, 7(16), 321-332.

Dubowicz, A. and P.J. Schulz. 2015. "Medical Information on the Internet: A Tool for Measuring Consumer Perception of Quality Aspects", Interact J Med Res, 4(1):e8, DOI: 10.2196/ijmr.3144.

Fehr, A.; C. Lange; J. Fuchs; H. Neuhauser; and R. Schmitz. 2017. "Health monitoring and health indicators in Europe." *Journal of Health Monitoring*, Vol 2(1), 3-20.

Freeman, L., (Ed.), 1989 *Research Methods in Social Network Analysis*. Fairfax: George Mason University Press.

Griggs, J. 2015. "Doctor you", *New Scientist*, Vol. 227, Issue 3029, 38-41.

Karnam, S. and P. Raghavendra, 2017. „Hybrid Doctors: The Need Risen From Informed Patients." *Journal of Clinical and Diagnostic Research*. Vol-11(2): ZI01-ZI04.

Kim, K.B; H.J. Park; and D.H Song. 2016. "Intelligent home disease pre-diagnosis system for Korean traditional medicine using neural networks." *Int. J. Information and Communication Technology*, Vol. 8, No. 1, 1-9.

Klüver C. and J. Klüver. 2013. "Self-organized Learning by Self-Enforcing Networks". In: *Proceedings of the 12th international work-conference on artificial neural networks (IWANN 2012),* I. Rojas; G. Joya; and J. Cabestany, (Eds.) Part I Lecture Notes in Computer Science, 7902, Springer, 518-529.

Klüver J. and C. Klüver. 2007. *Social Understanding, On Hermeneutics, Geometrical Models, and Artificial Intelligence*, Dordrecht (NL): Springer.

Klüver, C. 2016. "Steering Clustering of Medical Data in a Self-Enforcing Network (SEN) with a Cue Validity Factor." *IEEE Symposium Series on Computational Intelligence*, Athens, 2016, 1-8 DOI: 10.1109/SSCI.2016.7849883.

Klüver, C. 2016a. "Self-Enforcing Neworks (SEN) for the development of (medical) diagnosis systems," *International Joint Conference on Neural Networks (IJCNN). Proceedings of the IEEE World Congress on Computational Intelligence (IEEE WCCI 2016)*, Vancouver, 503-510. DOI: 10.1109/IJCNN.2016. 7727241.

Klüver, C. 2017. "A Self-Enforcing Network as a Tool for Clustering and Analyzing Complex Data. ICCS, Zürich. Procedia Computer Science (2017) Vol. 108, 2496-2500. DOI: 10.1016/j.procs.2017.05.169.

Klüver, C.; J. Klüver; B. Zurmaar. 2015. „OSWI: a consulting system for pupils and prospective students on the basis of neural networks." *Journal of AI & Society*. Vol. 30 Issue 1 23-30, Springer. DOI: 10.1007/s00146-014-0542-y.

Liu, C.; H. Sun; N. Du; S. Tan; H. Fei; W. Fan; T. Yang; H. Wu; Y. Li; and C. Zhang. 2016. "Augmented LSTM Framework to Construct Medical Self-diagnosis Android." *IEEE 16th International Conderence on Data Mining*, 251-260.

Lupton D. and A. Jutel. 2015. "'It's like having a physician in your pocket!' A critical analysis of self-diagnosis smartphone apps." *Social Science & Medicine*, Elsevier, 133, 128-135.

Molitch, M.; A. Hahr. 2017. „Management of Diabetes in the Pre-End-Stage Renal Disease and Chronic Kidney Disease." In: *Managing Diabetic Nephropathies in Clinical Practice.* G. L. Bakris, A. Hahr, R. Khardori, D. Koya, M. Molitch, F. C. Prischl, G. Schernthaner, B. Thajudeen (Eds), Springer International Publishing Switzerland Online 2017, 57-75.

Mueller, J.; Jay, C.; Harper, S.; Davies, A.; Vega, J.; and Chris Todd. 2017. "Web Use for Symptom Appraisal of Physical Health Conditions: A Systematic Review." *Journal of Medical Internet Research* 19(6):e202). DOI: 10.2196/ jmir.6755.

Peker, M. 2016. "A decision support system to improve medical diagnosis using a combination of k–medoids clustering based attribute weighting and SVM." *J Med Syst* 40:116.

Pombo, N.; N. Garcia; K. Bousson; and V. Fezizardo. 2015. "Aritificial Neural Learning Based on Big Data Process for eHealth Applications". In: *Artificial Intelligence Technologies and the Evolution of Web 3.* 0.T. Issa, P. Isaías (Eds.). Hershey: Information Science Reference (Inprint of IGI Global), 291-307.

Robertson, N.; M., Polonsky; and L. McQuilken. 2015. "Are my symptoms serious Dr Google? A resource-based topology of value co-destruction in online self-diagnosis." *Australasian Marketing Journal*, 22, 246- 256.

Rodríguez, J. H.; M. J. Rodríguez Conde; and F. J. Cabrero Fraile. 2016. "*Artificial Neural Networks applications in Computer Aided Diagnosis. System design and use as an educational tool*." Fourth International Conference on Technological Ecosystems for Enhancing Multiculturality TEEM'16, ACM, 1201-1208, DOI: http://dx.doi.org/ 10.1145/3012430.3012670.

Rosch, E. 1973. “Natural Categories.” *Cognitive Psychology*, 4, 328-350.

Shen, N.; M.J. Levitan; A. Johnson; J.L. Bender et al. 2015. "Finding a Depression App: A Review and Content Analysis of the Depression App Marketplace", *JMIR MHealth UHealth*, 3(1):e16.

Tkachenko, N.; S.Chotvijit; N. Gupta; E. Bradley; C. Gilks; W. Guo; H. Crosby; E. Shore; M. Thiarai; R. Procter; and S. Jarvis. 2017. “Google Trends can improve surveillance of Type 2 diabetes.” *Scientific Reports*, 7: 4993, DOI:10.1038/s41598-017-05091-9.

Vassis, D.; B. A. Kampouraki, P. Belsis, V. Zafeiris, N. Vassilas, E. Galiotou, N. N. Karanikolas, K. Fragos. 2015. “Using neural networks and SVMs for automatic medical diagnosis: A comprehensive review.” *AIP Conference Proceedings* 1644, 32, DOI: http://dx.doi.org/10.1063/ 1.4907814.

Yao, C.; Y. Qu; B. Jin; L. Guo; C. Li; W. Cui; L. Feng. 2016. “A Convolutional Neural Network Model for Online Medical Guidance.” *Special Section on Big Data Analytics for Smart and Connected Health*, IEEE Access, Vol. 4, 4094-4103.

Zuccon, G.; B. Koopman; and J. Palotti. 2015. "Diagnose This If You Can. On the effectiveness of search engines in finding medical self-diagnosis information". In *Advances in Information Retrieval, 2015* A. Hanbury, G. Kazai and A. Rauber (Eds.). Lecture Notes in Computer Science, Springer International Publishing, 562-567.

BIOGRAPHIES

Christina Klüver (Dr. phil., habil.) studied educational science and computer science. She obtained the Ph.D. in Communication Science and the Venia legendi in Computer Science. She is lecturer in computer science at the University of Duisburg-Essen, Germany, in the faculty of Economics and Business Administration. As a member of the research group COBASC (Computer Based Analysis of Social Complexity) her research interests are in applying models of nature analogous programming techniques like neural nets, cellular automata, and evolutionary algorithms to technical, social, medical, and economical problems.

Jürgen Klüver (Prof. Dr. phil.) studied Philosophy and Mathematics at the Universities of Kiel and Hamburg.
He is head of the research group COBASC (Computer Based Analysis of Social Complexity).
His main field of research is the development of theories on social and cognitive dynamics, based on the construction and investigation of according computer models, and the analysis of formal models suited for the simulation of social, cognitive, and economical processes.

Model of Modular IoT-based Bee-Keeping System

K. Dineva, T. Atanasova

Institute of Information and Communication Technologies, Bulgarian Academy of Sciences

KEYWORDS
Internet of Things (IoT), Agriculture, Bees, Heterogeneous data, Pure energy

ABSTRACT

The most common and popular approaches for solving problems with bees population by tools of Internet of Things (IoT) are analyzed. These solutions can be improved by using new type of hardware infrastructure components and modular architecture that are proposed in the paper. The proposed solution is directed to small bee producers with accent on economical and practical benefits.

INTRODUCTION

After analyzing statistical data concerning the worryingly low levels of bee populations, global environmentalists raise an alarm – bees are disappearing. Their colonies are decreasing and collapsing with a constant trend. If bee populations keep on decreasing with the same pace, environmentalists predict that they will have vanished from the face of the earth by 2035. The disappearance of bees endangers not only the extraction of honey, but also the production of fruits, vegetables, nuts and certain cereals [1].

The main reasons for this grave problem are the widespread use of pesticides, the insufficient bee activity in the nature, and the spread of diseases and parasites. Types of pesticides that would normally be used, unfortunately, can also harm bees. Additionally the pesticides usage can also pollute the honey, making it inappropriate for human consumption. A non-chemical approach for dealing with the problem is needed, and here the use of the Internet of Things (IoT) provides the necessary and long awaited solution [2]. With the help of related processing of heterogeneous data gathered from IoT devices in bee hives, predictive modeling of the behavior of bee families is possible, which would provide timely detection of different diseases and problems in bee colonies.

UPTODATE APPROACHES

A group of students studying Biology, Food Industry and Engineering Systems at Cork University College in Ireland are working on the challenge of creating a unique platform through which they can monitor, collate, and analyze bee colonies activity in an unobtrusive manner and on a large scale. The team integrates gas sensors to examine the impact of carbon dioxide, oxygen, temperature, humidity, chemical pollutants and air dust levels on honey bees. Generated data gives an opportunity to reveal how a number of environmental factors affect the behavior, health and productivity of bees [3].

The solution provided by BuzzBox is another remarkable approach [4]. The collection of audio data is the core of their solution. A system of processing and analysis of the data is created making it possible to determine accurately and in real time the condition of the beehive.

As studies show, measurement of the strength contained in certain ranges of the audio signal spectrum allows successful differentiation between the various hive states.

Gemalto in cooperation with Eltopia, is developing an innovative project to help resolve the crisis in the beekeeping sector using the help of so much needed new technologies in agriculture. Gemalto and Eltopia create a "smart hive frame" connected to the Internet that replaces the traditional frame in the commercial beehive. The "MiteNot" frame has built-in sensors and a heating element embedded. The breeding cycles of mites and the bees is the key for the effectiveness of the solution. By monitoring the temperature of about 32 elements in the hive, MiteNot identifies the exact time and place for applying heat to the hive to interrupt the mite propagation cycle before fertilization occurs and thus sterilizing the eggs to stop pest propagation. This approach is a good example of how the Internet of Things can help save bees [5].

Weather is essential to beekeeping. Weather correct prediction allows beekeepers to take proper care of their hives.

It is possible to take better decisions thanks to the usage of cognitive technologies which allow machines to learn from a wide variety of data sets, such as barometric reading of satellite images, tools for analyzing, summarizing the information and

measuring its impacts. IBM Watson is a global leader in development of such technology, and after the acquisition of The Weather Company (which provides up to 26 billion daily estimates for 2.2 billion locations worldwide), beekeepers are at the doorstep of a new era of weather forecast that changes the business models and everyday life [6].

MATERIALS AND METHODS

Various techniques and research methods are used to develop and enhance predictive modelling of the behaviour of bee families: study of problems associated with agriculture and environment management; processing, structuring and analyzing the information gathered about bee biological processes and ecological modelling; performing activities like observation, summary, discussions, graphic representation and table presentation of processed and summarized aggregated data; developing and practical testing of created hardware architectures that meet the research objectives; designing, development, installation and maintenance of an open source software platform using Cloud Service Provider for web-based simulation and bee hive management, predictive models of bees' biological process.

RESULTS

IoT means that everything related to the system can transmit data and communicate with other objects, devices and/or people. The result is that everything is able to be measured and traced all the time. A model of a self-learning and self-diagnostic system is proposed, that can be powered by clean energy sources. The system architecture is of a modular type (Fig.1).

The system can collect heterogeneous data like barometer levels, humidity, temperature, noise levels, picture, GPS, CO, CO2 and many more by using various sensors. There are external (apiary's perimeter) and internal sensors (integrated into bee hives). The result is a system that is able to make self-diagnostic of the collected data accuracy and the functional activity of the individual components.

The set of sensors forms a unified network relying on and using the Zigbee protocol which was originally designed to meet the needs for communication between IoT devices. Zigbee uses the IEEE 802.15.4 standard with a 2.4 GHz operating frequency. Thus it is compatible and can be used almost anywhere. The disadvantage of this standard is that two Zigbee profiles can interfere with one another while communicating. We propose using BLE (Bluetooth Low Energy) when needed. BLE is a standard that uses the same frequency as ZigBee and has a good data transfer rate of 1 Mb/sec. It connects to the network for only a few milliseconds, making it a suitable alternative to ZigBee. It should also be noted the small electricity consumption - the device can operate with a single battery charge for several years.

Sensor data are collected in an intermediate module that, on the basis of pre-defined set rules, transmits the required information to a module that has the capability to store it in an intermediate database and process it according to the specified parameters.

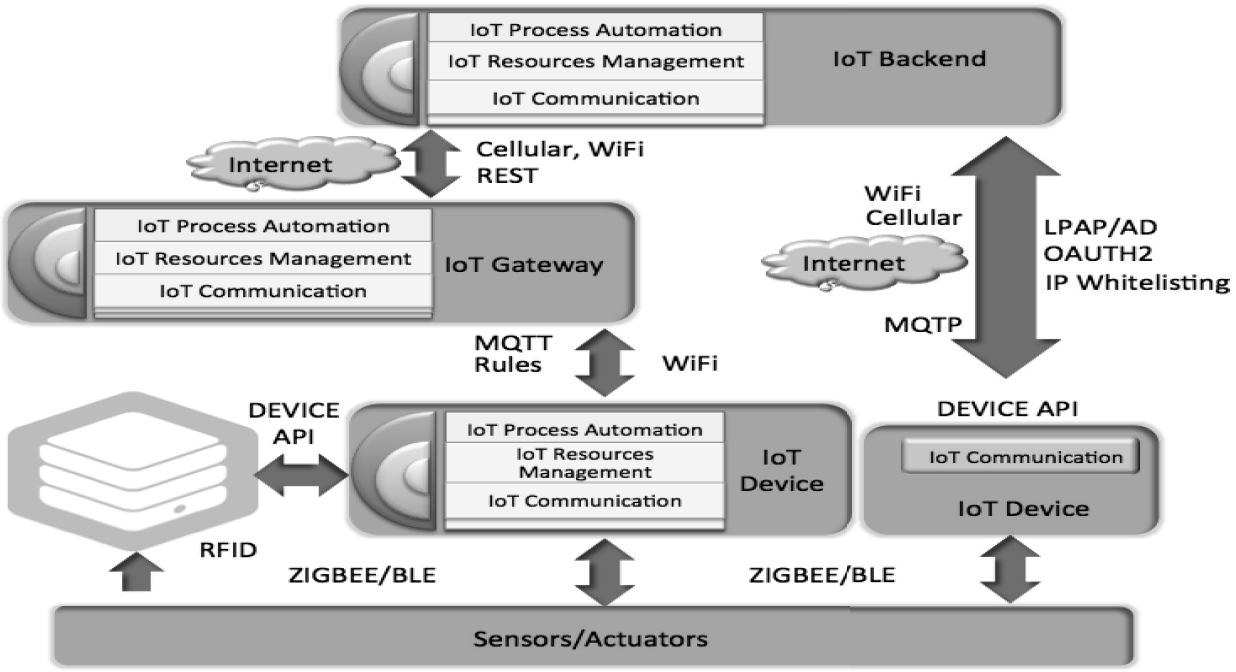


Fig. 1. The System Design

The Fog computing paradigm is used for distributed data transmission and timely processing. The information is then transmitted to a back-end module hosted in cloud service using MQTT protocol based on client-server architecture and TCP. It is primary designed by IBM for being used by distributed systems that have limited computational power and limited power supply. Devices on such network can easily send messages to others who are subscribed to receive them. Several devices can subscribe to receive messages from more than one device. This is controlled by set of rules implemented in the system module when working with information collected by various sensors.

An additional module is developed that communicates directly with a back-end. The communication is bidirectional and the purpose is to maintain control and validate the received data whenever such verification is needed. For example, if devices report temperatures of 35 degrees and control devices report a temperature of 25 degrees, this is an indication of a possible unit malfunction and the system can notify the user immediately for taking proper actions. Also, based on the information from the module, back-end logic can decide whether to collect or not data of a certain type and after that to apply a new rule by communicating with other modules.

The security of the communication between the modules and the back-end is accomplished by using several standards such as LDAP/AD, Oauth2 and IP Whitelisting. Their usage may vary on the specific needs.

The integration of IoT into a real environment has a relatively high price for the moment. This is one of the main reasons why the scalability has not been achieved so far at the desired rate. The current projects are developed mostly for commercial purposes, because along with the possible benefits of using them, there still remains the condition that they should be cost-effective. Cost-effectiveness in short period of time is the way such solutions can become interesting and accessible to beekeepers from amateurs to professionals.

There are a number of approaches and using them will lead to significantly reducing the cost of such an IoT system. Generally, most of these approaches also solve problems, from which typically IoT systems suffer [7], [8].

CONCLUSIONS

The short- and long-term benefits of monitoring bee colonies are related to the greater possibilities of using IoT technologies - allowing better decision making by providing real-time data access to connected devices integrated in bee hives.

In the case of research collaboration in a large international format, the vast amount of data collected from bee hives would be much more accurate and analyses. It will be much more thorough and useful for agricultural and scientific communities on a regional and global scale. This kind of research management could help scientists and researchers to corelate the health of bee colonies with external data such as weather conditions, farming patterns, usage of agricultural chemicals, and other to prevent and slow down the decrease of bee colonies and improve the bee genome.

The overall advantages over traditional approaches are the fact that industries are getting interconnected - data will not only be available and useful in a particular industry. It will be used in various businesses and industries enpowering further innovations. When this type of technology is applied to other similar industries, the impact will be enormous and in help in improving modern agricultural technologies. The ability to connect and share data has the potential to bring together different actions - such as growing crops and beekeeping - to make decisions that everyone can agree on. Opportunities are endless when machines, industries and people are connected and the results are an inspiration for further improvements.

REFERENCES

1. The global and European situation with bees and other pollinators, 2014, http://sos-bees.org/situation/
2. Ellis J. D., J. Klopchin, E. Buss, F. M. Fishel, W. H. Kern, C. Mannion, E. McAvoy, L. S. Osborne, M. Rogers, M. Sanford, H. Smith, P. Stansly, L. Stelinski, and S. Webb, Minimizing Honey Bee Exposure to Pesticides, ENY-162, April 2017, http://edis.ifas.ufl.edu/in1027
3. Reading Beehives: Smart Sensor Technology Monitors Bee Health and Global Pollination, 2015, http://www.libelium.com/temperature-humidity-and-gases-monitoring-in-beehives/
4. Theory Behind BuzzBox Audio Analysis, 2017, https://docs.opensourcebeehives.com/docs/theory-behind-audio-analysis
5. Gemalto's IoT Technology helps Solve Honey Bee Crisis, 2017, http://www.gemalto.com
6. Bee-2-b, 2016, https://paidpost.nytimes.com/ibm/bee-2-b.html
7. Landau D. M., Tiny Technology Helps Save the Honey Bees, *SCIENCE*, August 25, 2015.
8. Peter C., Saving Bees With The Internet Of Things, Forbes, JUL 7, 2016.

THE EFFECT OF SEXUAL NETWORKS ON FERTILITY LEVELS

Edinah Mudimu
College of Economic and Management Sciences
Department of Decision Sciences
University of South Africa, Muckleneuk Campus
P O Box 392, Pretoria, South Africa, 0003.
E-mail: mudime@unisa.ac.za

KEYWORDS

Agent-based, Fertility, Behavioural science, Parenthood

ABSTRACT

The majority of children are born within a couple relationship. Partnership formation plays a major role in the transition to parenthood. This paper presents an agent-based model where mate-search forms the foundation of the transition to parenthood. Women who give birth in the model are married or are in a sexual relationship with an opposite-sex agent. The mate-search rules determine the number of women that give birth. We analyse how changes in mate-search parameters impact on population growth. Results show that the probability that sexual relationships would be initiated, the likeability threshold and random search parameters cause a change in fertility levels observed in the model.

INTRODUCTION

There are a number factors that contribute to demographic changes observed in different countries. One of the factors is the structure of the social and sexual network in a given geographical location (Bernardi (2003)). Little is known about the extent to which social and sexual networking contribute to demographic changes. A lack of data on how social and sexual networks evolve makes it difficulty to model the interaction between network structures and demographic changes accurately. However, Kohler (2001) notes that differences and changes in social structures cause fluctuations in fertility levels in communities.

To explain this phenomenon, Kohler (2001) integrates social interaction theory with economic fertility models and observes changes in reproductive behaviour in a "local environment". Reproductive decision-making in the local environment is assumed to be affected by interpersonal information flow, customs, norms, social externalities and public policies within the social and sexual network. Mathews and Sear (2013) also note that the composition of a social network in the local environment influences reproductive behaviour.

More research on how types of social networks affect reproductive behaviour in human populations was carried out by Lois (2016). To understand the impact of different social networks on the reproductive behaviour, Lois (2016) identifies and analyses four social environments, namely the 'family-remote' network, the 'polarized' network, the 'disintegrated' network and 'family-centered' network, which all differ from one another in terms of their composition. According to Lois's (2016) definitions, the family-remote and polarized social networks are composed mainly of friends and acquaintances. Low emotional contagion, less effective social learning, and contrasting values and lifestyles are the major characteristics of such networks. The major differences between the two are that the polarized social network contains mainly relatives and is relatively larger than the family-remote social network.

Lois (2016) defines the disintegrated network type as a social network composed of individuals who are poorly integrated. In this social network there is little to no positive or negative network influence from others. The fourth social network identified by Lois (2016) is the family-centered network. The family-centered network is characterised by close long-term ties of family members with high opportunities for emotional contagion and social learning. The disintegrated network and the family-centered network have structural differences, but transition to parenthood has been found to be significantly higher in both networks than in the family-remote and polarized social networks. The presence of limited negative network influence in the disintegrated social network is believed to contribute to a higher transition to parenthood; whereas in the family-centered network, transition to parenthood is attributed to the presence of social pressure regarding family formation and strong childcare network support (Lois (2016)).

This article introduces an agent-based model which aims to explain the effect of social networking and couple relationship characteristics on fertility levels in a community. We use an agent-based model for social and sexual partnership formation developed by Mudimu and Engelbrecht (2015) as our base model. In the next section we present a brief literature review of some studies which have also used agent-based modelling.

REVIEW OF PREVIOUS WORK

Some empirical studies, Lois (2016) and Bernardi (2003), have identified social interaction factors that influence fertility. These studies confirm that social interaction, the nature and composition of the social network influence transition to parenthood and the number of children in a family. Computational and mathematical models, which use a top-down modelling approach, have been developed to demonstrate how social interaction contributes to child-bearing in communities (see for example Peristera and Kostaki (2007)). However, computational and mathematical models usually fail to fully explain the heterogeneity observed in social networks (Peristera and Kostaki (2007)) because of non-linearity in human behaviour.

Social and sexual network formation depends on non-linear human behaviour. To fully understand the process there is a need to develop individual-based models that closely replicate human behaviour in a given society. One modelling tool that has been found to model the emergence of macro-level patterns from non-linear micro-level population interactions closely is agent-based modelling (Billari et al. (2003)). Researchers like Mudimu and Engelbrecht (2015), Knittel et al. (2011) and Simao and Todd (2002) have shown how micro-interactions can lead to the development of social and sexual networks that replicate real-world settings. Singh et al. (2016), Yang (2016) and Diaz et al. (2011) have added mechanisms underlying fertility diffusion in social network structures in an attempt to understand how fertility is affected by social networks.

Diaz et al. (2011) consider a single-sex model where child-bearing depends only on the female agent and her social network composed of her mother, female friends and siblings. The selection of friends is based on education level and age. The model simulates different life stages of the agents. Agents leave the model through age-specific mortality rates. The child-bearing age is regarded as between 15 and 49 years. Agents older than the child-bearing age remain in the model as they influence fertile agents' decision-making about giving birth. In the model, fertility behaviour is influenced by three factors, namely education, age and parity. Data from Austria is used to calibrate the model. Results from the model show that social interactions can explain the shift in fertility but they excluded the influence that male partners can have on child-bearing.

Singh et al. (2016) and Sajjad and Ahn (2014) use data from Korea to calibrate their models which aim to explain differences in fertility levels. Only married females give birth in their models since cohabitation is not allowed in South Korea. Both studies used a reproductive age range between 16 and 45 years. Sajjad and Ahn (2014) consider a single-sex model with no social networking, using education as the only factor that influences fertility levels, and allow twin births; while Singh et al. (2016) have developed a population-data-fed two-sex agent-based model. To develop the marriage network, Singh et al. (2016) use income and age of agents. Agents can marry if their age difference is less than or equal to five years, with a maximumum difference of one in education levels. Three education levels are considered in the model (0 – high school; 1 – college; and 2 – university). Level of education is used to assign income to agents. In Korea, the government give a child incentive to couples with children. Therefore the decision to have children is based on a couple's surplus income and the government incentive. Results from the model developed by Sajjad and Ahn (2014) show that there is a higher probability of transition to motherhood for higher educated women than for lower educated women. The model developed by Singh et al. (2016) is work in progress, hence results from this study have not yet been available on the date of publication of their article. However, their model has improved knowledge in this research area by including male agents in child-bearing decisions.

In this paper we introduce a model which tries to replicate child-bearing in South Africa. The marriage rules and reasons why people decide to have a child differ from those in the papers cited here. In South Africa, cohabitation is common (Moore and Govender (2013)). Hence the rules that govern child-bearing are different from the rules uncovered by Singh et al. (2016). In the next section we present a summary of social and sexual partnership formation as referred to in this paper. For a detailed description of partnership formation, see Mudimu and Engelbrecht (2015).

SOCIAL AND SEXUAL NETWORK FORMATION

In the research article by Mudimu and Engelbrecht (2015), three agent-interaction networks are considered. These three networks are hierarchical in structure, with the social (friendship) network forming the foundation, followed by sexual relationship network ties and finally by marriage network ties. Agents are added to the population and exit through death according to mortality estimates adopted from Hontelez et al. (2013). Each agent has static and dynamic attributes. Some of the static attributes are gender, desire for sexual variety and maximum number of friends, date and sexual partners.

An agent's dynamic friendship network starts forming when the agent turns 15. Agents can have friends of any age, but the chances of connecting with a friend with an absolute age difference of less than five years is higher. The dynamic friendship network evolves based on three actions that happen at each time step. The three rules are: random friendship connection (a proportion af agents are randomly selected to make a random

friendship connection); friend of a friend connection (a proportion of agents are selected to form a friend of a friend connection); and the removal of a friendship link (a set of agents are selected to remove one friendship link). A friendship link can be removed if the two agents are not involved in a romantic relationship.

The likeability index is used to select potential candidates an agent can date from the friendship network. The likeability index is calculated using the age difference index, attractiveness index and aspiration level index. The age difference index is calculated based on the fact that male agents prefer a female partner of the same age or younger, while female agents prefer a male partner of the same age or older. The likeability index is the difference between one and the sum of the three indices (age difference index, attractiveness index and aspiration index). If the calculated likeability index is greater than the agent likeability threshold, the opposite-sex friend is added to the agent's list of potential partners. There is a 0.05 chance of randomly picking a potential partner outside an agent's friendship network. The likeability index of randomly selected partners is calculated in the same way.

At each time step, the asking agent (in this case the male agent) sends a date message to one of the potential partners. If the asking agent or the receiving agent is in a romantic relationship, a decision to send or receive a message must be made based on the agent's sexual desire value, the attractiveness of the potential date compared to the current partner and the agent's maximum number of dating partners. Each dating couple formed in the model goes through a courtship time period. Agents can engage in sexual activities during the courtship time period if they are older than their assigned sexually active age. Once a courting couple exceeds courtship duration, the couple may decide to marry. A marriage probability dependent on age and current marital status is used to make the marriage decision, since in the real world not all couples that exceed courtship duration will marry. The availability of the bride price (lobola) and the age or the marital status of the partners are some of the many factors that may affect the decision to marry.

Married agents do not exit the mate search pool, but have stricter rules on sending and receiving date proposals. Divorce is allowed in the model if a more attractive partner is encountered by a married agent. The attractiveness of the potential partner is weighted against the duration of the current marriage. Divorced agents re-enter the mate search pool as single agents. For a detailed description and a list of agent parameters of the mate search algorithm see Mudimu and Engelbrecht (2015). The following section explains how agents are added into the model through birth. The child-birth procedure and model results presented in this article form part of a thesis submitted by Mudimu (2016) to the University of South Africa.

CHILD-BIRTH PROCEDURE

In this paper we present a descriptive model that aims to resemble the structures and processes of child-birth in human societies. We use facts and rules to model the behaviour of the agents, hence the model can be classified as a declarative model. AnyLogic, a Java-based agent-based modelling platform, is used to implement the model. We utilise data available in literature to estimate the parameters used in the model. Where data is not available, plausible assumptions are used.

An agent in our model is either a female or a male individual. Dynamic and static attributes for agents in the model are represented by continuous or discrete state variables. Static variables remain fixed throughout the lifetime of an agent. Static attributes include gender, attractiveness level and aspiration level. Dynamic attributes, which include age, marital status and number of sexual partners, change in response to simulated events and ageing.

Agents are added to the social sexual network model through child-birth. Childbirth is dependent on the social and sexual partnering of agents in the model. In South Africa, cohabitation is common (Moore and Govender (2013)), hence in our model child-bearing is possible for a female agent in the child-bearing age group if she is involved in a sexual relationship or she is married. The child-bearing age group in this model ranges from 15 to 49 years. This is the age range used to calculate the general fertility rate (NAPHSIS (2012)).

An upper fertility-age limit is assigned to each female agent at creation. Female fertility declines with age, typically after age 30, and the actual age at which a female becomes infertile varies (Balasch (2010)). To accommodate this variation in our model, the upper fertility-age limit for each agent is sampled from a normal distribution with a mean of 39 years and a standard deviation of five years (Alam (2008)), with maximum and minimum cut-offs of 35 and 49 years respectively (normal(50,25,35,49)). To reduce the complexity of our model, we assume that during the fertile period, a female agent's fertility is uniform. Table 1 contains parameters used in the model for the childbirth procedure.

The time between the beginning of a sexual relationship and the first pregnancy (first birth interval) differs from one women to the next. Factors that contribute to the time variation to first pregnancy include type of sexual relationship, availability and knowledge of family planning methods, personal life planning, age at the time of initiating a sexual relationship, societal norms and at times biological elements (Amin and Bajracharya (2011)). According to a study carried out by Lofstedt et al. (2005) involving Chinese women aged between 15 and 64, the first birth interval is a minimum of 11 months and can be as long as 30 months or more.

Table 1: Child birth procedure parameters

Parameter	Default value	Description	Source
Fertility	normal(50,25,35,49)	Fertility upper limit for each female agent	Balasch (2010)
FirstPregProb	0.01	Probability of falling pregnant for the first time	Assumption
BirthPregProb	0.15	Probability of falling pregnant after birth	Assumption
Postpartum	six weeks	The first six weeks after birth	Catalyst (2002)
WaitingPeriod	normal(26,4,6,52)	Waiting time after birth before deciding to fall pregnant again	Assumption
PregDuration	normal(40,1,34,42)	Pregnancy duration	Kieler et al. (1995)

The study has established that cohorts that marry early seem to have longer first birth intervals than those that marry late. Similar results have been obtained from a study carried out by Amin and Bajracharya (2011). They studied the variation in first birth intervals for over 60 developing countries and found that the first birth interval varies mostly between 12 and 63 months, with extreme cases where pregnancy occurs immediately after marriage.

In our model we assume that there is a one percent chance to fall pregnant for the first time at each time step to accommodate the variation in the time to first pregnancy. In the South African culture it is not acceptable to fall pregnant before marriage. However, statistics show that females do fall pregnant and give birth before marriage in South Africa (Moore and Govender (2013)). We therefore assume that female agents who initiate a sexual relationship have a 0.2 chance of being put in the subset of agents who can fall pregnant during the course of a sexual relationship.

Another variable that is important when studying fertility is the time between births (birth interval), which differs from one female to the next. In South Africa, the median birth interval is between 27 and 33 months (Moultrie et al. (2012)). Different women prefer shorter or longer birth intervals (12 months to more than six years). The availability of contraceptives has been found to be one of the major factors that impacts on the length of birth intervals. According to the CATALYST Consortium (2002), the recommended optimal birth spacing is between 36 and 60 months. Most women will not fall pregnant within the first six weeks after childbirth. This period is called the "postpartum period". The six-week period gives enough time for the mother and the baby to undergo psychosocial adaptation to the new situation. In our model we therefore use a normal distribution truncated at six weeks (minimum) and 52 weeks (maximum), with a mean of 26 weeks and a standard deviation of four weeks, to model birth interval time. After the waiting period, there is a 0,15 chance of falling pregnant at each time step (weekly). Pregnancy duration usually varies between 38 and 40 weeks (Kieler et al. (1995)). In our model we use a pregnancy duration with a mean of 40 weeks and a standard deviation of one week, truncated at 34 and 42 weeks. A new agent is added to the model once the pregnancy duration lapses. The new agent is assigned static and dynamic attributes that define the agent upon birth.

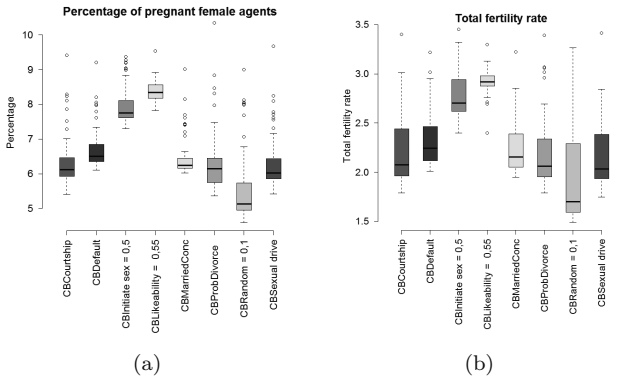
In our model, the decision to stop having children solely depends on the female agent. This decision is made immediately after birth if the female agent is still below her upper fertile-age limit. Data collected by Statistics South Africa (StatsSA (2003)) shows that approximately 40% of women in the 45 to 49 year age group had two or three children. About 0,9% of females in the same age group had 10 or more children. We have tried different probability values and selected the one that closely resembles the distribution of children as shown by the data collected by StatsSA (2003). A probability of 0,025 after each birth is used in this model. Results obtained for this birth procedure built into the social and sexual network are presented in the following section.

MODEL RESULTS

We have varied seven parameters in the social and sexual network model and analysed the effect they have on fertility rate. The seven parameters varied are: likeability threshold; courtship duration; probability that a sexual relationship would be initiated; divorce criteria; concurrency probability for married agents; sexual drive levels; and random partner search. Out of the seven parameters, decreasing likeability threshold, increasing the probability that a sexual relationship would be initiated and increasing random partner search resulted in significant changes in fertility levels. The other four parameters did not have an effect on fertility levels in the model.

Increasing the probability that a sexual relationship would be initiated and decreasing the likeability threshold respectively resulted in an increase in the percentage of pregnant females in the model. This consequently increased the fertility level in the model and the percentage of females waiting between childbirths. Another feature evident in the model analysis is a significant increase in the percentage of first marriages when the probability that a sexual relationship would be initiated is increased and the likeability threshold is decreased respectively.

Increasing the probability for random partner search resulted in a decrease in the percentage of never-coupled agents. However, this did not have the same effect on the percentage of pregnant women (Figure 1(a)); instead a decrease is observed in the percentage of pregnant females. An increase in the probability of random partner search increases the mixing chances of agents in the model, which in turn increases the chance of meeting more attractive partners. Hence, relationships formed when the probability for random partner search is increased may not last long enough to result in pregnancy.



(a) (b)

Figure 1: **Percentage of pregnant females**

An asymmetrical distribution that tails off to the right (positively skewed) is obtained for pregnancy waiting times using probabilities for falling pregnant as described in section entitled "Child-birth procedure". From our model results, pregnancy waiting time between births is on average 81 weeks, with a maximum of 532 weeks and a minimum of 15 weeks. Pregnancy waiting time from birth has a fixed minimum of six weeks set for all women regardless of the type of romantic relationship. Waiting time for female agents who are married at model initialisation, who get married during the simulation run and who are in a sexual relationship is on average 85 weeks, with a maximum of 617 weeks and a minimum of 0,5 weeks.

The distribution for the number of children born to a female before the end of her reproductive lifespan is positively skewed. The results show that, approximately 13% of female agents who have children in their lifetime have eight or more children. Most of the women (87%) have fewer than eight children. Our model result for women with more than eight children is higher by approximately 5,4% compared to the data obtained in literature ((StatsSA 2003, p. 49)). StatsSA (2003) notes that the data provided should be treated with caution since births might be under-reported. Our model counts the number of births for a female agent. We do not track whether the child survives or not. This may explain why we have a higher percentage of females with eight or more children than reported by StatsSA (2003).

DISCUSSIONS AND CONCLUSIONS

This article presents a model to study the dynamics of social and sexual networks and how they contribute to fertility levels in a population. The article expands the work presented by Mudimu and Engelbrecht (2015), which explains how agents select their sexual partners. We extend the model by allowing female agents to have children. We use data from South Africa to initialise our simulations. Agents in the model have properties such as age, attractiveness, aspiration level, gender, maximum number of dating and sexual partners.

Our model results show that fertility rates range between 1.9 and 2.9, with an average of 2.2. These values are within the range of the fertility levels observed in South Africa (StatsSA (2016)). We can conclude that our sexual network managed to capture important aspects of the social and sexual networks observed in South Africa. However, there is still a need to improve the way in which the social and sexual network as well as birth are modelled. There is also a need to investigate why a decrease in never-coupled agents does not automatically lead to an increase in pregnancy and a significant increase in agents in a sexual relationship when the random partner search parameter is increased. This result can be attributed to the quick break-up of couples who mate through random search or a delay in initiating sexual activities. We leave this as an area for further research.

For the future work, we intend to improve the way in which the child-bearing decision is made in our model. The decision to have children in our model depends on probabilities and only the female agent is considered. In real-life settings, having children, especially in a family set up, depends on both parties (husband and wife) as well as customs, norms, social externalities, family ties and public policies (Mathews and Sear (2013)). To improve the way in which we model birth we need to include the male agent and some of the social externalities and public policies in child-bearing decision-making. There is also a need to model the impact of contraception method, including the use of condoms explicitly. Our model does not take this into consideration and we leave this as an area that requires further investigation.

REFERENCES

Alam, S. J.: 2008, *Understanding Social Complexity in the Context of HIV/AIDS: A Case Study in Rural South Africa*, PhD thesis, Manchester Metropolitan University Business School.

Amin, S. and Bajracharya, A.: 2011, Marriage and First Birth Intervals in Early and Late Marrying Societies: An Exploration of Determinants, *2011 Annual Meetings of the Population Association of America Wash-*

ington, DC .

Balasch, J.: 2010, Ageing and infertility: an overview, *Gynecological Endocrinology* pp. 1–6.

Bernardi, L.: 2003, Channels of social influences on reproduction., *Population Research and Policy Review* **22**, 527–55.

Billari, F. C., Ongaro, F. and Prskawetz, A.: 2003, *Agent-based Computational Demography: Using Simulation to Improve Our Understanding of Demographic Behaviour.*, Physica-Verlag HD, Heidelberg, chapter Introducton: Agent-based Computational Demography, pp. 1–17.

CATALYST Consortium: 2002, Optimal Birth Spacing: New Research from Latin America on the Association of Birth Intervals and Perinatal, Mate and Adolescent Health, *CATALYST Consortium* .

Diaz, B. A., Frent, T., Prskawetz, A. and Bernardi, L.: 2011, Transition to parenthood: The role of social interaction and endogenous networks., *Demography* **48**, 559–579.

Hontelez, J. A. C., Lurie, M. N., Barnighausen, T., Bakker, R., Baltussen, R., Tanser, F., Hallett, T. B., Newell, M.-L. and de Vlas, S. J.: 2013, Elimination of HIV in South Africa through Expanded Access to Antiretroviral Therapy: A Model Comparison Study, *PLOS Medicine* **10**(10), 1–44.

Kieler, H., Axelsson, O., Nilsson, S. and Waldenstrom, U.: 1995, The length of human pregnancy as calculated by ultrasonographic measurement of the fetal biparietal diameter., *Ultrasound in Obstetrics and Gynecology* .

Knittel, A. K., Riolo, R. L. and Snow, R. C.: 2011, Development and evaluation of an agent-based model of sexual partnership, *Adaptive Behavior* **19**(6), 425–450.

Kohler, H. P.: 2000, Social interaction and fluctuations in birth rates, *Population Studies* **54**, 223–37.

Kohler, H. P.: 2001, *Fertility and Social Interaction: An Economic Perspective*, Oxford: Oxford University Press.

Lofstedt, P., Ghilagaber, G., Shusheng, L. and Johansson, A.: 2005, Changes in marriage age and first birth interval in Huaning Country, Yunnan province, PR China, *The Southeast Asian Journal of Tropical Medicine and Public Health* .

Lois, D.: 2016, Types of social networks and the transition to parenthood., *Demographic Research* **34**(23), 657–688.

Mathews, P. and Sear, R.: 2013, Does the kin orientation of a British woman's social network influence her entry into motherhood?, *Demographic Research* **28**, 313–340.

Moore, E. and Govender, R.: 2013, Marriage and Cohabitation in South Africa: An Enriching Explanation., *Journal of Comparative Family Studies* **44**(5), 623–639.

Moultrie, T. A., Sayi, T. S. and Timaeus, I. M.: 2012, Birth intervals, postponement and fertility decline in Africa: A new type of transition, *A Journal of Demography* .

Mudimu, E.: 2016, *On modelling thetransmissin of the human immunodeficiency virus (HIV) in a closed mixed society*, PhD thesis, Economic and Management Sciences, University of South Africa.

Mudimu, E. and Engelbrecht, G. N.: 2015, Agent-based model for social and sexual partnerships formation., *Adaptive Behavior* **23**((l)), 34–49.

NAPHSIS: 2012, Statistical Measures and Definitions. **URL:** *https://naphsis-web.sharepoint.com*

Peristera, P. and Kostaki, A.: 2007, Modeling fertility in modern populations, *Demographic Research* **16**(6), 141–194.

Sajjad, M. and Ahn, C. W.: 2014, Agent-based Model to analyse the Role of Women's Education on Fertility: The case of Korea, *Advanced Science and Technology Letters* **58**, 29–37.

Simao, J. and Todd, P. M.: 2002, Modeling mate choice in monogamous mating systems with courtship, **10(X)**, 1–24.

Singh, K., Sajjad, M., Paik, E. and Ahn, C.-W.: 2016, Simulating Demography - Dynamics of Fertility Using a Multi Agent Model, *2016 18th International Conference on Advanced Communication Technology (ICACT), Pyeongchang* pp. 787–791.

StatsSA: 2003, *South African Demographic and Health Survey SADHS 2003*, Statistics South Africa.

StatsSA: 2016, *Mid-year population estimates*.

Timaeus, I. M. and Moultrie, T. A.: 2008, On postponement and birth intervals, *Population and Development Review* .

Yang, Z.: 2016, An agent-based dynamic model of politics, fertility and economic development., *Proceedings of the 20th World Multi-Conference on Systemics, Cybernetics and Informatics (WMSCI 2016)* .

GRAPHICAL HUMAN BIO-ANALYSIS

Semi-automatic Brain Lesions Detection and Segmentation Method in MR Images

Carlos Segura Granados[1]
Volodymyr Ponomaryov[1]
Martha Hernandez-Cuellar[1]
[1]Instituto Politécnico Nacional, Santa Ana 1000, Col. San Fco. Culhuacan, 04430,
Mexico-city, Mexico
E-mails: fi.ing.carlossg@gmail.com, volodymyr.ponomaryov@gmail.com, mghcuellar@hotmail.com

KEYWORDS
Brain lesions, classification, MR images, feature extraction, k-means segmentation.

ABSTRACT

Magnetic Resonance Imaging (MRI) is one of the most widely used imaging studies for medical diagnosis because it provides important information about the body tissues and organs in a non-invasive way. Brain lesions detection - segmentation is a complex procedure because in some lesions the similarity with healthy tissue can be very similar. This study presents the development of a semi-automatic method for the detection and segmentation of brain lesions in MRI images. The training stage consists of feature extraction and classification into two class. During next test stage, we should identify an injury or it absent, following; the lesion is segmented showing the region of interest (ROI) for the specialist usage.

INTRODUCTION

An acquired brain injury usually causes a change in the neuronal activity, which can affect the areas of verbal communication, memory, attention and concentration that seem as the most important. Edema and necrosis are examples of more common brain lesions. The first one can be defined like as the increased water inside of the brain tissue, which causes an increased intracranial pressure (Jha 2003). On other hand, necrosis is defined as the morphological changes, which occur into the brain cells after their death (Syntichaki and Tavernarakis 2003). The goal of designing a CADe computer-aided detection system is to help the specialist to detect any possible cranial injury. This task is considered a difficult and slow. Because the practitioner have to analyze more than 150 images per study in a single patient.

RELATED WORKS

Nooshin Nabizadeh (Nabizadeh and Kubat 2015) studied an automatic tumor segmentation in the single-spectral MRI.
Khalid Usman (Usman and Rajpoot 2017) described a method called brain tumor classification from multi-modality MRI.
Ramakrishnan (Ramakrishnan and Sankaragomathi 2016) employed an estimation on the MRI brain tumor images.

PROPOSED METHOD

A flowchart of the proposed framework shown in next figure consists of several steps explaining below.

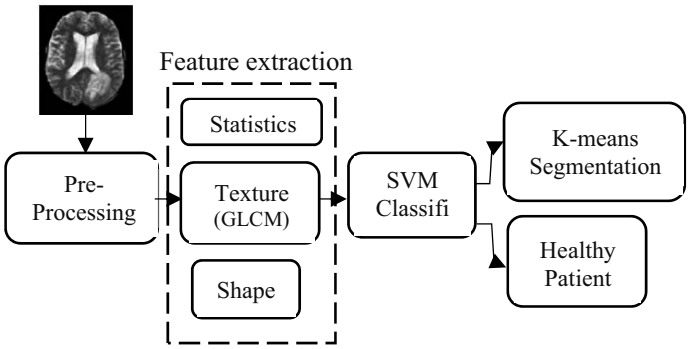


Figure 1. Scheme of the proposed method

Pre-processing

In this step, we use a median filter with 3x3-window size to reduce the noise of the MR images.

Feature Extraction

During next step, the most important features were extracted.
Statistical features. Haralick (Haralick et. Al.1973) has proposed some statistical characteristics. In this study, several important from them were used: *Mean, variance, kurtosis, max value.*

The shape of the region is important characteristics to determine whether there may be an injury into MR images. For the shape characteristics, the following ones were calculated: *Area, diameter, perimeter and symmetry*, defined in (Nixon and Aguado 2002).
The *texture* considers the spatial relationship of the pixels and can be characterized by the Gray Level Co-occurrence Matrix (GLCM).
Obtaining the normalized GLCM matrix, the next measures proposed by Haralic were extracted (Haralick et al. 1973):
Contrast, entropy, variance, auto-correlation, cluster shade, cluster prominence and sum average.

Classifier
In this study we have employed the *SVM Classifier*. A Support Vector Machine (Brereton and Lloyd 2010) estimates a function that should classify the input data into two different classes.

Segmentation method
Segmentation is the key part of the designed method because with it help we can obtain the highest accuracy and precision for a better specialist diagnosis.
The next block diagram exposes the proposed framework.

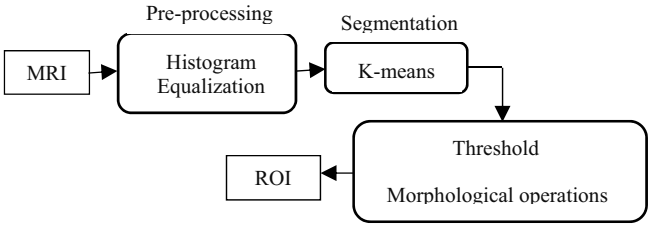


Figure 2. Block diagram of the segmentation stage method.

Histogram Equalization
This technique redistributes gray levels in order to obtain a uniform histogram. This allows to the areas with low contrast obtaining better contrast.

K-Means segmentation
The K-Means method (Tan et al. 2005) is an unsupervised learning algorithm that solves the clustering problem. This algorithm works by grouping each an object to an image of k different groups. The iterative stage of the algorithm begins, in which the centroid of each a group is recalculated by minimizing the least squares function as shown below:

$$J = \sum_{j=1}^{k} \sum_{i=1}^{n} \left\| x_i^{(j)} - C_j \right\|^2 \qquad (1)$$

where $\| x_i^{(j)} - c_j \|^2$ is the distance from any point $x_i^{(j)}$ to the centroid c_j; $c_j = (c_1, c_2, .., c_n)$ is the prototype vector of k clusters, according to $c_j = \frac{1}{|A_i|} \sum_{x_i \in A_i} x_i$

Extracting ROI
In the ROI extraction, the background of the image and the tissue surrounding the lesion should be removed with some tools as threshold and morphological operations. This particular ROI we call mask and it is multiplied by the original image.
Threshold
Getting the histogram can be proposed an optimal threshold to separate the brain tissue from the lesion. For the development of this stage, it is required to separate the lighter parts of the image (Zhou et al. 2010).

Morphological operations
Dilation (Zhou et al. 2010) consists of the expansion of pixels of a certain object. The filling of regions or voids is based on introducing white pixels into the region or object.

EXPERIMENTAL RESULTS AND DISCUSSION

The performance of developed method using the *BraTS 2015* database (The multimodal Brain Tumor Image Segmentation Benchmark) was evaluated.
Using studies of the patients with different types of Lesions (edema, necrosis and brain tumors), each an image has a size of 556 x 512 pixels. The database also provides the ground truth (GT) for each a study. In addition, it has the two types of MRI (T1 and T2). T2 is the type of image, which in this work was used.

Classifier Performance
For the evaluation of the classifier performance, the commonly used criteria are used as follows:

$$Sensitivity = \frac{TP}{TP + FN} \qquad (2)$$

$$Specificity = \frac{TN}{TN + FP} \qquad (3)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad (4)$$

Where:
TP is True Positives, TN is True Negatives, FN is false Negative, and FP is False Positives.

The classifier results are shown in table 1.

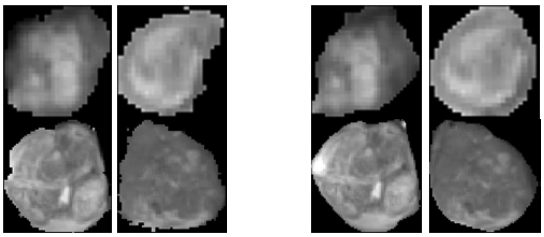
Table 1. Comparison of results with KNN classifier.

	Image type	Lesion type	Sens.	Spec.	Accu.
Nabizadeh	MR	Tumor	89.6	91.8	90.3
Ramakrishnan	CT	Tumor	84	99.3	91
Usman Khalid	MR	Tumor Edema Necrosis	88	87	94
Proposed method (SVM)	MR	Edema Necrosis	81	100	91

The classifier has been trained with 60 patients, 30 healthy patients and 30 patients with lesion and during the test experiments 22 news patients were used, (11of these were healthy patients and 11 patients had lesions). In addition, table 1 exposes the performance criteria and the results obtained by other existing proposals to compare them against with our work.

The classifier has been trained with 60 patients, 30 healthy patients and 30 patients with lesion and during the test experiments 22 news patients were used, (11of these were healthy patients and 11 patients had lesions). In addition, table 1 exposes the performance criteria and the results

obtained by other existing proposals to compare them against with our work.
As an example of the segmentation stage, the final image obtained by proposed algorithm is presented (figure 3).



a) b)

Figure 3. a) GT image, b) ROI obtained by our method.

Other metrics are used, such as the *Jaccard* index and the similarity coefficient. Where the index of similarity varies between zero and two, if the value is two it means total similarity

$$Ji = \left|\frac{Ax \cap Bx}{Ax \cup Bx}\right| * 100 \quad (5)$$

$$S = \frac{2(Ax \cap Bx)}{Ax \cup Bx} \quad (6)$$

The segmentation results were presented and compare them with similar obtained by other authors (see Table 2).

Table 2. Comparison of results for the segmentation stage.

	Proposed Method	MLT	EM
Precision %	98	-	100
Recall %	88	-	66.6
Accuracy %	**87**	-	83.54
S	**1.72**	1.48	-
Ji	**85.95**	74.01	-

CONCLUSIONS

The proposed framework appears to demonstrate good approximation of ROI that results in better performances in comparison with other similar methods presented by (MLT and EM). The similarity value is of 1.72, which equals 86% similarity between regions.
In addition, in the classifier results our work obtained good results with the SVM classifier; however, these results are below in comparison with (Nabizadeh) and (Usman). In spite of these results, we obtained the 100 percent value in specificity, do not detecting any false positive error. Future work is planned in application of more powerful classifiers: AdaBoos, expecting to increase the performance in the classification task In addition, improving the segmentation task we can raise the classification result. We will develop the principal components analysis (PCA) to get the same information but with less data, which it will help the classifier in training and test task.

Acknowledgements

Authors would like to thank Instituto Politecnico Nacional (México) and Consejo Nacional de Ciencia y Tecnología (México) (grant 220347) for their supports in this work.

REFERENCES

BraTS 2015. Kistler et. al, The virtual skeleton database: an open access repository for biomedical research and collaboration. JMIR, 2013. Available in 2016 at https://www.smir.ch/BRATS/Start2015

Brereton RG, Lloyd GR (2010) Support Vector Machines for classification and regression. Analyst 135:230–267. doi: 10.1039/B918972F

Haralick R, Shanmugan K, Dinstein I (1973) Textural features for image classification. IEEE Trans. Syst. Man Cybern. 3:610–621.

Jha SK (2003) Cerebral edema and its management. Med J Armed Forces India 59:326–331. doi: 10.1016/S0377-1237(03)80147-8

Nabizadeh N, Kubat M (2015) Brain tumors detection and segmentation in MR images : Gabor wavelet vs . statistical features q. 45:286–301. doi: 10.1016/j.compeleceng.2015.02.007

Nixon MS, Aguado AS (First Edition 2002) Feature Extraction and Image Processing. Vol. 2

Ramakrishnan T, Sankaragomathi B (2016) A professional estimate on the computed tomography brain tumor images using SVM-SMO for classification and MRG-GWO for segmentation. Pattern Recognit Lett. doi: 10.1016/j.patrec.2017.03.026

Syntichaki P, Tavernarakis N (2003) The biochemistry of neuronal necrosis: rogue biology? Nat Rev Neurosci 4:672–684. doi: 10.1038/nrn1174

Tan P-N, Steinbach M, Kumar V (2005) Chap 8 : Cluster Analysis: Basic Concepts and Algorithms. Introd to Data Min Chapter 8. doi: 10.1016/0022-4405(81)90007-8

Usman K, Rajpoot K (2017) Brain tumor classification from multi-modality MRI using wavelets and machine learning. Pattern Anal Appl 1–11. doi: 10.1007/s10044-017-0597-8

Zhou H, Wu J, Zhang J (2010) Digital Image Processing: Part II.

BIOGRAPHIES

CARLOS SEGURA GRANADOS graduated in 2016 by the National Autonomous University of Mexico in Mechatronics Engineering. He is currently studying the Master's Degree in Engineering Sciences in Microelectronics at the Higher School of Mechanical and Electrical Engineering (ESIME) at Culhuacan Unit of Instituto Politecnico Nacional (IPN) de Mexico.

VOLODYMYR PONOMARYOV received the Ph.D. in 1974 and D.Sci. degrees in 1981. His research interests include signal/image/video processing, real-time filtering, etc. He has published more than 500 international journal and conference scientific papers, and 23 patents of ex USSR, Russia and Mexico five scientific books in international editorials.

MARTHA HERNANDEZ-CUELLAR graduated from IPN (1993). Her research interests includes planning and evaluation of project . From 2000, she is lecturing in ESIME –Culhuacan (IPN).

TRANSCRIPTION INITIATION CONTROLS SKEWNESS OF THE DISTRIBUTION OF INTERVALS BETWEEN RNA PRODUCTIONS

Vinodh K. Kandavalli, Sofia Startceva, and Andre S. Ribeiro
Laboratory of Biosystem Dynamics, BioMediTech Institute and Faculty of Biomedical Sciences and Engineering, Tampere University of Technology, Finland
E-mail: andre.ribeiro@tut.fi

KEYWORDS
Transcription Initiation, Skewness in RNA production; Stochastic Models; Single-RNA measurements.

ABSTRACT

Most regulation in transcription controls when and with which intensity genes are expressed. However, recent evidence suggests that control is also exerted on the noiseness of this process. Here, we use an empirically validated stochastic multi-step model of transcription to explore how its steps kinetics affect the skewness of the distribution of intervals between consecutive RNA productions in individual cells. From the simulations, we show that skewness is independent of the mean transcription rate, while differing widely with the fraction of time that the RNA polymerase spends in the steps following open complex formation. Next, from qPCR and live, time-lapse, single-RNA microscopy measurements of multiple promoters, we validate our model predictions. Using the validated model, we then show that skewness affects, e.g., the fraction of time that protein numbers are below a threshold. We conclude that skewness in transcription kinetics can be tuned by the rate-limiting steps in initiation and, thus, may be an evolvable decision-making parameter of genetic circuits.

INTRODUCTION

In prokaryotic organisms, such as *Escherichia coli*, gene expression and, in particular, transcription, is the critical process where most regulation of the metabolism and responses to external fluctuations and signals occur (Ramos et al. 2001; López-Maury et al. 2008). It is, thus, not surprising that *E. coli* possesses a pletora of repression and activation molecules, along with other means to silence and activate specific genes (McClure 1985; Lutz et al. 2001). There are also various global regulation mechanisms, such as σ factors (Jishage et al. 1996) and DNA super-coiling (Menzel and Gellert 1983).

Nevertheless, bacterial cell populations exhibit single-cell heterogeneity in gene expression profiles (Leibler and Kussell 2010). This diversity has two sources. One is the stochastic nature of the chemical processes involving gene expression, due to the low number of regulatory molecules involved. The other is differences between cells in their numbers of various components, age, cycle stage, etc. (Elowitz et al. 2002).

This noise was found to affect multiple cellular functions, including stress response, metabolism, cell cycle, circadian rhythms and aging (Raj and van Oudenaarden 2008). Depending on its magnitude and on the function affected, the noise can be either beneficial or harmful. One of the main reasons for this wide range of influence is that it affects dynamics of genetic circuits and their ability to perform critical tasks, such as decision making, time counting, noise filtering, memory, etc. (McAdams and Shapiro 1995; Wolf and Arkin 2003).

Similarly to noise, temporal asymmetries in RNA and protein production are expected to have significant consequences, particularly since such asymmetries may determine whether a certain number of RNA or proteins is reached, allowing the crossing of a threshold 'used' by a genetic circuit for decision making. So far, such temporal asymmetries have not been quantified at the single cell level, but recent measurements of time intervals between consecutive RNA productions in individual cells of various promoters under various conditions suggest that they are not negligible (Tran et al. 2015; Lloyd-Price et al. 2016; Kandavalli et al. 2016; Oliveira et al. 2016; Häkkinen and Ribeiro 2015; Häkkinen and Ribeiro 2016).

Here, we investigate if and by which degree the evolvable rate-limiting steps in transcription initiation can tune asymmetries in the distribution of intervals between productions of RNA molecules in individual cells. For this, we consider a stochastic model of transcription initiation and investigate how the asymmetries in the distribution of intervals between consecutive RNA production events in individual cells differ within the realistic ranges of parameter values. Next, we perform experimental validation of these predictions using qPCR and live, time-lapse, single-molecule RNA microscopy measurements of the transcription kinetics of multiple promoters. Finally, we investigate whether changes in the degree of asymmetry in the distribution of intervals between consecutive RNA production events of a gene can have tangible consequences in the crossing of thresholds of protein numbers over time.

MATERIALS AND METHODS

Bacterial Strains and Plasmids, Media and Cell Growth, Microscopy, and Image Analysis

Microscopy data are from (Kandavalli et al. 2016; Oliveira et al. 2016). Briefly, *E. coli* cells carrying 2 plasmids were used: a low copy reporter plasmid expressing MS2-GFP controlled by the promoter P_{Lac} or P_{Tet}, and a single copy F-based plasmid, expressing the RNA with a 96 MS2-GFP binding site array followed by mRFP1 controlled by $P_{Lac\text{-}ara\text{-}1}$, P_{BAD}, or P_{TetA}. Cultures were grown in LB media overnight at 30 ºC in an orbital shaker with aeration of 250 rpm and

diluted to fresh LB media to initial OD_{600} of 0.05 (measured with Ultraspec 10 cell density meter). Next, they were incubated at 37 ºC at 250 rpm until reaching an OD_{600} of 0.25. To produce MS2-GFP, e.g. when under the control of P_{Lac}, cells are induced with 1 mM IPTG (for P_{Tet} we induce with 100 ng of aTc) and allowed to grow until OD_{600} of 0.5. For the target induction, 0.1% arabinose and 1 mM IPTG for $P_{Lac\text{-}ara\text{-}1(Full)}$, 0.1% arabinose alone for $P_{Lac\text{-}ara\text{-}1(ara)}$, 1 mM IPTG alone for $P_{Lac\text{-}ara\text{-}1(IPTG)}$ and 0.1% arabinose for P_{BAD} is used. For P_{TetA}, no induction is required, as the cells lack the gene coding for the repressor, TetR (Kandavalli et al. 2016).
For microscopy, a few µl of cells with the reporter and target plasmids were sandwiched between a coverslip and an agarose gel pad (2.5%), also containing the inducers. Prior to this, the chamber (FCS2, Bioptechs) was heated to 37 °C and placed under the microscope. Cells were visualized using a Nikon Eclipse (Ti-E, Nikon) inverted microscope, equipped with a 100x Apo TIRF (1.49 NA, oil) objective. Confocal images were obtained by a C2+ (Nikon) confocal laser-scanning system. To visualize fluorescence 'spots', we used a 488 nm laser (Melles-Griot) and an emission filter (HQ514/30, Nikon). Confocal images were taken every 1 min for 2 h, and phase contrast images were obtained every 5 min by an external phase contrast system and CCD camera (DS-Fi2, Nikon), using Nikon Nis-Elements software.
Analysis of the images was performed by the software 'CellAging' in four steps (Häkkinen et al. 2013): (i) cell segmentation from phase-contrast images, (ii) fluorescent RNA intensity detection from the confocal images, (iii) cell lineage construction, and (iv) RNA production estimation from the single-cell RNA intensity time series. We used the software to perform an automated segmentation of phase-contrast images, followed by manual correction. Next, from each segmented cell, at each time point, fluorescent spots are detected automatically. Finally, cell lineages are constructed, by establishing the relationships between cell masks in sequential frames. In these, time-series of fluorescent spots intensity were obtained for each cell. From those, the time points when novel RNA molecules ('spots') appear in each cell were estimated (Häkkinen and Ribeiro 2015). Finally, the time intervals between consecutive RNA productions in individual cells were estimated.

qPCR

In the case of P_{BAD} and P_{TetA}, qPCR data was obtained from (Kandavalli et al. 2016). In the case of $P_{Lac\text{-}ara\text{-}1}$, the data is from measurements performed here. To measure gene expression by qPCR, we grew cells as in (Kandavalli et al. 2016). Total RNA was isolated and quantified. The RNA samples were treated with DNase to remove residual DNA, followed by cDNA synthesis. cDNA samples were mixed with qPCR master mix containing iQ SYBR Green supermix (Biorad), with primers for the target and reference genes. The reaction was carried out in triplicates. For quantifying the target gene, we used the following primers: for mRFP1 (Forward: 5' TACGACGCCGAGGTCAAG 3' and Reverse: 5' TTGTGGGAGGTGATGTCCA 3'), and for the 16S RNA reference gene (Forward: 5' CGTCAGCTCGTGTTGTGAA 3' and Reverse: 5' GGACCGCTGGCAACAAAG 3'). The following conditions were used: 40 cycles of 95°C for 10 s, 52°C for 30 s and 72°C for 30 s for each cDNA replicate. We used no-RT controls and no-template controls to crosscheck non-specific signals and contamination. PCR efficiencies of these reactions were greater than 95%. The data from CFX Manager TM Software was used to calculate the relative gene expression and its standard error (Livak and Schmittgen 2001).

Model of Transcription

We model transcription as a multi-step process, represented in (1), following the empirically validated models in (McClure 1985). The level of detail of our model is based on the one we can reach in the measurements (i.e. intervals between RNA production events with measurements taken every minute).

From these intervals one can dissect the fraction of time of those intervals that is spent prior and after commitment to the open complex formation (Häkkinen and Ribeiro 2015). As such, transcription is modeled by the following multi-step process (McClure 1985):

$$P + R \xrightarrow{k_{cc}} RP_{cc} \xrightarrow{k_{oc}} RP_{oc} \xrightarrow{\infty} P + RNA + R \qquad (1)$$

where $k^*_{cc} = R \cdot k_{cc}$.
The process starts with RNAp, R, binding to a free, active promoter, P, and forming a closed complex, RP_{cc}, at the rate k_{cc}. k^*_{cc} stands for the inverse of the mean time that RP_{cc} remains in equilibrium with P and RNAp, until it starts forming a stable open complex. That is, this model does not explicitly represent the instability of RP_{cc}. As such, the first step is not an elementary chemical process. Rather, its rate represents the inverse of the time until a stable open complex forms, which depends on preceding events, such as binding and unbinding of the RNAp to the promoter (i.e. reversibility), 1D diffusive searches, etc. (Bai et al. 2006).
The second step in (1) represents the open complex formation, RP_{oc}, which is a nearly irreversible step (McClure 1985; Lloyd-Price et al. 2016; Kandavalli et al. 2016) and requires the RNAp to open the DNA double helix (Chamberlin.MJ 1974; McClure 1985). This is followed by promoter escape (after which P is released into the system), elongation and termination (i.e., the release of RNA and R). These latter steps are expected to be much shorter-length than the events in initiation (Herbert et al. 2008) and, thus, are not represented. Further, they would only affect the variance and not the mean duration of the intervals between transcription events.
Regardless of the complexity of the steps, recent studies suggest that, to a degree, the process can be well-modelled by two consecutive, independent exponential steps (Tran et al. 2015; Kandavalli et al. 2016). Thus, the probability density function (pdf) of the distribution of intervals between transcriptions is the convolution of their pdfs:

$$f_{\Delta t}(t) = \frac{k^*_{cc} \cdot k_{oc}}{k_{oc} - k^*_{cc}} \left(e^{-k^*_{cc} \cdot t} - e^{-k_{oc} \cdot t} \right) \qquad (2)$$

We assume also a first-order reaction modelling RNA degradation with a rate k_{d_rna} = 0.0033 s^{-1}, which is the median of the RNA degradation rate in *E. coli* (Bernstein et al. 2002):

$$RNA \xrightarrow{k_{d\_rna}} \varnothing \qquad (3)$$

For simplicity, we model translation as a single-step event which produces an unfolded protein Pro_{un} with a rate of k_{tr} = 0.0637 (Jones et al. 2007):

$$RNA \xrightarrow{k_{tr}} RNA + Pro_{un} \qquad (4)$$

Finally, proteins fold into functional at the rate k_{fold} = 0.0024, and degrade at the rate k_{d_pro} = 0.0017 (Cormack et al. 1996):

$$\mathrm{Pro}_{un} \xrightarrow{k_{fold}} \mathrm{Pro} \quad (5)$$

$$\mathrm{Pro} \xrightarrow{k_{d\_pro}} \varnothing \quad (6)$$

Skewness as a Measure of Asymmetry of the Intervals Distribution

As a measure of asymmetry of the distribution of transcription intervals, we use its skewness, *S*, as in (MacGillivray 1986):

$$S = \frac{m_3}{{m_2}^{3/2}}, \text{ where } m_r = \frac{1}{n}\Sigma(x_i - \bar{x})^r \quad (7)$$

More precisely, we estimate the sample skewness (S_s) of measured and simulated data distributions by applying a correction to increase the estimates precision for samples from asymmetric distributions (8) (Joanes and Gill 1998):

$$S_s = \frac{\sqrt{n(n-1)}}{n-2} \cdot S \quad (8)$$

To estimate the standard uncertainty of S_s, we performed non-parametric bootstrap as in (Carpenter and Bithell 2000). Namely, for each data set, we resampled the data randomly with replacement (using the original amount of samples) 10^5 times, and calculated the bootstrap sample skewness, S_{sb}. As the obtained S_{sb} distributions were well-approximated by a normal distribution, we estimated the standard uncertainty as the 68% percentile confidence interval of the S_{sb} distribution.

τ plots

In vitro and *in vivo* studies have demonstrated that the mean time between transcription events (Δt) can be altered by changing the free RNAp concentration (Shehata and Marr 1971; Lloyd-Price et al. 2016). Also, assuming (1), only the closed complex formation duration changes with the free RNAp concentration (McClure 1985). This change was shown to be linear for a given range of cell growth conditions (Lloyd-Price et al. 2016). As such, it is possible, within this range of conditions, to produce a τ plot by placing the inverse of the relative RNAp concentration in the x-axis, and the inverse of the relative rate of RNA production in the y-axis.

The relative rate of RNA production can be measured by qPCR. The relative RNAp concentration can be measured by Western Blot (Kandavalli et al. 2016). The data points are then fitted with a line. Scalling the production rates to the condition of interest, the intercept of the line with the y-axis equals ($\tau_{oc}/\Delta t$) of this condition, as it represents the media condition with infinite RNAp (and, thus, with infinitely fast closed complex formation).

Stochastic Simulations

To simulate the model (1)-(6), we use SGNS2 (Lloyd-Price et al. 2012), which is driven by the Stochastic Simulation Algorithm (Gillespie 1977), but allows also for multi-time-delayed reactions (Roussel and Zhu 2006). We accounted for individual cell observation times, as these affect the measured intervals (unlike in the theoretical predictions of the pdf of the distributions of intervals between RNA production events (2)). This single-cell observation time-lengths depend on (a) cell doubling time, (b) duration of the measurement, and (c) the degree of overlap between measurement time and cell doubling time. We measured such distribution of single-cell observation windows for each studied condition. Next, we set k_{cc} and k_{oc} according to the $\tau_{oc}/\Delta t$ obtain from qPCR and Western Blot. The absolute values of these rates are then fitted to match the mean of the measured distribution.

Truncated Gaussian Distribution of Intervals Between Consecutive RNA Productions

To obtain Gaussian distributions truncated at zero and with a given mean μ and squared coefficient of variation, CV^2, we use the following procedure. First, we obtained the best fit value of the standard deviation $\hat{\sigma}$ of the Gaussian with mean μ, which minimizes the difference between CV^2 and CV^2_{tr} of the truncated distribution. Next, we calculated a scaling coefficient $\alpha = \mu/\mu_{tr}$. Finally, we truncated at zero a Gaussian distribution with mean of $\alpha\mu$ and standard deviation of $\alpha\hat{\sigma}$, which results in the desired distribution.

RESULTS AND CONCLUSIONS

Skewness is Controlled by $\tau_{oc}/\Delta t$, but is Independent of the Mean Interval Between Transcription Events

We consider the model of transcription in (1). Its RNA production kinetics is determined by k^*_{cc} and k_{oc}. k^*_{cc} defines the inverse of the mean time for the RNAp to find the promoter and complete the closed complex (τ_{cc}). k_{oc} defines the inverse of the mean time for the completion of the open complex formation (τ_{oc}). Thus:

$$\Delta t = \tau_{cc} + \tau_{oc} \quad (9)$$

Given measurements of Δt in live *E. coli* cells of various active promoters (Häkkinen and Ribeiro 2016; Kandavalli et al. 2016; Lloyd-Price et al. 2016; Tran et al. 2015; Oliveira et al. 2016), we assume its realistic range of values to be between 10 and 2500 seconds. To study how changing the kinetics of the two rate limiting steps of the model allows tuning the asymmetry (as measured by *S*) of the Δt distribution, we vary $\tau_{oc}/\Delta t$ (by changing τ_{cc} and τ_{oc}) while maintaining Δt constant.

For each such combination of Δt and $\tau_{oc}/\Delta t$ values, from (7), we calculated *S* of the pdf of the Δt distribution. From this, we find that, first, *S* changes significantly with $\tau_{oc}/\Delta t$, being symmetric around 0.5, where it is minimal. On the other hand, it is independent from the mean value Δt, for any given constant value of $\tau_{oc}/\Delta t$ (Fig. 1).

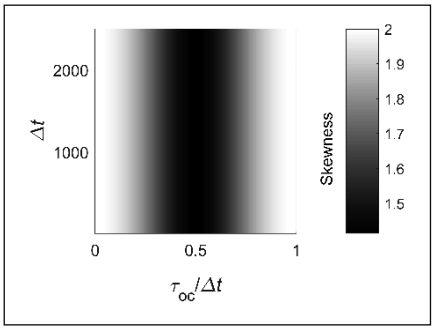


Figure 1: 2D plot of the model-based analytical prediction of the skewness of the distribution of intervals between consecutive RNA production events as a function of Δt and $\tau_{oc}/\Delta t$

Empirical Validation of the Model Predictions

To validate the above, we attained empirical data on the Δt distribution for various promoters and induction schemes (Kandavalli et al. 2016; Oliveira et al. 2016). Also, for each condition, by qPCR (Methods), we measured $\tau_{oc}/\Delta t$. In Fig. 2, we confront the empirical values of *S* as a function of $\tau_{oc}/\Delta t$ with the simulated predictions (Methods). Visibly, the model fits the data for a wide range of possible values of $\tau_{oc}/\Delta t$.

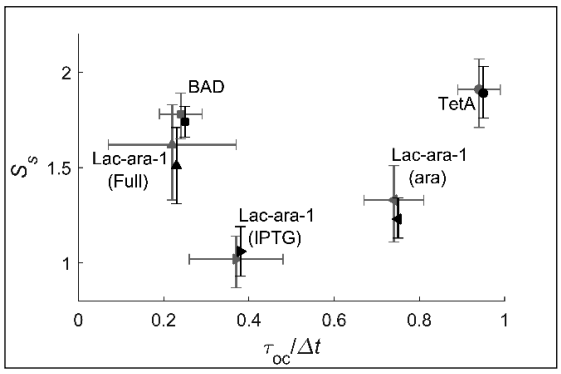


Figure 2: Sample skewness (S_s) as a function of $\tau_{oc}/\Delta t$ for measured data (grey points) and data obtained from simulations of the model (black points) along with standard uncertainties (error bars). In both sets of data, for each condition, 100 or more Δt intervals were extracted from a total of 100 or more cells. The data from simulations is shifted along the x-axis by 0.01, to assist visualization

Skewness in Transcription Initiation Kinetics Affects Threshold Crossing by Protein Numbers

We next explore *in silico* the potential role of *S* in tuning protein numbers over time, in particular, we study protein number threshold-crossing. For that, we quantify, as a function of $\tau_{oc}/\Delta t$, the fraction of time during the course of an *in silico* experiment that the protein numbers equal zero.

In addition to reactions (1)-(6), we accounted for RNA and protein dilution due to cell division as in (Goncalves et al. 2016), assuming mean cell lifetimes of 1h. We modeled 7 conditions, with the same mean RNA and protein numbers but differing in *S*, and measured the fraction of time that the protein numbers equal zero during a time series (each series being 100 hours long). We simulated 1000 time series per condition. In each, data from the first hour was omitted to exclude the transient state from the subsequent data analysis.

In Fig. 3, we present the results per condition, averaged over all simulations. Namely, we show the value set for $\tau_{oc}/\Delta t$ in each condition, along with the resulting *S* and CV^2 (Fig. 3) of the distribution of time intervals between consecutive RNA productions in individual cells. Also shown is the fraction of time that the model cells are absent of proteins produced by the gene of interest. This quantity is used here as a quantifier of the propensity of this gene expression system to cross a lower-bound threshold in protein numbers over time.

From Fig. 3, as $\tau_{cc}/\Delta t$ decreases within realistic parameter values (Fig. 2), causing S_s and CV^2 to change, we find that the fraction of time that proteins are absent from the model cells differs significantly between neighboring conditions.

As the model of transcription does not allow varying *S* and CV^2 independently, these results do not suffice to show that it was the change in *S* that caused the change in the fraction of time that the model cells spent without proteins produced by the gene of interest. To show that the threshold crossing can be affected by *S* alone, we performed an additional set of simulations where Δt follows a Gaussian distribution (truncated at zero) with the same mean and CV^2 as the Δt distribution of condition $\tau_{oc}/\Delta t = 0.5$. Results in Fig. 3 show that due to its lower *S* (0.85), as predicted, the fraction of time that proteins are absent in model cells with 'Gaussian-like' RNA production dynamics differs significantly from the control model cells with RNA production dynamics following (1), including when having the same CV^2.

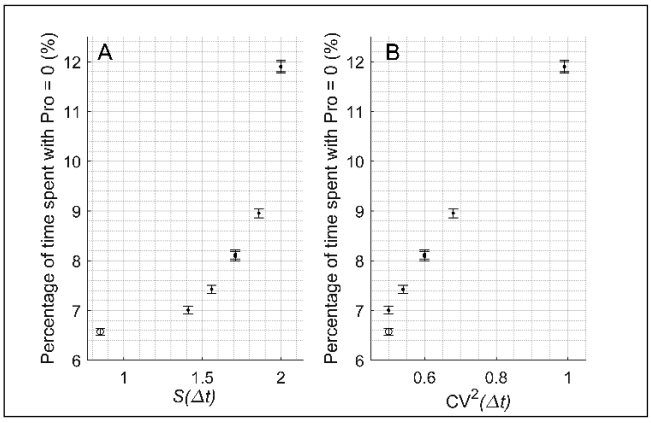


Figure 3: Skewness (*S*) and squared coefficient of variation (CV^2) of the distributions of intervals between consecutive productions of RNA molecules in individual model cells differing in τoc/Δt. Shown is the relative time that cells are absent of the protein of interest as a function of (A) *S* and (B) CV^2 of the time intervals distribution between consecutive RNA productions. Data from stochastic simulations (1000 cells per condition) and from a truncated Gaussian. The error bars are the 90% confidence intervals

We conclude that tuning *S* of the Δt distribution has tangible effects in RNA and protein numbers over time, even if the CV^2 is not or is only weakly affected. Importantly, according to model (1), this tuning can occur by regulating τ_{cc} and τ_{oc}, which are physical properties of the promoter that are both sequence-dependent (McClure 1985) and subject to external regulation, e.g., by transcription factors (Lutz et al. 2001) or global regulatory molecules, such as σ factors.

DISCUSSION

Here, based on a 2-step stochastic model of transcription and empirical data on the time intervals between consecutive RNA productions in individual cells from various promoters and induction schemes, we first made use of the model to investigate how tuning the relative duration of the steps prior and after commitment to the open complex formation allows tuning the skewness of the RNA production kinetics. We determined how this skewness changes as a function of these rate-limiting steps and, most interestingly, that it is minimized for equal duration of the two rate-limiting steps, and made use of the empirical data to validate these predictions. Finally, by tuning the skewness of the transcription initiation kinetics within realistic parameter value intervals we observed modifications in protein numbers dynamics strong enough to likely affect the behaviour of small genetic circuits.

Importantly, we expect S to be tunable via the regulation of τ_{cc} and τ_{oc}, which are sequence dependent and subject to external regulation, e.g., by transcription factors or global regulatory molecules, such as σ factors. Thus, this regulatory mechanism is expected to be both evolvable as well as adaptable to environmental changes.

In the future, we aim to expand our research, first, by studying more complex models of transcription and investigate how each rate-limiting factor influences the degree of skewness in RNA production kinetics. Second, we aim to investigate on how the tuning of skewness of the component genes allows attaining desired macro dynamics in various genetic circuits.

REFERENCES

Bai, L.; T.J. Santangelo; and M.D. Wang. 2006. "Single-Molecule Analysis of RNA Polymerase Transcription". *Annual Review of Biophysics and Biomolecular Structure*, 35, 343–360.

Bernstein, J.A.; A.B. Khodursky; P.H. Lin; S. Lin-Chao; S.N. Cohen. 2002. "Global analysis of mRNA decay and abundance in *Escherichia coli* at single-gene resolution using two-color fluorescent DNA microarrays". *Proc. Natl. Acad. Sci. U.S.A.*, 99(15), 9697–9702.

Carpenter, J. and J. Bithell. 2000. "Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians". *Statistics in Medicine*, 19, 1141–1164.

Chamberlin. M.J. 1974. "The selectivity of preparation". *Psychological Review*, 81(5), 442–464.

Cormack, B.P.; R.H. Valdivia; and S. Falkow. 1996. "FACS-optimized mutants of the green fluorescent protein (GFP)". *Gene* , 173, 33–38.

Elowitz, M.B.; A.J. Levine; E.D. Sigga; P.S. Swain. 2002. "Stochastic Gene Expression in a Single Cell". *Science*, 297, 1183–1186.

Gillespie, D.T. 1977. "Concerning the validity of the stochastic approach to chemical kinetics". *Journal of Statistical Physics*, 16(3), 311–318.

Goncalves, N.; S.M.D. Oliveira; V.K. Kandavalli; J.M. Fonseca; and A.S. Ribeiro. 2016 "Temperature Dependence of Leakiness of Transcription Repression Mechanisms of *Escherichia coli*". *Proc. of the Computational Methods in Systs. Biol.*, Sept. 21–23, Cambridge, U.K., 341–342.

Häkkinen, A. and A.S. Ribeiro. 2016. "Characterizing rate limiting steps in transcription from RNA production times in live cells". *Bioinformatics*, 32(9), 1346–1352.

Häkkinen, A.; A.-B. Muthukrishnan; A. Mora; J.M. Fonseca; and A.S. Ribeiro. 2013. "CellAging: A tool to study segregation and partitioning in division in cell lineages of *Escherichia coli*". *Bioinformatics*, 29, 1708–1709.

Häkkinen, A. and A.S. Ribeiro. 2015. "Estimation of GFP-tagged RNA numbers from temporal fluorescence intensity data". *Bioinformatics*, 31(1), 69–75.

Herbert, K.M.; W.J. Greenleaf; and S.M. Block. 2008. "Single-molecule studies of RNA polymerase: motoring along". *Annual Review of Biochemistry*, 77, 149–176.

Jishage, M.; A. Iwata; S. Ueda; and A. Ishihama. 1996. "Regulation of RNA polymerase sigma subunit synthesis in *Escherichia coli*□: intracellular levels of four species of sigma subunit under various growth conditions". *Journal of Bacteriology*, 178(18), 5447–5451.

Joanes, D.N and C.A Gill. 1998. "Comparing Measures of Sample Skewness and Kurtosis". *Royal Statistical Society,* 47(1), 183–189.

Jones, B.; D. Stekel; J. Rowe; C. Fernando. 2007. "Is there a Liquid State Machine in the Bacterium *Escherichia Coli*?" *Proceedings of the 2007 IEEE Symposium on Artificial Life*, 187–191.

Kandavalli, V.K.; H. Tran; and A.S. Ribeiro. 2016. "Effects of σ factor competition are promoter initiation kinetics dependent". *BBA - Gene Regulatory Mechanisms*, 1859(10), 1281–1288.

Leibler, S. and E. Kussell. 2010. "Individual histories and selection in heterogeneous populations". *Proc. Natl. Acad. Sci. U.S.A.*, 107, 13183–13188.

Livak, K.J. and T.D. Schmittgen. 2001. "Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method". *Methods*, 25, 402–408.

Lloyd-Price, J.; A. Gupta; and A.S. Ribeiro. 2012. "SGNS2: A compartmentalized stochastic chemical kinetics simulator for dynamic cell populations". *Bioinformatics*, 28(22), 3004–3005.

Lloyd-Price, J.; S. Startceva; V. Kandavalli; J.G. Chandraseelan; N. Goncalves; S.M.D. Oliveira; A. Häkkinen; and A.S. Ribeiro. 2016. "Dissecting the stochastic transcription initiation process in live *Escherichia coli*". *DNA Research*, 23(3), 203–214.

López-Maury, L.; S. Marguerat; J. Bähler. 2008. "Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation". *Nature Reviews Genetics,* 9(8), 583–593.

Lutz, R.; T. Lozinski; T. Ellinger; and H. Bujard. 2001. "Dissecting the functional program of *Escherichia coli* promoters: the combined mode of action of Lac repressor and AraC activator". *Nucleic Acids Research*, 29(18), 3873–3881.

MacGillivray, H.L. 1986. "Skewness and Asymmetry: Measures and Orderings". *The Annals of Statistics*, 14(3), 994–1011.

McAdams, H.H. and L. Shapiro. 1995. "Circuit Simulation of Genetic Networks". *Science*, 269(5224), 650–656.

McClure, W.R. 1985. "Mechanism and control of transcription initiation in prokaryotes". *Annual Review of Biochemistry*, 54, 171–204.

Menzel, R. and M. Gellert. 1983. "Regulation of the genes for *E. coli* DNA gyrase: homeostatic control of DNA supercoiling". *Cell*, 34(1), 105–113.

Oliveira, S.M.D.; A. Häkkinen; J. Lloyd-Price; H. Tran; V. Kandavalli; and A.S. Ribeiro. 2016. "Temperature-Dependent Model of Multi-step Transcription Initiation in *Escherichia coli* Based on Live Single-Cell Measurements". *PLoS Computational Biology*, 12(10):e1005174.

Raj, A. and A. van Oudenaarden. 2008. "Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences". *Cell*, 135(2), 216–226.

Ramos, J.L.; M.-T. Gallegos; S. Marques; M.-I. Ramos-Gonzales; M. Espinoza-Urgel; and A. Segura. 2001. "Responses of Gram-negative bacteria to certain environmental stressor" *Current opioin in Microbiology*, 4(2), 166–171.

Roussel, M.R. and R. Zhu. 2006. "Stochastic kinetics description of a simple transcription model". *Bulletin of Mathematical Biology*, 68(7), 1681–1713.

Shehata, T.E. and A.G. Marr. 1971. "Effect of nutrient concentration on the growth of *Escherichia coli*". *Journal of Bacteriology*, 107(1), 210–216.

Tran, H.; S.M.D. Oliveira; N. Goncalves; and A.S. Ribeiro. 2015. "Kinetics of the cellular intake of a gene expression inducer at high concentrations". *Molecular BioSystems*, 11, 2579–2587.

Wolf, D.M. and A.P. Arkin. 2003. "Motifs, modules and games in bacteria". *Current Opinion in Microbiology*, 6(2), 125–134.

SIMULATION IN HUMAN BIO-ANALYSIS

A MODELING APPROACH TO HEART FAILURE TREATMENT

Alexander Lassnig, 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, Integrated Care.

ABSTRACT

Demographic changes, increased life expectancy and associated rise in chronic diseases challenge public health care systems. New solutions based on optimized treatment methods and integrated concepts of care are essential to overcome future financial burdens. In this context models are feasible tools to evaluate potential advantages, estimate prospective implications and support decision makers in the healthcare environment. In this work a heart failure model based on discrete event and agent based methodologies is presented. The hybrid setting includes treatment procedures for outpatient as well as inpatient care on a patient individual level. Based on the conventional care, focusing on economic and health outcomes, alternative concepts of care can be assessed. With the verification steps completed, the model is in the phase of validation based on clinical data. Preliminary results underline the unfavorable prognosis of heart failure. The established model represents a sophisticated tool to evaluate sustainable solutions in heart failure treatment and will serve as a potential basis for decision-making, thus contributing to the formulation of holistic approaches in health care.

INTRODUCTION

Demographic changes through the so-called double aging effect and the associated rise in chronic diseases pose a challenge to public health care systems. New solutions based on optimized and individualized treatment methods and integrated care concepts are potential ways to manage future financial burdens. To convince health care providers to take these alternative approaches into regard, their beneficial outcomes have to be proven, analyzed and discussed in advance. In this context modeling offers sophisticated tools and serves as a basis for future decision-making. The presented model focuses on the treatment of heart failure patients. Based on comprehensive simulations of conventional care, predictions for outcomes of integrated care concepts, such as telemedical treatment or disease management programs, can be estimated.

The decision to model heart failure was based on its significance in the field of cardiovascular diseases and the correlation with the ageing society. Epidemiological data shows that heart failure is the leading cause of hospitalizations of patients older than 65 years (Zannad et al 2009) and necessary treatment expenditures account for 1-2 % of the total health care budget of western countries (Berry et al 2001). In addition to the financial impact, heart failure is associated with an unfavorable prognosis. A high disease-related mortality of 52.3 % within five years of the initial diagnosis of heart failure highlights the severity of the illness (Lloyd-Jones et al 2009). Likewise, a 30-day readmission rate of over 23 % after hospitalization indicates room for improvements in post-inpatient management (Ross et al 2010). Since the majority of expenses for heart failure patients arise due to hospitalizations and because of the clinical significance of rehospitalization rates, modeling inpatient care as precise as possible is crucial. Recent methods of care such as telemedical systems, disease management programs or their combination focus on the detection of symptoms and abnormal parameters to offer treatment at an earlier stage and thus stabilizing the health status and reducing unnecessary visits to the hospital (Koehler et al. 2011). Numerous studies focus on outcomes of those concepts of care; however, literature in this field is controversial in regard to the potential of new approaches. Therefore the following model was created to serve as an applicable tool for decision-making in this field.

METHODS

For simulation, the Java based software AnyLogic® (Version 8.1) was used. The two methodologies discrete event and agent based modeling were combined to describe the clinical pathway of heart failure patients. Both methods align naturally due to the structure of discrete event modeling with its entities passing through the virtual flow chart. Instead of entities, in this hybrid approach agents of the class patient with distinct features were used. Days served as the time basis for the model calculations and simulations were done with random seeds, making each run unique and requiring several iterations in order to attain significant results. Altogether 1000 iterations of the simulation were performed to create significant results. Simulation runs were done via an input mask allowing the user to specify the starting conditions such as number of patients, age, gender and their respective state of health.

Based on the discrete event model by Schroettner et al 2013, the model flow was further developed by introducing a patient collective to inpatient and outpatient care, allowing for more in-depth analyses of individual behavior through the agent based approach. Patients went through the treatment process on a daily basis and faced probabilities for certain events, such as visits to the physician, admissions or a change in medication.

To assess the current state of health New York Heart Association (NYHA) classes were used to differentiate four groups by the severity of heart failure. Each class correlated with different frequencies, lengths and costs of the treatment procedures. In addition age, gender and risk factors (e.g. smoking, obesity, comorbidities) can be taken into account to simulate specific collectives. Improvement or deterioration of the state of health can consequently be modeled by changes of the respective NYHA class.

AnyLogic® offers several tools to graphically present the simulation outcomes. For statistical data processing the information was gathered in a database as well as exported to a .csv file and processed in SPSS®, MATLAB® and R.

Patients

Contrary to the purely discrete event approach by Schroettner et al in 2013, the agent based approach allowed monitoring of the course of treatment in each instance of the class patient. The lower degree of abstraction of the hybrid approach could be realized based on the acquisition of new extensive data sets for outpatient and inpatient care of heart failure patients in Austria. Age, gender, state of health and risk factors could be adjusted and affected the decision tree alongside probability density functions for length of stay, costs, mortalities and class changes. Visits to the physician, medication, admissions, individual medical procedures, etc. were logged in a patient specific history file that also tracked NYHA class changes as well as arising costs.

Outpatient Care

The key elements of outpatient care included the physician, the specialist, the ambulance and the transport of the patient to the hospital. Expenses for the physicians, specialists and the outpatient clinic were reimbursed with a median value per visit and correlated with the severity of heart failure. Medication was represented by the Anatomical Therapeutic Chemical Classification System Codes (ATC) C03, C07, C09 and their subgroups. Costs for the medication were calculated with Weibull functions. Expenses for visits to the physician and for prescriptions were based on data, covering over 11000 heart failure patients, provided through cooperation with an Austrian health insurance provider. Outpatient mortality rates were derived, but lacked clear indication of the cause of death. To regard the state of health NYHA classes could change after visits to outpatient care providers.

Inpatient Care

Admissions, as well as stays at intensive care units with additional individual medical procedures, were implemented based on a scientific cooperation with Austrian health care providers. A cohort of over 6000 heart failure patients, treated over the course of 10 years, could be collected with detailed information on their clinical pathways. Cost estimations were based on the Austrian Diagnosis-Related-Groups system (DRG), where hospital stays are grouped into procedure-oriented diagnosis-related case flat rates. Every case flat rate is associated with a defined length of stay, with an allotted point score that is reimbursed to the hospital. Due to the nature of this calculation system, transgressions of the set treatment window had to be taken into account. Intensive care units (ICU) were modeled with Therapeutic Intervention Scoring System scores (TISS) that are associated with the grade of equipment of the ICU and thus the costs per day of treatment.

Median values or probability density functions can be chosen to simulate the lengths of stay, which correlate with the NYHA class. For the simulated scenarios the intensive care unit was simulated with a TISS score of 32 points, which equals a well-equipped ICU and lengths of stay for admissions were implemented as gamma functions. After the inpatient stay, potential improvement of the patients' state of health was considered. Additionally, treatment specific mortality rates and NYHA class changes after the inpatient care were implemented.

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 outcomes for conventional care. To outline some of the differences between the NYHA classes, a collective of 1000 patients was simulated over five years with 25 % being in each of the NYHA classes I to IV. Mortality rates for inpatient care strongly correlate with the NYHA class and have been implemented for the simulation runs. Outpatient mortality and NYHA class changes have been omitted for the simulated scenarios.

Figure 1 shows the development of the patient collective in terms of number of patients in each respective NYHA class. The dotted lines represent the standard deviation for the mean value, which is shown as solid lines.

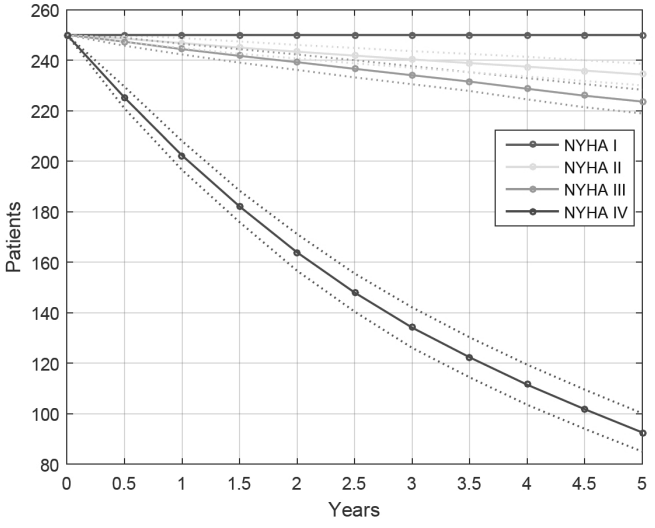


Figure 1: Number of Patients per Year and NYHA Class Considering Mortality (mean ± standard deviation)

Arising costs for the outpatient and inpatient care can be seen in Figures 2 and 3. The dash-dotted line represents mean values of simulation results calculated without inpatient mortality. Again, the solid lines (including mortality) stand for the mean values and the dotted lines show the standard deviation for the iterations.

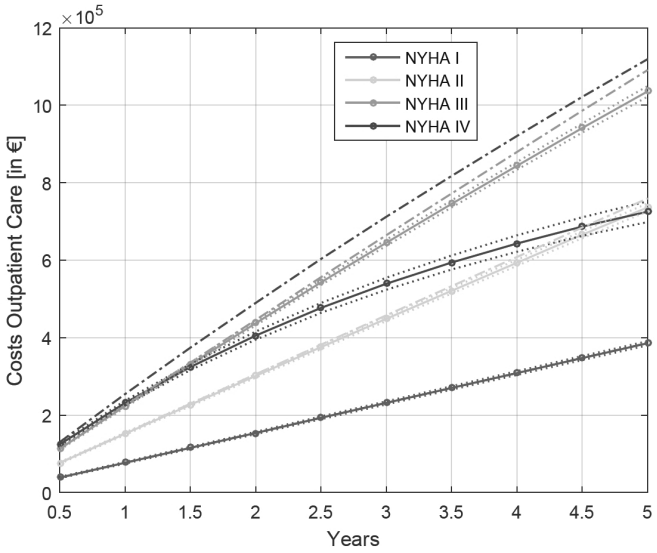


Figure 2: Cumulative Costs of Outpatient Care per Year and NYHA Class (mean ± standard deviation)

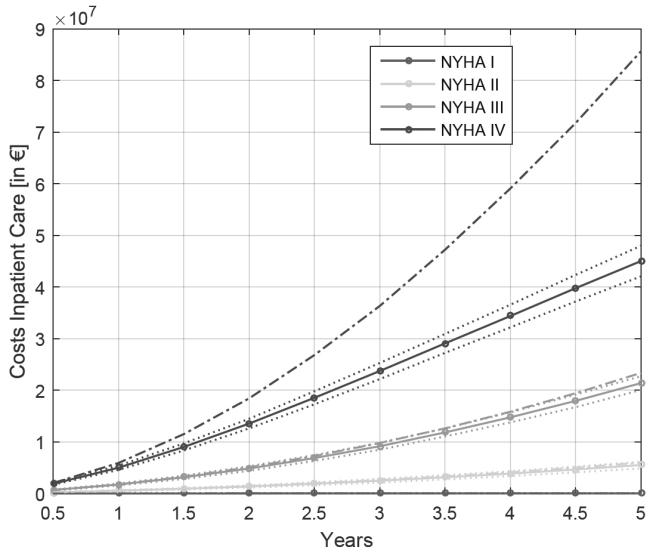


Figure 3: Cumulative Costs of Inpatient Care per Year and NYHA Class (mean ± standard deviation)

DISCUSSION

The equal distribution of NYHA classes for the simulation represents a severely ill collective of heart failure patients. Figure 1 shows that NYHA class IV patients have high inpatient mortality rates compared to the other classes. Of the initial 250 patients 93±7 are alive after five years. This rate would be even higher when accounting for outpatient mortality as well. In agreement with clinical data and literature, higher NYHA classes correlate with an overall increase in costs and mortality rates. Importantly costs for outpatient care display contrasting behavior. The difference of calculations with and without mortality is evident for NYHA class IV, which matches the expense level of NYHA class II patients after five years. Patients in NYHA classes II and III exhibit more consistent behavior and lower hospitalization rates. Due to calculations with only inpatient mortality rates, their costs are not as severely influenced. NYHA class I patients also had inpatient stays, which are negligible compared to the other classes. Compared to inpatient care costs for outpatient care do not correlate as strongly with the state of health. Since no evident information in the clinical data sets on the state of health of patients after visits to the hospital is available for outpatient care, the prescribed medication has to serve as a tool to distinguish different health states. In general heart failure has an unfavorable progression and thus a clear shift of the patients towards higher classes has to be expected.

The limitations of the simulated scenarios, such as the exclusion of outpatient mortality rates and NYHA class changes, are based on available information. Ongoing cooperation with health insurance and health care providers serve to overcome the lack of data in this field. The calculations of the conventional model have been verified and the model is in the phase of validation, based on clinical data and literature. Preliminary results show that the simulated scenarios correlate with findings in literature.

The implementation of the agent based approach is essential in order to observe model behavior and, as opposed to the model by Schroettner et al 2013, understand effects of different methods of care on individual level. Based on these improvements and the extensive data acquisition through mentioned cooperation, new research questions will be investigated after the validation phase, such as rehospitalization rates and comparisons of different telemedical and disease management approaches in dependency of patient profiles.

CONCLUSION

Simulated scenarios show that outcomes for the conventional care comply with findings in literature and clinical data. The built model represents a sophisticated tool to evaluate sustainable solutions for the treatment of heart failure patients. Ongoing cooperation with Austrian health insurance and health care providers are facilitating the validation process by comparing simulation results with clinical data. Following completion, the model will serve as a potential basis for decision-making in health care and thus contribute to the formulation and appropriation of holistic approaches in health care.

REFERENCES

Berry, C., Murdoch, D., and McMurray, J. 2001. “Economics of chronic heart failure.” *European Journal of Heart Failure*, No.3, 283-291.

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.

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., Chen, J., Lin, Z., et al. 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.

Zannad, F., Agrinier, N., and Alla, F. 2009. “Heart failure burden and therapy.” *Europace*, No.11 (Suppl. 5), 1-9.

Modeling Survival Times using Frailty Models

Liberato Camilleri, Roxanne Caruana, Alex Manche'
Department of Statistics and Operations Research
University of Malta
Msida (MSD 06) Malta
E-mail: liberato.camilleri@um.edu.mt

KEYWORDS
Heterogeneity, Shared and Unshared Frailty models, Kaplan Meier, Nelson Aalen, Cox Regression models.

ABSTRACT

Traditional survival models, including Kaplan Meier, Nelson Aalen and Cox regression assume a homogeneous population; however, these are inappropriate in the presence of heterogeneity. The introduction of frailty models four decades ago addressed this limitation. Fundamentally, frailty models apply the same principles of survival theory, however, they incorporate a multiplicative term in the distribution to address the impact of frailty and cater for any underlying unobserved heterogeneity. These frailty models will be used to relate survival durations for censored data to a number of pre-operative, operative and post-operative patient related variables to identify risks factors. The study is mainly focused on fitting shared and unshared frailty models to account for unobserved frailty within the data and simultaneously identify the risk factors that best predict the hazard of death.

1. Introduction

Survival analysis is a useful statistical method for answering questions that deal with the duration of events. Survival models have been used in several research fields to analyze data involving time to a certain event such as death, relapse and onset of a disease. Essentially the duration of a study is treated as the dependent variable and therefore proper definition of the investigation period plays a vital role in determining the number of deaths.

Although there are several types of non-parametric (Kaplan Meier and Nelson Aalen), semi parametric (Cox regression) and parametric techniques to analyze the survival times, these methods do not cater for unobserved heterogeneity. The introduction of frailty models overcomes this limitation. Fundamentally the same principles from survival theory apply, however a multiplicative term is incorporated in the distribution being considered in order to model the impact of frailty.

These models provide a novel approach to survival problems and they encompass two main types of models, namely the unshared and the shared frailty models. In the unshared case a dataset is analyzed assuming that each individual has an associated distinct random effect. The shared case assumes that persons sharing a common factor, such as children born to the same mother or patients with a common health condition, may be analyzed group-wise. Entities within each group are assigned the same frailty effect, but varying heterogeneity levels are expected to subsist among the clusters.

The word frailty was first coined by Vaupel et al., (1979) where it was presented in their research on mortality and later extended by Hougaard (1984). They illustrated that although individuals appear physically alike, they have different threats independently associated to them. In 1984, Hougaard further observed that the difference between the Gamma and Inverse Gaussian distributions is derived from frailty instability among those still alive. In the former case frailty remained steady but in the latter case frailty dropped as individuals grew older. It was further noticed that the random effect had an impact on the hazard equation, which led to the concept that frail persons are bound to decline faster. This unobserved random effect is discussed by several authors in various papers.

Frailty techniques are generally employed to estimate the variance of unobserved risks among individuals. In the univariate scenario, a frailty is assumed to have a unit mean and variance and operates multiplicatively on the baseline hazard. Failure times of particular occurrences are the central purpose for such analysis, as the interest lies in understanding the proneness to some specific occurrence, such as illness. For instance one might be concerned with the recurrence times of smoking after withdrawal, or the time it takes until heart failure sets in. Most often in clinical applications, frailty may be regarded as a means of describing the biological age rather than the chronological age, due to various factors.

The utility of shared frailty models was first highlighted by Clayton and Cuzick in the 70's where the authors emphasized the added benefit of including frailty when heterogeneity impact is common among individuals within a group. Each set has a distinct random effect, which in turn causes frailties to be interrelated. Furthermore, the distinction between a frailty model in the shared and the unshared case lies in the hazard function. Hougaard, and Whitmore & Lee enhanced developments on shared frailty models by addressing frailty by assuming a Weibull baseline hazard function and an Inverse-Gaussian frailty distribution. Flinn and Heckman in 1982 also made use of the lognormal distribution to address frailty.

Shared proportional hazard techniques were introduced primarily through the works of Therneau et al. (2000) and Ripatti and Palmgren (2000). These researchers implemented the penalized partial likelihood (PPL) method to elicit results on frailty models using either a Gamma or an Inverse Gaussian distribution. Subsequently in 2003, Klein and Moeschberger

presented an alternative approach to the PPL method by proposing the application of the expectation-maximization (EM) algorithm. The idea was to determine the variances of the maximum likelihood estimates from the information matrix. Moreover Therneau et al. (2003) proved a very important result, namely that the EM and PPL methods produce equivalent results for the gamma distribution. In fact this was confirmed in their studies which were implemented both in SAS and R.

In 2008, Jenkins developed an algorithm for STATA that allowed the inclusion of a univariate frailty term for discrete event times. He showed that despite the fact that the data comprises discrete event times it is possible to obtain reliable results similar to the continuous parametric techniques. A weakness of this method is that the heterogeneity term is only assumed to have a gamma or a normal distribution. Hence it is only possible to compare between discrete and continuous gamma frailty models. Some of the outstanding works on frailty used in this paper include Wienke (2011), Duchateau and Janssen (2008), Hanagal (2011), and Kleinbaum and Klein (2005).

2. Theory of unshared frailty models

The seminal work of Clayton and Cuzick in the late 70's highlighted the utility of shared frailty models and stressed the added benefit of adding frailty when examining associations between models. As highlighted in the introduction, there are two types of frailty models to analyze survival data in the presence of unobserved heterogeneity. In unshared frailty models, the frailty is introduced at the observation level as an unobservable multiplicative effect, α on the baseline hazard function $h_0(t)$ such that:

$$h\left(t\middle|\alpha\right) = \alpha h_0\left(t\right) \tag{1}$$

In this context, α is a non-negative random mixture variable where $E(\alpha)=1$ and $\text{var}(\alpha)=\sigma^2$. When σ^2 is small, the values of α are located close to 1; however the values of α are more dispersed when σ^2 is large, inducing larger heterogeneity in the individual hazards $\alpha h_0(t)$.

Let $S(t|\alpha)$ denote the survival function of a life conditional on the frailty α and let $\int_0^t h_0(s)ds = M_0(t)$ then

$$S\left(t\middle|\alpha\right) = e^{-\int_0^t h(s|\alpha)ds} = e^{-\alpha\int_0^t h_0(s)ds} = e^{-\alpha M_0(t)} \tag{2}$$

If observed covariates $\mathbf{X}$ are available then the hazard is proportional to the baseline hazard, where the constant of proportionality is the exponential term $\exp(\boldsymbol{\beta}'\mathbf{X})$. So model (1) becomes:

$$h\left(t\middle|\mathbf{X},\alpha\right) = \alpha h_0\left(t\right)\exp(\boldsymbol{\beta}'\mathbf{X}) \tag{3}$$

where $\mathbf{X}=(\mathbf{x}_1,\ldots,\mathbf{x}_p)$ and $\boldsymbol{\beta}=(\beta_1,\ldots,\beta_p)$ is the vector of regression parameters.

The two distributions that are normally considered for the probability density function $f(\alpha)$, of α are the gamma and inverse Gaussian distributions.

Given the simple Laplace transform of the Gamma distribution $\Gamma(k,\lambda)$, it is easy to derive the closed-form expressions of the survival and hazard functions. The exponential distribution is a special case of the Gamma distribution when the shape parameter $k=1$. If α has a Gamma distribution and $\alpha>0$, $\lambda>0$, $k>0$ its probability density function is given by:

$$f\left(\alpha\right) = \frac{\lambda^k}{\Gamma\left(k\right)}\alpha^{k-1}e^{-\lambda\alpha} \tag{4}$$

By setting $k=\lambda=1/\sigma^2$ ensures that the model is identifiable and ensures that $E(\alpha)=1$ and $\text{var}(\alpha)=\sigma^2$. Moreover, the unconditional survival and hazard functions are given by:

$$S(t) = \frac{1}{\left[1+\sigma^2 M_0\left(t\right)\right]^{\left(\frac{1}{\sigma^2}\right)}} \tag{5}$$

$$h(t) = \frac{h_0(t)}{1+\sigma^2 M_0\left(t\right)} \tag{6}$$

Moreover, if observed covariates $\mathbf{x}_i$ are available for life *i* then the mean frailty and frailty variance for a life dying beyond time *t* are given by:

$$E(\alpha\middle|\mathbf{X},T>t) = \frac{1}{1+\sigma^2 M_0\left(t\right)\exp(\boldsymbol{\beta}'\mathbf{X})} \tag{7}$$

$$\text{var}(\alpha\middle|\mathbf{X},T>t) = \frac{\sigma^2(1+\sigma^2)}{\left[1+\sigma^2 M_0(t)\exp(\boldsymbol{\beta}'\mathbf{X})\right]^2} \tag{8}$$

The Inverse Gaussian distribution is also considered as a frailty distribution because similar to the Gamma distribution, simple closed-form expressions exist for the unconditional survival and hazard functions. If α has an Inverse Gaussian distribution and $\alpha>0$, $\lambda>0$, $\mu>0$ its probability density function is given by:

$$f\left(\alpha\right) = \frac{\sqrt{\lambda}}{\sqrt{2\pi\alpha^3}}\exp\left[-\frac{\lambda(\alpha-\mu)^2}{2\mu^2\alpha}\right] \tag{9}$$

By setting $\mu=1$ and $\lambda=1/\sigma^2$ guarantees that the model is identifiable and ensures that $E(\alpha)=1$ and $\text{var}(\alpha)=\sigma^2$. The unconditional density function, the unconditional survival and hazard functions are given by:

$$S(t) = \exp\left(\frac{1-\sqrt{1+2\sigma^2 M_0\left(t\right)}}{\sigma^2}\right) \tag{10}$$

$$h(t) = \frac{h_0(t)}{\sqrt{1+2\sigma^2 M_0\left(t\right)}} \tag{11}$$

If observed covariates $\mathbf{x}_i$ are available for life *i* then the mean frailty and frailty variance for a life dying beyond time *t* are given by:

$$E(\alpha|\mathbf{X},T>t)=\frac{1}{\sqrt{1+\sigma^2 M_0(t)\exp(\boldsymbol{\beta}'\mathbf{X})}} \quad (12)$$

$$\text{var}(\alpha|\mathbf{X},T>t)=\frac{\sigma^2}{\left[1+\sigma^2 M_0(t)\exp(\boldsymbol{\beta}'\mathbf{X})\right]^2} \quad (13)$$

Possible choices for baseline hazard include the exponential, Weibull, Gompertz, log-normal and log-logistics distributions.

3. Theory of shared frailty models

A generalization of the unshared frailty model is the shared frailty model, where the frailty is assumed to be group-specific. Basically shared frailty arises when the heterogeneity impact is common among individuals within a group, yet each set has a distinct random effect, which in turn causes frailties to be interrelated.

Suppose there exist *n* groups and that group *i* comprises n_i observations associated with the unobserved frailty α_i for $1 \le i \le n$. Their hazard functions are given by:

$$h(t|\alpha_i)=\alpha_i h_0(t) \quad (14)$$

Let $S(t|\alpha_i)$ denote the survival function of a life conditional on the frailty α_i and let $\int_0^{t_{ij}} h_0(s)ds = M_0(t_{ij})$ then

$$S(t_{i1},...,t_{in_i}|\alpha_i)=\exp\left[-\alpha_i\sum_{j=1}^{n_i}M_0(t_{ij})\right] \quad (15)$$

If observed covariates $\mathbf{X}_i$ for $1 \le i \le n$ are available then the hazard is proportional to the baseline hazard, where the constant of proportionality is the exponential term $\exp(\boldsymbol{\beta}'\mathbf{X})$. Assuming that the survival times in group *i* are independent, then model (16) becomes:

$$h(t|\mathbf{X}_i,\alpha_i)=\alpha_i h_0(t)\exp(\boldsymbol{\beta}'\mathbf{X}_i) \quad (16)$$

where $\mathbf{X}_i=(\mathbf{x}_{i1},\ldots,\mathbf{x}_{in_i})$ and $\boldsymbol{\beta}=(\beta_1,\ldots,\beta_p)$ is the vector of regression parameters. The conditional survival function on frailty α_i which is shared by all individuals in group *i* is given by:

$$S(t_{i1},...,t_{in_i}|\mathbf{X}_i,\alpha_i)=\exp\left[-\alpha_i\sum_{j=1}^{n_i}M_0(t_{ij})\,e^{\boldsymbol{\beta}'\mathbf{x}_{ij}}\right] \quad (17)$$

The Gamma and Inverse Gaussian frailty models are often used mainly for their nice properties, particularly their simple Laplace transform. Popular choices for the baseline hazard include the Weibull and Gompertz distributions.

4. Application

The dataset used for the frailty model application comprised 365 Maltese patients who underwent aortic valve replacement from 1995 to 2014 at the cardiothoracic centre in a Maltese hospital. Although the ages of the patients ranged from 15 to 87, the vast majority were over 60. In fact it is well known that the risk of requiring heart surgery increases with age. All the patients were followed up after the operation. For those who died, the time of death was recorded in order to compute their survival duration. The majority of the patients were still alive by the end of the investigation period, and their survival times were set equal to the duration between the operation and the end of the investigation period. This type of censoring is non-informative, where observations are right censored.

The data for each patient was recorded by the surgeon conducting the operation. The predictors involved included pre- and post-operative factors, demographic and other patient related explanatory variables. Essentially, the dependent variable, **Time** is the survival duration after surgery recorded on a continuous scale. The variables **Status** is categorical indicating whether the patient died or survived by the end of the investigation period. This variable will be used to identify the censoring status of each patient.

The **Logistic Euroscore** estimates the predicted operative mortality for patients undergoing cardiac surgery. This risk measure of death has a metric scale. **Mechanical+Graft** is a categorical variable indicating the presence or absence of a mechanical valve during surgery with concomitant coronary artery bypass grafting. **Xeno+Graft** is a categorical variable indicating the presence or absence of a biological valve with artery grafting. The variable **Bleeding** records the blood volume, in millilitre, lost post-operatively until removal of chest drains. The variable **Transfusion** records the number of blood units transfused, where 1 unit corresponds to 250ml of blood. **IABP** is a categorical variable indicating whether an intra-aortic balloon pump was required to assist the heart to pump. **Dialysis** is a categorical variable indicating whether the patient was on dialysis due to kidney failure after the operation and the patient's **Age** is measured in years. **CTS** records the duration of patients in the central treatment suite after heart surgery. It is a categorical variable (1-4, 5-10, 11-16, 17 days or more) and will be used as the grouping variables in the shared frailty models.

5. Results of the unshared frailty models

All the fitted models in this section are implemented as proportion hazard models and assume a Gompertz baseline hazard function given by:

$$h_0(t)=\lambda_j e^{\gamma t_j} \quad (18)$$

where $\lambda_j=\exp(\beta_0+\beta_1 x_1+\ldots+\beta_p x_p)$ and γ is an ancillary parameter. Table 1 displays the parameter estimates, standard errors and p-values of the non-frailty model.

Table 1: Estimates of non-frailty model

Parameter	Coef.	S.E.	Z	$P>\|z\|$
Constant	-4.694	3.462	-1.356	0.175
Age	0.059	0.024	2.458	0.014
Logistic Euroscore	0.059	0.045	1.311	0.190
Mechanical+Graft	0.700	0.910	0.769	0.442
Xeno+Graft	0.788	0.913	0.863	0.388
Transfusion	0.192	0.029	6.621	0.000
Bleeding	0.001	0.003	0.333	0.739
IABP	1.358	0.419	3.241	0.001
Dialysis	1.392	0.351	3.966	0.000
Gamma γ	0.069	0.042	1.643	0.100
Log-Likelihood				-225.988
BIC				510.975

The non-frailty model identifies four significant predictors of survival duration. The parameter estimate of Age (0.059) indicates that for every 1-year increase in age the hazard of death increases by 6.1%; the parameter estimate of Transfusion (0.192) indicates for every 1-unit increase in transfused blood the hazard of death increases by 21.1%. The parameter estimate of IABP (1.358) indicates that for patients requiring an intra-aortic balloon pump after heart surgery the hazard of death is 3.89 times in patients who do not require this device. The parameter estimate of Dialysis (1.392) indicates that for patients on dialysis due to kidney failure the hazard of death is 4.02 times in patients who do not have this condition. The parameter estimates of Logistic Euroscore, Mechanical+Graft, Xeno+Graft and Bleeding are not significant because their p-values exceed the 0.05 level of significance. The log-likelihood of the non-frailty model is -225.99 and the estimate of the ancillary parameter (0.069) is not significantly different from 0.

Table 2: Estimates of unshared Gamma frailty model

Parameter	Coef.	S.E.	Z	$P>\|z\|$
Constant	-2.839	5.142	0.552	0.581
Age	0.096	0.043	2.233	0.025
Logistic Euroscore	0.048	0.077	0.623	0.533
Mechanical+Graft	1.574	1.427	1.103	0.270
Xeno+Graft	1.873	1.872	1.001	0.317
Transfusion	0.200	0.091	2.198	0.028
Bleeding	0.011	0.008	1.375	0.169
IABP	3.895	1.066	3.654	0.000
Dialysis	3.239	1.076	3.010	0.003
Gamma γ	0.317	0.089	3.562	0.000
Log (var α)	1.957	0.355	5.513	0.000
Log-Likelihood				-218.505
BIC				501.910

Table 3: Estimates of unshared Inv. Gaussian frailty model

Parameter	Coef.	S.E.	Z	$P>\|z\|$
Constant	-2.918	5.928	-0.492	0.623
Age	0.113	0.045	2.511	0.012
Logistic Euroscore	0.093	0.077	1.208	0.227
Mechanical+Graft	0.862	1.567	0.550	0.582
Xeno+Graft	0.995	1.568	0.635	0.525
Transfusion	0.297	0.058	5.121	0.000
Bleeding	0.003	0.006	0.500	0.617
IABP	2.650	0.743	-3.567	0.000
Dialysis	2.722	0.651	-4.181	0.000
Gamma γ	0.331	0.067	4.940	0.000
Log (var α)	4.484	0.976	4.594	0.000
Log-Likelihood				-216.260
BIC				497.419

To apply the theory described in section 2, unshared Gamma and Inverse-Gaussian frailty models were fitted using Stata streg directives. Table 2 and Table 3 show the parameter estimates, standard errors and p-values of these two unshared frailty models. For both models, the parameter estimates of Age, IABP, Transfusion and Dialysis are significantly positive complementing the results of the non-frailty model. Moreover, the estimates of the frailty variance of the Gamma (5.24) and Inverse Gaussian (88.59) model are both significant, which implies that the data exhibits substantial frailty. In fact, the BIC of the Inverse Gaussian (497.42) and Gamma (501.91) frailty models are considerably lower than the BIC of the non-frailty model (510.97).

6. Results of the shared frailty models

Table 4: Estimates of shared Gamma frailty model

Parameter	Coef.	S.E.	Z	$P>\|z\|$
Constant	-4.395	3.558	-1.235	0.217
Age	0.053	0.024	2.254	0.024
Logistic Euroscore	0.043	0.045	0.945	0.344
Mechanical+Graft	0.646	0.927	0.698	0.485
Xeno+Graft	0.803	0.935	0.859	0.390
Transfusion	0.203	0.029	6.855	0.000
Bleeding	0.002	0.033	0.074	0.941
IABP	1.120	0.426	2.626	0.009
Dialysis	1.387	0.351	3.956	0.000
Gamma γ	0.081	0.040	2.025	0.044
Log (var α)	1.601	0.643	2.488	0.013
Log-Likelihood				-220.813
BIC				506.525

Table 5: Estimates of shared Inv. Gaussian frailty model

Parameter	Coef.	S.E.	Z	$P>\|z\|$
Constant	-4.343	3.581	-1.213	0.225
Age	0.053	0.024	2.255	0.024
Logistic Euroscore	0.043	0.045	0.955	0.341
Mechanical+Graft	0.637	0.929	0.686	0.493
Xeno+Graft	0.791	0.939	0.843	0.399
Transfusion	0.203	0.030	6.844	0.000
Bleeding	0.002	0.033	0.083	0.934
IABP	1.124	0.426	2.636	0.008
Dialysis	1.389	0.035	3.959	0.000
Gamma γ	0.081	0.040	2.020	0.043
Log (var α)	2.803	1.342	2.089	0.037
Log-Likelihood				-218.559
BIC				502.017

To apply the theory described in section 3, shared Gamma and Inverse-Gaussian frailty models were fitted using Stata streg directives. The models are implemented as proportion hazard models and assume a Gompertz baseline hazard function. Table 4 and Table 5 show the parameter estimates, standard errors and p-values of these two shared frailty models. Both models, confirm that IABP, Age, Transfusion and Dialysis are significant predictors of the hazard of death. Moreover, the estimates of the frailty variance are both significant, which indicates the presence of substantial frailty. The unshared Inverse Gaussian frailty model yields the lowest BIC value (497.42) implying that it provides the best fit. On the other, the non-frailty model yields the highest BIC value (510.98) implying that it provides the poorest fit.

6. Conclusion

This paper presents two shared and two unshared frailty models assuming a Gamma or an Inverse Gaussian frailty distribution and a Gompertz baseline hazard function. This paper shows that in the presence of heterogeneous data these models provide a significantly better fit than non-frailty ones. For this data, the Inverse Gaussian assumption for the frailty distribution provided a better fit than the Gamma distribution.

An alternative approach to these parametric models is to fit semi-parametric frailty models, which do not require any assumptions on the baseline hazard function. These models can be implemented using the coxph directive in the R statistical software, where parameters are estimated using the EM (expectation maximization) algorithm, which iterates between two steps. The first step estimates the unobserved frailties and model parameters based on observed data. These estimates are used in the maximization step to obtain updated parameter estimates given the estimated frailties. The iterative procedure is continued until it converges. The likelihood includes both the observed data and unobserved frailties, which are assumed to be random. These models can also be implemented using the frailtypack in the R package, which uses the PPL (penalized partial likelihood) approach. However, this estimation method can yield different results when compared to the coxph approach. In frailty models these techniques work best when the random effects are significant. STATA has the facility to fit semi-parametric Gamma frailty models but not Inverse Gaussian models.

References

Clayton, D. G. & Cuzick J. (1985), Multivariate generalizations of the proportional hazard model. Journal of the Royal Statistical Society A, 148(2):82–117.

Dempster, A. P., Laird, N. M. & Rubin, D. (1977), Maximum Likelihood from Incomplete Data via the EM algorithm, *Journal of the Royal Statistical Society*, B, 39, 1-38.

Duchateau, L. & Janssen, P. (2008), The frailty model. New York Springer.

Flinn, C. & Heckman, J. (1982), New methods for analyzing structural models of labour force dynamics. Journal of Econometrics 18: 115-168.

Hanagal, D. D. (2011), Modeling survival data using frailty models. Chapman & Hall/CRC

Horowitz J. L. (1999), Semiparametric estimation of a proportional hazard model with unobserved hetrogeneity. Econometrica, 67(5):1001–1028.

Hougaard. P. (1984), Life table methods for heterogeneous populations: Distributions describing the heterogeneity. Biometrika, 71(1): 75–83.

Ripatti, S. & Palmgren, J. (2000), Estimation of multivariate frailty models using penalized partial likelihood. Biometrics 56: 1016-1022.

Therneau, T. M., Grambsch, P. M. &Pankratz, V. S. (2003), Penalized survival models and frailty. Computational and Graphical Statistics Journal, 12(1): 156-175.

Vaupel, J.W., Manton, K.G. & Stallard, E. (1979), The impact in individual frailty on the dynamics of mortality. Demography, 16(3), 439-454.

Wienke, A. (2010), Frailty models in survival analysis. Chapman and Hall/CRC

AUTHOR BIOGRAPHY

LIBERATO CAMILLERI studied Mathematics and Statistics at the University of Malta. He received his PhD degree in Applied Statistics from Lancaster University. His research specialization areas are related to statistical models, which include Generalized Linear models, Latent Class models, Multilevel models and Mixture models. He is an associate professor and Head of the Statistics department at the University of Malta.

AUTHOR LISTING

AUTHOR LISTING

Albayrak S. **356**
Albert V. **365**
Allison C. **336**
Arokiam A. **328**
Aschoff R.R. **283**
Atanasova T. **404**

Backfrieder C. **252**
Balsamo S. **288**
Barat S. **135**
Barbas H. **25**
Barbusiński K. **316**
Barn B. **135**
Barta J. **143**
Bäsig J. **295**
Baumgartner C. **425**
Behrisch M. **247**
Belaidi A. **328**
Bellemare J. **206**
Bezoušek P. **273**
Bhattarai S. **76**
Bieker-Walz L. **247**
Bismans F. **325**
Borovička A. **103**
Brandau C. **301**
Brinkjans J. **162**
Brožek J. **258**
Büchter H. **264**

Cakir V. **167**
Călinoiu C. **5**
Camilleri L. **428**
Cardoso J. **343**
Caruana R. **428**
Choi S.Y. **278**
Cichon T. **384**
Clark T. **135**
Clausen U. **233**
Colloc J. **13**
Conley W. **46/54**
Covaci F.L. **185**

Dahal K. **76**
de Freitas Almeida J.F. **212**
Dineva K. **404**
Dumond Y. **93**
Dupleac D. **336**

Elmer T. **174**
Elsen S. **63**
Ettlinger M. **356**

Fey D. **295**
Fikejz J. **258**
Foucher C. **365**
Furian N. **174/192/221**

Gimm K. **247**
Granados C.S. **415**
Grazebrook A. **350**
Gutschi C. **174**

Heinzl B. **157**
Hernandez-Cuellar M. **415**
Hernandez-Fragoso A. **88**
Higgins M. **328**
Holtkötter C. **162**
Hrabia C.-E. **356**

Irimia P.-C. **5**
Ivaschenko A. **227/390**

Jafri M. **288**
Janssens G.K. **201**
Jemai A. **81**
Junghans M. **247**
Juryca K. **273**

Kabašinskas A. **119**
Kalantar P. **162**
Kandavalli V.K. **418**
Kang H.S. **278**
Kapp K. **111**
Karaca T.K. **167**
Kargapolova N.A. **311**
Kastner W. **157**
Klüver C. **397**
Klüver J. **397**
Kohl J. **295**
Kolsanov A. **390**
Kulkarni V. **135**
Kwon H.J. **278**

Ladbrook J. **328**
Lassnig A. **425**
Latos B.A. **162**
Lefebvre D. **33**
Leobner I. **157**
Lindorfer M. **252**
Litvine I. **38/68/111/325**

AUTHOR LISTING

Manche A. **428**
Marin A. **288**
Mecklenbräuker C.F. **252**
Merkuryeva G. **201**
Michel C. **343**
Morais J.C. **212**
Mösl B. **192**
Motie Y. **148**
Mudimu E. **407**
Mulvany R. **328**
Musolino G. **238**
Mütze-Niewöhner S. **162**

Nakayama M.K. **49**
Naumann S. **264**
Navrátil J. **143**
Nazaryan A. **390**
Neubacher D. **174/192/221**
Nistor-Vlad R.-M. **336**
Nketsa A. **148/365**
Nunes G.M. **212**

Operacz A. **316**
Ostermayer G. **252**
O'Sullivan M. **221**

Palacios-Enriquez A. **88**
Petö J. **283**
Pinto L.R. **212**
Ponomaryov V. **88/415**
Poswayo S. **38**
Preyser F. **157**
Prisecaru I. **336**
Przybysz P.M. **162**
Puhalschi R. **5**

Rácz S. **283**
Radev D. **381**
Radeva S. **381**
Raich P. **157**
Ramaekers K. **201**
Rao D.M. **127**
Reichardt R. **119**
Ribeiro A.S. **418**
Rindone C. **238**
Roleček J. **273**
Rossmann J. **384**
Ryabchenko O. **68**

Santner P. **221**
Sarp B. **356**
Schmidt E. **127**
Schröttner J. **425**
Siron P. **343**
Sitnikov P. **227**
Smei H. **81**
Smiri K. **81**
Smolek P. **157**
Sokolov S. **381**
Startceva S. **418**
Studziński J. **316**
Šutienė K. **119**
Syusin I. **227**
Szabó G. **283**
Szeląg B. **316**

Thöndel E. **373**
Truillet P. **148**
Tutsch D. **301**

Valakevičius E. **119**
Vasiliu D. **5**
Vasiliu N. **5**
Vichare P. **76**
Vitetta A. **238**
Vössner S. **174/192/221**

Walker C. **221**
Wright S. **350**

Yoon S.J. **278**

Zhang D. **233**
Zhu M. **365**