15^{th} INTERNATIONAL INDUSTRIAL

SIMULATION CONFERENCE

2017

ISC'2017

EDITED BY

Janusz Kacprzyk

and

Jan W. Owsinksi

May 31- June 1, 2017

WARSAW, POLAND

A Publication of EUROSIS-ETI

15th Industrial Simulation Conference 2017

WARSAW, POLAND

MAY 31- JUNE 1, 2017

Organised by

ETI - The European Technology Institute

Sponsored by

EUROSIS

University of Skövde

SIEMENS

Co-Sponsored by

Ghent University

Hosted by

Systems Research Institute
Polish Academy of Sciences

Warsaw, Poland

INTERNATIONAL PROGRAMME COMMITTEE

Roland Wischnewski, (RIF) e.V., Dortmund, Germany
Farouk Yalaoui, Universite de Technologie de Troyes, Troyes Cedex, France

Simulation in Steel Manufacturing
Brian Hollocks, Bournemouth University, Bournemouth, United Kingdom

Simulation in Automotive Systems
Naoufel Cheikhrouhou, EPFL, Lausanne, Switzerland
Julien Richert, Daimler AG, GR/PAA, Sindelfingen, Böblingen, Germany

Simulation in Robotics
A. Chatzinikolaou, Athens, Greece
Andrzej Dzielinski, Warsaw University of Technology, Warsaw, Poland
Markus Koch, Orga Systems GmbH, Paderborn, Germany
Martin Mellado, UPV, Valencia, Spain
Bogdan Raducanu, Computer Vision Centre, UAB, Barcelona, Spain
Sergiu-Dan Stan, Technical University of Cluj-Napoca, Cluj-Napoca, Romania

Simulation in Electronics, Computer and Telecommunications
Teresa Alvarez, University of Valladolid, Valladolid Spain
Christos Bouras, University of Patras, Patras, Greece
Adnane Latif, Cadi Ayyad University, Marrakech, Morocco
Silvia Mirri, University of Bologna, Bologna, Italy
Maurizio Palesi, Universita di Catania, Catania, Italy
Marco Roccetti, University of Bologna, Bologna, Italy
Fernando Boronat Segui, UPV, Valencia, Spain
Renate Sitte, Griffith University, Gold Coast, Australia

Simulation in Electronics Manufacturing
Theresa Roeder, San Francisco State University, USA
Gerald Weigert, Dresden University of Technology, Germany

Simulation of Complex Multiprocessor Systems
Orhan Gemikonakli, Middlesex University, London, United Kingdom

Simulation in Computer Science
Lipika Deka, Tezpur University and IIT Guwahati, India
Ernst Kesseler, NLR, Amsterdam, The Netherlands
Wolfgang Kreutzer, University of Canterbury, Christchurch, New Zealand

Simulation in Logistics, Transport and Harbour Simulation
Maria Sameiro Carvalho, University of Minho, Guimaraes, Portugal
Isabel Garcia Guttierez, Univ. Carlos III de Madrid, Madrid, Spain .
Dmitry Ivanov, Chemnitz University of Technology, Chemnitz, Germany
Peter Lawrence, Swinburne University, Lilydale, Australia
Marie-Ange Manier, UTBM, Belfort, France
Roberto Montemanni, IDSIA, Manno-Lugano, Switzerland
Rosaldo Rossetti, University of Porto, Porto, Portugal
Hans Veeke, TU Delft, Delft, The Netherlands
Pengjun Zheng, University of Southampton, Southampton, United Kingdom

Simulation in the aviation sector
Gabriel Lodewijks, Delft University of Technology, Delft, The Netherlands

Hospital Logistics Simulation
Track Chair: Giorgio Romanin-Jacur, University of Padova, Vicenza, Italy
Antonio Abelha, Universidade do Minho, Braga, Portugal
Jose Machado, University of Minho, Braga, Portugal
Peter Summons, University of Newcastle, Australia

INTERNATIONAL PROGRAMME COMMITTEE

Complex Systems Modelling
Track Chair: Igor N Litvine, Nelson Mandela Metropolitan University, Port Elizabeth, South Africa
Frantisek Capkovic, Slovak Academy of Sciences, Bratislava, Slovak Republic
Alexandre Nketsa, LAAS, Toulouse, France
Miguel Rocha, University do Minho, Braga, Portugal
Alfonso Urquia, UNED, Madrid, Spain

Apparel and Textile Simulation
Track Chair: Jocelyn Bellemare, University of Quebec in Montreal (UQAM), Montréal (Québec) Canada

Simulation in Aerospace
Reza Azadegan, Urmia University, Urmia, Iran
Wolfgang Kuehn, University of Wuppertal, Wuppertal, Germany
Martin Spieck, DLR, Goettingen, Germany

Marine Simulation
Sergeij Kalashnkikow, DANFOSS, Austria

Simulation in Industrial Design and Product Design
Chiara Catalano, IMATI-CNR, Genoa, Italy
Yan Luo, NIST, Gaithersburg, USA
Catarina Rizzi, University of Bergamo, Bergamo, Italy

Simulation in Engineering Processes
Chrissanti Angeli, Technological Institute of Piraeus, Athens, Greece
Alejandra Gomez Padilla, University of Guadalajara, Mexico
Jan Studzinski, Polish Academy of Sciences, Warsaw, Poland
Joao Tavares, University of Porto, Porto, Portugal
Henk Versteeg, Loughborough University, Loughborough, United Kingdom

Civil and Building Engineering
Alistair Borthwick, Oxford University, Oxford, United Kingdom
Graham Saunders, Loughborough University, Loughborough, United Kingdom

Simulation in Energy and Power Systems
Sergeij Kalashnkikow, DANFOSS, Austria
Janos-Sebestyen Janosy, KFKI Atomic Energy Research Institute, Budapest, Hungary

Simulation in Multibody Systems
Ignacio Garcia-Fernandez, University of Valencia, Valencia, Spain

Simulation in Chemical, Petroleum and Mining Engineering
Mohamad R. Riazi, Kuwait University, Kuwait

Simulation in Military and Defense
Roberto de Beauclair Seixas, IMPA, Rio de Janeiro, Brazil
Carlos Palau, UPV, Valencia, Spain
Matthias Reuter, CUTEC Gmbh, TU-Clausthal, Clausthal, Germany

Virtual Reality and Graphical Simulation in Industrial Simulation
Track Chair: Guodong Shao, NIST, Gaithersburg, USA
Emilio Camahort, Universidad Politecnica de Valencia, Valencia, Spain
Anders Hast, University of Gavle, Gavle, Sweden
Fabrizio Lamberti, Politecnico di Torino, Turin, Italy
Christoph Laroque, University of Paderborn, Paderborn, Germany
Sudhir Mudur, Concordia University, Montreal, Canada
Marta Pla-Castells, Universidad de Valencia, Valencia, Spain
Marcos A.Rodrigues, Sheffield Hallam University, Sheffield, United Kingdom

INTERNATIONAL PROGRAMME COMMITTEE

Simulation of Complex Multiprocessor Systems
Peter Kvasnica, Alexander Dubcek University of Trencin, Trencin, Slovakia

Verification, Validation and Accreditation
Roberto Revetria, University of Genoa, Genoa, Italy
Agustin Yague, Technical University of Madrid, Madrid, Spain

Simulation and Training
Manuel Alfonseca, Universidad Autonoma de Madrid, Madrid, Spain
Wenji Mao, Chinese Academy of Sciences, Beijing, P.R. China
Eshan Rajabally, Loughborough University, Loughborough, United Kingdom
Gerhard Schreck, Fraunhofer IPK, Berlin, Germany

Workshops

Simulation-based evaluation in Living Labs
Stefan Hellfeld, FZI - House of Living Labs, Karlsruhe, Germany
Benjamin Herd, King's College, London, United Kingdom

Workshop on Intelligent Transport Systems
Track Chair: Anna Syberfeldt, University of Skovde, Skovde, Sweden
Paul Davidsson, Blekinge Institute of Technology, Ronneby, Sweden
Petr Hanacek, Brno University of Technology, Brno, Czech Republic
Jairo Montoya Torres, Universidad de la Sabana, Chia, Columbia

NANOSIM
Clemens Heitzinger, Cambridge University, Cambridge, United Kingdom
Yong K. Kim, University of Massachusetts Dartmouth, Dartmouth, USA
Javier Marin, ETSI, University of Malaga, Malaga, Spain

Workshop Augmented Reality and Pervasive Systems in Simulation
Alessandro Genco, University of Palermo, Palermo, Italy

Workshop on Simulation in Lean Manufacturing
Track Chair: Leif Pehrsson, Volvo Car Corporation, Skovde, Sweden
El-Houssaine Aghezzaf, Ghent University, Ghent, Belgium
Hendrik Van Landeghem, Ghent University, Ghent, Belgium

Simulation-based Optimization in Industry
Amos H.C. Ng, University of Skovde, Skovde, Sweden
Anna Syberfeldt, University of Skovde, Skovde, Sweden

OPENSIM
Philippe Geril, ETI Bvba, Ostend, Belgium

HORIZON2020
Valerian Croitorescu, University of Bucharest, Bucharest, Romania

INDUSTRIAL SIMULATION 2017

All articles published in these Proceedings have been peer reviewed

EUROSIS-ETI Publications are ISI-Thomson, IET, SCOPUS and Elsevier Engineering Village referenced
Legal Repository: Koninklijke Bibliotheek van België, Keizerslaan 4, 1000 Brussels, Belgium
CIP 12.620 D/2011/12.620/1

Selected papers of this conference are published in scientific journals.

For permission to publish a complete paper write EUROSIS, c/o Philippe Geril, ETI Executive Director, Greenbridge NV, 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, B-8400 Ostend, Belgium

Printed and bound in Belgium by Reproduct NV, Ghent, Belgium
Cover Design by Grafisch Bedrijf Lammaing, Ostend, Belgium

EUROSIS-ETI Publication
ISBN: 978-90-77381-98-4
EAN: **978-90-77381-98-4**

Preface

The present volume aims at complementing the Readers' knowledge, concerning the contemporary theories, methods and tools, as well as applications, in the truly broadly conceived domain of industrial systems simulation. The thus really broadly conceived domain includes not just all of those methodologies that can either directly or through appropriate extensions serve the literal purpose of simulating (emulating) the functioning of diverse aspects of the productive systems, but also all of those that can actually be used to represent, simplify, plan and forecast, optimize and control the systems of quite diverse characteristics, first of all, naturally, those that serve the productive aims.

This volume brings together not only various methodologies and theories, originating from a wide variety of scientific backgrounds, mathematical, statistical and purely pragmatic, not only a variety of potential, explicit or implicit, application fields, ranging from computer science, through logistics down to politics and the manipulation of the public scene, but, as well, a variety of institutional frameworks, within which both the respective research work, and the implementation efforts have been undertaken. In fact, we deal here with quite an overarching set of scientific studies and research teams.

Thus, first of all, the private body, having in mind the public interest, which is the main organiser of the series of conferences, here documented for its 2017 edition, namely EUROSIS, from the West of Europe, bent on bringing science to business, meets here a fully public organisation, devoted to fundamental research, from the East of Europe, that is – Polish Academy of Sciences, embodied in this case by the Systems Research Institute of the Academy. Second, in a similar vein, one easily perceives a characteristic composition of the teams of authors, international in a pronounced manner in quite an important part of cases. This international, and, indeed, inter-cultural aspect is, as well, reflect in quite adequate metaphoric manner the multifaceted globalisation of the present-day economic, social, and, well – also the political life.

A Reader is treated in this volume to a range of mature methodologies that have proven their usefulness in the definite domains of application, their origins being distributed over time across many decades, indeed, almost a century… If you are surprised by this statement, look at these: generalized nets, virtual reality, agent technologies, social networks, genetic algorithms, fuzzy logic, linear programming, graph theory,… and so on. In some cases we deal with outward use of pragmatic techniques that refer to what is called – not so justly, though – 'common sense' (the sense that we aim at being, actually, not exactly so 'common', as one can easily see).

Likewise, an interesting and, certainly, useful, choice of application domains is offered by the studies here presented (whenever we deal with actual application domains, which, actually, is the case of most, but not all, papers). From purely computer-science-oriented ones, and those, related to computer technology, through logistics and supply chain issues, internal cost analysis, information management, social network use, market research, up to hard mechanical production and operation questions. Many of the here contained studies touch upon very practical, down-to-earth issues, related to concrete industries and services, which are also often quite telling to an "average consumer", a "man-in-the-street" – whether we deal with the water supply and sewage system, or with consumer-oriented supply-demand management within a logistic system. This is also a telling determinant of this volume, in addition to the quite important methodological developments.

We certainly hope that this volume shall add in a significant manner to the body of knowledge in the respective domains, to the methodological state-of-the-art, and to the overall advancement of applicable know-how in the here approached, directly and indirectly, variety of domains.

We thank all the authors, the editor of this volume and the conference series, and we dedicate this selection of papers to the memory of all those, who have worked, discussed, planned, and fantasised in the context of both science and society in the airs of Palais Staszic[1] , where the meeting here, in a way, reported here, has been taking place, and to those of their contemporary spiritual descendants, who enrich the communities of their own choice and appurtenance, along with the entire world we live in.

Jan W. Owsiński, Janusz Kacprzyk

The Editors

[1] Stanisław Staszic was a scientist, clergyman, minister, as well as social and political activist of the turn of 18th and 19th centuries, a critical to Polish history, who marked in a distinctly positive way this period with his discoveries (e.g. in the field of geology), actions and practical initiatives.

CONTENTS

CONTENTS

SCIENTIFIC PROGRAMME

KEYNOTE

UNIVERSAL PETRI AND SLEPTSOV NETS

Dmitry A. Zaitsev
Computer Engineering Department
International Humanitarian University
st. Fontanskaya Doroga, 33
Odessa 65009, Ukraine
E-mail: daze@acm.org

KEYWORDS

Universal Sleptsov net; universal Petri net; cellular automata; concurrent programming; massively parallel computations

ABSTRACT

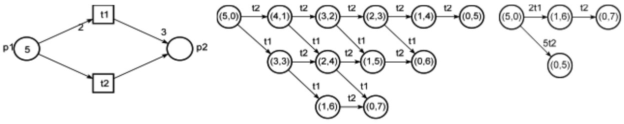
A universal Petri net represents a processor in the Petri net paradigm of computing; it executes (runs) a program specified by a Petri net which initial marking represents input data and final marking represents output data. A class of place/transition nets with multiple firing of a transition at a step, called Sleptsov nets, run fast compared to Petri nets that opens prospects for their practical application as a concurrent programming language. A series of universal nets have been constructed either directly or via simulating universal Turing machines and cellular automata. Examples of RSA encryption/decryption, solving Laplace equation, computing a fuzzy logic function, and fast discrete-time linear control illustrate principles of programming in Sleptsov nets.

INTRODUCTION

Petri nets are applied widely for modeling concurrent processes but their application as a concurrent programming language has been restricted to the area of the programmable logic controllers only (Zaitsev 2012; 2014b). Introducing a more general class of Sleptsov nets (Zaitsev 2016) made computations on them efficient that opens wide prospects for Sleptsov net computing (Zaitsev 2017c) where a universal net represents a prototype of a processor.

SLEPTSOV NETS RUN FAST

Sleptsov nets (SNs) (Zaitsev 2016), comparing to Petri nets (PNs), allow multiple firing of a transition at a step. Reachability graphs (RGs) of SN (c) and PN (b) obtained for the same substrate place/transition net (a), Fig. 1.



a) substrate b) RG of PN c) RG of SN

Figure 1: Behavior of PNs and SNs

According to (Zaitsev 2014a), SNs run exponentially faster comparing Petri nets implementing multiplication and division in logarithmic time.

SLEPTSOV NET PARADIGM OF COMPUTING

Using Sleptsov nets makes possible a uniform paradigm of computing (Zaitsev 2017c; 2014b) which does not involve other concepts for efficient massively parallel computations as shown in Fig.2.

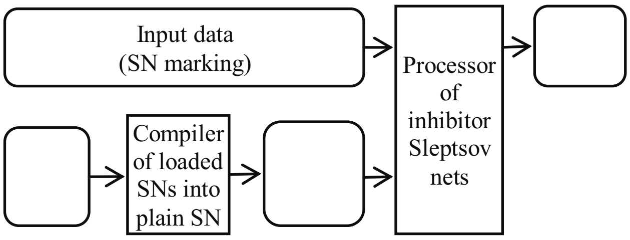


Figure 2: Sleptsov net computing

A program specified as a hierarchical inhibitor Sleptsov net, representing a composition of reverse control flows with data, is compiled into a plain net which is used as a machine language (Zaitsev and Jurjens 2016). The program runs on a Sleptsov net processor which implements a universal Sleptsov net concept.

UNIVERSAL NETS

Direct description of net behavior

A direct description of the net behavior by an inhibitor Petri/Sleptsov net produces rather bulky universal nets containing hundreds of nodes (Zaitsev 2012) which basic advantage consists in the conceptual unity.

Simulating universal Turing machines

Simulating small (weakly) universal Turing machines of Neary and Woods allowed the obtaining of smallest universal Petri/Sleptsov nets (Zaitsev 2017a; 2014a) containing about forty nodes shown in Fig.3.

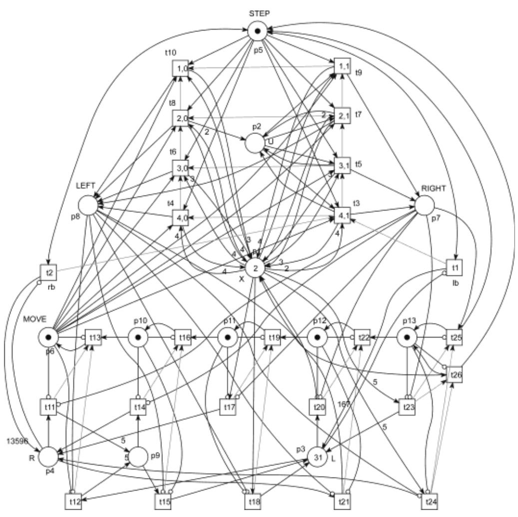


Figure 3: Universal Sleptsov net

Simulating universal cellular automata

Simulating universal cellular automata, particularly rule 110, especially with synchronous Petri nets (Zaitsev 2017b), lead to the smallest cell model consisting of 1 place and 2 transitions only shown in Fig.4.

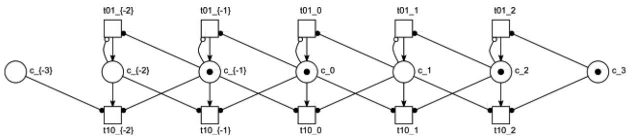


Figure 4: SN simulates universal cellular automaton

The replacing infinite row of cells with its finite segment of required length results in real-life massively parallel fast computations.

EXAMPLES OF SLEPTSOV NET PROGRAMS

RSA encoding/decoding

An SN which implements RSA encryption/decryption based on the expression $y = x^z \ mod \ n$ is represented in Fig.5.

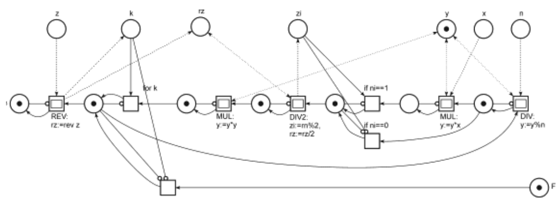


Figure 5: SN program of RSA encryption/decryption

Solving Laplace equation

An SN which solves Laplace equation $\frac{\delta^2\varphi}{\delta x^2} + \frac{\delta^2\varphi}{\delta y^2} = 0$ is composed as shown in Fig.6.

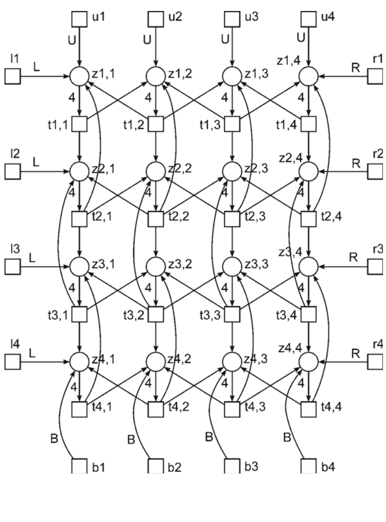


Figure 6: SN program solves Laplace equation

Computing fuzzy logic function

An SN that computes a fuzzy logic function $\varphi = x_1\bar{x}_2 \vee \bar{x}_1 x_2$ is represented in Fig.7.

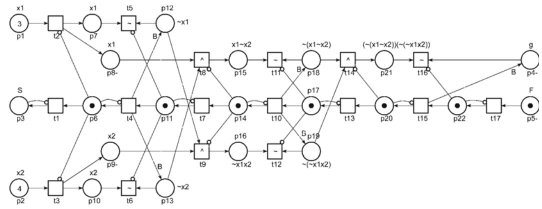


Figure 7: SN program computes fuzzy logic function

Discrete Linear Control

An SN which implements in two tacts discrete-time control, given with a system

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) \\ y(k) = Cx(k) + Dy(k) \end{cases}$$

is shown in Fig.8.

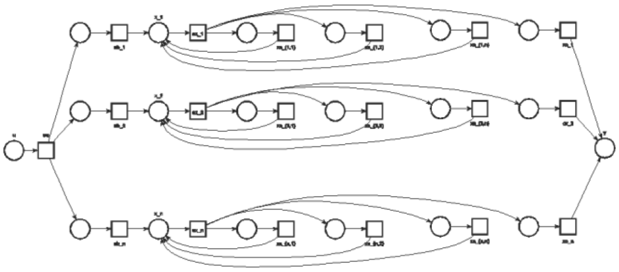


Figure 8: SN program of discrete linear control

CONCLUSIONS

Universal Sleptsov nets represent a prototype of processor in Sleptsov net computing – a massively parallel paradigm of computations. A series of universal nets have been constructed. Principles of programming in Sleptsov nets have been developed and illustrated with examples.

REFERENCES

Zaitsev, D.A. 2017c. "Sleptsov Net Computing". Chapter 672 in *Enciplopedia of Information Science and Technology*, Fourth Edition, IGI-Global: USA.

Zaitsev, D.A. 2017b. "Simulating Cellular Automata by Infinite Petri Nets". *Journal of Cellulr Automata*, 1-24.

Zaitsev, D.A. 2017a. "Universal Sleptsov Net". *International Journal of Computer Mathematics*. 1-13.

Zaitsev, D.A. and J. Jurjens. 2016. "Programming in the Sleptsov net language for systems control". *Advances in Mechanical Engineering*, 8(4), 1–11.

Zaitsev, D.A. 2016. "Sleptsov Nets Run Fast". *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 46(5) 682-693.

Zaitsev, D.A. 2014b. "Paradigm of Computations on the Petri Nets". *Automation and Remote Control*, 75(8), 1369–1383.

Zaitsev, D.A. 2014a. "Toward the Minimal Universal Petri Net". *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 44(1), 47-58.

Zaitsev, D.A. 2012. "Universal Petri Net". *Cybernetics and Systems Analysis*, 48(4), 498-511.

DMITRY A. ZAITSEV was born in May 13, 1964, in Konstantinovka, Ukraine. He received the Eng. degree in applied mathematics from Donetsk Polytechnic Institute, Donetsk, Ukraine, in 1986, the Ph.D. degree in automated control from Kiev Institute of Cybernetics, Kiev, Ukraine, in 1991, and the Dr.Sc. degree in telecommunications from Odessa National Academy of Telecommunications, Odessa, Ukraine, in 2006. Since 2009, he has been with the International Humanitarian University, Odessa, where he is currently a Professor of computer engineering. Since 2014, he has been with the Vistula University, Warsaw, where he is currently a Professor of computer science. He published a monograph, two book chapters and more than hundred papers. His major field of study is Petri net theory and its application in networking, production control, and computing. He developed universal Petri nets, the analysis of infinite Petri nets with regular structure, the decomposition of Petri nets in clans, and the method of synthesis of fuzzy logic function given by tables. Prof. Zaitsev is a senior member of the Association for Computing Machinery: http://member.acm.org/~daze

SOFTWARE TOOLS AND METHODS SIMULATION

A TOOLBOX FOR FUZZY LOGIC FUNCTIONS SYNTHESIS ON A CHOICE TABLE

Dmitry A. Zaitsev
Computer Science Department
Vistula University
st. Stoklosy, 3
Warsaw 02-787, Poland
E-mail: daze@acm.org

KEYWORDS

Fuzzy logic function; choice table; synthesis; constituent of maximum; disjunctive normal form; partitioning

ABSTRACT

A toolbox for fuzzy logic functions synthesis on a choice table, available for free download on GitHub, has been implemented in C language. The early published method of a continuous (fuzzy) logic function synthesis on a choice table has been adjusted for fast partitioning the source choice table with a set of fuzzy logic functions. The toolbox implements a command line style of programming using data located in textual files of simple intuitive formats. The toolbox can process big data rapidly and can be easily integrated into fuzzy logic frameworks as a synthesis engine for developing graphical environment of fuzzy (control) systems design. Formal aspects of optimality (minimalism) are directions for future research.

INTRODUCTION

Fuzzy sets and fuzzy logic find wide application in intelligent and control systems design (Kaufmann 1977; Kandel 1986; Novak et al 2016). Since a fuzzy logic function (FLF) of Zadeh (Zadeh 1965) takes value of one of its arguments or a negation of an argument, the function can be given by a choice table (CT) (Volgin and Levin 1990) where all the variants of ordering of arguments and their negations are listed and for each variant the function value is pointed out. Without loss of generality, we suppose that an FLF is represented in disjunctive normal form (DNF). The choice table is considered as an analog to the truth table of the conventional binary logic (Kleene 1967) because it gives a simple way to compare FLFs via building and comparing their CTs. However, not any choice table defines a (single) fuzzy logic function; a table can be partitioned and covered by a set of fuzzy logic functions valid on subdomains (Zaitsev et al 1998). Topics of fuzzy logic synthesis from the behavioral description have been studied in (Wielgus 2004).

In (Zaitsev et al 1998), a criterion to tell whether a given choice table defines a fuzzy logic function has been introduced and an algorithm of a fuzzy logic function synthesis has been constructed. The total time complexity of the algorithm is linear in the table length. However, the calculation of the criterion is square in time with respect to the table length. In case a table does not define a fuzzy logic function, a sophisticated procedure is stipulated to find subsets of the table rows which are not mutually overlapped. The time complexity of this procedure can be rather great taking into consideration that the table length is an exponent in the number of the function arguments. Fuzzy logic operations enter the state equation of Petri nets with multichannel transitions (Zaitsev and Sleptsov 1997) which have been recently applied as a concurrent programming language (Zaitsev and Jurjens 2016). Colored Petri nets (Zaitsev and Shmeleva 2006) use similar concepts; their recent application for modeling grid structures (Zaitsev et al 2016), based on a substrate (Shmeleva et al 2009), revealed complex deadlocks.

The present work introduces a simple heuristic technique which accomplishes the results of (Zaitsev et al 1998) regarding partitioning a given choice table into a set of subdomains (subsets of the table rows) and synthesizing a separate fuzzy logic function for each partition. The total time complexity of the technique is linear in the number of the table rows. A toolbox, available for free download on GitHub, has been developed; it contains the following tools: synthesize a DNF on a choice table; create a choice table on a DNF; check whether two functions/tables coincide; partition a choice table into subdomains supplied with a set of DNFs valid for each subdomain; generate a random choice table. The partitioning technique has been justified statistically on sets of random choice tables. The formal aspects of optimality (minimalism) are left beyond the scope of the present paper as a direction for future research.

BASIC CONCEPTS AND NOTIONS

For a *domain* represented with an interval of real numbers $D = [0,1]$, operations of *conjunction, disjunction, and negation*, are introduced as follows:

$$\begin{aligned} x \wedge y &= \min(x, y), \\ x \vee y &= \max(x, y), \\ \bar{x} &= 1 - x, \end{aligned} \qquad (1)$$

respectively, where $x, y \in D$.

A *fuzzy logic function* of n arguments is a function $f: D^n \to D$ obtained as a result of superposition of operations (1) on independent variables $x_1, x_2, \dots, x_n \in D$. Note that according to the above definition, an FLF takes value of an argument or a negation of an argument.

For FLFs, the following *basic laws* of Boolean algebra (Kleene 1967) are valid: commutative, associative, distributive for both conjunction and disjunction, absorption, double negation, idempotency of elements, Kleene and de-Morgan. However, the exception of the third law, $x \wedge \bar{x} \neq 0$ and $y \vee \bar{y} \neq 1$, is not valid. Though it follows from (1) that $x \wedge \bar{x} < M$, $y \vee \bar{y} > M$, where the central point $M = 1/2 = 0.5$ of the section D is called a *median*. *Disjunctive and conjunctive normal forms* (DNFs and CNFs) are introduced same as in Boolean algebra (Kleene 1967) with the only exception that a conjunct (disjunct) may contain a variable and its negation. A *minimization technique* for an FLF given by its DNF is studied in (Kabecode 1981; Kandel 1986; Wielgus 2014). To compare two given FLFs, we can compute and compare their values on all the variants of ordering arguments and their negations; the corresponding table is called a *choice table*. In Table 1, a choice table of function $f_1 = x_1\bar{x}_2 \vee \bar{x}_1 x_2$ is shown. We denote areas covering the function domain D^n as $A_i, 1 \le i \le L$, where L denotes the table's length. Fig. 1 shows the areas inside the unit square. In the general case, areas are formed as a result of hyperplanes $x_i = \bar{x}_j$ $(1 \le j \le n, 1 \le j \le n)$ intersection inside the unit hypercube.

Table 1: A choice table of function f_1

Number of area	Specification of area	Value of function f_1
1	$x_1 \le x_2 \le \bar{x}_2 \le \bar{x}_1$	x_2
2	$x_1 \le \bar{x}_2 \le x_2 \le \bar{x}_1$	x_2
3	$\bar{x}_1 \le x_2 \le \bar{x}_2 \le x_1$	$\bar{x}_2$
4	$\bar{x}_1 \le \bar{x}_2 \le x_2 \le x_1$	$\bar{x}_2$
5	$x_2 \le x_1 \le \bar{x}_1 \le \bar{x}_2$	x_1
6	$x_2 \le \bar{x}_1 \le x_1 \le \bar{x}_2$	x_1
7	$\bar{x}_2 \le x_1 \le \bar{x}_1 \le x_2$	$\bar{x}_1$
8	$\bar{x}_2 \le \bar{x}_1 \le x_1 \le x_2$	$\bar{x}_1$

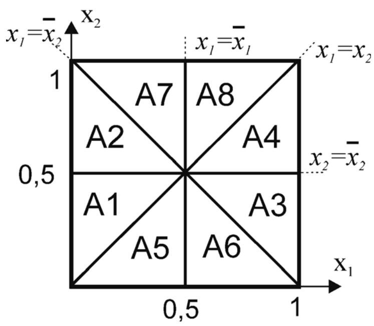


Figure 1: Areas of Table 1 inside the unit square (2-dimensional case)

In further notations, we omit symbol of variable "x" and consider vectors of indices. Denoting index of $\bar{x}_i$ as $-i$, we represent a sequence of indices, which specifies an area of the domain D^n, with a vector of 2n elements $\vec{\imath} = (i_1, i_2, \dots, i_{2n})$, $i_j \in \{-n, \dots, -1, 1, \dots, n\}$, $1 \le j \le 2n$. For brevity, we use negative indices i_j as well introducing the following notation:

$$x^i = \begin{cases} x_i, & i > 0 \\ \bar{x}_{|i|}, & i < 0. \end{cases} \quad (2)$$

Thus, a vector $\vec{\imath}$ represents a domain area specified with $x^{i_1} \le x^{i_2} \le \cdots \le x^{i_{2n}}$. A *choice table* T is a set of rows $T = \{t\}$, where the table row has the following form $t = (\vec{\imath}, a)$ which means that the function takes value x^a on the area specified by $\vec{\imath}$. When additional specifications are absent, we suppose that all the areas are listed in a choice table. Though studying a partially defined table could be useful, especially for the minimization purposes.

For example, the choice table, shown in Table 1, is specified as

$$T_1 = \{((1,2,-2,-1),2),((1,-2,2,-1),2), ((-1,2,-2,1),-2),((-1,-2,2,1),-2),((2,1,-1,-2),1), ((2,-1,1,-2),1),((-2,1,-1,2),-1),((-2,-1,1,2),-1)\}.$$

Note that a tuple $\vec{\imath}$ is symmetric with respect to its middle because $x \le y$ implies $\bar{y} \le \bar{x}$; thus, only one half of vector $\vec{\imath}$, for instance the first, can be actually stored. It means that the number of the choice tables of n arguments L(n) is defined by the number of permutations of n numbers multiplied by the number of variants for assigning signs to each of permutations; thus,

$$L(n) = n! \cdot 2^n. \quad (3)$$

Note that in examples we use functions of two arguments which CTs contain 8 rows; for 3 arguments the number of rows equals 48 and such examples are rather bulky; an example of synthesizing an FLF of 3 arguments is considered in Appendix. Asymptotically, the number of choice tables considerably exceeds the number of FLFs (Volgin and Levin 1990), which means that not any choice table defines an FLF. When the order of areas specified with vectors $\vec{\imath}$ is fixed, for instance same as in Table 1, we can represent a table as a vector of L(n) values $a_{\vec{\imath}}$. For example, the choice table $T_2 = (2,-1,2,-2,1,-1,-2,-1)$ does not define an FLF. However, each choice table can be partitioned in subdomains each of which defines an FLF (Volgin and Levin 1990). Note that not more than L(n) areas are required when specifying each function with a DNF consisting of a single argument (negation of an argument).

PARTITIONING A CHOICE TABLE WITH FUZZY LOGIC FUNCTIONS

We adjust the method of synthesis (Zaitsev et al 1998) to obtain a simple procedure for partitioning a given CT with a set of FLFs represented by DNF. The paper introduces a criterion when a given CT specifies an FLF based on a notion of *overlapping rows* of the table.

Here, we extended the technique (Zaitsev et al 1998) on partial tables and replace a formal application of the criterion by a simple iterative procedure which consists of the following steps:

1) Synthesize a DNF on a given table (part of table).
2) Build a CT on the obtained DNF and compare the function values.
3) If the function values coincide then the sought FLF is specified by the obtained DNF.
4) Otherwise proceed from the step 1) for the part of table which contains rows where the function values do not coincide (the "difference table").

In (Zaitsev et al 1998), it was proven that if a CT defines an FLF, then the FLF is represented with a DNF consisting of disjunctions of *constituents of maximum* on the rows of the CT, where a constituent of maximum for a table row is equal to conjunctions of variables starting from that which equals the function value to the last variable, inclusive. Using the CT notation from Section 2 for a table row $t = (\vec{1}, a)$, its constituents of maximum, denoted φ_t, is calculated as

$$\varphi_t = \bigwedge_{x^a \le x^i} x^i \,. \tag{4}$$

Then, the synthesized DNF has the following form

$$f = \bigvee_{t \in T} \varphi_t \,. \tag{5}$$

We leave the issues of formal minimization of FLFs (Kabecode 1981; Kandel 1986; Wielgus 2014) beyond the scope of the present paper only applying a simple DNF reduction with the tautology and absorption laws

$$x \vee x = x, x \vee (xy) = x.$$

Note that a constituent length does not exceed $2n$ and the maximal table length is $L(n)$ specified with (2). Thus, the algorithm complexity is $O(2n \cdot L)$. Though it is exponential in n, taking into consideration (2), its characterization as linear in the table length sounds more optimistic.

Let us consider examples of FLFs synthesis. First, we synthesize a function on the CT shown in Table 1. According to (5)

$$f_1 = \varphi_1 \vee \varphi_2 \vee \varphi_3 \vee \varphi_4 \vee \varphi_5 \vee \varphi_6 \vee \varphi_7 \vee \varphi_8 \,,$$

where according to (4)

$$\varphi_1 = x_2\bar{x}_2\bar{x}_1,\ \varphi_2 = x_2\bar{x}_1,\ \varphi_3 = \bar{x}_2x_1,\ \varphi_4 = \bar{x}_2x_2x_1,$$
$$\varphi_5 = x_1\bar{x}_1\bar{x}_2,\ \varphi_6 = x_1\bar{x}_2,\ \varphi_7 = \bar{x}_1x_2,\ \varphi_8 = \bar{x}_1x_1x_2.$$

Since φ_7 equals φ_2 , φ_6 equals φ_3, φ_2 absorbs φ_1 and φ_8, φ_3 absorbs φ_4 and φ_5, we obtain

$$f_1 = \varphi_2 \vee \varphi_3 = x_2\bar{x}_1 \vee \bar{x}_2x_1 \,,$$

that coincides with the initial expression which Table 1 has been constructed on.

Second, we synthesize an FLF (FLFs) on Table 2. We use a hint that the table defines a set of FLF introducing the following notation where the upper index specifies the function number in the source CT partitioning. We compose,

$$f_2^1 = \varphi_1 \vee \varphi_2 \vee \varphi_3 \vee \varphi_4 \vee \varphi_5 \vee \varphi_6 \vee \varphi_7 \vee \varphi_8 \,,$$

Where

$$\varphi_1 = x_2\bar{x}_2\bar{x}_1,\ \varphi_2 = \bar{x}_1,\ \varphi_3 = x_2\bar{x}_2x_1,$$

$$\varphi_4 = \bar{x}_2x_2x_1, \varphi_5 = x_1\bar{x}_1\bar{x}_2,\ \varphi_6 = \bar{x}_1x_1\bar{x}_2,$$
$$\varphi_7 = \bar{x}_2\, x_1\bar{x}_1x_2,\ \varphi_8 = \bar{x}_1x_1x_2.$$

Since φ_3 equals φ_4 , φ_5 equals φ_6, φ_2 absorbs φ_1, φ_5, φ_7, φ_8, we obtain

$$f_2^1 = \varphi_2 \vee \varphi_3 = \bar{x}_1 \vee x_2\bar{x}_2x_1 \,.$$

However, a CT of f_2^1 coincides with the source CT of function f_2 only on the areas A_2, A_3, A_4, A_6, A_8.

For the rest of the table, where the values of f_2^1 do not coincide with the corresponding values of f_2, we proceed with the same procedure composing

$$f_2^2 = \varphi_1 \vee \varphi_5 \vee \varphi_7.$$

Since φ_1 equals φ_7, we obtain

$$f_2^2 = \varphi_1 \vee \varphi_5 = x_2\bar{x}_2\bar{x}_1 \vee x_1\bar{x}_1\bar{x}_2.$$

Thus, we specify function f_2, given by CT shown in Table 2, as

$$f_2 = \begin{cases} f_2^1 = \bar{x}_1 \vee x_2\bar{x}_2x_1, & \vec{x} \in A_2, A_3, A_4, A_6, A_8, \\ f_2^2 = x_2\bar{x}_2\bar{x}_1 \vee x_1\bar{x}_1\bar{x}_2, & \vec{x} \in A_1, A_5, A_7. \end{cases}$$

A TOOLBOX FOR SYNTHESIS OF FUZZY LOGIC FUNCTIONS ON A CHOICE TABLE

Using the technique described in Section 3, a toolbox fzy_syn has been implemented in C language and placed for free download on GitHub. It uses the following abbreviations: “fzy” denotes fuzzy logic, “tab” denotes a choice table (either complete or partial), “dnf” denotes a DNF, “syn” denotes synthesis, “cmp” denotes comparison.

The toolbox contains the following command line tools:

- fzy_tab_syn partitions a given choice table with a series of synthesized fuzzy logic functions (DNFs);
- fzy_tab_dnf synthesizes a fuzzy logic function (DNF) on a given choice table;
- fzy_dnf_tab builds a choice table on a given fuzzy logic function (DNF);
- fzy_cmp_tab_dnf compares a choice table with a fuzzy logic function (DNF);
- fzy_gen_tab generates a random choice table for a given number of variables.

fzy_tab_syn implements a repeated combination of fzy_tab_dnf and fzy_cmp_tab_dnf. In the general case, a DNF synthesized with fzy_tab_dnf satisfies only a part of the CT. Using fzy_cmp_tab_dnf, the CT is partitioned into two tables: the first table contains rows where DNF coincides with the source table; the second table contains rows where DNF does not coincide with the source table. Then the process is repeated with the difference table until it will be empty.

Table 2: A choice table of function f_2

Number	Area	f_2	f_2^1	f_2^2
1	$x_1 \le x_2 \le \bar{x}_2 \le \bar{x}_1$	x_2		x_2
2	$x_1 \le \bar{x}_2 \le x_2 \le \bar{x}_1$	$\bar{x}_1$	$\bar{x}_1$	
3	$\bar{x}_1 \le x_2 \le \bar{x}_2 \le x_1$	x_2	x_2	
4	$\bar{x}_1 \le \bar{x}_2 \le x_2 \le x_1$	$\bar{x}_2$	$\bar{x}_2$	
5	$x_2 \le x_1 \le \bar{x}_1 \le \bar{x}_2$	x_1		x_1
6	$x_2 \le \bar{x}_1 \le x_1 \le \bar{x}_2$	$\bar{x}_1$	$\bar{x}_1$	
7	$\bar{x}_2 \le x_1 \le \bar{x}_1 \le x_2$	$\bar{x}_2$		$\bar{x}_2$
8	$\bar{x}_2 \le \bar{x}_1 \le x_1 \le x_2$	$\bar{x}_1$	$\bar{x}_1$	

On a given CT, fzy_tab_dnf builds a DNF. The CT can be either complete or partial. A DNF is reduced with the application of the tautology and absorption laws. It is possible that the obtained DNF be valid for a part of the table only. The corresponding test can be implemented with fzy_cmp_tab_dnf.

On a given DNF of an FLF, fzy_dnf_tab builds a complete CT. The CT contains all $L(n)$ areas specified with (3), where n is the number of FLF arguments. When enumerating the areas, it proceeds first with ascending order of permutations of n arguments, and second with all the combinations of arguments with and without negation; for a definite permutation, the ascending order of enumerating the negations corresponds to treating an argument without negation as a zero and an argument with negation as a unit. Thus, we enumerate n! permutations and for each permutation 2^n combinations of negations.

On the areas of a given CT, fzy_cmp_tab_dnf compares the function values with the values computed according to a given DNF. The CT can be either complete or partial. Two tables are written: the first table for the same values of function and the second table for different values of function taken from the source table.

fzy_gen_tab builds a complete CT with random values of function equal to arguments and negations of arguments. The CT specifies a complete function and contains all $L(n)$ areas specified with (3), where n is the number of FLF arguments.

The toolbox implements a command line interface with input and output data located in textual files having special format; who types of files are supported: a file of CT and a file of a DNF of an FLF.

Command lines to launch the tools, has the following format:

```
>fzy_tab_syn tab_file result_files_prefix
>fzy_tab_dnf tab_file dnf_file
>fzy_dnf_tab dnf_file tab_file
>fzy_cmp_tab_dnf tab_file dnf_file
        comm_tab_file diff_tab_file
>fzy_gen_tab n tab_file [ rand_seed ]
```

The command line parameters are specified as follows:

- n is the number of FLF arguments;
- tab_file is a file which contains a CT either complete or partial;
- dnf_file is a file which contains a DNF of an FLF;
- comm_tab_file contains rows of the source table where values of the functions coincide;
- diff_tab_file contains rows of the source table where values of the functions are different;
- result_files_prefix is a prefix to create file names of subdomains using the following suffices: "-tab-<i>" – a subdomain, "-dnf-<i>" – its DNF.
- rand_seed is an integer which is used as a seed to generate random values; on default, the system time is used as a seed.

Let us specify formats of files. A choice table file has the following format:

```
FZYTAB n
Ai    vi
...
FZYTABEND [(L rows)]
```

Where:

- "FZYTAB" is a label of the table file beginning;
- n is the number of arguments;
- L is the number of table rows;
- "FZYTABEND" is a label of the table file end;
- A_i is the current area ($1 \le i \le L$);
- v_i is the function value on the current area.

The current area A_i is specified as follows:

A_i : : j_1 j_2 ... j_{2n}

where the values range is

$$-n \le j_k < 0, 0 < j_k \le 2n, 1 \le k \le 2n.$$

The term j_k defines k-th item in the list A_i which equals to x_{j_k} if $j_k > 0$ and equals to the negation $\bar{x}_{|j_k|}$ if $j_k < 0$.

An example of the choice table file (fl_tab) for the function f_1 (according to Table 1) follows:

```
FZYTAB 2
 1  2 -2 -1      2
 1 -2  2 -1      2
-1  2 -2  1     -2
-1 -2  2  1     -2
 2  1 -1 -2      1
 2 -1  1 -2      1
-2  1 -1  2     -1
-2 -1  1  2     -1
FZYTABEND (8 rows)
```

A DNF file has the following format:

```
FZYDNF n
   Ci
...
FZYDNFEND [(L conjuncts)]
```

Where:

- "FZYDNF" is a label of the DNF file beginning;
- n is the number of arguments;
- L is the number of conjuncts;
- "FZYDNFEND" is a label of the DNF file end;
- C_i is the current conjunct ($1 \le i \le L$).

The current conjunct C_i is specified as follows:

$C_i :: \; j_1 \;\; j_2 \quad \dots \quad j_{m_i}$

where the values range is

$$-n \le j_k < 0, 0 < j_k \le 2n, 1 \le k \le m_i;$$

m_i is the number of items (variables and negations of variables) in the current conjunct C_i; the term j_k defines k-th item in the list C_i which equals to x_{j_k} if $j_k > 0$ and equals to the negation $\bar{x}_{|j_k|}$ if $j_k < 0$.
An example of the DNF file (f1_dnf) for the function f_1 follows:

```
FZYDNF 2
1 -2
-1 2
FZYDNFEND (2 conjuncts)
```

RUNNING TOOLS AND ANALYSING RESULTS

Let us consider examples of command lines. Suppose the source data file f1_dnf specifying DNF of f_1 has been created by hand in a text editor. The command line

```
>fzy_dnf_tab f1_dnf f1_tab
```

creates a CT of f_1 on its DNF file f1_dnf and saves the resulting CT in file f1_tab. We can check visually that it coincides with the table file shown above.
Suppose the source data file f2_tab specifying CT of f_2 has been created by hand in a text editor. The command line

```
>fzy_tab_dnf f2_tab f2_1_dnf
```

synthesizes a DNF file f2_1_dnf on the CT file f2_tab. The obtained DNF file f2_1_dnf follows:

```
FZYDNF 2
-1
2 -2 1
FZYDNFEND (2 conjuncts)
```

The command line

```
>fzy_cmp_tab_dnf f2_tab f2_1_dnf
              f2_tab_comm f2_tab_diff
```

compares a function given by CT file f2_tab with an FLF given by DNF file f2_1_dnf, writes the coinciding rows to the file f2_tab_comm, and writes the different rows to the file f2_tab_diff with the function values according to the source CT file (f2_tab). The obtained difference file f2_tab_diff follows:

```
FZYTAB 2
 1  2 -2 -1      2
 2  1 -1 -2      1
-2  1 -1  2     -2
FZYTABEND (3 rows)
```

It is not empty; consequently f_2 is not an FLF. The obtained coinciding part of the CT represented with the file f2_tab_comm follows:

```
FZYTAB 2
 1 -2  2 -1     -1
-1  2 -2  1      2
-1 -2  2  1     -2
 2 -1  1 -2     -1
-2 -1  1  2     -1
FZYTABEND (5 rows)
```

In fact it contains the rows where the function f_2^1 specified with f2_1_dnf represents the source function f_2. Then, we can synthesize a DNF for the difference file with

```
>fzy_tab_dnf f2_tab_diff f2_2_dnf
```

to obtain the second DNF file f2_2_dnf to cover the source CT of f_2:

```
FZYDNF 2
2 -2 -1
1 -1 -2
FZYDNFEND (2 conjuncts)
```

However, it is more convenient to use fzy_tab_syn which partitions the source CT automatically. The command line

```
>fzy_tab_syn f2_tab f2_syn
```

partitions the source CT f2_tab into subdomains and synthesizes an FLF represented with a DNF for each subdomain. The following files are created: f2_syn-tab-0, f2_syn-dnf-0 and f2_syn-tab-1, f2_syn-dnf-1 specifying the obtained partitioning. File f2_syn-tab-0 coincides with the above f2_tab_comm and f2_syn-dnf-0 coincides with the above f2_1_dnf. File f2_syn-tab-1 coincides with the above f2_tab_diff and f2_syn-dnf-1 coincides with the above f2_2_dnf.
To try the toolbox on big random data we use fzy_gen_tab to create random CTs of specified number of arguments. The command line

```
>fzy_gen_tab 6 F3_rand_tab
```

creates a CT of a function of 6 arguments with random values.

CONCLUSIONS

A toolbox for synthesis of fuzzy logic functions on a choice table has been implemented in C language; it is available for free download on GitHub. The method

described in (Zaitsev et al 1998) has been adjusted for fast partitioning the source choice table with a set of fuzzy logic functions. No formal minimization of fuzzy logic functions has been implemented though the DNF transformation using the tautology and absorption laws allows its considerable reduction. Formal aspects of optimality (minimalism) are a direction for future research.
The toolbox implements a command line style of programming using data located in textual files of simple intuitive formats. Though it lacks graphical user interface, the toolbox can process rather big data fast. One more benefit is possibility of its easy integration into fuzzy logic frameworks. Thus the toolbox can be used as a synthesis engine to develop graphical systems for fuzzy (control) systems design.

REFERENCES

Kabekode, V.S.B. 1981. “Minimization of disjunctive normal forms of fuzzy logic functions”. *Journal of the Franklin Institute*, 311(3), 171-185.

Kandel, A. 1986. *Fuzzy Mathematical Techniques with Applications*. Addison-Wesley.

Kaufmann, A. 1977. *Introduction a la theorie des sous-ensembles flous*. Masson.

Kleene, S.C. 1967. *Mathematical Logic*. Wiley.

Novak, V,; I. Perfilieva; and A. Dvorak. 2016. *Insight into Fuzzy Modeling*. Wiley.

Shmeleva, T.R.; Zaitsev, D.A.; and Zaitsev, I.D. 2009. “Verification of square communication grid protocols via infinite Petri nets”. In *Proceedings of 10th Middle Eastern Simulation Multiconference*, 53-59.

Volgin, L.I. and V.I. Levin. 1990. *Continuous Logic. Theory and Applications*. Tallinn, AS of Estonia.

Wielgus, A. 2014. "Heuristic algorithm of two-level minimization of fuzzy logic functions", In *Proceedings of 17th International Symposium on Design and Diagnostics of Electronic Circuits & Systems*, 302-305.

Wielgus, A. 2004. "Algorithms of multilevel synthesis of fuzzy functions", In *Proceedings of IEEE International Conference on Fuzzy Systems*, 929-934.

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

Zaitsev, D.A.; T.R. Shmeleva; W. Retschitzegger; and B. Proell. 2016. “Security of grid structures under disguised traffic attacks”. *Cluster Computing*, 19(3), 1183-1200.

Zaitsev, D.A. and J. Jurjens. 2016. “Programming in the Sleptsov net language for systems contro”. *Advances in Mechanical Engineering*, 8(4), 1–11.

Zaitsev, D.A. and T.R. Shmeleva. 2006. “Switched Ethernet Response Time Evaluation via Colored Petri Net Model”. In *Proceedings of International Middle Eastern Multiconference on Simulation and Modelling*, Alexandria, Egypt. 68-77.

Zaitsev, D.A.; V.G. Sarbei; anf A.I. Sleptsov. 1998. “Synthesis of continuous-valued logic functions defined in tabular form”. *Cybernetics and Systems Analysis*, 34(2), 1998, 190–195.

Zaitsev, D.A. and A.I. Sleptsov. 1997. “State equations and equivalent transformations for timed Petri nets”. *Cybernetics and Systems Analysis*, 33(5), 659-672

WEB REFERENCES

github.com/dazeorgacm/fzy_syn

APPENDIX: AN EXAMPLE OF SYNTHESIS OF AN FLF OF 3 ARGUMENTS

A given CT file (f3_tab):

```
 FZYTAB 3
 1  2  3 -3 -2 -1    -2
 1  2 -3  3 -2 -1    -2
 1 -2  3 -3  2 -1    -3
 1 -2 -3  3  2 -1     3
-1  2  3 -3 -2  1    -3
-1  2 -3  3 -2  1     3
-1 -2  3 -3  2  1     2
-1 -2 -3  3  2  1     2
 1  3  2 -2 -3 -1    -3
 1  3 -2  2 -3 -1    -3
 1 -3  2 -2  3 -1    -2
 1 -3 -2  2  3 -1     2
-1  3  2 -2 -3  1    -2
-1  3 -2  2 -3  1     2
-1 -3  2 -2  3  1     3
-1 -3 -2  2  3  1     3
 2  1  3 -3 -1 -2    -1
 2  1 -3  3 -1 -2    -1
 2 -1  3 -3  1 -2    -3
 2 -1 -3  3  1 -2     3
-2  1  3 -3 -1  2    -3
-2  1 -3  3 -1  2     3
-2 -1  3 -3  1  2     1
-2 -1 -3  3  1  2     1
 2  3  1 -1 -3 -2    -3
 2  3 -1  1 -3 -2    -3
 2 -3  1 -1  3 -2    -1
 2 -3 -1  1  3 -2     1
-2  3  1 -1 -3  2    -1
-2  3 -1  1 -3  2     1
-2 -3  1 -1  3  2     3
-2 -3 -1  1  3  2     3
 3  1  2 -2 -1 -3    -1
 3  1 -2  2 -1 -3    -1
 3 -1  2 -2  1 -3    -2
 3 -1 -2  2  1 -3     2
-3  1  2 -2 -1  3    -2
-3  1 -2  2 -1  3     2
-3 -1  2 -2  1  3     1
-3 -1 -2  2  1  3     1
 3  2  1 -1 -2 -3    -2
 3  2 -1  1 -2 -3    -2
 3 -2  1 -1  2 -3    -1
 3 -2 -1  1  2 -3     1
-3  2  1 -1 -2  3    -1
-3  2 -1  1 -2  3     1
-3 -2  1 -1  2  3     2
-3 -2 -1  1  2  3     2
FZYTABEND (48 rows)
```

A command line:

```
>fzy_tab_dnf f3_tab f3_dnf
```

The obtained DNF file (f3_dnf):

```
FZYDNF 3
-2 -1
2 1
-3 -1
3 1
-3 -2
3 2
FZYDNFEND (6 conjuncts)
```

The obtained FLF:

$$f_3 = \bar{x}_2\bar{x}_1 \vee x_2x_1 \vee \bar{x}_3\bar{x}_1 \vee x_3x_1 \vee \bar{x}_3\bar{x}_2 \vee x_3x_2 \,.$$

MODELLING MULTI-COMPONENT PREDICTIVE SYSTEMS AS PETRI NETS

Manuel Martin Salvador, Marcin Budka, Bogdan Gabrys
Bournemouth University
United Kingdom
E-mails: {msalvador,mbudka,bgabrys}@bournemouth.ac.uk

KEYWORDS

Predictive Systems, Petri nets, Process industry

ABSTRACT

Building reliable data-driven predictive systems requires a considerable amount of human effort, especially in the data preparation and cleaning phase. In many application domains, multiple preprocessing steps need to be applied in sequence, constituting a 'workflow' and facilitating reproducibility. The concatenation of such workflow with a predictive model forms a Multi-Component Predictive System (MCPS). Automatic MCPS composition can speed up this process by taking the human out of the loop, at the cost of model transparency (i.e. not being comprehensible by human experts). In this paper, we adopt and suitably re-define the Well-handled with Regular Iterations Work Flow (WRI-WF) Petri nets to represent MCPSs. The use of such WRI-WF nets helps to increase the transparency of MCPSs required in industrial applications and make it possible to automatically verify the composed workflows. We also present our experience and results of applying this representation to model soft sensors in chemical production plants.

INTRODUCTION

In many data mining problems one needs to sequentially apply multiple preprocessing methods to the data (e.g. outlier detection → missing value imputation → dimensionality reduction), effectively forming a preprocessing chain. Such data-driven workflows have been used to guide data processing in a variety of fields. Some examples are astronomy (Berriman et al. 2007), biology (Shade and Teal 2015), clinical research (Teichmann et al. 2010), archive scanning (Messaoud et al. 2011), telecommunications (Maedche et al. 2000), banking (Wei et al. 2013) and process industry (Budka et al. 2014) to name a few. The common methodology in all these fields consists of following a number of steps to prepare a dataset for mining. In the field of predictive modelling, the workflow resulting from connecting different methods is known as a Multi-Component Predictive System (MCPS) (Tsakonas and Gabrys 2012). At the moment, tools like WEKA (Frank et al. 2016), RapidMiner (Hofmann and Klinkenberg 2016) or Knime (Berthold et al. 2008) allow to create and run MCPSs including a large variety of operators. Each of these tools however uses a different representation of workflows. One of the goals of this paper is to establish a common framework for connecting multiple data processing components into workflows.

ALTERNATIVES TO REPRESENT MCPS

To formalise the notion of MCPSs under a common abstract framework, various approaches can be considered:

1. **Function composition**, where each component is a function $f : \mathcal{X} \rightarrow \mathcal{Y}$ that makes an operation over an input tensor $\boldsymbol{x}$ (i.e. a multidimensional array made of continuous or categorical values) and returns an output tensor $\boldsymbol{y}$. Several components can be connected by composing functions i.e. $f(g(\boldsymbol{x}))$. However, this notation can become tedious when representing complex workflows involving multiple components with parallel paths of different lengths. Moreover, it is not expressive enough to represent different states of a concurrent system.

2. **Directed Acyclic Graphs (DAG)**, consisting of a set of components (nodes) F and directed arcs A connecting pairs of components: $G = (F, A)$. This approach is very flexible and makes it easy to understand the structure of the MCPS. However, DAGs on their own are not expressive enough to model system execution (e.g. how data is transformed, what metadata is generated, iterations, and preconditions) or temporal behaviour (e.g. duration and delays).

3. **Petri nets (PN)**, which are a modelling tool applicable to many types of systems (Petri 1962). A Petri net $PN = (P, T, F)$ is a directed bipartite graph consisting of a set of places P and transitions T connected by arcs F. Depending on the system, places and transitions can be interpreted in different ways. In this paper, places are considered to be data buffers and transitions are data processing methods. PNs have been shown to be very useful to model workflows (van der Aalst 1998a) since they are very flexible, can accommodate complex process logic including concurrency and have a strong mathematical foundation (Murata 1989). Using workflow

algebra (Pankratius and Stucky 2005) one can modify and create PNs with relational operators like *selection* or *union*, which can be useful to adapt workflows. Analysis methods like (van der Aalst 2000) inspect PN structure to find potential design errors. An important advantage is that the graphical nature of PNs makes them intuitive and easy to understand for any domain expert. PNs are also vendor independent and, once composed, can be easily translated to any data mining tool. This approach has not been considered before to model MCPSs and it is proposed in this paper for the first time.

BACKGROUND

While Petri nets were introduced in (Petri 1962), the most recent definition of a Petri net, which has been adopted in this paper, was given in (Murata 1989) as the following tuple

$$PN = (P, T, F, W, M_0) \tag{1}$$

where $P = \{p_1, ..., p_m\}$ is a finite set of places, $T = \{t_1, ..., t_n\}$ is a finite set of transitions, $F \in (P \times T) \cup (T \times P)$ is a set of arcs, $W : F \to \mathbb{N}^+$ is a weight function, $M_0 : P \to \mathbb{N}$ is the initial marking (i.e. state of the net – number of tokens in each place). Additionally, a Petri net contains one or more tokens that represent units of the system to be processed. The lifetime of a PN is defined by a set of states $\mathcal{M} = \{M_0, ..., M_q\}$. Each state is the distribution of the tokens over P.
In the remainder of this section, we define a number of properties of PNs, which enable us to apply this formalism to defining the MCPSs.
In a Petri net, nodes are connected by arcs forming paths. Formally,

Definition 1 (Path) *A path C from a node $n_1 \in P \cup T$ to a node $n_k \in P \cup T$ is a sequence of nodes $\langle n_1 n_2 ... n_k \rangle$ such that $f_{i,i+1} \in F \mid 1 \leq i \leq k-1$. A path C is elementary iff, for any two nodes n_i and n_j on C, $i \neq j \Rightarrow n_i \neq n_j$.*

The behaviour of a Petri net is described by firing of transitions. A transition can be fired (i.e. activated) when each of its input places are marked at least with the number of tokens indicated by the value of the function $w(p, t)$ associated with the arc $p \to t$ (i.e. minimum number of tokens needed in p to fire the transition t). When the active transitions t are fired, the state of the net changes from M_n to M_{n+1} and tokens are transferred from input to output places for each transition.
The nodes of a Petri net (i.e. places and transitions) can have multiple output and input arcs. Depending on the behaviour of the node, there are four main constructs:

- **AND-split**: A token is produced for each of the output arcs.
- **XOR-split**: A token is produced for only one of the output arcs.
- **AND-join**: A token is received for each of the input arcs.
- **XOR-join**: A token is received for only one of the input arcs.

The input nodes of a node $n \in P \cup T$ are denoted as $\bullet n$, while output ones are $n \bullet$.
Despite of the fact that the classical Petri nets can cover a large range of applications (e.g. manufacturing (DiCesare et al. 1993), business process management (van der Aalst et al. 2000), hardware design (Yakovlev et al. 2000), molecular biology systems (Hardy and Robillard 2004), there are sometimes circumstances when new properties have to be defined in order to cover additional types of systems. For example, (van der Aalst 1998a) presented a new type of Petri net to represent business process logic called WorkFlow net (WF-net). WF-nets are sequential workflows with a single starting point and ending point. The simplest WF-net is shown in Figure 1, where transitions represent tasks, places could be seen as conditions, and tokens are cases (e.g. a patient, a document or a picture). Formally:

Definition 2 (WF-net) *A Petri net is a WF-net iff:*

a) there is only one source place $i \in P \mid \bullet i = \emptyset$;

b) there is only one sink place $o \in P \mid o \bullet = \emptyset$;

c) every node $n \in P \cup T$ is on a path from i to o.

The third point of this definition entails that if a new transition t connecting o with i is added, then the resulting Petri net is strongly connected.

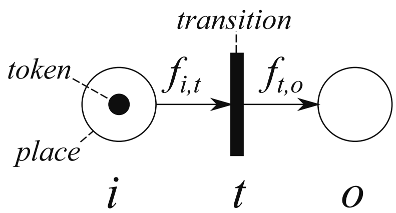


Figure 1: WorkFlow Net with a Single Transition

Definition 3 (Strong connectivity) *A Petri net is strongly connected iff for every pair of nodes (i.e. places and transitions) x and y, there is a path leading from x to y.*

The **soundness** property for WF-nets introduced by (van der Aalst 1997) implies that if a net has k tokens in the input place during the initial marking, it will have k tokens in the output place at the final marking (i.e. $M_0(i) = M_q(o)$). A WF-net with soundness is

guaranteed to terminate (i.e. it does not have deadlocks or livelocks).
In order to avoid deadlocks in WF-nets, (van der Aalst 2000) introduced the well-handled property that ensures the lack of bad constructions (e.g. XOR-split followed by AND-join block). Formally:

Definition 4 (Well-handledness) *A Petri net is well-handled iff, for any pair of nodes* $\{n_i, n_j\}$ *such that one is a place and the other a transition, and for any pair of elementary paths* C_1 *and* C_2 *leading from* n_i *to* n_j, $C_1 \cap C_2 = \{n_i, n_j\} \Rightarrow C_1 = C_2$.

Petri nets can become very large when defining complex processes (van der Aalst 1998b). To facilitate the representation, **hierarchical** Petri nets were introduced as an extension of PNs where a transition can be represented by another PN (called subnet) – see Figure 2. The action of replacing a transition by a subnet is called an iteration. Iterations are denoted as regular when the subnet has entrance and exit nodes acting as dummy nodes. This concept leads to a new type of Petri nets known as WRI-WF nets (Well-handled with Regular Iterations WF-net) presented in (Ping et al. 2004). Formally:

Definition 5 (WRI-WF net) *A Petri net is a WRI-WF net iff:*

a) the PN is a WF-net (see Definition 2);

b) the PN is well-handled (see Definition 4);

c) the PN is acyclic;

d) the iterations of the PN are regular.

WRI-WF nets are inherently sound (see (Ping et al. 2004) for proof).

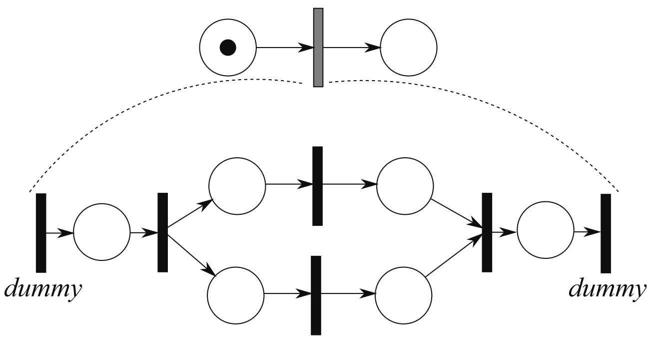


Figure 2: Hierarchical WF-net with Parallel Paths

MCPS DEFINITION

We propose to use WRI-WF-nets as the base for defining MCPSs. However, to comply with the predictive nature of MCPSs there are some additional restrictions that have to be added. Formally:

Definition 6 (MCPS) *A Petri net is an MCPS iff all the following conditions apply:*

a) the PN is WRI-WF-net (see Definition 5);

b) each place $p \in P \setminus \{i, o\}$ *has only a single input and a single output;*

c) the PN is 1-bounded, that is, there is a maximum of one token in each $p \in P$ *for every reachable state (i.e.* $M(p) \leq 1$*).;*

d) the PN is 1-sound (i.e. $M_0(i) = M_q(o) = 1$*);*

e) the PN is ordinary (i.e. $w = 1\ \forall w \in W$*);*

f) all the transitions $t \in T$ *with multiple inputs or outputs are AND-join or AND-split, respectively;*

g) any token is a tensor (i.e. multidimensional array).

In an MCPS, an atomic transition $t \in T$ is an algorithm with a set of hyperparameters λ that affect how the token is processed. An MCPS can be as simple as the one shown in Figure 1 with a single transition. For example, the token in i can be a set of unlabelled instances, and t a classifier which consumes such token from the arc $f_{i,t}$ and generates one token in o with the predicted labels through $f_{t,o}$.
An MCPS can however be hierarchically extended since each transition t can be either atomic or special (a subnet with additional starting and ending dummy transitions) – see Figure 2 where atomic transitions are black and special transitions are grey. As a consequence, an MCPS can model very complex systems with multiple data transformations and parallel paths (see e.g. Figure 5 for a multi-hierarchy example).
In predictive modelling, the semantics for transitions are: (1) preprocessing methods, (2) predictors, (3) ensembles and (4) postprocessing methods. Transitions representing (1), (2), and (4) can be either atomic or special. However, type (3) transitions are necessarily special since ensembles are made of several predictors and a combination method (e.g. voting).
Depending on the number of inputs and outputs, MCPSs can have any of the following types of transitions (see Figure 3):

- $1 \rightarrow 1$ transitions (e.g. a classifier that consumes unlabelled instances and returns predictions)
- $1 \rightarrow n$ transitions (e.g. a random subsampling method that consumes a set of instances and returns several subsets of data)
- $n \rightarrow 1$ transitions (e.g. a voting classifier that consumes multiple predictions per instance and returns a single prediction per instance)

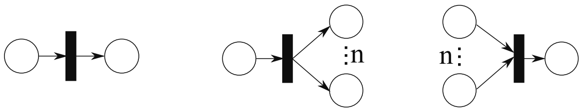


Figure 3: Types of transitions according to the number of inputs and outputs

APPLICATION IN PROCESS INDUSTRY

Processing plants have a large number of sensors that measure physical properties in different parts of the process. Values such as temperatures, pressures or humidity are easy to capture. However, acquiring other measurements is more expensive and often require human interaction. For instance, measuring the product concentration may require taking a sample and analysing it in the laboratory.

In order to improve production efficiency, a predictive model could deliver estimates of such hard-to-measure values based on the process state given by the easy-to-measure values from the sensors. This type of predictive models are called *soft sensors* because they can be seen as software or virtual sensors instead of physical.

Sometimes the first-principle models, that are based on the physical and chemical process knowledge, are available. Although such models are preferred by practitioners (De Assis and Maciel Filho 2000, Prasad et al. 2002), they are primarily meant for planning and design of the processing plants, and therefore usually focus on the steady states of the process (Chéruy 1997). Thus, such models can seldom be used in practice in a wide range of operating conditions. Moreover, often the process knowledge for modelling is not available at all. In such cases data-driven models fill the gap and often play an important role for the operation of the processes as they can extract the process knowledge automatically from the provided data. A review of data-driven soft sensors in the process industry is presented in (Kadlec et al. 2009). The most popular methods are multivariate statistical techniques like Principal Component Analysis (PCA) in a combination with a regression model (PCR) (Jolliffe 2002), and Partial Least Squares (PLS) (Wold et al. 2001). Other common approaches like Multi-Layer Perceptron (MLP) (Qin 1997) and Radial Basis Function (RBF) (Wang et al. 2006) are based on neural networks. (Kadlec et al. 2009) show that there are indeed dozens of methods to build soft sensors, each of them with various tunable hyperparameters. There is however no single method that is universally superior across all the problems (Wolpert and Macready 1997).

Although the most common application of soft sensors is online prediction, others include process monitoring, fault detection and sensor backup. In any of them, the main requirements in the process industry are:

- reliability – to provide truthful results;
- robustness – to work under any circumstances or inconvenience; and
- transparency – to be comprehensible by human experts.

Within a joint research project with a large chemical manufacturer we have worked on the development of data-driven soft-sensors. As a consequence, we developed guidelines for building MCPSs made of data preprocessing methods and predictive models (Budka et al. 2014). Nevertheless, designing and optimising a predictive system to work in a real environment requires a considerable human effort – e.g. (Martin Salvador et al. 2014). Reducing these labour-intensive tasks led us to investigate the automation of composing of MCPS (Martin Salvador et al. 2016b). However, the black-box nature of the automatic process reduces the transparency requirement from the industry. Representing MCPSs as Petri nets helped us to address such requirement.

In order to automatically compose and optimise MCPSs for a given classification or regression problem, we have adopted Auto-WEKA (Thornton et al. 2013), a tool for combined model selection and hyperparameter optimization. We have extended this software to generate a Petri net of the resultant MCPS in PNML (Petri Net Markup Language) format. This PN can then be analysed in any tool supporting this standard language (e.g. WoPeD (Freytag and Sänger 2014)). The software and source code are publicly available in our repository: `https://github.com/dsibournemouth/autoweka`.

EXPERIMENTAL RESULTS

We have carried out an experimental analysis of automatically composing and optimising MCPSs for 7 datasets representing regression tasks of real chemical processes. Four of these datasets have been made available by Evonik Industries as part of the collaboration within the INFER project (Musial et al. 2013), and have been extensively used in previous studies (Kadlec and Gabrys 2009, Budka et al. 2014, Bakirov et al. 2015):

- 'absorber' dataset which contains 38 continuous attributes from an absorption process. No additional information has been provided apart from this being a regression task;
- 'drier' dataset, which consists of 19 continuous features from physical sensors (i.e. temperature, pressure and humidity) and the target value is the residual humidity of the process product (Kadlec and Gabrys 2009);
- 'oxeno' dataset, which contains 71 continuous attributes also from physical sensors and a target vari-

able which is the product concentration measured in the laboratory (Budka et al. 2014); and

- 'thermalox' dataset, which has 38 attributes from physical sensors and the two target values are concentrations of NO_x and SO_x in the exhaust gases (Kadlec and Gabrys 2009).

Due to confidentiality reasons the datasets listed above cannot be published. However, 3 additional publicly available datasets from the same domain have also been used in the experiments. These are:

- 'catalyst' dataset consisting of 14 attributes, where the task is to predict the activity of a catalyst in a multi-tube reactor (Kadlec and Gabrys 2011);
- 'debutanizer' dataset, which has 7 attributes (temperature, pressure and flow measurements of a debutanizer column) and where the target value is the concentration of butane at the output of the column (Fortuna et al. 2005); and
- the 'sulfur' recovery unit dataset, which is a system for removing environmental pollutants from acid gas streams before they are released into the atmosphere (Fortuna et al. 2003). The washed out gases are transformed into sulfur. The dataset has five input features (flow measurements) and two target values: concentration of H_2S and SO_2.

Each dataset has been split into 70% training and 30% testing sets, unless partition was already provided. Auto-WEKA was run in parallel with 25 different initial seeds and taking the Root Mean Squared Error (RMSE) as optimisation measure. The best MCPS configurations for each dataset (i.e. lowest holdout error) are shown in Table 1. Ensemble methods (Bagging and RandomSubSpace), which train multiple predictors with different subsets of data, have been found to provide the best performance for all analysed datasets. In fact, the solutions found outperform the four most popular methods for building soft sensors (PCR, PLS, MLP and RBF) in 6 out of 7 datasets (see δ in Table 1). None of the most popular techniques have been selected among the best MCPSs, indicating the potential disadvantage of human bias towards well-known methods.
As a result of our new extension, we have been able to generate Petri nets like the ones shown in Figures 4 and 5, in which transitions represent the WEKA methods to process the datasets 'debutanizer' and 'sulfur', respectively.

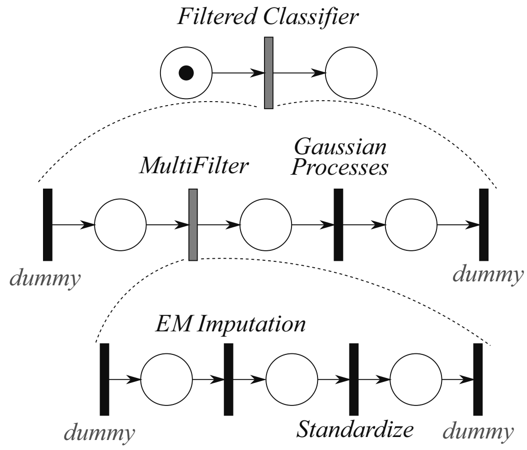


Figure 4: MCPS for 'debutanizer' Dataset

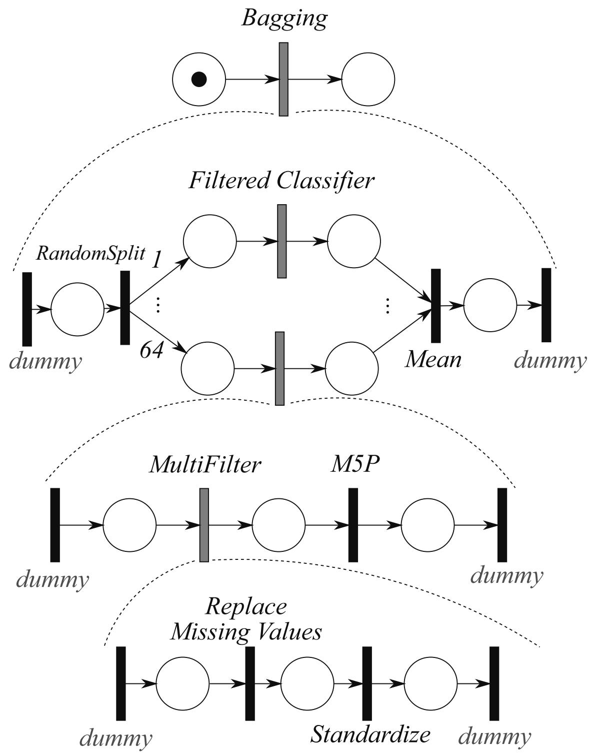


Figure 5: MCPS for 'sulfur' Dataset

CONCLUSION AND FUTURE WORK

In this paper we have proposed a novel definition of multi-component predictive systems (MCPSs) based on Petri nets. This vendor-independent formulation opens the door to formally verify that MCPSs are correctly composed (Sadiq et al. 2004), which is still an outstanding and non-trivial problem.

The experimental results show that it is feasible to automate the composition of MCPS for a real application such as process industry. The automatic generation of Petri nets helps to increase the transparency of the predictive system which a priori might be hidden under a black-box process. Moreover, since Petri nets are vendor independent, they can be easily translated to any data mining tool.

Petri nets are very useful to express semantic of complex

Table 1: Best MCPS for each dataset, holdout error $\mathcal{E}$ and difference with baseline δ ($\uparrow$ indicates an improvement). MV = missing value replacement, TR = transformation, DR = dimensionality reduction, SA = sampling.

dataset	**MV**	**TR**	**DR**	**SA**	**predictor**	**meta-predictor**	$\mathcal{E}$	δ
absorber	-	Wavelet	Rand.Subs.	-	KStar	Rand.SubSp.	0.8989	$\uparrow$ 0.0844
catalyst	Max	Normalize	-	-	GP	Bagging	0.0736	$\uparrow$ 0.1144
debutanizer	EM	Wavelet	-	-	IBk	Rand.SubSp.	0.1745	-0.0035
drier	EM	-	Rand.Subs.	Res.Samp.	M5P	Bagging	1.3744	$\uparrow$ 0.0573
oxeno	Zero	Normalize	-	-	M5P	Rand.SubSp.	0.0226	$\uparrow$ 0.0042
sulfur	Zero	Standardize	-	-	M5P	Bagging	0.0366	$\uparrow$ 0.0030
thermalox	Mean	Wavelet	-	-	GP	Rand.SubSp.	0.6904	$\uparrow$ 0.6170

systems. For example, we have also explored the use of coloured Petri nets to include an additional token representing meta-data when transitions of the MCPS need adaptation (Martin Salvador et al. 2016a). As future work, we would like to take advantage of workflow algebra (Pankratius and Stucky 2005) to model the adaptation of MCPSs. Moreover, it would be interesting to use Timed Petri nets to model task duration and delays.

REFERENCES

Bakirov R.; Gabrys B.; and Fay D., 2015. *On sequences of different adaptive mechanisms in non-stationary regression problems.* In *Proc. of the IJCNN.* 1–8.

Berriman G.B.; Deelman E.; Good J.; Jacob J.C.; Katz D.S.; Laity A.C.; Prince T.A.; Singh G.; and Su M.H., 2007. *Generating complex astronomy workflows*, Springer. 19–38.

Berthold M.; Cebron N.; Dill F.; and Gabriel T., 2008. *KNIME: The Konstanz information miner.* *Data Analysis, Machine Learning and Applications*, 1–8.

Budka M.; Eastwood M.; Gabrys B.; Kadlec P.; Martin Salvador M.; Schwan S.; Tsakonas A.; and Žliobaitė I., 2014. *From Sensor Readings to Predictions: On the Process of Developing Practical Soft Sensors.* In *Advances in IDA XIII.* vol. 8819, 49–60.

Chéruy A., 1997. *Software sensors in bioprocess engineering.* *Journal of Biotechnology*, 52, no. 3, 193–199.

De Assis A.J. and Maciel Filho R., 2000. *Soft sensors development for on-line bioreactor state estimation.* *Computers and Chemical Engineering*, 24, no. 2-7, 1099–1103.

DiCesare F.; Harhalakis G.; Proth J.M.; Silva M.; and Vernadat F.B., 1993. *Practice of Petri Nets in Manufacturing.* Springer.

Fortuna L.; Graziani S.; and Xibilia M.G., 2005. *Soft sensors for product quality monitoring in debutanizer distillation columns.* *Control Engineering Practice*, 13, 499–508.

Fortuna L.; Rizzo A.; Sinatra M.; and Xibilia M.G., 2003. *Soft analyzers for a sulfur recovery unit.* *Control Engineering Practice*, 11, 1491–1500.

Frank E.; Hall M.A.; and Witten I.H., 2016. *The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques".* Morgan Kaufmann.

Freytag T. and Sänger M., 2014. *WoPeD - An educational tool for workflow nets.* In *CEUR Workshop Proceedings.* vol. 1295, 31–35.

Hardy S. and Robillard P.N., 2004. *Modeling and simulation of molecular biology systems using petri nets: modeling goals of various approaches.* *Journal of Bioinformatics and Computational Biology*, 02, no. 04, 619–637.

Hofmann M. and Klinkenberg R., 2016. *RapidMiner: Data mining use cases and business analytics applications.* CRC Press.

Jolliffe I.T., 2002. *Principal Component Analysis, Second Edition.* Springer.

Kadlec P. and Gabrys B., 2009. *Architecture for development of adaptive on-line prediction models.* *Memetic Computing*, 1, no. 4, 241–269.

Kadlec P. and Gabrys B., 2011. *Local learning-based adaptive soft sensor for catalyst activation prediction.* *AIChE Journal*, 57, no. 5, 1288–1301.

Kadlec P.; Gabrys B.; and Strandt S., 2009. *Data-driven Soft Sensors in the process industry.* *Computers & Chemical Engineering*, 33, no. 4, 795–814.

Maedche A.; Hotho A.; and Wiese M., 2000. *Enhancing Preprocessing in Data-Intensive Domains using Online-Analytical Processing.* In *Proc. of DaWaK.* 258–264.

Martin Salvador M.; Budka M.; and Gabrys B., 2016a. *Effects of change propagation resulting from adaptive preprocessing in multicomponent predictive systems.* In *Proc. of KES.* vol. 96, 713–722.

Martin Salvador M.; Budka M.; and Gabrys B., 2016b. *Towards Automatic Composition of Multicomponent Predictive Systems*. In *Proc. of HAIS*. 27–39.

Martin Salvador M.; Gabrys B.; and Žliobaitė I., 2014. *Online Detection of Shutdown Periods in Chemical Plants: A Case Study*. In *Proc. of KES*. vol. 35, 580–588.

Messaoud I.; El Abed H.; Märgner V.; and Amiri H., 2011. *A design of a preprocessing framework for large database of historical documents*. In *Proc. of HIP*. 177–183.

Murata T., 1989. *Petri nets: Properties, analysis and applications*. 77, no. 4, 541–580.

Musial K.; Budka M.; and Blysz W., 2013. *Understanding the Other Side The Inside Story of the INFER Project*. *Smart Innovation, Systems and Technologies*, 18, 1–9.

Pankratius V. and Stucky W., 2005. *A Formal Foundation for Workflow Composition, Workflow View Definition, and Workflow Normalization based on Petri Nets*. In *Proc. of APCCM*. 79–88.

Petri C.A., 1962. *Kommunikation mit Automaten*. Ph.D. thesis, Universität Hamburg.

Ping L.; Hao H.; and Jian L., 2004. *On 1-soundness and Soundness of Workflow Nets*. In *Third Work. Model. Objects, Components, Agents*. 21–36.

Prasad V.; Schley M.; Russo L.P.; and Wayne Bequette B., 2002. *Product property and production rate control of styrene polymerization*. *Journal of Process Control*, 12, no. 3, 353–372.

Qin S., 1997. *Chapter 8 - Neural Networks for Intelligent Sensors and ControlPractical Issues and Some Solutions*. In *Neural Systems for Control*. 213–234.

Sadiq S.; Orlowska M.; Sadiq W.; and Foulger C., 2004. *Data Flow and Validation in Workflow Modelling*. In *Proc. of ADC*. vol. 27, 207–214.

Shade A. and Teal T.K., 2015. *Computing Workflows for Biologists: A Roadmap*. *PLoS Biol*, 13, no. 11, 1–10.

Teichmann E.; Demir E.; and Chaussalet T., 2010. *Data preparation for clinical data mining to identify patients at risk of readmission*. In *Proc. of CBMS*. 184–189.

Thornton C.; Hutter F.; Hoos H.H.; and Leyton-Brown K., 2013. *Auto-WEKA: combined selection and hyperparameter optimization of classification algorithms*. In *Proc. of SIGKDD*. 847–855.

Tsakonas A. and Gabrys B., 2012. *GRADIENT: Grammar-driven genetic programming framework for building multi-component, hierarchical predictive systems*. *Expert Systems with Applications*, 39, no. 18, 13253–13266.

van der Aalst W.M.P., 1997. *Verification of workflow nets*. In *App. and Theory of Petri Nets*. 407–426.

van der Aalst W.M.P., 1998a. *The Application of Petri Nets To Workflow Management*. *Journal of Circuits, Systems and Computers*, 08, no. 01, 21–66.

van der Aalst W.M.P., 1998b. *Three Good Reasons for Using a Petri-Net-Based Workflow Management System*. In *Information and Process Integration in Enterprises*. 161–182.

van der Aalst W.M.P., 2000. *Workflow Verification: Finding Control-Flow Errors Using Petri-Net-Based Techniques*. *Business Process Management*, 1806, 19–128.

van der Aalst W.M.P.; Desel J.; and Oberweis A. (Eds.), 2000. *Business Process Management*, *Lecture Notes in Computer Science*, vol. 1806. Springer.

Wang L.; Shao C.; Wang H.; and Wu H., 2006. *Radial Basis Function Neural Networks-Based Modeling of the Membrane Separation Process: Hydrogen Recovery from Refinery Gases*. *Journal of Natural Gas Chemistry*, 15, no. 3, 230–234.

Wei W.; Li J.; Cao L.; Ou Y.; and Chen J., 2013. *Effective detection of sophisticated online banking fraud on extremely imbalanced data*. *World Wide Web*, 16, no. 4, 449–475.

Wold S.; Sjöström M.; and Eriksson L., 2001. *PLS-regression: A basic tool of chemometrics*. *Chemometrics and Intelligent Laboratory Systems*, 58, no. 2, 109–130.

Wolpert D.H. and Macready W.G., 1997. *No free lunch theorems for optimization*. *IEEE Transactions on Evolutionary Computation*, 1, no. 1, 67–82.

Yakovlev A.; Gomes L.; and Lavagno L. (Eds.), 2000. *Hardware Design and Petri Nets*. Springer.

COMPUTING SPATIAL CHARGING NEEDS USING AN AGENT-BASED DEMAND MODEL

Daniel Krajzewicz, Michael Hardinghaus, Matthias Heinrichs, Sigrun Beige
Institute of Transport Research
German Aerospace Center
Rutherfordstr. 2
12489 Berlin, Germany
E-mail: {Daniel.Krajzewicz, Michael.Hardinghaus, Matthias.Heinrichs, Sigrun.Beige}@dlr.de

KEYWORDS
Charging stations placement, agent-based, demand model, microscopic simulation.

ABSTRACT

The buildup of electromobility necessitates strategies for deploying charging stations that match demand and allocate them in space. This paper presents a fine-grained approach for determining the charging needs and allocating the according charging infrastructure using the results of an agent-based traffic demand model. The usage of a fine-grained representation of each individual's daily trips delivers both driven distances and stay times at activity locations which are, respectively, used to compute the consumed energy and the times available for recharging. Based on these data, charging points are then allocated. An example of applying this method to the city of Berlin is presented.

INTRODUCTION

Traffic is responsible for around 20% of green house gas emissions in Germany (UBA 2012). Motivations for reducing the environmental impact of petroleum-based traffic come from different directions. Global climate change can only be stopped if less green house gases are emitted. Local administrations in Europe cope with the air quality regulations formulated in the "Directive 2008/50/EC on ambient air quality and cleaner air for Europe" (EP 2008) which limits the concentrations of air pollutants that affect the environment and the population's health. Finally, peak-oil is a long-discussed issue.
One of the currently pursued solutions is the encouragement of electromobility, which, if using renewable resources for generating electricity, could reduce traffic's impacts by decreasing both climate change and the pollution of the environment.
Besides developing these vehicles, proper strategies for deploying the necessary charging infrastructure are required. Not only the amount of the required charging stations has to be determined, but also the locations they should be positioned at. The lack of appropriate public charging infrastructure is identified as one major challenge when transiting to a system of electric mobility (Trommer et. al. 2015).
The computation of the needed infrastructure as well as its allocation in space has been addressed by research in the past years. Anderson et. al. determine the loading needs for one million electric vehicles envisioned by German government for the year 2020 using vehicle usage data from the MiD 2008 mobility survey (Anderson et. al. 2016), yet neglecting the allocation in space. Some approaches allocate charging infrastructure in space using road usage (Ip et. al. 2010) or by weighting activity locations (Gkatzoflias et. al. 2016). But one may as well find other approaches that evaluate single vehicle rides as done in the following. They use either data collected from the real world (Dong et. al. 2014) or ones generated by demand models (Xi et. al. 2013). Some approaches apply an optimization scheme for computing the best placement of a limited number of charging stations (Xi et. al. 2013, Dong et. al. 2014). Other investigate the relationship between charged vehicles and the energy grid (He et. al. 2013, Loisel et. al. 2014).
The remainder is structured as following. First, a short introduction into the used traffic demand and the used simulation settings are given. Then, the procedure of computing the charging demand is described. Afterwards, the computed charging needs are presented and discussed. Finally, the conclusions are given.

USED MODELS AND SETTINGS

In the following, the agent-based demand model used, TAPAS, is outlined first. Then, the representation of the region used for computing the charging needs and allocating the charging infrastructure is discussed.

Introduction to the demand model TAPAS

The presented investigations use the agent-based demand model "TAPAS" (Heinrichs et al. 2016). TAPAS uses a representation of a region's population where every person is modelled individually and is described by a set of socio-demographic attributes, such as his/her age, sex, employment status and information about the availability of mobility options (i.e., driving license, public transport season ticket, bicycle). Each person belongs to a household which has additional information about the available cars and the household's monthly income. Every household is located at a certain geo-location within the modelled region. A region as represented in TAPAS consists additionally of the locations of activities, such as work places, schools, shops, or recreation places, and matrices that resemble access, egress, travel times and costs for the regarded modes. TAPAS models the modes "walking," "bicycling," "car driver," "car passenger," "public transport," and "car sharing."

TAPAS processes the modelled population by iterating over households, first, then over the persons. For each person, a daily activity plan obtained from the German mobility survey "Mobilität in Deutschland 2008" ("MiD 2008"; Infras & DLR 2010) is chosen. Then, TAPAS computes the locations at which the activities take place as well as the mode of transport used to reach these locations. The daily activity plans are hierarchical; if, for example a trip chain contains the trip to work, approaching the work place is computed first. Additional actions, such as shopping performed on the way to or from work what is the main activity in this example, are computed afterwards. If the resulting trip chain extends given time limits or costs, it will be dismissed and a new plan is computed. Figure 1 shows the workflow of the TAPAS demand model.

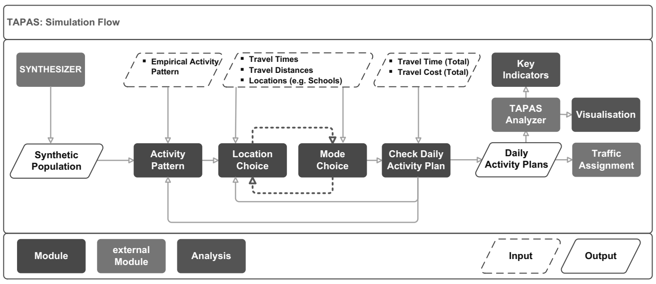


Figure 1: TAPAS workflow.

The used travel mode is determined using a multinomial logit model which is obtained by a regression of the mode choice options from the MiD 2008 survey against the attributes of persons, households, and the trip (chain). The locations are chosen either using a gravity-based (Hua and Porell 1979) approach or using so-called interviening opportunities (Stoufer 1940).

TAPAS delivers for each modelled person the trip chain for a complete average working day. This trip chain consists of single trips which are described by the departure and arrival places, times, the travel time, the used mode of transport, etc. A large set of evaluation tools can be employed to aggregate these result for obtaining, e.g., the information about certain user groups' behavior or the population's responses to regulatory, fiscal, infrastructural, or political measures (Krajzewicz et al. 2016). The generated trip chains can as well be further processed by the open source microscopic traffic flow simulation SUMO (Krajzewicz et al. 2012) for obtaining a user assignment, traffic flow measures, or pollution amounts.

To simulate electromobility, the representation of vehicles within TAPAS was extended by a maximum range. If the sum of the way length performed by an electric vehicle extends this range, the according plan is marked as being infeasible and is thereby dismissed.

The region of Berlin

In the following, a simulation setting that resembles the traffic demand in the city of Berlin in the year 2010 is used. Summarized information about the region is presented in Table 1. The population, including the persons and the households, was generated by applying the Iterated Proportional Fitting and the Iterated Proportional Updating algorithms to match data from Zensus 2011 and Mikro-Zensus 2011 (von Schmid et. al. 2016). The information about dwellings in Berlin used as input for the origins was supplied by the administration of Berlin and describes June 2012. Different data sources were used to model activity locations, including the NEXIGA data set which describes the year 2012 as well as OpenStreetMap data and other sources.

Table 1: Basic statistics of the simulated region.

Number of persons	3,287,530
Number of households	1,904,569
Number of vehicles	1,049,604
Number of dwellings	546,672
Number of Activity locations	351,289

The availability of driving licenses per age and sex and the car ownership per household are given in Figure 2.

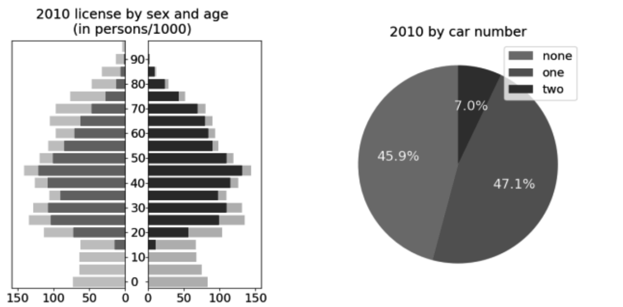


Figure 2: left: distribution of driving licenses per age and sex; right: car ownership per household.

The motorization rate in terms of vehicles per square kilometer and the modal split computed by the plain, as-is simulation of the region are shown in Figure 3. In the following, only rides performed by the motorized individual traffic mode ("driving a car") will be considered.

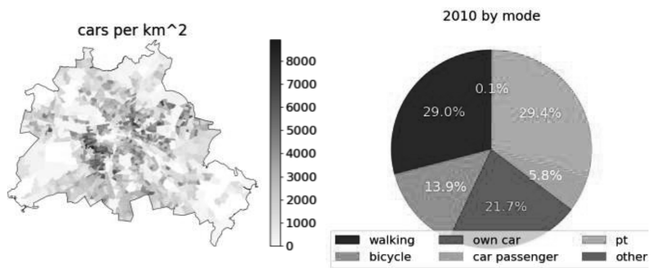


Figure 3: left: numbers of vehicle per square kilometer; right: the computed modal split.

CHARGING DEMAND

Electromobility and daily vehicle use

It should be noted, that neither the reduced price of using electric vehicles nor their lower range when compared to conventional vehicles are regarded in the following. This is motivated by an initial evaluation of distances covered by single trips and daily activities. As shown in Figure 4, 100% of single trips can be covered using the currently available ranges of electric vehicles of 150km and this even holds when taking a maximum range of 100km. As well, this is even the case for complete daily usage of vehicles as about 99.9% of distances traveled over a complete day can be covered without recharging. Please note that no commuting routes coming from regions beyond the modelled ones are considered in the given TAPAS settings. The solid lines in Figure 4 show the current common ranges of electrical vehicles of 100km and 150km. The dashed lines represent common distances driven by electric vehicle users as reported in (Frenzel et al. 2015, see Trommer et al. 2015 for a reduced English version) for reference.

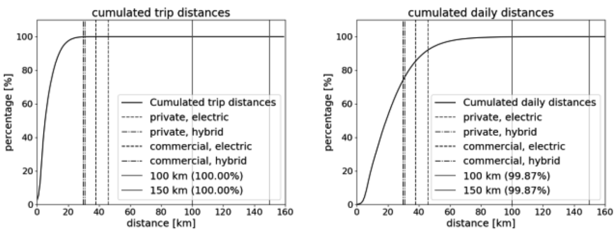


Figure 4: Cumulated distances of single trips (left) and complete daily car usage (right).

To summarize, one should assume that electromobility, given currently available ranges, can be employed in the regarded urban region without limiting mobility needs.
For simulating the maximum-fall, which in fact matches the envisioned plans for introducing electromobility (Zimmer et al. 2016), a 100% penetration rate of electric vehicles is used in the following evaluations.

Models for energy consumption and recharging times

It is assumed that a portion of the vehicle fleet is parked at privately owned places, such as a garage or a car port, and can thereby be charged at home. Socio-demographic, socio-economic and infrastructure data as well as data on the type of building at the level of a sub-traffic assignment zone ("TVZ – Teilverkehrszelle") have been used to compute the probability of a vehicle being charged at home using a linear regression. The resulting probabilities are shown in Figure 5.

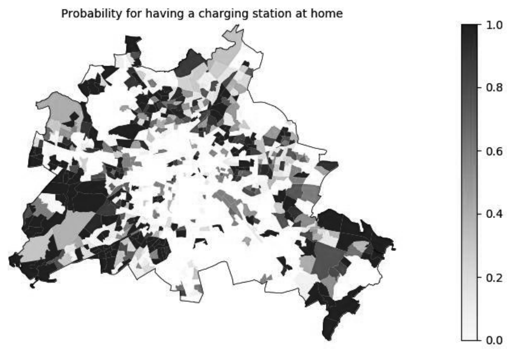


Figure 5: Probability for having a charging place at home (grey: missing values).

Still following the assumption that most of the daily trips can be completed without a recharge, the so determined 309,528 of overall 1,049,604 vehicles that can be recharged at home are disregarded in the following determination of the needed charging infrastructure. Vehicles within the 67 of 1223 zones that do not have a probability of home charging infrastructure supplied (shown grey in Figure 5) were assumed to have no loading infrastructure.
The energy consumption of the vehicles is not modelled explicitly. Rather, the amount of driven distance(s) is put against the recharging speed, which, itself, is computed as (re-)gaining a range within a given time span. The recharging speed is taken from (Hardinghaus et al. 2016). The respectively obtained recharging times for (re-)gaining a range of 1km are 5 minutes when using the slow (AC) and 1 minute when using the fast (DC) technology. One may note that these values are the measured plugged-in times, which include the time already completely recharged vehicles stay connected to the charging station. Theoretically, the AC technology should be capable to regain a single kilometer within 2 minutes and DC within 1/7min.
The recharging model assumes that at least 15 minutes are needed for recharging. All halts below this value are not used for recharging. From all halts above this duration, 15 minutes are subtracted to obtain the remaining charging time.

NEEDED CHARGING CAPACITIES

As motivated above, the charging needs and places for allocating the required infrastructure are determined by executing the model and evaluating the computed trip chains performed by car drivers. In the following, the traveled distances are put against the respectively following halting times to show that the energy lost during a trip can often be not recharged at the destination, first. Then, the method for allocating the needed charging infrastructure is described.

Energy consumption and recharging times

When looking at the relationship between traveled distances and the subsequent stays, as shown in Figure 6, one may note that the stop time is often not sufficient for recharging the energy that was used to approach the respective destination. This applies to both the slow AC and the fast DC technology.

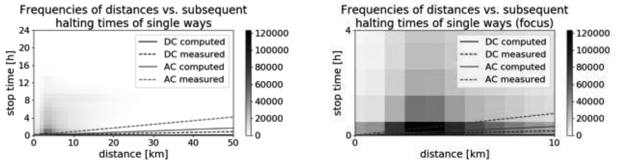


Figure 6: Occurrence distribution of driven distances vs. the subsequent halting time; left: all measurements, right: focus on small distances and travel times.

Charging needs allocation

For determining the charging needs, the trips performed by car are aggregated at their destinations. In a first step, vehicles are recharged using the slow AC technology at their destinations if their stop times allow to completely recharge them. The resulting amount of needed charging capacities is given in Figure 7, distinguishing between different types of subsequent activities.

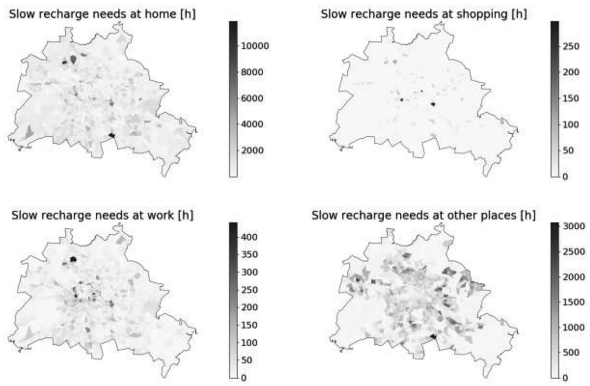


Figure 7: Needed slow (AC) charging capacities per TVZ; from left to right, bottom to top: at home, shopping, work, other.

Still, only a portion of the charging needs can be solved using the slow recharging AC technology. Vehicles that cannot be recharged using AC during the stay duration are assumed to recharge using the faster DC technology. Figure 8 shows the resulting DC capacity needs.

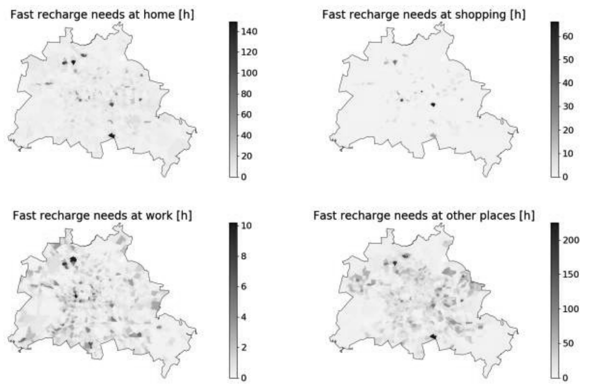


Figure 8: Needed fast (DC) charging capacities per TVZ; from left to right, bottom to top: at home, shopping, work, other.

The summarized needed recharging capacity is given in Figure 9 for the slow charging mode (AC) and in Figure 10 for the fast charging mode (DC).

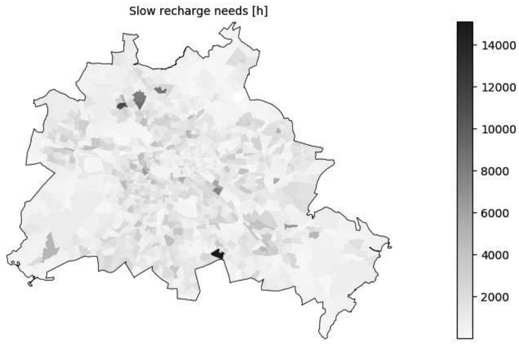


Figure 9: Needed slow charging (AC) capacities per TVZ.

Naturally, the distribution of the charging times correlate with the halting times and the distances traveled to reach the respective activity. Within the spatial distributions in Berlin, an area in the south shows to be a prominent one for all types of activities but work. This area is the "Gropiusstadt" – a dense populated area with Berlin's biggest shopping mall.

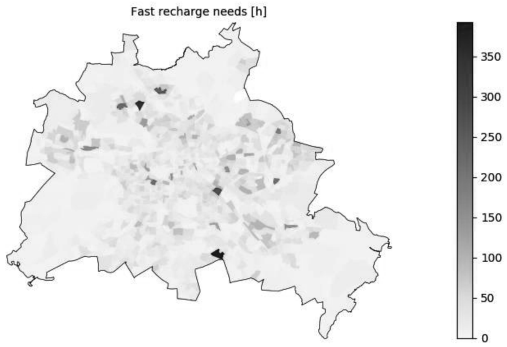
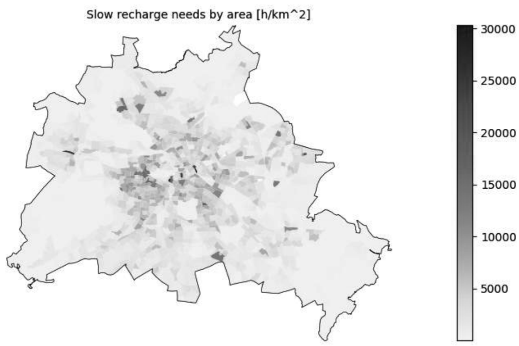


Figure 10: Needed fast (DC) charging capacities per TVZ.

Finally, the charging needs normalized by the area are given in Figure 11.



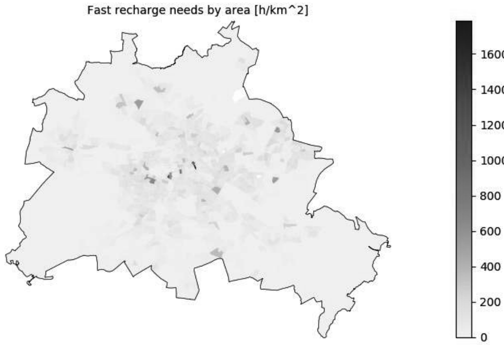


Figure 11: Needed slow (AC, top) and fast (DC, bottom) charging capacities per square kilometer per TVZ.

CONCLUSION

A method for computing the needs for recharging electric vehicles and for allocating the according infrastructure is presented. The method uses the results of the agent-based traffic demand simulation TAPAS which delivers for each modelled individual the trips performed during a usual working day.

Initially performed evaluations show that the ranges of current electric vehicles are sufficient for all single trips and for almost all distances covered during the complete day within the used Berlin test-case yet ignoring commuters. In contrary, if halts at home are neglected, many activities performed during the day are too short to regain the energy lost while approaching them.

Using the fine-grained results of the demand model, an allocation of charging needs at the level of traffic assignment zones could be determined. For this purpose, distances driven before accessing a destination were summed if the time the vehicle remains at the destination is sufficient for being completely recharged. The results show that in specific areas, a very high density of charging points is needed which one may assume to be hardly achievable.

Different extensions to the proposed methods should be performed in the future. First, the demand model should be extended by commuter trips from and to regions outside Berlin and the yet neglected changes in behavior when using electric vehicles should be investigated more deeply. Then, the available interfaces to a traffic flow simulation could be used for increasing the quality of the model for energy consumption. As well, the assumed threshold of 15 minutes below which a vehicle is not recharged should be discussed and revisited.

In addition, the resolution of placing the charging stations could be increased taking into account the given road and parking infrastructure at the TVZ. Finally, the given method computes the charging needs over a complete day. For delivering the number of needed charging stations, one should allocate recharging in time and design methods for scheduling recharging over the day.

REFERENCES

Anderson, J. E., N. Böttcher, T. Kuhnimhof. 2016. An approach to determine charging infrastructure for one million electric vehicles in Germany. In: *Conference Proceedings 1: Transportation Research Board*. (10.-14. Jan. 2016, Washington, D.C., USA).

Dong, J., C. Liu, Z. Lin. 2014. Charging infrastructure planning for promoting battery electric vehicles: An activity-based approach using multiday travel data. In: *Transportation Research Part C: Emerging Technologies*, Volume 38, January 2014, pp. 44-55, ISSN 0968-090X, DOI: 10.1016/j.trc.2013.11.001.

(EP) European Parliament and the Council of the European Union. 2008. "Directive 2008/50/EC on ambient air quality and cleaner air for Europe.", 2008.

Frenzel, I, J. Jarass, S. Trommer, B. Lenz. 2015. *Erstnutzer von Elektrofahrzeugen in Deutschland. Nutzerprofile, Anschaffung, Fahrzeugnutzung*. Project report

Gkatzoflias, D., I. Drossinos, A. Zubaryeva, P. Zambelli, P. Dilara, C. Thiel. 2016. Optimal allocation of electric vehicle charging infrastructure in cities and regions. Publications Office of the European Union, ISBN: 978-92-79-58008-6, DOI: 10.2790/353572.

Hardinghaus, M., M. Lehne, D. Kreyenberg. 2016. "Charging behavior and energy consumption of various electric vehicles in Berlin using different charging infrastructure (slow and fast) over one year." In: *EVS29 - 2016 Electric Vehicle Symposium* (19.-22.6. 2016, Montréal, Québec, Canada).

He, F., D. Wu, Y. Yin, Y. Guan. 2013. Optimal deployment of public charging stations for plug-in hybrid electric vehicles., In: *Transportation Research Part B: Methodological*, Volume 47, January 2013, pp. 87-101, ISSN 0191-2615, DOI: 10.1016/j.trb.2012.09.007

Heinrichs, M., D. Krajzewicz, R. Cyganski, A. von Schmidt. 2016. "Disaggregated car fleets in microscopic travel demand modelling." In: *The 7th International Conference on Ambient Systems, Networks and Technologies (ANT 2016)* (23.-26. Mai 2016, Madrid, Spain). DOI: 10.1016/j.procs.2016.04.111

Hua, C., F. Porell. 1979. "A Critical Review of the Development of the Gravity Model." In: *International Regional Science Review Journal*, Vol. 4, No. 2, pp. 97-126, DOI: 10.1177/016001767900400201.

Infras & DLR. 2010. "Mobilität in Deutschland 2008: Ergebnisbericht. Struktur - Aufkommen - Emissionen. " Technical Report. Available at http://www.mobilitaetindeutschland.de/pdf/MiD2008_Abschlussbericht_I.pdf

Ip, A., S. Fong, E. Liu. 2010. "Optimization for allocating BEV recharging stations in urban areas by using hierarchical clustering." In: *6th International Conference on Advanced Information Management and Service* (IMS), pp. 460-465.

Krajzewicz, D., J. Erdmann, M. Behrisch, L. Bieker. 2012. "Recent Development and Applications of SUMO - Simulation of Urban Mobility." In: *International Journal On Advances in Systems and Measurements*, 5 (3&4), 128-138. ISSN 1942-261x

Krajzewicz, D., R. Cyganski, M. Heinrichs, J. Erdmann. 2016. "Benefits of using microscopic models for simulating air quality management measures." In: *Transportation Research Board Annual Meeting* (11.-14. Jan. 2016, Washington, D.C., USA).

Loisel, R., G. Pasaoglu, C. Thiel. 2014. Large-scale deployment of electric vehicles in Germany by 2030: An analysis of grid-to-vehicle and vehicle-to-grid concepts, In: *Energy Policy*, Volume 65, Feb. 2014, pp. 432-443, ISSN 0301-4215, DOI: 10.1016/j.enpol.2013.10.029.

Stouffer, S.A. 1940. "Intervening Opportunities: A Theory Relating Mobility and Distance." In: American Sociological Review, Vol. 5, No. 6, 845-867.

Trommer, S., J. Jarass, V. Kolarova. 2015. "Early adopters of electric vehicles in Germany unveiled." In: *Proceedings of the*

28th International Electric Vehicle Symposium and Exhibition. (3.-6. Mai 2015, Kintex, Korea).

Xi, X., E. Sioshansi, V. Marano. 2013. Simulation–optimization model for location of a public electric vehicle charging infrastructure. In: *Transportation Research Part D: Transport and Environment*, Volume 22, July 2013, pp. 60-69, ISSN 1361-9209, DOI: /10.1016/j.trd.2013.02.014.

(UBA) Umweltbundesamt. 2012. *Daten zum Verkehr; Ausgabe 2012*.

von Schmidt, A., R. Cyganski, D. Krajzewicz. 2017. Generierung synthetischer Bevölkerungen für Verkehrsnachfragemodelle - Ein Methodenvergleich am Beispiel von Berlin. In: *HEUREKA'17 - Optimierung in Verkehr und Transport*, pp. 193-210. FGSV-Verlag. ISBN 978-3-86446-177-4.

Zimmer, W., R. Blanck, T. Bergmann, M. Mottschall, R. von Waldenfels, R. Cyganski, A. Wolfermann, C. Winkler, M. Heinrichs, F. Dünnebeil, H. Fehrenbach, C. Kämper, K. Biemann, J. Kräck, M. Peter, R. Zandonella, D. Bertschmann. 2016. „Endbericht RENEWBILITY III - Optionen einer Dekarbonisierung des Verkehrssektors“. Project report.

WEB REFERENCES

TAPAS web site:
http://www.dlr.de/vf/en/desktopdefault.aspx/tabid-2974/1445_read-29381/, last visited on 2nd of May 2017.

German Zensus 2011 is described at:
https://www.zensus2011.de, last visited on 2nd of May 2017.

The Micro-Zensus is described at:
https://www.destatis.de/DE/ZahlenFakten/GesellschaftStaat/Bevoelkerung/Mikrozensus.html, last visited on 2nd of May 2017.

A stochastic linear programming model for the reverse supply chain planning

A. Fall*, J. François*, R. Dupas*

**Univ. Bordeaux, CNRS, IMS, UMR 5218, 33405 Talence, France*

(Tel: 335-400- 8394; e-mail: julien.francois@u-bordeaux.fr)

Keywords: Reverse logistics, Tactical Planning, Optimization, Stochastic Linear Programming

Abstract: In this paper, a generic multi-period, multi-product stochastic linear programming model is formulated. It aims to maximize the total profit of the logistic sub-system studied (i.e. collect and sort) at the tactical planning level. This model encompasses classical production planning constraints and takes into account the uncertain quality of collected products through a stochastic optimization based approach. Finally, experimentation has been carried out in the field of a dismantling mechanical company and some preliminary results are shown in order to evaluate the interest of the stochastic modelling

INTRODUCTION

1.1 Context

Since nineties, reverse logistics which aims to get the best added value from used (end of life) products, emerges as a major economical and industrial topic. On a practical point of view, the reverse logistics flow starts with collect-and-sort operation of the input products, then disassembly, recycling of components and finally their elimination or disposal. A specific topic of the reverse logistics is the variable number of return products with different quality levels ranging from out of order to out of date.

Here are the main strategies to deal with the end of life products: i) reuse of products; ii) remanufacturing or refurbishing of used products; iii) disassembly and recovery of spare parts; recycling v) disposal or elimination of products.

In the literature, many different naming are used such as « reverse logistics », « reverse distribution » to refer to the same concept (Salema et al., 2007), (Riopel et al., 2011), (Jamshidi, 2011), (Zhou & Wang, 2008), (Salema et al., 2007), (Lee et al., 2009), (Hu et al., 2002), (Pishvaee et al., 2009), (Assavapokee et al., 2012), (Roghanian & Pazhoheshfar, 2014).

According to this literature, the reverse logistics can be defined as a reverse chain and a control flow of faulty products from customers to plants with many intermediary steps such as collect-and-sort, disassembly, while minimizing the impact on environment and the production costs.

This paper aims to: i) implement a generic planning model based on the stochastic linear programming dedicated to the planning of the reverse logistics operations; (ii) test this model on different scenarios of tactical planning encompassing the uncertainty of the quality levels of the input products.

STATE OF ART

Some papers in the field of reverse logistics planning are presented in this section. Determinist models, stochastic models and the indicators for evaluating stochastic approach are successively presented.

2.1 Deterministic Model

Many authors have proposed a deterministic planning model for reverse logistics.

(Hu et al., 2002) give a mixed integer linear model with multi-period and multi-product to minimize the total costs of reverse logistics including the following costs: collect, inventory, transportation of reuse or waste products, and the transportation to the disposal. Kim et al. (2006) propose a linear programming model for the remanufacturing supply chain in order to maximize the remanufacturing saving cost. These savings are measured by the difference between the procurement costs from the external suppliers and the remanufacturing costs of the collected products.

2.2 Stochastic Model

The reverse logistics modeling requires the use of a stochastic approach because of the multiple uncertain parameters existing in this field: the arrival dates, the quantities, the quality and the demands of the collected products, as well as the demands of components and raw material.

The stochastic modeling taking into account some random or uncertain parameters can be implemented with three different approaches (Sen & Higle, 1999): i) the scenario approach also called « recourse problem modeling » (RP): each scenario corresponds to a given value of the random parameter (or a set of values of the random parameters) and each scenario is associated to given probability; ii) the chance constrained programming approach: a mathematical formulation specifies the constraint satisfaction with a given probability; iii) the robust optimization: the model to be optimized takes into account the different uncertainties of the model.

Some models focus on the design of the supply chain networks (Salema et al., 2007), and are sometimes coupled with the direct supply chain networks (Pishvaee et al., 2009).

2.3 Performance Indicators

It is important to be able to evaluate a stochastic modelling in order to know if it is profitable and more pertinent in comparison with a deterministic model. Indeed a stochastic model can be evaluated using two complementary measures: the Expected Value of Perfect Information (EVPI) and the Value of the Stochastic Solution (VSS) (Birge & Louveaux, 2011), (Ayvaz & Bolat, 2014), (Lee & Johnson, 2014) et (Ramezani et al., 2013).

2.3.1 EVPI

This indicator measures the difference of results due to the unperfected information (data) which is represented through many scenarios. One of these scenarios has its parameter values which are equal to the mean values of all the scenarios; it is called the mean scenario and corresponds to the deterministic model. It corresponds to the scenario with perfect information. Indeed, the EVPI indicator calculates the difference of the objective functions from the deterministic model (perfect information case) « $\Re_{WS}$ » also called « Wait and See » (WS) model and the stochastic model (integrating many scenarios) «$\Re_{RP}$ ». This indicator measures the maximum cost that a decision maker (using the model) would accept to pay to get the perfect information (i.e. certain information).

2.3.2 VSS

This indicator measures the difference due to the stochastic evaluation. On a practical point of view, the implementation of this indicator needs to perform the three following steps: i) run the deterministic model with the mean scenario; ii) first run of the stochastic model integrating all the scenarios «$\Re$RP »; iii) second run of the stochastic model forcing certain variables to the values resulting from the evaluation of the deterministic model. This modified stochastic model is called « stochastic evaluation model » (EEV) and its result is labeled «$\Re$EEV ».

This indicator represents the difference of the objective function from the stochastic model «$\Re$RP » and from the stochastic evaluation model «$\Re$EEV ». VSS represents the cost of ignoring the uncertainties when making a decision. The more VSS is important, the more the stochastic modeling is useful (and reciprocally).

OBJECTIVE

Our work focuses on the collect-and-sort (i.e. logistics sub-system) of the reverse supply chain and aims to define a generic linear programming model for planning. Our model takes into account the uncertainties related to the qualities of the input products, that is why it is based on a stochastic approach. A RP modeling approach is used. This approach is based on a « Wait and See » (WS) deterministic model. In order to evaluate the results of our stochastic modeling, the stochastic evaluation model EEV will be used, as described in (Ramezani et al., 2012).

MATHEMATICAL FORMULATION

4.1 Overview of the model

A reverse logistics chain is made up of different functions (or sub-system) as in the case of the direct supply chain. A company in the field of reverse logistics can cover one or many of these functions. From an external point of view, related to the products flow, each function can be modeled in a similar way with the same number of input and output flows (except the barrier activity). However the internal function of each unit can be different. Figure 1 shows the general framework of the direct and reverse logistics including the different functions.

Our paper focuses on the sorting function (blue rectangle). Indeed our model can be considered as a generic model since it could be applied to other functionality in future works. Indeed a multi-product, multi-period, stochastic linear programming model is proposed. An overview of this model is presented in figure 2.

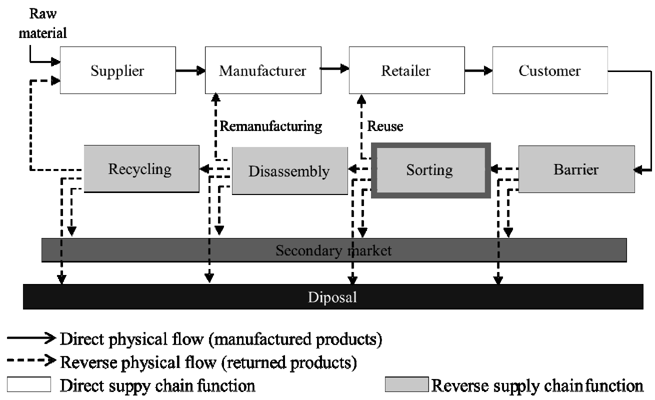


Fig. 1. General framework of direct and reverse logistics

There is one function (i.e. sorting) and four delivery destinations: U1 (direct flow), U2 (downstream function of the reverse flow), U3 (Secondary market for exportation or rebuilding products) and U4 (Disposal). Four quality levels are defined, each one respectively corresponding to a given destination: A (good), B (fair), C (poor) and D (very bad). This figure also shows the variables of the model expressed as quantities of products: input inventory $Ste_{t,p}$, sorted work in process $Tr_{t,p}$, output inventory $Sts_{s,t,I,q}$, delivery $L_{s,u,t,I,q}$, and backorders $B_{s,u\neq U4,t,i}$. The main parameters of the model are also presented in this figure: receipt of goods $C_{t,p}$, customer demands $Dmd_{u\neq U4,t,i}$ and random quantity of product of each quality $\alpha_{s,I,n,q}$. Notice that there are no backorders associated with the disposal destination as there is no disposal demand.

The sorting operation is inherently uncertain because the quality of input products can strongly varies during the time. Indeed an important hypothesis of our model is that the output inventories, the delivery and the backorders are considered as uncertain variables which are associated with scenarios and their probabilities.

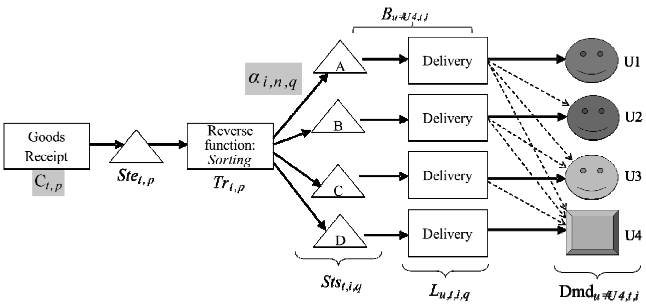


Fig. 2. Model overview: main variables and parameters

4.2 MILP formulation

Our model aims to maximize the profit (1) of the sorting function which is the difference between the revenue (2) resulting from the sales of products pertaining to all the destinations, and the costs (3) related to purchasing, inventory, processing, delivery and backorders.

The first constraints ((4) to (7)) are dedicated to the implementation in the form of linear equations of the random quality of each input products. The other remaining constraints ((8) to (25)) are usual constraints in modeling of production planning problems. Constraint (4) is a cumulative and recursive summation of products over the time. Constraint (5) is the balance of products transferred towards the inventory (6). Constraint (7) aims to increment internal variable *Cumul* for each new product. Constraint (8) controls the sum of products transferred so that they are equal to the number of work in process sorted products (i.e variable $Tr_{t,p}$), taken into account the bill of materials. Constraints (9) and (10) are respectively the processing and the delivery capacity. Constraints (11) and (12) correspond respectively to the maximum limits of input and output delivery. The maximum number of backorders is expressed in constraint (13). Constraints (14) and (15) expressed respectively the balanced of input and output inventory. Constraints (16) to (20) fix the initial values of the variables. Constraints (21) to (23) express the utilization rules of an input product according to its quality. For instance constraint (21) prevents a product with quality B or C or D, to be shipped towards destination U1 which is reserved for best quality products (i.e quality A). Constraint (24) aims to satisfy the customers demand. Constraint (25) is the backorder balance.

$$\text{Profit} = \text{Revenue} - \text{Costs} \tag{1}$$

$$\text{Revenue} = \sum_{s}\sum_{u \neq U4}\sum_{t}\sum_{i}\sum_{q} \pi_s * \text{PV}_{u,i} * \text{L}_{s,u,t,i,q} \tag{2}$$

$$\begin{aligned}\text{Costs} = {} & \sum_{s}\sum_{u}\sum_{t}\sum_{i}\sum_{q}\pi_s * \text{CL}_{u,i} * \text{L}_{s,u,t,i,q} + \sum_{t}\sum_{p}\text{CA}_p * \text{C}_{t,p} + \\ & \sum_{t}\sum_{p}\text{CSte}_p * \text{Ste}_{t,p} + \sum_{t}\sum_{p}\text{CTr}_p * \text{Tr}_{t,p} + \\ & \sum_{s}\sum_{t}\sum_{i}\sum_{q}\pi_s * \text{CSts}_i * \text{Sts}_{t,i,q} + \sum_{s}\sum_{u \neq U4}\sum_{t}\sum_{i}\pi_s * \text{Cret}_{u,i} * \text{B}_{u,t,i}\end{aligned} \tag{3}$$

$$Cumul_{t,i} = Cumul_{t-1,i} + \sum_{p} kc_{p,i} * Tr_{t-d,p} \quad \forall t, \forall i \tag{4}$$

$$\sum_{tp=1\ldots t} TF_{s,tp,i,q} - \alpha_{s,i,n,q} \leq M*(1 - y_{t,i,n}) \; \forall s, \forall t, \forall i, \forall q, \forall n = 1..100 \tag{5}$$

$$\sum_{tp=1\ldots t} \text{TF}_{tp,i,q} \leq \alpha_{i,n,q} \tag{6}$$

$$M * y_{t,i,n} = (n+1) - Cumul_{t,i} \quad \forall t, \forall i, \forall n = 1...100 \tag{7}$$

$$\sum_{q} TF_{s,t,i,q} = \sum_{p} kc_{p,i} * Tr_{t-d,p} \quad \forall s, \forall t, \forall i \tag{8}$$

$$\sum_{p} Tr_{t,p} \leq CapTr_t \quad \forall t \tag{9}$$

$$\sum_{u}\sum_{i} L_{s,u,t,i,q} \leq CapLiv_t \quad \forall s, \forall t, \forall q \tag{10}$$

$$\sum_{p} Ste_{t,p} \leq MaxSte_t \quad \forall t \tag{11}$$

$$\sum_{i}\sum_{q} Sts_{s,t,i,q} \leq MaxSts_t \quad \forall s, \forall t \tag{12}$$

$$\sum_{i} B_{s,u,t,i} \leq MaxBack_t \quad \forall s, \forall u \neq U4, \forall t \tag{13}$$

$$Ste_{t,p} = Ste_{t-1,p} + C_{t,p} - Tr_{t,p} \quad \forall t, \forall p \tag{14}$$

$$Sts_{s,t,i,q} = Sts_{s,t-1,i,q} + TF_{s,t,i,q} - \sum_{u} L_{s,u,t,i,q} \quad \forall s, \forall t, \forall i, \forall q \tag{15}$$

$$Ste_{0,p} = Ste0_p \quad \forall p \tag{16}$$

$$Sts_{s,0,i,q} = Sts0_{i,q} \quad \forall s, \forall i, \forall q \tag{17}$$

$$B_{s,u,0,i} = B0_{u,i} \quad \forall s, \forall u \neq U4, \forall i \tag{18}$$

$$Tr_{0,p} = Tr0_p \quad \forall p \tag{19}$$

$$Cumul0_i = 0 \quad \forall i \tag{20}$$

$$L_{s,U1,t,i,q} = 0 \quad \forall s, \forall t, \forall i, \forall q \in \{B, C, D\} \tag{21}$$

$$L_{s,U2,t,i,q} = 0 \quad \forall s, \forall t, \forall i, \forall q \in \{C, D\} \tag{22}$$

$$L_{s,U3,t,i,q=D} = 0 \quad \forall s, \forall t, \forall i \tag{23}$$

$$\sum_{q} L_{s,u,t,i,q} \leq Dmde_{u,t,i} + B_{s,u,t-1,i} \quad \forall s, \forall u, \forall t, \forall i \tag{24}$$

$$B_{s,u,t,p} = Dmd_{u,t,i} + B_{s,u,t-1,i} - \sum_{q} L_{s,u,t,i,q} \quad \forall s, \forall u \neq U4, \forall t, \forall i \qquad (25)$$

EXPERIMENTATION

The parameters of our model comes from a mechanical company based in the south west of France specialized in the procurement, the dismantling and the selling of spare parts of public works machines (bulldozer, scraper, excavator, loader, etc.) weighting more than10 tons. About 200 machines are bought each year.

Notice that an expert is needed to evaluate the potential reuse of products collected. He assigns a quality level to each product and therefore the most profitable destination to this product. However, according to the demand of products in the market and the level of inventory, he can decide to lower the quality of a product toward a less profitable destination in order to answer to a given demand while avoiding late penalties.

Two product numbers are considered (P1 and P2). Each product has a given quality level among the following set: {A,B,C,D}. The planning horizon has 10 periods and each period represents half of a day. Indeed the planning is done for a week of production.

The costs and selling prices used in the model are presented in Table 1. The processing time of the sorting operation is one period of time. Three products P1 and two products P2 are collected on the whole planning horizon. This relatively small value is in accordance with the low production capacity of the company (two sorted products per period) and the low maximum inventory level (five input products and five output products). The deliveries are limited to five products for each time period.

Parameter	Value		
CA_p	P1 = 9000€	P2 = 11000€	
$CSte_p$	P1 = 510€	P2 = 450€	
CTr_p	P1 = 33€	P2 = 35€	
$CSts_i$	P1 = 170€	P2 = 150€	
$CL_{u,i}$	**Destination**	**P1**	**P2**
	U1	1000€	1100€
	U2	0€	0€
	U3	800€	900€
	U4	600€	500€
$Cret_{u,i}$	**Destination**	**P1**	**P2**
	U1	340€	300€
	U2	300€	280€
	U3	270€	260€
$PV_{u,i}$	**Destination**	**P1**	**P2**
	U1	17000€	15000€
	U2	15000€	14000€
	U3	13500€	13000€

Table. 1. Costs and selling prices (per product)

At time zero (t=0), there are one product P1 in the input inventory, two products P1 in the sorted operation work in progress and also, one P2 (quality A) and one P1 (quality B). The following demands are considered: 2 P1 and 3 P2 for destination U1, 2 P1 and 3 P2 for destination U2 and finally 3 P1 and 2 P2 for destination U3.

5 scenarios are considered in our model (S1 to S5) with the same probability. Here are the quality distributions of products for each scenario.

	S1	S2	S3	S4	S5
A	60%	36%	35%	21%	23%
B	40%	50%	50%	52%	58%
C	0%	13%	13%	24%	15%
D	0%	1%	2%	3%	4%

Table. 2. Quality distribution of products for each scenario

The first level decisions are the input inventory and the workload. All the other variables are the second level decisions which are subsequent to the processing: output inventory, deliveries, and backorders.

Figure 3 shows the mean profits of the three models according to the five scenarios. For each scenario, ten replications (i.e. runs) are done with random sampling of the quality levels of products. In scenario 1, the profit is important since all the products have good quality (A or B) and therefore all the demands are satisfied except the last ones because the products supplied and the inventories are inferior to the demands. Hence some penalty costs are taken into account.

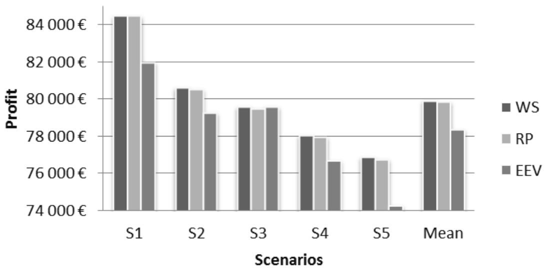


Fig. 3. Profits of the three models according to the senarios

The results of the stochastic model and the deterministic models are equal as it is shown by the zero value of indicator EVPI in figure 4. However, the evaluation model EEV gives a profit which is reduced by about 2000 euros compared to the deterministic model WS; indeed the value of indicator VSS is 3 % in figure 4.

In scenario 2 to 5, the quality of products decreases. Hence it is more and more difficult to satisfy the customers' demands and there are more late penalties accumulated over the horizon mainly in relation with destination U1 which is normally the more profitable selling. Moreover some quantities of output inventories of products with quality C arise and generate added inventories costs. Furthermore the limited capacity of inventories requires eliminating certain products to destination U4. Consequently there are some transportation costs without any gain as these products are not sold. As a result of all these remarks, the profits decrease from scenario 2 up to scenario 5.

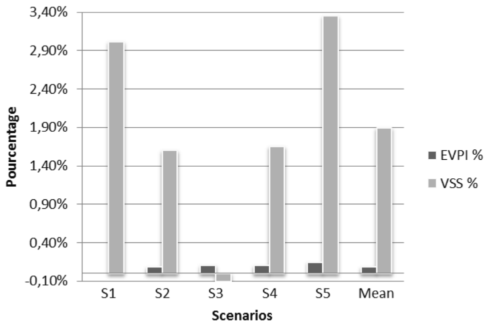


Fig. 4. EVPI and VSS indicators according to the scenarios

Figure 4 clearly shows that indicator EVPI is rather low (mean value of 0.1 %), proving that there is a very slight difference between the stochastic and the deterministic results. Otherwise indicator VSS has a mean value of 2 %, showing the potential interest of the stochastic approach.

Notice that in scenario 3, indicator VSS is negative. This can be explained by the fact that some variables of the stochastic evaluation model are constrained with the values of the mean scenario of the deterministic model (i.e. scenario 3). That is why in this case, the optimal solution of the deterministic model is obtained. And this observation can be generalized: more the scenarios are far from the mean scenario (S2 and S4, then S1 and S5), more the VSS indicator increases. This is because the constrained values with the values of the mean scenario correspond less to the other scenarios.

CONCLUSION

Due to the uncertain quality and quantity of products collected in the reverse logistics, a stochastic modeling is required to optimize flows. In this paper, a stochastic linear programming modeling using the scenario approach has been implemented to optimize the profit of an upstream company in the reverse supply chain. The interest of this modeling has been confirmed. The presented model is generic and has been applied to other functions of the reverse supply chain, not presented in this paper. For instance the disassembly of products has also been tested, taking into the inverted bill of material of the products disassembled. Another delivery strategy (i.e. utilization rules of used products) without the possibility to low the quality of products has also been investigated. In this case, a mean decrease of 28 % has been shown, compared to the model with the possibility to low the quality of products, presented in the paper. This shows that the approach with the possibility to low the quality is more interesting for the reverse logistics companies.

REFERENCES

Assavapokee, T., & Wongthatsanekorn, W. (2012). Reverse production system infrastructure design for electronic products in the state of Texas. (Elsevier, Éd.) Computers & Industrial Engineering, 62, 129-140.

Ayvaz, B., & Bolat, B. (2014). Proposal of a Stochastic Programming Model for Reverse Logistics Network Design under Uncertainties. International Journal of Supply Chain Management, 3, 33-42.

Birge, J., & Louveaux, F. (2011). Introduction to Stochastic Programming. (S. S. Engineering, Éd.) New York.

Hu, T.-L., Sheu, J.-B., & Huang, K.-H. (2002). A reverse logistics cost minimization model for the treatment of hazardous wastes. (Elsevier, Éd.) Transportation Research Part E, 38, 457-473.

Jamshidi, M. (2011). Reverse Logistics. Logistics Operations and Management.

Kim, K., Song, I., Kim, J., & Jeong, B. (2006). Supply planning model for remanufacturing system in reverse logistics environment. (Elsevier, Éd.) Computers & Industrial Engineering, 51, 279-287.

LLee, J.-E., Gen, M., & Rhee, K.-G. (2009). Network model and optimization of reverse logistics by hybrid genetic algorithm. (Elsevier, Éd.) Computers & Industrial Engineering, 56, 951-964

Lee, C.-Y., & Johnson, A. (2014). Proactive data envelopment analysis: Effective production and capacity expansion in stochastic environments. (Elsevier, Éd.) European Journal of Operational Research, 232, 537-548

Pishvaee, M. S., Jolai, F., & Razmi, J. (2009). A stochastic aptimization model for integrated forward/reverse logistics network design. (Elsevier, Éd.) Journal of Manufacturing systems, 28, 107-114.

Ramezani, M., Bashiri, M., & Tavakkoli-Mogaddam, R. (2013). A new multi-objective stochastic model for a forward/reverse logistic network design with responsiveness and quality level. (Elsevier, Éd.) Applied Mathematical Modelling, 37, 328-344.

Riopel, D., Chouinard, M., Marcotte, S., & Aït-Kady, D. (2011). Ingénierie et gestion de la logistique inverse. (Lavoisier, Éd.) Paris.

Roghanian, E., & Pazhoheshfar, P. (2014). An optimization model for reverse logistic network under stochastic environment by using genetic algorithm. (Elsevier, Éd.) Journal of Manufacturing Systems, 33, 348-356.

Salema, M. I., Barbosa-Povoa, A. P., & Novais, A. Q. (2007). An optimization model for the design of a capacitated multi-product reverse logistics network with uncertainty. (Elsevier, Éd.) European Journal of Operational Research, 179, 1063-1077.

Sen, S., & Higle, J. (1999). An Introductory Tutorial on Stochastic Linear Programming Models. *Institute for Operations Research and the Management Sciences, 29*, 33-61.

Zhou, Y., & Wang, S. (2008). Generic Model of Reverse Logistics Network Design. (Elsevier, Éd.) Journal of transportation systems engineering and information technology, 8, 71-78.

DATA MANAGEMENT SIMULATION

POLITICAL POLLS
MARKET RESEARCH AND BUSINESS OPTIMIZATION

William Conley
Austin E. Cofrin School of Business
480B L. G. Wood Hall
University of Wisconsin--Green Bay 54311-7001
ConleyW@uwgb.edu

KEYWORDS
Survey sampling, market research, elections, statistical optimization, Monte Carlo methods, business applications.

ABSTRACT

The 2016 United States of America (USA) presidential election was followed with considerable interest by people in the USA and in other countries too. One result is that many pollsters and others who tried to predict the outcome between the Democratic and Republican candidates had difficulty with their forecasts. The reason for this will be explored here. However, the important additional point will be made that market research is still very useful to business when it is done correctly, by experts. Also, the market research survey and sampling techniques can now be applied to general business optimization problems where it is fairly easy to use random number generators to survey the feasible solution space (population) for the correct answer.

INTRODUCTION

The 1948 USA presidential election between President Harry Truman (Democratic party) and Governor Thomas Dewey (Republican party) showed the importance of getting a random sample of the true population to be studied before making any predictions. A poll was taken with telephone calls to potential voters before the 1948 election and about 55 percent were for Thomas Dewey and about 45 percent were for Harry Truman. Then, one of America's most respected newspapers (*The Chicago Tribune*) ran a front page headline the day after the election announcing Thomas Dewey the winner. However, Harry Truman won the actual election. The problem was that a significantly lower percentage of Democrats had telephones than Republicans in 1948.

The follow up studies showed clearly that Dewey did get 55 percent of the vote from voters who owned telephones. Therefore, the sample was biased (unintentionally perhaps) toward the Republican candidate winning.

Then it was not too long after that 1948 election that most American families did have telephones, so political surveys (using telephones) in the USA might be fairly accurate (with a few concerns the professional pollsters had to watch out for). (Conley 2000) did a statistical analysis of the 2000 USA presidential election.

THE 2016 USA PRESIDENTIAL ELECTION

The author has lived in the USA for more than 65 years and taught business statistics in colleges and universities to about 10,000 students. However, America's telephone landscape has changed. The proliferation of cellphones and so many Americans have call ID, that if they do not recognize a telephone number as coming from a friend or an associate, they will not answer the phone. Also, data is being collected on Americans at a relentless pace. Therefore, many people will just not participate in surveys of almost any kind. If a political pollster placed 1,000 phone calls, perhaps only 800 could answer the phone. However, in the "political season" (protected by caller identification) only perhaps 100 answer it. Then when the pollster says it's a political survey 80 of them hang up the phone. Then if the question is poorly phrased, another ten will not answer. That leaves ten, and six of them are for candidate A and four are for candidate B. Therefore, should it be reported that in a survey of 1,000 Americans 60 percent of the respondents are for Candidate A and 40 percent are for candidate B? Of course not!

This is a hypothetical example, but surveying the American public on politics can be difficult today. Some of America's most respected pollsters declined to do political polls in the 2016 election because they could not be confident of the results. (Please see the AMERICAN MEDIA section herein.) These pollsters deserve praise and respect.

So, is market research in business in trouble? No! If it's done correctly by professionals it can be very useful (and even vital) to businesses worldwide. Also, the same principles of statistical survey sampling can be used in the new field of statistical optimization (Conley 2015). An example follows after some information by and about the American media.

AMERICAN MEDIA

Some of America's respected polling organizations told *The Wall Street Journal* that they would not do presidential polls in this (2016) election because they could not guarantee reasonable accuracy under current conditions (Barone 2015) and (McGinty 2016).

One of these pollsters is the highly respected Gallup organization. Their refusal to do the 2016 presidential polling survey speaks volumes about survey sampling should not be used when getting a truly random sample of the larger papulation is almost impossible.

On the day after the election a *USA Today* headline set the tone for the outpouring of criticism to follow: "TRUMP BLASTS PAST POLL PREDICTIONS" (Jackson 2016).

Perhaps what the American public really wants to know is are they really being informed or are they being manipulated. (Impoco 2016) reports that the Friday before the election, the highly respected *New York Times* predicted Hillary Clinton had an 84 percent chance of being elected president. (Hampson 2016) reveals the shock and surprise of so many Americans at the result. Marc Landy, a Boston College political science professor, provided the second part of Hampson's headline: "The pollsters have a lot to answer for".

A front-page *USA Today* headline (Callaway 2016) on Thursday, November 10, stated, "How did pollsters miss the mark?" Also, (Stanley 2016) reprints a Bloomberg News *USA Today* article with many comments from experts about what went wrong in the polling.

Could it be that given the realities of the media, telephones and computer age, it is virtually impossible to do political polls in a cost effective, accurate way? Yet market research and statistical optimization seem to be bright areas for survey sampling (Conley 2015).

MARKET RESEARCH ACTUARIAL SCIENCE AND SIMULATION

However, survey sampling (when it's done well) and various probability techniques are useful and viable in our computer age. The field of market research is so important when test marketing new products and discovering buyer trends and locating a company's customers (current and future ones).

Actuarial science is so dependent on careful and correct probability and statistics calculations. The insurance industry could not exist without it. (Conley and Conley, Sr. 1980) do optimization multiple integrals, and actuarial problems using survey-sampling techniques (instead of calculus) applied with computer simulation (which uses accurate random number generators (surveys)).

The articles in this proceedings volume give many examples of the usefulness of probability and statistics in conjunction with computer simulations. Also, quality control in manufacturing is another useful area for survey sampling.

STATISTICAL OPTIMIZATION

The same principles that govern survey sampling can be used (and slightly modified) to solve a wide variety of difficult multivariate optimization problems even when the functions are nonlinear.

Let us try survey sampling to find a solution to the nonlinear system of equations

$$X_2X_2X_2 + 7X_3X_4 + X_5 = 7{,}492{,}317 \quad (1)$$
$$X_1X_1X_1 + 3X_4X_6 + X_6 = 128{,}514 \quad (2)$$
$$X_3X_3X_3 + 5X_1X_2 + X_4 = 2{,}093{,}320 \quad (3)$$
$$X_4X_4X_4 + 6X_1X_6 + X_3 = 691{,}474 \quad (4)$$
$$X_5X_5X_5 + 3X_2X_5 + X_4 = 1{,}028{,}301 \quad (5)$$
$$X_6X_6X_6 + 4X_3X_6 + X_1 = 1{,}745{,}657 \quad (6)$$
$$X_2X_2X_2 + 5X_4X_5 + X_1 = 7{,}457{,}986 \quad (7)$$

subject to $0 \leq X_i \leq 200$ and X_is are all whole numbers for i = 1, 2, 3, 4, 5, and 6.

First we transform the system to minimize

$$f(X_1X_2X_3X_4X_5X_6) = |L_1 - R_1| + |L_2 - R_2| + |L_3 - R_3| + |L_4 - R_4| + |L_5 - R_5| + |L_6 - R_6| + |L_7 - R_7|$$

subject to $0 \leq X_i \leq 200$ and whole numbers where L_j and R_j are the left and right hand sides of equation j respectively for j =1, 2, 3, 4, 5, 6, and 7.

Then a multi stage Monte Carlo optimization (MSMCO) program is written to "survey sample" the feasible solution space for all whole number answers to solve the system of equations by minimizing $f(X_1X_2X_3X_4X_5X_6)$ down to zero. The first stage of the simulation will look at 10,000 feasible solutions selected randomly (controlled by a random number generator) and store the best answer found so far. Then centered about this best answer so far, in stage two, another sample of 10,000 feasible solutions will be selected and the best ones found will be stored and re-center the search that way. However, in step two the search region will be reduced in each dimension by a factor of 1.41 (or the square root of 2). In stage three another 10,000 feasible solutions will be looked at (always re-centering the search about the best answer found so far) in a further reduced search region whose dimensions are reduced by another factor of 1.41. Seventeen stages of 10,000 sample feasible solutions are done in this manner. Therefore, by the latter stages of the multi stage (MSMCO) simulation, this program funnels into the true optimal (minimum of f) of zero, solving the system. The answer produced in about three seconds of simulation run time on a desktop computer in the author's business school office is $X_1 = 46$, $X_2 = 195$, $X_3 = 127$, $X_4 = 87$, $X_5 = 99$, $X_6 = 119$, or $f(46, 195, 127, 87, 99, 119) = 0$ solving this system of equations. Figures 1 and 2 give partial geometric and statistical representations of this multi stage simulation at work. This system of equations is a hypothetical one. However, this multi stage Monte Carlo approach can work on economic equilibrium equations, production planning equations, transportation systems of equations and chemical yield equations (Conley, 2016) and equations in the electrical component industry (Conley, 1985). More business applications are mentioned in the conclusion.

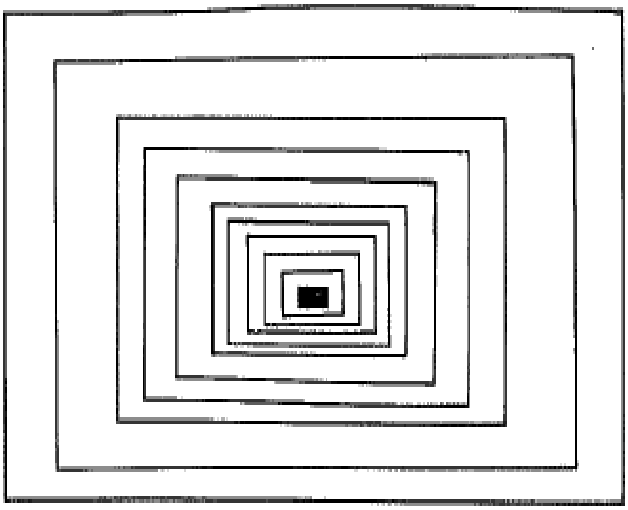


Figure 1: An eleven stage search for the optimal solution

Also, a manufacturing optimization and control example is presented in Conley (2015). Fifteen products are made and packaged by passing them through five departments with various machine and worker labor hours available. However, there are considerable returns to scale available as more of any one product is made. Therefore, a fifteen variable nonlinear system of equations arises and is solved using survey sampling techniques (multi stage Monte Carlo optimization) to find solutions to a factory production system of equations rather than estimate a probability in a public opinion poll.

Here the production manager does not have to worry about getting the wrong information from the public in a survey. Local optimals and the validity and accuracy of the system of equations are of a more immediate concern to the manager and the chief executive officer.

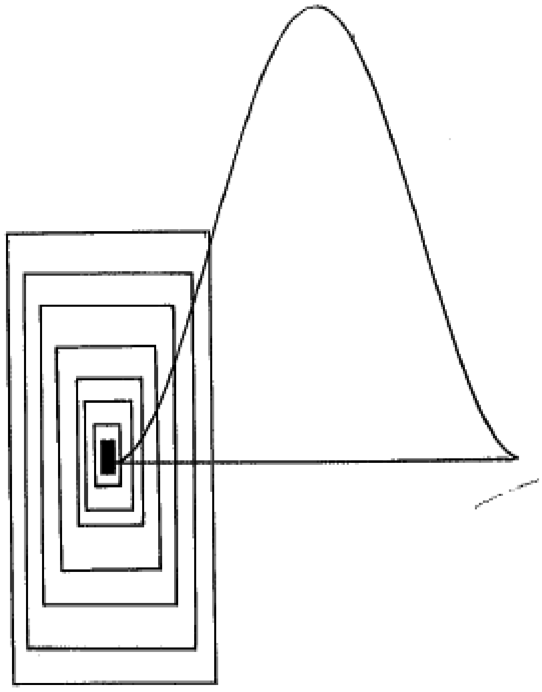


Figure 2: An eight stage MSMCO search for the minimum

CONCLUSION

The many political polls conducted about the 2016 presidential election in the United States of America make the case that polls have a considerable difficulty being accurate due to changes in telephone technology (problem of caller ID features and cellphones) and the reluctance of most Americans to answer questions about polls when so many organizations are now collecting data on Americans and creating and storing large data banks on the whole population.

Again the respected Gallup polling organization declined to do polls on the 2016 presidential election because they could not guarantee that they were getting an accurate random sample. Yet their announcement is a great service to the public that polls have to be done correctly to mean anything.

However, the good news about survey sampling is that market research surveys are still vital and useful and can be very accurate. Also, the new field of statistical optimization or multi stage Monte Carlo optimization (MSMCO) is useful in our computer age as long as the random processes can be controlled and focused to solve the problems under consideration.

The author's business statistics students were told repeatedly that if you doubt the ethics or integrity or competence of the person doing the statistical study, the results mean nothing or it could mean fraud, and also to be continually watching for samples that are not random if you are assuming randomness in the survey. They were also told that the "good news" is that most studies done with statistics are fairly accurate and the risks and potential errors are explained and discussed thoroughly.

However, Darrell Huff's bestselling book *How to Lie with Statistics* (Huff and Geiss 1993) does offer some sound advice (in spite of its ominous title) on what to watch out for when assessing the usefulness and accuracy of a statistical study. (Cochran 1977) also presents much good advice.

A hypothetical example would be a large news organization takes a political poll and reports its results on a "slow" news day. Now, is that reporting the news, or creating the news, or staging the news? How large was the sample and was it random? How many declined to respond? What were the questions asked, and how were they phrased?

Some additional examples of statistical optimization power and versatility can be found in (Conley 1986) and (Conley 2013) who works on multivariate inventory problems, (Wong1990), (Conley and Wichowski 1989) and (Wong 1996), among others. (Conley 2012a) approximates the solution to a one million variable nonlinear transportation problem using multi stage Monte Carlo optimization (MSMCO) simulations. Additionally, (Conley 2012b) uses multi stage Monte Carlo optimization (MSMCO) to optimize the trading of electric power between individual plants and also internationally between countries as supplies and demands for power change nationally and internationally. The adjustment and tuning of industrial machines is also worked on in Conley (2014) with MSMCO.

Please note that the (Conley 1988) article only deals with cost control in two and three dimensional (great circle airline distance calculations) space where (Conley 2017) deals with higher dimension shortest routes in correlation analysis.

Also, it should be noted that for Southern California's NBC4 TV station, veteran reporter and newscaster, (Ted Chen 2016), stated that the USC/*Los Angeles Times* polls that consistently predicted that Donald Trump would win the USA 2016 presidential election was scorned and criticized. Yet the USC/*Los Angeles Times* poll turned out to be about the only widely followed poll that made the correct predication. (USC is an acronym for the University of Southern California.) However, the USC/*Los Angeles Times* poll did use a different and unusual survey approach.

Despite forecasting differences and problems there remains the good news of great reporters and journalists from New York to Los Angeles, and Florida to Alaska, doing excellent and important reporting all over the United States. Examples of this great work are the *Milwaukee Journal Sentinel*'s Pulitzer prize-winning business reporter Kathleen Gallagher and reporter Mark Johnson (Gallagher & Johnson 2016).

REFERENCES

Barone, M. 2015. Why political polls are so often wrong. *Wall Street Journal*, New York, p. A17.

Bomey, N. 2016. "Soul-searching lies ahead for pollsters who blew it." *USA Today I*(November 10), Arlington VA.

Callaway, D. 2016, November 10. "How did pollsters get it so wrong?" *USA Today* (Nov. 10), Arlington, VA.

Chen, T. 2016. "How USC/LA Times poll predicted a Trump victory." *NBC4 TV Station News* (November 9)*,* Southern CA.

Cochran, W. C. 1977. *Sampling techniques* (3rd ed.). New York, NY: John Wiley and Sons.

Conley, W. C. 1984. "Making an adjustable resistor from single valued fixed resistors." *ntzArchiv. 7,* Berlin & Offenbach: VDE-Verlag GmbH, pp. 231-234.

Conley, W. C. 1986. Production planning with nonlinear cost considerations. *Journal of Cost Engineering 28*(6), June, pp. 16-20.

Conley, W. C. 1988. Controlling costs by finding the shortest route: A multi stage Monte Carlo optimization approach. *Journal of Cost Engineering 30*(11) (Nov.), pp 2038-2056.

Conley, W. C. 2001. "A lesson in statistics." *Milwaukee Journal Sentinel* (July 28), p. 12A.

Conley, W. C. 2012a. A million variable nonlinear transportation problem. Proceedings of the *2012 Industrial Simulation Conference ISC2012,* (June). Brno, Czech Republic, Ostend, Belgium: EUROSIS, pp. 113-121.

Conley, W. C. 2012b. An electric power optimization problem. Proceedings of the *2012 European Simulation and Modelling Conference ESM2012*, Oct. 22-24, Essen, Germany. Ostend, Belgium: EUROSIS, pp. 208-211.

Conley, W. C. 2013. Multivariate inventory and selected problems using multi stage Monte Carlo optimization. Proceedings of the *2013 Industrial Simulation Conference ISC2013*, Ghent University, Ghent, Belgium, May 22-24. Ostend, Belgium: EUROSIS, pp. 208-216.

Conley, W. C. 2014. Machine and engine adjustment. Proceedings of the *2014 European Simulation and Modelling Conference ESM2014*, University of Porto, Porto, Portugal, October 22-24. Ostend, Belgium: EUROSIS, pp. 321-326.

Conley, W.C. 2015. The other application area of survey sampling. Proceedings of the *2015 Industrial Simulation Conference ISC2015*, Valencia, Spain. Ostend, Belgium: EUROSIS, pp. 10-14; p. 3A..

Conley, W. C. 2016. Statistical optimization applied to chemical yield equations. Proceedings of the *2016*

Industrial Simulation Conference ISC2016, Universitatea Politehnica, Bucharest, Romania. Ostend, Belgium: EUROSIS, pp. 145-148.

Conley, W. C. 2017. Controlling costs by taking the shortest route to information. Under review.

Conley, W. C. & Conley Sr., W. C. 1980. Computer statistics versus calculus. Proceedings of *COMPSTAT.* Wurzburg, Germany: Physica Verlag.

Conley, W. C. & Wichowski, H. 1988, Simulation and model optimization of three biomedical systems. *Journal of the Japan Society for Simulation Technology* (March), pp.47-57.

Gallagher, K. & Johnson, M. 2016. *One in a billion: The story of Nic Volker and the dawn of genomic medicine.* New York, NY: Simon and Schuster.

Hampson, R. 2016. Divided nation united by shocker "The pollsters have a lot to answer for". *USA Today* (November 20), Arlington, VA. pp. 1A-2A.

Huff, D. and Geiss. 1993. *How to lie with statistics*. New York, NY and London: W. W. Norton and Company.

Impoco, J. 2016. "Polling error." *Newsweek Magazine*, November 8, Commemorative Edition, Topix Media Lab, New York. pp. 78-79.

Jackson, D. 2016. Trump blasts past poll predictions. *USA Today* (November 9), Arlington, VA. p. 5A.

McGinty, J. C. 2016. Polarized over polls: Internet or phone. *Wall Street Journal* (January 30-31), New York. P. A1.

Stanley, G. 2016. How did pollsters get election so wrong? *Milwaukee Journal Sentinel* (November 10), p. 11A.

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

Wong, J. T. 1996. A note on optimization in integers. *International Journal in Mathematical Education in Science and Technology, 27*, 865-874.

AUTHOR BIOGRAPHY

WILLIAM CONLEY was born in Lansing, Michigan, USA in 1948 and received a B.A. in mathematics from Albion College (with honors) and an M.A. in mathematics from Western Michigan University in 1971, an 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—Green Bay. The developer of multi stage Monte Carlo optimization and the CTSP multivariate correlation statics, he is the author of five books and more than 200 publications worldwide. He is a member of the American Chemical Society, a Michigan Scholar in College Teaching, a fellow in the Institution of Electronic and Telecommunication Engineers, and a senior member of the Society for Computer Simulation. Career highlights include presenting a paper at a National Aeronautics and Space Administration (NASA) conference in Houston and a speech at the NASA Goddard Space Flight Center outside of Washington, D. C. on the occasion of the tenth anniversary of their data center. Additional honors include Phi Beta Kappa (National Academic Honorary), Omicron Delta Kappy (ODM) National Leadership Honorary, and KME and ODE, The National Mathematics and Economics Honoraries.

CONTROLLING COSTS BY TAKING THE SHORTEST ROUTE TO INFORMATION

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
Optimization, costs, complex systems, transportation, statistical correlation.

ABSTRACT

The multistage Monte Carlo simulation solution technique is useful in solving general optimization problems and shortest route problems to reduce delivery costs for products and travel costs for people who must travel extensively. However, it is possible to expand this idea beyond two or three dimensions to finding shortest distances through spreadsheets of data (where the columns represent different variables). It turns out that if the shortest routes through spread sheets of data are shorter than through similar size and range random data, then the variables represented by the columns are correlated. Therefore, as our computer world continues to overwhelm us with spreadsheets of data, a shortest route statistical test can tell whether the variables are correlated or not. Practitioners can now take the shortest route to important information.

INTRODUCTION

Two examples are presented here, along with some explanation and drawings. (Anderson 2003), (Keller and Warrack 2003), (Hayter 2002) and (Black 2014) present and review the standard linear correlation coefficient r (and its multivariate linear Big R). However, the presentation here is more general because it can find linear or nonlinear multivariate correlation in our computer age with the new CTSP statistic for picking up correlations by finding shortest routes through the spreadsheets of data.

First consider a No Correlation in Example One. A business team looks at the n=49 lines of eleven columns of data in DATA SET ONE. There are 11x49=539 numbers there. However, do these numbers mean anything? Can they be reduced to an important piece of information? Everyone concerned knows that the n=49 lines represent 49 sample readings of the eleven variables (the eleven columns) in the study.

.

Table 1: Data Set One

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}
1	56	0	61	81	11	60	51	82	28	59	14
2	64	87	54	0	69	49	95	19	100	5	23
3	30	2	82	17	12	100	73	57	1	20	81
4	13	86	80	2	24	19	6	72	62	37	69
5	34	97	60	86	3	98	33	55	28	2	67
6	91	93	63	27	16	57	91	33	10	27	82
7	83	87	91	58	30	5	78	100	16	19	86
8	15	53	86	76	80	37	26	35	76	37	50
9	35	91	53	92	30	40	76	94	92	15	50
10	20	65	61	73	5	33	42	64	2	23	23
11	64	12	77	83	86	83	36	97	44	25	45
12	18	9	55	5	42	14	22	2	21	73	28
13	6	54	5	47	94	40	21	82	94	23	54
14	35	62	8	7	68	83	97	42	8	94	99
15	85	75	17	80	8	64	97	3	86	10	17
16	70	71	63	46	28	5	87	58	88	65	65
17	4	93	38	93	82	6	30	117	43	19	75

18	98	20	55	75	58	100	73	63	41	74	99
19	98	75	24	12	96	46	42	43	93	35	2
20	58	14	68	68	56	72	31	66	9	6	37
21	84	100	83	62	5	11	14	92	0	34	3
22	14	58	74	12	81	94	46	58	38	20	84
23	25	81	100	16	88	21	71	18	100	44	17
24	37	12	23	75	5	81	17	0	57	88	23
25	25	28	92	40	12	14	79	31	100	88	61
26	20	74	73	53	51	36	68	88	15	65	43
27	36	55	54	22	51	20	76	66	24	20	81
28	77	25	58	65	66	87	36	60	34	81	32
29	90	28	54	17	80	87	29	24	90	65	17
30	10	30	6	99	81	63	70	60	46	93	91
31	52	69	19	72	78	42	72	56	2	94	3
32	29	58	100	68	94	94	34	95	55	99	54
33	0	61	21	7	71	83	72	100	75	61	28
34	40	94	85	32	23	64	69	46	71	8	94
35	9	81	65	79	91	2	26	37	88	96	45
36	89	52	19	29	3	67	57	47	85	65	55
37	87	76	5	10	26	45	30	89	50	48	73
38	27	9	87	46	81	30	79	95	63	4	24
39	51	62	77	1	12	15	87	70	62	29	53
40	52	39	63	52	46	36	16	83	38	14	67
41	3	34	88	48	4	82	36	100	87	99	99
42	13	38	31	98	75	64	24	100	60	42	26
43	92	91	45	5	4	89	81	100	84	16	13
44	118	51	75	37	38	97	33	54	28	14	88
45	100	81	47	78	17	65	29	7	94	4	45
46	76	5	97	95	6	81	12	80	54	41	28
47	19	27	89	21	6	82	14	76	14	30	51
48	28	12	83	84	88	63	3	82	42	52	45
49	42	100	9	100	82	84	84	59	16	80	80

Therefore, to test the hypothesis that the variables are not correlated (H_o or null hypothesis) a shortest route connecting the n=49 lines of data is calculated using the multistage Monte Carlo optimization (MSMCO) shortest route algorithm (Conley 1991a) adjusted for eleven variables. (The Pythagorean distance formula is just generalized for 11 variables.)

The square root of $(x_1\text{-}y_1)^2 + (x_2\text{-}y_2)^2 + (x_3\text{-}x_3)^2 + (x_4\text{-}y_4)^2 + (x_5\text{-}y_5)^2 + x_6\text{-}y_6)^2 + x_7\text{-}y_7)^2 + (x_8\text{-}y_8)^2 + x_9\text{-}y_9)^2 + x_{10}\text{-}y_{10})^2 + x_{11}\text{-}y_{11})^2$ is used to measure the distance between two points $(x_1, x_2, x_3 \ldots x_{11})$ and $(y_1, y_2, y_3 \ldots y_{11})$ (in eleven dimensional space). The resulting shortest route through the data yielded a total distance of d.=4657.675 with the following point to point tour:

16	25	41	32	28	18	14	30	49	31
26	10	21	7	6	34	4	37	36	29
24	12	17	35	8	23	2	19	43	15
45	5	9	38	46	1	47	3	44	22
33	13	42	48	11	20	40	27	39	16

Therefore, in eleven dimensional space the short route taken was (from left to right top to bottom) from point 16 to point 25 to point 41 . . . to point 27 to point 39 and back to point 16.

Then four sets of random data (of similar ranges with k=11 columns and n=49 rows) were read into the multistage Monte Carlo optimization (MSMCO) shortest route algorithm and their shortest routes found were d_2=4500.262, d_3=4732.425, d_4=4514.775 and d_5=4528.076. The median of these four route distances is (4514.775+4528.076)/2=4521.426.

Now CTSP = d./median = 4657.674/4521.426 = 1.0301 which is comfortably in the CTSP range if there is no correlation between the variables. Therefore, the null hypothesis of no correlation cannot be rejected. Also note that this is a one-sided statistical test. H_o of no correlation can only be rejected if CTSP is significantly below one.

CTSP REASONING

The CTSP shortest route correlation statistic approach reasoning is that looking at the 12 pairs of the four random data shortest routes (d_2, d_3, d_4 and d_5) quotients, the range of them is .9509 to 1.0516.

4500.262/4732.425=.9509
4500.262/4514.775=1.0514
4500.262/4528.076=.994
4732.425/4500.262=1.0516
4514.778/4500.262=1.003
4528.076/4500.262=1.006
4732.425/4514.775=1.048
4732.425/4528.076=1.045
4514.775/4732.425=.954
4528.076/4732.425=.957
4514.775/4528.076=.997
4528.076/4514.775=1.003

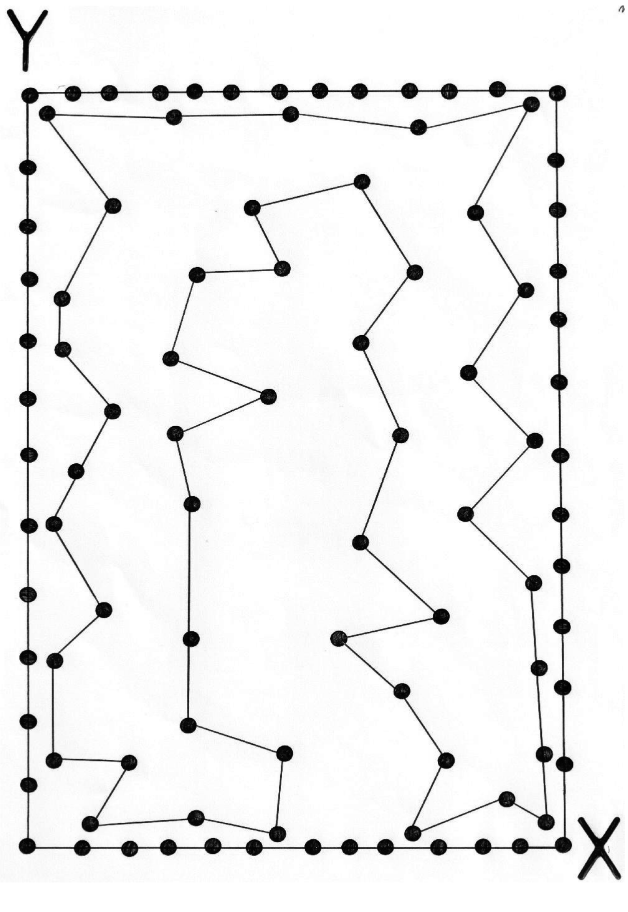


Figure 1: Points on Border are correlated with a Shorter Route than Uncorrelated Points Inside the Border with a Longer Route.

Therefore, if the real data that is being tested for correlation is not correlated (hence it is just random data) then the correlation statistic CTSP=d_1/(median of the random data shortest routes) calculated value should be around one, and not outside of the .9509 to 1.0516 range (more importantly not below the .9509 lower bound). Therefore, Ho of no correlation (the null hypothesis of no correlation) is not rejected. Figure 1 has two sets of n=48(x, y) points. The 48 Points inside the border are just random and a shortest route connecting them is much longer than the shortest route connecting the n=48 on the border. Hence, the x and y variables that produced the border points are highly correlated and CTSP would show that.

CORRELATION IN EXAMPLE TWO

The opinion in the firm is that this new spread sheet of data (please see Data Set Two) must be studied to see if the variables represented by the columns are related or not. Therefore, the data is read into the MSMCO shortest route algorithm program and a shortest route of d=4251.909 is found.

It has a total distance of d=4251.909 and the route goes from

44	18	34	13	15	5	39	2	49	47
45	33	3	36	25	20	11	21	40	28
26	37	38	30	9	1	35	23	43	48
24	14	7	19	6	42	41	24	31	29
16	32	10	17	22	8	46	12	4	44

Therefore, this short route (from this new 49 lines of data taken from eleven new variables) is to travel from point 44 to point 18 to point 34 . . . to point 12 to point 4 and back to point 44 in a complete tour through eleven dimensional space.

Now, because this data is still in the same ranges (0 to 100) and there are also n=49 points, it is possible to use d_2, d_3, d_4 and d_5 from the first problem to calculate the ranges. If n is not 49 and/or the ranges (0 to 100) were different, then new random data would have to be used for the CTSP calculations.

Therefore, the median = (4514.775+4528.076)/2 = 4521.426 and in this case CTSP = $\frac{4251.909}{4521.426}$=.9404 which is below the values where the range of the 12 quotients is (.9509 to 1.0516). Therefore, H_o can be rejected with a slight risk of probably less than 5%.

Table 2: Data Set Two

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}
1	33	100	19	71	45	50	99	39	81	1	30.2
2	24	62	46	48	30	84	16	82	26	57	31.0
3	49	100	72	9	84	95	16	83	42	34	43.6
4	28	34	90	31	16	61	12	48	66	50	59.8
5	72	37	66	32	19	30	9	97	79	100	77.6
6	18	3	6	76	26	36	83	66	67	84	58.8
7	15	33	70	96	50	11	37	18	91	78	61.9
8	62	10	100	100	27	43	16	21	85	1	91.6
9	3	97	18	79	100	68	97	74	97	27	27.8
10	16	35	65	47	62	26	23	2	21	73	58.4
11	100	90	50	3	47	3	8	24	68	66	76.9
12	82	20	100	29	92	88	14	30	49	28	83.7
13	23	94	100	0	58	100	87	26	100	74	93.7
14	69	76	89	41	53	29	12	15	90	55	61.1
15	83	11	100	34	58	49	90	97	79	91	100
16	27	2	49	64	56	60	63	46	21	31	44.2
17	24	21	61	24	33	0	19	3	54	45	69.1
18	86	13	95	53	13	95	28	50	45	92	100.0
19	27	63	13	80	18	2	60	26	37	93	54.7
20	90	100	54	20	35	3	14	55	100	35	60.5
21	98	98	24	11	30	15	29	34	47	42	57.1
22	79	37	91	35	11	24	40	22	93	16	83.8
23	36	58	9	36	61	91	96	75	1	42	42.8
24	71	54	47	72	5	69	48	92	8	54	55.3
25	62	88	74	21	94	4	19	76	72	39	40.7
26	64	41	40	100	83	42	28	100	44	39	28.1
27	84	96	75	34	79	26	78	23	67	38	54.1
28	86	51	18	62	60	26	38	97	30	72	38.2
29	45	4	57	33	34	67	94	62	47	39	60.7
30	39	98	34	66	100	93	72	13	63	33	34.2
31	45	25	35	79	3	68	64	87	49	18	45.6
32	37	15	96	91	79	56	39	45	25	70	65.1
33	56	67	78	8	79	86	15	48	49	78	56.7
34	69	26	97	20	33	99	84	30	60	81	100.0
35	83	79	1	66	23	86	97	13	44	41	63.7
36	44	65	66	4	92	42	57	71	12	44	40.8
37	65	27	36	83	77	66	22	63	69	100	49.2
38	55	62	21	61	92	75	44	5	53	100	51.5
39	60	88	63	54	24	91	8	70	31	100	63.6
40	85	62	9	25	68	20	30	100	31	43	35.3
41	95	23	32	76	25	43	25	79	39	35	51.7
42	48	20	4	65	31	34	28	68	51	64	34.3
43	8	60	5	3	84	37	98	43	11	5	47.4
44	94	38	94	56	23	42	47	2	14	57	98.8
45	62	45	11	0	75	81	22	32	96	64	52.9
46	76	1	93	36	39	26	17	74	81	23	78.7
47	23	50	31	10	64	79	2	38	99	92	56.1
48	36	61	46	99	95	16	79	15	8	15	48.3
49	16	62	29	10	36	76	25	97	41	85	42.1

It turns out that the cost equation function $f(x_1 \ldots x_{10}) =$

$$x_{11} = .00408(x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 + x_6^2 + x_7^2 + x_8^2 + x_9^2 + x_{10}^2)$$
$$+122.363$$
$$-.17136x_1 - .66096x_2 - .1224x_3 - .51408x_4 - .612x_5$$
$$-.38352x_6 - .31824x_7 - .58752x_8 - .47328x_9 - .22032x_{10}$$

fits the data quite well reflecting the idea that variables one through ten are correlated with (and or driving) the x11 variable which is a measure of the project cost.

FINDING THE MINIMUM COST SOLUTION

The firm can now see that because variable eleven x_{11} represents the changes in the firm's cost on a large project it works on frequently, variables x_1 through x_{10} are correlated with x_{11}. There may even be cause and effect (in addition to correlation) if some of the first ten variables are directly driving the x_{11} costs.

Therefore, the firm could use the MSMCO statistical optimization program to find the values for x_1 through x_{10} that produce the x_{11} minimum cost solution as a benchmark. The firm probably does not control all of the x1 through x_{10} variables. However, if they do then they would have the minimum project cost solution.

More likely, they would control some of them, and as the ones they do not control change, then they could use the multi stage Monte Carlo optimization (MSMCO) simulation technique to solve for the new optimal lowest cost solution under the new economic conditions.

THE MULTI STAGE MONTE CARLO APPROACH

Multi stage Monte Carlo optimization (MSMCO) (Conley 1981) is like the regular Monte Carlo or random search for an optimal in that a random sample of thousands of feasible solutions is looked at and the best one is stored. However, this is followed by a stage two where thousands more sample solutions are looked at (centered about the best answer so far) in a slightly reduced search region.

This is followed by a stage three where thousands more sample answers are looked at in a further reduced search region always centered around the best answer found so far. The MSMCO simulation process continues for as many stages as necessary to find the true optimal (minimum cost, etc.) or a useful approximate solution.

Please see Figure 2 for an artistic representation of the MSMCO process closing in on (just sort of funneling into) the region of the feasible solution space that contains the minimum cost solutions.

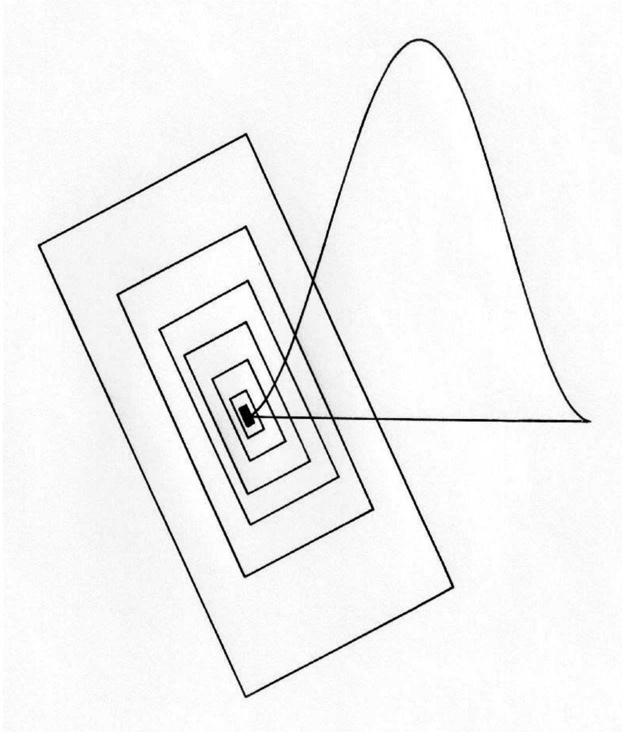


Figure 2: Multi Stage Monte Carlo Optimization Closing in on the Minimum Cost Solution

THE SAINT LOUIS ARCH PARABOLA

The St. Louis Gateway Arch (designed by Eero Saarinen) built on the banks of the Mississippi River in downtown St. Louis, Missouri is an elegant piece of American historical art commemorating the westward push of American settlers across North America in the 18th and 19th centuries.

However geometrically and algebraically the arch could be reviewed as a three dimensional parabola (approximately) that can help us to understand the new CTSP correlation statistic and its shortest route approach. The arch is 630 feet high (192 meters) and 630 feet wide at its base. Imagine that it sits on an x, y plane with the third dimension z axis perpendicular to the land the base of the arch sits on. Let us say that there is a sample of n=20 (x, y, z) three dimensional points that are on or near the arch. Also think of another n=20 (x, y, z) points that are randomly scattered around the arch inside a three dimensional cube that encloses the base. The cube has dimensions 630 by 630 x 630 feet.

Then when a shortest route is calculated connecting the n=20 points on or near the arch, its total distance will be much less than the distance of a shortest route connecting the other n=20 points that are randomly scattered around the arch. Therefore, the variables represented by the data on the arch are correlated (following a pattern) and the other variables are not correlated. Please see Figures 3 and 4 for a two dimensional projection of this idea. This is because correlation among several variables really means that one or more variables are implying or controlling other variables so that the dimensionality is really reduced topologically speaking.

The author has similarly tested hundreds of data sets of four through twelve dimensions (variables) using a k dimensional generalization of the Pythagorean distance formula and the multi stage Monte Carlo optimization shortest route algorithm. It shows overwhelming evidence that correlation in higher dimensions (more variables) leads to shorter shortest routes than for comparable random data shortest routes.

Therefore, the CTSP correlation statistic can and does work with more variables to find linear or nonlinear correlation even when the geometry is hard to picture.

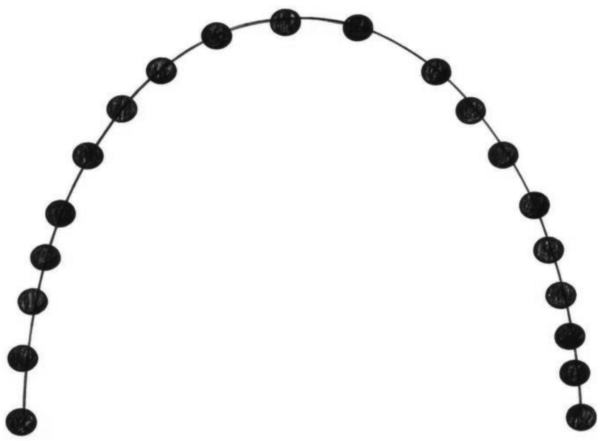


Figure 3: A Short Route, Therefore Correlation

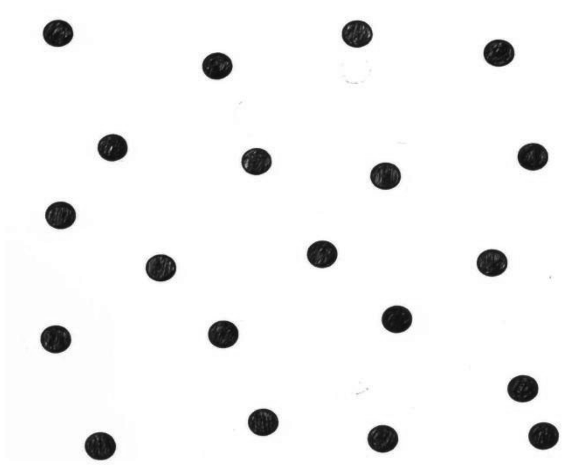


Figure 4: No Correlation, A Longer Route

Please note that Wikipedia (the free internet encyclopedia) presents a photograph of the Saint Louis Arch under the Gateway Arch entry on the internet.

CONCLUSION

Conley in (Conley 1988) found shorter shortest routes (using the MSMCO approach) on a test problem from the mathematics literature mentioned by (Lawler, Lenstra, Rinnooy Kan and Shmoys 1985). They also feature the famous test problems numbers 30, 31, 32, and 33 of 150 and 200 points from the mathematics literature along with the minimum shortest routes ever found using a variety of methods. Multi stage Monte Carlo optimization was then used on those four two dimensional problems and hundreds of shorter routes were found and presented by Conley (1991a), (Conley 1991b) and (Conley 1993).

However, this presentation says don't stop at two dimensions. Travel on shortest routes in three through eleven or more dimensions to useful information. Therefore, the eleven variable problems presented here use shortest route technology in our computer age to find out whether variables are correlated or not. This can help to reduce spreadsheets of data to meaningful information that can be acted upon by managers and decision makers.

If a shortest route is shorter than shortest routes through comparable random data, then the variables that produced the data are correlated. This fact can open up the field of correlation analysis for our computer age. (Anderson 2003), (Hayter 2002), and (Black 2014) present the traditional linear correlation analysis, while (Szarkowicz (1995), (Conley 1986) and (Wong (1996) present examples of multi stage Monte Carlo optimization (MSMCO) applied to engineering, cost control and nonlinear systems of equations as an example of MSMCO's versatility in the computer age.

Furthermore, (Kristensen 1997) reports his European company with 600 delivery trucks saving more than one million U.S. dollars a year on shipping with the multi stage Monte Carlo optimization approach to truck routing. Additionally, (Conley 1988) presents a CEO leaving her home office in Washington D.C. to visit the company's 21 branch offices (on six continents) to discuss business strategy with the branch managers. The shortest route to travel on was a distance of 40, 802 miles on the lengthy journey. As computer speeds and capacities increase the simulation technique multi stage Monte Carlo optimization can be applied to a wide variety of problems.

Additional reading on business statistics problems can be found in (Conley 1981), (Klibanoff, et al 2006) and (McClave, et al 2002).

REFERENCES

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. June 1986. "Production Planning with Nonlinear Cost Considerations." *Cost Engineering*, Vol. 28. No. 6., pp. 16-20.

Conley, W. C. November 1988. "Controlling Costs by Finding the Shortest Route: A Multi stage Monte Carlo Optimization Approach." *Cost Engineering*, Vol. 30, No. 11, pp. 20-24.

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. 2007. "Simulation Correlation and Optimization with Multi Stage Monte Carlo Optimization." *International Journal of Systems Science,* Vol. 38, No. 12, pp. 1013-1020.

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.

Klibanoff, P., Sandroni, A., Moselle, B., Saraniti , B. 2006. *Managerial Statistics: A Case Based Approach.* Thompson SouthWestern, Mason, Ohio.

Kristensen, S. E. 1997. Personal Communication.

Lawler, E., Lenstra, J., Rinnooy, Kan A., and Shmoys, D. 1985. *The Traveling Salesman Problem*. John Wiley and Sons. New York.

McClave, J. T., Benson, P. G., Sincich, T. 2002. *Statistics for Business and Economics*, Eighth Edition. Prentice Hall, Upper Saddle Rivr, New Jersey.

Szarkowics, D. 1995. "Investigating the brachistochrone with a multi stage Monte Carlo optimization." *International Journal of Systems Science*, Vo. 26, pp. 233-243.

Wong, J. Y. 1996. "A note of optimization in integers." *International Journal of Mathematics Education in Science and Technology*, Vo. l. 27, pp. 865-874.

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 and a senior member of the Society for Computer Simulation.

SIMULATION OPTIMIZATION

OPTIMIZING MAINTENANCE WORKS FOR HYDROELECTRIC POWER PLANTS BY GENETIC ALGORITHMS

Vlad Florin PÎRĂIANU, Constantin DRĂGOI, Nicolae VASILIU, Gabriel Lazaro PAVEL
University POLITEHNICA of Bucharest
313, Splaiul Independentei, Sector 6
RO 060042 Bucharest, Romania
E-mail: vlad.piraianu@upb.ro

KEYWORDS
simulation, maintenance, genetic algorithms optimization, hydropower plants,

ABSTRACT
Maintenance is part of the operation of any technological installation. In the energy domain, maintenance is a critical activity in any sector: energy transport and distribution or energy generation whether referred to a hydroelectric power plant, coal or nuclear power plant. The maintenance must be done according to the timeline established by the manufacturer of the equipment or whenever needed. Of course, different events may occur within equipment that will make mandatory the intervention on the machines. This paper aims to present an approach in the planning of the operation of hydroelectric power plants, when maintenance works are required. For the purpose of this article, an optimization program was developed based on genetic algorithms. Two objectives functions of optimization are analyzed in order to take into account the generated energy as well as the income maximization by planning the production correlated with the energy market with the scope of increased revenues. Results of different simulations are presented.

INTRODUCTION

Hydropower energy plays an important role in any modern energy grid system throughout the world, by providing clean and reliable energy while assuring the stability of the system. The ability to provide reliable auxiliary services (e.g. voltage and frequency balancing, reactive power regulation etc.) for any electrical grid is of great importance. Hydroelectric power plants (HPP) have certain advantages compared to any other energy production systems, whether we discuss about coal, gas, nuclear or large scale renewable energy systems like wind and/or photovoltaic. HPP have the ability to store energy in the form of hydraulic potential in their reservoirs (sometimes, even for a longer period of time), they are able to start very fast in order to compensate the need for energy, and stop when overproduction occurs. In order to properly operate any HPP, they must be in an adequate state of operation and all elements of the units and all necessary ancillary services are within normal operation parameters. Therefore, maintenance works are a very important activities that must to be done according to the specific requirements of each equipment. Predicative maintenance is a concept that involves certain activities and tests needed to properly evaluate the state of the machines. In the specific literature there are a number of research directions related to maintenance. The lifetime evaluation and fatigue estimation are considered to be useful tools for the planning of maintenance works (Dragoi et al. 2016, Piraianu et al. 2016). These are considered very important elements in the operation of any HPP and, taken into consideration the main objective of this domain - energy production, it is very important to plan correspondingly the maintenance with the scope of optimize the energy generated. Therefore, the optimal use of the water resource is becoming an important task, even if the HPP must enter into maintenance works, whether the works are planned or, as a result of a failure or defect. While planning these activities, the goal must be to maximize the energy and/or the output power. This article present a modern approach on the optimization of maintenance works for a HPP that involves specific works of the hydro mechanical equipment of the power plant with the main condition to have a certain degree of availability of the other unit(s) of the HPP, by case. A regular HPP could have two or more units of energy generation. The case study considered for this paper, presents an analysis for an existing layout of four HPP, with the schematization presented in Figure 1.

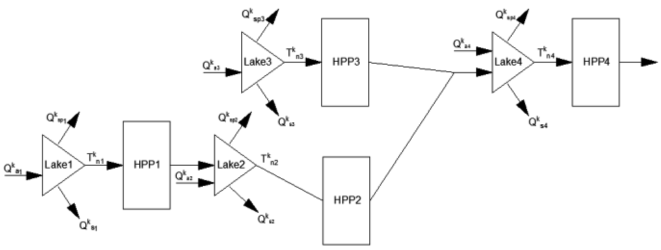


Figure 1: HPP layout configuration

Optimization is a powerful tools used in different sectors of any modern economy. Energy, as a vital sector of a competitive society, has numerous usages of optimization: in power generation, grid operation and transport. The optimal use of resource is a concept with high interest and significance, especially when it comes to the use of primary resources like water. The need to optimize the use of water has generated the evolution of modern optimization techniques such as: genetic algorithms (GA), linear programming (LP) and nonlinear programming (NLP), fuzzy logic (FL), dynamic programming (DP), etc. The genetic algorithms were used for different objectives for water retaining structures (Chau and Albermani 2003). Different techniques have been proposed to be used for the reservoir management, for the operation models of them (Yeh 1985) as well as for the development of optimization models and reservoirs system simulation (Wurbs 1993). The optimization techniques were applied successfully also in renewable energy (Banos et al., 2011). For the purpose of this article, GA will be used to optimize the renewable energy production for a HPP with the objective of increasing the production while assuring the necessary time to carry out maintenance works on different structures and/or equipment. GA optimization has been successfully used in the water reservoir optimization (Chiu et al. 2007, Chang et al. 2010, Mallipeddi et al. 2010, Holland 2012, Ferreira and Teegavarapu 2012). Also, different methods and models have been used to increase the efficiency and performance of a typical GA based program, by penalty functions or by hybrid optimization methods (Deb 2000, Lin et al. 2004 and Nanakorn et al. 2001). This article presents the results obtained by the use of optimization method based on GA combined with operation restrictions as well as penalty function for the exceeding of the level of the lakes, as a method of increasing the efficiency of the algorithm.

HYDROPOWER PLANT OPTIMIZATION

In order to optimize the maintenance works for a HPP, some planning must be done correlated with the type of works required by the intervention. Some works requires less than 12 hours, others up to 24 hours or in some cases even few days, or more. Even in cases of refurbishment of units, which takes few months, is important to plan the production correlated with the capabilities of the reservoirs of the power plants. For different scenarios, an algorithm based on genetic algorithm was developed. In order to evaluate the performance of the algorithm, the layout presented in Figure 1 was considered. The scope was to plan the production based on the availability of the lakes, of the hydrological regime and the state of operation for the units available for energy generation within each HPP. Of course, the unit that needs maintenance, will be not taken into consideration as available for operation. The targeted hydrological regime considered is a relatively reduced one, less than average, due to the fact that when the available flow is higher, then the HPP will run continuously and sometimes the decision to spill the excess of water is taken. The optimization algorithm was developed based on GA use with penalty function in order to adjust the minimum and maximum level of the water in the lakes and to limit the spills. The scope was to use the entire available capacity of the lakes without affecting the performance of the hydraulic turbine and with limitation in the spilled water, as secondary objective, which means losses not acceptable, up to a certain point. However, compared to similar cases of optimization that considered the objectives of maximizing either the generated energy, or power or efficiency, the algorithm developed for this study considers a comparative analysis between energy maximization and income maximization. A similar approach was done by the author for a cascade of hydroelectric power plants (Piraianu et al. 2014) without considering the maintenance works optimization. The maximization of the generated income will be done by planning the production in correlation with the Day Ahead Market (DAM), or the spot energy market. The DAM is the most competitive energy market in Romania, administrated by Romanian Energy Market Operator (OPCOM). Within this market, energy is traded on a daily basis (for the next day) for each hour of the day, with firm quantities and fixed prices. The bidders will submit offers on each trading intervals consisting of pairs of energy volume and price. The market closing price (MCP) is established for each hour in the intersection of prices for demand and supply offers for the predicted energy demand. Once the MCP is established, it becomes the reference price for an interval. Therefore, the scope will be to generate the maximum energy and corresponding income with the available capacity of the units, by planning the production in intervals with the highest prices, as much as possible.

MATHEMATICAL MODELLING

For the development of the mathematical model necessary for the optimization algorithm, the following notations and parameters were considered as shown in Figure 2.

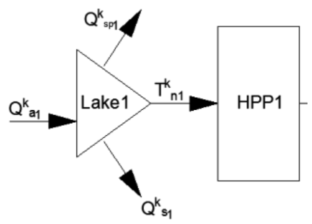


Figure 2: Notations used for each HPP

The main parameters are Q_{a1}^{k} - affluent volume in the lake for interval *k* for the hydrographic basin of the lake; T_{n1}^{k} - effluent volume from lake *n* in interval *k* that is used for power generation; Q_{s1}^{k} - volume of water considered for the servitude (ecological flow and other non-energy purposes like water supply, agricultural, etc.); Q_{sp1}^{k} - volume of spilled water. Furthermore, the mathematic equations describing the volume balance is presented in group of equations (1). The volumes of the lakes for each interval are expressed by the parameter V_i^k , included in the system of equations (1), with *i* representing the lake and *k* representing the interval of the day:

$$
\begin{aligned}
V_1^k &= V_1^{k-1} + Q_{a1}^k - Q_{s1}^k - Q_{sp1}^k - T_{n1}^k \\
V_2^k &= V_2^{k-1} + T_{n1}^k + Q_{a2}^k - Q_{s2}^k - Q_{sp2}^k - T_{n2}^k \\
V_3^k &= V_3^{k-1} + Q_{a3}^k - Q_{s3}^k - Q_{sp3}^k - T_{n3}^k \\
V_4^k &= V_4^{k-1} + T_{n2}^k + T_{n3}^k + Q_{a4}^k - Q_{s4}^k - Q_{sp4}^k - T_{n4}^k
\end{aligned}
\qquad (1)
$$

A number of constrains were considered due to technical limitations and mathematical considerations, like the used flow (with value between the minimum technical flow per turbine and maximum installed discharge per power plant, expressed as volume of water within time step), limitations in the speed of raising and lowering the level in the lakes. The availability of the HPP is induced by the coefficient α that is expressed in percentage, in equation (2), where Q_i is the installed discharge for the entire power plant. For example, if the HPP has two power units and one is not operational, then the coefficient will be equal with 50%.

$$T_n^k = \alpha \cdot Q_i \qquad (2)$$

For the energy calculation is used an equation which depends on the level of the lakes, the net head, hydraulic losses and the net head of the power plant, as well as the used flow (volume) for energy generation, as shown in (3):

$$E_n^k = a_n\left[b_n\left(V_n^{k-1}+V_n^k\right)^{c_n} - d_n\left(\frac{T_n^k}{Q_n}\right)^2\right]\cdot T_n^k \quad (3)$$

The parameters a_n, b_n, c_n, d_n are constants determined for each lakes, and are taking into account the parameters of the lake, configuration of the active volume of water and the capacity curve. In order to assure a higher precision of the algorithm, the entire capacity curve was considered in order to facilitate an analysis on all the available levels within the lakes. In Figure 3 is presented the capacity curve for HPP 1 lake.

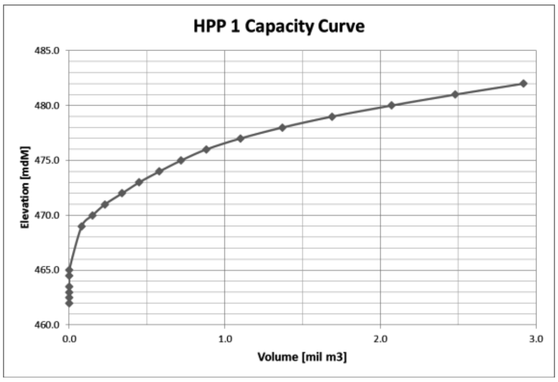


Figure 3: Capacity curve for HPP 1

Maximizing the energy production (objective 1 as declared) of the power plants, the equation (4) is used:

$$\max\left\{Energy = \sum_{n=1}^{4} E_n^k\right\} \quad (4)$$

where energy is calculated according to equation (3). The second objective taken into consideration, represents the maximization of the income, by planning the production in the trading intervals with highest prices. In order to introduce the market price in the equation, the coefficient p_k which represents the prices for each hour of the day.

Figure 4: Market Closing Price (data from www.opcom.ro)

The equation (4) used for the optimization of the energy production is modified according to equation (5) with the pricing element from the Figure 4 in order to achieve the second objective:

$$\max\left\{Income = \sum_{k=1}^{24} p_k \cdot \sum_{n=1}^{4} E_n^k\right\} \quad (5)$$

SIMULATION USING GENETIC ALGORITHMS

The algorithm for optimization based on Genetic Algorithms (GA) was developed under Pascal programming environment. The specific parameters of a typical GA program have been introduced in the program with the possibility to modify them in order to evaluate their influence in the performance of the GA to identify a maximal solution. The main parameters considered that can be selected are the selection type that could be roulette wheel or tournament, normalized geometric sorting; mutation type that can be: boundary or non-uniform, and several crossover techniques. Different simulation were completed to test the capability of the developed algorithm to generate a maximal solution for each objectives considering a different configuration of the availability of units within the HPP layout studied. The results presented in this article were obtained by using heuristic crossover, boundary mutation and normalized geometric sorting. Some parameters of the algorithm that were considered are: mutation probability 0.03, crossover probability 0.8, maximum generation used 24000. Although the trading interval of the DAM is one hour, due to technical consideration of the number of startups per day per each unit, the time step for these simulations is establish at three hours. This means that the operation interval as well as the shutdown periods will be of multiples of 3 hours. A higher number of startups per day will lead to increased fatigue and wear of the components subject to higher stress and strain. Several articles are commenting on this phenomena and are evaluating lifetime (Dragoi et al. 2013). However, the algorithm is designed to have the ability to have the simulation time step modified in case this is required. A three hour time step for simulation means that the variables that represent the solution of the optimization problem, the used flows (volumes) for power generation are eight per each power plant – for a day of production, leading to a total number of variables of this system of 32.

RESULTS

In order to demonstrate the capability of the algorithm, a special maintenance procedure was considered: the verification of the spiral case of the turbine. This is an operation that needs few days to complete and requires a complete empty case, the mount of access equipment and platforms, visual test and nondestructive tests of the metallic structures, etc. The targeted HPP is number 2 from the sketch presented in Figure 1, and it is equipped with two identical Francis turbines that generates up to 110 MW. In order to evaluate the performance of the algorithm the two case studies were done for two hypotheses in order to compare: the full availability of the units and unavailability of the second unit of HPP2. For each study case, simulation were done for each objective – energy maximization and income generation. For these simulations, the same initial conditions were used: the same values for the affluent flows in the lakes, same flows of servitudes, same pricing for the day ahead market, same levels of the lakes at the beginning and at the end of the day, etc. This was done in order to obtain comprehensive information about the planning of the operation. Taken into consideration that the planned maintenance operation is scheduled for approx. five days, the initial level of the lake at the beginning of the day (00:00), is the same at the end of the day (24:00) in order to be able to plan the operation on a daily basis while the maintenance works are completed and the unit is restored to an operation state. An adequate guideline for the variation of levels in the lake was given by restriction of the operations, in order to empty the lake based on the evolution of the prices of energy. Because HPP1 and HPP3 are supplied from the natural rivers, an average flow distribution per each time step of the simulation is considered according to Table 1.

Table 1: Hydrological regime for HPP1 and HPP3

Time step Flow [m^3/s]	1	2	3	4	5	6	7	8
HPP1	35	35	35	35	35	35	35	35
HPP3	100	100	100	100	100	100	100	100

For the financial data of the DAM taken into consideration in Figure 4, from OPCOM website, the average price for the energy for the entire day is approx. 52 EUR/MWh. Of course, for the purpose of the simulations, each trading interval was considered.
The results of the simulations, energy generated [MWh] per power plant, considering the energy maximization and assuming a full availability of the units is presented in Table 2.

Table 2: Energy Maximization – Full availability

Timestep	HPP1	HPP2	HPP3	HPP4	Total [MWh]	Income [EUR]
1	0.00	124.60	0.00	0.00	124.60	3,572
2	0.00	94.09	7.70	32.61	134.40	2,688
3	16.04	34.51	36.08	107.94	194.57	8,042
4	24.42	53.09	38.46	117.60	233.57	15,918
5	5.49	139.37	8.24	50.16	203.26	14,455
6	7.94	81.90	38.46	122.93	251.23	16,748
7	48.25	110.13	38.46	128.13	324.97	20,461
8	34.66	141.90	18.82	112.01	307.39	19,810
Sum	***136.80***	***779.59***	***186.22***	***671.38***	***1773.99***	***101,693***

Table 3: Energy Maximization – HPP2 unit 2 in maintenance

Timestep	HPP1	HPP2	HPP3	HPP4	Total [MWh]	Income [EUR]
1	0.00	113.67	7.44	21.20	142.31	4,079
2	0.00	45.10	0.00	0.00	45.10	902
3	16.04	46.82	36.08	110.21	209.15	8,645
4	24.42	66.92	38.46	120.15	249.95	17,034
5	5.49	147.17	8.24	51.62	212.52	15,113
6	7.94	92.05	38.46	124.80	263.25	17,550
7	48.25	119.09	38.46	129.78	335.58	21,129
8	34.66	148.88	18.82	113.30	315.66	20,342
Sum	***136.80***	***779.70***	***185.96***	***671.06***	***1773.52***	***104,794***

In order to compare the targeted objective, the simulations that takes into consideration full availability of HPP1, HPP3 and HPP4, and one unit from HPP2 in maintenance procedure, is presented in Table 3. To test the capability of the algorithm, the results of the simulations, in energy generated [MWh] per power plant, considering the income maximization and assuming a full availability of the units is presented in Table 4. In order to compare the targeted objective, the simulations that takes into consideration full availability of HPP1, HPP3 and HPP4, and one unit from HPP2 in maintenance procedure, is presented in Table 5.

Table 4: Income Maximization – Full availability

Timestep	HPP1	HPP2	HPP3	HPP4	Total [MWh]	Income [EUR]
1	0.00	0.00	7.44	0.00	7.44	213
2	0.00	0.00	0.00	0.00	0.00	0
3	16.04	0.00	36.08	93.40	145.52	6,000
4	24.42	323.12	38.46	167.39	553.39	37,712
5	42.61	343.77	36.88	174.16	597.42	42,484
6	0.00	108.35	13.22	125.52	247.09	16,472
7	16.75	0.00	30.94	21.08	68.77	4,329
8	34.66	0.00	18.82	85.86	139.34	8,979
Sum	***134.48***	***775.24***	***181.84***	***667.41***	***1758.97***	***116,190***

Table 5: Income Maximization – HPP2 unit 2 in maintenance

Timestep	HPP1	HPP2	HPP3	HPP4	Total [MWh]	Income [EUR]
1	0.00	0.00	7.44	0.00	7.44	213
2	0.00	0.00	0.00	0.00	0.00	0
3	16.04	0.00	36.08	93.04	145.16	6,000
4	24.42	173.26	38.46	139.72	375.86	25,614
5	42.61	173.30	36.88	153.97	406.76	28,925
6	0.00	173.23	14.90	116.26	304.39	20,293
7	16.75	87.03	28.95	46.91	179.64	11,310
8	34.66	173.13	18.82	117.76	344.37	22,192
Sum	***134.48***	***779.95***	***181.53***	***667.66***	***1763.62***	***114,548***

The used volumes of water for the energy generation for HPP2 are presented in chart forms. In Figure 5 are the results presented for the energy generation maximization and in Figure 6, for comparison, are shown the results for the income maximization.

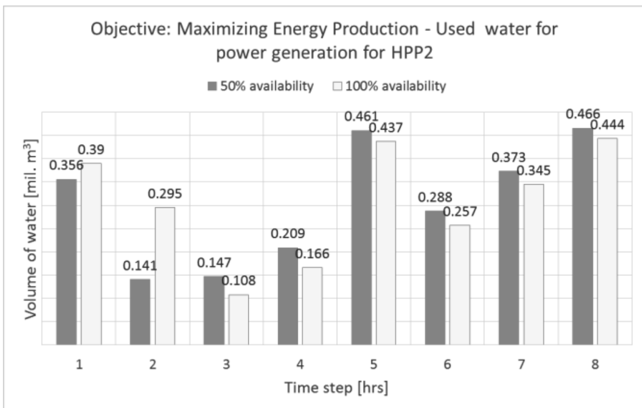


Figure 5: Used water for energy generation distribution per time step – energy maximization

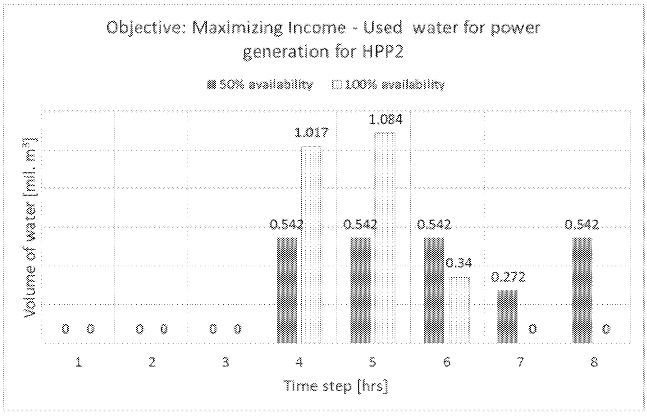


Figure 6: Used water for energy generation distribution per time step – income maximization

CONCLUSIONS

Planning the maintenance works is an important job of the operation activity. The planning has to take into consideration the seasonality of the hydrological regime with strong variations during summer and winter, and high precipitations in autumn as well as in spring when, the melting of snow occurs. Usually, the operation personnel plans important maintenance activities correlated with the weather. However, different maintenance operations cannot be scheduled with high precision, especially when different incidents may occur. For these scenarios, the developed algorithm could be a very good decision support system, to maximize the energy or income during maintenance planning. For example, as shown in the tables 2 and 3, even though one unit is not operational for HPP2, the generated energy is the same. Comparing the income obtained by the two different objectives, it is observed that although the generated energy is sensitively smaller, the revenues obtained by the HPPs is with 9.6% higher when the objective is to maximize income. If full availability of the HPPs is considered, the increase in income is even higher, up to 14.3%. The developed algorithm is a powerful instrument that can be of assistance to dispatchers.

REFERENCES

Banos, R.; Manzano-Agugliaro, F.; Montoya, F.G.; Gil, C.; Alcayde, A.; Gomez, J. (2011). "Optimization methods applied to renewable and sustainable energy: A review" *Renew. Sustain. Energy Rev.*, 15, 1753–1766

Chau KW, Albermani F (2003), "Knowledge-based system on optimum design of liquid retaining structures with genetic algorithms". *J Structural Engineering* ASCE 129(10):1312–1321

Chang, L.C.; Chang, F.J.; Wang, K.W.; Dai, S.Y. (2010). "Constrained genetic algorithms for optimizing multi-use reservoir operation". *J. Hydrology*. 2010, 390, 66–74

Chiu, Y.C.; Chang, L.C.; Chang, F.J. (2007). "Using a hybrid genetic algorithm-simulated annealing algorithm for fuzzy programming of reservoir operation". *Hydrology Process.*, 21, 3162–3172

Deb, K., (2000) "An efficient constraint handling method for genetic algorithms", 186, *pp 311–338*

Dragoi, C., Piraianu V.F., Husar C., Grovu M., (2016). "Lifetime evaluation of Hydropower Plants mechanical equipment for predictive maintenance", *SGEM 2016, Proceedings Paper, pp 465-472*

Dragoi, C., Piraianu V.F., Irimia P.C., (2013). "Lifetime evaluation of the mechanical elements for hydraulic turbines by numerical simulation", *Industrial Simulation Conferences, Eurosis, pp 83-87*

Ferreira, A.R.; Teegavarapu, R.S.V. (2012). "Optimal and adaptive operation of a hydropower system with unit commitment and water quality constraints". *Water Resources Management*, 26, *pp 707–732*

Holland, J.H. (2012)."Adaptation in Nature and Artificial System; The University of Michigan Press: Ann Arbor", MI, USA, 1975, *Energies*, 5 2910

Lin, C.Y.; Wu, W.H. (2004) "Self-organizing adaptive penalty strategy in constrained genetic search". *Struct. Multidiscip. Optim.*, 26, *pp 417–428*

Mallipeddi, R.; Suganthan, P.N. (2010) "Ensemble of constraint handling techniques" IEEE Trans. *Evol. Comput.*, 14, *pp 561–579*

Nanakorn, P.; Meesomklin, K. (2001). "An adaptive penalty function in genetic algorithms for structural design optimization". *Comput. Struct.*, 79, *pp 2527–2539*

Piraianu V.F., Dragoi, C., Vasilescu A., Vuta L.I., Dumitran G.E., (2016). "Lifetime evaluation of the mechanical elements for a Run-of-River Hydroelectric power plant", *SGEM 2016, Proceedings Paper, pp 473-480*

Piraianu V.F., Vasiliu N., Iana Ghe. (2014), "Modelling, simulation and optimization of hydropower plants cascade by genetic algorithms", *14th Middle Eastern Simulation and Modelling Multiconference, Eurosis, pp 90-94*

Wurbs RA (1993). "Reservoir-system simulation and optimization models". *J Water Resour Plan Manage.*, 119, (4): *pp 455–472*

Yeh WW-G (1985). "Reservoir management and operation models: a state-of-the-art review". *Water Resource Res* 21(12): *pp 1797-1818*

Romanian Energy Market Operator, www.opcom.ro

ACKNOWLEDGEMENT

This work has been funded by University Politehnica of Bucharest, through the "Excellence Research Grants" Program, UPB – GEX. Identifier: *UPB–EXCELENȚĂ–2016,* Energy Equipment Maintenance, Contract number 92/26.09.2016 (MECHE).

BIOGRAPHY

VLAD FLORIN PÎRĂIANU was born in Romania on the 4 of March, 1983. He received the degree in hydropower engineering from the University POLITEHNICA of Bucharest in 2007. He has a Ph.D. in hydropower engineering within the same university, and he is working as assistant professor in the Hydraulics, Hydraulic Machinery and Environmental Engineering Department in the Power Engineering Faculty.

STRUCTURAL OPTIMIZATION AND LAY-UP DESIGN FOR COMPOSITE SHELL-LIKE COWLING THROUGH VIRTUAL WIND TUNNEL TESTING AND PEAK STRESS REDUCTION

Sergey N. Shevtsov, Igor V. Zhilyaev, Paul A. Oganesyan, Natalia G. Snezhina and Elena N. Lyashenko
Don State Technical University, South Center of Russian Academy, Southern Federal University
344006, Rostov on Don, Russia
E-mail: sergnshevtsov@gmail.com
Jiing-Kae Wu
National Kaohsiung Marine University
Kaohsiung city, Taiwan R.O.C.

KEYWORDS
Composite structures, wind tunnel simulation, stress-strain analysis, structural optimization, lay-up design.

ABSTRACT

The main requirements for aircraft glass/carbon fiber shell-like cowls, which are exposed to non-uniform air pressure, are the specified mechanical stiffness to withstand the acting air pressure and reduce the level of noise caused by the airflow-induced vibrations at the constrained weight of the part. To increase the stiffness of the cowl the optimization of wall thickness and lay-up of the composite prepreg layers was developed and the optimization results are presented. Our approach assumes the FE wind tunnel simulation of the cowling model that was previously converted from the cowling CAD model to find the spatial distributions of air pressure applied to the shell surface at the different orientation of airflow. Transforming an external shape of the cowling to the elastic shell's model made of transversely isotropic composites we solve the structural mechanics problem to find the stress-strain distribution within the shell surface. At the formulation of optimization problem, the global and peak strain energies calculated within the optimized shell were assumed as the objectives. The wall thickness was varied using the parametric approach based on the parameters of a smooth function describing the thickening of the distributed wall.

INTRODUCTION

The glass/carbon fiber composites are widely used in the design of various aircraft and rotorcraft components such as fairings and cowlings, which have predominantly a shell-like geometry and are made of composite laminates which most often are adopted as quasi-isotropic. The main requirements for such composite parts are the specified mechanical stiffness to withstand the non-uniform air pressure at the different flight conditions and reduce the noise caused by the airflow-induced vibrations with limited weight of the part (Baker et al. 2004).

Most aircraft fairings, radomes and cowlings with open shell geometry, are manufactured using a laminating procedure of glass/carbon fiber prepreg and open-mould tooling. This process involves layering fiber or fabric over one mold surface and generally results in the formation of a laminate on the component surface (Baker et al. 2004; Jones 1998).

The main objective of the present study is the optimization of wall thickness and lay-up of the composite shell-like cowling. Material of the similar parts is usually assumed as quasi-isotropic laminates, having the same Young's modulus Y and same Poisson's ratio ν in all in-plane directions, but are not normal to the plane. Such quasi-isotropic, or more precisely, transversely isotropic laminates, are produced when the total number of unidirectional layers is $n \geq 3$; all the layers have equal orthotropic elastic constants and equal thickness; the orientation of the k th layer of a n-layered laminate is taken in the form $\theta_k = \pi \cdot (k-1)/n$ [2].

In recent years many scientific and engineering groups have developed a variety of approaches and numerical methods for solving the problem of composite structures optimization. Most of these approaches and optimization tools are developed by the key specialists and scientific groups (Ostergaard et al. 2011; Guillermin. 2001) and are presented in the studies (Gillet et al. 2010; Thomas et al. 2000; Gyan et al. 2012). As the objective or merit function, most optimization methods accept the highest stiffness to-weight ratio, maximization of the first natural frequency or minimization of the equivalent stress due to various stress hypotheses at the constrained part volume (weight), and with the restriction on the optimized external or internal shape.

In order to simplify the optimization problem by decreasing the number of design variables some parameterization of the geometry part as an analytic or pointwise function can be implemented (Gyan et al. 2012; Shevtsov et al. 2016).

This paper describes the structural optimization of the rotorcraft radar cowling with open shell geometry and undergoes the spatially distributed pressure from the airflow at the different flight modes.
The pressure distribution on the shell's surface is obtained by the solution of the CFD problem. Then we formulate the structural dynamics problem where the founded pressure field loads the shell. We parameterize the problem by adding an auxiliary sphere with a varied radius and coordinates of the centre, which are used as the design variables. The intersection of the shell and sphere defines the area where local wall thickness differs from the original values. The purpose of the optimization algorithm is to minimize the total and peak elastic strain with limited weight of the part. All stages of the structural optimization problem solving, the result of which is the distribution of shell walls thickness, were implemented in Comsol Multiphysics 5.1 environment.

VIRTUAL WIND TUNNEL TESTING

The studied rotorcraft composite cowling is mounted to the stiffer part of airframe. Such a cowling is exposed to a transient action of airflow, which can cause undesirable mechanical deformations, vibrations and noise. In order to determine the spatial distribution of the air pressure on the shell surface we neglected the fluid-structure interaction, and assumed that the shell is a non-deformable rigid body, and its external surface has the shape of the studied shell (see Fig. 1). All the dimensions were assumed to be small as compared to the dimensions of airframe. To avoid undesired airflow disturbance when tested in a wind tunnel, the geometry of the studied part after its import from the cowling's CAD to Comsol Multiphysics software, has been modified, so, that all the side surfaces have been extruded to the distances much greater than the real dimensions of cowling.
The problem has been formulated by means of the Reynolds averaged Navier-Stokes (RANS) equations (Wilkes et al. 2012). The $k-\omega$ turbulence model with two additional equations relative to the turbulent kinetic energy k and the dissipation per unit turbulent kinetic energy ω, enclosed the RANS equations, which simulate the transport of a turbulent kinetic energy k and ω.
In order to determine the distribution of the air pressure and the velocity field along the cowling's wall, the transient statement of the turbulence flow problem has been formulated. The airspeed in the inlet started from zero, slowly increased and reached its steady values of 83 m/s (cruise speed) and 167 m/s, which corresponds to the maximum speed of the helicopter at the oncoming wind gust. An outlet boundary condition was assumed to be $p=0$, whereas conditions in the form of the wall functions

$$\begin{cases} \mathbf{u}\cdot\mathbf{n}=0 \\ u_{\tan}=\mathbf{u}-(\mathbf{u}\cdot\mathbf{n})\mathbf{n} \end{cases}, \qquad (1)$$

where $\mathbf{n}$ is the unit normal vector to a wall, were used on the walls of the wind tunnel and also on the surface of cowling that are assumed as absolutely stiff (see Fig. 2, a). The cowling FE model was placed into the semicircular tunnel's section with different model orientation relative to corresponding airflow directions. The finite-element mesh consists of near 90,000 tetrahedron elements (see Fig. 2, b). The problem solving results for the blowing angles in the range ±10 degrees with the step equal to 2.5 degree, are the spatial distributions of the pressure field, which acts on the model surface (see Fig. 3).
These results have been stored in the form of text files for further use at solving the structural mechanics problem for the optimized shell-like cowling to identify the global elastic energy and peak stress within the shell.

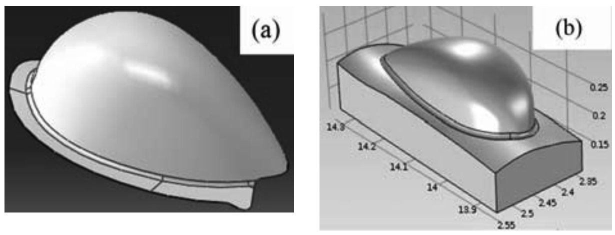


Figure 1: Cowling geometry imported from the CAD model (a) and its transformation to the external surface of a stiff 3D body (b)

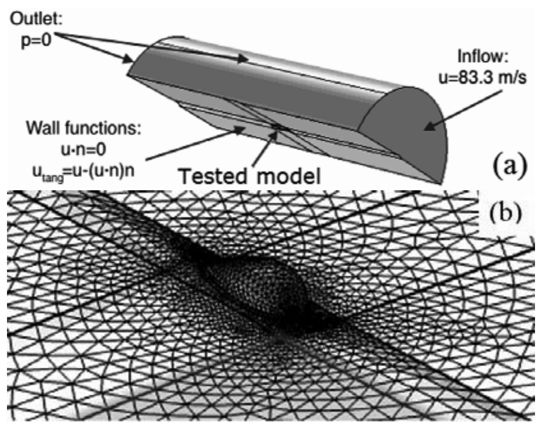


Figure 2: Geometry (a) and FE mesh (b) at the virtual wind tunnel testing of the cowling

THE FORWARD PROBLEM OF STRUCTURAL MECHANICS FOR THE OPTIMIZED COWLING

At the stage of determining the stress-strain state of the studied shell, we used two different materials: transversely isotropic with the in-plane elastic module E_{pl}=15 GPa, normal to plane elastic module E_z=8 GPa, out of plane shear module G_{pl_z}=3 GPa, and two Poisson ratios $\nu_{pl}=0.8$, $\nu_{pl\_z}=0.25$ respectively, and an isotropic one, which have only two independent elastic constants E_{pl}=15 GPa and $\nu_{pl}=0.8$. The density for both materials was accepted as $\rho=1850\,kg/m^3$.

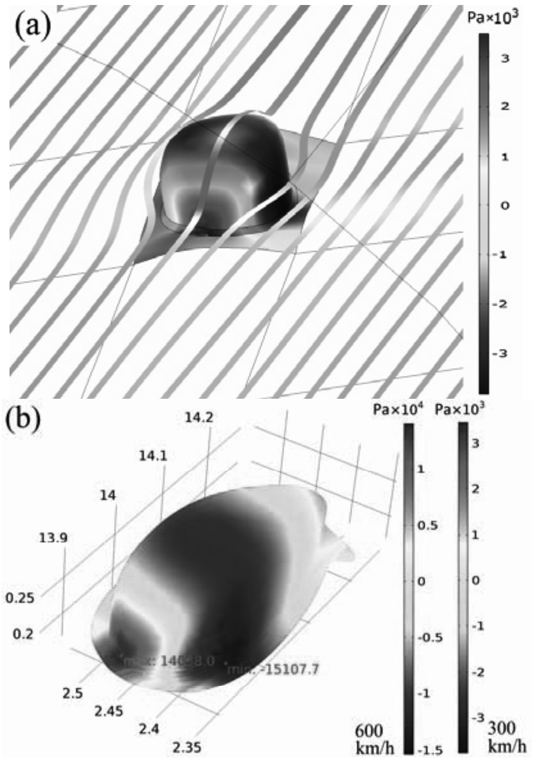


Figure 3: The examples of streamlines map (a) and pressure distribution on the model surface (b)

All these properties were preliminary determined experimentally. The finite element formulation of the forward problem uses the elastic shell model for the cowling structure (see Fig. 4) with the uniform wall thickness of 3 mm. Its geometry was converted from the CAD model. We formulated and solved the problem of linear elasticity without taking into account large displacements. At this stage of investigation we established most stressful case of the incident airflow orientation. This case corresponds to the blowing angle -10 deg, at which the total strain energy exceeds this value for other blowing angles by 30-50%. Our simulation results demonstrated a significant non-uniformity of von Mises stress distribution along the shell surface for both modeled materials. It requires the local thickening of the cowling's wall to eliminate the unsafe stress concentration.

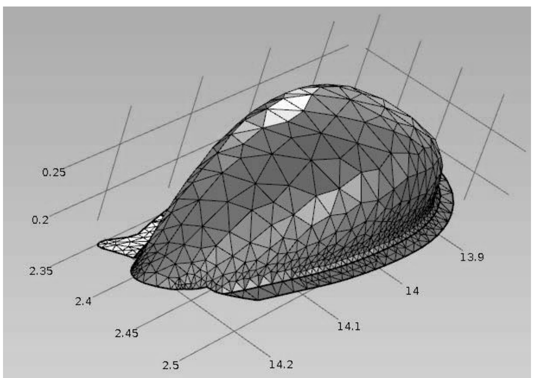


Figure 4: Finite-element meshing of the modeled shell at its stress-strain structural analysis

STRUCTURAL OPTIMIZATION OF THE COWLING

Our attempts to use known SIMP topology optimization methods with the softened penalty function to optimize the wall thickness of the cowling have not been successful due to the formation of stillbirths structures and very low numerical stability. We proposed the parameterization of the problem, which allows the use of only 4 design variables. This involves the introduction of a fictitious sphere, intersection of which with the shell surface defines the region with the modified thickness (see Fig. 5). The spatial distribution of the shell thickness in this area is constructed as a smooth function $K(x_p, y_p, z_p, R)$

$$\begin{cases} h = h_0(K+1) \\ K = \begin{cases} 0, D > R \\ \left((\mathrm{R} - \mathrm{D})/\mathrm{R}\right)^2 \end{cases} \\ D = \sqrt{(x - x_p)^2 + (y - y_p)^2 + (z - z_p)^2} \end{cases} \tag{2}$$

where $h_0 = 3\,\text{mm}$ is the wall thickness of the original shell, x_p, y_p, z_p and R are the coordinates of center and radius of the auxiliary sphere, respectively. The total elastic energy of the deformed shell, which was loaded by the pressure field acting at the blowing angle -10 deg, has been accepted as the objective function. Weight of the optimized shell was constrained by the value 1 kg. The built-in Comsol's optimizer implemented the cyclic call of the forward problem. The optimization results have been stored in the form of tables and text files, which contain the spatial thickness distributions for each set of design variables. By using the guess values for the design variables adopted from (Shevtsov S. et al. 2016) the optimization process was converged after 50-70 iterations.

The best optimization result for the cowling made of transversely isotropic laminate is presented in Fig. 6, which demonstrates the spatial wall thickness distribution. This optimum design reduces the peak von Mises stress to 2.7 times, total elastic energy – to 2.4 times, peak strain energy – to 7.5 times. It is worth to point out that reducing of the stress intensity for the transversely isotropic laminate is 10-15% more efficient relative to the isotropic one.

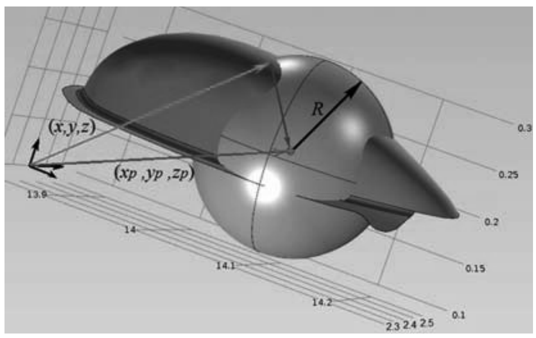


Figure 5: Defining the design variables for the shell optimization problem

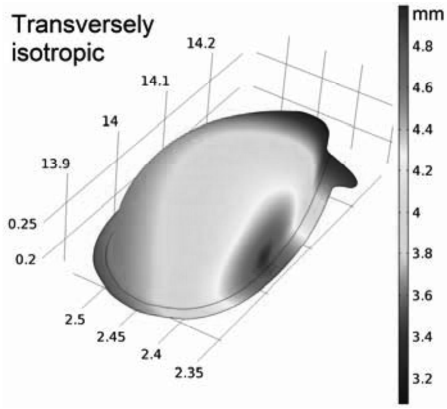


Figure 6: Spatial wall thickness distribution for the optimized shell-like cowling (colorbar denotes the values of wall thickness in mm)

A comparative analysis of the first 6 vibration eigenfrequencies of the original and optimized cowlings demonstrated the increases in these frequencies by 15-20% after optimization. Another optimization result is disappearance of some natural vibration modes.

We also tested the optimum set of design variables, which were obtained for the airflow angle -10 deg, for the other blowing cases to prove that the obtained optimum results are suitable not only for one airflow direction. This examination confirmed the efficiency of the optimum solution for all studied operating conditions, although the decrease of peak stress at these conditions does not exceed 1.5 times.

The obtained wall thickness distribution has been exported to MATLAB® and further used in the composite prepreg lay-up design that can be used during the cowling manufacturing.

DESIGN OF LAY-UP AND DRAPING ANALYSIS

In order to unwrap the shell into the plane surface and to start the lay-up process we specify two kinds of perpendicular geodesic lines lying on the shell surface. The origin of the new coordinates frame, which defines these lines, is placed on the highest point of the shell. To define the dimensions of the pattern of prepreg, the original curved surface of the shell should be flattened by approximation of the two sets of scattered points by the interpolation functions, which use the triangulation algorithm.

Our deployment algorithm includes a cyclic incision of the sloping shell surface by the planes, which pass through the vertical axis of new coordinate frame and are rotate around this axis with angular step $\Delta\varphi = 5$ degrees (see Fig. 7, a). The length of the section line $z(x(r,\varphi), y(r,\varphi))$ is calculated using the well known formula

$$l(r,\varphi) = \int_r \sqrt{1 + \left(z'_r(r,\varphi)\right)^2}\, dr \tag{3}$$

where the polar coordinates r,φ are determined in the area occupied by the shell. The shape features (slightly sloping walls) of the shell guarantees the unlikeness of $z(r,\varphi)$ for all the integration paths, that allows to describe the wall thickness as a function of the new coordinates $h(r,\varphi) \to \tilde{h}(l,\varphi)$ (see Fig. 7, a).

The contour lines for the thickness function $\tilde{h}(l,\varphi)$ in Figure 7, b, are shown with the step of 0.25 mm thickness, which corresponds to 2 prepreg's layers. Hence, 2 layers of fabric oriented according to the scheme [0^0; 90^0; $+45^0$; -45^0] can be placed within the areas surrounded by each two adjacent contour lines. This pattern makes it to easily optimize the final stacking sequence by the ply shuffling procedure. Due to the complexity of the lay-up pattern more than three knots must already be pinned onto the mould surface to fix the patches of the fabric in the exact placement. The last figure (see Fig. 7, b) can suggest the correct positions of these knots.

The obtained lay-up patterns for isotropic and transversely isotropic materials are very similar, while the second model of composite material, which is realistic, gives more reliable results.

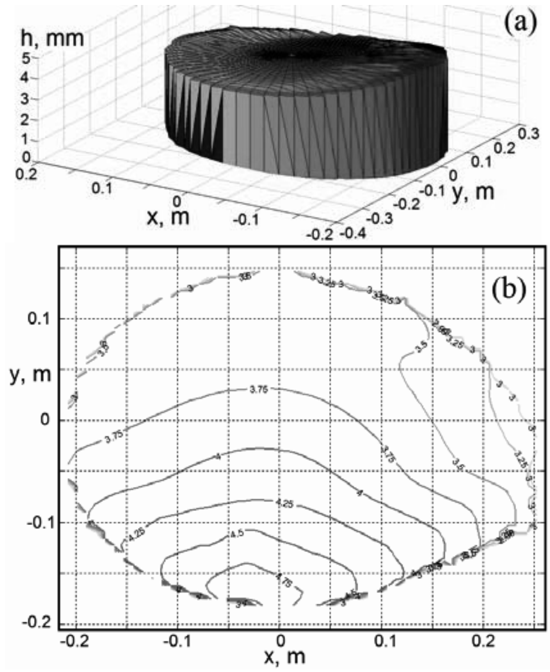


Figure 7: Unwrapped wall thickness distributions within the surface of optimized shell made of transversely isotropic material (a) and corresponding patterns of lay-up (b).

CONCLUSIONS

The structural optimization problem for the aircraft shell-like composite structures was solved using Comsol Multiphysics 5.1. Via a CFD study we determine a pressure field which impacts the shell during a highly stressed flight mode. The ascertained pressure distribution is used for numerically solving the forward structural mechanics problem. We propose the modification of the shell structure by a local thickening of the cowl wall to optimize the global stiffness and reduce a global elastic strain with limited total weight. The optimal solution provides a reduction in the global strain energy in 2.5 times, the von Mises stress reduction in 3 times, ~7 times reduction of the peak stress, up to 15% growth of the vibration frequencies on the first natural vibration modes. In addition, we present the mapping method of the lay-up of prepreg.

The proposed approach to the structural optimization problem for the composite shell-like parts has a specific application, but it can be effectively used for the slightly sloped shells made of quasi-isotropic or transversely isotropic multilayered composite laminates. Our results confirm that quasi-isotropic model of material is more realistic, and it allows eliminating the undesirable part deformations and failure at the most stressful critical load scenarios.

Acknowledgements

This work has been partially supported by the National Science Council of Taiwan, R.O.C. (NSC99-2923-E-022-001-MY3), Ministry of Science and Technology of Taiwan, R.O.C. (MOST103-2221-E-022-015), the Russian Foundation for the Basic Research (Grants 15-08-00849, 16-58-52013 MNT_a), and by Russian State Mission 007-01114-16 PR (project 0256-2015-0074).

REFERENCES

Baker, A., Dutton, S. and Kelly D. 2004. *Composite Materials for Aircraft Structures*, AIAA Eds. Virginia.

Gillet, A., Francescato, P. and Saffre, P. 2010. “Single- and Multi-objective Optimization of Composite Structures: The Influence of Design Variables.” *Journal of Composite Materials*, vol.44, 457-480.

Guillermin, O. 2001. “Computer-Aided Design and Manufacturing”. In *Composites. ASM Handbook*, vol. 21, ASM International, Materials Park, OH, 883-899.

Gyan, S., Gangul, R. and Naik, G. 2012. “Damage-tolerant design optimization of laminated composite structures using dispersion of ply angles by genetic algorithm.” *Journal of Reinforced Plastics and Composites*, vol.31, 799-814.

Jones, R.M. 1998. *Mechanics of Composite Materials*, Taylor & Francis, Philadelphia.

Ostergaard, M.G., et al. 2011. „Virtual Testing of Aircraft Structures.” *CEAS Aeronautical Journal*, 1, 83-103.

Shevtsov S. et al. 2016. “Material Distribution Optimization for the Shell Aircraft Composite Structure.” *Curved and Layered Structures*, vol.3(2), 214–222.

Thomas, H.L., Zhou, M., Shy, Y.K, and Pagaldipti, N. 2000. “Practical Aspects of Commercial Composite Topology Optimization Software Development”. In *Topology Optimization of Structures and Composite Continua*. Rozvany G.I.N. and Olhoff N. (Eds.), Kluver Academic Publishers, Dordrecht, Boston, London.

Wilkes, J.O. et al. 2012. *Fluid Mechanics for Chemical Engineers with Microfluidics and CFD*, 2nd ed., Prentice Hall, Westford, MA.

EVALUATION OF A SIMULATION-BASED PLC PROCESSOR OPTIMIZATION

Johannes Kohl, Wolfgang Bauer, Dietmar Fey
Department of Computer Science, Chair of Computer Architecture
Friedrich-Alexander-University Erlangen-Nürnberg (FAU), Germany
{johannes.kohl, wolfgangm.bauer, 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

The increasing complexity of industrial production lines and processes is a challenge for existing programmable logic controllers (PLC). For one the amount of data delivered from sensors to the controller and in addition to that observing production networks of an Industry 4.0 plant are challenges to communication networks as well as the PLCs. Secondly the production machine's complexity and speed are constantly increased. With the introduction of real time Ethernet based communication networks like ProfiNet and EtherCat the bandwidth issue is covered on the communication side. But to keep up with this development PLC computing power and data throughput must be increased accordingly. In this paper we are evaluating a simulation based optimization of a PLC processor architecture that allows a cycle accurate simulation of a superscalar instruction execution to increase the performance of existing programs. Additionally the simulation model which is show takes the required silicon area into account and allows a two objective optimization for speed and size.

INTRODUCTION

Most of the recent PLC designs use embedded general purpose processors in a system on chip (SoC) implementation. The designs can be further divided into Soft-PLC and Hard-PLC architectures. In the case of a Soft-PLC the application source code is compiled for and executed on a general purpose processor. The complex PLC instruction set is translated into sequences of native processor instructions. For highly complex instructions like indirect function calls the instruction sequences reach a length of over 100 instructions. The performance loss of this complex instruction set translation is compensated with a high system frequency and leads to higher energy consumption. The Hard-PLC on the other hand uses an additional acceleration core to execute the PLC application. Ideally each complex instruction is directly processed by the PLC processor without segmentation of the instruction. This enables the core to produce a high instruction throughput with a lower frequency.

Different approaches for hardware acceleration supporting PLC instruction code can be found in respective studies. Simple examples are Boolean module extensions that just speed up logic instruction throughput as presented in (Shuting-zeng and Zhijia-yang 2010) and (Mei-hua et al. 2005). A RISC architecture executes the PLC program in parallel to a general purpose processor computing all other instructions. All these approaches are based on instruction set usage patterns of programs written between 1990 and 2005 and assume that more than 60 % of the executed instructions are Boolean operations (Rho et al. 1995). But as Heller (Heller 2010) shows in her final thesis in 2010 the focus is shifting to more complex programs. These programs contain more flow control, integer and floating point arithmetic instruction and network communication. This shift leads to a higher level of synchronization between the two processors and slows down the system.

Chmiel et al. (Chmiel et al. 2016) present an FPGA based asynchronous dual core PLC with one core dedicated to Boolean instructions and the other one to integer arithmetic. The Boolean core has an up to 3.75 times higher clock cycle and reaches a significant increase in instruction throughput with 5.7 ns per operation. For the synchronization between the two cores a special hardware synchronization system is used. However for programs that constantly transfer data between the Boolean logic and integer arithmetic cores the system can not reach the high Boolean execution speed.

Milik et al. (Milik and Hrynkiewicz 2016) propose a direct compilation of PLC programs into a FPGA based multi-core architecture. The demonstrated *B2A1N* called triple core PLC shows fast execution speed of up to 4.1 ns per instruction and parallel execution of up to three independent instructions. On the down side only a limited part of the instruction set can be used leaving out indirect addressing and looping.

A more advanced SoC based approach is offered by the PLC7100 processors[1] that includes an ARM926EJ processor running an operating system and an additional PLC Processor execution the PLC program. In this system the PLC processor executes the entire PLC program and uses the ARM core only for sophisticated system library calls (e.g. TCP/IP communication). A PLC device utilizing the PLC7100 SoC architecture is the VIPA-SLIO-015[2]. Our architecture picks

[1] http://www.profichip.com/products/speed7-technology/plc-7100/?L=5
[2] http://www.vipa.com/en/products/control-systems/slio/cpus/cpu-

up this approach and focuses on increasing the instruction throughput with processor architecture improvements like caching, branch prediction and out-of-order execution. All extra processor components come with a set of parameters and must be adjusted to one another in order to achieve the highest possible performance. As an example a well designed data cache does not speed up the design as expected if the actual bottleneck is located within the instruction fetch. The question of how to find an optimal and well balanced processor configuration arises. Often designer's experience, best practice and rule of thumb approaches are applied to configure the processor architecture. But in this case a prototype is needed to evaluate the design decisions and at this point fundamental changes to the configuration are costly.
Therefore in this paper we evaluate a simulation based optimization approach with a fully configurable PLC processor model. The focus of the optimization is on the core's execution speed and the necessary silicon area. The principle approach is shown in figure 1. The design space of this processor architecture includes 10^{13} possible configurations. This huge amount of configurations and hour long verification runs do not allow complete design space verification coverage in advance. Therefore a set of tests is added to the evaluation process to identify model configurations containing implementation errors.

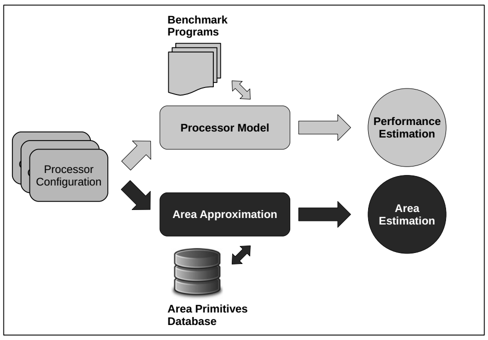


Figure 1: Workflow for execution time and silicon area estimation of a configurable processor model.

This paper is divided into six sections. In this section the idea of a heterogeneous PLC processor architecture and its requirements are shown. In the next section the already accomplished and ongoing work in the field of processor simulation is discussed. This is followed by a PLC processor architecture description in section three. A runtime simulation model verification technique and a processor optimization evaluation are presented in fourth and fifth section. The final section summarizes the results.

PROCESSOR SIMULATION TECHNIQUES

Simulation is an adequate way to cover a huge set of processor configurations. In order to find optimized processor architecture design a simulation based approach is chosen and therefore a suitable simulation environment is required. In academic research the three major processor simulation techniques instruction set simulator (ISS), mathematical modeling and cycle accurate simulation (CAS) are discussed.

Instruction Set Simulators

Many ISS with varying capabilities and purposes are available. Due to the lack of architecture information ISS are usually fast simulators reaching millions of instructions per second. These tools are usually used for software simulations and to verify instruction set changes or additions. Some ISS approaches like ArchC (Rigo et al. 2004) reach into the field of architectural simulation by applying optional processor pipeline model and instruction timings. Kohl et al. (Kohl et al. 2015) show that modeling a PLC processor with ArchC is possible. However the supported pipeline model is not sufficient enough for detailed processor architecture optimization. Other ISS implement an event based modeling approach described in (Luckham and Vera 1995) and reach high instruction throughput but not the required level of accuracy for architectural optimization.

Mathematical Models

The second types of simulators are mathematical processor models based on equations describing the processor behavior. These range from simple mathematical equations for specific processor elements like caches (Berg et al. 2006) to full featured systems of equations describing a complete architecture as shown in (Yashkov 1992) and (Hong and Kim 2009). These complex systems result in a time consuming design phase and are inflexible in case of changes to the design. Additionally the coverage of non-functional properties is not considered and therefore these modeling methods are discarded as well.

Cycle Accurate Simulators

Cycle accurate Simulators provide the highest level of details but at the same time require the longest simulation time. Models of this type show the actual behavior of a processor including all cycles the system is performing nothing due to instruction and data fetch. One benefit is that resource conflicts and pipeline hazards occur and their impact on system performance is covered. A CAS with a complex feature set is Gem5 (Binkert et al. 2011) that even supports area, timing and power estimation is extended by McPat (Li et al. 2009). Unfortunately Gem5 supports only a fixed set of predefined processor architectures that are not suitable for PLC processor simulation with the required level of details. Therefore a PACAS (Kohl et al. 2016) based approach with high abstraction level CAS modeling is chosen and extended with area estimation.

PLC PROCESSOR ARCHITECTURE

Instruction Set

The PLC architecture aims at the native execution of Step7 programs translated into the MC7-Binary format. It is a historically grown format that contains three basic types of instruction formats that are 16 bit and 32 bit and 48 bit long.

configurator/

The shortest on uses an 8 bit instruction identifier followed by either an 8 bit operand or an 8 bit second instruction identifier. As a second format a 16 bit identifier and a 16 bit operand is used. The third format type is 48 bit wide containing a 16 bit instruction identifier followed by a 16 or 32 bit operand. The instruction set allows bit addressing within bytes by including a 3 bit wide bit address segment into the instruction word. Many Step7 instructions encompass multiple execution steps and require complex execution units. To reduce the complexity of the processors design, decrease dependencies between instruction and allow higher clock rates the MC7 binary code is decoded into an architecture internal micro code. Therefore all macro instructions are fragmented into one or more execution steps and each of these steps is covered by one micro instruction.
An example for the micro code fragmentation is the direct addressed assign instruction `= A X.y`. This instruction assigns the result of the last executed binary conjunction to bit `y` in byte `X` of memory area `A`. The operation is fragmented into four steps starting with the memory address calculation that adds the byte offset `X` to the memory area's base address. Secondly the byte at the calculated address is read from the memory. Afterwards the conjunction result is assigned to the bit selected by the instruction's bit address field. Finally the manipulated byte is written back to the memory using the address calculated in the first step.

Pipeline Stages

A pipelined architecture with out-of-order execution is used to reach the required execution speed for the described instruction set. As shown in figure 2 the pipeline is structured in instruction fetch, instruction decode, register reservation, execution, reorder and write back. A more detailed description of each stage and its configuration options is given in the following sections.

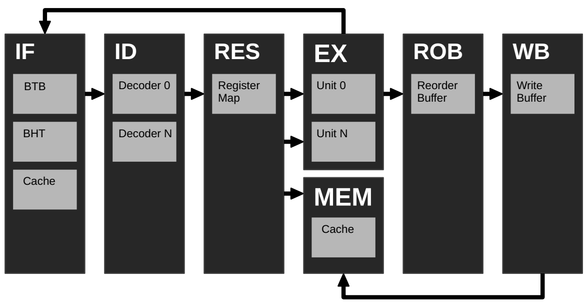


Figure 2: PLC processor architecture pipeline stage and component overview

Instruction Fetch (IF)
The instruction fetch contains the current program counter, a streaming buffer based instruction prefetching device and a memory interface with a transparent instruction cache. Additionally a branch behavior prediction and branch target prediction are performed by a branch history table (BHT) and a branch target buffer (BTB). The next program counter is calculated out of the current program counter, the instruction type at this address and the branch prediction results. In case of a conditional branch with dynamic target both the branch behavior and the branch target are speculated by the branch prediction unit. From this point on until the branch instruction reaches the execution unit the pipeline is filled with speculative instruction. In case of a misprediction of branch behavior or target or both the incorrectly fetched instructions must be swept out of the pipeline. The processor configuration contains seven parameters for the instruction cache allowing changes to the cache size, width, organization and replacement strategy. For the two prediction units BHT and BTB the size and organization are configurable. Moreover the amount, width and depth of the prefetching streaming buffers are controlled by the configuration.

Instruction Decode (ID)
The differences in binary format and an operand length exacerbate the instruction decoding. At first to extract the correct instruction flow from the fetched instruction line an instruction length decoder is necessary. It reads in the instruction line and splits it into up to three instructions, one for each decoder. The first instruction within the line is given by the current program counter. In order to find the next instruction start the length of the first instruction must be decoded. This leads to a length decoder cascade that increases the critical path through with increasing fetch line size. To cover at least one complete instruction with one fetch the minimal instruction fetch line must be 48 bit. To cover all cases and for a simple cache design the architecture's instruction line width is configurable to either 64 bit or 128 bit. After the instruction length is decoded the instruction line is passed on to the actual instruction decoders. A set of multiple parallel instruction decoders is necessary to provide the super scalar execution unit with a constant stream of instructions. The system distinguishes between complex instruction decoding for opcodes requiring a sequence of microcode and simple instruction decoding for opcodes translating into only one microcode. The first decoder in set is always complex and followed by a configurable amount of complex and simple decoders. To prevent disarray within the instruction stream the decoder results must be used in program order.

Register Reservation (RES)
After decode the instructions are passed on to the register reservation and renaming stage performing the Tomasulo-Algorithm. The register resources accessible by the program are mapped onto temporal general purpose registers. A program can access up to 26 registers at a time containing the four accumulators, the two address registers, the data block registers and the current block context. The width of a register depends on its purpose and is either 32 bit or 64 bit. Therefore the temporal register map is usable for both register widths. In addition to the programmable registers the injected microcode uses up to six registers for interim results. Taken together a minimum of 34 registers with a width of 32 bit are required. Therefore the temporal register file size can be configured between 34 and 64 registers. Moreover the amount of speculative program paths that are renamed and

executed is configurable. For each of these paths a snapshot of the register map before the branch is taken that is used to restore the register content in case of a misprediction. After register reservation and renaming the instructions are stored into the issue queues based on the instruction category. Each issue queue is connected to one or two execution paths. If all required input resources are available the issue queue sends the instruction to the correlated execution path. The amount of issue queues, their storage capacity and the issuing order are controlled by the processor configuration. Additionally the amount of rename operations per cycle can range from one up to the amount of issues queues.

Execution Unit (EX)
The execution unit contains the parallel execution paths and the load store memory unit. Each path is able to perform a certain groups of operations either in one cycle, in multiple cycles or in a pipeline mode. The micro instruction set is split into 15 groups of operations based on their functionality. Exceptions are the floating point add, multiply and divide operations that are each in their own group. The mode and the extent of instruction groups covered by each path are defined by the configuration. To generate an operational processor configuration at least each instruction group must be covered by one execution path. Additionally the first instruction path is reserved for branch instruction and checks for mispredictions.

Memory Unit (MEM)
The load store unit contains the first level data cache and handles the cache misses plays a special roll. It is implemented as a pipeline and allows out-of-order load store operations to hide the cache miss penalty. Only if all positions in the cache miss buffer are used no further load store operation is processed. Additionally the stage handles the programs memory write operations through its transparent cache. The configuration of the load store unit includes the data cache size, width, organization and replacement strategy as well as the miss handler size and the data bus configuration.

Instruction Reorder Buffer(ROB)
Reordering of the executed instructions is accomplished in a reorder buffer (ROB). It stores all executed instructions and their results and commits these results back to the permanent register map in correct program order. In case of a branch misprediction all instruction of the falsely chosen program path are removed from the ROB. The model allows the configuration of the ROBs size and the amount of commits to the permanent register map per cycle.

Write Back (WB)
Memory write operations are handled by the write back stage that contains a write buffer to delay writes. This reduces the amount of cache access conflicts between load instructions and write instructions. Write accesses are postponed to a time in program execution with less load operations. In case of a full write buffer load instructions are delayed to execute the store instructions. To ensure data consistency the write buffer content is forwarded to all load instructions. The write back stage configuration contains the two parameters write buffer size and write threshold. The threshold must be reached before the write buffer issues a cache access request.

Area Estimation

Shah et al. (Shah et al. 2014) describe a database based approach that requires configurable RTL descriptions of primitive processor components. For each of the primitives the synthesis results of the required parameter sets are stored in the database. To estimate the processors area the architecture description is partitioned into the predefined primitives and their configurations. Based on this information the precalculated area values are read from the database and added up. We pick up this technique and partition all PLC processor components into configurable primitives for area estimation. The primitives' functionality ranges from simple registers, multiplexers and adders to more profound components like streaming buffers, issue queues and cache miss handlers to entire functional units for data forwarding and floating point execution. Memory components are divided into ROM, RAM and register based storage like the register map and issue queues. To meat the timing requirements the delay path for each pipeline stage and the resulting primitives is constructed. The delays of all components in each path add up to the path delay and the longest path defines the system frequency. The Synopsys Design Compiler and a UMC 110 nm high speed process is used to generate the area and timing data. All components are designed with regard to a maximal path delay of 5 ns to reach 200 MHz clock frequency.

VERIFICATION

To ensure the correct functional behavior of the described model for all possible configurations two verification approaches are applied. The first one is an isolated module test and the second one utilizes system level test benches. On module level the test designs are stimulated with constrained random input signals utilizing the verification methodology UVM (Accellera 2015). Using randomized input stimuli provides finding unexpected issues and misbehavior in the system model where directed and static tests are only capable of test cases thought by the verification engineer. Assertions and reference model data are applied to the output of the design under test to ensure the functional and timing correct behavior. In order to determine when the design is verified sufficiently suitable coverage goals for each module must be defined. Due to the smaller amount of parameters for one module compared to the complete system often all resulting models can be verified. For more versatile modules combinations of extreme and typical configuration values are used for verification.
Advancing to system level tests the huge amount of combinations of sub module configurations makes it impossible to verify each allowed model configuration individually. Therefore a randomized selection of configurations is used for detailed tests. Similar to the module level verification randomized input stimuli are used shown in figure 3. The program generator utilizes the methods described in (Bauer

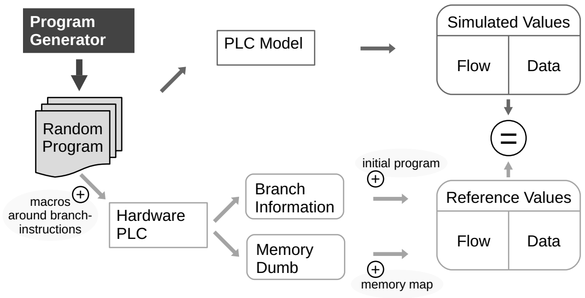


Figure 3: Random program based processor test workflow with comparison of program flow and data results.

et al. 2016) and is able to generate randomized programs with randomized input data based on the instruction behavior description of PLC. Thereby it is possible to trigger different branch behaviors and cache access pattern and stimulate the pipeline with different instruction sequences which is used for cover goal purposes as well. By tracking the usage of the results during program generation it can be ensured that each executed instruction affects either the flow of the program or memory data. So the correct system behavior is assured by comparing the program flow and the memory fields. The reference program flow is extracted from a former generation hardware PLC system by inserting macro code for each branch instruction into the program code which transmits the current processor status via TCP connection. After the program execution the program flow is reconstructed using all received processor status data and the relevant memory fields are extracted from the hardware.

PROCESSOR SIMULATION RESULTS

Due to the 62 available PLC processor architecture parameters and the resulting possible configurations a complete simulation of the entire design space cannot be conducted in a reasonable amount of time. Prior to a large scale design space exploration a model evaluation with randomized processor configuration is performed. During simulation the program flow and result data is extracted from the model and compared to the references after simulation. Measurements show that in this case providing both the program flow and the result data in combination with the final comparison to the references increases the simulation time by 18.7 % in average. A processor simulation parameterized with one particular configuration which executes a benchmark program is counted as an individual. The two criteria chip area and cycles per instruction (CPI) are used as design goals. The results of a 1,000,000 individual processor configuration simulations are shown in figure 4. A processor cluster of 32 AMD Opteron dual core processors required 45 hours to execute the simulations. Each dot in this diagram represents an individual. In addition to that the black colored dots show the resulting Pareto-front. Owing to the high complexity in the processor model description some generated configuration combinations show misbehavior detected by the applied runtime verification. This extends the total simulation time to more than the tenfold but increases the confidence in the reached results.

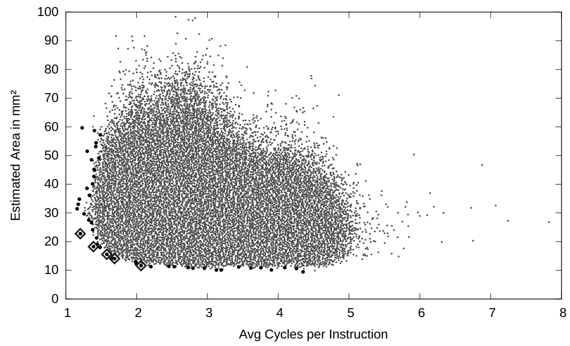


Figure 4: Estimated areas and execution speeds of 1 million random processor configurations.

Due to the long simulation time of the model (up to 10 minutes for one program) tests that cover multiple benchmark programs are impossible in reasonable time. Therefore only a PID controller program is executed by each model instance. To evaluate the quality of the processor configurations found by the random design space exploration additional programs are applied to five individuals on the Pareto-front. The selected five configurations are visualized by a diamond frame in the diagram. For the instruction classes binary, integer and floating point benchmark programs are executed on the resulting processor individuals and compared to a real PLC device. The binary program implements a 256 bit adder, as integer test a quick sort algorithm is executed and the PID program covers the floating point instructions. Extracting the CPI value from the PLC device is not possible. Instead the execution time is used to compare the device with the simulation. To ensure comparability between the simulations and the device the program execution time is extracted from the device by lightweight software tracing. The application code is extended with a system library function block (SFB 7) call that starts a micro second counter at the beginning of the main function and a timer stop at the end. The result of this measurement in comparison to the simulation model is shown in figure 5. To reduce the comparison to only the processor architecture improvements the simulation model is set to the devices clock frequency of 96 MHz.

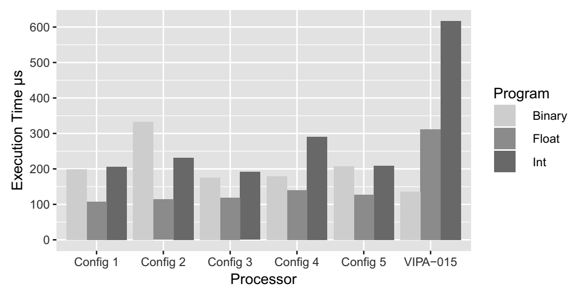


Figure 5: Comparison of five random out-of-order PLC processor configurations to a real PLC device.

The applied optimization method is based on a simulation executing a floating point biased program and produces a Pareto-front optimized for this kind of program. Therefore the integer and floating point performance of the selected individuals exceed the one of the PLC device by between two and six times. The binary program execution time on the other hand is less efficient with up to twice as long execution times than the device. This shows that the program executed during an optimization process must cover the entire instruction set.

CONCLUSION

In this paper the evaluation of a simulation based PLC processor optimization is shown. To increase the architectures instruction throughput out-of-order execution, caching and branch prediction are applied. All these components have their own set of parameters and increase the resulting design space further. Therefore a design space optimization based on a cycle accurate simulation model is considered. This makes it even more difficult if not impossible to reach full simulation model test coverage of all parameter combinations in advance. A reference program flow and result data based lightweight verification scheme is shown that increases the confidence in the optimization results. Furthermore a comparison between a random set of simulated processor configurations and a real PLC device shows that the applied design flow is suitable but requires an adjusted benchmark program.

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

Accellera, 2015. *Universal Verification Methodology (UVM) 1.2 User's Guide*. Accellera SYSTEMS INITIATIVE.

Bauer W.; Kohl J.; Bäsig J.; Rübesam S.; and Fey D., 2016. *Generation of Executable Runtime Constrained Random Programs Functional Processor Verification*. In *ESM 2016 The 2016 European Simulation and Modelling Conference., Proceedings*. 256–263.

Berg E.; Zeffer H.; and Hagersten E., 2006. *A statistical multiprocessor cache model*. In *IEEE International Symposium on Performance Analysis of Systems and Software*. 89–99.

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.

Chmiel M.; Kloska W.; Polok D.; and Mocha J., 2016. *FPGA-based two-processor CPU for PLC*. In *International Conference on Signals and Electronic Systems (ICSES)*. 247–252.

Heller S., 2010. *Entwicklung eines Programmanalysetools als Basis für Prozessoroptimierungsstrategien*. Diplomarbeit, Fakultt Elektrotechnik Feinwerktechnik Informationstechnik der Georg-Simon-Ohm-Hochschule Nürnberg.

Hong S. and Kim H., 2009. *An Analytical Model for a GPU Architecture with Memory-level and Thread-level Parallelism Awareness*. In *Proceedings of the 36th Annual International Symposium on Computer Architecture*. ACM, ISCA '09, 152–163.

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.

Kohl J.; Reichenbach M.; Demel C.; and Fey D., 2016. *A SystemC Based Framework for Cycle Accurate Processor Simulation and Parameter Analysis*. *IFAC-PapersOnLine*, 49, no. 25, 92 – 97. 14th IFAC Conference on Programmable Devices and embedded Systems.

Li S.; Ahn J.H.; Strong R.; Brockman J.; Tullsen D.; and Jouppi N., 2009. *McPAT: An integrated power, area, and timing modeling framework for multicore and manycore architectures*. In *42nd Annual IEEE/ACM International Symposium on Microarchitecture, MICRO-42*. 469–480.

Luckham D. and Vera J., 1995. *An event-based architecture definition language*. *IEEE Transactions on Software Engineering*, 21, no. 9, 717–734.

Mei-hua X.; Feng R.; Zhang-jin C.; Shu-feng K.; and Runguang L., 2005. *IP Core Design of PLC Microprocessor with Boolean Module*. In *Conference on High Density Microsystem Design and Packaging and Component Failure Analysis*. 1–5.

Milik A. and Hrynkiewicz E., 2016. *Distributed PLC Based on Multicore CPUs - Architecture and Programming*. In *Proceedings of 14th IFAC International Conference on Programmable Devices and embedded Systems, PDeS*. 422–428.

Rho G.S.; hoon Koo K.; Chang N.; Park J.; gi Kim Y.; and Kwon W.H., 1995. *Implementation of a RISC microprocessor for programmable logic controllers*. *Microprocessors and Microsystems*, 19, no. 10, 599 – 608.

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.

Shah S.A.A.; Wagner J.; Schuster T.; and Berekovic M., 2014. *A lightweight-system-level power and area estimation methodology for application specific instruction set processors*. In *24th International Workshop on Power and Timing Modeling, Optimization and Simulation (PATMOS)*. 1–5.

Shuting-zeng and Zhijia-yang, 2010. *A high performance architecture design of PLC dedicated processor*. 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE).

Yashkov S.F., 1992. *Mathematical problems in the theory of shared-processor systems*. *Journal of Soviet Mathematics*, 58, no. 2, 101–147.

ENGINEERING SIMULATION

MODELLING MIXED LIQUOR SUSPENDED SOLID AND SUBSTRATE LOAD ON THE BASIS OF WASTEWATER QUALITY INDICES AND OPERATIONAL PARAMETERS OF THE BIOREACTOR: DATA MINING APPROACH

Bartosz Szeląg[1], Jan Studziński[2]
[1] Kielce University of Technology, Faculty of Environmental, Geomatic and Energy Engineering
Al. Tysiąclecia Państwa Polskiego 7, 25 – 314 Kielce, Poland
E-mail: bszelag@tu.kielce.pl
[2] Polish Academy of Sciences, Systems Research Institute (IBS PAN)
Newelska 6, 01-447 Warsaw, Poland
E-mail: studzins@ibspan.waw.pl

KEYWORDS
Support Vector Machines, Mixed liquor suspended solid, Wastewater treatment, Substrate load, Classification Trees.

ABSTRACT

In the present study the Support Vector Method (SVM) is used to analyze the dependency between input variables (quantity and quality of wastewaters at the inflow and operational characteristics of an Activated Sludge Tank (AST) and a result of predicted mixed liquor suspended solid (MLSS) and substrate loads (F/M). Computations revealed, that the highest errors for MLSS are present if only the load of organic compounds susceptible to chemical degradation was included in the input data and lowest when load of coal, ammoniacal nitrogen, suspension and ASCh's operational characteristics were used as the input. Moreover, it appeared that indices of wastewater quality at the inflow to the treatment plant can be simulated on the basis of the measured discharge and temperature of wastewaters and in a result it is possible replacing these measured indices with modeled ones to simulated MLSS and F/M. The lowest errors of predicted substrate loads, computed on the basis of modeled using SVM indices values were obtained for coal, ammoniacal nitrogen, suspension loads and AST's operational characteristics - sludge temperature and pH and methanol dosage.

INTRODUCTION

It is a complex task to operate the wastewater treatment plant (WWTP). That requires maintaining an appropriate level of bioreactor technological parameters to ensure the required degree of contaminant reduction. Observations of the treatment facility operation showed that, due to variations in the quantity and quality of wastewater flowing into the WWTP, to be able to maintain constant process conditions, it is necessary to control on-line the operational parameters in the activated sludge tanks (Zealand and Russell 1984, Chua et al. 2000, Henze et al. 1999, 2002). That is aimed at maintaining constant age of the sludge, and appropriate ranges of substrate load and biomass content. Substrate load (F/M) describes the value of the load of compounds susceptible to biodegradation relative to biomass quantity in the activated sludge tank. In practice, substrate load can be treated as a factor that limits the performance of activated sludge (Jenkins et al. 2004, Barbusiński and Kościelniak 1995). The mode of the plant operation must be adjusted to the task of maintaining the required degree of nitrogen and phosphorus compounds reduction relative to the sedimentation unit size. The reduction degree is specified in the relevant regulations. Consequently, it necessary to maintain symbiosis between microorganisms to ensure good quality of the processes of removal of biogenic compounds. Wastewater treatment methods can be classified as low- and high-load systems (Henze et al. 2002, Mąkinia 2010, Medina and Neis 2007). As regards the former category, F/M value is not higher than 0.05 g BOD/g MLSS·d. Conversely, when aerobic sludge stabilisation takes place during the time, in which newly formed micro-organisms, stored in the cells, are oxidised and highly loaded, the sludge needs anaerobic stabilisation, and F/M value ranges 0.4÷1.5 g BOD/g MLSS·d. As regards typical systems for the removal of carbon, nitrogen and phosphorus compounds, it is recommended that F/M value should not exceed 0.10 g BOD/g MLSS·d and not be lower than 0.05 g BOD/g MLSS·d due to sedimentation problems caused by filamentous bacteria (Madoni et al. 1999, Santos et al. 2015, Eikelboom 2000). Therefore, substrate load must be kept within a specified range to maintain a required degree of reduction of biogenic compounds. Currently, the settings are usually corrected in real time, and also when the bioreactor malfunction occurs, which is not an optimal solution from both economic and environmental standpoint. Therefore, due to random character of inflow rate, and thus also biochemical oxygen demand, anomalous events may occur, in which it is necessary to correct the settings in different parts of the treatment facility to obtain the optimum range of F/M quantities. In view of the above, it is reasonable to try to predict, sufficiently in advance, the values of loads of biodegradable compounds. That allows the staff to increase the load of compounds susceptible to biodegradation by means of adding external supplemental carbon source (methanol, ethanol, leachate, etc.) to the treatment system, or by specifying adequate mixed liquor suspended solid. The latter is conditioned by the degree of recirculation, concentration of recirculated sludge, the amount of excess sludge, and others. To make predictions, statistical or physical-mathematical models are employed. The latter are based on the systems of differential equations that describe runoff (Zawilski and Brzezińska 2013, Gironás et al. 2010) and biochemical reactions that occur in wastewater flowing through the sewers (Hu et al. 2006, Huisman et al. 2000).

Due to the fact that the development of a physical model is very expensive, and numerous input data need to be collected, data mining methods are often used to model the treatment facility (Szeląg and Gawdzik 2016, 2017, Minsoo et al. 2016, Han et al. 2015) and its catchment area (Dellana and West 2009). Those methods are relatively quick to devise and the specialist knowledge of the physics of the phenomenon analysed is not required. At the so-called training stage, the model structure is built, which provides a basis for the prediction of a given parameter. The models include, among others, the following: Artificial Neural Networks (Rojek and Studziński 2014), Support Vector Machines (Guo et al. 2015), Random Forests (Verma et al. 2013), Boosted Trees (Kusiak et al. 2013), k-Nearest Neighbour, etc. For the sake of substrate load simulations, it is necessary to build, at the same time, three separate models to simulate wastewater inflow, biochemical oxygen demand and mixed liquor suspended solid. On the basis of literature review (Wei and Kusiak 2015), it can be concluded that the models developed for the simulation of inflow to WWTP over a time horizon from 1 to 7 days show satisfactory predictive abilities. As regards the simulations of mixed liquor suspended solid, it is necessary to perform costly investigations geared towards determining values of wastewater quality indices (including BOD, COD, TSS, and NH4) at the inflow to WWTP (Bagheri et al. 2016, Güçlü and Dursun 2010). Also, the measurements, often taken online, of the reactor technological parameters (degree of recirculation, recirculated sludge concentration, quantity of returned excess sludge, sludge temperature and pH, etc.) generate extra operating costs. As regards the determination of indices of pollution, biochemical oxygen demand determination is particularly problematic due to the time it takes, i.e. 5 days.

This study presents the methodology for substrate load modelling on the basis of statistical models developed to simulate inflow rate, biochemical oxygen demand and mixed liquor suspended solid. Models for BOD determination were devised on the basis of the results of inflow and wastewater temperature measurements, whereas the determination of the inflow rate was based exclusively on the results of measurements that preceded the modelled value. Due to high costs of continuous measurements of mixed liquor suspended solid, the impact of the load of contaminants, i.e. biogenic compounds (BOD, COD, SS, NH_4^+) and individual parameters of the bioreactor operation (pH, T_{sl}, DO, WAS, RAS), on the prediction of the parameter of concern was analysed in detail. That is important from the engineering standpoint, because it is possible to evaluate the effect of disregarding selected parameters on the results of mixed liquor suspended solid simulations. Additionally, due to technical problems related to the measurements of quality indices, the possibility of simulating those on the basis of flow values and inflowing wastewater temperature measurements was taken into consideration. From the practical perspective, it is an important fact as it ultimately allows the assessment of simulations of both mixed liquor suspended solid and substrate load on the basis of wastewater quality indices computed from basic parameters measured at the wastewater treatment facility.

OBJECT OF INVESTIGATION

The investigations concerned the wastewater treatment plan (WWTP) located in the terrain of the commune of Sitkówka – Nowiny. The plant receives sanitary wastewater from the separate wastewater system of the city of Kielce, the commune of Sitkówka – Nowiny, and a part of the commune of Masłów. 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 pretreated using step screens and aerated grit chambers, with separate grease traps (Fig 1). Next, wastewater is pumped to four primary settling tanks (O.1–O.4), from which it is delivered to the biological unit (bioreactor with separate denitrification and nitrification chamber). After dephosphatation process, wastewater is conveyed to denitrification chamber for the removal of phosphorus compounds.

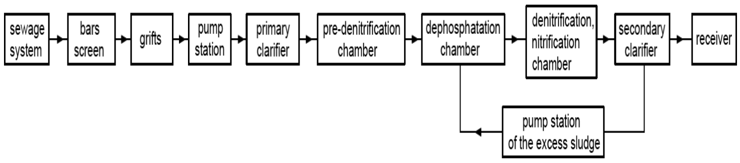


Figure 1. Technological diagram of the Sitkówka – Nowiny treatment plant.

Then, wastewater together with activated sludge is transferred to four secondary settling tanks, from where after clarification, it flows to the receiving water, i.e. the river Bobrza.

METHODOLOGY

In this study, to compute substrate load, three separate statistical models were devised, namely those for predictions of inflow rate, biochemical oxygen demand and mixed liquor suspended solid As to wastewater inflow to the treatment plant, autoregressive model was developed, based on the results of the last measurements that were taken. Simulations of biochemical oxygen demand values were carried out on the basis of measured values of inflow, wastewater temperature, and their combinations from last measurements that preceded the modelled values.

Because of high costs of determining wastewater quality indices and the possibility of limiting the scope of investigations into biogenic compounds, the impact of organic and nitrogen compounds, and that of suspended solids on the error in prediction of mixed liquor suspended solid was thoroughly examined in the analyses. In the next stage, the impact of the reactor operational parameters, including oxygen concentration (DO), sludge temperature (T_{sl}) and pH, degree of recirculation (RAS), the amount of returned excess sludge (WAS) on the simulations of MLSS values was also analysed. As a result, the analyses mentioned above made it possible to specify the impact of independent variables on the simulation error for the parameter of concern. Ultimately, that also allowed a reduction in the scope of measurements of wastewater quality indices and

activated sludge tank (AST) operational parameters, which lowered the costs of investigations.
Due to the fact that substrate load depends on biochemical oxygen demand load, the simulations also included the determination of the BOD prediction error. Because of high cost of measurements of wastewater quality indices at the inflow to WWTP and technical problems related to obtaining series of constant resolution, simulations of quality indices of $C(t)_j$ were considered. That could be done on the basis of measurements of temperature and inflow rate of wastewater delivered to the treatment plant. The analyses quoted above are intended to develop a model form that would make it possible to predict mixed liquor suspended solid on the basis of flow functions $Q(t-i)_j$ describing values of quality indices $C(t)_j$, and also of operational parameters of the bioreactor. To identify variables in the models describing BOD, COD, NH_4^+ and SS, i.e. flow quantities and wastewater temperature from last measurements, the classification trees method was applied. Classification trees were used to determine the so-called predictors importance (IMP) that expresses the impact of independent variables on the parameter of concern, individual wastewater quality indices in this case. Next, the results of simulations, i.e. $C(t)_j$, and also $T_{sl}(t)$ and $Q(t)$, were substituted into the model expressed by formula (1) and substrate load was determined:

$$MLSS(t) = F\{Q(t), T_{sl}(t), C(t)_j, pH, WAS, RAS, DO\} \tag{1}$$

$$C(t)_j = f(Q(t-i)_j) \tag{1a}$$

$$T_{sl}(t) = f(T_{sl}(t-l)) \tag{1b}$$

$$Q(t) = f(Q(t-k)) \tag{1c}$$

Then, computational results were compared with those measured ones. That is important for practical applications, because it is also possible to predict values of quality indices on the basis of Q, T quantities, thus reducing the costs of carrying out continuous determinations. Additionally, from the technical standpoint, it is possible to simulate mixed liquor suspended solid, and thus to control, in appropriate advance, the values of WAS, RAS, DO, and pH. That could contribute to increase in AST operation efficiency and keep the processes taking place in WWTP within the range that ensures the required reduction of biogenic compounds.
In this paper, the possibility of using Support Vector Machines method to predict wastewater quality indices and other parameters was taken into consideration. To make the training process correct, and then to properly assess the performance of the statistical models applied, five-fold cross-validation was used. Simultaneously, the data were partitioned into the training set (75%), and the testing and validating set (25%). Prior to the start of the construction of models, data standardisation was performed by means of min-max transformation expressed by formula:

$$\overline{A_i} = \frac{A_i - \min A}{\max A - \min A} \tag{2}$$

where: $\overline{A_i}$ - normalized, with the min-max method, value of the *i*-th element of set *A*, A_i - value of the *i*-th element of set *A* recorded in measurements, max *A* – maximum value of a single element in the set of parameter *A*, min *A* - minimum value of a single element in the set of parameter *A*.
Support Vector Machines (SVM) cover a group of methods developed by Vapnik (1998) first 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 feature space of a larger size is applied. In the SVM method, in the goal function definition, the error function with insensitivity threshold ε is used (Burges 1998):

$$L^{\varepsilon}(d, y(x)) = \begin{cases} 0 & gdy \quad |d - f(x)| \le \varepsilon \\ |d - y(x)| - \varepsilon & gdy \quad |d - f(x)| > \varepsilon \end{cases} \tag{3}$$

where: ε – assumed model accuracy, x – input vector, y(x) – value of the model output signal expressed by dependence (Burges 1998):

$$y(x) = \sum_{j=1}^{K} w_i \cdot \varphi_i(x) + b = w^T \cdot \varphi(x) + b \tag{4}$$

in which: $w=[w_1,\ldots, w_k]^T$ – transposed vector of weights, $\varphi(x)=[\varphi_1(x),\ldots, \varphi_k(x)]$ – vector of basis functions.
In the SVM method, the aim of training is to properly select a vector of weights w, number of neurons, basis functions and their parameters so that the error function in the form given below could be minimised:

$$E = \frac{1}{2}\sum_{i=1}^{p} L_{\varepsilon}(d_i, y(x_i)) \tag{5}$$

in which: p – number of training pairs (x_i, d_i).
Taking into account complementary variables ξ_i and $\xi_i^{'}$, the problem of SVM network training can be formulated as minimisation of the values of the network weights *w* and of variables ξ_i, $\xi_i^{'}$ written as follows (Burges 1998):

$$\min\{\phi(w, \xi, \xi')\} = \frac{1}{2} \cdot w^T w + C\left[\sum_{i=1}^{p}(\xi_i + \xi_i^{'})\right] \tag{6}$$

under the following functional constraints:

$$\begin{cases} d_i - w^T \varphi(x) - b \le \varepsilon + \xi_i \\ w^T \varphi(x) + b - d_i \le \varepsilon + \xi_i \\ \xi_i \ge 0 \\ \xi_i^{'} \ge 0 \end{cases} \tag{7}$$

Employing Lagrange multipliers (α_i) method, the system of equations described by dependences (7) and (8) can be converted to a dual problem and expressed in the form dependent on the so-called kernel function K (Burges 1998):

$$y(x) = \sum_{i=1}^{N_{sv}} (\alpha_i - \alpha_i^{'}) K(x, x_i) + w \tag{8}$$

in which: N_{sv} – number of support vectors dependent on C and ε. Linear, polynomial and Gaussian kernels are the most frequently used kernel functions. In this study, Gaussian kernels were employed and it was assumed ε=0.01, whereas C was specified at the stage of analyses.
The evaluate the predictive abilities of the mathematical models developed for the study to compute inflow rate, sludge temperature, quality indices, mixed liquor suspended solid and substrate load, the following parameters were used:

- mean absolute error:

$$MAE = \frac{1}{n} \cdot \sum_{i=1}^{n} \left| y_{i,obs} - y_{i,pred} \right| \tag{9}$$

- mean relative error:

$$MAPE = \frac{1}{n} \cdot \sum_{i=1}^{n} \left| \frac{y_{i,obs} - y_{i,pred}}{y_{i,obs}} \right| \cdot 100\ \% \tag{10}$$

where: n – data set size, $y_{i(obs,\ obl)}$ – measured and computed value of dependent variable, $\bar{y}_{obs\ ,pred}$ – arithmetic mean of measured-to-computed value of dependent variable.

RESULTS

On the basis of the results of measurements of the quantity and quality of inflowing wastewater and bioreactor operational parameters, the range of parameter variation was identified. (Table 1). That is essential with respect to operation of the models used in this study because it provides information on the range of applicability of those models. The data in Table 1 show that the quantity and temperature of inflowing wastewater ranged vastly, which led to considerable differentiation in the content of biogenic compounds in the inflow. In the Sitkówka-Nowiny treatment plant, high variation in the load of contaminants conveyed to AST contributed to substantial differentiation in mixed liquor suspended solid (1.19÷5.89 kg/m^3), degree of recirculation (44.6÷167.7%), or dissolved oxygen concentration (0.55÷2.78 mg/dm^3). The fact that methanol was periodically dosed into wastewater (maximum 4.56m^3) confirms a varied content of biodegradable carbon compounds in the inflow to the treatment plant, which affects, among others, the pattern of processes of removal of nitrogen compounds from wastewater. Based on the results of measurements of the wastewater quantity and quality and the bioreactor operation, statistical models for predicting mixed liquor suspended solid were developed.
The models accounted for the possibility of modelling a given parameter on the basis of the load of organic compounds, nitrogen compounds, and solids, and additionally AST parameters (RAS, WAS, pH, T_{sl}, DO). Table 2 presents the results of MLSS simulation using the SVM method. For statistical models (Table 2) devised with the SVM method, C value included in formula (6) ranged 5÷8.
The data in Table 2 indicate that errors in MLSS prediction based on organic compounds load expressed in terms of BOD and COD are equal to MAE=0.792 kg/m^3 and MAPE=20.10%. When the load of ammoniacal nitrogen and solids is accounted for in the models, mean absolute error and mean relative error decrease and amount to MAE=0.733 kg/m^3, MAPE = 17.19% and also MAE=0.700 kg/m^3, MAPE=6.54%, respectively. Taking into account sludge parameters, namely temperature and pH in the model for MLSS simulations leads to a reduction in errors in mixed liquor suspended solid prediction to respective values: MAE=0.672 kg/m^3, MAPE=13.36%, and MAE=0.617 kg/m^3, MAPE=12.65%.

Table 1. Range of variation of parameters describing wastewater quantity (Q), quality (BOD, COD, SS, TN, NH4) and bioreactor operation (T_{os}, pH, DO, MLSS, RAS, WAS, m_{met}, F/M).

Variable	Minimum	Mean	Maximum
Q, m^3/d	32564	40698	86592
T_{in}, ^{0}C	8.4	16.6	20.9
T_{sl}, ^{0}C	10.0	15.9	23.0
pH	7.2	7.7	7.8
MLSS, kg/m^3	1.19	4.26	5.89
RAS, %	44.6	90.70	167.6
m_{met}, m^3	0.00	1.35	4.56
WAS, kg/d	3489	11123	19194
DO, mg/dm^3	0.55	2.56	5.78
F/M, gBOD/g s MLSS·d	0.03	0.07	0.15
BOD, mg/dm^3	127	309	557
COD, mg/dm^3	384	791	1250
SS, mg/dm^3	126	329	572
NH_4^+, mg/dm^3	24.4	7.8	65.9

Table 2. Parameters of fit of computations with SVM models for mixed liquor suspended solid to the results of measurements.

Variables	training		testing	
	MAE	MAPE	MAE	MAPE
Q, COD	0.821	20.772	0.845	20.98
Q, BOD, COD	0.776	19.678	0.792	20.10
Q, BOD, COD, NH_4^+	0.703	17.150	0.733	17.19
Q, BOD, COD, NH_4^+, SS	0.686	16.304	0.700	16.54
Q, BOD, COD, NH_4^+, SS, T_{sl}	0.669	12.950	0.672	13.36
Q, BOD, COD, NH_4^+, SS, T_{sl}, pH	0.600	12.288	0.617	12.65
Q, BOD, COD, NH_4^+, SS, T_{sl}, pH, m_{met}	0.557	11.970	0.580	12.01
Q, BOD, COD, NH_4^+, SS, T_{sl}, pH, m_{met}, DO	0.528	11.223	0.546	11.47
Q, BOD, COD, NH_4^+, SS, T_{sl}, pH, m_{met}, DO, RAS	0.469	10.832	0.472	11.05
Q, BOD, COD, NH_4^+, SS, T_{sl}, pH, m_{met}, DO, RAS, WAS	0.378	9.385	0.392	9.73

Due to a decrease in BOD value at the inflow to WWTP, a drop in mixed liquor suspended solid in AST during abnormal events occurs. That, together with the necessity to ensure large reduction in pollutant loads make it necessary to dose methanol, being an external supplemental carbon source. As a result of taking methanol into account in the statistical model, the values of MLSS prediction errors decrease to MAE=0.580 kg/m^3 and MAPE=12.01%. The data in Table 2 also indicate that when additional parameters, such as dissolved oxygen concentration, or a degree of recirculation are considered in the statistical model based on the load of organic compounds, nitrogen compounds, solids and reactor parameters (pH, T_{sl}, m_{met}) MLSS simulation errors are reduced to MAE=0.546 kg/m^3, MAPE=11.47% and MAE=0.472 kg/m^3, MAPE=11.05%, respectively. The lowest computational error values (MAE=0.392 kg/m^3 and MAPE=9.73%) for mixed liquor suspended solid in the treatment facility of concern are obtained when independent variables are loads of organic compounds, nitrogen compounds and suspended solids and also AST technological parameters (pH, T_{sl}, RAS, WAS, m_{met}, DO).

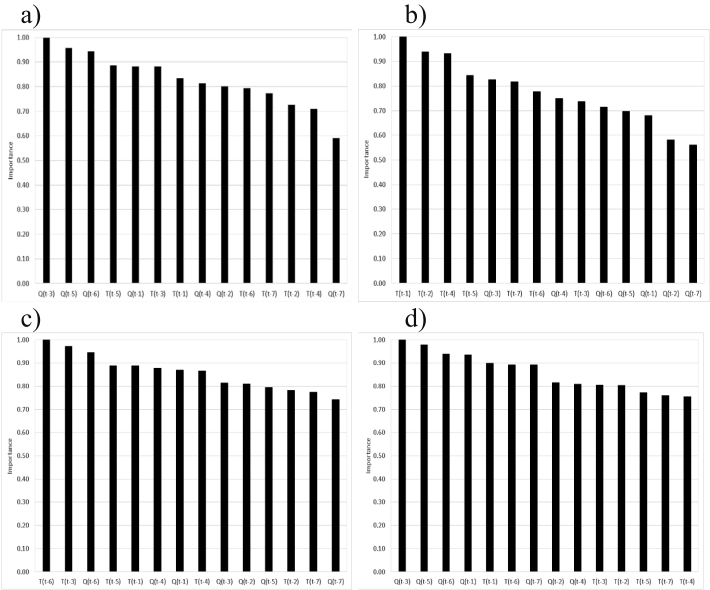


Figure 2. Results of computations of the importance of predictors Q(t-i), T(t-j) for wastewater quality indices: a) BOD, b) COD, c) SS, d) NH_4^+.

To identify the structure of the model described by eqs. (1) and (2), the possibility of simulating wastewater quality indices on the basis of inflowing wastewater temperature and quantity was analysed. Additionally, the analyses also included autoregressive statistical models developed for the prediction of inflow Q(t) and temperature $T_{sl}(t)$ found in eq. (1). Results of computations of importance of independent variables, and the impact of individual predictors on the quality indices of concern and also Q, T_{sl} obtained with classification trees, are presented in Figs. 2 ÷ 5. The data in Figs. 2a ÷ 2d show that as regards the wastewater quality indices, values of importance of individual predictors are greater than IMP>0.8 for variables that describe both temperature and inflow rate.

For instance, as to COD, IMP quantities reach the value equal to at least 0.8 for six independent variables, namely T(t-1), T(t-2), T(t-4), T(t-5), T(t-7) and Q(t-3). In the remaining cases (BOD, NH_4^+), the number of independent variables, for which importance value exceeds IMP = 0.8, is greater than for SS and amounts to ten variables for BOD and eleven for NH_4^+. Analyses carried out for inflow rate and sludge temperature (Table 3) demonstrated that values of importance of independent variables are the highest for the last two measurements preceding the modelled quantity and are equal to at least 0.80.

On the basis of the computations above, models were developed for predicting wastewater quality indices (BOD, COD, SS, NH_4^+). They were based on the last measured values of Q and T that preceded the modelled value. Additionally, statistical models were devised using the SVM method to predict daily inflow and sludge temperature. The results of the analyses are presented in Tables 4 and 5.

The data in Table 4 indicate the statistical models obtained with the SVM method, used for predictions of quality indices of concern, show satisfactory predictive abilities, which is confirmed by the computed error values. For instance, with respect to the model for do BOD simulations, the values of mean errors of prediction of the analysed index are MAE=46.78 mg/dm^3 and MAPE=13.76%. For SS and NH4, they are equal to MAE=41.74 mg/dm^3, MAPE=17.66%, and MAE=4.28 mg/dm^3, MAPE=8.28%, respectively. With respect to statistical models for Q, T_{sl} computations (Table 5), values of mean errors of prediction of the analysed variables are MAE=2588 m^3/d, MAPE=6.34% and MAE=0.77^{0}C, MAPE=5.19%. Using the computed values of Q and BOD, it was found that the mean absolute and relative errors in the simulation of the load of biochemical oxygen demand are MAE=3119 kg/d and MAPE=27.15%, whereas the mean value of BOD daily load is 12339 kg/d. As regards statistical models (Table 4 and 5) developed using the SVM method, quantity C ranged from 6 to 9.

The results of simulations presented above, i.e. quality indices (BOD, COD, SS, NH_4^+) and Q, T_{sl} were substituted into the model described by eq. (1). Mixed liquor suspended solid, and then substrate load were determined. Parameters of fit of MLSS and F/M computational results to measurements can be seen in Table 6. Figures 3 and 4 show the comparison of the results of simulations of the parameters analysed for selected computational variants (where: variant 1 – Q, BOD, COD, SS, NH4, pH, T_{sl}, RAS, WAS, m_{met}; variant 2 – Q, COD; variant 3 – Q, BOD, COD, SS, NH_4^+, pH, T_{sl}.) listed in Table 2.

Table 3. Impact of individual independent variables on Q and T_{sl} values, expressed by means of predictor importance.

Q		T_{sl}	
Variable	Importance	Variable	Importance
Q(t-1)	1.00	T_{sl}(t-1)	1.00
Q(t-2)	0.80	T_{sl}(t-2)	0.96
Q(t-3)	0.67	T_{sl}(t-3)	0.85
Q(t-4)	0.53	T_{sl}(t-4)	0.72
Q(t-5)	0.52	T_{sl}(t-5)	0.63
Q(t-6)	0.50	T_{sl}(t-6)	0.62
Q(t-7)	0.49	T_{sl}(t-7)	0.48

The data in Tables 6 make it possible to state that for the model described by eqs. (1), (1a), (1b), (1c), the values of errors in the sludge concentration predictions are greater compared with the situation when those input data (Table 2) directly represent wastewater quality indices (BOD, COD, NH_4^+, SS), flow, and also sludge temperature and pH measured at the instant t. In the case considered here (Table 6), as previously (Table 2), the lowest error values (MAE=0.439 kg/m^3 and MAPE=10.89%) in MLSS simulations were found for the variables including the loads of organic, nitrogen, and solids in the inflow to the WWTP, and also the bioreactor parameters.

Table 4. Parameters of fit of SVM method simulations to the results of measurements of wastewater quality indices (BOD, COD, SS, NH_4^+) on the basis of Q and T values.

Quality indices	training		testing	
	MAE	MAPE	MAE	MAPE
BOD	42.86	14.94	46.78	13.76
COD	89.77	12.26	90.26	13.02
NH_4^+	4.01	8.25	4.28	8.28
SS	39.02	13.76	41.74	17.66

Table 5. Parameters of fit of SVM method simulations to the results of measurements of Q, T_{sl} values in the activated sludge.

Variables	training		testing	
	MAE	MAPE	MAE	MAPE
Q	2711	6.45	2588	6.34
T_{sl}	0.49	2.16	0.51	2.36

The data in Table 6 demonstrate that the worst computational results concerning mixed liquor suspended solid were produced for the variables that express the loads of organic compounds, nitrogen compounds and solids in the inflow to the WWTP and also activated sludge parameters such as pH and T_{sl}. In this case, the values of errors in MLSS predictions were equal to MAE=0.931 kg/m^3 and MAPE=23.11%. It should be noted that minor (maximum 2 %) variation in the values of mean errors, i.e. MAE=0.821÷0.841 kg/m^3 and MAPE=20.38 ÷ 20.88% was found for statistical models (Table 6) for the simulations of mixed liquor suspended solid that were based only on the load of organic compounds susceptible to chemical degradation (MAE=0.835 kg/m^3 and MAPE=20.73%).

The other models showing low error variation were those developed on the basis of variables describing the loads of organic compounds, nitrogen compounds (MAE=0.841 kg/m^3 and MAPE=20.88%) and, additionally of solids (MAE=0.832 kg/m^3 and MAPE=20.66%), and also the reactor operational parameters such as pH, T_{sl} and the amount of methanol dosed (MAE=0.821 kg/m^3 and MAPE=20.38%).

Additionally, the data listed in Table 6 show that the lowest error values in substrate load simulations (MAE= 0.0174 gBOD/g MLSS·d and MAPE=26.63%) were observed for the data set for the variables describing the loads of BOD, COD, SS, NH4 and the parameters of the bioreactor operation, such as pH, T_{sl} and the quantity of the methanol dosed to provide an external source of carbon. Conversely, the highest values of errors in F/M prediction, equal to MAE=0.0205 gBOD/g MLSS·d and MAPE=32.71% were found when independent variables of the parameter of concern were loads of organic compounds, nitrogen compounds and solids. It should be noted that small (only 1 %) variation in the values of mean absolute error was observed for statistical models developed on the basis of variables describing the loads of organic compounds, nitrogen compounds and solids, and also the reactor operational parameters including pH, T_{sl}, m_{met}, RAS, WAS and DO (MAE=0.0187 gBOD/g MLSS·d and MAPE=29.3%).

Table 6. Comparison of the parameters of fit of MLSS and F/M computational results obtained on the basis of eq. (1) to measurements.

Variables	MLSS		F/M	
	MAE	MAPE	MAE	MAPE
Q, COD	0.835	20.73	0.0189	28.47
Q, BOD, COD	0.804	19.96	0.0188	28.22
Q, BOD, COD, NH_4^+	0.841	20.88	0.0196	30.32
Q, BOD, COD, NH_4^+, SS	0.832	20.66	0.0205	32.71
Q, BOD, COD, NH_4^+, SS, T_{sl}	0.792	19.66	0.0191	29.60
Q, BOD, COD, NH_4^+, SS, T_{sl}, pH	0.931	23.11	0.0200	28.38
Q, BOD, COD, NH_4^+, SS, T_{sl}, pH, m_{met}	0.821	20.38	0.0174	26.63
Q, BOD, COD, NH_4^+, SS, T_{sl}, pH, m_{met}, DO	0.718	17.82	0.0186	28.51
Q, BOD, COD, NH_4^+, SS, T_{sl}, pH, m_{met}, DO, RAS	0.560	13.90	0.0180	27.76
Q, BOD, COD, NH_4^+, SS, T_{sl}, pH, m_{met}, DO, RAS, WAS	0.439	10.89	0.0187	29.30

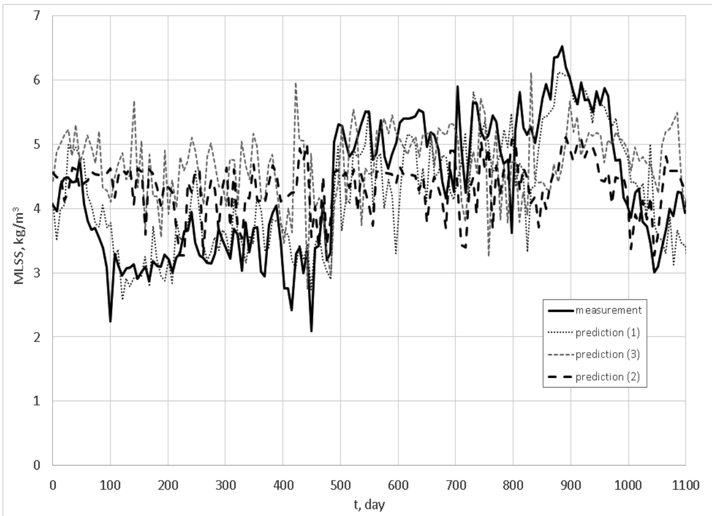


Figure 8. Comparison of the results of mixed liquor suspended solid measurements and computations with the models 1, 2 and 3 developed using the SVM method for the period analysed.

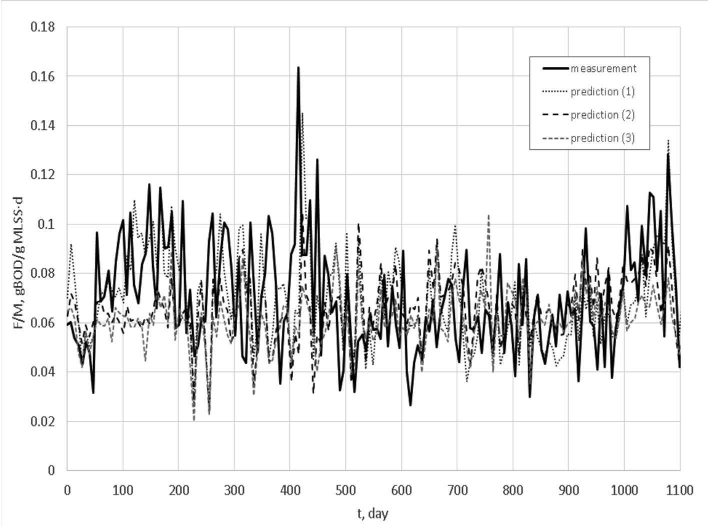


Figure 9. Comparison of the results of substrate load measurements and computations with the models 1, 2 and 3 developed using the SVM method for the period analysed.

The same was true for the models based exclusively on COD load (MAE=0.0189 gBOD/g sMLSSd and MAPE=28.47%). In the remaining statistical models, the variation in the mean absolute and relative errors in the data set was not wide, as values changed within a small range MAE=0.018 ÷ 0.020 gBOD/g sMLSSd and MAPE=27.76 ÷ 30.22%.
Taking into consideration the values of errors in MLSS prediction, it is possible to state that the dependence expressed by eq. (1) can provide a basis for computations of the technological parameter of concern in practical applications. As regards statistical models developed to simulate substrate load, based on the loads of BOD, COD, SS and NH_4^+ and also the reactor operational parameters, the values of the mean absolute and relative errors ranged as follows: MAE=0.0174÷0.0205 gBOD/g MLSSd and MAPE =26.63 ÷ 32.71%.

CONCLUSIONS

On the basis of computations performed for the study, it can be stated that the accuracy of prediction of mixed liquor suspended solid is substantially affected by both wastewater quality indices (BOD, COD, SS, NH_4^+) and also the bioreactor operational parameters. The analyses demonstrated that the highest values of mean errors in MLSS simulations were found for the statistical models devised using the SVM method, which were based on the load of biogenic compounds in the inflow to the treatment facility. Lower values of errors in mixed liquor suspended solid predictions were obtained for the models based on the load of inflow pollutants and also AST technological parameters such as RAS, WAS, pH, T_{sl}, DO, $m_{met.}$.
Additionally, computations performed using the statistical models developed with the SVM method demonstrated it is possible to simulate the values of wastewater quality indices on the basis of quantities of wastewater flow and temperature obtained from last measurements that preceded the predicted value. From the practical standpoint, that is essential as it offers the possibility of making predictions, and thus controlling the bioreactor operational parameters in appropriate advance, which is necessary to optimise the operation of the treatment facility. The analyses conducted for the study show high potential of the derived dependence for mixed liquor suspended solid simulations. The dependence is based on the operational parameters (WAS, RAS, pH, DO, etc.), measured on-line, and on the values of wastewater quality indices (BOD, COD, SS, NH_4^+) obtained from computations using models relying on Q and T values. In this case, the highest values of errors in MLSS simulations were found for the model taking into account the load of biogenic compounds in the inflow to WWTP and the reactor parameters, namely pH and T_{sl}.
In practical applications, the statistical models developed for the prediction of mixed liquor suspended solid allow limiting the scope of measurements of quality indices and the bioreactor operational parameters. The models also permit to assess the influence of simplifications on the simulation results. A relevant practical aspect of the analyses conducted for the study is the possibility of making predictions of mixed liquor suspended solid, and consequently, of the substrate load of activated sludge on the basis of the determined flow functions that describe individual wastewater quality indices. That is of key importance, because in the study analyses, the control and optimisation of the values of F/M and MLSS in the WWTP can be based exclusively on the results of inflow and wastewater temperature measurements and AST operational parameters selected by the facility operator.
Substrate load computations performed for this study, based on the results of simulations flow rate, and the values of biochemical oxygen demand and mixed liquor suspended solid demonstrated that for the SVM–based models, extreme values of MAE and MAPE errors did not differ more than 17.81% and 22.83%, respectively. The results indicate a relatively narrow range of variation in the absolute error and relative substrate load, despite much differentiation in independent variables and predictive abilities of models for mixed liquor suspended solid simulations. The lowest error values were obtained for the model, in which the input data were BOD, COD, TSS and NH_4^+ loads and reactor parameters (pH, T_{sl}, m_{met}). Conversely, the highest error values were found for the model based exclusively on the load of biogenic compounds in the inflow to the WWTP. Considering prediction accuracy with respect to BOD load and MLSS value, it can be concluded that with respect to substrate load, the load of biodegradable carbon compounds was decisive for the error in the F/M value simulations.

REFERENCES

Bagheri, M.; S.A. Mirbagheri; M. Ehteshami; Z. Bagheri Z. and A.M. Kamarkhani. 2016. “Analysis of variables affecting mixed liquor volatile suspended solids and prediction of effluent quality parameters in a real wastewater treatment plant.” *Desalination and Water Treatment*, Vol. 57, No. 45, doi: 10.1080/19443994.2015.1125796.

Barbusiński, K. and H. Kościelniak. 1995. “Influence of substrate loading intensity on floc size in activated sludge process.” *Water Research*, Vol. 29, pp. 1703–1710, doi:10.1016/0043-1354(94)00326-3.

Burges, C. 1998. “A tutorial on support vector machines for pattern recognition.” *Data Mining and Knowledge Discovery*, 2, 121, doi:10.1023/A:1009715923555.

Dellana, S.A. and D. West. 2009. “Predictive modeling for wastewater applications: Linear and nonlinear approaches.” *Environmental Modelling and Software*, Vol. 24, pp. 96-106, http://dx.doi.org/10.1016/j.envsoft.2008.06.002.

Eikelboom, D.H. 2000. “Process control of activated sludge plants by microscopic investigation.” IWA Publishing, London, UK.

Gironás, J. L.A. Roesner; L.A. Rossman and J. Davis. 2010. “A new applications manual for the Storm Water Management Model (SWMM).” *Environmental Modelling & Software*, Vol. 25, pp. 813–814, doi: 10.1016/j.envsoft.2009.11.009.

Güçlü, D. and Ş. Dursun. 2010. “Artificial neural network modelling of a large-scale wastewater treatment plant operation.” *Bioprocess and Biosystems Engineering*, Vol. 33, No. 9, pp. 1051-1058, doi:10.1007/s00449-010-0430-x.

Guo, H.; K. Jeong; J. Lim; J. Jo; Y.M. Kim; J.P. Park; J.H. Kim and K.H. Cho. 2015. “Prediction of effluent concentration in a wastewater treatment plant using machine learning models.” *Journal of Environmental Sciences*, Vol. 32, No. 1, pp. 90–101, http://dx.doi.org/10.1016/j.jes.2015.01.007.

Han, H. G; L. Ying; Y.N. Guo and J.F. Qiao. 2016. “A soft computing method to predict sludge volume index based on a recurrent self – organizing neural network.” *Journal Applied Soft Computing*, Vol. 38, No. C, pp. 477–486, http://dx.doi.org/10.1016/j.asoc.2015.09.051.

Henze, M.; W. Gujer; T. Mino; T. Matsuo; M.C. Wentzel; G.v.R. Marais and M.C.M. van Loosdrecht. 1999. “Actived Sludge Model No. 2d, ASM2d.” *Water Science and Technology*, Vol. 39, No. 1, pp. 165–182, doi:10.1016/S0273-1223(98)00829-4.

Henze, M.; P. Harremoes; E. Arvin and J. LaCour Janses. 2002. *Wastewater Treatment, Biological and Chemical Processes*, Springer–Verlag, Berlin, ISSN, pp. 1431-2492.

Jenkins, D.; M.G. Richard and G.T. Daigger. 2004. “Manual on the causes and control of activated sludge bulking, foaming, and other solids separation problems.” 3rd ed. Lewis Publisher, NY, USA, ISBN 9781566706476.

Kusiak, A. A. Verma and X. Wei. 2013. „A data–mining approach to predict influent quality.” *Environemntal Monitoring and Assessment*, Vol. 185, pp. 2197–2210, doi: 10.1007/s10661-012-2701-2.

Madoni, P.; D. Davoli and G. Gibin. 1999. “Survey of filamentous microorganisms from bulking and foaming activated-sludge plants in Italy.” *Water Res.*, 34, 1767–1772, doi: 10.1016/S0043-1354(99)00352-8.

Mąkinia, J. 2010. “Mathematical Modeling and Computer Simulation of Actived Sludge Systems.” IWA, London. doi: 10.2166/9781780401683

Medina, M. and U. Neis. 2007. “Symbiotic algal bacterial wastewater treatment: effect of food to microorganism ratio and hydraulic retention time on the process performance.” *Water Science & Technology*, Vol. 55, No. 7, pp. 165–171, doi: 10.2166/wst.2007.351.

Minsoo, K.; K. Yejin; K. Hyosoo; P. Wenhua and K. Changwon. 2016. “Evaluation of the k–nearest neighbour method for forecasting the influent characteristics of wastewater treatment plant.” *Frontiers of Environmental Science & Engineering*, Vol. 10, No. 2, pp. 299-310, doi: 10.1007/s11783-015-0825-7.

Rojek, I. and J. Studziński. 2014. “Comparison of different types of neuronal nets for failures location within water-supply networks.” *Maintenance and Reliability*, Vol. 16, No. 1, pp. 42–47, YADDA: bwmeta1.element.baztech-d4e78b27-a393-45e5-b9db-9d2388f8d881.

Santos, L.A.; V. Ferreira; M.M. Neto; M.A. Pereira; M. Mota and A. Nicolau. 2015. “Study of 16 Portuguese activated sludge systems based on filamentous bacteria populations and their relationships with environmental parameters.” *Applied Microbiology and Biotechnology*, Vol. 99, No. 12, pp. 5307–5316, doi: 10.1007/s00253-015-6393-8.

Szeląg, B. and J. Gawdzik. 2016. “Application of Selected Methods of Artificial Intelligence to Actived Sludge Settleability Predictions.” *Polish Journal on Environmental Study*, Vol. 25, No. 4, pp. 1709-1714, doi: 10.15244/pjoes/62262.

Szeląg, B. and J. Gawdzik. 2017. “Assessment of the Effect of Wastewater Quantity and Quality, and Sludge Parameters on Predictive Abilities of Non-Linear Models for Activated Sludge Settleability Predictions.” *Polish Journal on Environmental Study*, Vol. 1.

Vapnik V. 1998. “Statistical Learning Theory.” John Wiley and Sons. New York.

Verma, A.; X. Wei and A. Kusiak. 2013. “Predicting the total suspended solids in wastewater: A data – mining approach.” *Engineering Applications of Artificial Intelligence*, Vol. 26, No. 4, pp. 1366–1372. http://dx.doi.org/10.1016/j.engappai.2012.08.015.

Wei, X. and A. Kusiak. 2015. “Short-term prediction of influent flow in wastewater treatment plant.” *Stochastic Environmental Research and Risk Assessment*, Vol. 29, No. 1, pp. 241–249, doi:10.1007/s00477-014-0889-0.

Zawilski, M. and A. Brzezińska. 2013. “A real rainfall intensity distribution over an urban area and its effect on a combined sewerage system.” *Urban Water Journal*, Vol. 11, No. 7, pp. 532-542, doi: 10.1080/1573062X.2013.831909.

Hu, B.; O.K. Matar; G.F. Hewitt and P. Angeli. 2006. “Population balance modelling of phase inversion in liquid–liquid pipeline flows.” *Chemical Engineering Science*, Vol. 61, 4994–4997.

Huisman, J.L.; S. Burckhardt; T. Larsen; P. Krebs and W. Gujer. 2000. “Propagation of waves and dissolved compounds in a sewer.” *Journal of Environmental Engineering*, ASCE, Vol. 128, No. 1, pp. 12–20.

Zealand, J. and M. Russell. 1984. “Simple control equations using F:M ratio for operation of activated sludge aeration basins.” *Environmental Technology Letters,* Vol. 5, No. 1-11, pp. 49-56.

Chua, H.; P.H. Yu; S.N. Sin and K.N. Tan. 2004. “Effect of food:microorganism ratio in activated sludge foam control.” *Twenty-First Symposium on Biotechnology for Fuels and Chemicals*, May 2–6, 1999, Fort Collins, Colorado, pp. 127–135.

DESIGN OF A UWB COPLANAR FED ANTENNA FOR BIOMEDICAL APPLICATIONS

Benchakroun Hamza
Laboratory of technology of Information and modeling
National School of Applied Sciences, Marrakech, Morocco
hamza.benchakroun@ced.uca.ma

Mahlaoui Zakaria
Laboratory of technology of Information and modeling
National School of Applied Sciences, Marrakech, Morocco
zakaria.mahlaoui@ced.uca.ma

Latif Adnane
Laboratory of technology of Information and modeling
National School of Applied SciencesMarrakech, Morocco
a.latif@uca.ac.ma

Henrique Manuel
Department of Electrical and computer Engineering
INESC TEC –INESC science and Technology,
University of Porto, Portugal
Hsalgado@fe.up.pt

KEYWORDS: **Ultra wideband , coplanar antenna, microwave imaging , UWB, CPW, circular antenna.**

ABSTRACT: In this paper, a new ultra-wide band planar antenna is presented for microwave imaging, the antenna is designed to operate in a frequency band from 2.9 GHz to 10.8 GHz. The antenna was designed to be adaptable for multi-viewing imaging due to it simple form, low cost and ease to be manufactured. The paper features results of the simulation the new ultra large band antenna and a performance comparison whit other UWB antennas.

I- INTRODUCTION

Since the FCC allows the use of the UWB spectrum, the development of new systems that will adapt to this band has been growing. Much attention has been given to developing new systems that fit with UWB applications as radar detection, biomedical imaging. The proposed UWB antenna is more likely to be used in the biomedical imaging applications as a system to detect breast cancer.

Breast cancer is the second deadliest cancer for women, for more efficiency and an early detection, the biomedical field needs new systems that should be safe, comfortable and sensible [1][2]. The medical field already has its methods to detect breast cancer like X-ray mammography, Magnetic Resonance Imaging (MRI), as efficient as those techniques are they are missing between 10% to 30 % [2] of very early breast cancer stages.

The UWB microwave imaging is one of the techniques that has been developed, in the hunt for the next techniques that will detect breast cancer in its early stages. The basis for microwave detection is to compare between the dielectric properties of a normal tissue and cancerous tissue; the result of the comparison gives a prediction of malignant tissue characteristics as its size, shape, placement… [3].

The paper is organized as follows, the first part is going to introduce the UWB technology. The second unit is dedicated to the design and modeling of the antenna. The third and the last section is for discussing the results and performance of the antenna and comparing its performances with other UWB antennas.

The objective of this paper is to show a new UWB antenna, more likely to be used in microwave biomedical imaging. CST Microwave studio 2014 has been used for the design and the simulation.

II- ULTRA WIDE BAND

UWB is defined as a system with a very large band, this large spectrum comes usually with some advantages as a low power, high debit of data, high time resolution, low-cost and an ease of implementation, resistance to interference and so on. Those advantages opened a wide range of UWB application to radar detection, biomedical imaging, and HD communication.

The definition of UWB isn't a special one, the FCC (Federal communication Commission) [4] defines it as a system with a bandwidth larger than 500 MHz or larger than 20% if we are working wiht the relative bandwidth W/fc (W is the width of the band and fc is the carrier frequency). The UWB is being defined by its very large band that is 7.5GHz between 3.1GHz to 10.6GHz for a limitation of the power emission level -41.3 dBm/MHz [4].

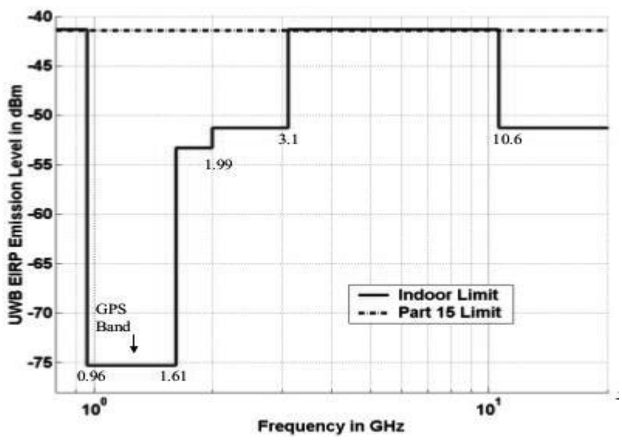


Figure ;1 : limit of the EIRP for UWB application

The fig.1 shows the power limitation required for an UWB system (FCC 2002) [4], the limitation is given in EIRP by frequency. EIRP standard for equivalent isotopically radiated power and it is the power that a theory isotropic antennae emits to reach the main maximum power density in the direction of maximum gain of the antenna. The EIRP expression is:

$$EIRP = P_t + L + G \qquad (1)$$

Where P_t is the power transmitted from the radio, L is the cable losses and G is the gain of the antenna.
Because the UWB systems have a low power spectral density, the interference between other system that use the same band can be ignored and the low power enables the resolution of the UWB receiver, that makes the system more efficient for ranging and positioning so more likely for the pulse to penetrate walls and ground for radar detection and penetrate the skin for biomedical decisions.
For an antenna to be called an UWB antenna or to be implemented in a UWB application, it should follow some requirements listed in the following unit [5].

III- MATCHING AND EFFICIENCY

The UWB communication systems uses a very short pulse duration of tens of hundreds of nanoseconds, and as the pulse and the bandwidth are inversely proportional, the shorter the pulse is the wider the spectrum is going to be.
Antennas would match the UWB requirements if they have a bandwidth greater than 500 MHz defined at -10dB according to the FCC, or have a relative bandwidth of more than 20%[4][5].The formula of the relative bandwidth is defined as the following:

$$BW_r = \frac{2(f_h - f_l)}{(f_h + f_l)} \qquad (2)$$

Where f_l and f_h are the low and the high frequency respectively
The efficiency of an UWB antenna can be evaluated by the specter efficiency, the evaluation matching has to be over the whole range of frequencies. The expression of the spectral efficiency is defined as follow [5]:

$$\eta_{rad(\%)} = \frac{\int_0^{\infty} P_t(\omega)\left(1-|\Gamma_t(\omega)|^2\right)d\omega}{\int_0^{\infty} P_t(\omega)d\omega} * 100\% \qquad (3)$$

Where P_t is the power at the antenna's terminal transmit, $\Gamma_t(\omega)$ is the normalize coefficient of reflection

IV- SIGNAL DISPERSION ANS DISTORTION

The signal passing through the UWB antennas is a very short pulse, the shorter the pulse is, the more likely the UWB antenna's response is going to be distorted and delayed due to the ripple after the pulse called the rippel effect as shown in fig. 2[5].

The rippling effect is caused by the geometry of the antenna, and it causes frequency translation, dispersion or delay on the transition reducing the speed of data transmission [5].

V- ANTENNA DESIGN

This section is dedicated to the design and the performance of the proposed antenna. The design proposed is a notched coplanar antenna. The geometry of the antenna consisted of a 25*30 mm2, Rogers RT5850 substrate with ε_r=2.2 and thickness h=3.175 mm. the top part of the antenna is a circle with two cutouts of 90° each. The structure of the antenna is shown in the fig.3:

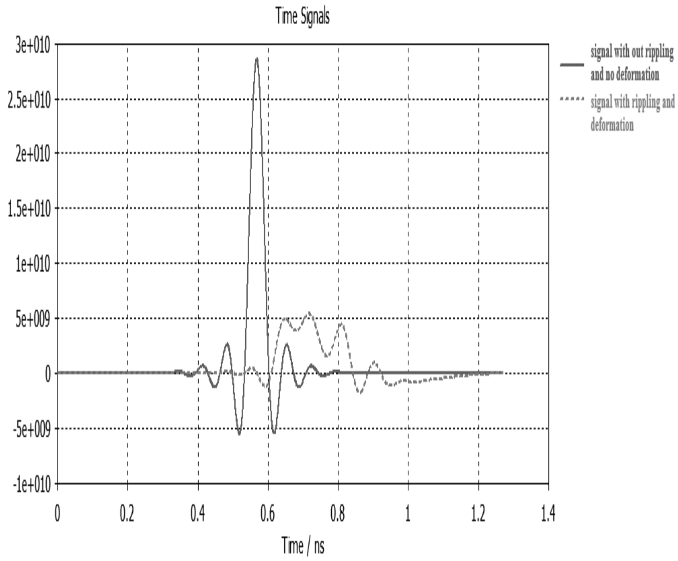


Figure 2 in signal and out signal with rippling

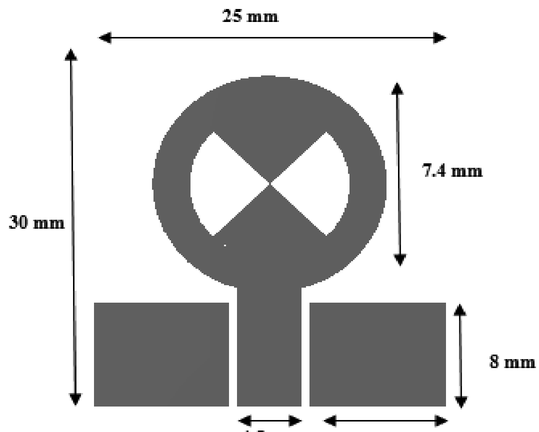


Figure 3 : geometry of the antenna proposed

The two main characteristics that effect the performance of the antenna were the inner radius of the circle R and the laminate thickness. The gap between the transmission line and the ground structure was optimized so that the antenna's impedance can match a 50Ω SMA port.

V-1 THEORETICAL RADIUS OF THE PATCH

To calculate the radius of the patch the theory from [6] (banalise, 1982) has been used, the used formulas did not take into consideration the fringing effect.

$$a = \frac{F}{\left\{1+\frac{2h}{\varepsilon_r F \pi}\left[\ln\left(\frac{\pi F}{2h}\right)+1.7726\right]\right\}^{1/2}} \quad (4)$$

$$F = \frac{8.791*10^9}{f_r\sqrt{\varepsilon_r}} \quad (5)$$

V-2-PERFORMANCE OF THE ANTENNA

Fig.4 shows the return loss or more known as the parameter S11 of the antenna, the spectrum of the Fig.4 contains the UWB frequency band spectrum. The spectrum's antenna range is from 2.8 GHz to 10.9 GHz and contains tow resonate frequencies 6.1 GHz and 9.8 GHz with the S11 parameters -46 dB and -35 dB respectively. The antenna can be easily used in the UWB (3.1 GHz – 10.6GHz) applications [4].

V-2-1 VSWR

VSWR stand for Voltage Standing Wave Ratio, it is a parameter that describes the power that is reflected by the antenna. The parameter WVSR is a function of the reflection parameter and it is defined by the following expression equ. (6):

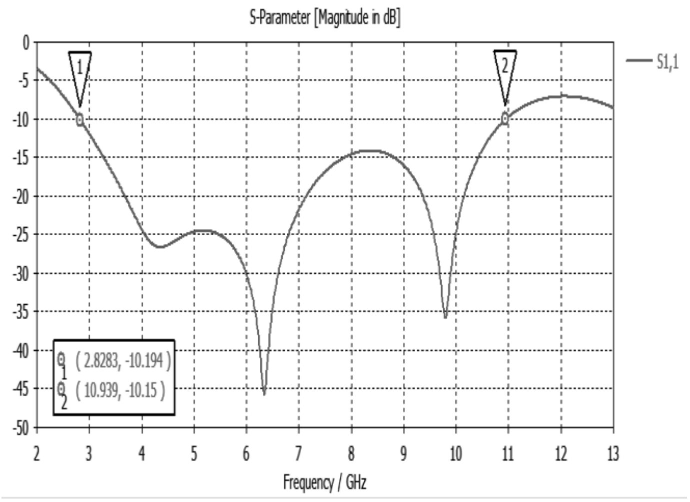


Figure 4: The return loss S11

$$VSWR = \frac{1+|\Gamma|}{1-|\Gamma|} \quad (6)$$

The fig.5 shows the graph of the VSWR of the antenna proposed, we can tell from the graph that the VSWR is under tow in all the bandwidth so the VSWR can be considered as good.

V-2-2 RADIATION PATTERN

The main focus of the paper is to design a UWB antenna that can be easily implemented in UWB application with a small dimension and a larger frequency spectrum. The fig.6 shows the radiation pattern of the antenna proposed for the frequency of 9.8 GHz, the color red defined the higher range of the gain and the green refers to the lower part [7].

The parametric study is done in two parts; the first part was to choose the laminate and then the value of the radius of the patch. The focus of the first part of the study is to compare the proposed geometry on three different Rogers laminates.The table.1 shows the characteristics of the three laminates. Fig.7 presents the return loss of the antenna with the different laminates, according to the results of fig.7, the laminate Rogers with a thickness of 3.175 mm was chosen to complete the parametric study.

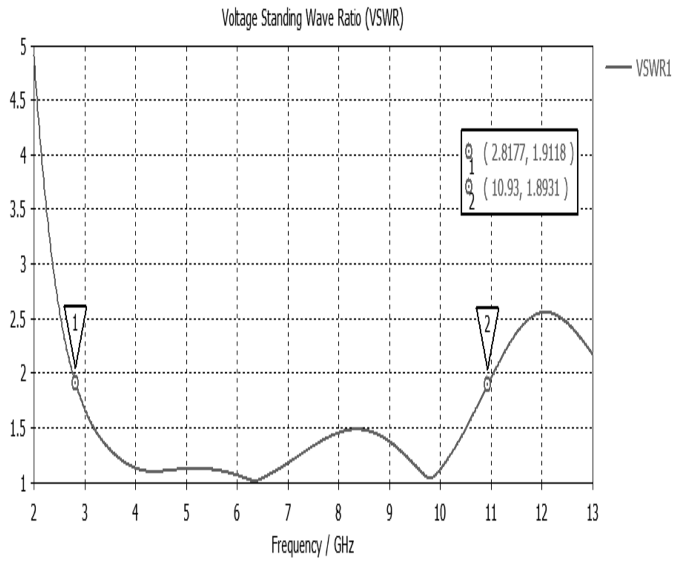


Figure 5 : Graph of VSWR

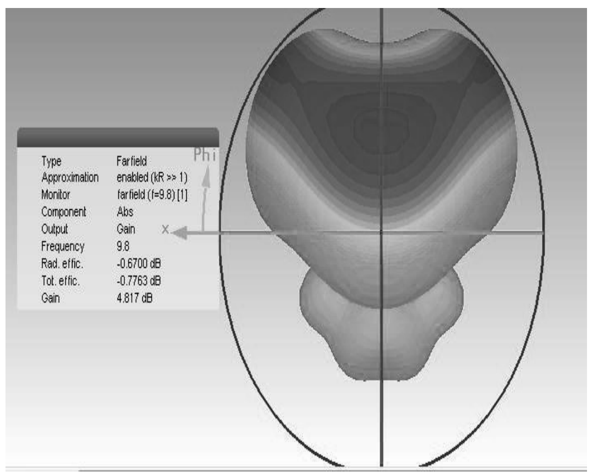


Figure 6: Radiation pattern of the antenna

Table 1 : Standard laminates characteristics

laminates	**Laminate characteristics**	
	Standard thickness	***Standard copper cladding***
Laminate 1	0.787 mm	35µm
Laminate 2	1.57 mm	35µm
Laminate 3	3.175 mm	35µm

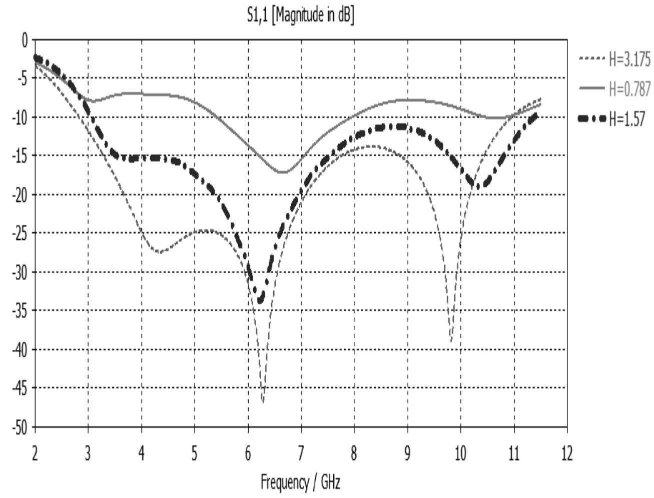


Figure.7 : Effect of the thickness of the substrate

The variation of the radius is altered around the theoretical radius to optimize the performance of the antenna. The patch of the antenna is separated in two parts, first where the cutouts are and the second part is the ring that contours the cutouts, this parametric study consisted on varying the inner radius of the patch, the radius of the circle that contains the cutouts.
The fig.8 shows a relation between the radius and the depth of the return loss, R=4.8 mm shows a better S_{11} parameter.

VI-COMPARISON WITH OTHERS UWB ANTENNAS

The antennas that were compared to the antenna proposed in the paper are: a coplanar Microstrip Antenna with Defected Ground structure for UWB Applications [7], A Printed UWB Antenna with Full Ground Plane also for WBAN Applications [8] and CPW-fed Slot Patch Antenna [9]

The table 2 shows the difference between four UWB antennas, the proposed antenna has a wider band, includes all UWB frequencies, the highest gain is around 5.6 dBi.

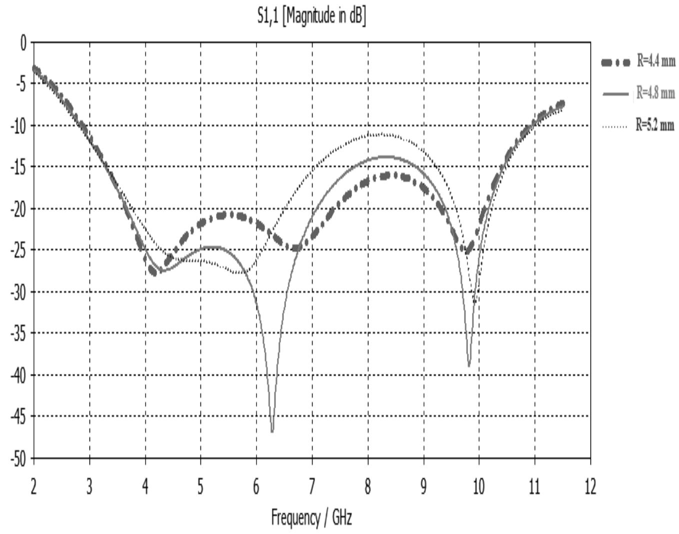


Figure.8 :Effect of the inner radius on the antenna

Table 2 : Performance comparison

Antennas	**Characteristic of the antennas**		
	Dimension (w*L)	bandwidth	Gain
Antenna proposed	25mm*30mm	2.7GHZ to 10.98GHZ	3.5 dBi
Antenna [7]	30mm*32mm	3.1GHz to 9.9GHz	-
Antenna [8]	75mm*85mm	4GHz to 9.5 GHz	-
Antenna [9]	24mm*30mm	4.82GHz to 8.87 GHz	3.9 dBi

VII-CONCLUSION

The new coplanar antenna with cutouts on a circular patch demonstrates advantages: larger frequency spectrum, a good VSWR, improved parameter S11. The geometry of the patch was calculated and design with CST microwave studio. The antenna was designed for UWB application with likely microwave imaging in the range from 3.1GHz to 10.6GHz. Moreover the antenna is smaller than traditional antennas, and as a coplanar it can be easy to implement. Also, enhancing gain will be a good improvement by making this antenna in an array due to its small size.

REFERENCES

Shao, W. and B. Zhou, “UWB microwave imaging for breast tumor detection in inhomogeneous tissue,” *Proceedings of the 2005 IEEE Engineering in Medicine and Biology, 27th Annual Conference, Shanghai, China, 1496–1499, 2005.*

X. Xiao and T. Kikkawa, “Early breast cancer detection by ultrawide band imaging with dispersion consideration,” *Jpn. J. Appl. Phys., vol. 47, no. 4 PART 2, pp. 3209–3213, 2008.*

S. C. Hagness, A. Taflove, and J. E. Bridges, “Three-dimensional FDTD analysis of a pulsed microwave confocal system for breast cancer detection: design of an antenna-array element,” *IEEE Trans. Antennas Propag., vol. 47, no. 5, pp. 783–791, 1999.*

Ultra-Wideband Operation FCC Report and Order, Tech.Rep.US 47 CFR Part 15, 2002.

D. Valderas, J. I. Sancho, and D. Puente, Design and Applications. .

Balanis C.A(1982) Handbook of Microstrip Antennas. *John Wiley and Sons New York.*

L. Cheema and K. Sherdia, “Design of Microstrip Antenna with Defected Ground structure for UWB Applications,” *Int. J. Adv. Res. ..., vol. 2, no. 7, pp. 2525–2528, 2013.*

R. B. V. B. Simorangkir, S. M. Abbas, and K. P. Esselle, “A Printed UWB Antenna with Full Ground Plane for WBAN Applications,” *pp. 127–130, 2016.*

V. youtube. com/watch?v=DcAhAHF7pq. Niranjan, A. K. Saxena, and K. V Srivastava, “CPW-fed slot patch antenna for 5.2/5.8GHz WLAN application,” *Prog. Electromagn. Res. Symp. PIERS 2012 Kuala Lumpur, pp. 1350–1352, 2012.*

BIOGRAPHY

Adnane LATIF was born in Morocco in 1973, has a PHD in Telecommunications and Networks, professor at Cadi Ayyad University, National School of Applied Sciences, Marrakesh, Morocco. Researcher in the areas of Antennas, Microwaves and Wireless communications.

Author of some fifty publications in international conferences and international journals.

Founder and coordinator of the Spring School 2010 "Wireless and Emerging Technologies."

Founder and President of the Moroccan Association of Telecommunications and Electronics Technologies.

Member organizer of several International Conferences and Spring Schools.

Member of International Program Committee of several international conferences.

Reviewer and Board member of several International Journals

Author of two books.

VIRTUAL REALITY ENVIRONMENTS

ANYCRANE: TOWARDS A BETTER PORT CRANE SIMULATOR FOR TRAINING OPERATORS

Rustam Abubakirovich Fayzrakhmanov
Ivan Sergeevich Polevshchikov
Aydar Fadanisovich Khabibulin
Filipp Igorevich Shkliaev
Department of Information Technologies and Automated Systems
Perm National Research Polytechnic University
29 Komsomolsky prospekt, Perm, Russia
fayzrakhmanov@gmail.com

Ruslan R. Fayzrakhmanov
Department of Computer Science
Oxford University
Wolfson Building, Parks Road
Oxford OX1 3QD
ruslan.fayzrakhmanov@cs.ox.ac.uk

KEYWORDS

Computer training simulation complex, computer-aided training system, modelling, training, professional skills.

ABSTRACT

In this paper, we present the results of our ongoing project, AnyCrane. The project is dedicated to the development of training simulators for port crane operators. In contrast to existing well-known systems, our research is two-fold: 1) we focus on the simulation of different processes and situations involving portal crane operator and port personnel; 2) we introduce an adaptive self-training approach integrated into the simulator to help a trainee to acquire required skills.

INTRODUCTION

The effectiveness and productivity of complex technological processes in modern ports and terminals is mainly defined by the competence and professionalism of the personnel. In particular, port crane operators are required to have specific skills to perform complex technological operations timely and accurately, excellent knowledge of the process, as well as fast and proper reaction in different regular and extreme situations. Training of crane operators on real technological objects is quite expensive and, due to the danger of the process, it puts a high responsibility on an instructor. Thus, training of beginners typically starts with simulators.

There are several world leading simulators developed by companies such as “CM Labs[1]”, “Liebherr[2]”, “Global Sim[3]”, “SPM[4]”, and “Forward[5]”. Their products are mainly focused on recreating a realistic environment, including rendering the virtual environment, modeling various objects in ports and terminals, as well as using control panels and chairs similar to those of real cranes. In particular, CM Labs’ port equipment training simulators for different cranes render 3D models of the scene with their textures,

Figure 1: Interaction of a trainee with the virtual environment, AnyCrane

different weather conditions, as well as features and physics of various objects with high resolution and impressive precision.

Simulators of the mentioned companies also implement a training component, that is typically comprises set of exercises. Results of these exercises are represented in terms of data sheets and graphs to be analysed by the instructor and used both to improve the effectiveness of the training (e.g., adapting exercises) and assess the performance of the student.

To the best of our knowledge, regardless of the indicated advantages of modern port crane simulators, they do not pay enough attention to other participants of the cargo handling process. For example: locomotion of the transport and personnel that can explicitly or by implication influence the process, or solely used to model the environment. In addition, in smaller ports and terminals with less rigorous requirements to cargos, a slinger is an indispensable part of the handling process. He hooks, unhooks the load, and coordinates the work of the crane operator. Thus, his role and participation should be correctly defined and incorporated into the training. Our research is aimed at modeling these aspects, which we implement in our simulator, presented in this paper.

Furthermore, according to our research and the feedback received from instructors using contemporary training simulators, the training component should be enhanced and be more “intelligent” (Bouhnik and Carmi 2012; Ashtiani el al. 2013;). In particular, it should be able to analyse the progress the student makes, its performance and, based on the goals

[1] https://www.cm-labs.com
[2] https://www.liebherr.com/
[3] http://www.globalsim.com
[4] http://specprofmat.ru
[5] http://www.autotrenajer.ru

specified, provide relevant supportive hints and accordingly modify the training process. The system should additionally provide the student and instructor with information regarding the development of specific skills and abilities, instead of “raw” exercise-specific data. On the one hand, this allows the student to get better control on his achievements and be more independent, and, on the other hand, allows the instructor to concentrate on more fundamental aspects of the learning process related to knowledge, abilities, and skills. We implement this in our training component of the simulator.

THE CRANE TRAINING SIMULATOR

Our simulator consists of seven main components to simulate a dynamic visual environment: physics and mathematics module, logical module, graphical module, audio module, engine, display, and control panel (see Fig. 2).

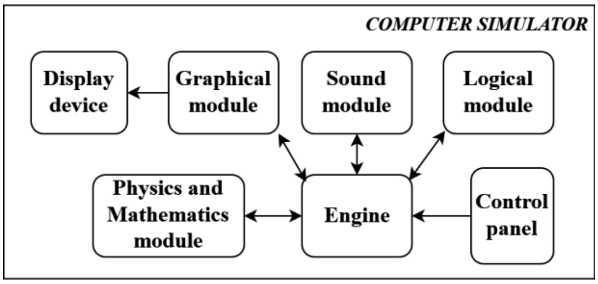


Figure 2: Structure of the training simulator

The physics and mathematics component models physical features of various objects that can be found at the port (be it a cargo, wire rope, or liquids) and the reaction of the virtual environment and its objects to signals coming from the control panel. Our training simulator has a control panel and levers similar to those of a typical port crane (see Fig. 1). In accordance with the specification of the crane, signals from controls are transmitted to the crane electric actuators, described by the second order differential equations. Electric actuators, in turn, drive the structural elements of the crane, setting in motion of the crane (e.g., turning the tower or moving the boom). Any movement of structural elements changes the position of the suspension point and causes oscillations of the cable-cargo system. Thus, this system plays a crucial role in the cargo handling processes.

We model the cable-cargo system with two main goals in mind: 1) to reflect an adequate physical behavior and 2) render realistic visual model. A cable is interpolated by a set of interconnected points with a specific set of applied restrictions (e.g., on the length between points) (Khabibulin and Shklyaev 2015). This model is suitable for realistic simulation of the physical features of the wire rope (e.g., oscillations or cable break). The virtual environment simulating the dynamics of the overload process is shown in Fig. 3.

A scenario is realized and composed by the logical module. It determines traffic routes and the frequency of the transport appearance. We also model a port staff,

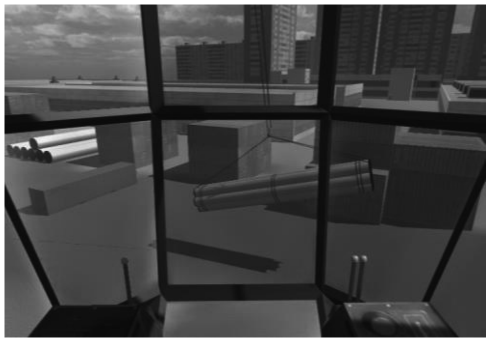


Figure 3: Simulating the dynamics of the overload process

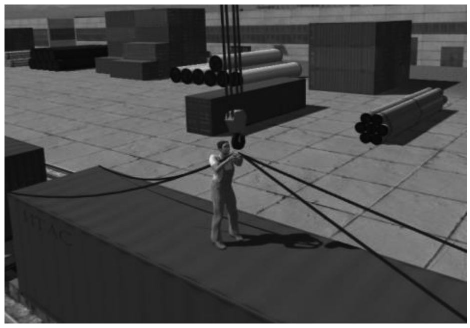


Figure 4: Simulation of the process of slinging a cargo container

an essential part of the handling process, performing slinging and showing specific signals to the operator.

The A* path search algorithm is used to simulate locomotion of the staff across the scene. Fig. 4 illustrates the process of slinging a cargo.

The graphic module creates a realistic image. It uses the created 3D models to visualize all objects in the scene. The sound module allows the simulation of the audio environment. It can be vehicle whistles, engine noise, sound from cargo collisions, etc. The engine unites all the components of the simulator and provides a unified API for integrating third-party modules.

THE TRAINING COMPONENT

The development of professional sensorimotor skills in AnyCrane is performed with the help of exercises, modeling real technological operations in the context of a cargo loading and unloading processes.

The training component of our simulator collects and processes all relevant data reflecting the progress a trainee makes, analyzing its professional knowledge and skills. These data is obtained in real-time from the engine and specify various relevant characteristics of the system, such as the lifting speed of load, the tension of the cable, etc. A schematic structure of the training system and its relation with the computer simulator are shown in Fig. 5.

(1) manages initial settings of the system which correspond to the initial expertise of the student and its goal. Settings are set via a user interface before the training process. The training plan with specific sequence of exercises is formed in (2) according to the initial settings. The functional block (3) analyses

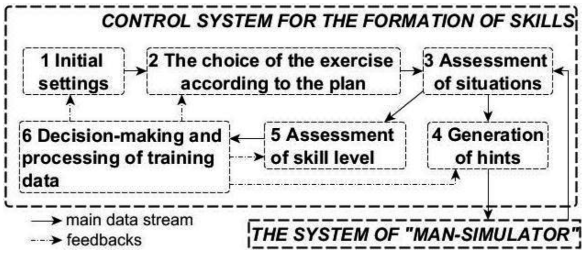


Figure 5: Structure of the control system

various situations identified during exercises and identifies problematic cases. (4) leverages this information and creates set of visual and audio hints and advises to be integrated into the exercise to support the student and enhance the learning process. (5) analyses the performance of the trainee in terms of its skills, abilities and knowledge. The assessment is made based on the evaluation of different situations identified in exercises, individual and integral quality criteria. The decision-making module (6) is activated after the student performed an exercise; it adapts system parameters and defines further training scenario, e.g.: a) repeat the exercise, b) adapt skill assessment criteria, c) correct the set of selected hints, and d) refine initial settings.

The introduced approach ensures good information support and feedback for the student and convey both exercise-specific (3) and training-specific (5) information to the instructor. We believe that this system can considerably enhance the training processes helping the student be more independent from the trainer and have better control of the progress he makes. In **(**Fayzrakhmanov et al. 2017) we present promising results we achieved with the use of our training simulator and its training component (mentioned as Ganz TSC). In particular, our case study shows that students trained with the use of our approach move cargos 15% faster, on average, preserving the required quality of work.

CONCLUSIONS AND FUTURE WORK

In this paper, we introduced a training simulator, results of our ongoing project, AnyCrane. In contrast to existing systems, we focus on modeling the overall variety of relevant and important components of the handling process (including transportation and a slinger). Furthermore, we developed and integrated a training component which is oriented on the improvement of skills, abilities, and knowledge of a student.

Our future work is to enrich a slinger with additional functionality, signals, coordinating the work of the crane operator, and, thus, make interaction between a crane operator and a slinger more realistic. We also plan to develop a port scene constructor for instructor to create a port utilizing existing components (e.g., cranes, slingers, and trucks). Furthermore, we plan to pay more attention on different weather conditions (e.g., wind, rain, and snow) which can affect the cargo handling process.

REFERENCES

Ashtiani M.J., M. Nomanof, B.S. Bigham and A. Madadi. 2013. "Computer assisted assessment (CAA) and electronic problem based learning". *Life Science Journal*, V.10, No.1, 726-730.

Bouhnik D., G. Carmi. 2012. "E-learning Environments in Academy: Technology, Pedagogy and Thinking Dispositions". *Journal of Information Technology Education: Research*, V.11, 201–219.

Fayzrakhmanov R.A., I.S. Polevshchikov. 2013. "Increased of Efficiency in the Automated Training of Fuelling Machine Operators Using Iterative Simulation Learning". *World Applied Sciences Journal,* No.22 (Special Issue on Techniques and Technologies), 70-75.

Faizrakhmanov R.A., I.S. Polevshchikov. 2016. "Automated control simulation of professional skills formation for production system operator". *Scientific and technical journal of information technologies, mechanics and optics*, V.16, No.1, 181–190.

Fayzrakhmanov R.A., I.S. Polevshchikov, A.F. Khabibulin. 2017. "Computer Simulation Complex for Training Operators of Handling Processes". *Proceedings of the 5th International Conference on Applied Innovations in IT*, No.1, 81-86

Khabibulin A.F., P.I. Shklyaev. 2015. "Modeling of cable oscillations for gantry crane simulator". *Innovacionnye tehnologii: Teorija, Instrumenty, Praktika,* No.1, 82-88.

Expert system based opportunistic mobile social network for the historical places virtual re-construction utilizing populace's acquainting photos.

Haifaa O. Elayyan
Arab Open University
Amman, Jordan
Haelayyan@gmail.com

Philippe Geril
ETI
Ostend, Belgium
pgeril@yahoo.co.uk

Keywords-
Opportunistic mobile social network, Expert systems, panoramic photos and 360 photos, Virtual Heritage Places reconstruction, walkthrough, Datamining , IoT .

Abstract-

We propose a novel opportunistic mobile social network that exploits a place with all details, photos and social characteristics, such as similarities, mobility patterns for people who are interested to perform the message routing and data sharing. In such networks, the users with mobile devices are moving digital treasure with a cost that goes to zero as they are able to form on-the-fly social networks to communicate with each other and share data objects, mainly places. The new intelligent social network motivates people to upload photos for places not only their faces and friends. These photos can be 2d / 3d and panoramic ones. The photos will be used later to construct the corresponding place virtually as an enormous knowledge related-photos base which is expected to be built / trained / classified with data Mining tools and Image processing expert systems . The photos will result in a well maintained visual database for that place. This Enterprise knowledge-based social network will be available for people to practice and enjoy a walk through as an 360 virtual visit. Basic features of a social network will be still taken into consideration such as comments and sharing. The system will have a recommender subsystem to recommend places according to the people's experience. The system will be aiming to reconstruct heritage places as its first goal and priority.

I. Introduction

The historic environment is one of our most precious assets. It contributes to our economy, to our cultural identity and to our sense of place. Without it, any country would not be that special country. Historic environments support job opportunities and contribute billions to the national gross value added (GVA).`There is tourism, obviously, but there are also jobs and work for local businesses in their care and restoration. It was forecast that billions of visits would be made to historic sites in 2017 and despite of the war, people are still heart-linked with these places all over the world [1]

I.A. Six Practical Reasons to Save Historical places

What is historic, and worth saving, varies with the beholder, but some definition is urgent. Simply put, "historic" means "old and worth the trouble." It applies to a building that is part of a community [1]

1) Historical places have intrinsic value.
2) When you tear down a historical place, you never know what is being destroyed.
3) Statistics have shown that most of new businesses prefer to have a database of historical places on their websites.
4) Historical places attract people.
5) Historical places are reminders of a city's culture and complexity
6) Regret goes only one way.

Perceptible past. And though it may surprise cynics, old buildings can offer opportunities for a community's future.

This project examines both the cultural and practical values of historical places and looks at why preserving them is beneficial not only for a community's culture, but also for its local economy.
War, by its very nature, is a destructive force. And the magnificence of an ancient building, the history of which was forged over centuries, can vanish with the push of a button or the squeeze of a trigger. History, Heritage and Archaeology, [Scotland, 2017][2] website highlights 10 historically and culturally significant sites completely lost to the ravages of war, some as recently as 2017. The list contains many placed distributed among contents. Worth to mention that these places are completely destroyed and there is lack of documentary resources for these places especially during the war period [2]

Since years the Middle East has suffered a massive deconstruction of historical places as most of its countries have had revolutions and internal fights that destroyed almost 80% percent of its historical places. Syria has had the hugest share in deconstruction, in which unfortunately most of its old and historical places were destroyed. The project is aimed to motivate people all around the world to upload their photos of Aleppo city before and after the war, using their mobiles with the new technology of 360 panorama tic photos as well as short movies from millions of people. In doing so do you not think it will be manageable to build this city virtually?

Keep in mind that you will have thousands of photos for a corner so there is no possibility of missing small details of this corner as the photos will be from different views.
The Middle East has been ranked as one of the top regions in the world for ancient sites, but unfortunately many countries in the region are at war and many historical sites are in danger [1]

I.B Virtual heritage

Virtual heritage or cultural heritage and technology is the body of works dealing with information and Communication technologies (ICT) and their application to cultural heritage, such as *virtual archaeology*. .Virtual heritage and cultural heritage have independent meanings: *cultural heritage* refers to sites, monuments, buildings and objects "with historical, aesthetic, archaeological, scientific, ethnological or anthropological value", whereas *virtual heritage* refers to instances of these within a technological domain, usually involving computer visualization of artefacts or Virtual Reality environments.
One technology that is frequently employed in virtual heritage applications is augmented reality, which is used to provide onsite reconstructions of archaeological sites or artefacts. Many virtual heritage projects focus on the tangible aspects of cultural heritage, for example 3D modelling, Graphics and animation. In doing so they often overlook the intangible aspects of cultural heritage associated with objects and sites, such as stories, performances and dances. The tangible aspects of cultural heritage are not inseparable from the intangible and one method for combining them is the use of virtual heritage serious games, such as the 'Digital Songlines" and 'Virtual Songlines" which apply virtual reality to preserve, protect and present the cultural heritage of Australian Aborigines.

II. Innovation

A typical social network is a social assembly made up of a set of social performers (such as individuals or organizations), sets of dyadic ties, and other social relations between actors. The social network standpoint runs a set of methods for investigating the structure of entire social bodies as well as a diversity of theories explaining the pragmatic patterns in these structures.

Social networks and the analysis of them is an inherently interdisciplinary academic field which emerged from social psychology, sociology, statistics, and graph theory. Georg Simmel authored early structural theories in sociology emphasizing the undercurrents of triads and "web of group affiliations". Jacob Moreno is credited with developing the first sociograms in the 1930s to study interactive relationships [13]. The diffusion of ideas and innovation studies focus on the feast and practice of ideas from one actor to another or one philosophy to another. This line of research seeks to elucidate why some become "early adopters" of ideas and innovations, and links social network composition with facilitating or impeding the spread of an innovation. Many concepts of a social network can be re-deployed in our idea and the most trending are social capital and mobility benefits. The sociological concept which refers to the value of social relations and the role of cooperation and confidence to achieves positive outcomes. The term refers to the value one can get from their social ties. For example, newly arrived immigrants can make use of their social ties to established migrants to acquire jobs they may otherwise have trouble getting (e.g., because of unfamiliarity with the local language). Studies show that a positive relationship exists between social capital and the intensity of social network use. While the mobility can benefit within any organization, targeting its members tends to focus their activities inside their own groups, which stifles creativity and restricts opportunities. A player whose network bridges structural holes has an advantage in detecting and developing rewarding opportunities. Such a player can mobilize social capital by acting as a "broker" of information between two clusters that otherwise would not have been in contact, thus providing access to new ideas, opinions and opportunities. British philosopher and political economist John Stuart Mill, writes, "it is hardly possible to overrate the value of placing human beings in contact with persons dissimilar to themselves. Such communication [is] one of the primary sources of progress."Thus, a player with a network rich in structural holes can add value to an organization through new ideas and opportunities. This in turn, helps an individual's career development and advancement [14].

In this paper, we are proposing a new concept of social networks and deploying recent technology in a collaborative platform. The proposal aims mainly to revive historical places and build a huge second virtual life with avatars that can be utilized later for academic/gaming /video clips purposes and documentaries. The proposed systems suggests a novel expert system with new research fields including mobile applications, image processing, Artificial Intelligence, internet Of thing concept, Clouding, Opportunistic computing, Recommender systems, Mobile data offloading Simulators, Data mining and classification concepts. Neural networks and heuristics are the main concepts in our system. The system still presents the concept in a simple way that non-expert users can enjoy sharing their experience and motivating them to be a part of constructing the knowledge –based system and many other features and innovation re-presenting usual utilities over the web. Some are summarized in the following section:

- Promote the concept of virtual tours with simple tools that people can design by themselves.
- Hosting the biggest database ever for each single spot with short time and simple budget utilizing people readiness to share information and experience in different way than sharing personal photos for fun.
- Creating and building of virtual tours for damaged and destroyed sites. Even non-destroyed sites.
- This will give a chance to visit the site and build a deep Knowledge without the annoyance of travelling or depending on limited resources of having photos of the aforementioned site. It will be easier for studying; consulting; suggesting enhancement or even maintenance practice.
- SMART Tourism
- The experience of people is scattered among the different applications on Internet (WhatsApp; Twitter; Face book) and yet this experience of visiting sites and having pictures is still not as professional as required. The new social network will still retain the friendly communication between friends but yet upgrade this communication to a better view of these sites.

• Most of hotels around the world have the 360 tour on their websites but it is this view from their side that lacks the credibility as they post only the good things. Giving people a chance to record their experience of visiting a place and uploading thousands of photos of a specific corner; reception; suit and rooms will give a fair chance for other to decide whether to visit this place or not .
• Architecture: This is an important point as in this world we have a variety of internal and outdoor designs that cannot be reached or visited as they are far away or expensive or unknown for others . This will energize the exchange ability of the architectural design.
• Reviving historical places and building a huge second virtual life with avatars that can be utilized later for academic/gaming /video clips and documentaries.
• Suggest a novel Expert system with new research fields including mobile applications, image processing, Artificial intelligence, the internet of thing concept, Clouding, Opportunistic computing, Recommender systems , Mobile data offloading Simulators, data mining and classification concept. Neural networks and heuristics are the main concepts in our system.
• Propose the concept in a simple way that non-expert users can enjoy sharing their experience and motivating them to be a part of constructing the knowledge –based system.
• Promote the concept of virtual tours with simple tools that people can design by themselves.
• Hosting the biggest database ever for each single spot within a short time and simple budget utilizing people's readiness to share information and experience in different way than sharing personal photos for fun.

III. Proposed System:

Expert system based opportunistic mobile social network for the historical places virtually re-construction utilizing populace's acquainting photos

The mutual design of our system goes actually through the different stages that can be divided for details' purposes. The system mainly deploys three important components:

1- Opportunistic mobile social network.
2- Cloud – based Expert systems
3- Internet of things platforms.

The following part of this paper examines and clarifies the use of these systems and how they are interconnected to each other formulating a huge platform for a new concept of social networks having the "the Place " as a main actor .

III.A. Why use an opportunistic mobile social network:

Opportunistic mobile networks comprise of human carried mobile devices such as smart phones that interconnect with each other in a "store-carry-forward" style, reduce the conforming communication overhead without any infrastructure. Opportunistic mobile networks present diverse challenges compared to traditional fixed networks, such as the Internet, that assumes the accessibility and convenience of a contemporary, reasonably low propagation delay/ interruption, low packet loss rate path between the two end points that communicate. In opportunistic networks, disconnections and highly adjustable delays caused by the mobility of mobile devices moving into the wireless range are the norm. Another major contest in opportunistic communication arises from the small format influence of mobile devices that introduce resource boundaries compared to static computing systems. Furthermore, employment and deployment of actual opportunistic mobile networks, systems and submissions is stimulating, very often classy and time-consuming as mobility itself is a noteworthy problem in mobile networking. Opportunistic mobile networks can be seen as a oversimplification of DTNs (Delay Tolerant Networks)[3]

Explicitly, in opportunistic mobile schemes such as in DTNs, mobile social applications and location-based services neither a priori knowledge is assumed about the conceivable points of stoppages, nor the presence of separate Internet like sub networks is assumed. Opportunistic mobile networks are designed by individual nodes that are per chance connected for long time intermissions, and that opportunistic character exploits any exchange with other nodes to forward messages using routing protocols, such as DSR (Dynamic Source Routing). The steering approach between predictable DTNs and opportunistic mobile networks is therefore quite different. As in DTNs, constant end-to-end connectivity may never be obtainable as it is concerned with intersecting highly heterogeneous networks, the possible points of disconnections (and, sometime, the duration of disconnections) are known, routing can be achieved along the same lines used for conventional Internet protocols, bearing in mind the duration of the disconnections as a supplementary cost of the links [3][4]

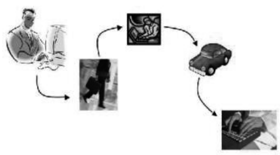


Figure 1: DSR (Dynamic Source Routing

The goal for opportunistic mobile networks is to use opportunistic algorithms for exploiting the throughput that can be provided by opportunistic algorithms. The global acquaintance of networks including information framework would enable prime and optimal routing. Opportunistic mobile social networks are a form of mobile ad hoc networks that exploit the human social characteristics, such as similarities, daily routines, mobility patterns, and interests to perform the message routing and data sharing.[1][2] In such networks, the users with mobile devices are able to form on-the-fly social networks to communicate with each other and share data objects.[3] The following figure gives an imaginary scheme of it.

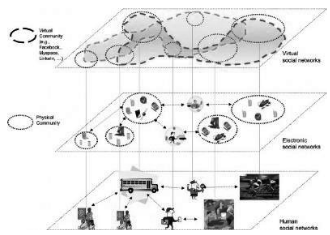


Figure 2. Opportunism in social networks

Opportunistic mobile social networks has many features that can be employed to get the best out of this social network as the following:
a. Social Matrices: The most common social metrics that are widely utilized to shape message communications by exploiting social networks are betweenness centrality, degree centrality, and closeness centrality[3]
b. Opportunistic computing which utilizes the shared resources, content, services, applications, and computing resources, by the devices connected in an opportunistic mobile social network, to provide a platform for the execution of distributed computing tasks.[11] However, opportunistic computing requires middleware services to cope with the intermittent connectivity and delay of the opportunistic communication environments.
c. Recommender systems such as a novel application area of opportunistic mobile social networks. Such systems track the user activities, mobility patterns, and utilize the user's contextual information to provide recommendations on variety of items.
d. Mobile Data Offloading: An active area of research in the applications of opportunistic mobile social networks is mobile data offloading. With large number of ever increasing smart phone users, most of the 3G networks are overloaded.[13][14] Several research works are performed to utilize opportunistic mobile social networks in offloading of mobile data traffic to reduce the load on 3G networks. \

III.B. Opportunistic IoT: Exploring the harmonious interaction concerning human and the internet of things

The traditional view of Internet of Things (IoT) challenges to connect all the physical objects to build a global, infrastructure-based IoT. The project is based on the first level on IoT as our first interactive platform with users and their different devices to collect information with ease and in a limited time. Reasons behind using the IoT can be summarised as follows depending on the great difference IoT has participated in the ITC. [5][6]

Internet of Things has been recognized as one of the developing technologies in IT as noted in Gartner's IT Hype Cycle (see Fig.2). A Hype Cycle [10] is a way to represent the emergence, adoption, maturity, and impact on applications of specific technologies. It has been forecasted that IoT will take 5–10 years for market acceptance.

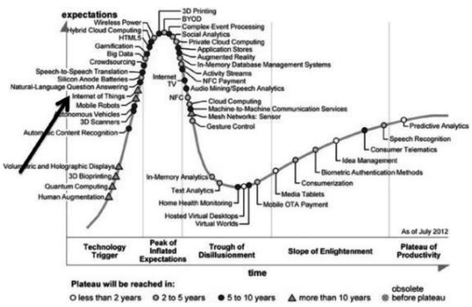


Figure 3. Gartner 2012 Hype Cycle of emerging technologies.

Main Characteristics being clearly used in our project are [11]:

a. Radio Frequency Identification (RFID)
b. Secure Data Aggregation
c. Visualization
d. Numerous application
e. Quality of Service.
f. Data Mining
g. GIS based Visualization
h. Cloud Computing
i. International activities.

III.C. Expert System:

The objective using an expert system is to show, explicit expertise can help solving complex image processing problems, and build a based knowledge that converts the system by time to be a completely automated one . Expert systems for image processing are classified into four categories, and their objectives, knowledge representation, reasoning methods, and shortcomings are discussed. The categories are:
(1) consultation systems for image processing;
(2) knowledge-based program composition systems;
(3) rule-based design systems for image segmentation algorithms;
(4) goal-directed image segmentation systems [5]

The importance of choosing effective image analysis approaches is crucial. Two methods are proposed for characterizing image analysis strategies: one from a software engineering viewpoint and the other from a knowledge representation viewpoint. The expert system never forgets, it can store and retrieve more knowledge than any single human being can remember, and makes no errors, provided the rules created by the subject matter experts accurately model the problem at hand.

Typically an expert system is composed of two major components, the Knowledge-base and the Expert System Shell[6]. The Knowledge-base is a collection of rules encoded as metadata in a file system, or more often in a relational database. The Expert System Shell is a problem-independent component housing facilities for creating, editing, and executing rules. The shell portion includes software modules the purpose of which is to:

- Process requests for service from system users and application layer modules;
- Support the creation and modification of business rules by subject matter experts;
- Translate business rules, created by a subject matter experts, into machine-readable forms;
- Execute business rules; and
- Provide low-level support to expert system components (e.g., retrieve metadata from and save metadata to a knowledge base, build Abstract Syntax Trees during rule translation of business rules, etc.).

III.D. The Mutual design

This section shows the Mutual design of the proposed system.

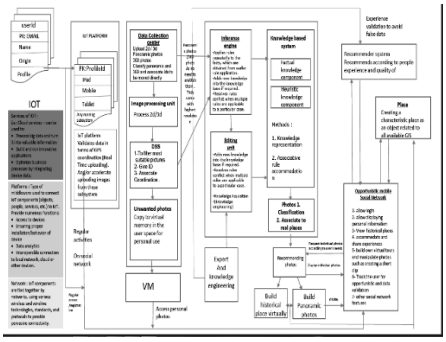


Figure 4. Proposed system

In the following section we indicate briefly to the most trending competitors in this field to show the differences and on which they had based their work and researches.

I.V Previous Studies and Related Work

A. BigSFM: Reconstructing the World from Internet Photos

Our group is working on building a 3D model of the world from online photo collections, and our research spans several areas, from image features, to large-scale image matching, to structure-from-motion optimization, to applications such as location recognition. This page summarizes our work, has links to code and datasets we have made available, and has a description of each project.

The seriatim project impels location recognition and image localization method that leverages feature correspondence and geometry estimated from large Internet photo collections. Such recovered structure contains a significant amount of useful information about images and image features that is not available when considering images in isolation. For instance, we can predict which views will be the most common, which feature points in a scene are most reliable, and which features in the scene tend to co-occur in the same image. Based on this information, we devise an adaptive, prioritized algorithm for matching a representative set of SIFT features covering a large scene to a query image for efficient localization. Our approach is based on considering features in the scene database, and matching them to query image features, as opposed to more conventional methods that match image features to visual words or database features. We find this approach results in improved performance, due to the richer knowledge of characteristics of the database features compared to query image features. We present experiments on two large city-scale photo collections, showing that our algorithm compares favourably to image retrieval-style approaches to location recognition.

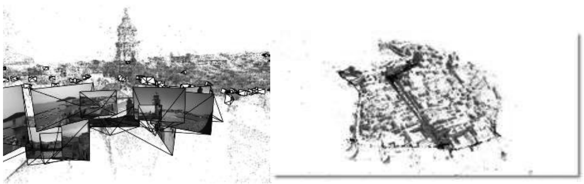


Figure 5. This work was supported in part by NSF (grant IIS-0713185), Intel, Microsoft, Google, and MIT Lincoln Laboratory. We also thank Flickr users for use of their photos
Cornell Participants (and alumni): Noah Snavely (Cornell), David Crandall (Indiana University), Daniel Hauagge (Cornell), Kyle Wilson (Cornell), Song Cao (Cornell), Yin Lou (Cornell), Yunpeng Li (EPFL), Andrew Owens (MIT), Johannes Gehrke (Cornell) Dan Huttenlocher (Cornell) Other Collaborators: Sameer Agarwal (University of Washington and Google), Brian Curless (UW), Yasutaka Furukawa (Google), Steve Seitz (UW),Ian Simon (UW), Rick Szeliski (Microsoft Research)

B. Saving History with 3D Eyes Project :

The project is an effort by the British Institute of Digital Archaeology in cooperation with Dubai Foundation for the Future and UNESCO to document archaeological sites**.**

Figure 6. Reconstruction of places relying on technology in repairing damaged monuments

J. Time is known to leave scars on historic sites and human monuments. Changing climate factors and wars have their devastating impact on archaeological heritage with an important historical and cultural significance.

It is not new information that time leaves on the face of the ruins. Everyone knows about the apparent damage to the face of the Sphinx in Egypt, and the grazing holes in the Inca town of Machu Picchu in Peru. Alexei Karenauska, director of technology at the Institute of Digital Archaeology in London and a researcher at Magdalen College, confirmed the increasing violence against historic sites by extremist groups, according to the website of the UAE Future Observatory.

"There is a growing awareness among all those interested in cultural heritage and heritage of the dangers of various organized cultural cleansing programs, which are being seen by various terrorist and terrorist organizations around the world," Kareenowska said in interviews. Karenauska and her team at the Institute of Digital Archaeology seek to preserve historical artefacts from damage through 3D imaging. Kareenowska collaborates with the Dubai Foundation for the Future, UNESCO and a host of other partners in capturing millions of 3D images of endangered objects around the world. Through these efforts, the group hopes to have a publicly available database, which anyone can use, being made for scientific research or for fun by strolling through history.

"The sites targeted and destroyed, some of them well documented, but many are not," she says. "So our vision was a project that uses stereoscopic camera technology to document threatened sites." Images taken in the team's portal are saved to the open-source million-image database. The team employs a group of indigenous people on a daily basis for areas with damaged and destroyed monuments, familiar with their area and its historical and cultural value, taking pictures and providing them with the team.

"Local people basically take photographs with cameras we provide them for stereoscopic images and then return them to the UK for processing," Kareenowska explains. "The process is as simple as it seems, yet, despite its simplicity, it is a project that takes a long time. Of course, keeping all our antiquities is a huge challenge."

The Kareenowska team provides state-of-the-art technologies to carry out the mission. The cameras are expensive, light-weight, eye-catching and easy to use. About 5,000 cameras are scheduled to be distributed in 2016 only. "We are very grateful for the support of the Government of the United Arab Emirates (UAE) and our other partners at many levels," she said. "This support has allowed us to support the support of the Government of the United Arab Emirates and our other partners at many levels. Expanding the scope of the project is much greater than we originally expected." Karinowska acknowledges that the unification of efforts allowed the completion of a real database of history, an achievement that was difficult to imagine just a few years ago.

"Technology has a lot to offer to preserve cultural heritage and appreciate culture. Today we can use analytical techniques. Even 10 years ago, it looked like something only in Star Trek. We can use this technology to know the age of things," she says. , Their origin, the type of material used in their manufacture, and so on. " Karinowska describes the work as a "technological tool for the preservation and study of cultural heritage and archaeology," where three-dimensional imaging allows the restoration of life to the ruins and its historical preservation through photographs.

C. AirPano.com project

AirPano is a not-for-profit project created by a team of Russian photo enthusiasts focused on taking high-resolution aerial panoramic photographs. Although we usually photograph from a helicopter, we also like to shoot from an airplane, a dirigible, a hot air balloon, and a radio-controlled helicopter. Today AirPano is the largest resource in the world – by geographical coverage, number of aerial photographs, and artistic and technical quality of the images — featuring 360° panoramas of the highest quality shot from a bird's eye view.
Since the beginning of the project, the AirPano team has created panoramas of a majority of the most interesting and significant places and cities of the Earth. There are about 3000 aerial panoramas on our website showing more than 300 famous locations of our planet, including the North Pole, Antarctica, the Mariana Trench, and even the Earth's view from the stratosphere. New virtual tours are published on www.Airpano.com every week. Unlike viewing a traditional photograph limited by its frame, or watching a film that follows its director's idea, a spherical panorama gives you a freedom of being on location, where you can turn around 360° and have a closer look at every detail. When viewed on a large screen the panorama technology creates a complete illusion of a personal presence on the spot, it makes you feel as if you are in the sky above New York City, underwater, above the raging waterfall, and even inside a microwave oven. In 2012 the AirPano project won the "RuNet Awards 2012" contest. In 2013 we received a grant from the Russian Geographic Society (RGO) for further development of the project. Russian President V. V. Putin personally presented the grant to our colleague Sergey Semenov. The AirPano photographs and panoramas were published in "National Geographic", "GEO", "Der Spiegel", "Daily Mail", and many other publications on multiple occasions. Year after year we continue winning numerous international and national photo contests. During the whole five years since 2011 to 2015 AirPano has been permanently winning in the prestigious international panoramic photography contest "Epson International Pano Awards". These events are covered in our News section in detail.

Figure 7. Lucerne, Switzerland (ID:7908)

II.V Conclusion

It is fact that people love to share and exchange information and the mobility has contributed a lot to this habit. The project is suggesting and proposing a new social network concept that takes the place as an object of knowledge and culture and experiences of people supposed to build this place virtually using their experience. The proposed system will be an additional data source specially with wars happening everywhere destroying places and vanishing valuable culture we do not have enough documentation of . Mobility facilities help a lot to reach places that we cannot reach or it will cost a fortune to cover . People are the main partners in this social network and it will have all facilities and interesting features as other social networks. The new concept should be able to be integrated into the social networks and should be working in parallel with that information transfer and more convenient use.

The system will present new fields and tools that help people to construct places and imagine walk through to any place they like, for example building a walk-through from the airport to the place they want their guest to visit . The proposed system is a new version of scattered experiences and computer tools on internet that comes in a more moderated platform.

References

[1] Virtual Reconstruction of Lost Architectures from the TLS Survey to AR Visualization. R. Quattrini a*, R. Pierdicca b , E. Frontoni b , R. Barcaglioni a a DICEA, Dept. of Civil and Building Engineering and Architecture, Polytechnic University of Marche, Ancona, Italy – r.quattrini@univpm.it, roberta_ba76@tin.it b DII, Dept. of Civil and Building Engineering and Architecture, Polytechnic University of Marche, Ancona, Italy – (r.pierdicca, e.frontoni)@univpm.it Commission VI, WG VI/4 The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLI-B5, 2016 XXIII ISPRS Congress, 12–19 July 2016, Prague, Czech Republic.

[2] Post Disaster Virtual Revival: 3D CG Manual Reconstruction of a World Heritage Site in Danger Elham Andaroodi, Mohammad Reza Matini and Kinji Ono Additional information is available at the end of the chapter http://dx.doi.org/10.5772/46406.

[3] 3D Research Challenges in Cultural Heritage II [edited by Sander Münster, Mieke Pfarr-Harfst, Piotr Kuroczyński, Marinos Ioannides] Available Online https://books.google.com.sa/books?id=Az5QDQAAQBAJ&pg=PA275&lpg=PA275&dq=bigsfm+:+reconstruction+the+word+from+internet+photos&source#v=onepage&q&f=false.

[4] Behrouz Jedari and Feng Xia, "A Survey on Routing and Data Dissemination in Opportunistic Mobile Social Networks,"http://arxiv.org/abs/1311.0347.

[5] O. Khalid, S. U. Khan, S. A. Madani, K. Hayat, L. Wang, D. Chan, and R. Ranjan, "Opportunistic Databank: A Context-aware on-the-fly Data Center for Mobile Networks," in Handbook on Data Centers, S. U. Khan and A. Y. Zomaya, Eds., Springer-Verlag, New York, USA.

[6] J. Kolodziej, S. U. Khan, L. Wang, N. Min-Allah, S. A. Madani, N. Ghani, and H. Li, "An Application of Markov Jump Process Model for Activity-Based Indoor Mobility Prediction in Wireless Networks," in *9th IEEE International Conference on Frontiers of Information Technology (FIT),* Islamabad, Pakistan, December 2011, pp. 51-56.

[7] Opportunistic mobile social networks: From mobility and Facebook friendships to structural analysis of user social behaviour, A. Socievole, F. De Rango, A Caputo [http://dx.doi.org/10.1016/j.comcom.2016.04.025]

[8] Nikolaos Vastardis and Kun Yang, "Mobile Social Networks: Architectures, Social Properties, and Key Research Challenges," IEEE Communications Surveys and Tutorials, vol. 15, no. 3, 2013, pp. 1355.

[9] Expert Systems with Applications. An International Journal, ISSN: 0957-4174.

[10] The Application of Expert Systems: A Review of Research and Applications. C. F. Tan1 , L. S. Wahidin1 , S. N. Khalil1 , N. Tamaldin1 , J. Hu2 and G.W. M. Rauterberg2 1Integrated Design Research Group, Centre for Advanced Research on Energy, Universiti Teknikal Malaysia Melaka, Durian Tunggal, Melaka, Malaysia 2Designed Intelligence Group, Department of Industrial Design, Eindhoven University of Technology, Eindhoven, Netherlands E-Mail: cheefai@utem.edu.my

WEB ADDRESSES

[11] http://www.cs.cornell.edu/projects/p2f/

[12] [http://www.middle-east-online.com/?id=230708]

[13] http://www.airpano.com/Project-AirPano.php

[14] Wikipedia.com

Books :

[12] Opportunistic Mobile Networks Advances and Applications [2016] . Sudip Misra, Barun Kumar Saha, Sujata PalISBN: 978-3-319-29029-4 (Print) 978-3-319-29031-7 (Online)

[13] The Diameter of Opportunistic Mobile Networks, Augustin Chaintreau , Abderrahmen Mtibaa Laurent Massoulie Christophe Diot Thomson, Paris, France. email: firstname.lastname@thomson.net.

LATE PAPERS

Driving Simulator Development
Phase IV – Defining Hexapod Kinematics by Driving Simulation

Valerian CROITORESCU, Rafael HARITON

Scientific Research and Continuous Training Center for Sustainable Automotive Technologies
University POLITEHNICA of Bucharest, 313 Splaiul Independentei st., 6th Sector, 060042, Room JC004, Bucharest, ROMANIA

E-mail: valerian.croitorescu@upb.ro

KEYWORDS
Vehicle dynamics, motion control, driving simulator, inverse kinematics, forward kinematics, trajectory

ABSTRACT

The driving simulation plays an important role for defining the hexapod kinematics. The driving simulator uses a 6 DOF hexapod. The complex motion of the hexapod needs a mathematical description for forward and inverse kinematics. This paper presents the approach to define hexapod kinematics using algorithms and geometric computation starting from the already investigated driving simulator motions and behavior. Based on this approach, the actuators active length is also determined.

INTRODUCTION

Driving simulators are complex mechatronic systems that are able to develop and to define the vehicle dynamics and vehicle behavior in very difficult conditions almost impossible to be reproduced in the real world. The very complex motions defined by the upper platform and by the actuators is difficult to be reproduced virtually and then implemented as predefined driving cycles.
The driving simulator is based on a hexapod containing an upper platform and a lower platform. The upper platform supports the vehicle and the dome, if existing. The upper platform is linked to the lower platform using six different actuators. Both the upper and the lower platforms are using joints and both actuators heads, for being able to define the motion according to the vehicle dynamics.
The upper platform motion is following the active lengths for the actuators and is running using the classic motion algorithm, presented in (Tuca et al, 2015-1). The upper platform allows operations for all six degrees-of-freedom of the Cartesian inertial frame.
The motion platform dimensions correspond to the vehicle type, dimensions and mass. The vehicle is placed on top of the motion platform and all the driving profiles are realized by the actuators. The upper platform can have different shapes, but an equilateral triangle is the main geometrical shape that the upper platform is based on. The most important geometrical point of the motion platform is the motion reference point MRP (figure 1), as defined in (Tuca et al, 2015-1).The motion reference point (MRP) denotes the point in space at which the platform translations and rotations are centered (Nahon and Reid, 1990). The motion reference point is defined to be the geometrical center of the equilateral triangle, being responsible to act as the origin of the entire system, where all the forces, torques and X-Y-Z displacements are sensed. For the hexapod system, although the location of the MRP can vary, it is usually located with respect to the geometry of the motion platform. Most commonly, it is defined as the centroid of the two triangles formed at the upper joint rotation points. The motion platform consists of the Stewart platform, a hexapod including an octahedral assembly of joints (figure 2). The upper platform rigidly holds the motion platform as far as the both upper and motion platforms are one system, while driving the simulator operations. The upper platform recreates the vehicle behavior during different driving scenarios, including acceleration, braking, cornering.
The actuators are defined as moving legs because of their active displacement stroke that is being able to control the tilt angles and the position of the upper platform. The motion platform defines the tilt angles and the motion for entire operations on all trigonometrical axis: three translational motions (surge, sway, and heave) and three rotational motions (pitch, roll, and yaw). The actuators and the joints limit both the tilt angles, the rotation and the displacements of the upper platform because the actuators are able to operate only for the active strokes and the joints have an active area and a fixed point being a part at least of the upper platform that limits the rotation to less than the 360 degrees on all directions.

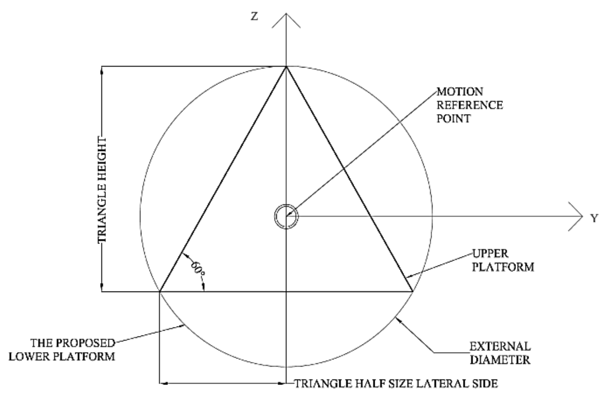


Figure 1: The upper platform and the motion reference point (Croitorescu and Hariton, 2016)

Figure 2: The hexapod sketch – part of the driving simulator virtual model (Croitorescu and Hariton, 2016)

The kinematics consists of studies regarding motion for a complex system, without taking into account the forces and the moments that cause the motion. The hexapod motion is based on kinematics and dynamics, motion planning and control, using mathematical equations and algorithms. The hexapod kinematics include the forward and the inverse kinematics (figure 3).

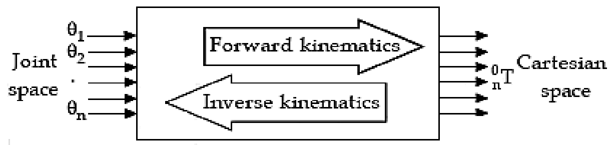


Figure 3: The forward and inverse kinematics basics (Kucuk and Bingul, 2006)

The hexapod kinematics include the kinematic chain (joints, links, flanges, actuators etc.) which act as an open or a closed loop. The hexapod forward kinematics consist of the known position or velocity for each joint and the aim to find out the end position or velocity during and after driving simulator operations. The hexapod inverse kinematics consist of the known desired position or velocity of the end point where the joint must reach and the aim to find out the position or velocity of each joint. The inverse kinematics task is very complex. Solving it includes trigonometrical equations, complex matrix and a huge work amount of time.

This paper aims to highlight the approach on how the hexapod kinematics can be defined using the driving simulator with six degrees of freedom (6 DOF) and using that as a basic approach, on how the actuators active stroke can be determined.

FORWARD AND INVERSE KINEMATICS

Following (Omran et al., 2009), two different strategies are able to be implemented for investigating the kinematics, i.e. both forward and inverse kinematics, for achieving an appropriate control strategy.

The hexapod kinematics control model consists of a joint space control model and a task space strategy control model. The difference between these two models consists of the number of input-output systems that controls the model. Therefore, the joint space control model (figure 4) has only one single-input and one single-output system, while the task space strategy control model (figure 5) has multi-input and multi-output systems. The \`a\` refers to task space displacements and \`q\` refers to joint space displacements.

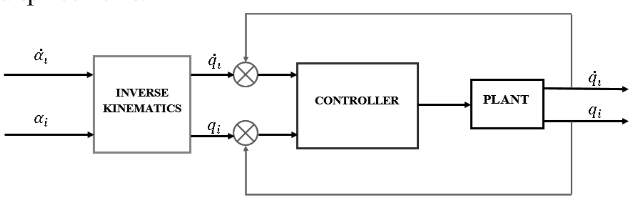


Figure 4: The joint space control model (Omran et al., 2009)

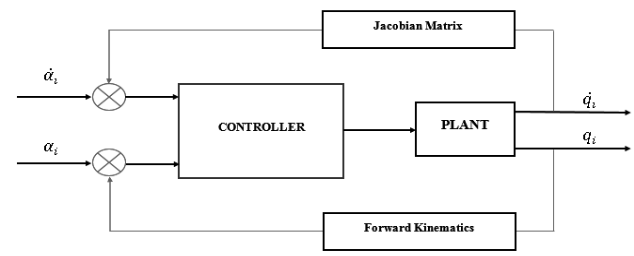


Figure 5: The task space strategy control model (Omran et al., 2009)

The joint space control model includes the inverse kinematics model. The hexapod has a non-linear coupling behavior, therefore the joint space control model is not as accurate as needed.

The inverse kinematics model is able to determine the link lengths in terms of desired/specified upper platform position and angular displacements.

The *forward kinematics* of the hexapod has a mathematical background definition for the position of the upper platform. For investigating the forward kinematics, a coordination frame is attached to each joint.

The upper platform has three different pairs of spherical joints, each pair has two different joints. The spherical joints have axial rotation, their maximum rotation being constrained by their own design. The ground platform has another three different pairs of spherical joints, each pair has two different joints. Each one by one joint from the upper platform is connected with another one by one joint from the lower platform (figure 6). Each joint has its own center of gravity, J_{CoG} to whom the forces and the moments are expressed from the motion rotation point. All these forces and torques are responsible for the translational motion and the inclination angles.

The upper platform is linked to the bottom platform by linear cylinders acting as actuators. The hexapod ground fixed platform is linked by six identical kinematic links to the upper mobile platform. The kinematic links consist of six spherical joints and six universal joints. The joints are located at the ends of the actuators that consist of a screw ball and a rod, working inside a shell (Croitorescu et al, 2015). Three different pairs of joints are linking another pair of two neighboring actuators to the upper platform. The bottom platform integrates another three pairs of different joints, each one for a different pairs of actuators.

The actuators are using defined control corresponding to all the vehicles motions. All the forces and torques acting on the vehicle are being reproduced via the driving simulator actuators. Therefore, the hexapod motions are directly related with the vehicle motion and behavior.

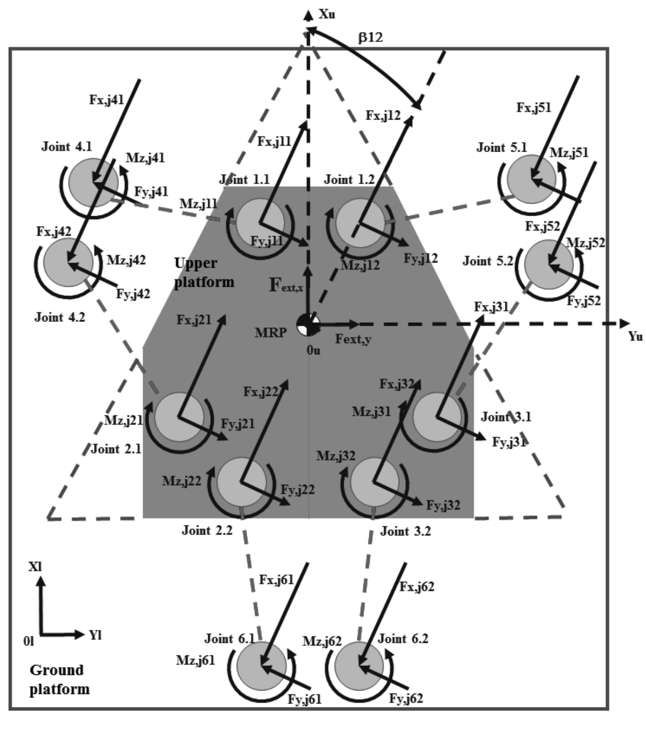


Figure 6: Forces acting on simulator with respect to the coordination system

The upper platform is making the same body motion with the motion platform being defined by the motion reference point position along the coordination axles and by the motion platform inclination angles that represent the roll, the pitch and the yaw motions (figure 7).
The surge motion consists of the forward and backward translations along the X-axis.
The sway motion consists of the sideways translation along the Y-axis.
The heave motion consists of the vertical translation along the Z-axis.
The pitch motion consists of the tilting rotation around the Y-axis.
The roll motion consists of the tilting rotation around the X-axis.
The yaw motion consists of the horizontal rotation around the Z-axis.

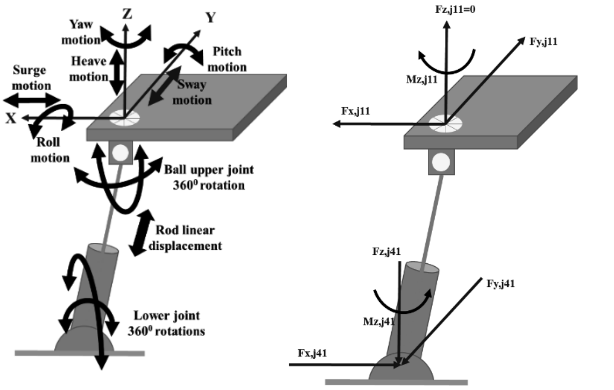


Figure 7: The motion platform behavior related to one actuator and the forces acting for the upper and lower platforms joints (Croitorescu and Andreescu, 2016)

Based on the six different actuators and on the 12 different joints, the upper and the lower platform are defining two different vectors, one for each platform:

$$\vec{u_i} = \begin{bmatrix} u_{xi} \\ u_{yi} \\ u_{zi} \end{bmatrix} \tag{1}$$

$$\vec{l_i} = \begin{bmatrix} l_{xi} \\ l_{yi} \\ l_{zi} \end{bmatrix} \tag{2}$$

where $\vec{u_i}$ is the vector corresponding to the upper platform and $\vec{l_i}$ is the vector corresponding to the lower platform, and i =1…6 is corresponding to the investigated joint.
Since all forces and torques coming and going to the motion platform act in the MRP, the motion platform behavior is also defined by a vector of the MRP translations:

$$\overrightarrow{MRP} = \begin{bmatrix} MRP_x \\ MRP_y \\ MRP_z \end{bmatrix} \tag{3}$$

A universal rotation vector is defining the conversion from the roll-pitch-yaw angles according to (Paul, 1981):

$$\vec{R} = \theta \cdot \vec{k} \tag{4}$$

where $\vec{k}$ is the rotation vector, θ is the rotation angle.

Inverse kinematics is the mathematical process that reproduces and determines the joint parameters that provide the desired position for each of the hexapod moving components. The motion is transformed in joint actuator trajectories.
The inverse kinematics are needed to define the hexapod control, even if it is computationally expansive and takes a lot of time in the real time control. By using the Cartesian coordinates system, the actuators are working in the joint space. The inverse kinematics are able to be solved geometrically and analytically.
The geometrical approach consists of decomposing the spatial geometry of the hexapod into several plane geometry problems, for example by investigating only a joint and the actuator length corresponding to that joint motion. Being a plan geometry problem, all the equations are easy to be determined using trigonometrical functions as cos and sin, for the X0Y reference system, allocated to each joint.
The analytical approach consists of using matrix and algebraically equations, including nonlinear equations. Several solutions for some trigonometric equations are already given by literature. The solution for inverse kinematics is given by the hexapod structure, therefore the driving simulator complexity has to be taken into account.
Two different coordination frames are used, one for the upper platform, the Xu0uYuZu, and one for the lower platform, the Xl0lYlZl, both located in the centers of the upper, respectively the lower platform, in both cases the Z axis being pointed outward. Each joint center is being in line with the center of the platform (upper or lower) and the resulted line Jij with the X axis is defining an angle, βij, where i=1…6 represents the number of the joint and j=1,2 represent one joint from each pair of two (figure 6). For example, the angle β12 is the angle defined by Xu0u and the joint 12 line, J12 (figure 6).
The center of the joints corresponding to both platforms are disposed on imaginary circles having different radius size, Ru for the upper platform (figure 8) and Rl for the lower platform. The upper platform joints Jij (J11, J12, J21, J22, J31, J32) position is defined by:

$$\vec{J_{ij}} = \begin{bmatrix} {J_{ij}}_{xi} \\ {J_{ij}}_{yi} \\ {J_{ij}}_{zi} \end{bmatrix} = \begin{bmatrix} R_u \cos \beta_{ij} \\ R_u \sin \beta_{ij} \\ 0 \end{bmatrix} \tag{5}$$

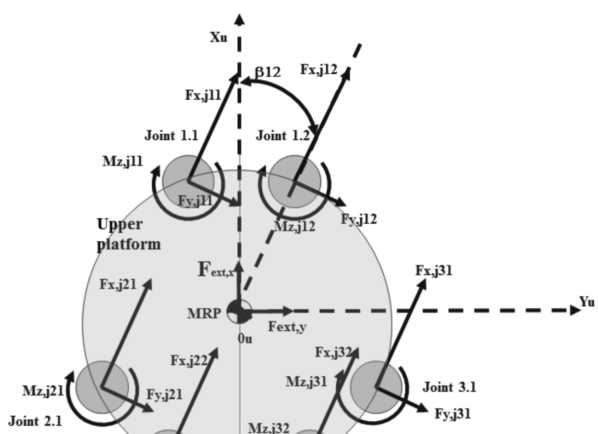


Figure 8: The imaginary circle corresponding to the center of the joints from the upper platform

By using the same approach, in the same manner, the lower platform joints are defined as J41, J42, J51, J52, J61, J62.
The upper platform rotation defines a rotation matrix, ROT, by the roll, pitch and yaw angles, including the following: ROTx(α) is the rotation of α about the fixed X axis, ROTy(ɵ) is the rotation of ɵ about the fixed Y axis, ROTz(γ) is the rotation of γ about the fixed Z axis:

$$\mathrm{ROTz}(\gamma)=\begin{pmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{6}$$

$$\mathrm{ROTy}(ɵ)=\begin{pmatrix} \cos & 0 & \sin \\ 0 & 1 & 0 \\ -\sin & 0 & \cos \end{pmatrix} \tag{7}$$

$$\mathrm{R_x}(\varphi)=\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix} \tag{8}$$

$$ROT = ROT_Z(\gamma)ROT_Y()ROT_X(\varphi) = \tag{9}$$

$$\begin{bmatrix} cos\gamma cos & cos\gamma\ sin\varphi\ sin - cos\varphi sin\gamma & sin\varphi sin\gamma + cos\varphi\ cos\gamma\ sin \\ cossin\gamma & cos\varphi\ cos\gamma + sin\varphi\ sinsin\gamma & cos\varphi\ sin\ sin\gamma - cos\gamma sin\varphi \\ -sin & cossin\varphi & cos\varphi cos \end{bmatrix}$$

With respect to the fixed lower platform, the upper platform also defines the translation vector of its origin (equation 3).

Therefore, the locations for the upper platform joints, Jij-u | i=1,2,3, j=1,2| and for the lower platform joints Jij-l |i=4,5,6, j=1,2|, together with the rotation matrix, ROT, and the vector that denotes the translation of the origin of the upper platform with respect to the lower platform, MRP, defines the vector LLi of the link i:

$$LL_i = ROT \cdot \mathrm{Jiju} + MRP - \mathrm{Jijl} \tag{10}$$

The rod available linear displacement is defined by the maximum rotation angle of the spherical joints. The length varies as the upper platform changes the equilibrium position. The force feedback that comes from each actuators depends on the load capacity, being different distributed to each actuator. Therefore, the driving simulator uses a complex function in order to reproduce the vehicle behavior:

$$f_{motion} = f_{motion}(l_j, \tau_j, G_j) \tag{11}$$

where l_j is the maximum available linear displacement for the `j` actuator, τ_j is the angle that can be obtained between the rod axle and the upper platform geometrical plan, G_j is the load for the `j` actuator.
For the known position and orientation of the upper platform, each actuator maxim active length is defined as follows:

$$\begin{aligned} {AL_{max}}^2 = {} & [MRPx - Jijlx + Jijux \cdot cos\gamma cos + Jijuy \\ & \cdot (cos\gamma\ sin\varphi\ sin - cos\varphi sin\gamma)]^2 \\ & + [MRPy - Jijly + Jijux \cdot cossin\gamma + Jijuy \\ & \cdot (cos\varphi\ cos\gamma + sin\varphi\ sinsin\gamma)]^2 \\ & + [MRPz + Jijux \cdot (-sin) + Jijuy \\ & \cdot cossin\varphi]^2 \end{aligned} \tag{12}$$

The task space coordinates (surge, sway, heave, roll, pitch, and yaw) are defined as functions of joint space coordinates (length of each actuator).
The actuators are identical kinematic chains that achieve different active lengths being able to be controlled separately. Therefore, the upper platform is performing all the requested motions.

DISCUSSIONS

The hexapod kinematics are strong related to the vehicle motion used inside the simulator. As far as the vehicle is following different driving scenarios, its behavior is very complex. Electric actuators are used to enhance vehicles dynamics performances. As a result of the fast response the actuators are able to give under several requests from the driver, the vehicle follows strategies for secure and safe operation between two different destinations while successfully negotiating the traffic and obstacles.
The vehicle`s behavior is assimilated in the upper driving simulator platform. Originating from the driver making control inputs from the vehicle cab, each sub-system plays a vital role in forming the perception of the virtual driving environment. While the vehicle accelerates, the upper platform will change its position (front side will go higher, while the back side will go down) and the actuators from the front will be decompressed while the actuators from behind will be compressed. This study case refers to the active length of the actuators that are determined by using forward and inverse kinematics.
Taking into account all the requirements for the driving scenarios assistance and for the optimization of the actuators response, the major challenge is the dedicated control of the entire system in order to achieve operation close to reality. But the complicated control is strong linked with the complicated hexapod motion that is bringing difficultness for investigating the kinematics.
The hexapod kinematics uses the advantage of its strong and robust mechanical structure, but meets difficulties in operation because of its dimensions and limited actuator active length. The active actuators length is one of the most difficult tasks to be solved even if it can be independently determined.
The existence of the several kinematic chains, all connected by the motion platform, requires a complex description of its dynamics. The hexapod is divided into two different parts, the motion platform (the upper platform) and the actuators (the moving legs). Both divided systems have to follow the trajectories according to the vehicle driving tasks. Therefore, by a mathematical approach, the different trajectories can be computed.
In order to define the hexapod dynamics from a new perspective, without using any vehicle, the driving simulator's behavior offers understanding and learns how all the key sub-systems of the driving simulators are working and the importance of each of them is highlighted; the static limits for the motion platform, the actuating cylinder and joints operation boundaries, and how to simulate different moving scenarios starting from a sketch.
The understanding of the hexapod kinematics using the driving simulator is very useful from small to large scale laboratory work and teaching. The advantages consist also of learning how to set up and how to implement the instrumentation, how to understand the mathematical modeling (algebraically and geometrically) and where the reference sensors positions for better investigation are chosen, needed for testing.

CONCLUSIONS

The purpose of the different approaches is to perform a definition for hexapod kinematics.
The algorithms and the motion platform kinematics can be optimized as a demand of all the basic vehicle motions, taking into account the necessities of assisting the driving and optimizing the simulator control systems.
The objectives include the definitions for all the basic hexapod motions that define the translations and the rotation angles.
Both the forward and inverse kinematics for the hexapod were investigated and defined.
The task for the forward kinematics consisted of tracking the position for the motion platform and its inclination/orientation. The mathematical approach may sometimes skip to an equivalent trajectory. It cannot be proven that the defined solution is complete and it is able to define all the motion point of the trajectory.
The forward kinematics concerns itself with the relationship between the individual joints and the position and orientation for the motion platform. By determining the positon and the orientation of the motion platform, the forward kinematics investigated the joint variables that are the angles between the rotational joints.
The inverse kinematics has been studied, representing a strong request for complex systems such as hexapods are. Inverse kinematics investigations needed lots of time and included geometric solution and algebraic computations. The hexapod produced multiple solutions during inverse kinematics investigations, which may not always correspond to the physical solutions. Therefore, each set of resulted values should be checked in order to determine if they are identical or not with the physical limitation. The matrices and the trigonometrical transformations used represent a difficult task.
The inverse kinematics definition was made starting from the vehicle driving simulation during defined cycles.

REFERENCES

Croitorescu, V., Tudorache, Gh., Stamin, Ș., Jiga, G. (2015), Screw Ball Actuator Stress Behavior, *Applied Mechanics and Materials* Vols. 809-810 (2015) pp 616-621, Trans Tech Publications, Switzerland, DOI:10.4028/www.scientific.net/AMM.809-810.616, Proceedings of Innovative Manufacturing Engineering International Conference – IManE 2015, Iași, Romania

Croitorescu, V.; Andreescu, C., (2016), *Driving Simulator Development: Phase III – Defining Basic Vehicle Dynamics by Driving Simulation,* Proceedings of the 14th Annual Industrial Simulation Conference, Bucharest, Romania, EUROSIS-ETI Publication, ISBN: 978-90-77381-93-9, EAN: 9789077381939, INSPEC referenced, BDI indexed in SCOPUS, 2016

Croitorescu, V., Hariton, R., (2016), Proceedings of International Conference on Innovative Technologies, IN- TECH 2016, Prague, 6- 8.9.2016, Publisher: Faculty of Engineering University of Rijeka, ISSN 1849-0662

Kucuk, S., Bingul, Z., (2006), *Industrial-Robotics-Theory-Modelling-Control, Robot Kinematics: Forward and Inverse Kinematics*, ISBN 3-86611-285-8, pp. 964, ARS/plV, Germany, December 2006, Edited by: Sam Cubero

Nahon, M. A., Reid, L. D. (1990), Simulator Motion-Drive Algorithms: A designer's perspective, *AIAA Journal of Guidance, Control and Dynamics*, 13(2), pp. 356-362, 1990

Paul, R.P., (1981), Robot Manipulators, The MIT Press, Cambridge, 1981

Omran, A., Bayoumi,, M., Kassem, A., El-Bayoumi, G., (2009), Optimal Forward Kinematics Modeling of Stewart Manipulator Using Genetic Algorithms, *Jordan Journal of Mechanical and Industrial Engineering*, Pages 280 – 293, ISSN 1995-6665

Tucă, A., Croitorescu, V., Brandmeier, T., (2015) *Driving Simulator Development: Phase I – From State of Art to State of Work*, Proceedings of Industrial Simulation Conference, Universidad Politecnica de Valencia, Spain

BIOGRAPHY

VALERIAN CROITORESCU is Associate Professor of University POLITEHNICA of Bucharest, graduating his PhD degree in hybrid electric vehicles development topic. His main fields of interest are powertrain and alternative powertrain (hybrid, electric) design and development, vehicle dynamics, modeling and simulation of engineering systems, real-time simulation, automotive driving behavior and driving simulation. His strong involvement in international cooperation activities and his academic records include numerous awards and certifications. From 2013, he is Founder and Executive Director of the Scientific Research and Continuous Training Center for Sustainable Automotive Technologies (SRCTCSAT), organized in 2013 inside University POLITEHNICA of Bucharest

LINEAR LAYOUT DESIGN USING THE SOFTWARE TOOL VISTABLE

Nikola Banduka, Ivan Peko, Marina Crnjac
Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture
University of Split
Ruđera Boškovića 32, 21000 Split, Croatia
nikola.banduka90@gmail.com

Igor Bošnjak
Faculty of Mechanical Engineering and Computing
University of Mostar
Matice Hrvatske bb, 88000 Mostar, Bosna and Hercegovinia

Nikola Banduka, Stefan Đurić
Faculty of Engineering
University of Kragujevac
Sestre Janjić 6, 34000 Kragujevac, Serbia

KEYWORDS

Static facility layout, Layout designing, Linear layout, Layout optimization, visTABLE.

ABSTRACT

Some studies show that 30-75% of total cost of a product are material-handling costs. This problem relates to facility layout problems. The best solution to avoid this cost is preventing it in the early stage of the design phase. In this paper linear layout designing of new manufacturing facility for three different products is presented. It is given a comparison of theoretical linear layout model with practical realization, also. Practical realization in 2D environment is realized with software visTABLE for layout modelling, planning and optimization. Three cases of transformation from theoretical linear layout model to 2D environment, are presented. After transformation results are presented, the best solution is chosen. Each of cases are presented in 3D view for better objectivity of manufacturing facility.

INTRODUCTION

Some of the main occupations for industries are reducing of cost and staying competitive on turbulent market of nowadays. Designing layout activity consists of finding of optimal solution for locating manufacturing facilities. The main goal of this activity is cost reduction. Study presented by (Sule 2008) points out that 30-75% of total costs of a product are material-handling costs. Furthermore, 20-50% of total operating costs in manufacturing companies are material-handling costs (Tompkins et al. 2010). One of ways to reduce material-handling costs is optimally locating manufacturing facilities. This activity is a part of a layout designing which addresses to facility layout problems. Optimal layout designing simplifies management, reduces manufacturing cycles and the number of employees, and as the most important reduce surface required for system set up or production arrangement (Dolgui and Proth 2010). Reducing of this activities usually leads to quality and system reactivity increment, cost and work-in-progress (WIP) reducing, increase of productivity and effectiveness, etc.

There are two models of manufacturing layouts: Static facility layout (SFL) and dynamic facility layout (DFL). The topic of this paper is mainly oriented to SFL. Interests for layout optimization increased in 1980s. At that time, layouts still was steady with stable material flow and deterministic operating times. From then until now, there were many papers written. Many different approaches and methods were presented. But still there is a lot of space for research because the need for facility layout problems diverges (Drira 2007). Development and fast progress of concurrent engineering technology lead to need for including of a third dimension in layout planning and designing. 3D Software tools on global market are more and more suitable for layout designing and optimization. According to (Drira 2007) combination of the layout designing and optimization with graphical tools would result with more efficient and attractive solutions. One of this software solutions is software visTABLE for layout modelling, planning and optimization. This software has a combination of layout optimization tool and virtual reality (VR) graphical solution.

Therefore, in this paper will be presented layout designing and optimization of linear layout type using the software visTABLE. Achieved layout model will be presented in a full virtual tour, also. This study addresses the problems of transformation from theoretical layout model to real environment.

LAYOUT DESIGNING

The idea was to transform predefined theoretical linear layout model to 2D environment, which is not simple. The biggest problem is missing of starting point. Some authors used different methods to solve this problem, but there are still problems during the transformation phase. Finding of the optimal solution is hard and time-consuming job, also.

Review studies by (Singh and Sharma 2006) and (Drira et al. 2007) point out many different studies for layout improvement during the layout design. (Dolgui and Proth 2010) also gave a many solutions for layout transformation and space utilization. (Xiao et al. 2016) used linear programing and annealing algorithm to address these problems. (Centobelli et al. 2016) reduced lead time by improving the factory layout and material flow optimization in digital factory, using a semi simulation software which has VR solution. (Kanduc and Rodic 2016) used a force generated graph algorithm and simulated annealing for machine layout optimization.

However, none of these solutions are comprehensive. Comprehensive solution has to respond on three main

demands. First, material handling problems or cost reduction, which is the main problem that the FLP should address. Second is space utilization which in many cases is not realistic, because with mathematical models and optimization usually can be found optimal solution, but it doesn't necessarily means that it is realistic one. And third is the functionality and realistic solution. Solution found by theoretical calculation has to be transformed into a functional practical solution. This is the biggest problem during the layout designing because of the influence of many different factors. This is the main reason why some authors recommend using 3D tools during the layout design. 3D can give better objectivity and more realistic overview. Any software which have 2D and 3D simulation options can be used for this purpose. Some of these softwares are: Siemens Technomatix, Dassault Systemes DELMIA, Simio, etc.
In this case of linear layout designing, software tool visTABLE was used. Steps to process realization are presented in Figure 1.

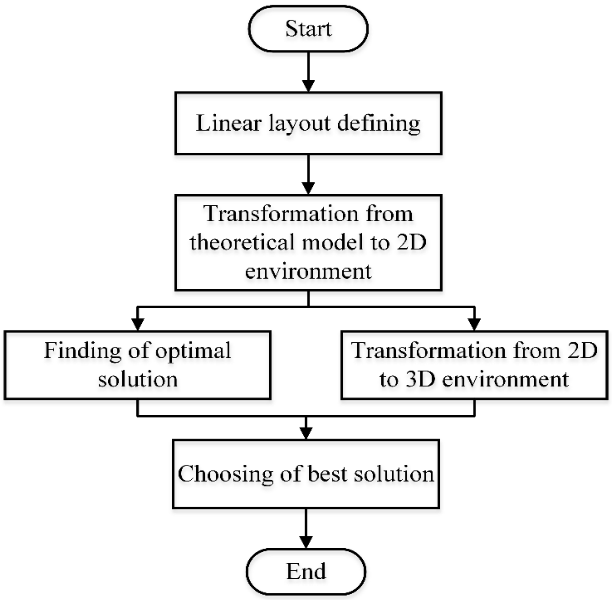


Figure 1. Steps to process realization.

First will be defined theoretical model of linear layout. After that, this theoretical layout will be transformed to 2D environment in software visTABLE. Optimal solution should be found and 2D environment should be transformed to 3D environment for better objectivity. After all these steps, best solution should be chosen.

Linear manufacturing layout

In literature three types of SFL are usually mentioned: linear layout, functional department and cellular layout (Dolgui and Proth 2010). Linear layouts are the most suitable for productions with less variety of products. One of important parameters is stable volumes of products, also. It is desirable for linear layouts to be automatized. High degree of automation gives better productivity, but decreases agility of the system, and reverse. Anyway, the combination of automatized parts of production lines with non-automatized parts of production lines would not affect production system. Therefore, an optimal solution should be found. In the Figure 2 flow chart example of one linear layout is presented.

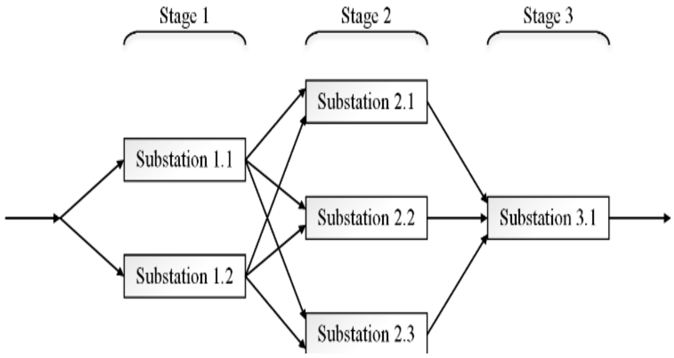


Figure 2: Linear layout (Dolgui and Proth 2010).

The main problem with any type of layout model is a transformation from theoretical to practical model. Therefore, in future, researches on topic related to manufacturing layout should be oriented to solving this problem.

Software visTABLE

This software has integrated planning functions for instant validation of layouts. So, its main purpose is factory layout planning. The second function is very user friendly VR 3D modelling system. A VR can provide a practical (virtual) objectivity in combination with theoretical model. The use of VR for factory layout modelling is very important. Often happens that theoretical models do not fit into the existing factory layout after practical realization. Therefore, by the use of virtual factory models, wastes and additional costs can be avoided, and it is less time-consuming. Furthermore, it can also be used for other purposes, for example, education. visTABLE's architecture provides distribution of business via the internet, which is useful for group planning and modelling, also (Müller and Spanner-Ulmer 2006). The areas of application proposed by visTABLE®touch company are: factory and layout planning, material flow analysis, implementation of the results of a value stream mapping, optimization of logistics processes, assembly planning, scenario planning, running route organization, workshops, etc. Software visTABLE is also recognized in the industry, especially in Germany (Banduka et al. 2017).
There are various authors who have used visTABLE in their research. (Horejsi and Polcar 2013) used visTABLE's automatic converter during layout projecting, and compared it with another developed automatic converter. (Černý, Z. and Bureš 2008) used it as one of three tools for virtual layout projection and analysis with ergonomic aspect. Research was realized in a group of workplaces and whole company layout, but they avoid using it for individual workplace modelling and analysis. (Horbach et al. 2011) used visTABLE in planning and visualization of the concept of building blocks for adoptable factory systems in the experimental and digital factories. (Spath et al. 2006) did layout planning for assembling workstations with workers. VisTABLE was also used in a study with integration of ERP systems, for virtual modelling of ERP systems by (Szendrei et al. 2010). This software was concretely used for making of

one virtual structure of OEM and its internal transport systems. The authors mentioned that bistable saved a lot of time and money by avoiding the necessity to make a physical environment. Beside this industrial purpose, visTABLE has been also widely used for education (Horejsi and Polcar 2013). In general software visTABLE has already been widely used in literature and especially for layout modelling and optimization.

EXAMPLE

Example of a linear layout designing was presented for three different products: Shaft with gear, movable jaw and spindle. For designing of manufacturing layout of these three products, required surfaces was predefined. Surface Input data are presented in: Table 1 for production, Table 2 for storage and Table 3 for administration. The total surface of the building is 864.4m^2. Data on transportation intensity for each of the products are predefined, also. Transportation intensity for a shaft with gear is 11 units per day, for movable jaw 21 units per day and for spindle 31 units per day.

Table 1: Surfaces of manufacturing layout

Production		
Workplace No.	Operation	Surface (m^2)
1	sawing	15.84
2	grinding	7.5
3	drilling	2.91
4	milling	27.22
5	drilling	13.2
6	scraping	42.84
7	grinding	18.36
8	grinding	25.74
9	milling	13.05
10	scraping	28.8
11	manual work	9
12	manual work	9
13	manual work	9

Table 2: Storage surface

Storage	
Type	Surface (m^2)
input	62.03
output	64.5

Table 3: Administration surface

Administration	
Room type	Surface (m^2)
sanitary	25
office	9
tool room	9
control room	9

Linear layout flow was designed in 2D for three before mentioned products. Workplace numbers for flow chart presented in Figure 3, were taken from Table 1.

In Figure 3 material flow across workstations for every of the three products is set. Material flow for three different products is marked with three different colours. Material flow for a shaft with gear is marked with blue coloured arrows, movable jaw with green arrows and spindle with red arrows. This flow chart of material flow was transformed to 2D environment in software visTABLE. The transformation will be realized for three cases. In case 1 workstations with real surface will be transformed to 2D environment in visTABLE. Arrangement of workplaces is set according to the flow chart of Figure 3. As it is mentioned this is ideal case. Therefore, other cases with the more realistic manufacturing layout have to be realized.

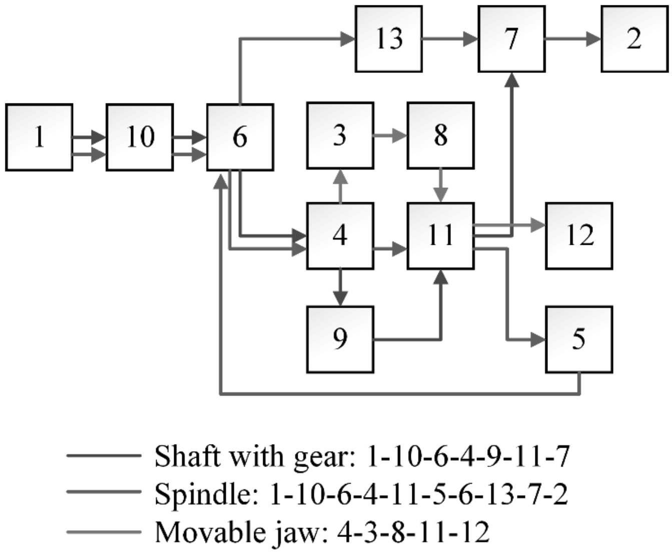


Figure 3: Flow chart of materials for three products for linear layout

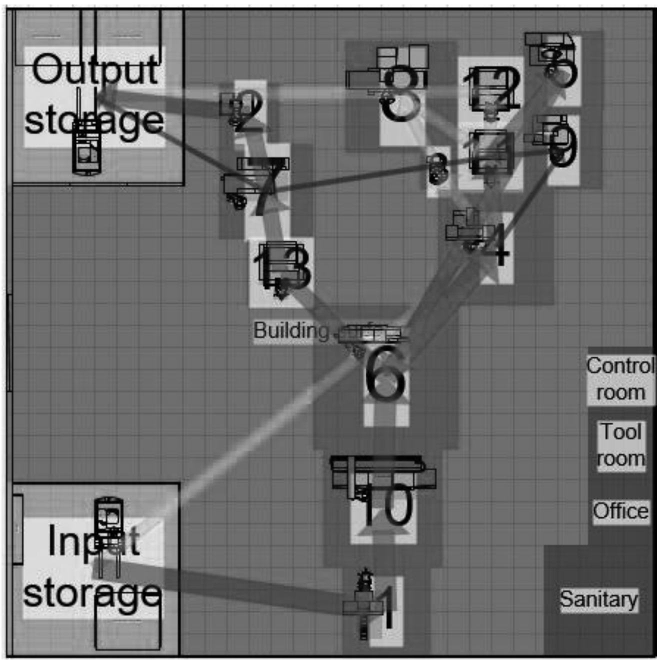


Figure 4: 2D Transformation of linear layout to visTABLE – case 1

In Figure 4 is presented transformation to visTABLE – case 1. As it can be noticed, material flow is not adaptive and functional. There was no space for transport roads. Another problem is that the shortest traveling distances which were taken into account, are not realistic ones. This is one of the reasons why it is the ideal case.
Case 2a presented in Figure 5 is more realistic than case 1. As it can be noticed in case 2a, transportation roads are included and optimal rearrangement of workplaces was founded manually. But traveling distances still were set as the shortest. This is not possible in real industry cases because the material flow has to go true precisely defined routes. Therefore, it has to be set an optimal solution for traveling distances.
Our optimal solution is presented in Figure 6. This is the most realistic case for one real industrial environment – case 2b. The problem with the shortest traveling distances is solved with software visTABLE optimization option. This kind of traveling distances representation is called spaghetti diagram and represents traces for products or activities flow through the production process.
Software tool visTABLE has a great option for optimization of spaghetti diagram. In Figure 6 can be seen additional arrows in the same colour as material flow arrows. These arrows are part of that option. Arrows start from input storage and finish at output storage, true the way of material flow.

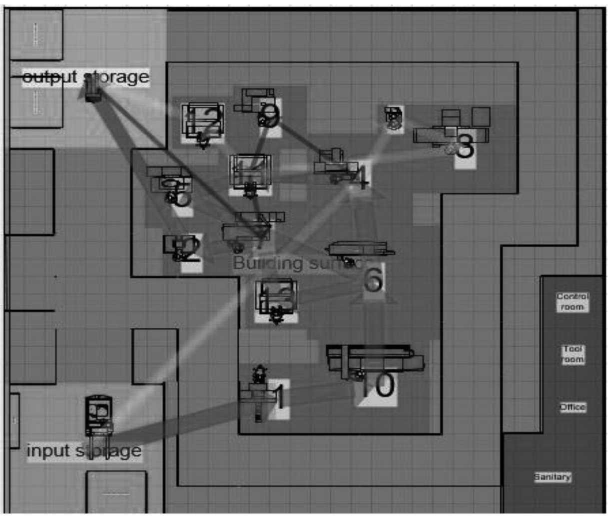


Figure 5: Transported linear layout with roads included and rearranged workplaces – case 2a

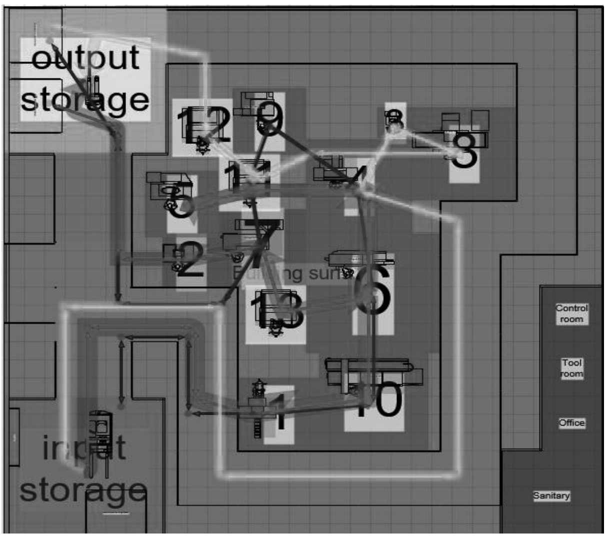


Figure 6: Transported linear layout with our optimal solution of shortest traveling distances – case 2b

RESULTS AND DISSCUSION

The main difference between theoretical and practical realization of the layout is in the absence of reality and objectivity. Theoretical model represents the ideal case which is in the large number of cases impossible. In this paper is presented one transformation of that kind. Measures for relating to layout profitability are traveling distance (in m) and intensity of material flow (in units). During the creation of layout material flow distances between storages and workstations should be short as it is possible.
After transformation from theoretical model to 2D environment in software visTABLE, three cases were realized. Achieved results are presented with three diagrams for every three cases. In Figure 7 are presented results from case 1, in Figure 8 results from case 2a and in Figure 9 results from case 2b.

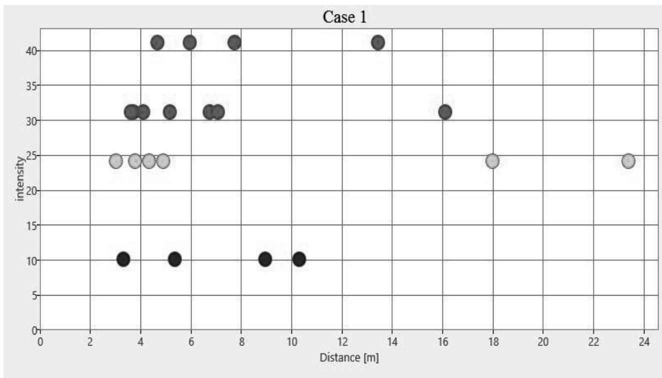


Figure 7: Diagram with case 1 results

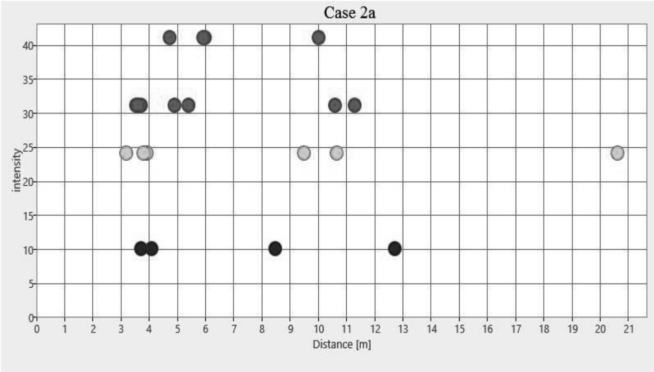


Figure 8: Diagram with case 2a results

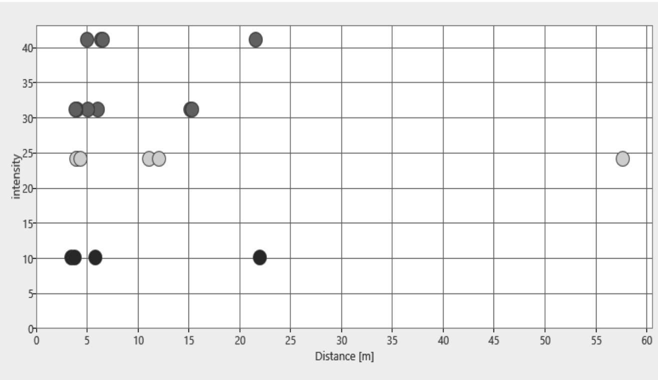


Figure 9: Diagram with case 2b results

In this three cases best solutions can be defined by two aspects (see Table 4).

The first aspect is traveling a distance. Best solution from this aspect is case 2a, then case 2b and the last one is case 1. These results are example of absence of a theoretical model for linear layout designing.

The second aspect is realistic aspect. A realistic solution is one with properly defined traveling routes (not the shortest one) and with functional and adaptive material flow. According to this aspect only realistic solution is case 2b.

Table 4. Ranging of each of the cases

Ranging	Distance	Realistic
1	case 2a	case 2b
2	case 2b	case 2a
3	case 1	case 1

Realistic aspect can be proved with a 3D virtual reality overview. For 3D visualization it will be taken just two overviews, because the case 2a and the case 2b are the same in 3D visualization form, the difference is only in virtual arrangement of material flow. In Figure 10 3D view of case 1 is presented. It can be noticed that this solution isn't practical and functional. There's no space for transportation roads and lots of surface isn't used properly.

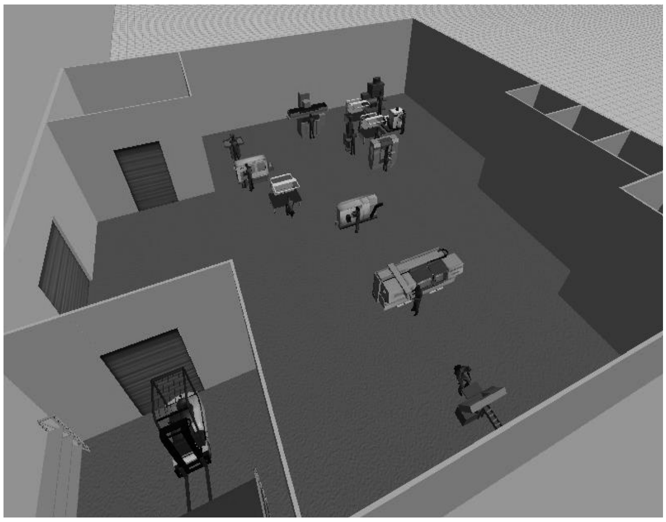


Figure 10: 3D view of case 1

In Figure 11 is presented 3D view of case 2a and case 2b with rearranged positions of workplaces and added transportation roads. With included transportation roads and rearranged workplaces, industry environment gets a realistic dimension.

With this 3D visualization of manufacturing layout, user can get a realistic view of workplace arrangement. Therefore, machines, inventory, logistics, tools, etc. can be manipulated in the goal to achieve the best solution.

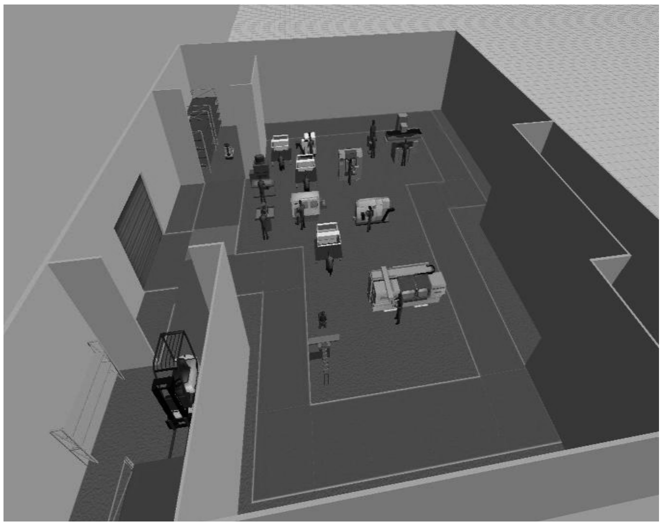


Figure 11: 3D view of case 2a and case 2b

CONCLUSION

Main advantages were achieved with the use of software tool visTABLE. Using of software tool visTABLE helped mostly in transformation phase. First in 2D environment, where a theoretical model of linear layout was transformed. Then, with a 3D option where the functionality of the layout was checked in a virtual environment with a high degree of objectivity. Therefore, on this way was solved main constraint of transformation from theoretical model to practical model. One more constraint is that the theoretical model calculates the shortest traveling distances, only. This problem was also solved with software visTABLE with its optimization function. Use of software tool visTABLE was very user friendly and efficient because optimal solution was found with no use of complicated methods for transformation. In this research are identified few constraints, also. First during the transformation phase, transformation was done manually by users opinion according to the theoretical linear layout model. Second, the optimal solution is our optimal solution which means that does not necessarily mean that it is the best solution.

Future researches should be mainly oriented to finding a better optimal solution for the transformation of theoretical layouts to 2D environment. These solutions are usually complex and demanding, so it may increase time-consuming and complexity of the transformation process.

Future of layout designing, planning and optimization lays in this kind of software tools. Therefore, advanced transformation methods should be combined with some of these software tools. Especially for functionality and suitability checking. 3D options and VR should be used in future, also. Especially for objectivity increment, during the design phase.

ACKNOWLEDGEMENT

This paper has been supported by Croatian Science Foundation under the project Innovative Smart Enterprise – INSENT (1353). This publication also has been supported by the European Commission under the Erasmus Mundus project Green-Tech-WB: Smart and Green technologies for innovative and sustainable societies in Western Balkans (551984-EM-1-2014-1-ES-ERA MUNDUS-EMA2).

REFERENCES

Banduka, N.; Mladineo, M. and Eric, M. 2017. “Designing a layout using a Schmigalla method combined with software tool visTable.” *International journal of simulation modelling*.

Centobelli, P.; Cherchione, R.; Murino, T. and Gallo, M. 2016. “LAYOUT AND MATERIAL FLOW OPTIMIZATION IN DIGITAL FACTORY”. *International journal of simulation modelling*, No. 15 (2), 223-235.

Černý, Z. and Bureš, M. 2008. “SYSTEM APPROACH FOR PRODUCTION ORGANISM DESIGN”. In *Proceedings of the 12th International Research/Expert Conference – TMT 2008,* 1313-1316.

Dolgui, A. and Proth, J. M. 2010. *Supply chain engineering: useful methods and techniques.* Springer Science & Business Media, London.

Drira, A.; Pierreval, H. and Hajrl-Gabouj, S. 2007. “Facility layout problems: A survey.” *Annual Reviews in Control*, No. 31 (2), 255-267.

Gäse, T.; Krauß, A.; Leidholdt, W.; Mauesberger, E.; Swoboda, F. and Urbanneck, C. 2006. “Methoden und Werkzeuge für die vernetzte Fabrikplanung – Ergebnisse des Projektes netPLAN”. In *proceeding of the Vernetzt planen und produzieren VPP 2006.* 145-155.

Horbach, S.; Ackermann, J.; Müller, E. and Schütze, J. 2011. “Building blocks for adaptable factory systems.” *Robotics and Computer-Integrated Manufacturing*, No. 27 (4), 735-740.

Horejsi, P. and Polcar, J. 2013. “An unconventional Software Environment For Factory Layout Design And Automated Converter.” *MM science journal*, No. 6 (4), 444-447.

Kanduc, T and Rodic, B. 2016. “OPTIMIZATION OF MACHINE LAYOUT USING A FORCE GENERATED GRAPH ALGORITHM AND SIMULATED ANNEALING”. *International journal of simulation modelling*. No. 15 (2), 275-287.

Singh, S. P. and Sharma R. R. K. 2006. “A review of different approaches to the facility layout problems”. *International Journal of Advanced Manufacturing Technology*, No. 30, 425-433.

Spath, D.; Müller, F.; Scholtz, O. and Busch, M. 2006. “Visuelle mitarbeiterorientierte Montageplanung”, *Werkstattstechnik*, No. 8 (9), 582-586.

Sule, D. R. 2008. *Manufacturing facilities: location, planning, and design.* CRC press, Northwest, Washington, D.C.

Szendrei, D.; Teich, T.; Unger, K. and Militzer, J. 2010. “Eine integrierte betriebswirtschaftliche Fallstudie mit SAP ERP und deren Einbindung in das SAP LEGO Labor”. In *Proceedings of the Multikonferenz Wirtschaftsinformatik – MKWI 2010*, 1637-1647.

Tompkins, J. A.; White, J. A.; Bozer, Y. A. and Tanchoco, J. M. A. 2010. *Facilities planning.* John Wiley and Sons, Danvers.

Xiao, Y.J.; Zheng, Y.; Zhang, L.M. and Kuo. Y.H. 2016. “A combined zone-LP and simulated annealing algorithm for unequal-area facility layout problem”. *Advances in production engineering & management*, No. 11 (4), 259-270.

AUTHOR BIOGRAPHY

Nikola Banduka was born in Belgrade, Serbia. He went at the University of Kragujevac, Faculty of Engineering. There he finished bachelor studies in 2012 and master studies in 2014. After master studies he enrolled PhD studies in 2014 on the same faculty, and employed in one Slovenian automotive company – Grah Automotive, also. On second year of PhD studies in 2015 he got an Erasmus Mundus scholarship for three years and enrolled University of split, Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture. Despite Erasmus Mundus project he was accepted on another project funded by the Croatian Science Foundation– INSENT (Innovative smart enterprise) 1353. On INSENT project he is working on making of Croatian production system which would fit industry 4.0. Now he is writing a PhD thesis related to reliability. His main topics of interests are Lean, Digital factories, Digital manufacturing, Factory layouts, Failure mode and effect analysis. He published some papers on these topics, also.

AUTHOR LISTING

AUTHOR LISTING