21ST EUROPEAN CONCURRENT ENGINEERING CONFERENCE 2015

ECEC'2015

11TH FUTURE BUSINESS TECHNOLOGY CONFERENCE

FUBUTEC'2015

19TH EUROMEDIA CONFERENCE

EUROMEDIA 2015

EDITED BY

Paulo Jorge Sequeira Gonçalves

April 27-29, 2015

LISBON, PORTUGAL

A Publication of EUROSIS-ETI

Printed in Ghent, Belgium

21ST EUROPEAN CONCURRENT ENGINEERING CONFERENCE

10TH FUTURE BUSINESS TECHNOLOGY CONFERENCE

19TH EUROMEDIA CONFERENCE

2015

LISBON, PORTUGAL

APRIL 27-29, 2015

Hosted by

IST

Organised by

ETI

Sponsored by

EUROSIS

University of Skövde

Ghent University

IST

EXECUTIVE EDITOR

PHILIPPE GERIL
(BELGIUM)

Editor

Paulo Jorge Sequeira Gonçalves

ECEC General Conference Chair
Paulo Jorge Sequeira Gonçalves
Polytechnic Institute of Castelo Branco
Lisbon, Portugal

FUBUTEC General Conference Chair
Philippe Geril
ETI Bvba
Ostend, Belgium

EUROMEDIA General Conference Chair
Paulo Jorge Sequeira Gonçalves
Polytechnic Institute of Castelo Branco
Lisbon, Portugal

Local Programme Committee

Susana Vieira, Instituto Superior Tecnico, University of Lisbon, Portugal
Nuno Fernandes, Polytechnic Institute of Castelo Branco, Portugal
Paulo Brito, Polytechnic Institute of Portalegre, Portugal
Ana Madureira, Polytechnic Institute of Porto, Portugal
Maria José Madeira, Universidade da Beira Interior, Portugal
Paulo Gonçalves, Polytechnic Institute of Castelo Branco, Portugal

ECEC Honorary Chair
Uwe Baake
Daimler Buses, Evobus GmbH
Neu-Ulm, Germany

ECEC International Programme Committee

1. E-Business in CE
Chimay J. Anumba, The Pennsylvania State University, University Park, USA
Heiko Duin, BIBA, Bremen, Germany
Yo-Ping Huang, Tatung University, Taipei, Taiwan

2. Organization and Management
Eduardo Castellano, IK4-IKERLAN, Mondragon-Arrasate, Spain
Iris Graessler, Cologne University of Applied Sciences, Cologne, Germany

3. Supporting Technologies
Sarka Kvetonova, Brno University of Technology, Brno, Czech Republic

4. Formal Methods
Yan Luo, NIST, Gaithersburg, USA
Stefano Marrone, Seconda Universita di Napoli, Caserta, Italy

5. Embedded Systems
Petr Hanacek, Brno University of Technology, Brno, Czech Republic
Jose Machado, University do Minho, Braga, Portugal
Stefano Marrone, Second University of Naples, Caserta, Italy

ECEC International Programme Committee

6. Engineering Data Management and Information Modeling
Uwe Baake, Daimler Buses, Evobus GmbH, Neu-Ulm, Germany
Jiri Vondrich, Czech Technical University, Prague, Czech Republic

7. Engineering Process Management and Modelling
Antonio Carvalho Brito, FEUP, University of Porto, Porto, Portugal
Imed Kacem, University of Technology of Troyes, Troyes, France
Carlo Meloni, Politecnico di Bari, Bari, Italy
Roberto Montemanni, IDSIA, Manno, Switzerland
Maurizio Palesi, University of Catania, Catania, Italy
Christian Andreas Schumann, University of Zwickau, Zwickau, Germany

8. Collaborative CE environments for virtual teams
Veronique Baudin, LAAS-CNRS, Université de Toulouse, Toulouse Cedex, France
Khalil Drira, LAAS-CNRS, Université de Toulouse, Toulouse Cedex, France
Paul Stratil, Stuttgart, Germany

9. CE Enhanced Lean Manufacturing
Christian Hillbrand, V-Research GmbH, Dornbirn, Austria
Jean-Claude Hochon, AIRBUS, Blagnac, France
Rik Van Landeghem, Ghent University, Zwijnaarde, Belgium
Rob Walker, Anglia Ruskin University, Chelmsford, United Kingdom

10. Practical Applications and Experiences
Florin Pop, University POLITEHNICA Bucharest, Bucharest, Romania

11. Factory Planning and Control for Small and Medium sized Enterprises (SME's)
Sílvio do Carmo Silva, University of Minho, Portugal
Paulo J.S. Gonçalves, Polytechnic Institute of Castelo Branco, Portugal
Nuno O. Fernandes, Polytechnic Institute of Castelo Branco, Portugal
Veronique Limere, Ghent University, Ghent, Belgium

12. Digital Direct Manufacturing
Philippe Geril, ETI, Ostend, Belgium
Peter K. K. Loh, Nanyang Technological University, Singapore

13. Patient Centered Health Care with Meaningful Use of IT
Gurdev Singh, State University of New York at Buffalo, Buffalo NY, USA

14. Student Session
Prof Algirdas Pakstas, London Metropolitan University, London, UK

FUBUTEC International Programme Committee

Statistical Analysis and Data Mining of Business Processes

Paulo Alexandre Ribeiro Cortez, University of Minho, Guimaraes, Portugal
Selwyn Piramithu, University of Florida, Gainesville, USA
Inna Pivkina, New Mexico State University, USA
Danny Van Welden, KBC Bank, Brussels, Belgium

Discrete Event Simulation and Queueing Systems in Business and Economics

J. Manuel Feliz Teixeiro, FEUP-University of Porto, Porto, Portugal
Leonid Novickis, IDC Information Technologies, Riga, Latvia
Kamelia Stefanova, University of National and World Economy, Sofia, Bulgaria

Agents in Business Automation and Economics

Martin Fredriksson, Blekinge Institute of Technology, Karlskrona, Sweden
Markus Koch, Orga Systems GmbH, Paderborn, Germany
Isabel Praca, Polytechnic of Porto, Porto, Portugal

Simulation in Business and Economics

Eduardo Castellano, IK4-IKERLAN, Mondragon-Arrasate, Spain
Sudhir Mudur, Concordia University, Montreal, Canada
Yan Luo, NIST, Gaitherburg, USA
Simon See, Oracle Systems, Singapore

Simulation in Business Games

Renate Sitte, Griffith University, Gold Coast, Australia
David Wortley, Coventry University, Coventry, United Kingdom

Simulation in OR and Knowledge Management

Carmen Bobeanu, Ghent, Belgium
Chrissanthi Angeli, Technological Institute of Piraeus, Athens, Greece
Imed Kacem, Université Paul Verlaine Metz, Metz, France
Christophe Roche, Universite de Savoie, Le Bourget du Lac Cedex, France

Emergency Management and Risk Analysis Management

Ismail Bouassida, LAAS-CNRS, Université de Toulouse, Toulouse, France
Khalil Drira, LAAS-CNRS, Université de Toulouse, Toulouse, France
Konstantinos Kirytopoulos, University of the Aegean, Greece
Gurdev Singh, UB Patient Safety Research Center, State University of New York at Buffalo, Buffalo NY USA
Lode Vermeersch, Delcredere, Brussels, Belgium

Simulation in E-Management, E-Government, E-Commerce and E-Trade

Vrassidas Leopoulos, Nat. technical University of Athens, Athens, Greece

Simulation Tools for Business Management and Business Intelligence Simulation

Carlo Meloni, Politecnico di Bari, Bari, Italy

Geo-Simulation and Big Data

Marwan Al-Akaidi, Arab Open University, Kuwait

Measuring the Costs and Effects of Environmental Policies for Business

Dietmar Boenke, Hochschule Reutlingen, Reutlingen, Germany
Elias A. Hadzilias, Universite Catholique de Lille, Lille, France
Roberto Razzoli, University of Genova, Genova, Italy

Ecological sustainable development and Innovative technologies for a bio-based economy

Philippe Geril, ETI Bvba, Ostend, Belgium

EUROMEDIA International Programme Committee

General Program Chair

Ashwin Ittoo, ULG, Liege, Belgium

WEBTEC Programme Committee

António Carlos Abelha, Universidade do Minho, Braga, Portugal
Riadh Ben Halima, LAAS-CNRS, Université de Toulouse, Toulouse Cedex, France
Boguslaw Butrylo, Bialystok Technical University, Bialystok, Poland
Khalil Drira, LAAS-CNRS, Université de Toulouse, Toulouse Cedex, France
Tom Guérout, LAAS-CNRS, Université de Toulouse, Toulouse Cedex, France
Markus Koch, Orga Systems GmbH, Paderborn, Germany
Joern Loviscach, Fachhochshule Bielefeld, Bielefeld, Germany
Wenji Mao, Chinese Academy of Sciences, Beijing, China P.R.
Thierry Monteil, LAAS-CNRS, Université de Toulouse, Toulouse Cedex, France
Lorenzo Motta, Ansaldo Segnalamento Ferroviaro s.p.a. Genova, Italy
Joachim Nern, Aspasia Knowledge Systems, Dusseldorf, Germany
Carlos E. Palau, Universidad Politecnica de Valencia, Valencia, Spain
Stephan Reiff-Marganiec, University of Leicester, Leicester, UK
Prof. Paola Salomoni, Universita di Bologna, Bologna, Italy

MEDIATEC Programme Committee

Track Chair

Ignazio Infantino, Phd, ICAR-CNR, Palermo, Italy

ipc

Ali Arya, Carleton University, Ottawa, Ontario, Canada
Helena Barbas, CENTRIA-UNL, Lisbon, Portugal
Fernando Boronat Segui, Universidad Politecnica de Valencia, Gran de Gandia, Spain
Vincent Charvillat, IRIT-ENSEEIHT, Toulouse cedex, France
Jehan Francois Paris, University of Houston, Houston, USA
Florin Pop, University POLITEHNICA of Bucharest, Bucharest, Romania
Marco Roccetti, Universita' di Bologna, Bologna, Italy
Leon Rothkrantz, Delft University of Technology, Delft, The Netherlands
Leonid Smalov, Coventry University, Coventry, United Kingdom

COMTEC Programme Committee

Marwan Al-Akaidi, Arab Open University, Kuwait
Christos Bouras, University of Patras, Patras, Greece
Boguslaw Butrylo, Bialystok Technical University, Bialystok, Poland
Sophie Chabridon, Institut Télécom, Télécom SudParis, Evry Cedex, France
Jonathan Dukes, O'Reilly Institute, Trinity College, Dublin 2, Ireland
Mohammad Riaz Moghal, Ali Ahmad Shah-University College of Engineering and Technology, Mirpur, Pakistan
Oryal Tanir, Bell Canada, Montreal, Canada

APTEC Programme Committee

Juan Carlos Guerri Cebollada, Universidad Politecnica de Valencia, Valencia, Spain
Peter Kvasnica, Alexander Dubcek University of Trencin, Trencin, Slovakia
Viorel Nicolau, "Dunarea de Jos" University of Galati, Romania
Jeanne Schreurs. Hasselt University, Diepenbeek, Belgium
Ramiro Velazquez, Universidad Panamericana, Aguascalientes, Mexico

E-TEC Programme Committee

Jose Machado, University do Minho, Braga, Portugal
Paulo Novais, University do Minho, Braga, Portugal
Selwyn Piramuthu, University of Florida, Gainesville, FL, USA

Workshops

Cloud Computing

Jan Broeckhove, RUCA-UA, Antwerp, Belgium

Knowledge Management, E-Business management and E-Mobility

General Chair

Ricardo Chalmeta, Universidad Jaume I, Castellon, Spain
Jeanne Schreurs, Hasselt University, Diepenbeek, Belgium

In-Car Applications

Petr Hanacek, Brno University of Technology, Brno, Czech Republic

Virtual Reality Applications

Pieter Jorissen, Karel de Grote-Hogeschool, Hoboken, Belgium
Mircea Popovici, Ovidius University, Romania
Marcos A Rodrigues, CCRC Sheffield Hallam University, Sheffield, United Kingdom

Cybersecurity

Fernando de la Cuadra y de Colmenares, ONTINET.COM

Computer Graphics

Roberto de Beauclair Seixas, IMPA, Rio de Janeiro, Brazil
Philippe Geril, ETI, Ostend, Belgium

Student Track

Goreti Marreiros, Instituto Superior de Engenharia, Porto, Portugal

All articles published in these Proceedings have been peer reviewed

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

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

EUROSIS is a Division of ETI Bvba, The European Technology Institute, Torhoutsesteenweg 162, Box 4, B-8400 Ostend, Belgium

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

EUROSIS-ETI Publication

ISBN: 978-90-77381-88-5
EAN : 978-90-77381-88-5

Preface

It is our great pleasure to welcome you at IST Tecnico Lisboa, for our annual European Concurrent Engineering Conference (ECEC 2015), the FUture BUsiness TEChnology Conference (FUBUTEC 2015) and EUROMEDIA‘2015, which are held from April 27th-29th, 2015 in the beautiful city of Lisbon, Portugal

The accepted presentations for the conferences cover: Data Management, Logistics and Routing Optimization, Engineering Applications, VR Based Engineering, New Media Frontiers and Remote Education. The presentations provide an open forum for researchers from academia and other research communities to present, discuss and exchange related ideas, results, and experiences in these areas.

For the benefit of the participants and authors, we have endeavoured also to make the event an interesting one by inviting two high-level tutors to expand upon their research and share their knowledge for the general benefit of all those present.

The tutorials are by André Bigand of the Littoral University, Calais, France, who will give a tutorial on “Image Quality Assessment for Global Illumination Methods based on Machine Learning“, and by Paulo Goncalves of IDMEC/LAETA and Instituto Politecnico de Castelo Branco, Portugal, who will give a tutorial on “Robotic Assisted and Computer-Integrated Surgery“.

We would like to thank all the authors for submitting their research work within the conferences, as well as to the authors of accepted papers for their participation and presentation of the papers during this year‘s event.

Furthermore, we also hope you have some free time after the conferences to enjoy the beautiful city of Lisbon and its surroundings.

We look forward to meet you all again in April 2016 for next year’s edition of the conferences!

Paulo Jorge Sequeira Gonçalves
and
Philippe Geril

FUBUTEC-ECEC-EUROMEDIA‘2015
Conference Chairs

CONTENTS

CONTENTS

SCIENTIFIC PROGRAMME

DATA MANAGEMENT

BUSINESS BANKRUPTCY PREDICTION BASED ON HYBRID CBR MODEL

Bruno Trstenjak
Department of Computer Science
Polytechnic of Medimurje in Cakovec
Bana J.J. 22b, Cakovec, Croatia
Email: btrstenjak@mev.hr

Dzenana Donko
Department of Computer Science
Faculty of Electrical Engineering
Zmaja od bosne bb, Sarajevo
Bosnia and Herzegovina
Email: ddonko@etf.unsa.ba

KEYWORDS

Bankruptcy Prediction, Hybrid CBR model, Case-based reasoning, Domain driven model.

ABSTRACT

Predicting bankruptcy and early identification of the financial crisis nowadays has become a field of particular interest in which various studies have been conducted. The consequences of business bankruptcy have a negative impact on the whole society. Signs of financial distress can be detected much earlier before bankruptcy occurs. For this reason, a variety of scientific methods have been developed for timely detection of the difficulties in the business.

The main purpose of this paper is to analyze quality prediction of business bankruptcy using the novel hybrid Case-based reasoning (CBR) model with a new approach of data classification. The hybrid model in its algorithm has integrated several methods of machine learning: Information Gain, K-means and Case-based reasoning. The principle of work of the hybrid model is based on the experts knowledge base, library of past cases and determined weight values of object features. The model in the process of prediction uses the entire information as a basic value and does not use features reduction methods. jCOLIBRI is used as a framework for the development of our hybrid CBR model and algorithm. Our experiment has achieved promising results and provides a good basis for future research.

INTRODUCTION

Today in the modern world business bankruptcy has become commonplace. Business bankruptcies and their frequency always have a significant negative impact on the whole society, especially on the social and economic component. The causes of business bankruptcy have been discussed by many authors, trying to identify indicators that will signal the possibility of negative scenarios in the business. Paul J. FitzPatrick identified five stages leading to business problems: incubation, financial embarrassment, financial insolvency, total insolvency, and confirmed insolvency (Fitzpatrick P.J. 1932). All these and similar efforts are focused at defining features that describe the business object and on the basis of their values predict business bankruptcy. In order to prevent and predict such negative events different kind of research has been conducted.

The machine learning is one of the scientific fields in which the concept of prediction has been explored and various algorithms have been developed in order to archive better prediction accuracy (Baldi et al. 2000), (Pandya and Virparia 2013). Modern bankruptcy prediction models often use various statistical analysis and various methods of machine learning in the field of data classification. In the studies based on statistical analyses, the methods such as logical regression, multivariate discriminant analysis and factor analysis were used (Ohlson 1980), (West 1985). Somewhat different approach in prediction are provided by the use of machine learning methods such as neural networks (Ravisankar et al. 2010), support vector machines (Jong and Kwon 2010), Bayesian network models (Sun and Shenoy 2007) and many others. To achieve the highest prediction accuracy, models that use several machine learning techniques fused in a single algorithm were developed. Such models are called hybrid models, and their complex algorithms regularly achieve better classification results than individual methods (Ping and Yongheng 2011), (De Andrés et al. 2011).

In order to achieve the highest possible prediction accuracy we have developed newer hybrid model and algorithm based on Case-based reasoning (CBR) method of classification. Our hybrid model provides a slightly different classification approach. The model has embedded three distinct processes with associated algorithms: generating weight values of features, clustering and classification.

Information Gain (IG) is a filter selection algorithm that is used for ranking the input attributes. The large number of features that are included in the prediction model can influence on the speed of prediction. The irrelevant features can also have influence on the prediction accuracy. IG measures entropy of features and ranks them according to the measured values (Beniwal and Arora 2012).

For the purposes of classification process K-means algorithm is used (An et al. 2008). K-means is one of the simplest unsupervised learning algorithms used for solving clustering problems. Clustering is the process of dividing data into smaller groups, clusters, associated on the basis of common properties. K-means algorithm is used to optimize the clustering data, prepare for classification phase, all in order to improve the classification accuracy.

The third process in the hybrid model is classification of objects. For the purposes of classification CBR method was used. CBR method in the classification process uses the methodology of solving new problems by determining the similarity from the past. Experiences are essential for CBR method and principle of its work. All collected experiences are written in the form of cases. CBR investigates cases from the past, and on the basis of similarity, functions propose solution to the current situation (Hullermeier and Cheng 2013). The CBR method based on the four cycles proposed by Aamodt and Plaza [1994] is a commonly accepted process model.

The rest of the paper is organized as follows; Research objectives section briefly describes research objectives, Overview of hybrid CBR model section deals with the concept of a novel hybrid algorithm and the knowledge database. In the Case study section presents the achieved results of the case study and Conclusion section presents conclusions and directions for further development.

RESEARCH OBJECTIVES

The main goal of this study is to analyze quality prediction of business bankruptcy using the novel hybrid CBR model with a new approach of data classification and implement classification of objects based on the features weight values and database knowledge from the domain of problem. In the process of predicting, we attempt to preserve all the objects information without reducing the number of their features. We wanted to measure maximum prediction accuracy and explore the dependence between the accuracy of prediction and the size of the dataset. Detect weak and the good characteristics of the hybrid model in the proces of prediction.

OVERVIEW OF HYBRID CBR MODEL

During the years of research in various domains of problems we have witnessed that experience and knowledge can play an important role in making the final decision. Sometimes decisions based on experience can give better results than decisions based on scientific mathematical methods and algorithms. All this led to the idea of combining the experience and knowledge with the machine learning methods, all with the goal to achieve high accuracy in prediction of bussiness bankruptcy.

The algorithm of our hybrid model prediction is carried out in three internal processes: generating weight values of features, clustering and classification. The process of generating the weight values is performed during model initialization. A generator uses the IG method for ranking the features of objects and performs aggregation of results with the data from domain driven model. The hybrid algorithm with the aggregation of information forms a weight vector which forwards in the process of classification. Two first processes in the hybrid algorithm provide additional information and thereby strives to increase the accuracy of classification. Figure 1 shows the concept of a hybrid CBR model.

The weight generator

The main purpose of the generator is to determine the weight values of features, the features which describe the objects in the dataset. The generator is composed of two components: IG method for ranking object features and domain driven model which contains knowledge of experts from the problem domain.

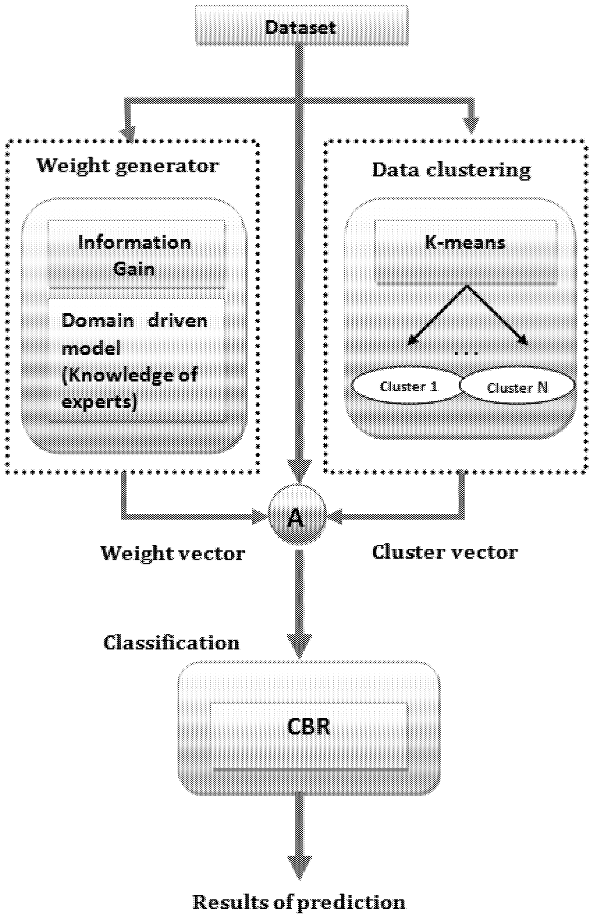


Figure 1: The structure of the hybrid CBR model

The domain driven model presents the knowledge of experts and their opinions about the importance of features that describe the business object. In the knowledge base weight value for each feature is written. The value indicates the importance of each prediction feature based on the opinion of experts; strength impact of features on the final result of prediction. Various studies in the classification field showed that the features have different influence on prediction results (Janecek and Gansterer 2008). Weight values in domain driven model were determined by experts, based on their experiences and opinions. In our study, the values were determined by surveying experts from various fields of economy who work in our institution. The values are in the range from 0-1, Table 1 shows meanings of the values.

The experiences which we have obtained during our research has shown that sometimes the method of ranking the importance of features and their weight values could be in-correct. Therefore, Domain driven model was included in the algorithm as a corrective factor in the process of defining weight vector. The hybrid algorithm performs correction of x weight values, which has computed using IG method.

Table 1: Values in Domain driven model

Value	Description
0	The feature is not important for the process of prediction and is not used. The attribute is ignored.
0< x <1	The value represents the corrective factor in the process of generating a weight vector.
1	The feature has a maximum impact on the result of prediction. The feature must come into the process of prediction.

Entropy is a measure of disorderliness of the system. Information Gain (IG) method calculates the value of the features information. Value is defined as the amount of information, provided by the feature items for the class. IG uses the following expression for the calculation:

IG(Class, Feature)=H (Class) – H(Class |Feature)

where H is entropy, which is defined by using the following equation (1).

$$Entropy(S) = -\sum_{j=1}^{m} p_j \log_2 p_j \,, \qquad (1)$$

where p is probability, for which a particular value occurs in the sample space S.

Entropy value ranges from 0 to 1. Value 0 means that all variable instances have the same value, value 1 equals the number of instances of each value. Entropy shows how the attribute values are distributed and indicates "purity" of features. High entropy indicates the uniform distribution of attributes. Opposite to that, low entropy indicates the distribution which is concentrated around the point (Lei 2012).

Data clustering

K-means is one of the simplest unsupervised learning algorithms used for solving clustering problems. Let X={xi, i=1,..n} be the set of n dimensional objects which should be classified into k clusters, C={cj, j=1,..,k}. The algorithm determines the quality of the clustering calculating square error between the mean of a cluster and the points in the cluster. The goal of algorithm is to minimize the sum of the squared error over all K clusters. The quality is determined by following the error function, as shown equation (2) (Jain 2010):

$$E = \sum_{j=1}^{k} \sum_{x \in C_j} \left| x_i - \mu_j \right|^2 \,, \qquad (2)$$

where E is sum of the squared error of all objects, μj indicates the average of cluster Cj. $|x_i-\mu_j|2$ is a chosen distance measure between data point xi and centroids value. The algorithm can use different methods to calculate the distance (Euclidean, Manhattan, Minkowski, etc.) (Kouser and Sunita 2013).

Data classification

The principle of CBR method is based on solving new problems by observing the similarity with the previous solved problems. CBR method uses a problem-solving approach analogous to the way of problem solving by man when he draws on his experiences (Klein 1999). Each CBR system contains embedded library of past cases that have been resolved in the past. This is something like collecting life experiences in the domain of the problem. Each case represents a description of the problem with its associated solution. CBR method with built-in function of similarities tries to find the most similar case from the library. The retrieved cases from the library are used to suggest a solution. If the proposed solution is not satisfactory, method tries to revised selected cases and find a new solution. The method adds a new revised case to the cases library and thereby expands the knowledge base. The whole execution cycle of algorithm can be divided into four main steps (Richter and Weber 2013):

1. *Retrieve* – retrieval is the first step in the cycle. The algorithm tries to find the best matching case(s). The case that will be selected among the retrieved cases depends on the similarity function which is used.
2. *Reuse* - If the new problem situation is exactly like the previous one then algorithm reuses the old case. If the retrieved cases do not offer acceptable solution for a new problem, the algorithm performs the adaptation of retrieved cases
3. *Revise* - This step starts when a solution is proposed to solve a new problem. The aim of revise is to assess the acceptability of proposed solutions, the newly formed case.
4. *Retain* - In the retain step there are new useful cases in the case base for future reuse. In this way the CBR system has learned a new experience and retains the knowledge gained from solving the new problem.

Which cases to be retrieved is decided based on a given similarity threshold. CBR performs measurement of similarity on the local and global level. Local similarity refers to the measurement of similarity between pairs of features. Global similarity refers to a comparison of the similarity between all the features that make up the object. Measuring similarity can be shown by the following equation (3) (Najib and Ahmad 2012):

$$Similarity(T,S) = \sum_{i=1}^{n} f(T_j, S_i) \times w_i \,, \qquad (3)$$

Where
T= target case
S= source case
n= number of features in each case

I= individual feature from 1 to n
f= similarity function for features I in cases T and S
w= importance weighting of feature I.

CASE STUDY

Dataset

Business bankruptcy is a financial failure when a company or organization is unable to pay its financial obligations (Ohlson 1980). In order to implement the process of prediction it is necessary to detect the features by which the model will classify objects in the analysis. Features of objects which are used in prediction can be of financial and non-financial nature. In this research we used a validated and prepared datasets created by Jeffrey S. Simonoff (Simonoff 2003). Business companies have described with seven features. For the testing purposes we used a dataset of 50 instances. Table 2 shows an overview of the features (Simonoff 2003):

Table 2: Overview of features in the dataset

No	Feature	Description
1	CN	Name of business entity.
2	WC/TA	Working capital as a percentage of total assets; the difference between current assets and liabilities, and is thus a measure of liquidity.
3	RE/TA	Retained earnings as a percentage of total assets; a measure of cumulative profitability over time, and is thus an indicator of profitability.
4	EBIT/TA	Earnings before interest and taxes as a percentage of total assets; a measure of the productivity of a firm's assets.
5	S/TA	Sales as a percentage of total assets; the standard capital turnover ratio, indicating the ability of a firm's assets to generate sales.
6	BVE/TL	Book value of equity divided by book value of total liabilities; ratio measures financial leverage, being the inverse of the debt to equity ratio
7	Class	B-Bankruptcy, NB-Non-Bankruptcy

Generating a weight values

In the initial phase of prediction, hybrid model begins with the generation of weight values for each feature from the dataset. Generating was performed in a few steps. IG method performs ranking of features and determines their importance, presence in the dataset. Parallel with this activity, algorithm collects information from domain driven model. In this study, the knowledge base is formed through a questionnaire conducted in the Polytechnic of Međimurje in Čakovec, Croatia. A survey was conducted over experts from various fields of economic science. Each expert assessed the importance of the features in the range from 0 to5. Value 0 denotes that the feature is irrelevant for prediction and has no influence on business bankruptcy. Value 5 in the table indicates the attributes that are most important in the process of determining bankruptcy. On the basis of their opinions we formed a study knowledge database. The algorithm performs data aggregation by merging the results obtained by the IG method and the data from Domain driven model. Table 3 shows an overview of the aggregation results.

Table 3:Overview of the aggregation results

		Values		
No	Feature	IG	DDM	Aggre.
1	CN	1.00	0.0	-
2	WC/TA	1.18	0.73	0.96
3	RE/TA	0.55	0.60	0.58
4	EBIT/TA	0.45	0.67	0.61
5	S/TA	0.00	0.87	0.43
6	BVE/TL	0.36	1.00	0.68

Table 3 shows: IG - values calculated with IG method, DDM - values obtained from a questionnaire survey, AGG - results of aggregation and final weight values of features.
Indicative situation can be observed in Table 3, the situation where the IG method marked the name of the business entity as an important feature in the dataset. In the domain driven model, that feature was marked as irrelevant for the prediction business bankruptcy. In the process of determining the final weight value of feature, the marked feature was excluded from the prediction process.
The opposite situation was created for BVE/TL feature. The results of conducted survey indicate a BVE/TL feature as something that should be necessarily used in the process of prediction. For other features, the algorithm performs a correction by reducing or increasing their values.
In order to evaluate contributions of the hybrid model, before testing the hybrid model an initial data classification has been performed. Classification of raw data was conducted using the Bayesian network method.

Table 4: Results of the initial classification

=== Bayesian network classification ===	
=== Stratified cross-validation ===	
=== Summary ===	
Correctly Classified Instances	**86 %**
Incorrectly Classified Instances	14 %
Kappa statistic	0.72
Mean absolute error	0.14
Root mean squared error	0.3198
Relative absolute error	28.42 %
Root relative squared error	63.67 %

During the initial classification of raw data from a dataset it was determined that the feature "Company name" is an irrelevant feature. We performed classification with and without "Company name" feature. The obtained measurement results are identical in both cases. When logically considering the situation, a personal name can not have an impact on the the outcome of prediction. Precisely because of such situations, a hybrid model uses knowledge and Domain driven model. To prevent incorrect classification due to an error in the definition of the input data. Table 4 shows obtained results of initial classification using Bayesian network method.

Data clustering

Parallel with the process of generating the weight values, hybrid algorithm performs clustering of objects using the K-means method. To achieve high accuracy of prediction it is necessary to define the optimal number of clusters. The initial number of clusters was determined using a simple rule of thumb shown in equation (4) (Madhulatha 2012):

$$k \approx \sqrt{\frac{n}{2}}, \qquad (4)$$

where n is the number of objects.
According to equation (4) the initial number of clusters was determined and its value was 7. For measuring distances we used Manhattan function. K-means using a distance function measures dissimilarity between objects. Results of objects clustering are shown in Table 5:

Table 5: K-means clustering results

== kMenas ==	
=== Model and evaluation on training set ===	
Sum of within cluster distances: 62.47	
Missing values globally replaced with mean/mode	
Cluster No	Clustered Instances
0	11 (22%)
1	3 (6%)
2	10 (20%)
3	5 (10%)
4	7 (14%)
5	8 (16%)
6	6 (12%)

The results present the distribution of objects, i.e. a similarity between the business entities based on their features values. The results show that the objects are relatively uniformly represented in the dataset and dataset does not have extreme cases.

Data classification

Such prepared data were forwarded to the final stage of prediction, classification with the CBR method.
For the purpose of the case study, we used jCOLIBRI framework (Recio-García at el. 2008) adapted to the domain of the problem. In the jCOLIBRI framework we have integrated our own classes adapted for measurement prediction of bankruptcy. Classification results show that the initial value of the number of clusters was well determined. The classification with such prepared data shows very good final prediction results, presented in Table 6.
The results from Table 6 show that better results are achieved using the hybrid model compared to the initial classification obtained by the Bayesian network method.The hybrid model has achieved extraordinary good classification results. The achieved improvement of prediction exceeded our initial expectations. The prediction accuracy over 99% is excellent.

Table 6: Classification results obtained with Hybrid CBR model

=== Stratified cross-validation === === Summary ===
Business Bankruptcy Accuracy: 99.6
Number of Cycles: 50
Time per Cycle: 15.26 ms
Total time: 763 ms

At the end of our study, we measured sensitivity of the hybrid model to the size of the dataset and number of cases. The original dataset consists of 50 objects. At the beginning of the testing and measurement accuracy of classification, we have reduced the number of objects to 40%. After that, the number gradually increased per step of 10 objects. Figure 2 shows obtained results of classification dependent on the number of objects.

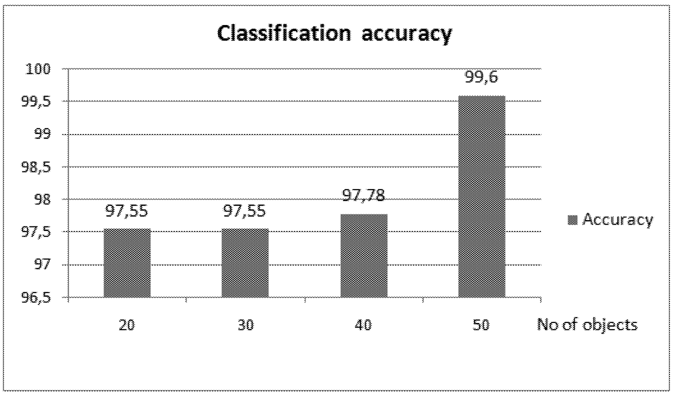


Figure 2: The accuracy of prediction depending on the number of objects

From the achieved results it is evident that hybrid model achieves high accuracy of prediction also with a smaller number of objects. During gradual increase of the number of objects, the model has increased the prediction accuracy. The results present a relatively low sensitivity of the hybrid model to the dataset size. These model characteristics ensure its application in various fields of research.

CONCLUSIONS

In this paper we present a novel concept of hybrid model for prediction. The aim of our research was the development of a new hybrid and Domain driven model based on CBR method. The model was focused on business bankruptcy prediction. According to the results achieved with a hybrid model, we can conclude that we have successfully designed a concept of a hybrid algorithm, combining the classical machine learning methods and experts knowledge from the problem domain.

The final results showed that with this model we can achieve significantly higher prediction accuracy compared to the conventional approach. The hybrid algorithm has proved stability in prediction. During testing, the model showed sensitivity in the prediction accuracy based on the defined weight values. Also, the model achieves high accuracy of prediction using a relatively small dataset. This feature of the model gives the possibility of its use in different areas which are not explored and where there is no large knowledge database. Our concept is certainly a good starting point for further development with various types of predictions. Future testing of different types of datasets will indicate the elements that will need to be modified, with the main purpose to achieve better results.

REFERENCES

An Y., J. Baek, S. Shin, M. Chang and J.Park. 2008. "Classification of Feature Set Using K-means Clustering from Histogram Refinement Method." *Networked Computing and Advanced Information Management*, 320 - 324.

Anil K. Jain. 2010. "Data clustering: 50 years beyond K-means." *Pattern recognition letters*, 651-666.

Baldi P., S.Brunak, Y.Chauvin , CA.Andersen and H.Nielsen. 2000. "Assessing the accuracy of prediction algorithms for classification: an overview." *Bioinformatics* , 412-424.

Beniwal S. and J. Arora,. 2012. "Classification and feature selection techniques in data mining." *International Journal of Engineering Research & Technology.*

De Andrés J., P. Lorca, F. Javier de Cos Juez and F. Sánchez-Lasheras. 2011. "Bankruptcy forecasting: A hybrid approach using Fuzzy c-means clustering and Multivariate Adaptive Regression Splines (MARS)." *Expert Systems with Applications*, 1866-1875.

Fitzpatrick P.J. 1932. "A comparison of ratios of successful industrial enterprises with those of failed companies." *Expert Systems with Applications*, 1866-1875.

Klein G.A. 1999. "How people make decisions." *1st MIT Press pbk, London,Cambridge.*

Hullermeier E. and W. Cheng 2013. " Preference-based CBR: General Ideas and Basic Principles." *Proceedings of the Twenty-Third international joint conference on Artificial Intelligence* , 3012-3016.

James A. Ohlson. 1980. "Financial Ratios and the Probabilistic Prediction of Bankruptcy." *Journal of Accounting Research* , 109-131.

Janecek A.G.K. and W.N.Gansterer. 2008. "On the Relationship Between Feature Selection and Classification Accuracy." *Journal of Machine Learning Research* , 90-105.

Jeffrey S. Simonoff. 2003. "*Analyzing Categorical Data." New York: Springer-Verlag.*

Yoon J.S. and S. Kwon Young. 2010. "A practical approach to bankruptcy prediction for small businesses: Substituting the unavailable financial data for credit card sales information." *Expert Systems with Applications* , 3624-3629.

Kouser K. and Sunita. 2013. "A comparative study of K Means algorithm by different distance measures." *International Journal of Innovative Research in Computer and Communication Engineering* .

Lei S. 2012. "A feature selection method based on Information Gain and Genetic algorithm." *2012 International Conference on Computer Science and Electronics Engineering*, 355-358.

MacQueen J.B. 1967. "Some Methods for classification and analysis of multivariate observations." *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability*, 281-297.

Madhulatha S.T. 2012. "An overview on clustering methods." *IOSR Journal of Engineering* , 719-725.

Najib M.S. and M.U. Ahmad. 2012. "Agarwood Classification: A Case-based Reasoning Approach Based on E-nose." *IEEE 8th International Colloquium on Signal Processing and its Application* , 120-126.

Pandya S. and P.V. Virparia. 2013. "Comparing the applications of various algorithms of classification technique of data mining in an Indian." *International Journal of Advanced Research in Computer Science and Software Engineering* , 1023-1026.

Ping Y. and L. Yongheng. 2011. "Hybrid Classifier Using Neighborhood Rough Set and SVM for Credit Scoring." *Expert Systems with Applications* , 738-753.

Ravisankar P.,V. Ravi and I. Bose. 2010. "Failure prediction of dotcom companies using neural network–genetic programming hybrids." *Elsevier Journal of Information Sciences* , 1257–1267.

Recio-García J.A., B.Díaz-Agudo and P.González-Calero. 2008. "*jCOLIBRI2 Case-Based Reasoning Framework."* Universidad Complutense De Madrid.

Richter M.M. and R. Weber. 2013. "Case-Based Reasoning: A Textbook." *Springer Science & Business Media.* 17-34.

Sun L. and P.P.Shenoy. 2007. "Using Bayesian networks for bankruptcy prediction: Some methodological issues." *European Journal of Operational Research* , 738-753.

West R.C. 1985. "A factor analytic approach to bank condition." *Journal of Banking and Finance* , 253—266.

BIOGRAPHIES

BRUNO TRSTENJAK is a lecturer at the Polytechnic of Medimurje in Cakovec, Croatia. In 2012, he completed MSc degree in the field of Computer sciences at the Faculty of Electrical Engineering in Sarajevo, BiH. Since 2008, he has worked as an academic member at the Medjimurje University of Applied Sciences, where he has been involved in lecturing several courses in fields of Computer Sciences. His area of interest covers: programming and application development, field of artificial intelligence, machine learning and data mining. He has been involved in the several EU projects in the field of education, as a project manager or a member. Recently, he is head of department at professional study program - Computer Engineering.

DZENANA DONKO is an associated professor at the Faculty of Electrical Engineering University of Sarajevo, BiH. She received her Ph.D. from the Faculty of Electrical Engineering, University of Sarajevo in 2004. Ms. Donko participated in many project and studies, some of them related to the UNDP development projects in BiH. Her current research interests include areas of object oriented analysis and data mining. She is author of many papers published at different journals and conferences. In several occasions she participated as a plenary speaker at the conferences presenting subject on data mining. She is also a member of many scientific and program committees on the different international conferences.

FIRST PRINCIPLES SIMULATION MODEL IDENTIFICATION BASED ON REAL INDUSTRIAL PROCESS DATA

Igor Slastenov
Moscow Institute of Physics and Technology
(State University)
Institutsky side str., 9
141700, Dolgoprudny
Russia
E-mail: igor.slastenov@phystech.edu

Salimjan Tasanbaev
M.O. Auezov South Kazakhstan
State University
pr. Tauke-khana, 5
486050, Shymkent
Kazakhstan
E-mail: tasanbayev@mail.ru

Victor Dozortsev
ZAO "Honeywell"
Kievskaya str., 7
121059, Moscow
Russia
E-mail: victor.dozortsev@honeywell.com

KEYWORDS

Identification, simulation model validation and verification, simulation-optimization, training system, industrial process.

ABSTRACT

A method of simulation model identification based on experimental industrial process data is suggested. Series of experiments are carried out. The results are analyzed. Using of first principles models for checking and adjustment of advanced identification algorithms is discussed.

INTRODUCTION

Last years are characterized by wide spreading of simulation tools which permit to create accurate models of complicate industrial processes (Dozortsev and Kreidlin 2010). It opens a prospect to solve important applied tasks requiring online intervention in operation of industrial object which is very undesirable or even impossible because of safety reasons, economic losses, amount of needed experiments, etc.

The identification of such models may be considered in paradigm of simulation models verification and validation (V&V) (Sargent 2011). Many of V&V methods are applicable to first principles industrial plant simulation models (face validity, traces, tuning test, etc.). At the same time the using of other techniques, such as event validation, historical data validation, sensitivity analysis (Iooss and Lemaitre 2015) are limited. The reasons are the huge number of variables (many dozens of input and many hundreds of output ones), the high dynamic complexity, and the occurrence of measured and not measured disturbances.

This paper is the first attempt to apply the real industrial process data to first principles simulation model tuning. Technically the proposed approach is close to the simulation–optimization one (April et al. 2003; Ammeri et al. 2011) with the special criterion used for the evaluation of identification quality.

Two tasks are considered in this work. The first one is simulation model tuning based on data acquired from real industrial object. Model tuning method has been suggested and experimental research has been performed. The second task is input-output model parameter identification. A brief description is given for active identification using precise simulation model. Further development of effective algorithms in both fields is a task for the future.

TASK OF SIMULATION MODEL PARAMETER TUNING

Tightening of safety and product quality regulations leads to high demand for operator training systems (OTS). At once user requirements become stronger, and OTS producers have to continuously improve their solutions. In this work an approach described which permit to simplify simulation model tuning and reduce OTS price in some cases.

There are two types of OTS simulation models: standard and custom. Standard models are to simulate some process in principal without reference to any particular implementation. Custom models are created with account of real process configuration and should reproduce quantitative characteristics of the process accurately. Therewith extremely rigid demands are made on accuracy of custom training model because an incorrect training model may make operators acquire "false skills".

Creation of custom models is an expensive process. In some cases standard model functionality may be enough for customer requirements, but it should be tuned to correspond to behavior of the modeled process. Another situation occurs if an existing simulation model stop to imitate the object behavior enough accurately because of equipment replacement or operation mode changing. A procedure is described below which allow to approach behavior of a standard or a custom simulation model to behavior of a real process. Therefore, a compromise between model modernization cost and accuracy restrictions may be achieved.

Besides OTS application a high-accuracy simulation model may replace a real object during identification and checking of relatively simple regression models. Such models are applied, for example, in advanced process control (Coward 2008). Utilizing of a precise first principal model in this case permit to reduce affecting an actual process and hence to cut costs and to raise safety.

SIMULATION MODEL PARAMETER IDENTIFICATION

Consider a simulation model which is described by the following equation system

$$\begin{cases} \mathbf{x}_{i+1} = \mathbf{f}_i(\mathbf{x}_i, \mathbf{u}_i; \boldsymbol{\beta}), \\ \mathbf{y}_i = \mathbf{g}_i(\mathbf{x}_i, \mathbf{u}_i); \end{cases} \tag{1}$$

where $\mathbf{x}_i = \left(x_i^{(1)},...,x_i^{(r)}\right)$ is the state vector, $\mathbf{y}_i = \left(y_i^{(1)},...,y_i^{(s)}\right)$ is the output vector, $\mathbf{u}_i$ is the control vector, $\boldsymbol{\beta}$ is the parameters vector.

Define a sequence $\mathbf{U} = \{\mathbf{u}_i\}_{i=0}^{n}$, which is called a control sequence. If the initial state $\mathbf{x}_0$ is determined, then $\mathbf{y}_i$ depends on the sequence $\mathbf{U}$ and the parameter vector $\boldsymbol{\beta}$ only. Define a sequence $\mathbf{Y}(\mathbf{U};\boldsymbol{\beta}) = \{\mathbf{y}_i(\mathbf{U};\boldsymbol{\beta})\}_{i=0}^{n}$, which is called a response of the model (1) to control sequence $\mathbf{U}$.

Let the real object described by the model (1) has the output vector $\mathbf{z}_i(\mathbf{U}) = \left(z_i^{(1)}(\mathbf{U}),...,z_i^{(s)}(\mathbf{U})\right)$. Define a sequence $\mathbf{Z}(\mathbf{U}) = \{\mathbf{z}_i(\mathbf{U})\}_{i=0}^{n}$, which is called a response of the object.

The task of identification of the model (1) is to find optimal value of the parameter vector $\boldsymbol{\beta}$, which provide the best fit of the model output to the object output:

$$\hat{\boldsymbol{\beta}}(\mathbf{U},\mathbf{W}) = \arg\min_{\boldsymbol{\beta}} C(\mathbf{U},\boldsymbol{\beta},\mathbf{W}), \tag{2}$$

where $C(\mathbf{U},\boldsymbol{\beta},\mathbf{W})$ is the loss function. For example, a quadratic criterion may be used:

$$C(\mathbf{U},\boldsymbol{\beta},\mathbf{W}) = E_{eq}(\mathbf{Z}(\mathbf{U}),\mathbf{Y}(\mathbf{U};\boldsymbol{\beta}),\mathbf{W}) = \sum_{i=0}^{n}\sum_{j=1}^{s} w_i^{(j)} \cdot \left(z_i^{(j)}(\mathbf{U}) - y_i^{(j)}(\mathbf{U};\boldsymbol{\beta})\right)^2 \tag{3}$$

where $w_i^{(j)}$ – elements of weight matrix $\mathbf{W} \in \mathrm{R}^{s \times n}$.

MODEL TUNING BASED ON EXPERIMENTAL DATA

A model of a debutanizer (type of a distillation column) has been used as an example of the tuned model, because operation of a distillation column is studied well, and such equipment is wide-spread in industrial chemistry.

A distillation column consists of a set of trays (Fig. 1). Vapor phase contacts with liquid one on a tray, and mass exchange occurs. Light fractions in the liquid boil up to the vapor flow, and heavy fractions in the gas phase condense to the liquid flow.

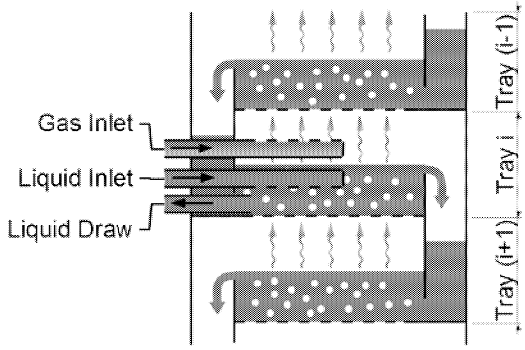


Figure 1: Distillation Column Operation

Efficiency of mass exchange between the gas and liquid phases plays a very important role for the process of distillation. The model of a debutanizer has a parameter β, which is called tray efficiency and has been chosen as the tuned parameter in this model. This parameter cannot be known a priori by a developer creating the simulation model. In addition, the efficiency may change during column operation. Therefore, tuning of the parameter β is required.

Tray efficiency β ranges from 0 to 1. If $\beta = 0$, there is no mass exchange, and compositions of the gas and liquid phases passing through a tray remain the same. If $\beta = 1$, phase equilibrium on a tray is attained instantaneously.

In order to conduct numerical experiments, a procedure for iterative search of the parameter $\boldsymbol{\beta}$ value has been written in MATLAB. It allows choosing a particular search algorithm depending on dimension of the parameter vector, type of considered process, a priori information, etc.

In this work a simple case of a scalar parameter (tray efficiency) tuning is considered. It is assumed, that there is no a priori information about dependence of model outputs on this parameter. In this case the most effective methods of parameter search are one-dimension ones. In suggested implementation the golden section technique

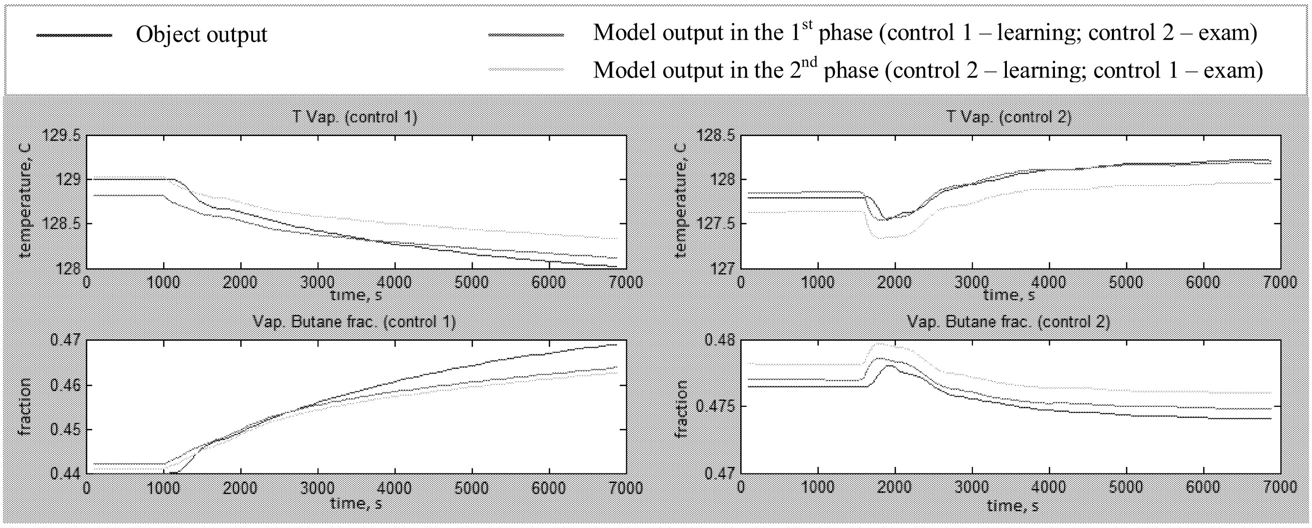


Figure 2: Learning and exam samples have different nature.
Control 1 – feed stream composition variation. Control 2 – feed stream temperature variation

(Mathews and Fink 2004) is used, because in comparison with other methods of this type it requires minimum amount of optimized function value calculations. For model adequacy checking cross-validation is applied for two control sequences.

As carrying out experiments on a real object is expensive and dangerous, an auxiliary simulation model has been used instead. This model was considered as a black box, i.e. no information about the model structure and its parameter values was used during the conducted experiments. Further by the "object" is meant the mentioned auxiliary simulation model.

EXPERIMENTAL STUDY RESULTS

Several series of experiments have been performed. Two control sequences $\mathbf{U}_1$ and $\mathbf{U}_2$ in each experiment and two object responses $\mathbf{Z}(\mathbf{U}_1)$, $\mathbf{Z}(\mathbf{U}_2)$ were utilized respectively. Each experiment consists of two phases. In the first phase the sequence $\mathbf{U}_1$ is used for determination of parameter $\boldsymbol{\beta}$ (*learning*), and the sequence $\mathbf{U}_2$ is used for checking model adequacy (*exam*). In the second phase the sequences swap over: $\mathbf{U}_2$ is used for learning, and $\mathbf{U}_1$ is used for exam. The final value of $\boldsymbol{\beta}$ is chosen according to the following expression:

$$\boldsymbol{\beta} = \hat{\boldsymbol{\beta}}_m, m = \arg\min_{j=1,2} \max_{i=1,2} C\left(\mathbf{U}_i, \hat{\boldsymbol{\beta}}_j, \mathbf{W}\right), \qquad (4)$$

where $\hat{\boldsymbol{\beta}}_j$ is the parameter value obtained by learning from the sequence $\mathbf{U}_j$.

The composition and the temperature of the feed stream have been chosen as the control variables. Output variables in all experiments were the temperature and the butane fraction of the gas stream leaving the top tray.

Each series of experiments was aimed at testing how certain conditions influence identification quality.

1. The same simulation model both as the "object" and as the tuned model.

This series was performed in order to ascertain that the chosen algorithm works correctly and converges at different initial estimates. Other series dealt with various conditions which can affect identification precision:

2. Different number of considered output variables.
3. Additive output noise.
4. Structure divergence of "object" and "model".
5. Using different types of learning and exam control samples.

The loss function (3) was used to evaluate identification quality. The elements of the weight matrix were determined in the following way:

$$w_i^{(j)} = \left[(n+1)\left(\Delta^{(j)}\right)^2\right]^{-1}, i = 0, \ldots, n, j = 1, \ldots, s \qquad (5)$$

where $\Delta^{(j)}$ is the acceptable mean square tolerance for the output variable $y^{(j)}$. The value of $\Delta^{(j)}$ is determined by the accuracy requirements for given process and depends on the scale of the corresponding output variable variance. Then the identification quality criterion may be specified as

$$C(\mathbf{U}, \boldsymbol{\beta}, \mathbf{W}) \leq 1. \qquad (6)$$

The plots of the "object" and model outputs for one of the experiments in series 5 (different types of learning and exam samples) is shown on Fig. 2. From the *face validity* point of view (Sargent 2011) the model behavior meets "object" one. The quantitative results of the experiments are shown in Table 1. The first column corresponds to the series' numbers listed above. Column Base shows the average value of the loss function (3) for the model before tuning. The last column displays the ratio of the loss function values on exam to those on learning. This parameter permits to evaluate the consistency of the parameter estimated value (if this ratio is much greater than 1, then overfitting occurred during learning, and the founded parameter value cannot be relied on).

Table 1: Average Values of the Loss Function

#	Basic	Learning	Exam	Exm / Lrn
1	1.27E+03	6.06E-03	6.08E-03	1.00
2	1.57E+03	5.91E-03	5.99E-03	1.01
3	1.71E+02	3.75E-01	3.89E-01	1.04
4	1.43E+03	1.72E-01	2.51E-01	1.46
5	1.45E+03	2.06E-01	2.89E-01	1.40

Experimental data demonstrate that although the quality of identification decreases with experimental conditions' toughening, applying of the proposed tuning procedure may significantly improve the model accuracy (Exam loss function values in comparison with Basic ones).

CHECKING IDENTIFICATION ALGORITHMS

Once an accurate first principles simulation model is built it may be used for engineering support of identification of input-output plant models and their testing. It can be especially important for active identification algorithms which require costly intervention into the object operation. During active identification a control signal is formed in such a way as to conduce to collecting information about the object along with leading to certain mode. These aims contradict each other; therefore the solution of the problem should be gained by compromise between them (Bar-Shalom and Tse 1976).

Consider following model sensitivity index to tuned parameter:

$$J = \sum_{i=1}^{n} \left(\frac{\partial \mathbf{y}_i(\mathbf{U};\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} \right)^{\mathrm{T}} \overline{\mathbf{W}}_i \left(\frac{\partial \mathbf{y}_i(\mathbf{U};\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} \right) \qquad (7)$$

where $\overline{\mathbf{W}}_i = \mathbf{w}_i(\mathbf{w}_i)^{\mathrm{T}}$ – weight matrix similar to (3).

The sensitivity index is associated with Fisher information matrix (Mehra 1974). The more sensitive some output variable to evaluated parameter the more information can be obtained by variation of this variable.

Let control be chosen for inequation $J \geq r$, where r – minimal pace of collecting information, to be satisfied. Such control leads to dual effect (Bar-Shalom and Tse 1976) and provides conditions for an active experiment. But applying active identification to real object is dangerous and leads to economical losses. So it is useful to test new identification algorithms using simulation model instead of real object. It should be noted that similar approach is used in advanced predictive control for input-output models to be created before process running (Coward 2008) or for intervention to process to be minimized (Dozortsev et al. 2000).

FIRST RESULTS AND FUTURE TASKS

Analyses of experimental results shows possibility of simulation model tuning based on experimental data in principal. Influence of factors usually occurring in practice on identification quality has been explored.

The main prospects in this field:

- testing of tuning algorithms by different model types;
- using created identification tools for model tuning based on data from real process and applying them in actual OTS projects;
- further checking new identification algorithms.

REFERENCES

Ammeri, A. et al. 2011. "A Comprehensive Literature Review of Mono-Objective Simulation Optimization Methods". In *Advances in Production Engineering and Management.* Vol. 6. No. 4. Pp. 291-302.

April , J. et al. 2003. 'Practical Introduction to Simulation Optimization". In *Proceedings of 2003 Winter Simulation Conference*. Pp. 71-78.

Bar-Shalom, Y. and E. Tse. 1976. "Concepts and Methods in Stochastic Control". In *Control and Dynamic Systems: Advances in Theory and Applications*, Vol. 12 (C. T. Leondes, ed.), Academic Press.

Coward, A. 2008. "Dynamic simulation can accelerate APC payback". In *Process Engineering*. Vol.89. No.2. Pp.27-30.

Dozortsev et al. 2000. "Model Predictive Control Using Object's Rigorous Dynamic Simulation". In *Proceedings of International Conference "System Identification and Control Problems (SICPRO'2000)"*. Pp. 44-50.

Dozortsev, V.M. and E.Yu. Kreidlin. 2010. "State-of-the-art automated process simulation systems". In *Automation and Remote Control*. Vol. 71. No 9. Pp. 1955-1963.

Iooss, B. and P. Lemaitre. 2015. "A Review on Global Sensitivity Analysis Methods". In *Uncertainty Management in Simulation-Optimization of Complex Systems: Algorithms and Applications* (C. Meloni and G. Dellino, eds.), Springer.

Mathews, J.H. and K.D. Fink. 2004. Numerical Methods: Using Matlab. NJ: Prentice Hall.

Mehra, R. 1974. "Optimal Input Signals for Parameter Estimation in Dynamic Systems – Survey and New Results". In *IEEE Transactions on Automatic Control*. Vol. AC-19. No 6. Pp.753-768.

Sargent, R.G. 2011. "Verification and Validation of Simulation Models". In *Proceedings of 2011 Winter Simulation Conference*. Pp. 183-198.

XML-BASED MODELING LANGUAGES AND DATA BINDING FOR COLLABORATIVE DESIGN IN MULTIDISCIPLINARY TEAMS

Martin Hoppen
Malte Rast
Michael Schluse
Juergen Rossmann
Institute for Man-Machine Interaction
RWTH Aachen University
Ahornstrasse 55, 52074 Aachen, Germany
E-mail: hoppen@mmi.rwth-aachen.de

Lars Pfotzer
FZI Research Center for
Information Technology
Haid-und-Neu-Strasse 10-14
76131 Karlsruhe, Germany
E-mail: pfotzer@fzi.de

Thomas Meschede
TU Berlin, Department of
Aeronautics and Astronautics
Marchstrasse 12-14
10587 Berlin, Germany
E-mail:
thomas.meschede@tu-berlin.de

KEYWORDS

Collaborative Environments, Modeling Languages, 3D Simulation, XML, XSD, Data Binding

ABSTRACT

3D simulation applications are often developed or employed in distributed and multidisciplinary teams. An important basis is a common language for data modeling and exchange to simplify or even permit a common understanding. The Extensible Markup Language (XML) provides an excellent basis technology to define a common modeling language. Such a language can either be customized to a specific problem domain like modular and reconfigurable satellites and their on-orbit-servicing (OOS) in a distributed R&D space project. Alternatively, a standardized modeling language like X3D or AutomationML can be used if it fits an application's requirements. A common modeling language shall, however, not only be used for data exchange, but also as an application schema for participating systems to avoid "friction losses" due to translations. For that purpose, XML data binding techniques can be used to integrate XML schemata and data into object-oriented runtime systems. This allows for a tighter integration than generic interfaces like the Document Object Model (DOM) or the Simple API for XML (SAX).

INTRODUCTION

When collaboratively working on 3D simulation projects in distributed teams, special care has to be taken regarding the incorporated data. A well-defined, common data modeling language is needed to allow for consistent data management and exchange. It must be flexible and particularly extensible to accommodate to the different requirements of various applications and heterogeneous teams using different tools. For that purpose, XML provides an excellent basis technology. On the other hand, for applications from the 3D context like Geographic Information Systems (GIS), Computer-Aided Design (CAD), or other 3D software, an object-oriented modeling is advisable, as the corresponding data usually consists of a huge number of hierarchically structured parts with interdependencies (Elmasri and Navathe 2010).

To integrate both requirements, XML data has to be made available within object-oriented runtime systems. This can be done generically using techniques like DOM or the SAX. More intuitive is the so-called XML data binding approach. Here, the schema of the XML data is mapped to object-oriented classes and their attributes and XML data is made available in terms of corresponding objects. However, the challenge is the lack of standard mappings due to ambiguities in between the two domains (XML and object-orientation). Thus, the intended data model can only be preserved when the modeling conventions within the XML schema are understood and reasonably translated to object-oriented structures.

In this paper, we present two use cases of XML-based modeling languages and corresponding data binding solutions from the context of 3D simulation projects. The first considers the case of a customized base language for modeling general object-oriented 3D data in XML. Here, modeling conventions and object-oriented mappings are adopted from the well-defined Geography Modeling Language (GML). As a second use case, we present XML data binding approaches for established standard modeling languages using the example of the VRML successor X3D and AutomationML, a language for modeling plant engineering data.

As an example for the customized modeling language, we present a multidisciplinary R&D space project. Today's satellites are mostly monolithic systems without the possibility for maintenance, repair or reconfiguration. The lifetime of individual components or the mission duration determines the lifetime of the whole satellite. At the end of its lifetime, the satellite becomes space debris that endangers other satellites, the international space station or even populated areas on the earth. To overcome this problem, the project iBOSS

(Weise et al. 2012) develops a concept for modular satellites that can be reconfigured, tugged to other orbits or deorbited in a controlled crash. Figure 1 shows the idea of the reconfiguration of such a modular satellite under investigation in a virtual environment.

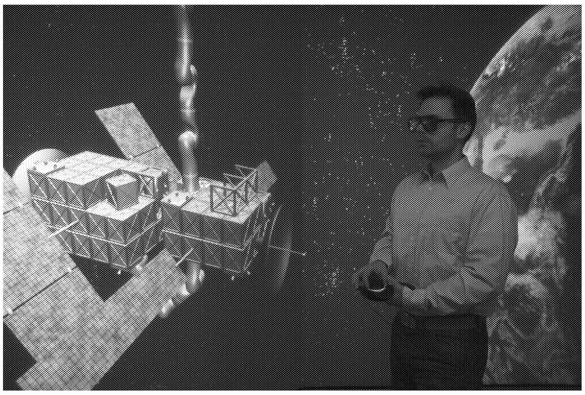


Figure 1: Satellite Model based on the iBOSS Concept (Weise et al. 2012) in Virtual Reality

Part of the project is the generation of an extensible catalog describing all satellite components and payloads. Satellite configurations can be constructed from this catalog and analyzed with respect to technical constraints or reconfiguration scenarios. With a common XML-based modeling language for this catalog and the models constructed from it, the iBOSS project partners can effectively join their expertise on spacecrafts, robotics, lightweight construction, 3D simulation, and Virtual Reality.

STATE OF THE ART

XML is a widespread format for data modeling and exchange. XML data builds a tree-like hierarchical structure with complex elements as inner nodes and simple elements as leaf nodes.

Given a schema description, an XML document is called valid if it is well-formed (regarding the XML standard) and its contents comply with the schema. Common methods for describing an XML schema are Document Type Definitions (DTDs) and XML Schema Definitions (XSDs) (Skulschus and Wiederstein 2004). For various reasons, including their non-XML syntax, DTDs have been superseded by the more powerful yet more complex XSDs. The XML Schema Definition Language for describing an XSD is based on XML itself providing the advantage of tool reuse and rich modeling features. An XSD can define global elements for documents, complex or simple types, complex or simple content models for types, local elements and attributes of complex types (including their type and multiplicity), inheritance of types (by extension or restriction), substitutability between elements, and abstractness of elements and types.

To generically access and process XML documents, two common methods exist. Using DOM, an XML document is completely parsed into a tree structure in memory where it can be accessed and manipulated using a standard API. While this approach is convenient, it may be bad for large documents. Another method is SAX, an event-based approach where the calling program is notified of each start or end tag providing stream-based processing.

Using DOM or SAX, XML data can be accessed in a very generic way. A more intuitive approach, however, is to map the XML-represented data and the corresponding schema to an object-oriented representation, i.e., to object and class structures of the utilized runtime system. Such a mapping is called XML data binding (Bourret 2001; 2011). In a process called unmarshalling or deserialization, an XML document's content is mapped to corresponding objects. To provide the necessary object-oriented structures, appropriate classes and attributes have to be derived, typically from a corresponding XSD. Different tools for XML data binding exist. However, the mapping from XSD to object-oriented structures is not as straightforward as it may seem.

This so-called impedance mismatch can be depicted with the example of XSD complex types and elements and their ambiguous mapping to the object-oriented class concept. Figure 2 gives a simplified XSD metamodel excerpt. On schema level, an `Element` has a `ComplexType`. The latter defines inheritance and a content model for the former. Both can independently be either abstract or not. On XML document level, instances of elements (`ElementInstance`) are used. Different approaches can be used to map `Element` and `ComplexType` to object-oriented class(es). A detailed survey can be found in (Bourret 2001).

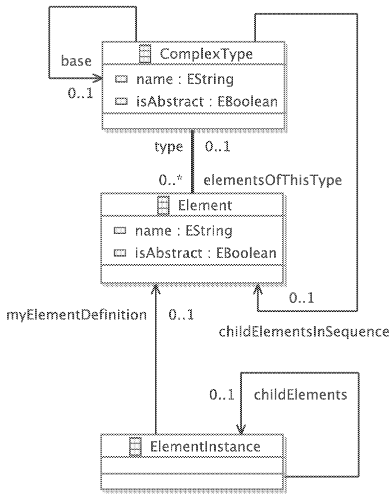


Figure 2: Simplified XSD Metamodel: Relationship between XSD Complex Type and XSD Element Definitions as well as Element Instances in Documents

As a conclusion, a completely *generic* mapping between XML/XSD and object-oriented structures is inappropriate. Instead, in practical modeling languages, a limited set of modeling concepts and coding conventions is used for which specific semantics are defined. This can be seen in different standardized modeling languages in various fields of application like X3D for 3D and Virtual Reality data (Web3D Consortium 2014, Schwellenbach 2014), AutomationML for industrial automation (AutomationML association 2014, Konieczny 2014, Drath et al. 2008), and GML for geography information systems (Open Geospatial Consortium (OGC) 2014).

X3D is the XML-based successor of the Virtual Reality Modeling Language (VRML). An X3D document (or model) contains a scene graph describing the structure of the scene. X3D's object model comprises complex objects called nodes and their properties called fields describing a node's current state. A field either holds simple property values (that can also be used for event interchange) or contains other nodes. The latter is the basis for building up scene graph hierarchies. Using the special DEF and USE statements, nodes can be given unique names (DEF) by which they can be referenced (without copying) from other nodes (USE) allowing a node to have more than one parent. X3D object types are structured by inheritance in an interface hierarchy where only leaf classes are non-abstract. X3D's XML encoding is defined by an XML Schema Definition.

The Automation Modeling Language (AutomationML) is an XML-based language for modeling and interchanging engineering data for automation technology. It combines the existing languages CAEX (Computer Aided Engineering Exchange) for object structure and language integration (Fedai and Drath 2005), COLLADA (Collaborative Design Activity) for geometry and kinematics, and PLCopen XML for logic. CAEX provides various concepts of object-oriented modeling.

GML provides basic structures for modeling and interchanging geographic data using XML. It usually serves as a foundation for customized modeling languages or application schemata, e.g., CityGML (CityGML 2014) for city data or ForestGML (Rossmann et al. 2013a) for forest data. For that purpose, it provides base types and well-defined modeling conventions with specific semantics.

USE CASE 1: CUSTOMIZED MODELING LANGUAGE

As we have seen, there is a variety of XML-based languages for specific purposes. In our iBOSS application example, we have to combine spacecraft design, robotics and 3D simulation. None of the existing XML-based languages exactly fits our needs. Therefore, we developed a custom modeling language for modular satellites and their robotic on-orbit reconfiguration. In order to reuse our concept for other application scenarios, we separated the modeling conventions in three levels in the following sections. The basis is a generic modeling language for representing object-oriented data in XML. On the second level, we add advanced techniques that may be of use in various engineering fields and, finally, we add an application specific layer.

General Modeling Conventions

As previously stated, in general, the concepts of XML and object-orientation do not match one-to-one. However, in this context the aim is not to map arbitrary XML documents with each and every possible XML modeling technique to an object-oriented representation and vice versa (also called round-trip). Instead, a common schema definition is agreed on, where modeling techniques can be limited and corresponding semantics be defined. For this purpose, the approaches from the GML specification as mentioned above can be adopted. In a generalized form, GML's UML mappings are the conceptual basis for the following XSD modeling conventions.

One convention is to always define pairs of XSD global elements and corresponding complex types (Listing 1). In an object-oriented context, such a pair is mapped to a single class. Its base class is defined by the corresponding complex type's extension base type. XML child elements are used to model the class's attributes. In contrast, XML attributes are rather used for special properties of class attributes or objects like ids, id-based references or units of physical quantities. In the example, a class *Satellite* is defined using a global element definition and a corresponding complex type. The class is derived from class *Spacecraft* using the extension mechanism on type level. Here, a *Satellite* has two properties: an *orbit* (of type string) and a list of *components* (containing *Component* objects).

Listing 1: Class Definition in XSD

```
<xs:element name="Satellite" type="SatelliteType"/>
<xs:complexType name="SatelliteType">
  <xs:complexContent>
    <xs:extension base="SpacecraftType">
      <xs:sequence>
        <xs:element name="orbit" type="xs:string"/>
        <xs:element name="components"
           type="ComponentCompositionType"
           minOccurs="0" maxOccurs="unbounded"/>
      </xs:sequence>
    </xs:extension>
  </xs:complexContent>
</xs:complexType>
```

In a mechanism for typed reference properties, one can differentiate between references with or without transitive deletion characteristics. The former corresponds to UML composite aggregations while the latter to UML shared aggregations or "normal" UML associations. Other than UML, relations are typically modeled without independent association types but rather using references from one object to another. Also adopted from GML, two modeling techniques for such references are used.

As a component is usually only referenced by one composite at a time, the proposed composite reference mechanism only allows inline definition of target elements (Listing 2).

Listing 2: Composite Reference Type in XSD

```
<xs:complexType name="ComponentCompositionType">
  <xs:sequence minOccurs="0" maxOccurs="unbounded">
    <xs:element ref="Component"/>
  </xs:sequence>
</xs:complexType>
```

In contrast, for shared aggregations or simple cross references, targets may be referenced from several properties. Thus, an id-based mechanism is necessary. Adopted from GML, a generic reference type is defined that uses the XML Linking Language (XLink) (The World Wide Web Consortium (W3C) 2014) to refer to the target of the reference (Listing 3).

Listing 3: Class Definition with Non-Composite Reference in XSD

```
<xs:element name="Component" type="ComponentType"/>
<xs:complexType name="ComponentType">
  <xs:sequence>
    <xs:element name="neighbor"
        type="ReferenceType" minOccurs="0">
      <xs:annotation>
        <xs:appinfo>
          <targetElement>Component</targetElement>
        </xs:appinfo>
      </xs:annotation>
    </xs:element>
  </xs:sequence>
</xs:complexType>
```

An exemplary instance document conforming to the presented XSD definitions is given in Listing 4. The *id* XML attribute is defined in a super base class to consistently identify objects. The *mass* property shall be derived from a base class using an additional *unit* XML attribute (see next section).

Listing 4: Example for Instance of Class Satellite

```
<Satellite id="satellite1">
  <mass unit="kg">1000</mass>
  <orbit>LEO</orbit>
  <components>
    <Component id="component1">
      <neighbor xlink:href="#component2"/>
    </Component>
    <Component id="component2">
      <neighbor xlink:href="#component1"/>
    </Component>
    <Component id="component3"/>
  <components>
</Satellite>
```

Advanced Modeling Techniques

Besides the aforementioned basic modeling rules, some more advanced techniques were also integrated into the approach.

A variation of the previously introduced document-internal references are document-external references. In the context of 3D simulation models, such references can be used as an include mechanism for sub models from a catalog XML file. References are usually defined similar to compositions because sub models are interpreted as being instantiated as a part of a composite.

Another more advanced modeling technique similarly used in GML is to define physical quantities with units. Within the base schema, special derived simple types are defined, which combine a floating point value with a unit enumeration. An example is given in Listing 4 (*mass*).

Application-specific Modeling

To support integration, a common data modeling language was defined using an XSD following the aforementioned conventions. This schema is used to create valid XML documents for data exchange between project partners. To modularize construction, the iBOSS application schema defines an abstract satellite building block. It has common properties like position, orientation, and its interfaces. An interface is another type that can be used to logically connect blocks. Therefore, an interface element has a non-composite reference to its neighbor interface it is connected to. Using the building block base type, specialized blocks like a thruster block or a payload block are derived. A catalog XML file is used to offer a predefined choice of building block instances and components.

Currently, we use a component catalog that is referenced from the building block catalog, which again is referenced from actual satellite models (see figure 3).

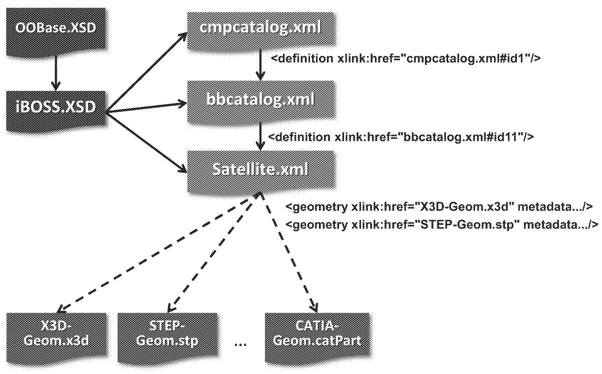


Figure 3: Structure of Schema Definitions and Component / Building Block Catalogs

All catalog and model files comply with the application-specific XSD, based on the generic XSD, and can be nested in unrestricted depth. References to objects in submodels can be encoded as idpaths, e.g. *xlink:href="#id1#id2#id3"*, where *#id1* references a submodel in the current file, *#id2* references a subsubmodel in the submodel and *#id3* references the target object in the subsubmodel. Other file types can be included as well, e.g., for geometric models (X3D, STEP, etc.). It depends on the software tool and its import capabilities whether these includes are resolved.

3D Simulation and Virtual Reality System

In the context of the R&D project iBOSS, data is modeled according to the modeling conventions presented in the last sections. To demonstrate the usability, we present three tools for different purposes that access the data with XML data binding approaches.

In 3D Simulation, data is typically structured hierarchically, i.e., in a scene graph. To represent arbitrary spatial and logical structures, the Versatile Active Simulation Database (VSD) offers graph-based object-oriented data structures. The VSD is part of the Virtual Environments and Robotics Simulation System (VEROSIM) (Rossmann et al. 2013b) we use in our projects (Figure 1).

VSD provides a meta system to access its schema (i.e., a reflection API). This facility is needed for dynamic XML data binding as described in the next section. In Figure 4, its (simplified) structure is shown. Each object class in VSD is represented by a *MetaInstance* containing *MetaMethods* and *MetaProperties* for describing its methods and attributes. Simple data types are represented by *MetaTypeVals*.

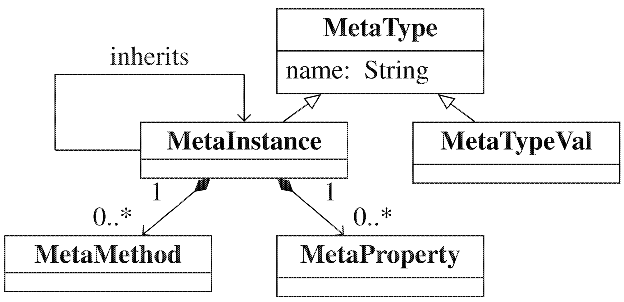


Figure 4: Simplified Structure of VSD's Meta System.

All classes of the database (each described by a *MetaInstance*) are derived from a single class *Instance*. This base class provides mechanisms for the communication between instances and access to the meta system, i.e., the complete class hierarchy with properties and methods. For a concrete simulation model, these classes are instantiated.

XML Data Binding

To integrate XML schema and data into VEROSIM, an XML data binding approach was realized. The XML schema file is used to adopt the schema in VSD. The pairs of XML global elements and complex type definitions are either mapped to existing VSD *MetaInstances* or new ones are created (and mapped as well) where necessary. Accordingly, XML child element definitions are mapped to VSD *MetaProperties*. These can either be value or reference *MetaProperties*. For value *MetaProperties*, the XML type is translated to the corresponding VSD *MetaTypeVal*. E.g., an *xs:integer* type is mapped to an *int* within VSD. Likewise, typed references are identified in their XSD representation and mapped to *MetaProperties* with the appropriate *MetaInstance* target type and with or without transitive deletion characteristics. The special types for physical quantities with units are mapped to value *MetaProperties* with units.

After adopting the schema from an XSD file, corresponding XML files can be unmarshalled to VSD instances. For each XML element, a loading mechanism retrieves the corresponding *MetaInstance* from the schema mapping. Accordingly, for child elements, corresponding *MetaProperties* are retrieved. The *MetaInstance* is instantiated to a VSD instance and value *MetaProperties* are loaded. Composite references are processed recursively to build up the object hierarchy. In a downstream secondary pass, non-composite references are resolved. Document-external references are resolved by opening the corresponding XML file, loading the referenced portion and attaching it to the object hierarchy.

Saving the contents of the VSD to an XML document (marshalling) also works recursively, beginning at the root instance. Each instance's *MetaInstance* is retrieved and using the schema mapping an appropriate XML element is created. For each *MetaProperty*, XML child elements are added, which either contain simple values, further child elements representing target instances for composite references, xlink attribute values for internal non-composite references, or xlinks to external sub model elements.

Computer Aided Satellite Design (CASD)

The automatic construction of mission specific satellite configurations from modular building blocks is one goal of the iBOSS project. Concerning the high complexity and huge variety of modular and reconfigurable satellites, we developed the Computer Aided Satellite Design (CASD) tool chain (Göller et al. 2012). The CASD tool enables operators and planners of new satellites to automatically generate satellite configurations from a standardized catalog of building blocks. The CASD synthesis process itself is internally divided into several steps to reduce the complexity of the overall optimization problem. Data has to be transferred and exchanged between the individual stages shown in Figure 5.

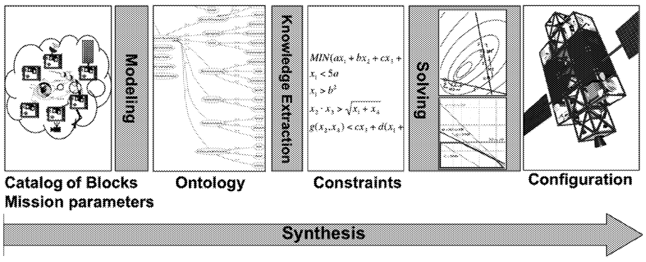


Figure 5: CASD Synthesis Process (Göller et al. 2012)

In the first step, mission specific parameters and resource specifications are used to select the required building blocks from the previously defined catalog. The outcome of the module selection is an XML file containing a list of selected building blocks referencing the catalog. The reasoning step uses this list of selected building blocks to infer rules and constraints for their placement. In the last step, the module arrangement uses these constraints to generate and optimize the satellite configuration consisting of the building blocks. The result of this optimization is a set of best fitting satellite configurations, which are written to separate XML files. With this representation, it is possible to easily exchange satellite configurations between all users of the overall tool chain. Furthermore, these satellite configurations are used to generate reconfiguration plans to describe the transformation from one configuration to another. These plans are stored in an XML file as well (Rühl et al. 2014).

For intuitive and simple handling of the XML data binding, we use the open-source and cross-platform tool CodeSynthesis XSD (Code Synthesis Tools CC 2014). It is a data binding compiler that generates C++ classes from XML Schema Definitions. The process of generating and compiling the C++ classes is integrated into our cmake based build framework. The generated data structure can be used to load, verify and save XML files complying to the XSD.

Automated Model Generation & Parameterization

The development of a modular satellite is in fact a variant of an optimization task: A module catalog has to be optimized until it fulfills the requirements of certain reference missions. This task is very complex and can only be done by iteratively building and testing satellites with different modules. As this process would take too much time when using real hardware, a sophisticated simulation is required, which can automate the task.

The simulation needs to be run with a large list of parameters for the configuration of the whole compound of the satellite as well as the building blocks themselves. The XML database contains these parameters ordered in three catalogs of modules, components and satellites.

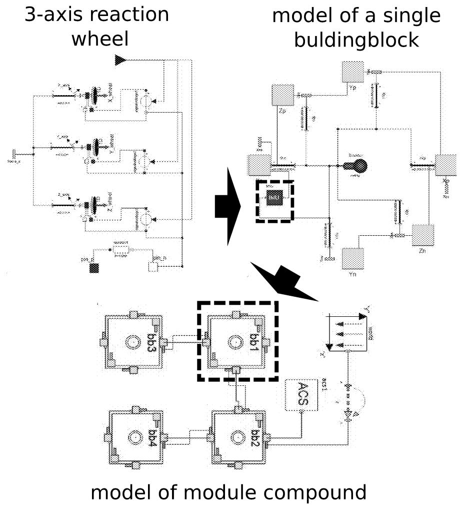


Figure 6: Model Hierarchy within Modelica

The simulation model is based on the model description language *Modelica*. This language allows for a systematic and hierarchical integration of simulation models making the simulation architecture very modular. Figure 6 shows the model hierarchy used in the project. A library of components for subsystems of spacecraft like EPS (Electronic Power Subsystem), AOCS (Attitude and Orbit Control Subsystem), TCS (Thermal Control Subsystem) and DHS (Data Handling Subsystem) has been defined for this purpose in terms of Modelica packages.

Using the XML files, simulation models can now be generated automatically: First, a *build tool* parametrizes the models from the library with the information from the XML files and then stitches them together to form individual modules (for example, a model for a reaction-wheel module). In a second step, a *connect tool* analyzes the structure of a module compound and connects the individual modules accordingly to form a satellite. In a last step, the module is automatically compiled using a Modelica compiler.

(De-)serialization of the database between Python and XML is based on standard libraries (xml.etree.ElementTree) provided with Python 3.4. We do not use XSD-based generation tools (generateDS, PyXB, pyxsd), but wrote our own object-relational mapper (ORM) making extensive use of Python's dynamically typed nature. XSD is mainly used for validity checks with our project partners. With this approach, we have more flexibility embedding the database into our own development chain.

USE CASE 2: STANDARD MODELING LANGUAGES

As a second use case, the modeling languages X3D and AutomationML were integrated into the simulation system VEROSIM. To apply an XML data binding approach, the basic modeling conventions of the respective XSDs have to be identified. Subsequently, they have to be mapped to VSD's modeling primitives, i.e., *MetaInstance*, *MetaProperty*, and *MetaTypeVal*. Given this general mapping, a corresponding data binding approach as presented above can be utilized to adopt the schema in the VSD and to load data from (and save data to) schema-compliant XML files.

After unmarshalling an instance document to the VSD, its contents are available to the components of the simulation system. However, the semantics or functionality of these objects are still unknown. For example, X3D or COLLADA geometry representations cannot immediately be interpreted by the simulation system's render engine. Following (Hoppen et al. 2012), a functional data synchronization component can be developed to translate such representations online to the simulation system's native format.

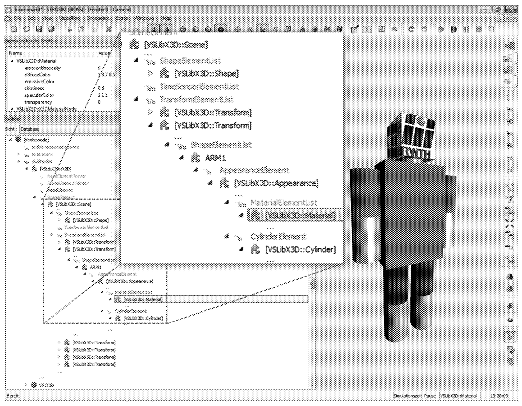


Figure 7: Simple Exemplary X3D Document loaded in VEROSIM.

X3D

Following the analysis of the X3D format in (Schwellenbach 2014), abstract X3D node types are represented by global abstract XSD complex types. In turn, concrete X3D node types are represented by global XSD element definitions with an embedded, anonymous complex type to realize the content model and inheritance. Inheritance is modeled using XSD complex type's extension mechanism. X3D fields representing simple (i.e., non-node) values are represented by XML attributes with special XSD simple types derived by restriction from build-in basic types (e.g., SFBool derived from xs:boolean). X3D fields that contain other nodes are modeled as XSD child elements referencing the corresponding node's global element definition. XSD groups and attribute groups are used to model recurring content models and reference them from respective complex type content models (mostly sequences). Finally, the aforementioned DEF/USE statements are modeled using XSD ID/IDREF attributes that are members of every base complex type and, thus, of every type and global element. This allows each and every complex child element to become a reference to another existing element of the same type that was declared somewhere else using a DEF statement.

These modeling concepts are then appropriately mapped to the VSD meta system. Using these mappings, the X3D schema and corresponding data can be loaded into the VSD. Based thereupon, a functional data synchronization approach was developed for a downstream translation of geometry representations, transformations, materials, and texture definitions. An example is given in Figure 7.

AutomationML

As for X3D, the first step for an XML data binding approach for AutomationML is an analysis of the modeling conventions within the corresponding XML Schema Definitions. In (Konieczny 2014), this is conducted with a focus on the CAEX root schema.

In general, CAEX uses a single global element definition `CAEXFile` with an anonymous complex type as an entry point to every document. Alongside, some basic global XSD complex types are defined. Within `CAEXFile`'s complex type, local elements with anonymous or named complex types are interleaved in a "russian-doll" design. Besides, local elements as well as attributes with simple types are used to model non-complex properties. There are no abstract element or type definitions in the CAEX XSD.

Again, these modeling concepts are mapped to appropriate VSD structures. Using these mappings, the CAEX schema can be adopted within the VSD and corresponding instance documents can be unmarshalled.

CONCLUSION AND OUTLOOK

We present an approach how XML-based modeling can be used to provide a common language for distributed teams working on or with 3D simulation applications. Depending on the application's requirements, a customized or a standardized modeling language can be more advantageous. Furthermore, we show how XML data binding techniques can facilitate the integration of such a language into the different tools in heterogeneous teams. In particular, we illustrate how an XML Schema Definition for a customized language can be build up or how a standardized schema has to be analyzed to allow for a reasonable mapping to object-oriented structures. Altogether, this provides an excellent collaboration basis for distributed interdisciplinary research teams to cope with similar problems of data modeling and exchange.

Our aim is to develop standards for the integrated development of complex systems with 3D simulation techniques. We intend to support the whole life cycle of complex systems with electronic media, starting from system specification to design up to operation and maintenance. As the system specification is the first step that forms the basis, we currently work on a consistent evolution of a system design process starting from the specification. This evolution naturally contains iterations and parallel development, such that a consistent data management is a challenge. To this end, we want to combine the presented data structures for simulation with SysML for specification.

Acknowledgements

Parts of this work were developed in the context of the research projects iBOSS-2 and INVIRTES. Supported by the German Aerospace Center (DLR) with funds of the German Federal Ministry of Economics and Technology (BMWi), support codes 50 RA 1203 (iBOSS-2) and 50 RA 1306 (INVIRTES).

REFERENCES

AutomationML association, 2014. *AutomationML specification*. URL `https://www.automationml.org`. (2015-01-22).

Bourret R., 2001. *Mapping W3C Schemas to Object Schemas to Relational Schemas*. URL `http://www.rpbourret.com/xml/SchemaMap.htm`. (2015-01-22).

Bourret R., 2011. *XML Data Binding Resources*. URL `http://www.rpbourret.com/xml/XMLDataBinding.htm`. (2015-01-22).

CityGML, 2014. *CityGML specification*. URL `http://www.citygml.org`. (2015-01-22).

Code Synthesis Tools CC, 2014. *CodeSynthesis XSD*. URL `http://www.codesynthesis.com/products/xsd/`. (2015-01-22).

Drath R.; Luder A.; Peschke J.; and Hundt L., 2008. *AutomationML—the glue for seamless automation engineering*. In *Int. Conf. on Emerging Technologies and Factory Automation (ETFA)*. IEEE, 616–623.

Elmasri R. and Navathe S., 2010. *Database Systems: Models, Languages, Design, And Application Programming*. Prentice Hall International. ISBN 978-0-13-214498-8.

Fedai M. and Drath R., 2005. *CAEX—a neutral data exchange format for engineering data*. *ATP International Automation Technology*, 1, 3.

Göller M.; Oberländer J.; Uhl K.; Roennau A.; and Dillmann R., 2012. *Modular Robots for On-Orbit Satellite Servicing*. In *Proceedings of the 2012 IEEE International Conference on Robotics and Biomimetics*. Guangzhou, China, 2018–2023.

Hoppen M.; Schluse M.; Rossmann J.; and Weitzig B., 2012. *Database-Driven Distributed 3D Simulation*. In *Proceedings of the 2012 Winter Simulation Conference*. 1–12.

Konieczny M., 2014. *Integration of AutomationML into a 3D simulation system using a flexible XML-Schema/XML-based interface*. Unpublished seminar paper (in German).

Open Geospatial Consortium (OGC), 2014. *Geography Markup Language (GML)*. URL `http://www.opengeospatial.org/standards/gml`. (2015-01-22).

Rossmann J.; Hoppen M.; and Bücken A., 2013a. *Semantic World Modelling and Data Management in a 4D Forest Simulation and Information System*. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-2/W2, 65–72. doi:10.5194/isprsarchives-XL-2-W2-65-2013.

Rossmann J.; Schluse M.; Schlette C.; and Waspe R., 2013b. *A New Approach to 3D Simulation Technology as Enabling Technology for eRobotics*. In *1st Int. Simulation Tools Conf. & EXPO (SIMEX)*. 39–46. ISBN 978-90-77381-78-6.

Rühl S.W.; Oberländer J.; Roennau A.; and Dillmann R., 2014. *Using a Robot Workspace Description for Planning On-Orbit Servicing Tasks on Modular Satellites*. In *International Symposium on Artificial Intelligence, Robotics and Automation in Space*. To appear.

Schwellenbach P., 2014. *X3D as an Interchange Format for Model Data in 3D Simulation*. Unpublished bachelor thesis (in German).

Skulschus M. and Wiederstein M., 2004. *XML Schema: Grundlagen, Praxis, Referenz*. Galileo Press, Bonn. ISBN 978-3-89842-472-1.

The World Wide Web Consortium (W3C), 2014. *XML Linking Language (XLink)*. URL `http://www.w3.org/TR/xlink/`. (2015-01-22).

Web3D Consortium, 2014. *Extensible 3D (X3D)*. URL `http://www.web3d.org/x3d/specifications/`. (2015-01-22).

Weise J.; Brieß K.; Adomeit A.; Reimerdes H.G.; Göller M.; and Dillmann R., 2012. *An Intelligent Building Blocks Concept for On-Orbit-Satellite Servcing*. In *Proc. of the Int. Symp. on Artificial Intelligence, Robotics and Automation in Space (iSAIRAS)*.

LOGISTICS AND ROUTING OPTIMIZATION

A SIMULATION STUDY OF THE MAKE-TO-STOCK AND MAKE-TO-ORDER PRODUCTION

Nuno O. Fernandes
Polytechnic Institute of Castelo Branco
Av. do Empresário 6000-767, Castelo Branco, Portugal
nogf@ipcb.pt

Sílvio Carmo-Silva
Dept. of Production and Systems, University of Minho
Campus de Gualtar 4710-057, Braga, Portugal
scarmo@dps.uminho.pt

KEYWORDS

MTS-MTO production; Workload control, Simulation.

ABSTRACT

In order to keep competitive, it is common for manufacturing companies in several industrial sectors to offer the possibility for product customization. However, the delivery times may become quite long, particularly if the utilization of resources is high. Therefore, in order to maintain delivery times short, products may be manufactured in two stages. In the first, standard components are made-to-stock. In the second, these components are customized into different end products. The paper investigates the impact of load-based order release to the second stage, when combined with a simple replenishment control policy at the first stage. We also investigate the impact of unbalanced workloads between the two stages. Results of a simulation study show that extra capacity at the first stage have a greater positive effect on the percentage of tardy orders, while extra capacity at the second stage, only slightly influence this.

INTRODUCTION

Worldwide competition is compelling manufacturing companies to shorten delivery times (Morikawa et al. 2014), while offering high product variety. The hybrid make-to-stock/make-to-order (MTS-MTO) manufacturing allows companies to exploit the benefits of delayed differentiation, reducing delivery times and inventory costs in comparison to the pure MTO and MTS strategies, respectively.

In this paper we consider the production carried out in two stages. In the first, standard components are manufactured and stocked as semi-finished products inventory. In the second, semi-finished products are assigned to orders for customization according to specific customer requirements. The semi-finished products inventory acts as a decoupling point between the two stages.

The hybrid MTS-MTO production is challenging. High inventory of semi-finished products means high holding costs, whereas low inventory may increase the waiting time of orders for semi-finished products. In addition to inventory decisions, order release and dispatching decisions at both stages determines the company capability to quote short and reliable delivery times, a strategy to keep up with market competition. In spite of this, most research on inventory positioning in supply chains ignores the intricacies of scheduling, typically assuming that orders are processed in the sequence in which they arrive to the production system (Kaminskya and Kayab 2008).

In this paper, we apply load-based order release combined with a simple CONWIP replenishment policy in order to satisfy customer orders within competitive delivery times and avoiding the stock-outs of semi-finished products. Load-based order release ensures that orders are not released too early or too late to the second stage, while maintaining the workload at low and stable levels. CONWIP (Spearman et al. 1990) ensures that the semi-finished products' buffer is filled to the required level, without a rigid order release plan to the first stage.

Most of the load-based order release models found in the research literature refer to job shops and make-to-order (MTO) manufacturing, see e.g. Tatsiopulos et al. (2008). Research effort in the area of workload control (WLC) for MTS-MTO flow shops is rather limited. Recently Fernandes et al. (2014) showed that WLC can be successfully applied to hybrid MTS-MTO general flow shop systems. The paper extends this work to pure flow shop systems. The paper investigates the impact of load-based order release when workloads between the two stages of production are balanced and when they are not. The paper employs discrete event simulation to model and analyse the behaviour of this manufacturing system.

SIMULATION STUDY

Production System Configuration

We consider a hypothetical two stage production system with an intermediate buffer of semi-finished products. Stage one is a manufacturing stage of standard components and stage two is a customization stage for the components made, based on the customer requirements and specifications. Both stages have three machines organized as pure flow shops.

As customer orders arrive to the production system their due dates and operation times are identified. It is assumed that all orders are accepted and sufficient raw material inventory is always available in the beginning of the first stage. Orders inter-arrival times follow an exponential distribution and, as in previous studies (e.g. Oosterman et al. 2000), and due dates are market driven and set by assigning a uniformly distributed allowance at the time of order arrival. In this study, the allowance varies between 30 and 45 time units. This results in approximately 7.5% of orders being tardy with evenly balanced workloads between production stages, unrestrictive release and 90% fill rate. This has been verified through preliminary simulation tests.

Operation processing times follow a truncated 2-Erlang distribution, with a mean of one time unit, and a maximum of four times the mean value, also adopted by Thürer et al. (2014). The arrival rate combined with the routings and processing times ensures that the average machine utilisation is 90%. Moreover, we make some common assumptions: machines capacity remains constant over time and no breakdowns are considered; set-up times are assumed to be sequence-independent and included in the operation processing times; distances and transportation times between work centres are assumed to be negligible.

Order Release and Dispatching

In a MTS-MTO system, customer orders are satisfied from the semi-finished products buffer, here also referred as standard components. However, when a customer order arrives and finds this buffer empty, the order is backordered. Backorders are filled whenever the standard components become available at the intermediate buffer. Arriving orders consume standard components from the intermediate buffer and flow to a pre-shop pool, waiting its release to the second stage of production. They are considered for release according to their urgency, i.e., due date, and are released only if the resulting workload at machines of stage two does not exceed predefined workload norms or limits. The release procedure is as follows: (1) At fixed time intervals, of four time units in our study, workloads on the machines of the second stage are computed and orders are release until no further releases would allow the workload to remain below the norms in every machine; (2) between periodic releases an order 'pulled' release is carried whenever the workload of the first machine in this stage falls to zero. Machines' workloads are calculated by the corrected aggregate load approach (Oosterman et al. 2000).

When an MTO order flows to the second stage for customization an MTS order is generated to the first stage for the replenishment of the semi-finished product buffer. A CONWIP control policy is applied at the first stage of processing. Thus, each time a semi-finished product is consumed by an order, the release of a new job (i.e., MTS order) to the first stage is authorized. CONWIP uses several cards of single type to control the total amount of work-in-process (WIP) allowed in the first processing stage. Cards are attached to the job at the beginning of the first stage and detached from the job (i.e., semi-finished product) when it is consumed by an MTO order.

Shop floor dispatching at the first stage is based on the *first-come-first-served* rule and at the second stage is based on the s*hortest operation due date* rule.

Experimental Design

The experimental factors and simulated levels of the study are summarised in Table 1. Forty-two (7 workload norms; 2 levels of the fill rate; and 3 scenarios for workload balance between stages) simulation cases are tested, and each test case runs 100 replicates. The time horizon for a simulation case is 33000 time units and only data of the last 30000 time units are collected, i.e., a warm-up period of 3000 time units is considered.

Table 1: Experimental factors and levels

Experimental Factor	Levels		
Workload balance between stages	scenario I	scenario II	scenario III
Order fill rate	90%		95%
Workload norms (N)	7 levels of restriction		

Workload balance between production stages was tested under 3 scenarios, by varying the level of protective capacity in each stage, namely: (I) 0% protective capacity; (II) 15% protective capacity at stage one (i.e., 5% per machine); and (III) 15% protective capacity at stage two (i.e., 5% per machine). Protective capacity is defined as a given amount of extra capacity used to protect against statistical fluctuations (Cox and Blackstone 2002).

The order fill rate, defined as the percentage of orders that is filled from the semi-finished products buffer, was tested at two levels, namely 90% and 99%. Note that the fill rate is expected to approach 100% as the semi-finished products inventory increases and tends to infinity. Thus, CONWIP aims at determining the minimum inventory at the intermediate buffer to achieve the desired fill rate, i.e., 90% and 99%. This was determined through pre-test simulation runs.

The method applied to release orders to the second stage of processing was tested at 7 workload norm levels of restriction by tightening the norm stepwise down from infinity.

RESULTS AND DISCUSSION

This section discusses the results of the simulation study described in the previous section. Tables 2 to 4 summarises the performance results in terms of the shop throughput time (T_t), the total throughput time (TT_t), the standard deviation of lateness (SD_l) and the percentage of tardy orders (P_{tardy}) across scenarios I to III. T_t refers to the time that elapses between order release and order completion. This is computed from the moment the order is released to stage two until it is completed. TT_t is the sum of the T_t, the waiting time before the release at the pre-shop pool and the waiting time for semi-finished products availability.

From Table 2 it can be observed that increasing the fill rate, from 90% to 99%, leads to a lower P_{tardy}. A higher fill rate means a higher probability of orders being filled from the intermediate buffer of semi-finished products, and therefore, lower throughput times and lower percentages of tardy orders are likely to be obtained. However, this is obtained at the cost of having a higher WIP (semi-finished products or raw materials that have been released to stage one, but have not yet completed processing at this stage) on stage one. Once it is assumed that raw material inventory is allays available, the number of cards exactly equals WIP. Thus, increasing the fill rate from 90% to 99% requires increasing WIP about 68%, i.e., from 31 to 52 units.

From Table 2 it can also be concluded that if workload norms are carefully set, this allows to reduce both, TT_t and P_{tardy}, independently of the level of the fill rate.

Table 2: Performance Results for Scenario I

Fill Rate	C[1]	Release Method	T_t	TT_t	P_{tardy}
99%	52	N=infinity	17.7	19.7	7.2
		N=14.9	16.5	19.4	6.4
		N=12.4	15.9	19.3	6.2
		N=10.4	15.1	19.1	5.9
		N=8.6	14.2	19.0	5.8
		N=7.2	13.2	18.9	5.9
		N=6	12.2	18.9	6.1
90%	31	N=infinity	17.1	19.8	7.5
		N=14.9	16.2	19.5	6.9
		N=12.4	15.6	19.4	6.7
		N=10.4	14.9	19.3	6.4
		N=8.6	14.0	19.1	6.2
		N=7.2	13.0	19.0	6.3
		N=6	12.1	19.1	6.5

Table 3: Performance Results for Scenario II

Fill Rate	C[1]	Release Method	T_t	TT_t	P_{tardy}
99%	37	N=infinity	17.7	19.6	7.1
		N=14.9	16.5	19.3	6.4
		N=12.4	15.9	19.2	6.1
		N=10.4	15.1	19.1	5.8
		N=8.6	14.1	19.0	5.7
		N=7.2	13.1	18.9	5.8
		N=6	12.3	18.9	6.1
90%	22	N=infinity	17.4	19.7	7.3
		N=14.9	16.3	19.5	6.7
		N=12.4	15.8	19.4	6.4
		N=10.4	15.0	19.2	6.1
		N=8.6	14.1	19.1	6.0
		N=7.2	13.1	19.0	6.0
		N=6	12.2	19.1	6.3

Table 4: Performance Results for Scenario III

Fill Rate	C[1]	Release Method	T_t	TT_t	P_{tardy}
99%	52	N=infinity	12.3	14.2	1.21
		N=14.9	12.0	14.1	1.07
		N=12.4	11.8	14.1	1.00
		N=10.4	11.4	14.0	.93
		N=8.6	11.0	13.9	.91
		N=7.2	10.3	13.8	.88
		N=6	9.8	13.7	.88
90%	31	N=infinity	11.9	14.5	1.52
		N=14.9	11.7	14.4	1.41
		N=12.4	11.5	14.4	1.36
		N=10.4	11.2	14.3	1.28
		N=8.6	10.7	14.2	1.23
		N=7.2	10.2	14.1	1.20
		N=6	9.5	14.0	1.18

[1]C = CONWIP cards for achieving the specified fill rate.

Analysing the influence of workloads between production stages, Tables 3 and 4, it can be concluded that a greater positive effect on P_{tardy} results when there is protective capacity at the MTO stage, than when there is protective capacity at the MTS stage. In fact, increasing the level of protective capacity at this stage directly contributes to reduce Tt, and thus, TTt and Ptardy, while increasing capacity at the MTS stage does not influence these measures of performance. However, a higher protective capacity at the MTS stage allows for reducing the level of WIP on stage one. In particularly, increasing protective capacity from 0% to 15% reduces WIP about 29%, from 52 to 37, for a fill rate of 99% and from 31 to 22, for a fill rate of 90%.

CONCLUDING REMARKS

This study investigates the impact of load-based order release in a hybrid MTS-MTO two stages flow shop system. The study analyses the impact of load-based order release when workloads between stages is unbalanced. Results led us to conclude that: (1) Load-based order release contributes to reduce the total throughput time and the percentage of tardy orders in this production system; (2) The effect of protective capacity in the MTS and MTO stages of the production system is not symmetric. A higher protective capacity in the MTO stage has a greater positive effect on the percentage of tardy orders, than a higher protective capacity at the MTS stage.
Future research work should evaluate the impact of different replenishment control policies and shop configurations in context of this hybrid production.

ACKNOWLEDGEMENTS

This study had the financial support of FCT-Fundação para a Ciência e Tecnologia of Portugal under the project PEst2015-2020: UID/CEC/00319/2013.

REFERENCES

Cox, J. F., III, and J. H. Blackstone, Jr, eds. 2002. APICS Dictionary Alexandria, VA: APICS.

Fernandes N.O., Silva, C. and Carmo-Silva, S. (2014), Order release in a hybrid MTO-MTS two-stage production system, *Pre-print of the 18th International Working Seminar on Production Economics,* February 24-28, Innsbruck, Austria.

Kaminskya, P. and O. Kayab. 2008. "Inventory positioning, scheduling and lead-time quotation in supply chains." *International Journal of Production Economics* 114 (1) 276–293.

Land, M.J., M. Stevenson and M. Thürer. 2014. "Integrating load-based order release and priority dispatching." *International Journal of Production Research*, 52, 4, 1059-1073.

Morikawa, K., K. Takahashi and D. Hirotani. 2014. "Make-to-stock policies for a multistage serial system under a make-to-order production environment." *International Journal of Production Economics*, 147, 30–37.

Oosterman, B.J., M. Land, and G.J.C. Gaalman. 2000. "The influence of shop characteristics on workload control." *International Journal of Production Economics*, 68 (1), 107–119.

Spearman, M. L., D.L. Woodruff, and W. J. Hopp. 1990. "CONWIP: a pull alternative to kanban." *The International Journal of Production Research*, *28*(5), 879-894.

Tatsiopulos, I.P., A. Batsis, and A. G. Papadopoulos. 2008. "Orders Loading and Release in Flow Shops including Outsourcing Networks.", *Lean Business Systems and Beyond, IFIP – The International Federation for Information Processing*, 257, 75-92.

Thürer, M., M. Land, M. Stevenson, L.D. Fredendall, and G. Filho. 2014. "Concerning Workload Control and Order Release: The Pre-Shop Pool Sequencing Decision, *Production & Operations Management",* (in Press: DOI: 10.1111/poms.12304)

PICKUP AND DELIVERY SELECTION WITH COMPULSORY REQUESTS

Katrien Ramaekers
An Caris
Gerrit K. Janssens
Tabitha Maes
Research Group Logistics, Faculty of Business Economics
Universiteit Hasselt – campus Diepenbeek
Agoralaan 1, 3590 Diepenbeek, Belgium
E-mail: {katrien.ramaekers,an.caris,gerrit.janssens,tabitha.maes}@uhasselt.be

KEYWORDS
vehicle routing, pickup and delivery problem, meta-heuristics, optimisation

ABSTRACT

The operational planning decisions of a carrier consist of accepting transport requests of customers and constructing daily vehicle routes. Several customers may have less-than-truckload requests to be transported between two specified locations: these customers are consolidated into vehicle tours by the carrier. However, a carrier has only a limited capacity within his own vehicle fleet. Therefore he can only serve a selection of clients. Transport requests of clients are only accepted if they contribute to a higher total profit. A paired pickup and delivery selection problem is hardly investigated in literature. Due to long term contract or for other commercial reasons, some of the requests cannot be neglected, even if they do not contribute fully to profit. This practical aspect is modeled with the Pickup and Delivery with selection of customers. A mixed-integer programming formulation is given. It is indicated how the problem is solved by means of a meta-heuristic, more specifically a tabu-embedded simulated annealing algorithm.

INTRODUCTION

Road transport is performed in two modes: the 'own account' mode and the 'hire and reward' mode. In the 'own account' mode, the owner of the vehicles and of the moved goods are identical. But many companies outsource their transport activities to reduce overhead costs. Logistics service providers are hired to execute the transport. Such companies are called freight carriers, which operate in the 'hire and reward' mode. Operational planning problems relate to decisions as: crew scheduling (assignment of crews to vehicles or transshipment facilities); empty balancing (preparation of the operations for the next planning period, and vehicle routing and scheduling (scheduling of the services for the pickup and delivery phases).

Pickup and delivery planning problems comprise the allocation of resources for fulfilling the tasks within a region for a given planning period. A transportation plan many times is solved as a vehicle routing and scheduling problem. But, in fact, before the problem is solved, the carrier has to decide whether a certain transport request is accepted for completion and, if yes, also the mode of completion (own equipment or another carrier who receives a charge) is selected.

The request selection is evaluated by means of corresponding revenues and costs. The cost evaluation of a request requires the determination of a transportation plan (routing). In some cases, it is possible to decide for each transport request whether it is fulfilled by the carrier or by a third party LSP. But sometimes contracts prohibit to outsource specific transport requests to a LSP. The customer requires fulfillment of the transport request by the carrier. Such a request is called a 'compulsory request'.

Several actors are involved with the transport of goods. To model freight transport, the actors involved in the decision making process have to be represented. In Maes et al. (2011) a conceptual framework is presented to model freight transport. The key actors in the framework include firms, carriers, and forwarders. These actors allow the model to work on an activity-based level, focusing on the different activities of each actor. The decision making process of carriers is one of the key aspects in modelling logistic decisions in a behaviour based transportation model. When modelling at an activity-based level, the behaviour of carriers needs to be taken into account. One of the decisions a carrier has to make is whether or not to accept transport requests he receives. Furthermore, he needs to plan the sequence of pickups and deliveries to optimize the use of vehicles given time and capacity limitations.

A carrier faces the daily problem of optimally scheduling his transport orders. Each day a carrier receives transport requests from his clients, which have to be executed within a certain time period. To obtain a maximal profit, the carrier has to group certain orders and create an optimal sequence of paired pickup and delivery tasks. In literature, this problem is called a pickup and delivery problem (PDP). Within a PDP mostly it is assumed that all requests have to be fulfilled. In reality a carrier may refuse a transport order, when he believes the order is not profitable. Sometimes non-profitable orders are accepted, due to reasons of competition or long-term commitment to a client. In such a case a carrier accepts the transport order which is less or non profitable, because it will generate other requests with a profit sufficiently high to offset the loss of the first transport order. In our conceptual framework only current requests are taken into account and the possible loss of future requests is ignored. Actors take decisions for one simulation period at a time. If a request is accepted, it will generate revenue when the transport is completed. When a carrier has to decide whether to accept a certain request, the problem is defined as a Pickup and Delivery Selection Problem (PDSP). This problem has been introduced by Schönberger et al. (2002).

LITERATURE REVIEW

The PDP is a generalization of the vehicle routing problem (VRP), which is a generalization of the traveling salesman problem (TSP) (Mitrovic-Minic, 1998). All of these problems have been widely investigated and numerous extensions have been developed. In a VRP generally all trip requests either originate or terminate at a single depot. In a PDP the trip requests are made between two locations that are outside the depot. In this section the division between paired and unpaired pickup and delivery points is used as in Parragh et al. (2008). Pickup and delivery vehicle routing problems are characterised with unpaired pickup and delivery locations. In this case an identical load is considered, and each unit picked up may be used to serve a delivery request. A classical pickup and delivery problem on the other hand, has paired pickup and delivery locations. Every request is associated with a paired origin and destination location and a specified load.

The focus of the literature review is on a specific case of the PDP, the pickup and delivery selection problem. In a PDSP not all transportation requests have to be fulfilled. A carrier receives transportation requests during the entire day. When additional requests are received, a decision has to be made whether the carrier will take the responsibility of the transport or not. The PDSP is NP-hard as it is a generalization of the travelling salesman problem. In literature this problem is not often investigated, but several variations on the problem exist. Two main bodies of routing literature are relevant for the PDSP. On the one hand the VRP with profits and on the other hand literature concerning PDP. The PDP is more relevant to the problem presented, however profit maximization has been more applied to VRP. In the next subsection first several vehicle routing problems with profits are presented, as the PDSP may be seen as a variation of these problems. Next, techniques used on the PDP are discussed which might be useful for the PDSP. To end this section the available literature on PDSP and variants of the problem are given.

VRP with profits

Feillet et al. (2005) give an overview of the TSP with profits. A distinction is made between three problem types, depending on the objective function. A first problem is called the profitable tour problem (PTP), which has as objective to simultaneously find a tour that minimizes travel cost and maximizes the collected profit. The problem studied in this paper, the PDSP, may be situated in this category. The second problem, the orienteering problem (OP), has as objective to maximize the collected profit while travel costs do not exceed a preset maximum cost. The last problem is known as the prize-collecting TSP (PCTSP). Here the collected profit is defined as a constraint, ensuring that the profit may not be smaller than a preset value.

Also vehicle routing problems exist for which it is not necessary to visit every node on the graph. The VRP with profits is the extension of a TSP with profits to multiple vehicles. Aksen and Arras (2005) study the single-depot capacitated VRP with profits and time deadlines (VRPPTD), in which it is not necessary to visit all customers. Their objective is to find the number and routes of vehicles to maximise the total profit.

The Multiple Tour Maximum Collection Problem (MTMCP) of Butt and Ryan (1999) is closely related to the VRP with profits. Due to limited availability of time not all nodes may be visited. Only nodes which give the highest contribution in terms of profit are selected. An optimal solution procedure for the MTMCP is described. This procedure is based on a generalized set-partitioning formulation and uses constraint branching and tour storage techniques to improve solution time.

Pickup and delivery problem

The pickup and delivery problem is used to find optimal routes, for a fleet of vehicles, to satisfy a set of transportation requests. Almost every practical PDP problem is restricted by several time constraints (Mitrovic-Minic, 1998). First, time windows determine when a load may be picked up or delivered at a certain location. Next, drivers of vehicles are restricted in their use by time windows. In most countries drivers may only drive a certain amount of time and are obligated to respect rest moments. In this section the main characteristics of paired or one-to-one pickup and delivery problems are presented. This means that each request originates at a single location and is meant for another destination (Cordeau, 2008).

Pickup and delivery problems may be divided into dynamic and static problems Savelsbergh and Sol (1995). In a dynamic problem not all request are known in advance, but may be received during the entire simulated period. Routes are constructed with the requests known at that time. When a new transportation request becomes available at least one route has to be adjusted. In a static problem all requests are known when the routes are constructed and no later adjustments to the planning are required. In this chapter the PDSP is defined as a static problem, with all requests known in advance.

Within the PDP various objective functions are used, depending on the purpose or criteria of the research. In Savelsberg and Sol (1995) an overview is given. The most common objective functions used by single vehicle problems are mainly related to minimizing duration, completion time, travelled time or client inconvenience. Problems with multiple vehicles mostly try to minimize the number of vehicles or maximize profit, this while minimizing the distance travelled or the travel time. This is also the case in Li and Lim (2001) their objective function exists of four elements. First the number of vehicles is minimized, than total travel cost and total schedule duration and finally the drivers' total waiting time is minimized. The maximization of profit, leads us to the pickup and delivery selection problem, which is only rarely applied within PDP.

Pickup and delivery selection problem

Few research articles investigate 'Paired pickup and delivery problems' in which profits determine the acceptance of an order. The PDSP adds complexity to the traditional PDP as it requires selecting which subset of nodes in the graph to visit, as well as determining the order of visits in each tour. Another difficulty is added when the nodes are constrained by time windows. Schönberger et al. (2002) considers the PDP in which not all nodes have to be visited to maximize profit. In this problem orders may be less-than-truckload. A hybrid algorithm is presented to solve the problem. The hybrid algorithm is composed of a genetic algorithm that is seeded by a parallel route construction heuristic. The construction heuristic generates a feasible solution by assigning requests to vehicles based on their order on a time axis. Requests that violate the capacity or time window restrictions are removed from the routes. After the construction heuristic, improvements are made using a genetic algorithm. Schönberger (2005) divides requests into two categories: tactical requests and operational requests.

'Tactical request' acceptance problems require a general decision about the future acceptance of different requests. Mostly, this type comprises all requests of a certain customer. Due to the long-term acceptance of certain requests, it may be necessary to require medium- or long-term investments for additional transport or transshipment resources. The general acceptance is recommended only if the agreed revenues cover the sum of necessary investment and operation costs.

'Operational request' acceptance problems require that the carrier company has to decide about the acceptance of particular requests, which are not part of long-term contracts. Such a request is accepted if expected revenues cover expected additional costs caused by this additional request. If a carrier refuses a customer demand, it may be expected that also all other requests of this customer are lost for this carrier. According to Schönberger (2005) lost revenues cannot be adequately incorporated into the calculation of the profitability of a request. In Arda et al. (2008) a profitable PDP with time windows is presented. The authors study orders of full truckload and try to maximize global profit while respecting time windows. To solve this NP-hard problem genetic algorithms are used. First a parent of an ordered set of transportation orders is made. A feasible solution can be extracted by choosing successively the first order that fits the time windows constraints. Schönberger et al. (2002) and Arda et al. (2008) use homogenous vehicle fleets and propose static models. They do not take into account a fixed cost of using an additional vehicle. Also Frantzeskakis and Powell (1990) investigate the PDSP. In their case a dynamic aspect is added and only full truckloads are considered. The carrier decides which loads he will accept or refuse and how many vehicles to relocate in order to maximize the total expected profit over a planning horizon.

Verweij and Aardal (2003) study in their merchant subtour problem the selection of optimal locations to buy and sell products. A selection is made so that the merchant may optimize his profit. The problem is a variation of the PDSP, in which only a single vehicle is considered.

Kleywegt and Papastravou (1998) propose a problem in which transport is done between terminals in long haul shipments. Vehicles are either at a terminal or en route between two terminals. The selection of clients happens at a terminal after which their loads are consolidated into vehicles and direct transport to a terminal is conducted. This leads to a full truckload problem in which clients are concentrated at terminals and not spread out in the area.

In Ting and Liao (2012) a selective pickup and delivery problem (SPDP) is formulated. This may be seen as a variant of the PDSP in the case of 'unpaired pickup and delivery nodes'. In the SPDP the constraint that all pickup nodes must be visited is relaxed. The objective is to find the shortest route for visiting all delivery nodes, without necessarily visiting all the pickup nodes as the nodes are not paired and only a single commodity is taken into account. The problem is solved using a memetic algorithm that allows to simultaneously deal with the selection of pickup nodes and the visiting order of nodes.

Another option within the PDP instead of not visiting all nodes is to outsource some of the requests to a third party logistic player. Schönberger (2005) investigates the possibility to make use of a logistic service provider (LSP). In this case all requests are divided between either the own vehicle fleet or the LSP. Routes have to be established for the own vehicles and the sum of charges to be paid for all externalized requests has to be minimized. Krajewska and Kopfer (2009) also study a PDP where the carrier has the possibility to outsource transport requests. They make use of a tabu search algorithm to solve their Integrated Transportation Planning Problem (ITPP). The main difference with other studies that include outsourcing is the use of three different outsourcing types instead of one. A first group of subcontractors works nearly exclusively for the carrier and is paid on tour basis. The second group of exclusively employed subcontractors is paid on a daily basis. The last group consists of independent subcontractors which are not employed exclusively.

This paper offers the following novelties compared to existing research. The traditional PDP is extended to a PDSP by allowing a selection of transportation requests. This leaves the carrier with the option to discard transportation requests which lead to a lower total profit. Next to the planning and scheduling of vehicles into routes as in a classical PDP, a selection within the transportation requests has to be made. The problem at hand considers more than one commodity and paired pickup and delivery locations. This is different to the study of Ting and Liao (2012) where a single commodity is considered and pickup and delivery are unpaired. Furthermore, multiple vehicles are considered and transport loads are less-than-truckloads.

In the study of Verweij and Aardal (2003) only a single vehicle is assumed and in the work of Arda et al.

(2008), Frantzeskakis and Powel (1990) and Kleywegt ad Papastravou (1998) full truckloads are investigated. The paired pickup and delivery locations, together with the multiple vehicles and less-than-truckload requests make the PDSP very hard to solve. The only paper that studies a PDSP with similar problem characteristics but in a different problem context is Schönberger et al. (2002). Their heuristic results are not compared to exact solutions or lower bounds and reported results are only briefly described. This hinders the comparison of computational results.

PROBLEM FORMULATION

In this section a mathematical representation of the problem is given. First, the key characteristics of a PDSP are described. Next, all symbols that are used are presented and the objective function and corresponding constraints are formulated. The problem is defined as a static PDSP problem. The formulation is an adaptation of the PDPTW formulation of Mitrovic-Minic (1998) and is extended to include the selection of requests.

Key characteristics of a PDSP

To represent logistic decisions within an activity based freight transportation model, the decisions of a carrier have to be modelled. First, the key characteristics related to this problem are presented. This allows formulating a PDSP model in the next subsection.

First of all, not all requests have to be accepted. Every fulfilled request leads to revenue. If a request is accepted, a reward is achieved when the transport is done successfully. For every request a time window is assigned to the pickup and delivery location. In the problem definition at hand, only hard time windows are considered. A request consists of less-than-truckloads. Furthermore, pickup has to occur before delivery of each request (Precedence constraint) and pickup and delivery have to be performed by the same vehicle (Pairing constraint). In our model multiple vehicles are used and it is assumed that capacity is the same for all vehicles. All vehicles depart from and return to a depot of the carrier. Finally, travel costs and travel times for each link are known and assumed to be constant.

Introduction of symbols

Requests A carrier receives a set P of requests. Because the set of requests is equal to the number of pickup locations the same symbol is used in both cases. Each request $r \in P$ consists of a pickup location p_i, a delivery location d_{i+n}, a quantity to be shipped q_i and a revenue Rev_i if the request is completely satisfied. So each request is given as a quadruple. The quantity q_i may either be a positive or negative number, depending on the type of operation, either a pickup or a delivery task.

Locations Three different types of locations may be distinguished, each with their own time window. A set of pickup locations $P=\{1,...,n\}$ and a set of delivery locations $D=\{n+1, n+2,...,2n\}$ are included, each with an earliest operation time e_i, a latest operation time l_i and a quantity q_i that needs to be shipped or delivered. A single depot O is available, where each vehicle starts and ends his route. If this is a start location, the depot is denoted as node 0. For an end location the notation $2n+1$ is used for the node.

Network A network $G(A,N)$ is given, with $N = P \cup D \cup O$, the set of nodes and A a set of undirected arcs. Within the network the distance between two nodes i and j is given as d_{ij}. The travel cost ct, expresses the charge for travelling a single distance unit. The cost to travel a certain link is expressed as $ct \cdot d_{ij}$. The last variable on the network is t_{ij} which stands for the time needed to travel from node i to node j.

Vehicles The carrier has a given homogenous fleet K of own vehicles. Each vehicle k has a maximum capacity of Q_{max}. Vehicles are bound in time by their driver, who is subject to legal driving time restrictions. Only the total amount of driving hours is checked, not the daily rest requirements. Making the assumption that a carrier has the same amount of drivers as vehicles, each vehicle has a start time s_k and a finish time f_k. The difference between f_k and s_k may not exceed the legal driving time of the driver. To keep track of the content of the vehicle, so that it will not exceed the maximum capacity, load variables L_i^k are introduced.

Operations A vehicle has to perform several operations on its route. Each pickup and delivery task takes a certain amount of operation time ot_i to perform per unit that needs to be handled. The total time a vehicle spends at the pickup or delivery location, can be found as follows: $ot_i \cdot q_i$. Due to the hard time windows applied at each pickup and delivery location, a vehicle cannot start his pickup operation until after time e_i. The vehicle is allowed to arrive earlier at the location, but must then wait until the start of the time window. A vehicle may never arrive to a location after the end of the time window l_i. Different waiting protocols may be defined depending on the solution heuristic used. It may be preferred to drive first and wait at the arrival location or to wait first at the previous location and then drive. The empirical study of Mitrovic-Minic and Laporte (2004) shows that the wait first strategy has the potential to build shorter routes compared to drive first in case of dynamic planning problems. Waiting at the starting positions results in more requests being known at the time they leave and a better potential to optimize the route. On the other hand the study revealed that the wait first strategy requires much more vehicles for the same set of locations. Therefore, a new waiting strategy (advanced dynamic waiting) was introduced. Here the drive first and wait first strategy are combined by serving locations in one service zone according to the drive first strategy and apply the wait first strategy between different service zones. This strategy was able to outperform the common used drive-first waiting strategy (Mitrovic-Minic and Laporte, 2004). For static vehicle routing, as the problem at hand, drive first is the most commonly used waiting strategy and will be used for the PDSP.

Variables For this problem two sets of binary variables are defined.

$$X_{ij}^k = 1 \; if \; vehicle \; k \; travels \; from \; i \; to \; j; 0 \; else$$

$$Y_i^k = 1 \; if \; vehicle \; k \; performs \; request \; i; 0 \; else$$

Next to these binary variables, two sets of continuous variables are introduced.

$$T_i^k = the\ time\ of\ vehcile\ k\ after\ node\ i\ is\ served$$
$$L_i^k = the\ load\ of\ vehicle\ k\ after\ node\ i\ is\ served$$

Objective function

The objective of the PDSP is to maximize the profit collected along the vehicle tours. Profit is defined as the sum of the total revenue collected on all the tours minus the total cost of performing the tours.

Profit = Revenue - Cost

For the revenue *Rev* a table is created which stipulates the price of the transport order in function of the distance to be travelled. The total revenue is found by accumulating all revenues of the requests that are accepted and executed .

$$Rev_{tot} = \sum_{k\in K}\sum_{i\in P} Rev_i . Y_i^k$$

The total cost (C_{tot}) is calculated as the sum of the costs of each link travelled by a certain vehicle k. In this case the cost is only related to the distance being travelled. No fixed cost component is enclosed for the use of a vehicle. The assumption is made that a carrier has a fixed vehicle fleet at his disposition and no extra cost is imposed for the use of the vehicles.

$$C_{tot} = \sum_{k\in K}\sum_{i\in N}\sum_{j\in N} ct . d_{ij} . X_{ij}^k$$

The objective function which needs to be maximized is:

$$max[Rev_{tot} - C_{tot}].$$

Constraints

In this section the constraints to which the objective function is subjected are formulated. The constraints are grouped according to their function.

Flow conservation constraints This constraint is introduced to make sure that vehicles entering a location will also leave this location.

$$\sum_{\substack{j=1\\ j\neq i}}^{N} X_{ij}^k - \sum_{\substack{j=1\\ j\neq i}}^{N} X_{ji}^k = 0, \forall i \in N, \forall k \in K$$

Vehicle constraints Each vehicle starts and ends his tour in the depot O. If a vehicle is not used it stays at the depot. This is represented by the following constraints.

$$\sum_{j\in P} X_{0j}^k \leq 1, \forall k \in K$$
$$\sum_{i\in D} X_{i(2n+1)}^k \leq 1, \forall k \in K$$

Every request may only be executed by at most one vehicle. Let the index i^{comp} be used for the requests which are compulsory and i^{Ncomp} for those which are non-compulsory. The constraints to express this are:

$$\sum_{k=1}^{K} Y_{i^{Ncomp}}^k \leq 1, \forall i \in P$$

$$\sum_{k=1}^{K} Y_{i^{comp}}^k = 1, \forall i \in P$$

A next constraint states that a vehicle cannot load more freight than its maximum capacity.

$$L_i^k \leq Q_{max}$$

To keep track of the load of a vehicle at a certain moment, the following constraints are necessary. Each vehicle leaves and returns to the depot empty.

$$L_0^k = 0, \forall k \in K$$
$$L_j^k - L_i^k - |q_j| \geq M_1.(1 - X_{ij}^k),$$
$$\forall i, j \in N\ and\ i \neq j, \forall k \in K$$

Time window constraints

Each node has to be served within its time window. The start of the operation, as well as the end of the operation has to fall within the time window.

$$e_i + ot_i.|q_i| \leq T_i^k \leq l_i, \forall i, j \in N\backslash 0, \forall k \in K$$

To keep track of time, a time variable is introduced. To start, the time variable is set equal to the start time of the vehicle.

$$T_0^k = s_k, \forall k \in K$$

A vehicle may not exceed his finish time.

$$s_k \leq T_i^k \leq f_k, \forall i \in N, \forall k \in K$$

The time variable is increased after every operation. The time after service at a certain node, is found by adding the travelling time and operation time to the time variable after serving the previous node. Also the time windows have to be respected, so that the arrival time at a node may not precede the earliest operation time allowed on that location. This is specified in the following constraint:

$$T_i^k + t_{ij} - T_j^k + ot_j . |q_j| \leq (1 - X_{ij}^k) . M_2,$$
$$\forall i, j \in N, \forall k \in K$$

Due to the time window constraint on T_i^k, it is assured that the operation does not start before e_j.

Pairing and precedence constraints

If a request is performed, then vehicle k has to finish its operations at the pickup location i before it can visit the associated delivery location $n+i$. This is known as the precedence constraint, expressed as:

$$T_i^k + t_{i(n+i)} - T_{n+i}^k \leq (1 - Y_i^k) . M_2, \forall i \in P, \forall k \in K$$

It is not allowed to split a request over multiple vehicles. A vehicle has to perform both the pickup and the delivery activity. This is known as the pairing constraint, expressed as:

$$\sum_{j\in N\backslash O} X_{ij}^k = Y_i^k, \forall i \in P, \forall k \in K$$

and

$$\sum_{j\in N\backslash O} X_{j(n+i)}^k = Y_i^k, \forall i \in P, \forall k \in K.$$

In the formulation two big M-values are used, where M stands for a sufficiently large number. The value of M_1 is set equal to the maximum load capacity of the vehicles, and the value of M_2 is set equal to the maximum length of a working day.

SOLUTION METHOD

The problem under study is solved by means of a metaheuristic. While this paper does not intend to explain detail the design of the metaheuristic, the experiments and the computational results, briefly a description is given.

The heuristic is based on the tabu-embedded simulated annealing algorithm of Li and Lim (2001). The algorithm starts with an insertion heuristic to create a first feasible solution. This solution is further improved by an improvement heuristic. Instead of repeating the tabu search until the procedure terminates, it is restarted from the current best solution after several iterations without improvement. At the same time the global annealing temperature is reset. After a number of restarts without improvement the algorithm is ended. The generation of new best solutions is done via a TABU Search algorithm. To avoid cycling, the visited solutions are recorded into a tabu list, which contains the total profit of a solution. As the probability of two different solutions having the same total profit is very small, it is sufficient to only keep track of total profit.

CONCLUSIONS

A variant of the Pick-up and Delivery Vehicle Routing Problem is investigated, in which the carrier can make a selection of customers to be served, depending on which transport requests offer profit. Furthermore a number of requests are compulsory and are not open to the question of selection. We have been able to formulate the optimization problems as a mixed-integer linear programming problem and have developed a meta-heuristic to solve the problem.

ACKNOWLEDGEMENTS

This work is supported by the Interuniversity Attraction Poles Programme initiated by the Belgian Science Policy Office (research project COMEX, Combinatorial Optimization: Metaheuristics & Exact Methods).

REFERENCES

Aksen D. and N. Aras (2006). “Customer selection and profit maximization in vehicle routing problems”. In: Operations Research Proceedings 2005, (eds.) Haasis, H-D. d Kopfer, H. and Schönberger, J., part III, 37-42.

Arda Y., Y. Crama and T. Pironet (2008). “A profitable pickup and delivery problem with time windows”. ORBEL 22(22nd National Conference of the Belgian Operations Research Society). Booklet of abstracts, Brussels, 26-28 January, pp. 76-77 (http://orbel22.hallot.net/docs/BookletORBEL22.pdf

Butt S.E. and D.M. Ryan (1999). “An optimal solution procedure for the multiple tour maximum collection problem using column generation”. *Computers & Operations Research*, 26, 427-441.

Feillet D., P. Dejax and M. Gendreau (2005). Travelling salesman problems with profit. *Transportation Science*, 39(2), 188-205.

Frantzeskakis, L.F. and W.B. Powel (1990). A Successive Linear Approximation Procedure for Stochastic, Dynamic Vehicle Allocation Problems. *Transportation Science*, 24(1), 40-57.

Kleywegt A.J. and J.D. Papastravou (1998). Acceptance and dispatching policies for a distribution problem. *Transportation Science*, 32, 127-141.

Krajewska M.A. and H. Kopfer (2009). Transportation planning in freight forwarding companies: Tabu search algorithm for the integrated operational transportation planning problem. *European Journal of Operational Research*, 197, 741-751.

Li, H. and A. Lim (2001). A metaheuristic for the pickup and delivery problem with time windows. *Proceedings of the 13th IEEE International Conference on Tools and Artificial Intelligence,* November 2001, Dallas TX, USA, pp. 160-167.

Maes T., K. Ramaekers, A. Caris, G.K. Janssens and T. Bellemans (2011). Simulation of logistic decisions within a freight transportation model, In: S. Balsamo (ed.), *Proceedings of the 2011 Industrial Simulation Conference (ISC'2011)*, Venice, Italy, pp. 185-191.

Mitrovic-Minic S. (1998). “Pickup and delivery problem with time windows: a survey”, Technical report TR1998-12, School of Computers Science, Simon Fraser University, Burnaby, BC, Canada (ftp://fas.sfu.ca/pub/cs/techreports/1998).

Mitrovic-Minic S. and G. Laporte (2004).Waiting strategies for the dynamic pickup and delivery problem with time windows. Transportation Research Part B, 38, 635-655.

Parragh S.N., K.F. Doerner and R.F. Hartl (2008). A survey on pickup and delivery problems. Part II: Transportation between pickup and delivery locations, *Journal für Betriebswirtschaft*, 58(2), 81-117.

Savelsberg M.W.P. and M. Sol (1995). The general pickup and delivery problem. Transportation Science, 29:1, 17-29.

Schönberger J., H. Kopher and D. Mattfeld (2002). “A combined approach to solve the pickup and delivery selection problem”, Operations Research Proceedings 2002, In: Leopold-Wildburger, U., Rendl, F. and Wascher, G., pp. 150-155.

Schönberger J. (2005). *Operational freight carrier planning. Basic concepts, optimization models and advanced memetic algorithms*. Springer, Berlin (ISBN -540-25318-1).

Verweij B. and K. Aardal (2003). The merchant subtour problem, Mathematical Programming, 94(2-3), 295-322.

Ting C.-K. and X.-L. Liao (2013). The selective pickup and delivery problem: formulation and a memetic algorithm. *International Journal of Production Economics*, 141(1), 199-211.

A BI-OBJECTIVE SIMULATION-OPTIMIZATION APPROACH FOR SOLVING THE DYNAMIC VEHICLES ROUTING PROBLEM IN BICYCLE SHARING SYSTEMS

Ahmed A. Kadri
ECAM-EPMI Cergy-Pontoise
LCOMS, Université de Lorraine
Ahmed-abdelmoumene.kadri
@univ-lorraine.fr

Karim Labadi
ECAM-EPMI, ECS-Lab
13, Bd de l'Hautil, Cergy-Pontoise
k.labadi@ecam-epmi.fr

Imed Kacem
LCOMS
Université de Lorraine
Imed.kacem@univ-lorraine.fr

KEYWORDS

Bicycle sharing systems, Discrete-event simulation, Modeling, Performance optimization, Stochastic Petri nets, Genetic algorithm, vehicles routing.

ABSTRACT

Bicycle Sharing Systems represent an ecological urban mode of transportation allowing people to rent a bicycle at one of many automatic rental stations scattered around the city. In order to meet the users' demands, bicycle sharing systems must be rebalanced regularly. The problem is to design (near-) optimal vehicles routes to visit and rebalance the stations so as to minimize the number of users who try to collect bikes from empty stations or to deliver bikes to full stations. This paper deals with the vehicles routing problem in the dynamic case by extending our previous work based on stochastic Petri net model coupled with a genetic algorithm. We consider the minimization of the uncomfortable situations overall network stations and the total operational cost simultaneously. The proposed simulation-optimization method is tested on a large set of instances and then applied to Cristolib, a real self-service bicycle system of Creteil city, France.

INTRODUCTION

Bike Sharing Systems (BSS) are a very popular means providing bicycles to users in a cheap and self service way. They are one of the solutions to face many public transportation problems, including traffic congestion, air pollution, global oil prices, and global warming. The systems can be described as a set of stations scattered around a city usually located 300 meters apart, and each one is equipped with a terminal and a limited bicycle stands. This systems are very useful for short travels around a city, a user can pick a bicycle at one of any station and return it to the same or any other station, provided that there is an empty locking berth.

Although their apparent success in the world, the exploitation of such transport systems implies crucial challenges and one of them is to ensure users that they will be able to find a bicycle or to park it at each station all the day long. Indeed, some stations have more demand than others, especially during peak hours. Bicycles also tend to collect in stations located in the city centers and stay there. In these situations, if no action is taken by the service provider they rapidly fill or empty, thus preventing other users from collecting or delivering bicycle. At an operational level, in order to be able to meet users demands with a reasonable standard of quality, vehicles are used to balance the bicycles among the stations. The objective is to minimize the number of users who cannot be served, i.e., the number of users who try to take a bike from an empty station or to return it to a full station. In the practice, the rebalancing operation can be carried out in two different modes (Raviv et al. 2011). The first one is static, the bicycle redistribution operation can be carried out during the night when the usage rate of the BSS system is very low and the bicycle repositioning is performed based on the status of the system at that time and the demand forecast for the next day. This problem has been recently addressed in some works (Benchimol et al. 2011; Chemla et al.2012; Kadri et al. 2014; Kadri et al. 2015; Raviv et al. 2013; Shu et al. 2010; Forma et al. 2010). The second mode is dynamic, the bicycle redistribution operation can be carried out during the day when the usage rate of the BSS is significant and bicycle repositioning is performed based on the current state of the station (Kadri et al. 2013; Kadri et al. 2014; Contardo et al. 2012; Sayarshad et al. 2012).

In our previous works (Labadi et al. 2010; Labadi et al. 2012), we developped a stochastic Petri nets based approach for modelling, control and performance analysis of BSS considered as discrete event systems. However, such models must be parameterized to provide a good perfomance of such systems. Therefore, we proposed in (Kadri et al. 2013; Kadri et al. 2014) an optimization approach cosidering these models with the aim of maximizing the QoS.

This paper deals with the vehicles routing problem in the dynamic mode and takes into account both the users satisfaction and the total operational cost. This work can be considered as an extension of our previous work (Kadri et al. 2013; Kadri et al. 2014), where we developed an integrated stochastic Petri net and genetic algorithm approach suitable for performance modeling and optimization for control and balancing purposes of such complex dynamical systems. To the best of our knowledge, except our works no other works has been undertaken on BSS modeling and performance optimization from a stochastic and discrete-event point of view. In (Kadri et al. 2013; Kadri et al. 2014), the objective was to optimize the quality of service in BSS by investigating the impact of the regulation thresholds on users satisfaction

and by considering a fixed route for the regulation vehicles. In this work, we consider a known intervals of thresholds for each station of the system. In practice, they can be defined according to statiscial analysis of data provided from a BSS survey during a given period of normal functioning.
The remainder of this paper is as follow: In section 2 we present a mathematical formulation for the problem and we define the objective and the constraints. We describe in section 3 our BSS Petri net model presented in our last works (Labadi et al. 2010; Labadi et al. 2012; Kadri et al. 2014). In section 4, we present an optimization approach based on coupling the Petri net model and a genetic algorithm. In section 5, we present the experimental results and discuss the efficiency of our approach. Finally, we conclude the paper with some discussions and perspectives.

MODEL FORMULATION

The problem is to design a dynamic routes for vehicles visiting the stations in order torebalance them, so as to minimize the number of users who try to collect bikes from empty stations or to deliver bikes to full stations. Hence, our bi-objective function (1) includes a term which represents the users' satisfaction defined as a sum of the critical times (empty and full stations), as well as a term representing operating cost which is proportional to the total distance traveled by the repositioning vehicles. The sum of the two components represents the objective to be minimized.
In this work, we consider only a single vehicle routing problem, since the regulation is done by area in real situations. We assume the existence of a depot as the starting point of each vehicle's route. However, we note that the depot may be viewed as a regular station, except that it typically has a relatively large capacity, large initial inventory and no demand. The above decisions are subject to capacity constraints of the vehicles, the stations and the depot, as well as time constraints concerning the total traveling, loading and unloading times. The latter two components are assumed to be linear with the number of bicycles loaded/unloaded.
Additional modeling choices are associated with the permissible actions of the fleet of vehicles performing the repositioning operation, in particular, limitations on the allowed set of routes which the goal is to visit each station one and only one time in a tour.
Before presenting the mathematical formulation of our problem, we introduce first the following notations, variables, and parameters :

N	A set of nodes including the network stations, indexed by $i = 1,.., n$.
N_0	A set of nodes, including the network stations and the depot (denoted by $i=0$), $i = 0,.., n$.
E_i	Current number of bicycles at the station i before the repositioning operation.
R_i^-	Minimal level of the required number of bicycles at a station i.
R_i^+	Maximal level of the required number of bicycles at a station i.
$d_{i,j}$	Vehicles traveling distance from a station i to a station j.
t_i	Arrival time of a vehicle to the station i, $i = 1,..,n$.
$_i^-$	Represents the gap between Ei and R_i^-, where: $_i^- = R_i^- - E_i$ **if** $E_i < R_i^-$.
$_i^+$	Represents the gap between Ei and R_i^+, where : $_i^+ = E_i - R_i^+$ **if** $E_i > R_i^+$.
Cv	Capacity of redistribution vehicle.
Dv_i	Number of available bicycles in the vehicle at station i.
pt_i	Parking time of the vehicle at a station i.
t_l	Time required to remove a bicycle from a station and load it to the vehicle.
t_u	Time required to unload a bicycle from the vehicle to a station.
T_{sim}	Total time of simulation.
$M_i=0$	Represents an unavailability event in a station i.
$M_i=C_i$	Represents a saturation event in a station i.
$T_{(Mi=0)}$	Total time of unavailability of bicycles at station i during the simulation (Empty station).
$T_{(Mi=Ci)}$	Total time of saturation event at station i during the simulation (Full station).
ΔT	Event duration.
$x_{i,j}$	Binary variable which equal to 1 if a vehicle travels from i to j, 0 otherwise.

Based on the notations, the following mathematical model can be formulated. first, we introduce the objective function that contains two components: first, the sum of the remaining times in critical situations (empty, full) over all stations $S_i \in N$, and second the total travel distance for the redistribution vehicle:

$$\text{Minimize } \omega_1 \sum_{i=1}^{n} \left(T_{(M_i=0)} + T_{(M_i=C_i)} \right) + \omega_2 \sum_{i=0}^{n} \sum_{j=0}^{n} d_{i,j} x_{i,j} \qquad (1)$$

S.C.

$$\sum_{j=1}^{n} x_{0,j} = 1 \qquad (2)$$

$$\sum_{j=1}^{n} x_{j,0} = 1 \qquad (3)$$

$$t_0 = 0 \qquad (4)$$

$$t_j \geq t_i + x_{i,j} d_{i,j} + t_u \partial_i^- + t_l \partial_i^+ + pt_j + (x_{i,j} - 1) * M$$
$$\forall i \neq j \ , \ j \in N \ , \ i \in N_0 \qquad (5)$$

$$\sum_{i=0}^{n} x_{i,j} = 1 \qquad \forall j \in N_0 ,\ i \neq j \qquad (6)$$

$$\sum_{i=0}^{n} x_{j,i} = 1 \qquad \forall j \in N_0 ,\ i \neq j \qquad (7)$$

$$\partial_i^+ \leq C_v - Dv_i \qquad \forall i \in N \qquad (8)$$

$$Dv_i \geq \partial_i^- \qquad \forall i \in N \qquad (9)$$

$$\mathrm{T}_{(\mathrm{M_i}=0)} = \sum_{t=0}^{\mathrm{Tsim}} (\mathrm{M_i} = 0) * \Delta \mathrm{T} \qquad i \in N \qquad (10)$$

$$\mathrm{T}_{(\mathrm{M_i}=\mathrm{C_i})} = \sum_{t=0}^{\mathrm{Tsim}} (\mathrm{M_i} = C_\mathrm{i}) * \Delta \mathrm{T} \qquad i \in N \qquad (11)$$

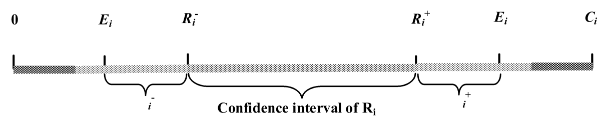
$$x_{i,j} \in \{0,1\} \tag{12}$$

$$\omega_1, \omega_2 > 0 \tag{13}$$

The equation (1) minimizes the weighted objective consisting of the sum of the remaining times of stations in the critical situations (empty and full) and the total distance traveled by the vehicle. The constraint (2) and (3) ensures the start and the end of a tour at the deposit. Constraint (4) represents the departure time of the vehicle from the depot, and constraints (5) insure the minimal time required for the displacement of the vehicle from a station *i* towards a station *j* including parking time when arriving to destination and unloading times. The latter two components are assumed to be linear in the number of bicycles picked/delivered. Constraints (6) (resp.,(7)) are vehicle flow-conservation equations (a station can be visited only one time by a vehicle, resp., only one vehicle can exit a station). constraints (8) (resp.,(9)) are vehicle capacity constraints (available places in vehicle to pick up bicycles from a station, resp., available bicycle to delivery to a station). Constraints (10) (resp.,(11)) represent the remaining times of the stations in critical situations (empty stations, resp., full stations). Finally, (12) are binary constraints. The weighting values of the objective function are linked to the bike rental pricing and the vehicles cost travel measured by the total distance. This weights can be calculated using the Pareto techniques.

PETRI NET MODEL FOR BSS

As introduced, we already proposed a Petri net approach dedicated for BSS modeling and analysis for control purposes. We consider a BSS with *n* stations noted by $S = \{S_1, S_2, \quad , S_n\}$. Each station ($S_i$) is equipped with C_i bicycle stands (the capacity of a station S_i). The system requires a constant control which consists in transporting bicycles from stations having excess of bicycles to stations that may run out of bicycles soon. In the general way, the main objective of the control system, performed by using redistribution vehicles, is to maintain a bicycle safety level at each station S_i to ensure the availability of a minimal number of bicycles for pick up denoted (R_i^-), thus the availability of (C_i - R_i^+) attachment points for return bicycle at each station.



The Petri Net model consists of three subnets (modules) representing three different functions indicated as follows: (a) the “station control” subnet; (b) the “bicycle flows” subnet; and (c) the “redistribution circuit” subnet. The main function of each subnet is described as follows:

- A “bicycle flows” subnet represents the bicycle traffic flows between the different stations of the network.
- A “station control” subnet represents the control function of the stations to ensure bicycles are available for pick up and vacant berths available for bicycle return at every station S_i.
- A “redistribution circuit” subnet represents the path (circuit) to be followed by the redistribution vehicle in order to visit the different stations of the network.

More details on the dynamic of the developed modeling approach can be found in our previous papers (Labadi [12]-[13]).

PN-GA OPTIMISATION APPROACH

Motivated by the randomness of BSS and the NP hardness of the studied problem, we developed an approach based on the integration of the Petri net model for BSS and a genetic algorithm in order to solve the vehicles routing problem. Although Petri nets itself is not an optimization tool, it can provide a practical decision making tool, if it is combined with and enhanced by some optimization algorithms (Kadri et al. 2013; Kadri et al. 2014; Nodhi et al. 2012; Sauer and Xie 1993; Sharda and Bnejee 2013, Sadrieh et al. 2007). In this part of this work, a genetic algorithm is coupled with the Petri net model in order to solve the problem formulated previously. Our discrete event simulation/optimization approach is illustrated by the Figure1.

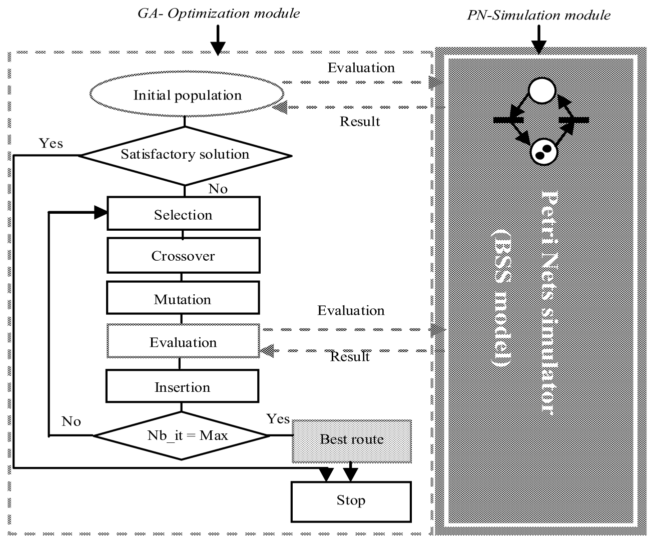


Figure 1: PN-GA optimization approach

From an optimization point of view, our genetic algorithm follows classical genetic algorithm steps. The solutions are coded as chromosomes (Holland 1975), which may contain one or more parts of parameters of the modeled system. Since we investigate the vehicle routing problem, coding of chromosomes is introduced by a sequence containing stations in their visiting orders. An initial population is generated and evaluated via the simulation of BSS-Petri Nets model, which returns the result of the objective function for each chromosomes. After evaluation, genetic algorithm performs the operators of crossovers, mutations and insertions until the stop conditions of algorithm. Particularly. Figure 2 illustrates the chromosome presentation and the different operators performed by the genetic algorithm.

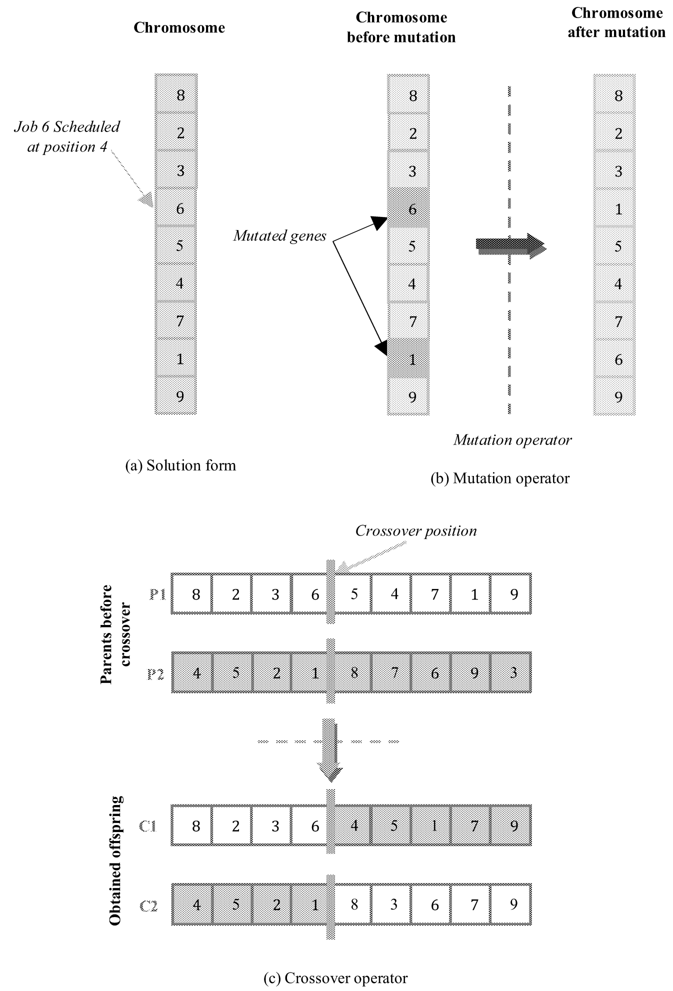


Figure 2: Genetic Operators

As shown in Figure 2(a), the chromosome representing a vehicle route constitutes sequence, the genes contain the assignments of jobs for a vehicle in their execution order, thus a job represents a displacement of a vehicle from a station *i* to *j*. Different genetic operators are applied as crossovers and mutations under fixed probabilities. After evaluation, each individual has an adjustment level which presents the value of the bi-objective function defined in equation (1) of the mathematical formulation. The evaluation of the objective is done via the simulation of PN-model and the evaluation of simulation results. We note that this step is the most expensive in terms of execution time of our GA.

A REAL APPLICATION CASE

We describe in this section the numerical experiments carried out in order to evaluate our approach. To integrate the Petri net model and the genetic algorithm, we developed in C ++ language a tool for both modeling, simulation, performances evaluation and parameters/decisions optimization of BSS. The experimentations were carried out on an Intel Core i7 2.7 GHz processor and 8 GB of RAM, under Windows 8 environment.

Data setting

We tested our approach for a real case study for the BSS (Cristolib) of Creteil city, France (http://www.cristolib.fr). As shown in Figure 3, the network consists of 10 stations with a total of 130 bicycles. Each station is characterized by its capacity and an average regulation threshold (R_i^-, R_i^+), as indicated in Table 1. The travel times of vehicles between the different stations including the deposit, estimated using Google map application, are shown in Table 2. For this application, the capacity of the vehicle is fixed to 30 bicycles and the times allocated to a pickup or dilivery operation are fixed to 10 seconds per bike. The average frequentation of the stations, given in Table 3, represents user requests measured during a 10 days of normal functioning (excluding summer season, special events,...). We also specify that the total simulation time for the execution of the PN representing the system is fixed to two weeks. This time is relatively long, but it allows us to obtain a reliable estimation of the average performance of the system during one day's service.

Table 1 : Stations Capacities (Source: http://www.cristolib.fr).

Station	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
Capacity	26	22	24	40	18	24	20	34	22	20
Aerage thereshold	13	11	12	20	9	12	10	17	11	10

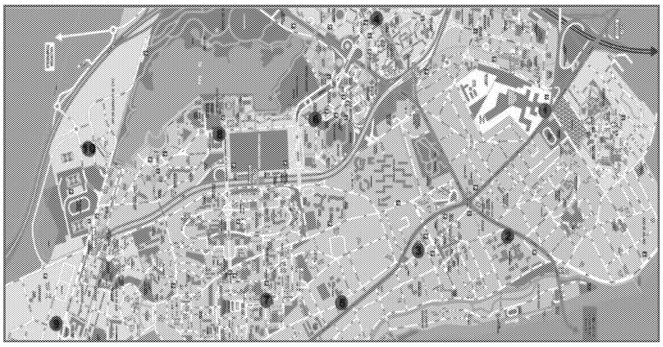


Figure 3: Distribution of bike stations in the network (Source: http://www.cristolib.fr)

Table 2 : Estimation of displacement delays of a redistribution vehicle (in minutes), source Google map.

Station	*D*	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
D	0	5	8	9	12	14	16	12	15	18	20
S_1	5	0	3	3	4	3	6	6	7	7	9
S_2	8	3	0	3	6	4	6	7	7	9	10
S_3	9	3	3	0	6	2	6	4	7	7	11
S_4	12	4	6	6	0	6	4	6	6	6	8
S_5	14	3	4	2	6	0	6	2	6	6	9
S_6	16	6	6	6	4	6	0	5	3	4	5
S_7	12	6	7	4	6	2	5	0	4	6	8
S_8	15	7	7	7	6	6	3	4	0	5	4
S_9	18	7	9	7	6	6	4	6	5	0	4
S_{10}	20	9	10	11	8	9	5	8	4	4	0

Table 3: Average travel demand matrix, source: statistical average of user requests for a day.

Station	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
S_1	6	6	4	18	13	4	4	14	10	10
S_2	11	8	2	4	17	11	8	13	6	12
S_3	15	15	5	11	13	19	17	13	14	4
S_4	15	5	25	3	22	8	8	10	4	14
S_5	6	18	3	31	4	6	18	9	7	8
S_6	22	12	5	16	4	2	12	4	16	11
S_7	10	14	17	14	9	18	15	5	16	17
S_8	14	18	17	15	6	10	9	8	17	12
S_9	13	9	30	5	8	8	11	13	2	8
S_{10}	11	6	5	4	17	11	8	8	6	12

Algorithm Setting

We note that the variations of GA parameters can impact on the quality of solution. In this application, we set the GA parameters as following : the initial population size is set to 20, we consider a one point crossover with a probability (P_{cross}=0,9) and we perform one mutation by swapping two elements in chromosomes with a probability (P_{mut}=0,1). The iterations number is limited to 200.

We remember that setting the weights ω_1 (resp. ω_2) of the bi-objective function weighting the critical times (resp. total distance) is a pricing cost problem since they are linked to the bike rental pricing and (resp. the operational cost of a vehicle which is measured by the total traveled distance).

One of the advantages of our Petri net model is its associated performance analysis methods which allow an efficient analysis and evaluation of the system behavior during simulation time. In these tables, several performance indices are given for each station such as:

The average time (%) where a station remains in empty (resp. full) state is calculated by using the equations :

$$Empty\ rate\ (S_i) = 100 * \frac{T_{sim} - \sum(\tau_k - \tau_{k+1})_{M(PS_i)\,=\,0}}{T_{sim}} \tag{14}$$

$$Full\ rate\ (S_i) = 100 * \frac{T_{sim} - \sum(\tau_k - \tau_{k+1})_{M(PS_i)\,=\,C_i}}{T_{sim}} \tag{15}$$

$$Average\ number\ of\ bicycles\ (S_i) = \frac{\sum M(PS_i) * \tau}{T_{sim}} \tag{16}$$

The average critical time (%) where a station remains in empty or full states (i.e., sum of the equations (14) and (15)). We define also the quality of service QoS for each station as the average service time (%) defined by the equation (17). The quality of service of the network is then ($\sum QoS$) / n.

$$QoS = \left(1 - \left(\frac{T_{sim} - \sum(\tau_k - \tau_{k+1})_{M(PS_i)\,=\,0}}{T_{sim}} + \frac{T_{sim} - \sum(\tau_k - \tau_{k+1})_{M(PS_i)\,=\,C_i}}{T_{sim}} \right) \right) * 100 \tag{17}$$

Where T_{sim} is the total simulation time; (τ_k - τ_{k-1}) represents the cycle duration where the station S_i is empty (resp. full) (i.e., $M(PS_i) = 0$ resp. $M(PS_i) = C_i$ (C_i represents the capacity of the station S_i).

Table 4 : System performance with the best vehicle route.

Station	Empty time	Saturation time	Critical time	QoS	Availability rate
S_1	0	19.44	19.44	80.56	20/26
S_2	0	12.71	12.71	87.29	15/22
S_3	4.96	1.53	6.49	93.51	10/24
S_4	0	0	0	100	20/40
S_5	2.29	4.32	6.61	93.39	10/18
S_6	1.99	0.89	2.88	97.12	10/24
S_7	12.29	0	12.29	87.71	5/20
S_8	11.20	0	11.2	88.8	8/34
S_9	4.12	0.35	4.47	95.53	8/22
S_{10}	0	9.55	9.55	90.45	14/20
Average	3.68	4.87	8.56	91.4	-

Algorithm convergence and results analysis

Table 4 provides the system performances when the total distance is favored (ω_1=0; (ω_2=1). Different metrics of performances are given for each station such as empty times, full times, critical times, availability rate of bicycles and finally the *QoS*. The terms that comprise the objective (i.e., critical times and total distance) are assessed separately in table 5. However, the provided solution is not necessary optimal due to the used approach and the NP-hardness of the studied problem, but it can be a good initiative to optimize the objective in such stochastic systems.

To test the impact of the weights associated to the objective function on the quality of solutions, we test some variation of ω_1 and ω_2 by favoring one over the other, then we evaluate separately the cost of each term of the objective function (i.e. the total critical times and the total distance). The grounds that the quantity of the first term of the objective function is very smaller than the second term, the weighting of this last must be less than the first term except for cases where the *QoS* is not favored. For the considered tests, we keep the same input data. The different configurations and results are given in Table 5.

Table 5 : Evaluation of the weightings impact on the bi-objective function.

Approach	Weights ω_1	ω_2	Critical times (% of total service time)	Total distance (vehicle)	Full Objective
PN-GA	1	0	**3.21**	5410	3.21
PN-GA	1500	1	5.49	4970	13205
PN-GA	800	1	4.91	4320	8248
PN-GA	100	10	8.56	**3300**	3300
PN-GA	0	1	8.56	**3300**	3300
Dijkstra - PN	-		8.56	**3300**	-

One can observe that the variation of weights can affect the objective, as we can observe it particularly when one of the weighting is set to zero. The weights impact definitely on the decisions related to the vehicles routes and the value of the objective function. Since, setting of weights depend on pricing policies that are linked to profits of each objective.

However, adjusting these settings manually is invalid. Hence, setting the weights becomes an auxiliary *NP*-hard problem. This can be separately solved by finding a Pareto optimal points for the corresponding bi-objective function (Mostaghim and Teich 2005), which allow us to obtain a meaningful solution that meets the objectives. When solving this problem, the parameters related directly to the objective function must be known, such as the pricing rental of bicycles and the cost of route (for example 2€ for 1 Km).

In this work, we limit ourselves to the presentation of the modeling-optimization approach and we certainly intend to solve the Pareto optimal in future works.

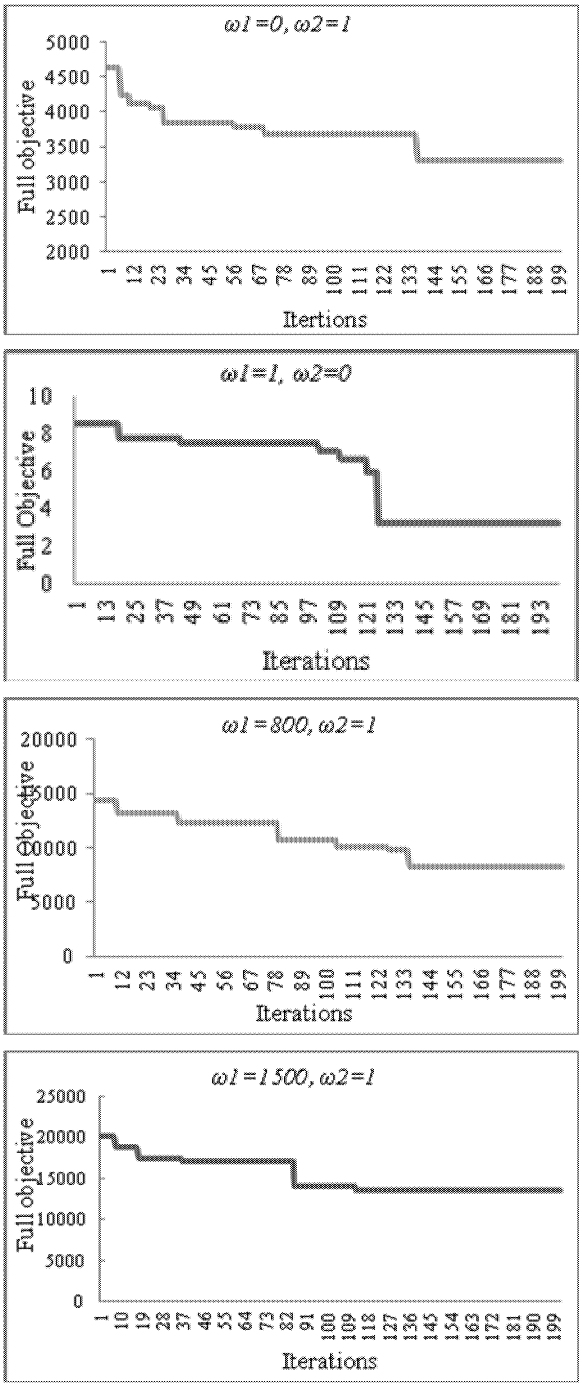


Figure 4: convergence of the optimization Algorithm

As shown in figure 4, we observe the evolution of the best solution for the different configurations during the iteration number and therefore improving the objective. In our approach, we consider both the users satisfaction measured by the critical times and the operational cost measured by the total traveled distance by a rebalancing vehicle, and considering the stochastic demand of users.

CONCLUSION

Bicycle sharing systems have emerged around the world as a viable urban mobility alternative, being already widely spread. These systems have been quickly evolving in the last decades and currently they are integrated with other existing transportation modes in many cities. This study investigates the vehicles routing problem in bicycle sharing systems considering the dynamic case. We define and formulate a mathematical model taking into account the time required to perform the operation with the aim to minimize both the critical situations through the networks stations as well as the operational cost for the service providers.
Motivated by the randomness of BSS and the NP-hardness of the studied problem, we developed an integrated approach based on coupling between a genetic algorithm and BSS-stochastic Petri Net model in order to find the best possible routes for the redistribution vehicles that minimize the objective. The developed approach integrates the users satisfaction, the quality of service, as well as the operational cost. Our approach was tested for a real problem. According to the best of our knowledge, except our previous work, this approach is the first in literature based on coupling genetic algorithm with Petri nets model for solving the vehicles routing problem in BSS, such an approach, should be ranked among the predictive approaches investigating the operational problems in BSS. Finaly, this work addresses both the design and the management of bicycle sharing systems considering the stochastic behaviour of such systems.

In future works, we aim to present some mathematical formulations inverstigating other operational problems in bicycle sharing systems and thereby the developpement of efficient metaheuristic methods and/or exact algorithms to solve problems of large sizes.

REFERENCES

Benchimol, M.; Benchimol, P.; Chappert, B.; Taille, A.D.L.; Laroche, F.; Meunier, F. and Robinet, L. 2011. “Balancing the stations of a self-service "Bike Hire" system”. *RAIRO Operations Research*, No.45, 37–61.

Chemla, D.; Meuniera, F. and Wolfler Calvo, R. 2012. “Bike sharing systems: Solving the static rebalancing problem”. *Discrete Optimization*, No.10 (2), 120–146.

Contardo, C., Morency, C., Rousseau, L. 2012. “Balancing a Dynamic Public Bike-Sharing System”.Technical report TR-12-09. CIRRELET, Montral, Cannada.

DeMaio, P. 2009. Bike-sharing: History, impacts, models of provision and future, *Journal of Public Transportation*, No.14 (4), 41-56.

Di Febrarro, A.; Giglio, D. and Sacco, N. 2004. “Urban traffic control structure based on hybrid Petri nets”. *IEEE Transactions on Intelligent Transportation systems*, No.5 (4), 224-237.

Forma, I.; Raviv, T. and Tzur, M. 2010. “The Static Repositioning Problem in a Bike-Sharing System”. *Proceedings of TRISTAN VII, 7th Triennial Symposium on Transportation Analysis* (Tromsø, Norway, Jun. 20-25). 279- 282.

Holland, J. H. 1975. “Adaptation in Natural and Artificial Systems”. *University of Michigan Press*.

Kadri, A. A.; Labadi, K. and Kacem, I. 2013. “An Integrated Stochastic Petri Net and GA Approach for rebalancing of self service bicycle-sharing systems”. *43rd International Conference Computers & Industrial Engineering* (Hong Kong, Oct. 16-18) No.2, 1304-1311.

Kadri, A. A.; Labadi, K. and Kacem, I. 2014. “An integrated Petri net and GA-based approach for performance optimization of bicycle sharing systems”. *Accepted to appear in the European Journal of Industrial Engineering.* No. xx, xx-xx.

Kadri, A. A.; Kacem, I. and Labadi, K. 2014. “Metaheuristic methods for solving the static balancing problem in public bicycle sharing systems”. *International Conference on Computational Intelligence for Modeling, Control and Automation - CIMCA'2014* (Online conference, Dec 10-12). ISBN-10:0646935941 - ISBN-13:978-0-646-93594-2

Kadri, A. A.; Kacem, I. and Labadi, K. 2015. “A branch and bound algorithm for solving the static balancing in bicycle sharing systems”. *Working paper.*

Labadi, K.; Benarbia, T. and Darcherif, M. 2010. “Sur la régulation des systèmes de vélos en libre-service : approche basée sur les réseaux de Petri . *the 8th ENIM IFAC International Conference of Modelling and Simulation, MOSIM'10*, (Hammamet, Tunisie, Mai 10-12).

Labadi, K.; Benarbia, T.; Hamaci, H. and Darcherif, M. 2012. “Petri Nets Models for Analysis and Control of Public Bike-Sharing Systems . In *Petri Nets - Manufacturing and Computer Science*. Pawel Pawlewski (Eds.), 21-28.

Mostaghim, S. and Teich, J. 2005. “A data structure for storing pareto sets in multiobjective evolutionary algorithms with elitism . In *Evolutionary Multiobjective Optimization*. Springer (Eds.), 81–104.

Nodhi, A.N.; Divrazm,F.R.; golkhayli, T.; Falahnzhad, H. and Rezai, S,M. 2012. “A hybrid algorithm Scheduling flexible manufacturing system consideration cost maintenance . *Recent Researches in Information Science and Applications*, 19-23.

Raviv,T.; Tzur M. and Forma, I. 2011. “The static repositioning problem in a bike-sharing system . *Working paper*.

Raviv T.; Tzur M. and Forma I. 2013. “Static Repositioning in a Bike-Sharing System: Models and Solution Approaches . *EURO Journal on Transportation and Logistics*, No.2 (3), 187-229.

Sadrieh S. A.; Ghaeli M.; Parisa Bahri A. and Peter L. Lee. 2007. “An integrated Petri net and GA based approach for scheduling of hybrid plants . *Computer in Industry*. No. 8(6), 519-530.

Sauer, N. and Xie, X. 1993. “Marking Optimization of Stochastic Timed Event Graphs . *Application and Theory of Petri Nets.* 357-376.

Sayarshad, H.; Tavassoli, S. and Zhao, F. 2012. “A multi-periodic optimization formulation for bike planning and bike utilization”. *Applied Mathematical Modeling*, No.36 (10), 4944-4951.

Sharda, B. and Banerjee, A. 2013. “Robust manufacturing system design using multi objective genetic algorithms, Petri nets and Bayesian uncertainty representation”. *Journal of Manufacturing Systems*, No.32 (2), 315–324.

Shu, J.; Chou, M.; Liu, Q.; Teo, C.-P. and Wang, I-L. 2010. “Bike-sharing system : deployment, utilization and the value of re distribution”.working paper, http://www.bschool.nus.edu.sg/Staff/bizteocp/BS2010.pdf. Accessed Feb. 15, 2012.

Takagi, R.; Goodman, C,J. and Roberts, C. 2003. “Modelling passenger flows at a transport interchange using Petri nets”. *Journal of Rail and Rapid Transit*, No. 217(2), 125-134.

Vogel, P.; Greisera, T. and Mattfelda, DC. 2011. “Understanding Bike-Sharing Systems using Data Mining: Exploring Activity Patterns”. *14th EWGT & 26th MEC & 1st RH, Procedia Social and Behavioral Sciences*, No. 20, 514–523.

BIOGRAPHIES

AHMED ABDELMOUMENE KADRI received his Engineer degree in Industrial IT from the Université d'Oran (Es Senia), Algeria in 2010 and his Master’s degree in Computer sciences from the Université de Lorraine, France in 2012. He is currently working toward his PhD in computer engineering at the "Laboratoire LCOMS" of the "Université de Lorraine" (Metz, France) and the Graduate School of Electrical Engineering and Industrial Management (ECAM-EPMI), Cergy-Pontoise, France. His research interests include operational research, combinatorial optimization, modeling, performance evaluation and optimization of stochastic systems such in logistic and transportation field.

KARIM LABADI is currently an Assistant Professor at the School of Electricity, Production and Industrial Management (EPMI) and member of the ‘Equipe Commande des Systèmes’ (ECS-Lab, EA.3649) at the ‘Ecole Nationale Supérieure d'Electronique et ses Applications’ (ENSEA), Cergy-Pontoise France, since 2006. He obtained his M.S degree in Applied Automation and Informatics from the IRCCyN, ‘Ecole Centrale de Nantes (ECN)’, France in 2002, and PhD in Industrial Engineering and Optimization from the University of Technology of Troyes (UTT), France in 2005. His research interests include Petri nets and their applications to modelling, control and performance optimization of stochastic discrete event systems such as production, transportation and logistic systems.

IMED KACEM is a Full Professor at the University of Lorraine (France) where he created the LCOMS laboratory in 2011. He is the Head of this multidisciplinary laboratory (*http://lcoms.univ-lorraine.fr/*). He is a nominated member at the National Council of the Universities (Computer Science Section CNU27, France), Associate or Guest Editor for several referred journals (such as EJIE, JSSSE (Springer), Computers & Industrial Engineering (Elsevier) …) and organizer of major international conferences (CoDIT14, IEEE/CIE39, IEEE/ICSSSM06, WAC/ISIAC06). His scientific activity is in a transversal and interdisciplinary domain: the Operational Research. More precisely, his contributions are related to the design of exact and approximate algorithms with a guaranteed performance for the NP-hard combinatorial problems. He obtained the « 2010 Great Award of Research » from the Universities of Lorraine and the « 3rd Robert Faure Award 2009 » from the French Society of Operational Research and he has regularly the PEDR or the PES Premium (level A) since 2006.

ADAPTIVE MOBILE AD-HOC NETWORK (AMAN): A QOS FRAMEWORK FOR MOBILE AD-HOC NETWORKS

Benny M Nyambo
Computer Science Department,
Faculty of Science
University of Zimbabwe
Harare, Zimbabwe
nyambo@science.uz.ac.zw

Gerrit K Janssens
Research Group
Logistics
Hasselt University
Diepenbeek, Belgium
gerrit.janssens@uhasselt.be

Wim Lamotte
Expertise Center for Digital
Media
Hasselt University – tUL - iMinds
Diepenbeek, Belgium
wim.lamotte@uhasselt.be

KEYWORDS
MANET, Quality of Service, Multimedia, Adaptation

ABSTRACT
Quality of Service (QoS) support for Mobile Ad-hoc NETworks (MANETs) is a challenging task due to the dynamic topology and limited resources in MANETs, and the QoS model should be the first matter to consider as a system goal. The INSIGNIA framework, ASAP and DiffServ model can provide system-level QoS support for MANETs, but each have pros and cons in service precision and scalability. In this paper, we propose a hybrid QoS model for MANETs, called AMAN, which combines the per-flow granularity of INSIGNIA and ASAP and the per-class granularity of DiffServ, to provide a responsive, scalable, and flexible QoS support for MANETs. The simulation results show that AMAN can achieve effective service differentiation and offer the best QoS to the per-flow service under various mobility conditions.

INTRODUCTION

Mobile ad-hoc networks represent future generation wireless networks, with a high degree of versatility and robustness, capable of being deployed quickly and economically at places lacking any infrastructure. The characteristics of these networks such as bandwidth scarcity and fluctuations, node mobility, hostile working conditions and battery power constraints, have hindered their development. Projected applications of such networks include defense-based, disaster relief operations, commercial applications and sensor networks. Many of these applications require a distinction in the quality of connections being supported in terms of bandwidth availability, end-to-end delay and jitter. As multimedia traffic finds its way into wireless networks, the use of the UDP transport layer protocol alone does not suffice to accomodate the needs of users.

Most multimedia traffic require a more stable throughput for them to be regarded useful. Therefore there is a dire need for a QoS model integrated within the nodes for such applications. The model must be able to distinguish flows based on their QoS needs and have mechanisms that work to meet those requirements. Since there is no central coordinator, the QoS model must operate in a fully distributed manner. The QoS framework should allow layers of the OSI model to interact and cooperate in providing a comprehensive quality of service to the traffic

We propose a QoS model that differentiates traffic into classes and attempts to provide bandwidth and delay guarantees to flows of highest priority class. The framework relies on its strong adaptation algorithm to give more access to resources for the highest priority class of traffic whilst the least priority traffic cedes its share of bandwidth allocation to the highest priority until the time resources are available.

RELATED WORK

The SWAN (Ahn, 2002) model differentiates traffic into real-time UDP traffic and best effort TCP traffic. It is a stateless and fully distributed model that provides soft QoS assurances to real-time traffic. It uses admission control for real-time traffic, rate control of TCP traffic and ECN congestion control mechanisms to ensure that real-time packets meet QoS bounds. Each node comprises an admission controller that maintains information about the status of the outgoing link in terms of the available bandwidth and amount of congestion. It does this by promiscuously listening to all packet transmissions within its range. The source node sends a probe message to the destination to find the bottleneck bandwidth along the path. If this value is greater than the requirements plus a threshold value, the flow is admitted; otherwise it is rejected and marked as best-effort. All TCP flows are considered as best-effort.

INSIGNIA (Lee, 98), (Lee 2000) is an in-band signaling system that supports adaptive reservation-based services in ad-hoc networks. All the control information is carried in-band with the data and encoded using the IP option field in datagrams. The signaling system supports a number of protocol commands that drive fast reservation, fast restoration and an end-to-end adaptation mechanism. This in-band information is snooped as data packets traverse intermediate nodes/routers and used to maintain soft-state reservations in support of flows/micro-flows. To establish reservation-based flows between source-destination pairs, source nodes initiate fast reservations by setting the appropriate fields in the INSIGNIA IP option field before forwarding packets. Reservation packets (i.e. data packet with the appropriate IP option set) traverse intermediate nodes, executing admission control modules, allocating resources and establishing soft-state reservation at all intermediate nodes between source-destination pairs. The reservations need to be periodically refreshed by the packets of the flows. In the event of a change in the path resulting from movement of the nodes, the first packet along the new path makes fresh reservations along this path thereby performing a fast restoration. Reservations made along the old path are removed on a timeout. Flows in the network are expected to be adaptive to bandwidth availability. A flow that was allocated a MAX amount of bandwidth initially could be downgraded to a MIN amount or even to best-effort in the event of rerouting of a flow or if network conditions change. ASAP (Xue, 2003) is an adaptive reservation QoS protocol. When a flow requires service, it is soft-reserved first and later hard reserved. When the route has been established the resources are then hard-reserved to the flow. ASAP employs an adaptation method that will allow some flows to release bandwidth if the allocated bandwidth is more than their required minimum. It also has a route repair mechanism that allows self-healing by finding alternative routes when a route is broken, which happens frequently in wireless networks. ASAP adaptation helps to stabilise the throughput of flows in wireless ad-hoc networks.

Ramraj(2010) presented a bandwidth management protocol to demonstrate fair bandwidth sharing between nodes. One node is identified as the bandwidth manager and the bandwidth management protocol is invoked at flow establishment, flow teardown and when bandwidth adaptation is required. When the bandwidth manager node is inaccessible, a new manager node is elected. In our proposed approach we differ with Ramraj (2010) by performing the bandwidth management on all nodes so that we remove the problem of single point of failure. Pattanayak (2009) proposed a distributed cluster scheme for MANETs, in harsh environments, based on the concept of survivability to support QoS requirements and to protect bandwidth efficiently. With the incorporation of clustering algorithms in survivability technology, they employ a simple network configuration and expect to reduce occurrences of faults in MANETs. In (Bushehrian, 2014), the author proposed a reliability aware scheme in which clients rank service providers on their reliabiltity and response time. They use a Hidden Markov Model for the service selection and choose the best service providers whilst avoiding poor providers; this managed to lower violations of the client service level agreement. We need a framework that optimizes assignment of bandwidth to flows according to the type and priorirty of traffic. The types of traffic have to be multiple unlike in some of the other frameworks which classifiy traffic into two classes only.

PROPOSED QoS FRAMEWORK

This section describes the proposed framework architecture, AMAN whose proposed architecture is illustrated in figure 1. The proposed QoS architecture has five basic modules namely Bandwidth Estimation, Bandwidth Adaptation, Congestion Control, Admission Control and Reservation.
The implementation of the proposed traffic management scheme is supposed to fulfill these requirements:

i. Admission of a new flow into the network if there is enough bandwidth to carry the flow without interfering with other ongoing traffic.
ii. Increase in the available bandwidth by reducing the allocation of other ongoing applications in order to incorporate a new flow.
iii. Deny of a new flow, whose requested bandwidth is greater than the available bandwidth after Bandwidth Adaptation has been done.

Traffic Differentiation

For traffic differentiation we adopt the DiffServ model, (Nichols, 1998), for classification and marking packets. DiffServ provides QoS by dividing traffic into a number of classes and allocating network resources on per class basis. The class is marked directly on the packet in the 6 bit DiffServ Code Point (DSCP) field, which is part of the original type of service (ToS) field in the IPV4 header. The DiffServ field is split into the 6-bit DSCP field and a 2-bit field which is used for Explicit Congestion Notification (ECN) mechanisms.

We propose assigning priorities to different types of traffic, which are going to affect the treatment of traffic at nodes. We adopt the CISCO QoS Baseline as our classification model (Szigeti, 2013), see Table 1.

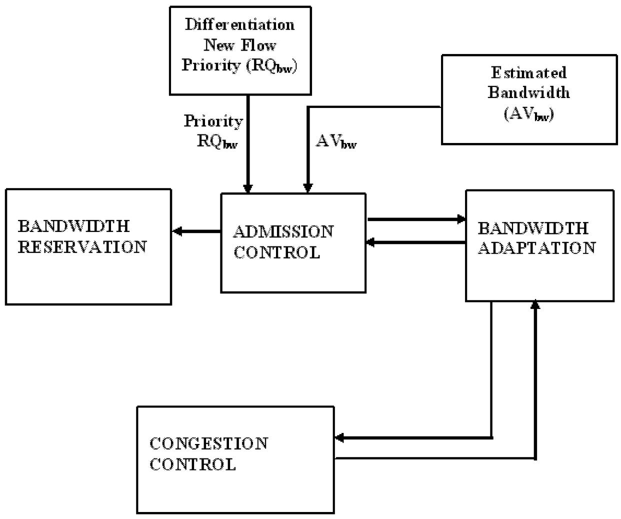


Figure 1: Traffic Management Framework Architecture

Bits	7	6	5	4	3	2	1	0
DiffServ Field	DSCP	DSCP	DSCP	DSCP	DSCP	DSCP	ECN	ECN
Function	Class Selector			Drop Probability		0	ECN	

Figure 2: The functions of Differentiated Services (DS) bits in class creation

Admission Control

A new call for transmission will need to be admitted at each and every node in a path from source and a destination of the call. This can only happen if the requested bandwidth of the new flow is smaller than the available bandwidth. Figure 3 below demonstrates what happens in the Admission Control Module. When a call for transmission is made by a new flow, its requested bandwidth (MinBw) is compared with the available bandwidth (AV_{bw}). If the requested bandwidth is smaller than the available bandwidth then the flow is automatically admitted, otherwise the Bandwidth Adaptation Module is invoked to try to release extra bandwidth being used by other flows so that enough resources are freed to allow the new flow. The Adaptation module returns a modified (AV_{bw}) available bandwidth value which is a sum of the old AV_{bw} value together with the released bandwidth (RL_{bw}) values. If the new value of available bandwidth is greater than the minimum requirement of the new flow, then it will be admitted, else the flow will be denied.

Table 1: Cisco QoS Baseline Classification, Marking, and Mapping Recommendations

Application	Binary DSCP	CLASS OF SERVICE
IP Routing	110000	6
Voice	101110	5
Interactive Video	100010	4
Streaming Video	100000	4
Locally Defined Mission Critical Data	011010	3
Call Signalling	011000	3
Transactional Data	010010	2
Network Management	010000	2
Bulk Data	001010	1
Scavenger	001000	1
Best Effort	000000	0

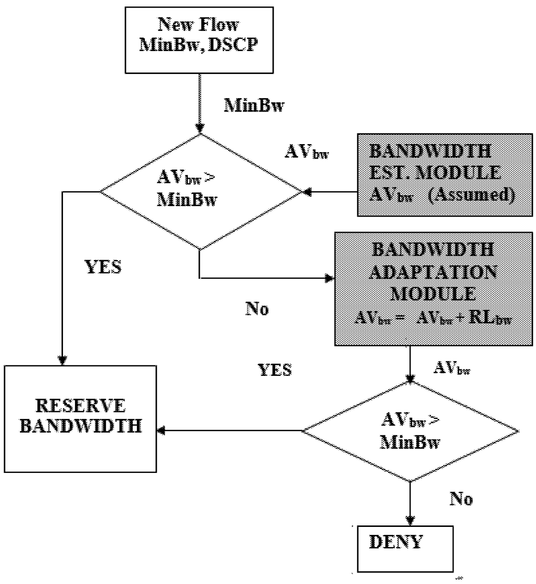


Figure 3: The function of the Admission Control Module

QoS Signalling Method

AMAN is an in-band signalling system it uses the packet's header to carry all its control information. In case of IPv6 this information can be transmitted within the base header and/or within any extension headers. AMAN uses the eight bits of the CLASS field to transmit its Message Type indicator and congestion notification. The IPv6 Hop-by-Hop options extension header will carry the request for reservation (RES), the minimum and maximum required bandwidth (MinBw and MaxBw) and the bandwidth reserved by a node for the specific flow (ActualBw). The QoS option has four fields, the reservation indicator (RES), Minimum and Maximum bandwidth fields (MinBw and MaxBw) and the Actual allocated bandwidth field (ActualBw).

The RES bit is set to 1 if the source is seeking a reservation for a new flow. At every intermediate node, there is an admission control procedure to ascertain whether the resources available can sustain a new flow without hindering ongoing communications. When the available bandwidth is not enough, then an intermediate node should not admit a new flow, otherwise it will interfere with on-going flows whilst it does not meet its own minimum requirements.

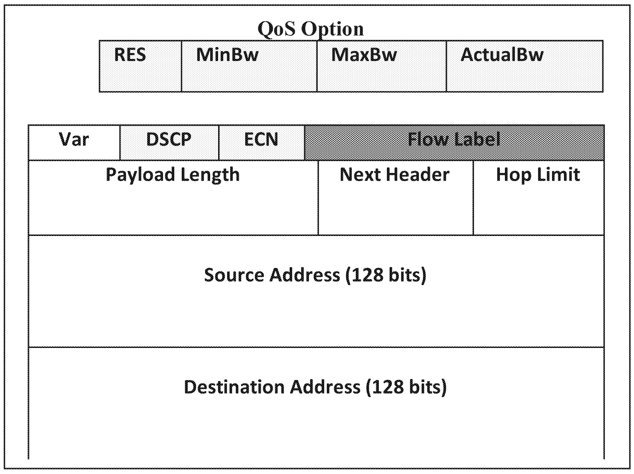


Figure 4: Signalling messages embedded in the IPv6 Header

When the source node initiates a new flow, it fills all the fields with data and forwards the packet to the next node. Every intermediate node checks the MinBw and ActualBw and compares them with its AvailBw. If there are enough resources (that is AvailBw > MinBw), and if the ActualBw from the previous node is greater than AvailBw, the ActualBw field is updated to AvailBw otherwise the field is not changed. Every node in the network keep a QoS table to store the flow label, source address, DSCP and bandwidth allocated to the flow. Various values of ActualBw are reserved at various nodes along the path. This means that, for some time, there would be over-reservation of resources at some nodes. This is corrected by unicasting a RES report (QR) packet from destination node to the source node showing the bottleneck bandwidth for the flow. Figure 5 shows a source S sending data to destination D along a certain path selected by the routing protocol. The bottleneck for the path is 150 Mbps at node X. Node D sends a RES report (QR 150) packet to the source node showing the bottleneck bandwidth. The RES report message does not have to follow the same path as the reservation message since this may not be possible in MANETs. The source will then change the ActualBw field to 150 MBps and change the RES field to 0. All the nodes receiving the packets will now allocate this new value and release extra bandwidth they had allocated.

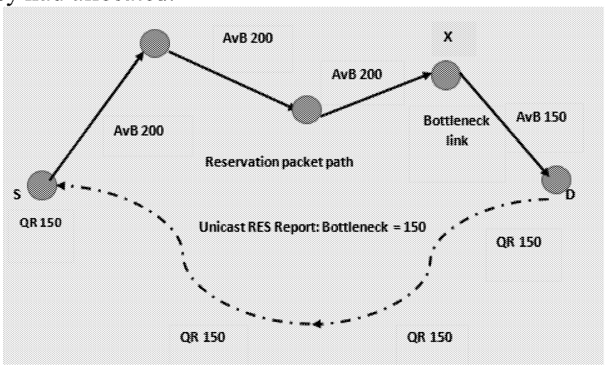


Figure 5: Fast reservation showing reservation request and reservation report for bottleneck bandwidth.

Re-routing (Fast restoration).

Reservation-based flows are often rerouted within the lifetime of ongoing sessions due to node mobility, as illustrated in Figure 6. A node moves out of radio contact and the flow is re-routed through mobile node *Y*. The minimum reservation is immediately restored along the new path, while reservations along the old path are timed out and automatically removed and there is no change along the *common path.*

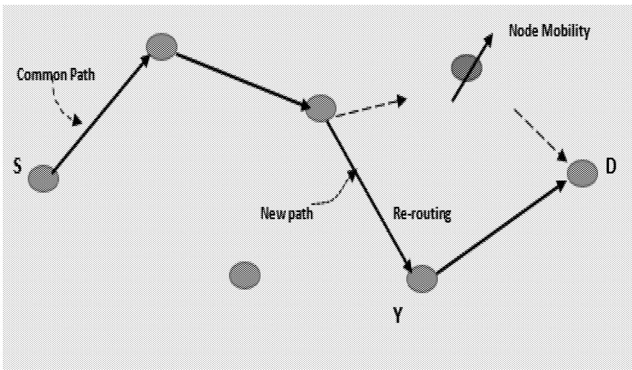


Figure 6: Fast restoration showing how an intermediate node help find new routes when a link is broken by mobility.

Rerouting active flows involves the MANET routing protocol (to determine new routes), admission control, and resource reservation for nodes along the "new path." Soft-state timers are continually refreshed and reservations maintained as long as packets associated with a particular flow are periodically received at intermediate routing nodes between source-destination pairs.

Bandwidth Adaptation

Adaptation ensures that a new flow can be admitted by accepting bandwidth released by other applications already in the network. Figure 7 illustrates the bandwidth adaptation algorithm. When the request for adaptation arrives, the priorities of ongoing applications are checked to get the flow with the least priority. The adaptation module must degrade the least priority flows to their minimum required bandwidth in order to free resources for the new flow. The module must be able to calculate the amount of bandwidth to be released by low priority flows before degrading them. If the highest priority flow asks for resources when the remaining resources are not enough, then we have to pause some of the least priority traffic to release resources.

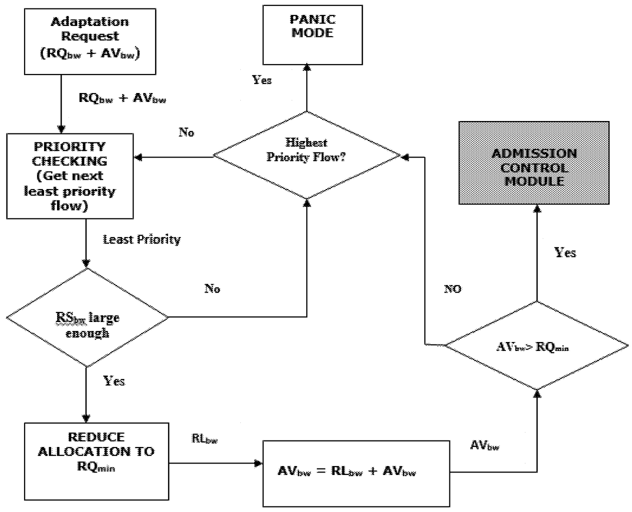


Figure 7**:** Bandwidth Adaptation Algorithm

In case of network congestion notification, we can adapt the ongoing flows to reduce the allocations until the channel is no longer congested. The destination node informs the source of this congestion through a QoS report and the source scales down the transmition rate of flows to MinBw starting with the least priority. If congestion persists, then flows will be paused starting with the least priority flow. More flows will be stopped until the system has been restored to a non-congested state. We call this the panic mode. After a random back-off time a source with a throttled or paused flow can attempt to increase or re-admit the traffic flow. Node mobility or session dynamics may cause a flow routed via Y to be scaled up from minimum to maximum required service.

PERFORMANCE MODELLING OF AMAN

AMAN was evaluated through simulation. The system is written in C++ code that consists of several functions which will be called to perform different tasks including computing the available bandwidth. The code consists of a data structure called flows with five attributes namely:

- *Flow Priority* –Every flow is given a randomised priority value to resemble different real-time traffic, with the flow with the lowest value having the highest priority.

- *Flow ID/Address* - Each flow that enters the network has a Flow ID to uniquely identify all the flows in the network and also to keep track of the flows which were reduced to required minimum during bandwidth adaptation.
- *Reserved Bandwidth (RBW)* –When the flow is admitted into the network it is given its required bandwidth and this becomes the Reserved Bandwidth (RBW).
- *Required Minimum Bandwidth (RQmin)* – This is the minimum required bandwidth requirements of a flow in the network.
- *Releasable Bandwidth (RLBW)* – This is the difference between the reserved bandwidth and the required minimum. The bandwidth that can be released for other new flows to be admitted.

All these values will be randomised for the sake of the experimental studies. This random behaviour tries to mimic the random nature of traffic arriving and leaving the node. An array of up to 100 flows is defined and this will be used for different flows in the simulations. The Channel Capacity is assumed to be 2Mbps. A new flow will go through admission control, reservation and sometimes it might have to go through bandwidth adaptation first if the available bandwidth is not sufficient. As the simulations are run the results are written to output files where data was extracted for analysis. A flow can be admitted without need for adaptation, admitted after some adaptation has been done or denied admission even when adaptation has been done.

RESULTS AND FINDINGS

Data from the simulations help us to deduce various things among of which is to calculate the success rate of bandwidth adaptation in the resource reservation scheme. The success rate of the Bandwidth Adaptation process gives us a rough probability that when a new flow is initiated in a loaded network, the flow is admitted. We compared our results from adaptation simulations with a control experiment without adaptation for the purposes of comparison. We estimate the admission success rate of a scheme as:

$$S = \frac{No.of\ flows\ admitted}{Total\ No.of\ flows}.100\ \%$$

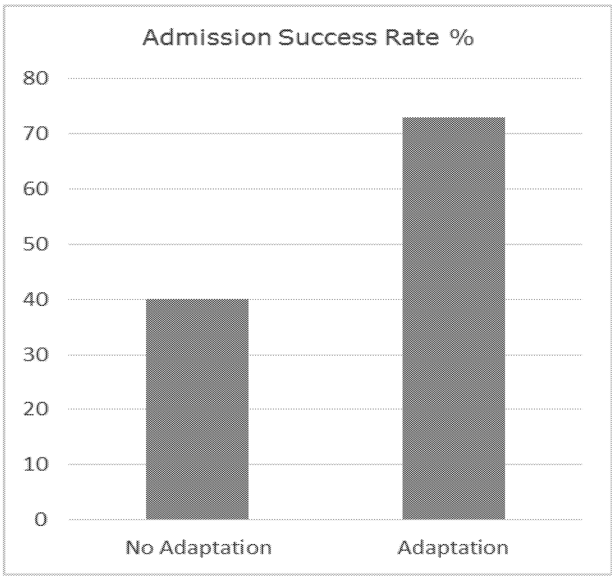


Figure 8: Comparison of the success rate of admitting a new flow for a network that does not do bandwidth adaptation to one that does adaptation.

The success rate of the adaptation process is determined by the number of flows admitted after adaptation has taken place, and was found to be 73 %. The success rate of a new flow to be admitted without adaptation was found to be 40% as shown in figure 8. This shows primarily that adaptation improves the chance of admitting a new flow by more than 30%. The results shows that most simulations with a new flow's bandwidth requirements above a certain threshold value (in our case 0.75Mbps) have to undergo the adaptation process despite the size of the network. In this case network size is represented by the number of flows already in the network when a request is done.

The results also show that of those simulations, which required adaptation (22 out of the 30 simulations carried out), most of the simulations with new flow's requirements above 1Mbps resulted in the new flow being denied admission. These simulation results highlight that the Adaptation process increases the chances of admission of flows that could have been denied by a factor of 0.73 or success rate of 73% as shown by the estimates of the Success rate of a scheme with bandwidth adaptation.

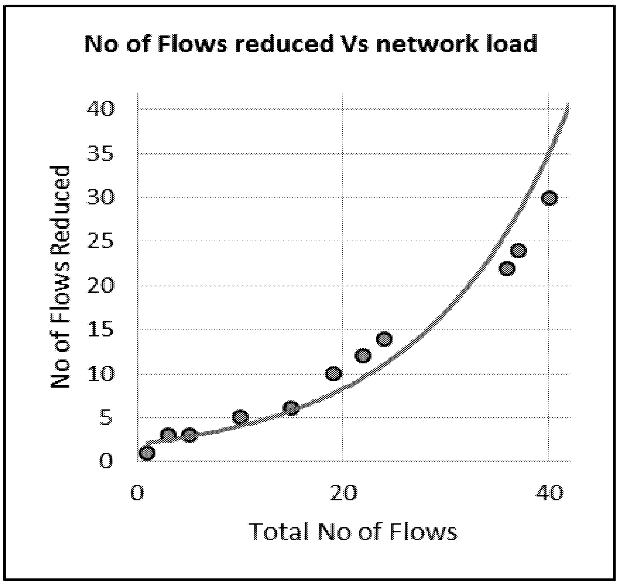


Figure 9: The variation of number of flows whose bandwidth is reduced to their required minimum during adaptation as the load in the network increase.

The results in Figure 9 show that the network size is exponentially proportional to the number of flows that have their reservation reduced to their minimum bandwidth requirement. The general trend highlights that as the network size increases the number of flows that will be reduced to working minimum also increases exponentially and this might have a negative effect on the quality of service of the flows reduced. However the performance will remain within acceptable range since the adaptation does not go below the minimum required bandwidth. Under congestion, in case of rerouted flows, adaptation and flow pausing might impact heavily on the network performance since some applications will be forced to stop operation.

Figure 10 shows the relationship between the number of flows in a network and the consumed bandwidth. For the network that does not do adaptation, a small number of nodes will consume a lot of bandwidth. This is attributed to the fact that flows will continue consuming bandwidth close to their maximum requirements, taking no regard to the congestion happening in the network. With adaptation, the network can accept more flows at the same time satisfying their minimum bandwidth requirements.

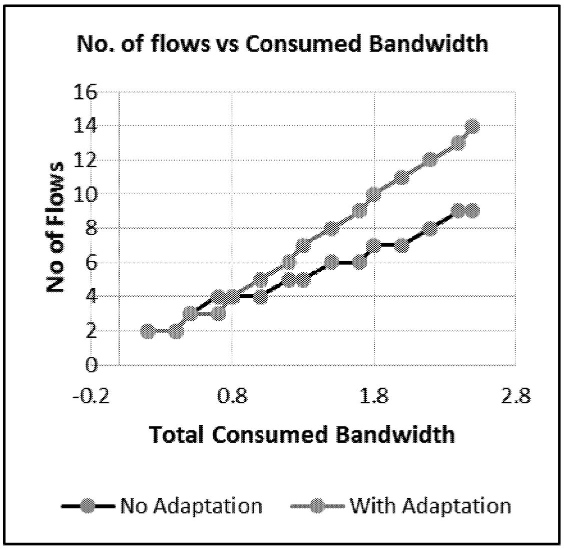


Figure 10: The relationship of the number of flows that the network can sustain, to the amount of consumed bandwidth.

CONCLUSIONS

In this research we presented a scheme for traffic differentiation and management in Mobile Ad-hoc Networks (MANETs). The proposed scheme is intended to efficiently manage the reservation of bandwidth in MANETs based upon the available bandwidth within the network. The scheme also employs a Bandwidth Adaptation process, which is the novel part of the research, to increase the chances of admission of a new flow into the network. We have discussed in detail how the proposed scheme works and how it increases the available bandwidth in the network by reducing the allocations of other flows to required minimum requirements according to the priorities of the flows. The research also includes an empirical analysis of the behaviour of the scheme by conducting simulations under varying conditions of required bandwidth and priority of new flow, available bandwidth and the size of the network. The results presented for the proposed scheme demonstrates that Bandwidth Adaptation enhances the admission of flows with bandwidth requirements greater than the available bandwidth. Our scheme has an added attribute of a multi-level priority scheme, which treats traffic of various classes differently and gives an advantage to the highest priority traffic. In other schemes like INSIGNIA, ASAP and SWAN, traffic is differentiated into real-time and best-effort traffic only. This does not respect the idea that multimedia traffic that is flourishing in the internet these days have varying characteristics and varying importance. Our framework also differs with ASAP and INSIGNIA in that it does not allow traffic to be degraded to best effort in case of congestion but stop flows that fall below the minimum acceptable bandwidth. This is done to ensure that no multimedia traffic is allowed to flow in the network unless it can be found useful to the end users.. The scheme has the potential of coming up with efficient and realistic reservations.

REFERENCES

Ahn G., A. T. Campbell, A. Veres, and L. Sun. 2002. "Supporting Service Differentiation for Real-Time and Best-Effort Traffic in Stateless Wireless Ad Hoc Networks (SWAN)". *IEEE Transactionson Mobile Computing, Vol.1, No. 3, (JUL-SEP)*

Bushehrian O, "Model-Based Service Selection For Reliable Service Access In MANET". Malaysian Journal of Computer Science. Vol. 27(4), 2014

Chaudet C., O. Festor, I. Gu´erin Lassous, and R. State, (2004) "A Managed Bandwidth Reservation Protocol for Ad Hoc Networks". *SAPIR 2004, LNCS 3126, Springer-Verlag Berlin Heidelberg* (pp. 13–20)

Lee S, Ahn G, X. Zhang, and A. Campbell (2000) INSIGNIA: An IP Based Quality of Service Framework for Mobile ad Hoc Networks, *Journal of Parallel and Distributed Computing*, 60, (pp 374-406)

Mohapatra P., J. Li, AND C. Gui. 2003 "QoS in mobile ad hoc networks" *IEEE Wireless Communications* (Jun)

Nichols K., S. Blake, F. Baker, D. Black (1998)," Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers (RFC 2474)" Network Working Group

Pattanayak B, K, Jagadev A. Mishra M, K. Nayak M. R, "A Distributed Cluster Scheme for Bandwidth Management in Multi-hop MANETs". International Journal of Computer Science and Network Security, VOL.9 No.10, October 2009

Ramraj R, "Bandwidth Allocation and Routing Information for Wireless Mobile Ad-hoc Networks" Seventh International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2010)

S. Lee , A. Campbell, "INSIGNIA: In-band signaling support for QoS in mobile ad hoc networks", *MoMuC 98, Berlin, Germany,* Oct. 1998

Szigeti T., C. Hattingh, R. Barton, K. Briley (2013) "End-to-End QoS Network Design: Quality of Service for Rich-Media & Cloud Networks (2nd Edition)", Cisco Press (Nov)

Veres A., A. T. Campbell, M. Barry, and L. H. Sun. 2001. "Supporting Service Differentiation in Wireless Packet Networks Using Distributed Control." *IEEE Journal on Selected Areasin Communicaions, Vol. 19,* (OCT)

Xue J., P. Studi and G. Alonso. "ASAP: An Adaptive QoS Protocol for Mobile Ad Hoc Network". *Proc IEEE PIMRC* 2003, Beijing

AUTHOR BIOGRAPHIES.

BENNY M NYAMBO is currently studying towards a PhD in Computer Science at the Hasselt University, Belgium. He has been a lecturer in Computer science at the University of Zimbabwe. He has research interests in performance modelling of communication networks and in embedded systems.

GERRIT K JANSSENS is professor Operations Research Management and Logistics at Hasselt University within the Faculty of Business Economics. His main research interests include the development and application of operations research models in production and distribution logistics.

WIM LAMOTTE is professor Computer Science at Hasselt University, teaching courses in networking and multimedia. He leads the Multimedia Communications research group of the Expertise Centre for Digital Multimedia. He is a member of IEEE and ACM.

ENGINEERING APPLICATIONS

FEM AS A TOOL FOR CONCURRENT CIVIL ENGINEERING

Ion Manea
Flower Power USA Inc.
902 4th St., SW Suite B, Auburn
WA, 98001, U.S.A.
manion@seanet.com

Ștefan Sorohan, Daniela Vasiliu, Nicolae Vasiliu, Florin Rădoi
Power Engineering Faculty
University POLITEHNICA of Bucharest
RO 060042, Bucharest, Romania
vasiliu@fluid-power.pub.ro

KEYWORDS

Composite materials, civil buildings, water storage tanks, FEM analysis.

ABSTRACT

The strong penetration of the composite materials in different fields as aerospace and automotive industries, civil engineering, and many other domains obliged the designers to assess the real static and dynamic behaviour of different types of mechanical components, buildings, bridges, petrochemical tanks etc. Obviously, the results of any sophisticated computations are far removed from the experimentally ones. The main source of such a discrepancy comes from the uncertain hypothesis on the co-operation between the reinforcement (shell), which is strong but brittle, and the matrix (or core), which is much weaker, but much lighter and flexible. A wide diversity of such materials is now spreading in the light buildings field, due to many qualities: small overall specific weight, good overall strength for flexure, very small specific heat, moderate price, and a lot of quick assembling possibilities by adhesives or fasteners.

INTRODUCTION

The aerospace and automotive extensive use of different types of composite materials generated a special interest for many types of applications in different unexpected field like civil engineering. The high degree of thermal insulation, the very low specific weight, the capability of mixing different components for obtaining different mechanical, thermal, and even aesthetic advantages led the engineering community to extensive studies of these modern materials. Now, composites are used in many industries today to enable high-performance products at economic advantage. These industries range from space to sports and include manufactured products for aircraft, transportation, energy, construction, sports, marine, and medical use. There are many materials, economic and to using composites, but a solid knowledge of the physical properties, including the mechanics, tooling, design, inspection & repair, and manufacturing options is required for working in this medium as they are intrinsically linked. (edX Inc 2015). This paper describes an attempt to explore the possibility of using composite materials for some important civil engineering objectives like water storage aerospace tanks for different purposes as in isolated residential houses (fig.1 and fig. 2). One of the most important properties of the double shielded materials in this case is the good resistance of the humidity and microorganisms, commonly found in water technology from any application. The practical computation was performed for a 6 cell water tank made of composite panels, in order to find out the resistance limit of a common composite material produced in U.S.A. (Manea, Sorohan and Vasiliu 2010).

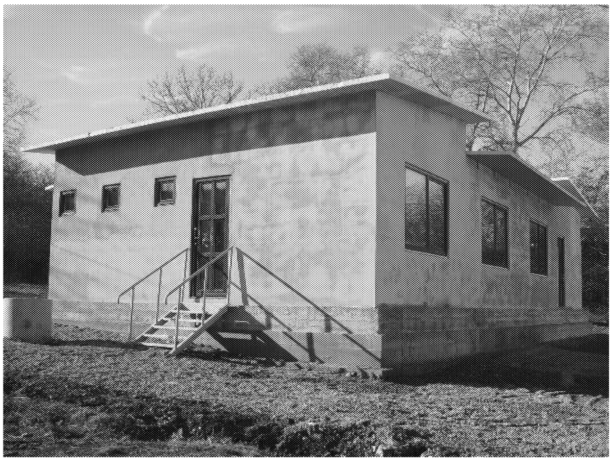


Figure 1: House from composite panels only, designed as kindergarten (Manea, Sorohan and Vasiliu 2010)

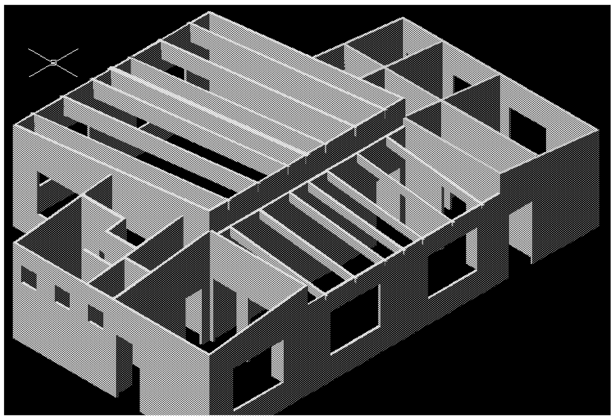


Figure 2: The internal structure of the kindergarten from composite elements only (Manea, Sorohan and Vasiliu 2010)

SANDWICH BEAM COMPUTATION

In theory it can be considered that a composite beam, with weak core made out of polyurethane foam can ensure flexural strength through the hard other shell, while the inner core ensure continuous support for the linings. The core sustains the shear stress and overcomes the buckling of the outer shell, or residual deformation in case of hard solicitation. The adhesive used, must withstand the stretch and shear stress that occurs at the interface between shell and core (Adams 1997). A realistic assessment imposes a good knowledge of the material properties. This information must be obtained experimentally. Even so, for a start, a simplified calculus can be made on the hypothesis that composite materials are similar to the isotropic material. For a sandwich beam, simply supported, and loaded with a uniformly distributed load (fig. 3), for which $d/t > 6$ and $\frac{E_f t d^2}{E_c c^3} > 17$, results the equation of maximum deformation (Jones1998).

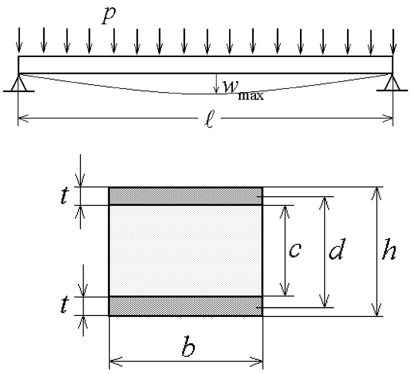


Figure 3: Load diagram for a composite beam with a week core

$$w_{\max} = \frac{5}{384}\frac{p\ell^4}{(EI)_{ech}} + \frac{1}{8}\frac{p\ell^2}{(GA)_{ech}} \tag{1}$$

Here $(EI)_{ech} \cong \frac{bt(c+t)^2}{2} E_f$ and $(GA)_{ech} \cong bcG_c$,

where E_f is the Young's modulus of the shell and G_c is the shear modulus of the core.

Maximum normal stress in the shell and core, can be determined using the equations:

$$\sigma_f^{\max} = \frac{M}{btc}\frac{h}{c}; \qquad \sigma_c^{\max} = \frac{M}{btc}\frac{E_c}{E_f}, \tag{2}$$

The maximum shear stress in the core can be determined using the relationship:

$$\tau_c^{\max} = \frac{T}{bd}. \tag{3}$$

For the following dimensions, b=1 m, c = 0.0814 m, and t = 0.0013 m, the material properties are:

a) shell:

E_f = 30000 MPa; ν_c =0.3 ρ_f = 1600 kg/m^3

b) core:

G_c = 3.5714 MPa; E_c = 10 MPa; (ν_c =0.4) ρ_c = 32 kg/m^3

From equation (1) we can obtain the load for which a maximum deformation of $\ell/300$ (max allowed) is obtained. Furthermore, the maximum stress is obtained from equation (2) and (3). Table 1 presents synthetically the results obtained for various lengths of the beam.

Length	Load "p" for which max. deformation is $\ell/300$	Max stress in the shell	Max. stress in the core	Max. shear stress in the core
ℓ [m]	p [N/m]	$\sigma_f^{\max}$ [MPa]	$\sigma_c^{\max}$ [kPa]	$\tau_c^{\max}$ [kPa]
1.0000	6317	7.7013	2.4876	38.197
1.5000	3420	9.3818	3.0305	31.022
2.0000	**2031**	**9.9043**	**3.1992**	**24.562**
3.0000	849	9.3147	3.0088	15.400
4.0000	418	8.1588	2.6354	10.117

Table 1. Maximum load, and stress for various lengths of the composite beam

If the mechanical properties (E_c and G_c) of the core are reduced by 10 times, then the maximum load that produces the same deformation is reduced by 5.65 times. If the mechanical properties (E_c and G_c) of the core are increased by 10 times, then the maximum load that produces the same deformation is increased by 1.89 times. If the mechanical properties of the shell (E_f) are reduced by 3 times, then the maximum load that produces the same deformation is reduced by 1.95 times. If the mechanical properties of the shell (E_f) are increased by 3 times, then the maximum load that produces the same deformation is increased by 1.46 times. The results in bold from the table above have been obtained using FEM tool (ANSYS 2014). FEM analysis can be an effective and realistic tool, if the mechanical properties of the polyurethane foam are known. For the shell, the estimations are much easier (Caillet et al. 2007, Hodgkinson 2000, Sorohan and Părăuşanu 2005), as the materials used to make it have a much narrow spectrum of mechanical properties than the foam.

Preliminary analysis of a 6 cell water tank made of composite panels

We consider a tank of 1200 m^3 consisting of 6 vats made out of composite material, built on top of the ground. The tank is designed for storage and treatment of waste water, extracted from deep shafts. The purpose of the analysis is to determine maximum deformation and stress when all the vats are full with water. We consider composite panels, with a total 100 mm width, and 1.5 mm shells thickness.

The material is considered isotropic, with below properties:

Material	Young's modulus [MPa]	Poisson ratio [-]	Density [kg/m^3]
Shell	30000	0.3	1600
Core	10	0.4	32

FEM computational model

In Fig. 2 is presented the basic layout of the tank, with its main components: base, honeycomb structure, vats and cover. We consider the middle plane of all the panels, which is modelled using "sandwich" type elements of an average element size of 250 mm. The complete computational model consists 58121 finite elements and 162918 nodes. The computational model considers the structure continuous at joints (glued together). The mathematical model does not include joint details. This model cannot offer information about the crumpling of the shell, and the stress concentrators, even though part of the concentrator's effect is included. The nominal stress in each panel is correctly estimated, especially in areas far from the mathematical joints. The mass of the model is 22986 kg. We consider a load generated by fluid with a density of 1200 kg/m^3, from a height of 5.05 m from the base. The weight of the structure has also been included.

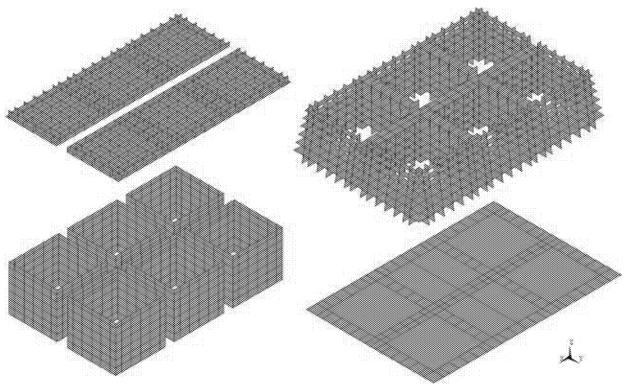


Figure 4: FEM – computational model of the 4 components: base, honeycomb structure, vats and cover

It is considered that the base is perfectly sustained by the concrete foundation (no point of the base leaves the foundation, even though this phenomenon could occur at the edges of the model). The base is anchored by the concrete with screws fixed along the red edges presented in Fig. 5, so the calculus model considers that this lined does not suffer any displacements on X and Y axis.

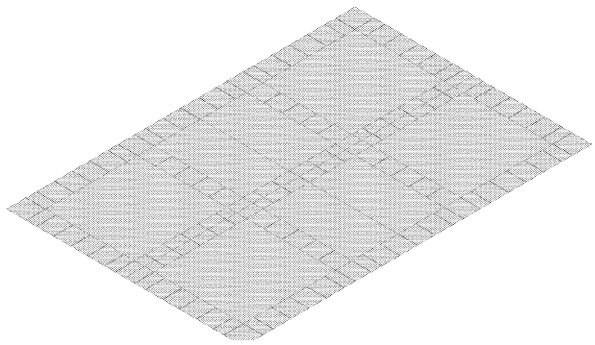


Figure 5: Complete fixture (bolted) to the ground; the rest floor is simply supported by the concrete foundation and no displacement along Z-vertical axis is allowed

Results obtained from the simulation

All the units have been considered in SI: m, kg and sec, so the results will be in m – for the displacements, N – for the forces, Pa – for the pressures and stresses. The load from waste water pressure and also the fixed displacement of some nodes is presented in Figure 6. Total vertical reaction force is 1331.6 x 10^4 N from the applied pressure and structural weight. Figure 7 presents the total elastic displacement distribution due to the applied load.

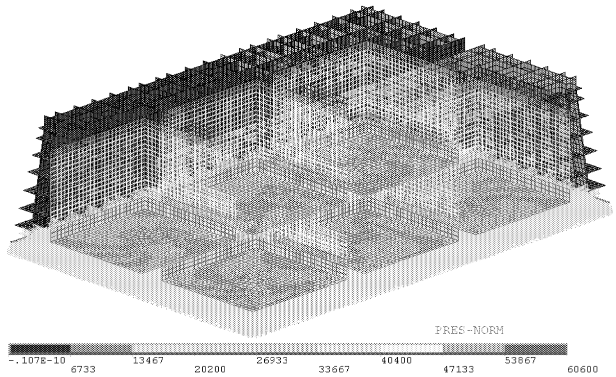


Figure 6: Hydrostatic pressure load in Pa, from all 6 vats full with waste water.

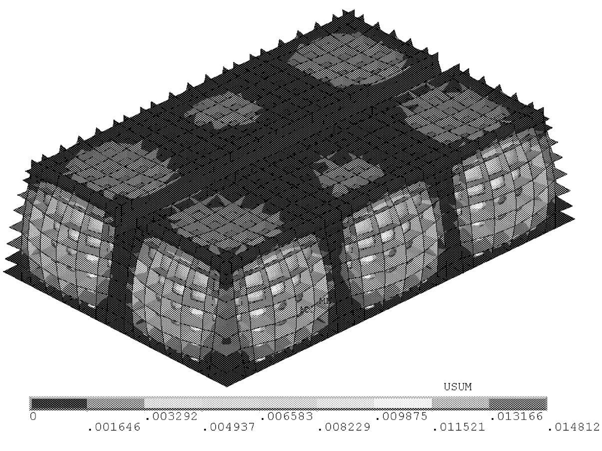


Figure 7: Total deformation of the model in m (displacement scale factor is 100)

Figure 8 shows the maximum von Mises stress distribution in the panels for all parts of the model, whereas figures 9…11 show the same stress but only for components.

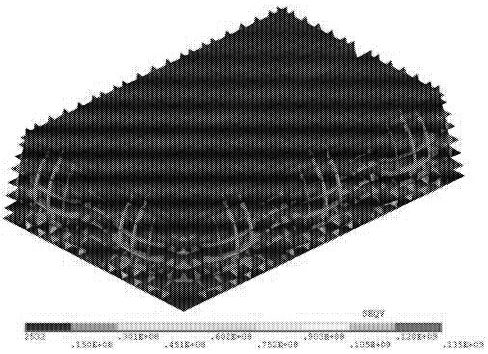


Figure 8: Equivalent stress (von Mises) represented on the middle plane of the composites, as a maximum value from the two outer shells

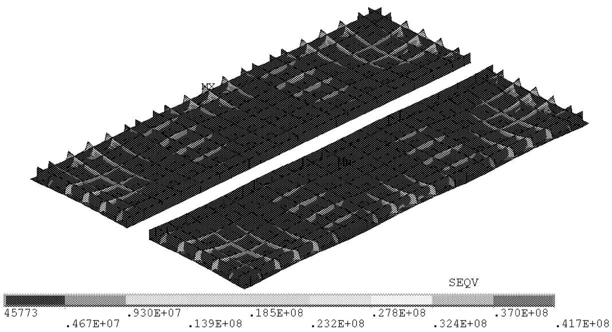


Figure 9: Equivalent stress (von Mises) represented on the middle plane of the composites, as a maximum value from the two outer shells – only upper cover

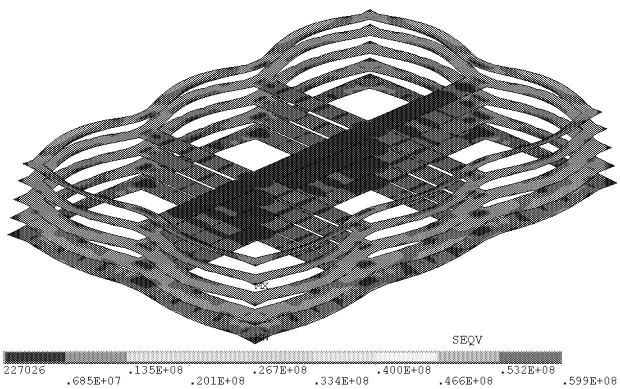


Figure 10: Equivalent stress (von Mises) represented on the middle plane of the composites, as a maximum value from the two outer shells

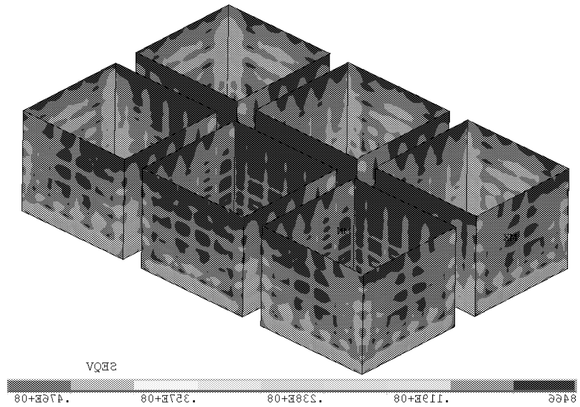


Figure 11: Equivalent stress (von Mises) represented on the middle plane of the composites, as a maximum value from the two outer shells – only vats panels

Figure 12 presents the maximum principal stress distribution on the horizontal plates of the honeycomb structure to see where the structure is predisposed to crack development. Figure 13 presents the minimum principal stress distribution on the horizontal plates of the honeycomb structure to see where the sandwich faces are predisposed to elastic buckling.

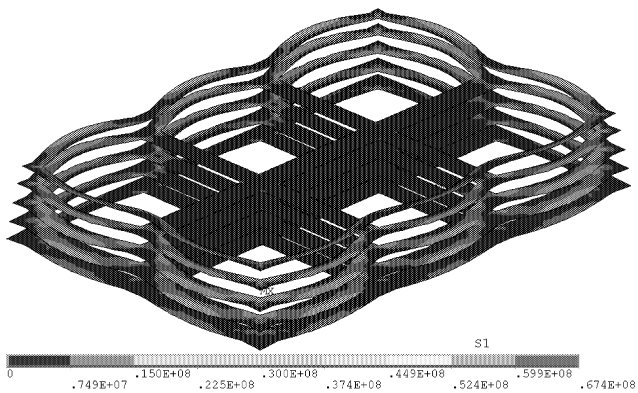


Figure 12: Stretch stress (S1), represented on the middle plane of the composites,
as a maximum value from the two outer shell
(the view deformation is amplified 100 times) –
only horizontal plates of the honeycomb structure

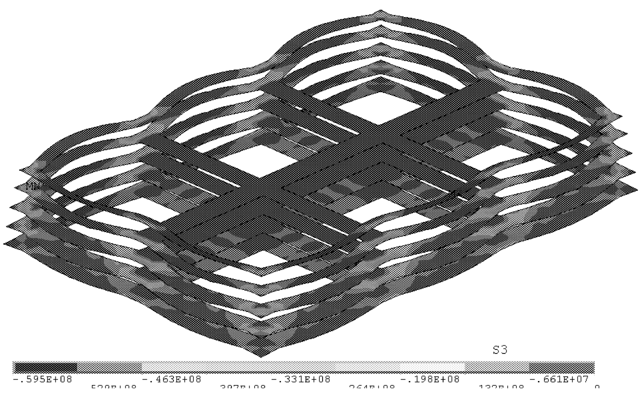


Figure 13: Compression stress (S3), represented on the middle plane of the composites, as a maximum value from the two outer shell–only horizontal plates of the honeycomb structure

CONCLUSIONS

Maximum deformation in the model is around 15 mm, and maximum equivalent stress is 135 MPa. This value is obtained in the area where the honeycomb structure is fixed to the base. The main stress is compression, and in reality in that area the tendency is the floor to sink in the base. The computational model is not fully adequate for 100% trustworthy results. In a honeycomb structure, the stress (stretch and compression) is under 70 MPa, value that should be tolerated by a quality composite panel. This is a theoretical analysis, and in order to fully trust the results we have to completely characterise the composite material and know its mechanical properties: foam material, joint procedure, glue procedure between the core and the shell, yield strength and ultimate strength for both core and shell etc. The main problem of using composites in practice is their sensitivity to low velocity impact (Henkhaus and Ramirez 2004, Her and Liang 2004, Sekine el al. 1998, Shyr and Pan 2003).

REFERENCES

Adams R.D., Comyn A., Wake W.C. (1997). "Structural Adhesives in Engineering", Second edition, Chapman and Hall.

*** ANSYS. (2014). "Structural Analysis Guide and Advanced Analysis Techniques Guide", 3rd Edition, SAS IP".

Caillet, J., Carmona, J.C., Mazzoni, D. (2007). "Estimation of Plate Elastic Moduli through Vibration Testing", Applied Acoustics 68, pp. 334-349.

*** edX Inc. (2015). "Composite Materials Overview for Engineers".

Henkhaus K. and Ramirez, G. (2003). "Overview of research on composite material impact behaviour", 16th ASCE Engineering Mechanics Conference, July 16-18, 2003, University of Washington, Seattle.

Jones, T.S. (1998). In "Handbook of Composites", 2nd edition, edited by S. T. Peters, Chapman & Hall, London, New York (1998), p. 838.

*** MIL-HDBK-17. (2002). "Composite Materials Handbook".

Henkhaus K, Ramirez G. (2004). "Overview of research on composite material impact behaviour". www.ce.washington.edu/em03/proceedings/papers/86.pdf.

Her, S.C. and Liang, Y.C. (2004). "The finite element analysis of composite laminates and shell structures subjected to low velocity impact". Composite Structures, (6) pp.277-285.

Hodgkinson, J.M. (2000). "Mechanical testing of advanced fibre composites". Cambridge: Woodhead Publishing Limited.

Manea, I., Sorohan, Şt., Vasiliu, D., and Vasiliu, N. (2010). "Using advanced fibre composites in civil engineering". National Romanian Research Project IT-FPS No.12-104.

Sekine, H., Hu, N., Fukunaga, H and Natsume, T. (1998). "Low-velocity impact response of composite laminates with a delamination". Mech. Comp. Mat. Struct. 1998; (5): 257-278.

Shyr, T-W., Pan, Y-H. (2003). Impact resistance and damage characteristics of composite laminates. Composite Structures 2003; 62:193-203.

Sorohan, Şt. and Părăuşanu, I. (2005). "Identification of elastic properties of a composite laminate plate", International Conference on Structural Analysis of Advanced Materials, ICSAM 2005, 15-17 sept. 2005, Bucharest, pp. 75-80, ISBN 973-8449-98-7.

http://www.reinforcedplastics.com/view/20099/composites-in-construction-advance-in-new-directions/

BIOGRAPHIES

ION MANEA is a professional mechanical engineer from Romania. He was a superintendent and led a team of 80 in the construction site of a major hydroelectric dam on Danube River. After settlement in USA, he worked as Chief Engineer for a refrigeration Contractor in Seattle, and then as Plant Engineer for J. R. Simplot. Since 1988, as the owner of a design/built food facility consultant and original equipment contractor, he had the opportunity to know first-hand Washington food producing areas and as well Alaska fishing community. In Alaska he was trusted to provide technical services for a fast track project fish plant of $ 70 Million. Ion Manea has worked as a program manager and advisory board member for University of Washington Education Outreach, where he was developing, marketing and managing new educational programs in the area of applied technologies. He was recognized as organizer, marketer and implementer with can do attitude. In 2007 he was the author of two grant applications that have been awarded an aggregate funding of $250,000 from USDA for "Planning for marketing farm-produced sunflower non-ester renewable fuel and biodiesel feedstock" and "Technical Assistance for small minority farm production of sunflower oilseeds and non-ester renewable fuel and animal feed". Currently, he is pursuing research interests with Flower Power USA on biofuels quality assurance and with Polytechnic University of Bucharest on biofuels injection.

ŞTEFAN SOROHAN received the M.Sc. degree in Mechanical Engineering, in the Transportation Faculty, Railway Systems from the University POLITEHNICA of Bucharest in 1989. He became a Ph.D. in Finite Element Model Reduction in 2001 and now he is professor at Strength of Materials Department within the same university.

His main activity consist in finite element modelling of structural, thermal and multiphysics problems.

DANIELA VASILIU graduated in mechanical engineering in 1981 and prepared the Ph.D. thesis in the field of the numerical simulation of the hydrostatic transmissions. She is currently professor in the Department of Hydraulics, Hydraulic Machines and Environmental Engineering, head of Fluid Power Laboratory of the University POLITEHNICA of Bucharest with works in the field of modelling, simulation, and experimental identification of the electro hydraulic control systems.

NICOLAE VASILIU graduated in Hydropower Engineering from University POLITEHNICA of Bucharest in 1969. He became a Ph.D. in Fluid Mechanics after a research stages in Ghent State University and Von Karman Institute from Bruxelles. He became state professor in 1994, leading the ENERGY & ENVIRONMENT RESEARCH CENTRE from the University POLITEHNICA of Bucharest. He managed five years the Innovation Romanian Agency. He is member of the Romanian Technical Science Academy, promoting the numerical simulation as engineering tool.

FLORIN RADOI received a Bachelor degree in Automotive Engineering (2011), and a M.Sc. in Advanced Hydraulics and Pneumatics Systems, both from University POLITEHNICA of Bucharest. He started to work at Parker Hannifin since early college, and now is serving as application engineer in the field of complex hydraulics and pneumatics automation systems. He is also Ph.D. candidate within UPB, where he is developing R&D activity in the field of automotive auxiliary heating systems.

MODERN CONCURRENT APPROACH FOR WIND TURBINE DRIVETRAIN DYNAMICS AND DURABILITY PREDICTION

Dan Predică
SISW Romania
Bdul. Garii nr. 13A
RO 500203 Brasov, Romania
E-mail: dan.predica@siemens.com

Nicolae Vasiliu
University POLITEHNICA of Bucharest
313, Splaiul Independentei, Sector 6
RO 060042 Bucharest, Romania
E-mail: vasiliu@fluid-power.pub.ro

KEYWORDS

Wind Turbine, Structure Durability, Multi-Body System, Noise and Vibrations

ABSTRACT

The wind energy industry continually evolves, and industry professionals have streamlined drivetrain design to a consensus configuration. This configuration and its design iteration have existed for many years; consequently, design and manufacturing flaws have been minimized sequentially. Regardless of the maturity of the drivetrain design and design process, however, most wind turbine downtime is attributed to gearbox-related issues. Moreover, driveline replacement and lubrication accounts for 38% of the parts cost of the entire turbine. The authors present an approach to modeling and simulation of the dynamic/life behavior of drivelines in wind turbines, based on practical experience gained in research projects related to the previously mentioned aspects. The following steps involved in the dynamic/life simulations of the wind turbine drivelines are explained by the following:
-Finite Element model updating of a full wind turbine, with the accent on the driveline model by *component level correlation*, and *system level operational correlation*;
-Comparative approaches for accurate drivetrain loads determination/application;
-Damage/Life prediction calculation applied to drive train components.

INTRODUCTION

Nowadays, manufacturers are pressured to deliver more complex products with increased quality in shorter development cycles. Engineering the performance of mechanical designs with traditional test only-based development processes is no longer an option. The only valid alternative is evaluating functional performance attributes on a CAE virtual prototype, employing methodologies previously validated by real file measurements. Any use of computer software to solve engineering problems is termed as Computer-Aided Engineering (CAE). A typical CAE program is made up of a number of mathematical models encoded by algorithms written in a programming language.
The natural phenomena being analyzed are represented by an engineering model. The physical configuration is described by a geometric model. The results, together with the geometry, are made visible via a user interface on the display device and a rendering model. A wide range of technologies are used during different design phases in simulation driven product development, including structural finite element analysis, acoustics, crash analysis, fatigue and failure analysis, and computational fluid dynamics.
The main aspects of simulation and testing can be grouped as follows:
-3D Virtual simulation: gear modeling, system modeling for noise, vibration and durability, fatigue life prediction;
-1D Multi-Domain simulation: electric architecture comparison, control strategies studies, power grid connection optimization;
-Testing solutions: vibro-acoustic measurement and analysis, driveline operational measurements.

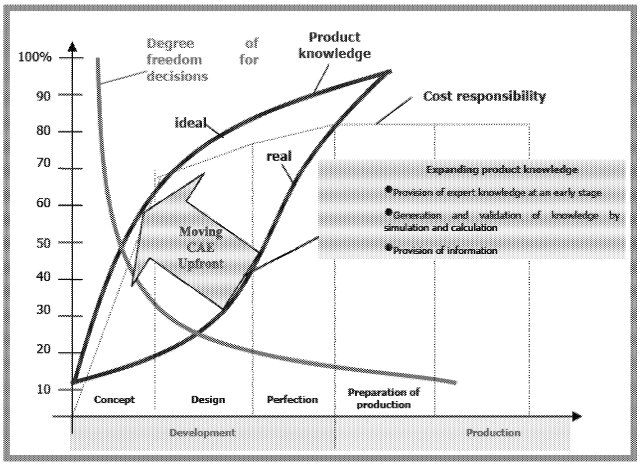


Figure 1: Evolution of applying CAE methods in the design process

FINITE ELEMENT MODEL UPDATING

The purpose is to validate individual/assembly Finite Element (FE) models starting for an initial FE model. The full wind turbine preliminary modelling methods were validated/refined in order to make sure that the FE model accurately estimates the dynamics of the actual wind turbine.

Component level correlation

Component level correlation was performed at either Modal Assurance Criterion (MAC) or Frequency

Response Function Assurance Criterion (FRAC) level. The MAC expresses the degree of relationship between a pair of vectors. These vectors can be the modal vectors (normal or complex mode shapes) or response vectors from a forced response for instance. It is practically a mathematical criterion, that enables comparison of the FE calculated and measured normal modes characterizing the dynamics of a component/structure.
The FRAC provides a measure of the correlation between a set of test and FE frequency response functions. Besides the visual comparison of curves, the relationship between the functions can be expressed through the FRAC. This function expresses the comparison between the experimental and analytical functions. The FRAC results in a Frequency Factor, which is an indication of the global over or under estimation of the stiffness or mass of an FE structure.

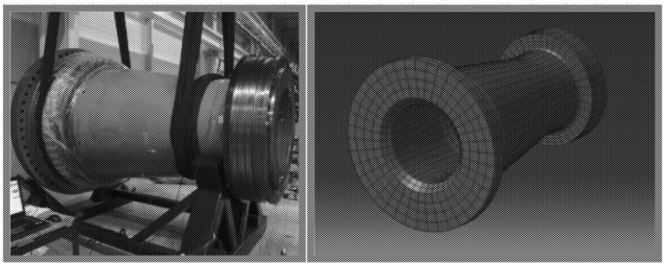


Figure 2: Physical and FE model of the Main shaft

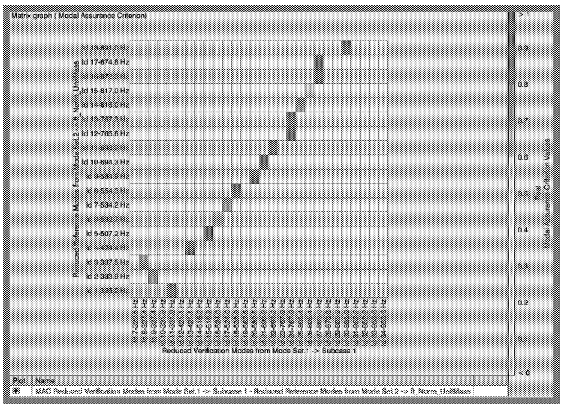


Figure 3: MAC matrix for the main shaft

The MAC matrix for the Main Shaft, diagonal and values above 0.7 is indicative of a quite good correlation. Low frequency radial pumping modes were inverted with respect to test measured data, this being due to the fact that the respective mode is a repeated mode occurring due to axial symmetry of the shaft.

Blade correlation

As the MAC is best fit for low/medium modal density structures and given that the blades were expected to yield a rather large number of modes in the frequency range of interest, the FRF level correlation was chosen. The number and location of points used for measurement was determined using a Pre-Test Analysis. It allows defining an optimal measurement set-up in terms of number and location of sensors and excitation devices based on information from an existing FE model.
Even though the FE model used is not perfect, the technique generates good results. It makes sure that enough measurement points are use in order to capture the relevant modes of the structure (by evaluating the AutoMAC), as well as that the proper force input points are being used in order to excite the dynamic modes of interest (by evaluating the Driving Point Residues). The test configuration location selected for the blades is presented in the Figure 4. Blade root is positioned to the left of the picture.

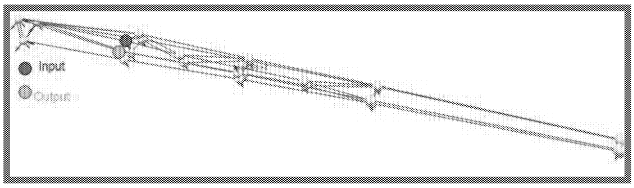


Figure 4: Blade test wireframe

For the correlation evaluation, individual FRF, as well as the overall sum of the measured FRFs were considered. Both individual and sum of FRFs showed a good correlation, both in terms of amplitude and phase of the FRFs, as presented in Figure 5.

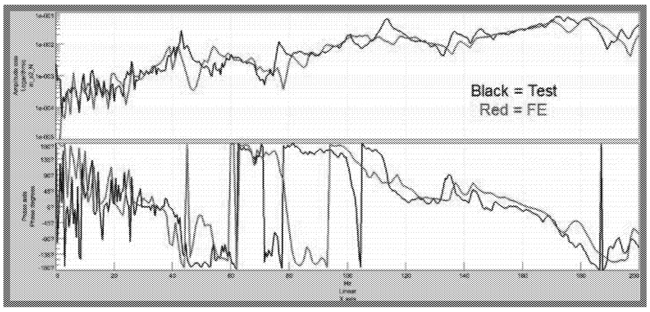


Figure 5: Blade Frequency Response Function comparison (amplitude and phase)

System level operational correlation

Having validated the individual FE models for the most relevant parts of the wind turbine, the correlation in operational conditions was performed to further validate the virtual model. For evaluating the results, visual comparison and MAC applied to the forced response vectors (displacements, accelerations) were used. Overall correlation related to the drive train proved to be positive. Figure 6 presents a comparative deformation shape of the wind turbine drivetrain. The methods used for determining and applying the loads in performing the actual forced response, will be described in the Comparative Approaches for Drivetrain load Determination/ Application section.

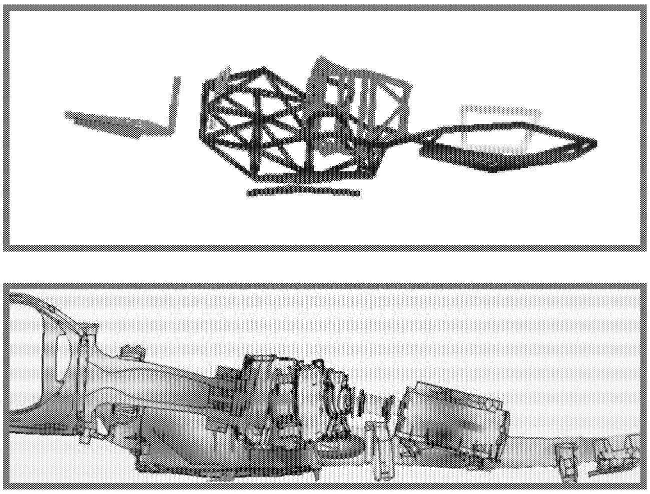


Figure 6: Comparison of Operational Deflection Shapes of the Test (upper image) and FE (lower image)

COMPARATIVE APPROACHES FOR DRIVETRAIN LOAD DETERMINATION AND APPLICATION

Purpose: validating the force spectra to be applied to the previously updated FE model of the wind turbine.
Several methods for the calculation and application of the loads the following approaches were considered: Rigid Multi Body calculation, Flexible Body calculation and Inverse Load Identification with test based accelerations.
Additional information pertaining to these methods are presented next.

Rigid Multi Body (MBS) simulation technique

MBS consists in creating a virtual model comprising a collection of interacting bodies having the dynamic behavior described using six degree of freedom (DOF) attached to each body's Centre of Gravity.
Mass and Centre of Gravity location are sufficient for each body's definition. In this kind of model, the different drive train components - shafts, gears, planets, planet carrier and housing are considered to be rigid.
Bearings and gear contact meshing forces are modelled by means of discrete flexibilities. This type of model is referred to by the term: rigid multi-body model with discrete flexibility.
The system of differential-algebraic equations of motion solved in time domain for the MBS Dynamic analysis type is:

$$\begin{bmatrix} M & 0 & \phi_r^{\ T} \\ 0 & J' & \phi_\pi^{\ \prime T} \\ \theta_r & \phi_\pi' & 0 \end{bmatrix} \begin{bmatrix} \ddot{r} \\ \dot{\omega}' \\ \lambda \end{bmatrix} = \begin{bmatrix} F^A \\ n'^A - \varpi J \omega' \\ \gamma \end{bmatrix}$$

where:

M - system mass matrix
J' - inertia matrix of the body
F^A - bodies applied force matrix
ω' - angular velocity matrix
Φr, Φπ, γ - terms of the acceleration equation
λ - Lagrange multipliers
n'^A – applied torques applied to the bodies

Flexible Multi Body (MBS) simulation technique

It overcomes the shortcomings of the Rigid MBS approach, offering the best insight in the local deformations and stress levels. This technique is no straightforward adaptation of the traditional method, since it implies the inclusion of additional generalized DOFs to represent the deformations of an individual body. These DOFs are introduced as component modes, which are calculated by an FE analysis and subsequently synthesized into a global set of modes for one body, through the process known referred to as Component Mode Synthesis (CMS).
CMS uses the following type of modes to describe the dynamic behavior of a component:
-Normal modes: resulting from an eigen values problem and describing the dynamic flexibility of the component;
-Constraint/Attachment modes: describe the local stiffness of the connections of a body to the neighboring ones;
-Rigid modes: the rigid-body modes represent the gross motion of a substructure.
To model the gears meshing inside the gearbox, which is one of the main excitation sources in the drivetrain, two main approaches are available:
-an analytical one based on the CAI/ISO methodology for meshing stiffness calculation;
-an advanced, discrete method which combines a series of FE cases for calculating teeth deformation and an analytical definition of the tooth micro-geometry (fig. 7).

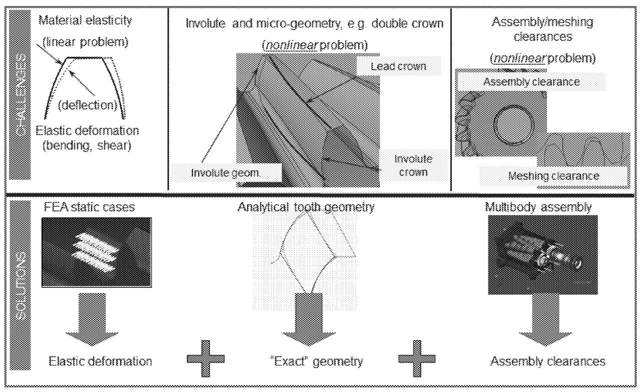


Figure 7: Advanced method for calculation of the gear contact force

Load Identification technique

Load Identification problem is the inverse one of Forced Response: knowing the system responses and the transfer functions, the loads that caused this response are identified from the responses and the transfer or frequency response functions (FRFs). They are either measured, or computed, based on FE dynamically validated models, as a relationship between the structural responses (displacements, velocities or accelerations) due to unit force at the input points. Finally, they can be assembled in a transfer function set. The transfer function set matrix is then inverted and combined with structural/acoustical operational responses gathered in a response function set, in order to obtain force estimation. For the identification of loads Lj at the j-th excitation DOF whereby $1 \leq j \leq N$ for given operational responses Ri at the i-th response DOF whereby $1 \leq i \leq M$, this gives:

$$\begin{bmatrix} H_{11} & \dots & H_{1N} \\ .. & \dots & \dots \\ H_{1M} & \dots & H_{MN} \end{bmatrix}^{+} \begin{bmatrix} R_1 \\ \dots \\ R_M \end{bmatrix} = \begin{bmatrix} L_1 \\ \dots \\ L_M \end{bmatrix}$$

Hij is the transfer function between the i-th response DOF and the j-th excitation DOF. The sign + denotes the pseudo inversion operator. For square matrices, this corresponds to a classical inversion operator. In order to avoid numerical problems in the matrix pseudo inversion, singular value decomposition methods are used. The number of responses (M) should at least be equal to the number of input forces to be estimated (N), in order to have a unique solution for the operational forces. However, taking more responses than the unknown loads (M > N) allows you to over-determine the set of equations. This allows you to obtain a more accurate, least squares estimate result for the operational loads.

The following combinations of the previously presented methods were investigated:
- Load Identification using Fully Rigid MBS model;
- Load Identification using Partially Flexible MBS model;
- Load Identification using Test Measured Accelerations.

For the Load Identification based on Test measured accelerations, several acceleration sets combinations were considered: Main Shaft, Gear Box, Generator Area, and Blades. They all yielded similar results, with the highest responses for frequencies associated with gearbox operation.

Several operation conditions accelerations, like running at nominal speed, running at intermediate speeds and run-up were considered.

For Load Identification using Partially Flexible MBS model, only relevant parts from mass and stiffness perspective, were included as flexible parts: Blades, Main Shaft, Gearbox Housing.

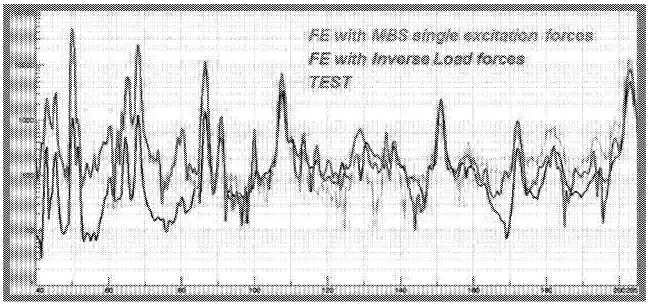


Figure 8: Comparative overview force spectra for different methods

By comparing the acceleration responses obtained for different cases, it can be concluded that the Partially Flexible MBS simulation results were the closest to the ones measured on the actual wind turbine.
Since it was validated that the gearbox related frequency content was most present in the identified forces, the method of applying drivetrain forces to the full FE model of the turbine was scrutinized.
Two options were considered: applying the forces generated by each stage of the gearbox to separate point, and applying the forces gearbox generated forces to a single center point, connected to the relevant force transfer areas in the gearbox: bearings, gear shafts etc.
After evaluating the results, it was found that the applying the loadings at separate points, one for each stage of the gearbox resulted in the best predicted acceleration levels.
Be that as it may, the results of one point load application scheme was considered to be sufficiently accurate and it was preferred as it simplifies the load application for durability calculation, being fit to the single random load vibration fatigue approach .

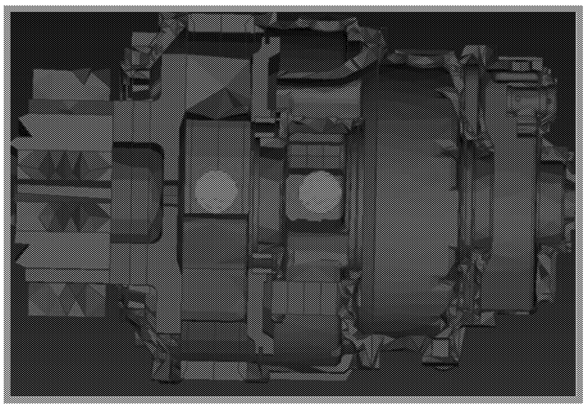


Figure 9: Force application points for forced response

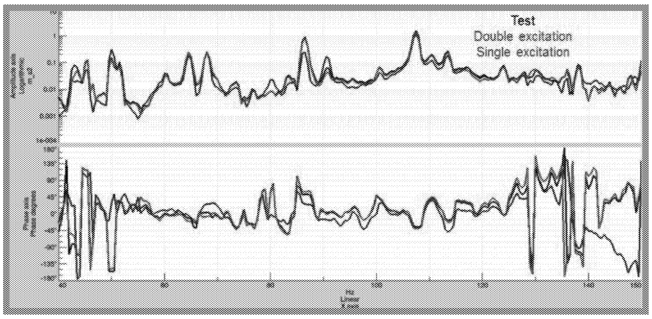


Figure 10: Comparative overview of the acceleration response along the drivetrain points (amplitude and phase)

DURABILITY CALCULATION

The main steps involved in a virtual prototyping process for *durability* are as following.
Step 1: external loads to internal forces; from given external loads compute internal forces acting on components.
Step 2: internal forces to stress tensor spectra (fig.11); from given forces acting on the component, determine time history of the local pseudo stress tensor based on theory of elasticity.
If the loading frequency content is *below* lowest natural frequency we can use *Quasi-static superposition* approach. In this case, the influence coefficients c, which account for the stress concentration factors, due to individual loads, in the equivalent analytical formulation, are determined (figure 12).
If the loading frequency is *above* lowest natural frequency *Modal superposition method* is used. In this case the modal participation factors and modal influence coefficients c, are determined (figure 13).

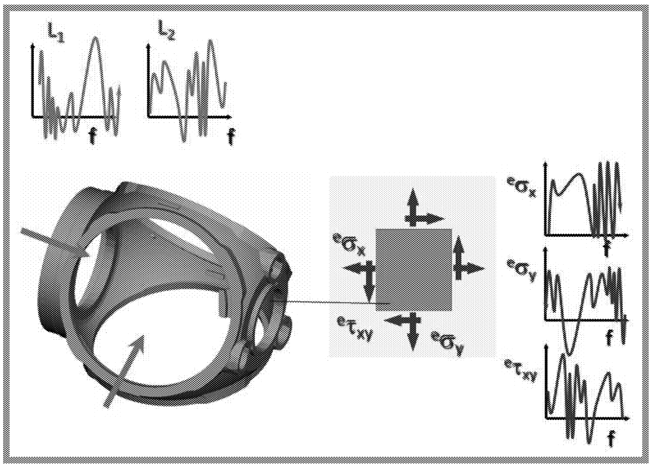


Figure 11: Internal forces converted to stress tensor spectra through the test validated FE analysis

Step 3: stress tensor to damage; fatigue damage is initiation and growth of short cracks, but different levels of detail are required for different applications. Two analysis approaches available are possible: long life (high cycle fatigue) approached by stress-life, and shorter life (low cycle fatigue) approached by strain-life.

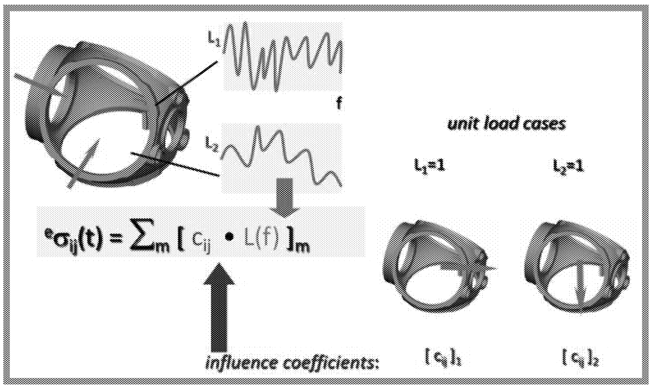


Figure 12: Quasi-static influence coefficients determination by unit load application

For both methods, the stages involved are: determination of the local load histories from FE analyses; calculation of local Rainflow matrix; endurance limit correction; mean stress correction; damage calculation.

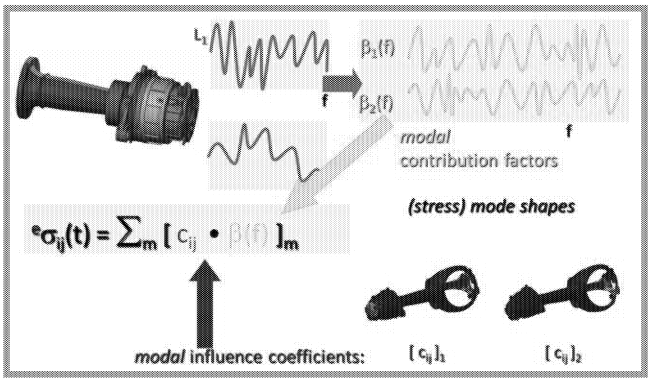


Figure 13: Modal Superposition method: influence coefficients and modal participation factors are combined

Regarding the dimensionality of the stress used for damage calculation, two approaches are available: equivalent stress and critical plane.
Equivalent stress and strain approaches reduce the multi-axial stress or strain states to a single scalar value that can be used to calculate fatigue damage just as for a uniaxial state of stress.
When a structure is under cyclic multi-axial loading, it is necessary to use multi-axial fatigue criteria that account for the multi-axial loading. Critical plane analysis refers to the analysis of stresses as they are experienced by a particular plane in a material, as well as the identification of which plane is likely to experience the most extreme damage.

Figure 14: Safety factor plot: blue and green areas are of low safety factor

IEC 61400-1 Wind turbine generator systems – Part 1 Safety Requirements standard, identifies four different classes of wind turbines to suit differing site wind conditions, with increasing class designation number corresponding to reducing wind speed, 1 to 4. A typical wind turbine is subjected to a severe fatigue-loading regime. The rotor will rotate some $2x10^8$ times during a 20 year life, with each revolution causing a complete gravity stress reversal in the low speed shaft. Given the long life design requirements applicable to wind turbines, stress-life approach is appropriate for most of the components.
For the durability calculations, Virtual Lab Durability /FALANCS Vibration Fatigue was used. Several random load quantification and stress/strain processing methods for damage calculation are available: Signed von Mises with Goodman correction, Full 3D critical plane with Goodman correction, etc . They allow calculation and display of durability related quantities like: damage, fatigue life, safety factor, as showed in figure 14.

CONCLUSIONS

A complete CAE based method for simulating the dynamics/durability of drivetrain related components was presented. This allows identification of critical loads and critical design parameters; enables optimization runs based on damage results; permits easy switch of components and subsystems; it is scalable: evolve model from concept to detailed.
Further work will be focused on: investigating the influence on the calculated load of including additional flexible bodies in the MBS calculation; facilitating the co-simulation between the 3D MBS simulations and 1D multi-physics platforms (Imagine.Lab, Matlab) for control system integration and system level engineering.

ACKNOWLEDGEMENTS

The authors are grateful for all the technical support received from the SISW COMPANY in different manners: free licenses, free technical training for the research team members, and many other facilities.

REFERENCES

Van der Auweraer, H., Leuridan, J. 2012. "Innovation Engineering by Simulation". ECEC'2012 Conference, Bucharest, 18-20 April 2012.
Hermanns, O., Hack, M., Gerard, F., Bosmans, I. 2007. "MBS and Durability Solutions for Wind Turbines", LMS International Workshop, August 2007.
Heege, A., Radovcic, Y., Betran, J. 2006. "Fatigue Load Computation of Wind Turbine Gearboxes by Coupled Structural, Mechanism and Aerodynamic Analysis", DEWI Magazine Nr. 28, February 2006.
Oyague, F., Butterfield, C.P., Sheng, S. 2009. Gearbox Reliability Collaborative Analysis, WINDPOWER 2009 Conference, Chicago, Illinois, May 4 – 7, 2009.
Burton, T. Jenkins, N. Sharpe D., Bossanyi, E. 2011. Wind Energy Handbook – Second Edition, John Wiley & Sons Ltd., 2011.

WEB REFERENCES

http://lmsintl.com
http://www.nrel.gov/
http://www.techdesignforums.com/
http://www.nafems.org/

BIOGRAPHIES

DAN PREDICA received the M.Sc. degree in Mechanical Engineering, Aircraft Engineering Department, from the University Transilvania of Braşov in 1998. He is currently holding the position of Team Leader at SISW Engineering Services in Romania, having extensive knwoledge of MBS and FE based CAE codes. He is Ph.D. student of the University POLITEHNICA of Bucharest, in the CAE based wind turbine design methods.

NICOLAE VASILIU graduated in Hydropower Engineering from University POLITEHNICA of Bucharest in 1969. He became a Ph.D. in Fluid Mechanics after a research stages in Ghent State University and Von Karman Institute from Bruxelles. He became state professor in 1994, leading the ENERGY & ENVIRONMENT RESEARCH CENTRE from the University POLITEHNICA of Bucharest. He managed five years the Innovation Romanian Agency. He is member of the Romanian Technical Science Academy, promoting the numerical simulation as engineering tool.

THERMAL AND STRUCTURAL ANALYSIS OF AN AUXILIARY VEHICLE HEATER SYSTEM

Rădoi Florin, Sorohan Ştefan, Vasiliu Nicolae, Călinoiu Constantin
University POLITEHNICA of Bucharest
313, Splaiul Independentei, Sector 6
RO 060042 Bucharest, Romania
E-mail: {radoi, sorohan, vasiliu, calinoiu}@fluid-power.pub.ro

KEYWORDS
Model design, numerical simulation, multiphysics FEM analysis, concurrent engineering by compatible software packages.

ABSTRACT

The paper deals with a steady-state thermal analysis and thermo-elastic stress analysis of an auxiliary vehicle heater system using finite element modeling. A detailed FE model has been created for this purpose. The FE model consists of the main parts of the assembly, including a description of thermal and mechanical loads, and contact interaction between its parts and external boundary conditions. The model considers heat transfer coefficients from current literature, which are iteratively corrected to estimate the temperature distribution in the main parts according to the target power of the heater estimated to be around 4kW. The design was developed out using Catia v5 r18, and analysis was carried out using the FE program ANSYS Workbench v15. The analysis is a very useful step in a concurrent design process of such a device.

INTRODUCTION

This paper presents a three dimensional simulations of the temperature distribution within an automotive additional heater, and then the thermo-elastic stresses that appears due to the non uniform temperature distribution. The computational approach attempts to avoid the challenging aspects of complicated physical and chemical interactions of the burning and flow aspects.
The design and analysis of gas auxiliary vehicle heating systems is based on combined theoretical and empirical approach and the design of combustion chamber is less than exact science. Modern auxiliary vehicle heaters (fig. 1) are very compact and have an extremely high energy conversion rate. Operating temperature range is usually between -40 to +40 °C.
When the system is used as parking heater, it is activated by a remote control device, and it goes into start mode. The heater fan (blower) "washes" the burner with fresh air. Than the fuel is aspirated by a solenoid volumetric pump from the petrol tank and sent to the combustion chamber of the heater. The fuel-air mixture is ignited by a glow plug and the combustion chamber starts to warm up. A centrifugal circulating pump moves the cooling agent along a closed loop: from the parking heater to the car heater‘s heat exchanger, then into the engine and back to the parking heater. The car heater‘s heat exchanger send the heat into the cabin via the car‘s fan. The cooling agent is also introduced in the car engine and warms it up as well.

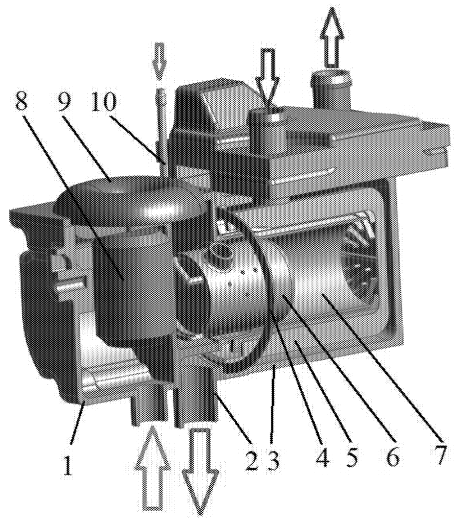


Figure 1: Auxiliary water heater system (WEBASTO).

A detailed complex FE analysis provides valuable information about temperature distribution and mechanical stresses in the overall assembly of the auxiliary vehicle heater when are used together multi-physic (thermal and structural), combustion analysis and CFD. The biggest part of the thermal energy produced by the combustion process is transported with the exhaust gases and transferred to the heat exchanger boundaries of the combustion chamber by different heat transfer mechanisms: radiation, forced convection and conduction. For a heating output of about 4 kW, fuel consumption in a heating phase of 20 minutes at full load is around 0.2 l. The lower consumption of the preheated engine compensates up to 40 % of the heater fuel consumption.

MODELLING BY FEM

The target of the finite element analysis was to get the temperature distribution and maximum stresses generated in the components of the water heater system. Because the measuring of the temperatures is a very complex task and the CFD analysis is not yet available, for the first phase of the design a simplified analysis was done. The complete temperature distribution in structural components can be obtained using an adequate finite element model in thermal analysis and neglecting the structural effects. Then, using the obtained temperature distribution as loads (Fig 2) a sequentially multi-physics coupled analysis may be developed.

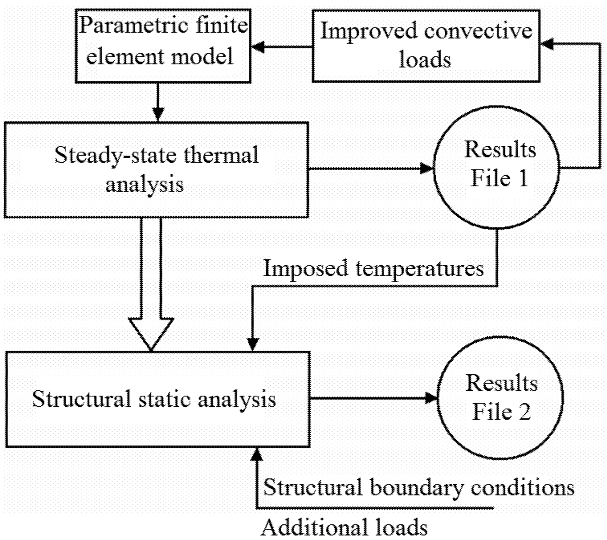


Figure 2: Sequential coupling of thermal and structural problem

A sequentially coupled physics analysis is the combination of analyses from different engineering disciplines which interact to solve a global engineering problem (Ansys 2004). The term sequentially coupled physics refers to solving one physics simulation after another. Results from one analysis become partially inputs for the next analysis. If the analyses are fully coupled, results of the second analysis will change some input to the first analysis. Due to some unknowns concerning the real boundary conditions in the thermal analysis, the steady state thermal analysis has a supplementary loop for convective load improvement to fit the imposed thermal boundary condition for partial validate the thermic balance.
The steady-state thermal analysis was simplified, only conduction and equivalent surface convection from the air, water and combustion gases being included in the thermal balance. Radiation and mass transfer were neglected in theoretical formulation but they are considered through adequate surface convection coefficients. The global equation of the thermal model, in matrix form, is (Cook et al. 1989)

$$[K_t]\{T\} = \{Q\}, \tag{1}$$

Here $[K_t]$ is the thermal characteristic matrix, $\{T\}$ is the vector of unknown nodal temperatures and $\{Q\}$ is the thermal load vector. The thermal characteristic matrix has components from conductivity and from surface convection, while $\{Q\}$ is obtained only from surface convection coefficients and bulk temperatures.
The convective normal heat flow q is obtained from the relation

$$q = h(T_S - T_B) \tag{2}$$

where h is the film coefficient, T_B is the bulk temperature of the adjacent fluid and T_S is the temperature at the surface of the model. A steady-state thermal analysis may be either linear, with constant material properties, or nonlinear, with material properties depending on temperature. The thermal properties of most materials do vary with temperature, so the analysis is usually nonlinear.
The static structural analysis is described by an equation similar to (1),

$$[K]\{U\} = \{F\}, \tag{3}$$

where $[K]$ is the stiffness matrix, $\{U\}$ is the vector of unknown nodal displacements, and $\{F\}$ is the global load vector obtained from internal pressure and thermal loading. Even for small displacements, the stiffness matrix is nonlinear because it is a function of mechanical material constants like Young's modulus E, Poisson ratio ν etc., all of them depending of temperature.
The thermal expansion coefficient α of a material is function of temperature, and the thermal strain on a linear direction is obtained by

$$\varepsilon_{th} = \alpha_s\left(T - T_{ref}\right), \tag{4}$$

where α_s is the secant coefficient of thermal expansion, T is the current temperature and T_{ref} is the reference (strain-free) temperature.
In the loop of improved convective loads module (see Fig. 2) of thermal analysis, we supposed that the temperature of fluids inside and outside the heater is unknown, also the convection film coefficients inside and outside the walls are unknown but their range correspond to the current literature (Baukal 2000; Gulder 1986; Heywood 1988; Incopera and DeWitt 2002; Kays and Crawford 2005). Starting with some guess values we can improve the unknowns to fit the estimated power of the heater. The project schematic of the analysis is presented in Figure 3.

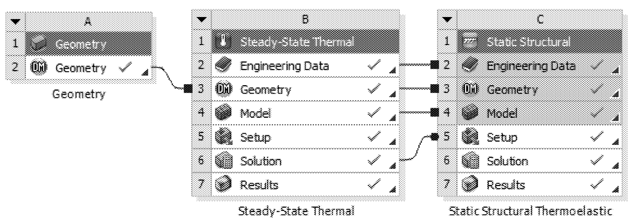


Figure 3: Workflow of the sequentially coupled physics analysis

POWER OUTPUT ESTIMATION

Net heating value of light oil (or diesel) is around of 36 MJ/l. If the fuel consumption in the water heating system is around of 0.65 l/h, results an average net generation power from fuel burning around of 6.5 kW. The desired heat exchanger power is 4.3 kW, so the effective power of the heat exchanger is estimated to be around of 66 %.

MATERIAL DATA

Design work of modern auxiliary vehicle heater requires accurate determination of temperature, in order to optimize the correlations between size, shape and properties of materials used for parts on one hand, and the thermo-mechanical applications, on the other.
Adopting the aluminum alloy for heat exchanger and its housing can effectively reduce the total weight but the elastic modulus of aluminum alloy is far less than that of cast iron, the deformation amount of the aluminum and the corresponding stress can be too large for a proper design.
In the combustion chamber, there are high values of combustion temperature in the order of 2000K - 2500K (Lefbvre and Ballal 2010; Matarazzo and Laget 2011). The maximum temperature of the proximity material is much lower and the regions around the combustion chamber need to be securely cooled to prevent overheating. These are key information in order to optimize the design of the heater.
The combustion chamber has to be tested under these operational conditions. This procedure is necessary but expensive, so numerical simulation is used to reduce costs and prevent costly design mistakes.
Important information was also found in (Serth 2007) which stated that the highest temperature of any point in each component must not go more than 66% of the melting point temperature of the component material.
In the finite element model the heat exchanger was considered from an aluminum alloy but the flame guide and burner tube were considered from a refractory steel. Except the thermal conductivity and Young's modulus which were considered as temperature functions (Figs. 4 - 7), the remaining properties were considered constant.
The Poisson coefficient was considered 0.28 for steel and 0.33 for aluminum alloy, and the thermal expansion coefficient was considered $17*10^{-6}$ 1/K for refractory steel and $23*10^{-6}$ 1/K for aluminum alloy.

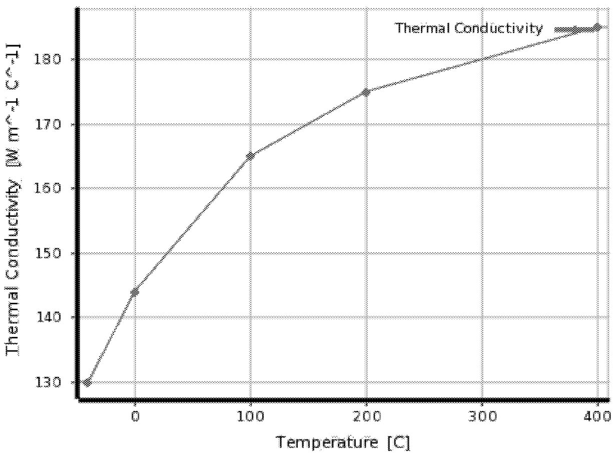


Figure 4: Thermal conductivity of aluminum alloy

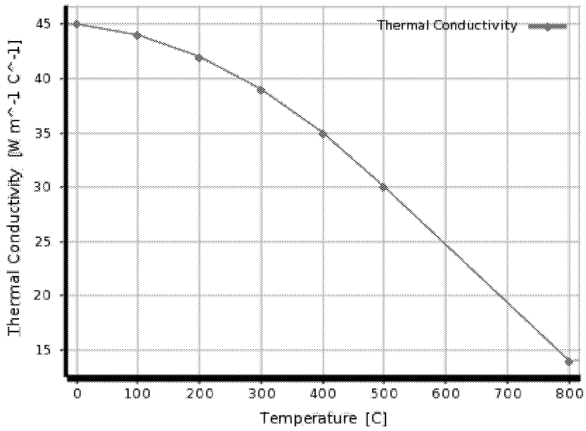


Figure 5: Thermal conductivity of refractory steel

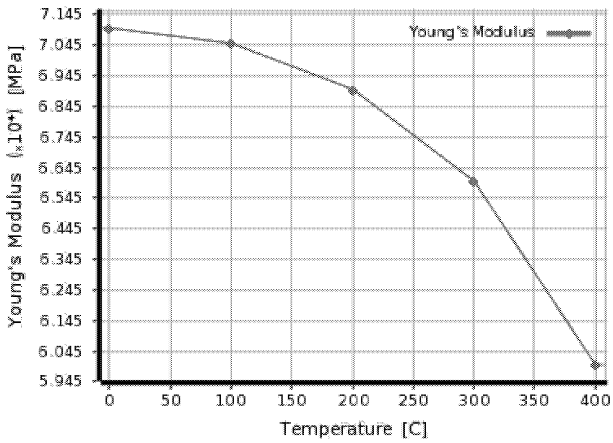


Figure 6: Young modulus of aluminum alloy

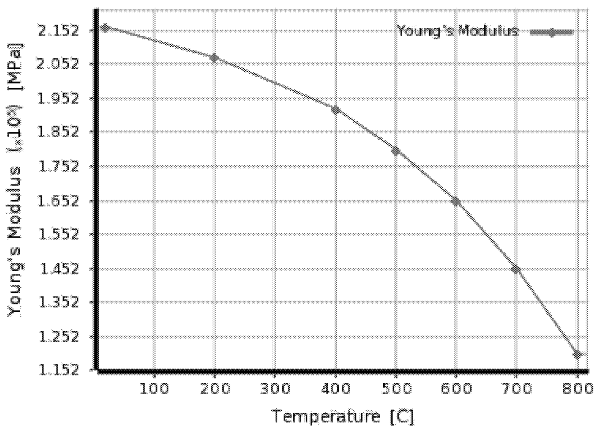


Figure 7: Young modulus of refractory steel

MESH DESCRIPTION

To optimize the analysis, we included in the finite element model only the main parts 1-7 from Figure 1 which indeed are important for thermal and structural analysis. Some parts as burner motor, blower, fuel lines, plastic covers, electronic devices, sensors etc. were neglected because their effect is not very important. Combustion gases or heat transfer medium were not explicitly modeled. Their effect upon structural parts was considered only as boundary conditions. Due to irregular shape of the parts we adopt for mesh a tetrahedral finite element type, Solid87 for thermal analysis and Solid187 for structural analysis. A total of 205491 of nodes and 112178 finite elements result after the automatic mesh using maximum relevance of the mesh algorithm (Fig. 8). Some parts were further refined using a 2 mm element size criterion. The same mesh was used for thermal and for structural analysis.

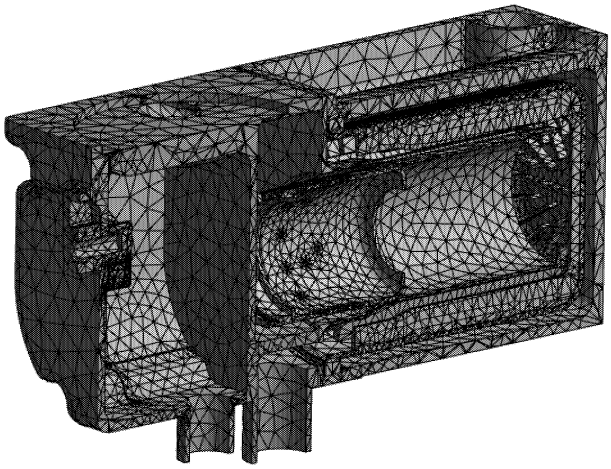


Figure 8: Mesh of the finite element model using ten nodes tetrahedral elements

CONTACT ELEMENTS

Thermal interactions between parts were modeled by perfect non thermal resistence contact elements, whereas the structural contact elements were considered according to real interactions, (i.e. bonded or no separation). Linear contacts were introduced between involved surfaces, for example, the Figure 9 marks linear no separation contact between port housing and burner tube. The element type of contact used in Ansys was Conta174 and for target element was used Targe170.

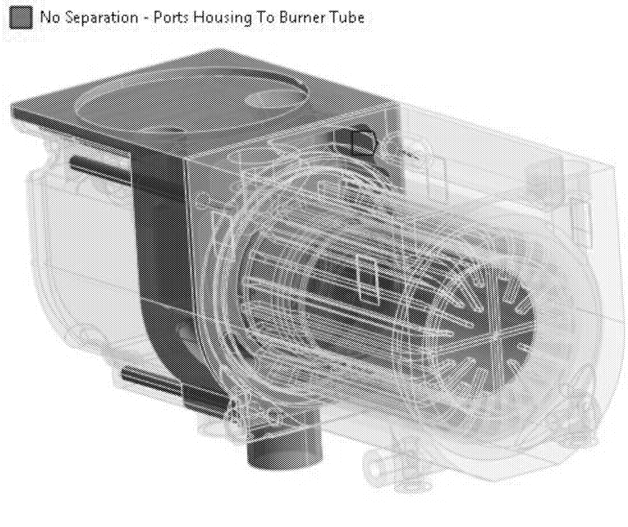


Figure 9: Example of contact between bodies

THERMAL BOUNDARY CONDITIONS

The thermal loads were applied iteratively on surfaces, with the updated values mentioned in Table 1. The model was iterated until the desired result was obtained - the heat exchange power for heat exchanger around of 4 kW.
Figure 10 shows the areas of convective boundary condition between combustion gases and the heat exchanger and Figure 11 shows the areas of convective boundary condition between heat transfer medium-water and the heat exchanger body.

Table 1. Convective boundary conditions

Part	Location	h $\left[\frac{W}{m^2\,°C}\right]$	T_B [°C]
1-Blower cover	Interior	45	0
	Exterior	25	0
2-Ports housing	Interior near motor	45	0
	Exterior	25	0
	To burner	50	0
	To outlet	100	400
3-Heat exchange housing	Exterior to air	25	0
	Interior to water	1500	100
5-Heat exchange body	Exterior to water	1500	100
	Interior to combustion gases	150	700
6-Flame guide	Exterior to admission air	50	0
	Interior to flame	200	720
7-Burner tube	Interior to admission air	50	0
	Interior to combustion gases	150	700
	Exterior to combustion gases	150	700
	Flange to outlet	100	500

Figure 10: The convective boundary condition between combustion gases and the heat exchanger

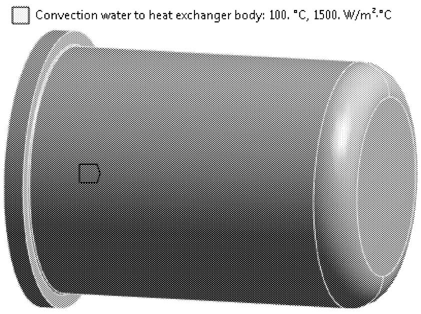


Figure 11: The convective boundary condition between heat transfer medium-water and the heat exchanger body

THERMAL RESULTS

For the boundary conditions from Table 1 it was obtained a power of 4.53 kW for convective heat transfer between combustion gases and heat exchanger body and 4.16 kW for convective heat transfer between heat transfer medium (water) and heat exchanger body. The temperature distribution in structural parts varies between 77 ºC and 700 ºC (Figure 12).

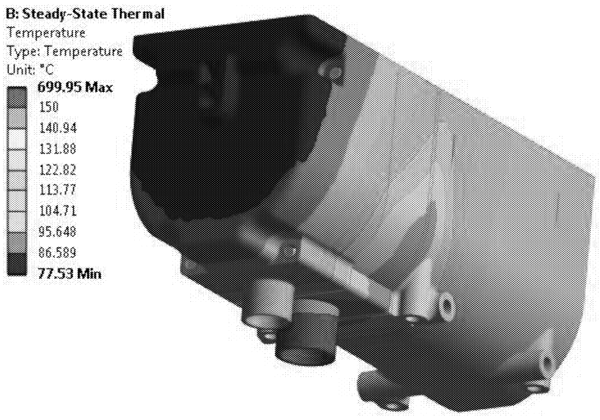


Figure 12: Temperature distribution on all model. (Maximum temperature is inside; here the top legend is not uniform, it was modified for improving clarity)

Figures 13 - 15 present the temperature distribution in the main components of combustion chamber. The gray shadow correspond to unselected parts for which the results are not plotted.

STRUCTURAL BOUNDARY CONDITIONS

The mechanical connection using long bolts between blower cover and heat exchange housing were included as special connection using Beam188 for bolt and constraint equations to transfer loads to area of direct contact between parts.

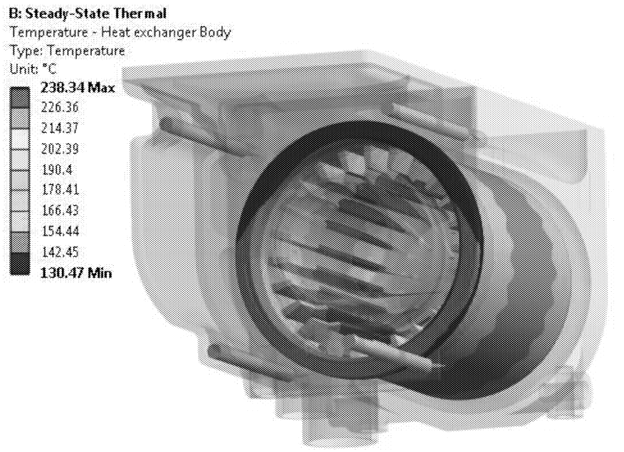


Figure 13: Temperature distribution on the heat exchanger only

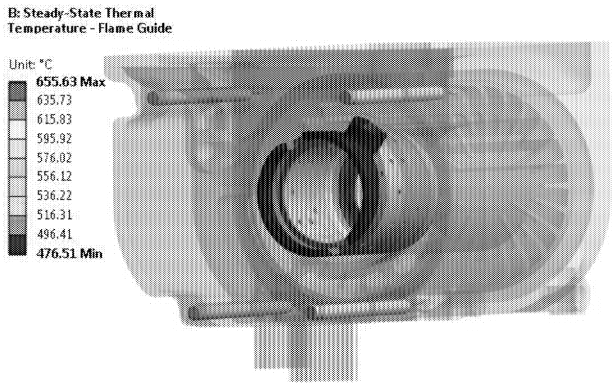


Figure 14: Temperature distribution on the flame support part

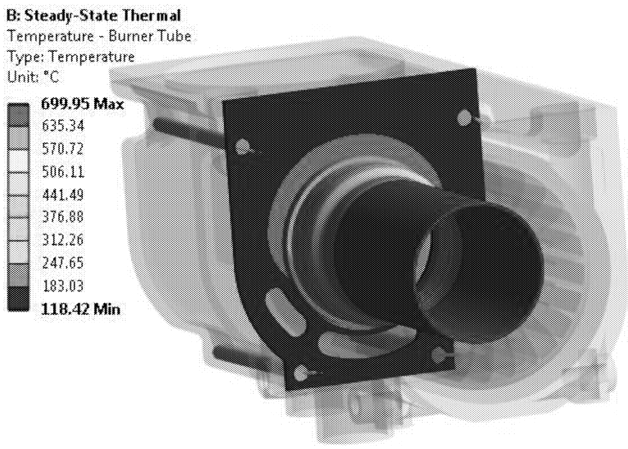


Figure 15: Temperature distribution on the burning tube part

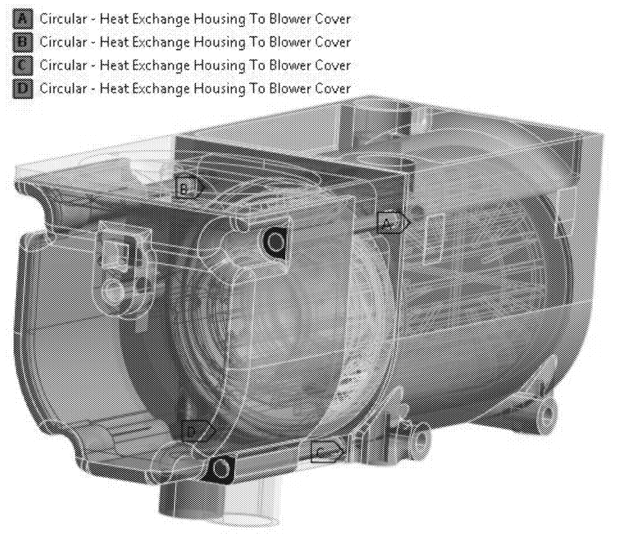


Figure 16: The four studs used for two parts connection modeled as beam elements

To remove the rigid body motion of the model and to simulate a real mode of fixity of the heater without introducing supplementary mechanical strain due to the fixed area, four points, with remote displacement condition, were used (Fig. 17). Supplementary, the maximum pressure 2.5 bar of the heat transfer agent, was included in applied static structural boundary conditions. Because there are no friction in the model, to keep the burner tube in central position four weak springs (Combin14 element type in Ansys) were introduced between burner tube and heat exchanger house.

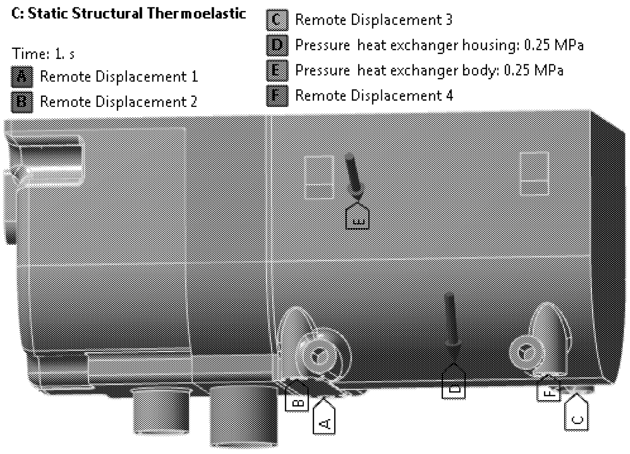


Figure 17: Position of the fixed points and the pressure load from the liquid in the heat exchanger

STATIC STRUCTURAL RESULTS

The reaction forces were very small (max 32 N) due to adequate fixing the model.

The total displacement distribution for thermo-elastic analysis for all components is presented in figure 18 where the legend is not uniform because the maximum displacement is inside of the heater. It can be seen where one fixed point on three direction was chosen (dark blue area).

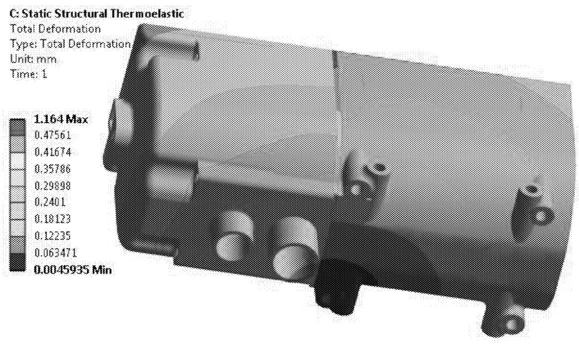


Figure 18: Total displacement distribution for thermo-elastic analysis

The axial deformation of the burner tube is around of 1 mm (Fig. 19) because the maximum temperature in the cylindrical tube is large.

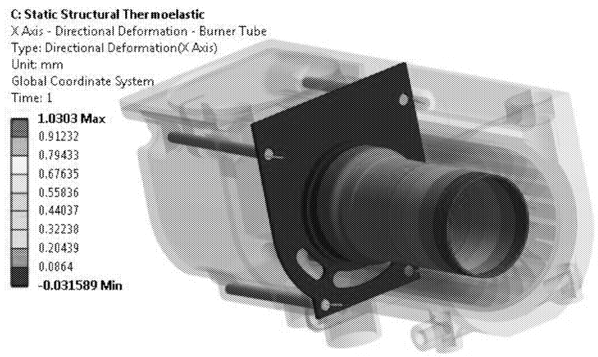


Figure 19: Axial displacements for burner tube

The maximum von Mises stress in the heat exchanger body is not very large, only some picks of 103 MPa can be observed (Fig. 20), whereas in the burning tube, due to the high gradient of the temperature, the maximum von Mises stress is around of 306 MPa (Fig. 21).

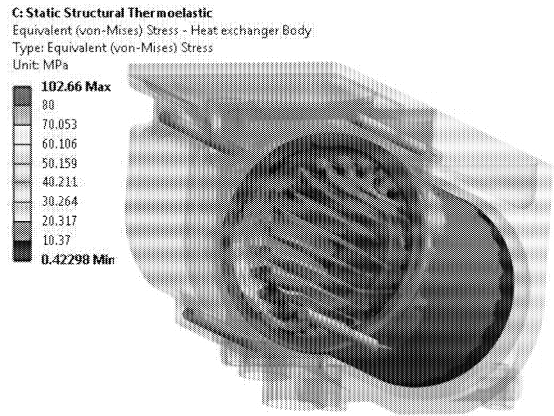


Figure 20: Von Mises stress distribution for heat exchanger body

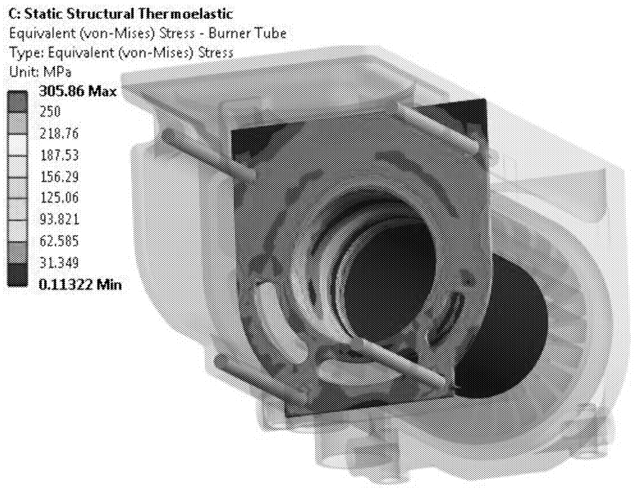


Figure 21: Von Mises stress distribution for burner tube

CONCLUSIONS

A very simple procedure, using finite element modeling and simulations, for temperature distribution in a water heater system is presented and the obtained results are in accordance with the current literature. Based on this temperature distribution, the stress distribution in the model is effortless to obtain. The presented methodology can be improved if some experimental data became available. Then the convection coefficients and bulk temperature can be considered spatially variables and an optimization module can be used for power balance of heat flow in the model. At this stage important results were obtained only by careful trials of different thermal boundary condition from the literature until the desired power output was reached (Jonathan et al. 2007).

REFERENCES

***ANSYS Coupled-Field Analysis Guide. 2004. ANSYS, Inc., Canonsburg.

Ashby, M. 2005. *Material Selection in Mechanical Design.* Butterworth-Heinemann.

Baukal, C. E. jr. 2000. *Heat Transfer in Industrial Combustion*. CRC, pp 65-117.

Cook, R. D., Malkus, D. S., Plesha, M. E. 1989. Concepts and Applications of Finite Element Analysis, 3rd ed., John Wiley & Sons, University of Wisconsin-Madison.

Gülder, O. L. 1986. Flame temperature estimation of conventional and future jet fuels. *J. Eng. Gas Turbine Power*, *108*, 376–380.

Heywood, J. B. 1988. “Internal combustion engine fundamentals”, New York: McGraw-Hill.

Illán, F., Alarcón, M., 2010. ”Numerical analysis of combustion and transient heat transfer processes in a two-stroke SI engine”, Applied Thermal Engineering vol. 30, 2469-2475.

Incropera, F. P., DeWitt, D. P. 2002. *Introduction to Heat Transfer*, Fourth Edition, Wiley.

Jafari, A., Hannani, S. K. 2006. ”Effect of fuel and engine operational characteristics on the heat loss from combustion chamber surfaces of SI engines”, Int. Communications in Heat and Mass Transfer vol. 33, pp. 122– 134.

Kays, W. M., Crawford. M. E. 2005. *Convective heat and Mass Transfer* McGraw-Hill.

Lefbvre, A. H., Ballal, D. R. 2010. *Gas Turbine Combustion - Alternative Fuels and Emissions*, 3rd ed.; CRC Press: Boca Raton, FL, USA.

Matarazzo, S. and Laget, H. 2011. *Modelling of the Heat Transfer in a Gas Turbine Liner Combustor*, MCS 7, Chia Laguna, Cagliari, Sardinia, Italy, September 11-15.

Serth, R.W. 2007. “Process Heat Transfer Principles and Applications”, Amsterdam: Elsevier.

Strakey, P. A., Eggnspieler, G. 2010. Development and validation of a thickened flame modeling approach for large eddy simulation of premixed combustion. *J. Eng. Gas Turbine Power*, *132*, 1–9.

Tuediger, G. Heat exchanger of an auxiliary heater. 2005. US Patent 6 932 151 B2.

Jonathan, Y., Douglas, Th., Faccone, P.A. 2007. Vehicle heater and controls here for. US Patent 7 270 098 B2.

WEB REFERENCES

www.webasto.com
www.parkingheater.com
www.eberspaecher.com
www.ansys.com

ACKNOWLEDGEMENTS

The work has been funded by the Sectorial Operational Program Human Resources Development 2007-2013 of the Ministry of European Funds through the Financial Agreement POSDRU/159/1.5/S/134398.

BIOGRAPHIES

FLORIN RADOI received a Bachelor degree in Automotive Engineering (2011), and a M.Sc. in Advanced Hydraulics and Pneumatics Systems, both from University POLITEHNICA of Bucharest. He started to work at Parker Hannifin since early college, and now is serving as application engineer in the field of complex hydraulics and pneumatics automation systems. He is also Ph.D. candidate within UPB, where he is developing R&D activity in the field of automotive auxiliary heating systems.

ŞTEFAN SOROHAN received the M.Sc. degree in Mechanical Engineering, in the Transportation Faculty, Railway Systems from the University POLITEHNICA of Bucharest in 1989. He became a Ph.D. in Finite Element Model Reduction in 2001 and now he is associate professor at Strength of Materials Department within the same university. His main activity consist in finite element modeling of structural, thermal and multiphysics problems.

NICOLAE VASILIU graduated in Hydropower Engineering from University POLITEHNICA of Bucharest in 1969. He became a Ph.D. in Fluid Mechanics after a research stages in Ghent State University and Von Karman Institute from Bruxelles. He became state professor in 1994 and a member of the Romanian Technical Science

Academy in 2008. He was always involved in solving complex concurrent engineering problems in fluid power systems.

CONSTANTIN CALINOIU graduated in Hydropower Engineering from University POLITEHNICA of Bucharest in 1975. In 1980 he became a member of the Hydraulic Laboratory from the Romanian Aerospace Institute. In 1998 he became Ph.D. and associated professor in the Fluid Power Laboratory from U.P.B. He is working mainly in modeling, simulation, and identification of the hydraulic and electro hydraulic control systems.

VR
BASED
ENGINEERING

SUPPORTING PRODUCT ENGINEERING BY TECHNOLOGIES OF VIRTUAL AND AUGMENTED REALITY

Prof. Dr.-Ing. Iris Gräßler
M.Sc. Patrick Taplick
Heinz Nixdorf Institute-University of Paderborn-Chair for Product Creation
Fürstenallee 11
D-33102 Paderborn, Germany
E-mail: Patrick.Taplick@hni.uni-paderborn.de

KEYWORDS

Product engineering, Virtual Reality, Virtual Prototyping, Virtual Engineering, Augmented Reality, Hololens

ABSTRACT

The article deals with the capabilities and utilization of Virtual and Augmented Reality technologies in the product creation process.
Starting with defining Virtual Reality (VR) and Augmented Reality (AR), possibilities and features of these technologies are illustrated.
The paper contains the previous integration of the VR and AR technologies in the process orientated approach of Virtual Engineering.
On the basis of broadly applied technologies, the potential of new AR technologies, in detail such as Microsoft Hololens, will be drafted. Several applications in conjunction to the V-model of mechatronic engineering development methology will be mentioned and explained.

INTRODUCTION

Engineers in the product creation process face new challenges in the development of innovative and complex products. Especially mechatronic manufactures have many domain specific requirements and properties of mechanics, electronics and informatics for example (VDI 2206 2004).
Because of the cross-disciplinary character of the development it increases the importance to structure and support the product creation process with approaches like Systems Engineering and Virtual Engineering.
Another important point is the increased competition in industries which focuses the product creation process on lower development costs and shorter time-to-market for complex products (Stark et al. 2010).
This aim can be reached by the substitution of physical by virtual prototypes which are developed with different software solutions like Computer Aided Design (CAD), Finite Element Method (FEM), Computational Fluid Dynamics (CFD) and Multi Body Simulation (MBS) (Mandić and Ćosić 2011). For the concept of Virtual Prototyping, the VR and AR technologies have a crucial role. With the help of these technologies several purposes like the visualization of the prototype for property validation can be handled. In the product creation process of complex products with many interdisciplinary parts VR and AR are necessary tools which help to check the status and progress of development continuously. The implementation of product behaviors and properties for the virtual product reduces the necessary amount of physical prototyping.
In automotive industry crash tests of virtual prototypes to validate the concept of the product with a minimal quantity of physical prototypes are used (Zorriassatine et al. 2003).
To establish VR/AR technology and approaches of Virtual Prototyping and Virtual Engineering in industrial development, this article shows technologies which are already used in the product creation process.
Especially the using of VR/AR technologies which includes all VR devices during the whole process of product engineering shows a high relevance for the actual and future processes in industries. Development potential is seen in a new AR visualization system, the Microsoft Hololens, which can support the product creation process by using it for communication, design and production.
Finally several application scenarios are described for the multiple usage of the Hololens. Thus the potential of this new AR technology is illustrated.

STATE OF THE ART

Virtual Reality

Virtual Reality is characterized by BURDEA and COIFFET as a real-time simulation using computer graphics. The aim is the creation of a realistic environment which should be generated by visual, auditive, tactile and olfactoric modalities (Burdea and Coiffet 2003).

The I³ model, which was generated by BURDEA and COIFFET, shows the important columns which define the concept of VR: Immersion, Interaction and Imagination (Figure 1).
Immersion describes the rate of the user inclusion in the virtual environment. Interaction is the possibility to interact in a virtual environment. Imagination means the ability to immerge the user into a virtual world.
Compared to other visualization technologies, VR is a real-time operating system. The expression real-time means the accomplishment to predict results and provide them in a short, defined and constant interval (Dörner et al. 2013). The virtual rendering is computed continuously with graphical algorithms.

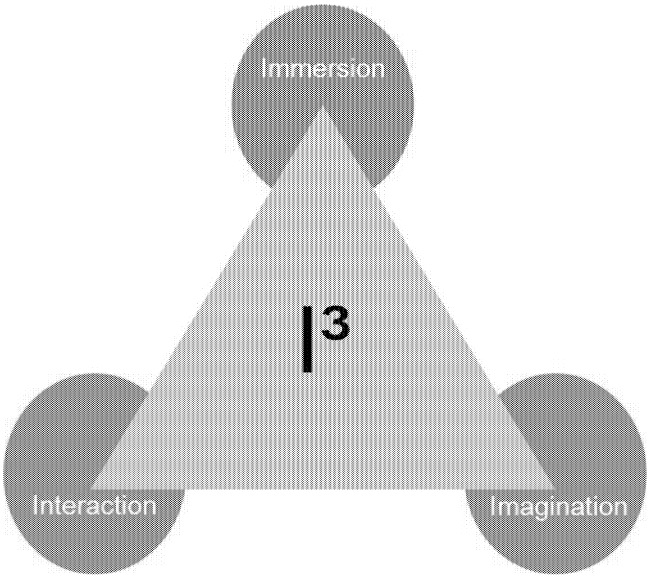


Figure 1: I³ model for VR (Burdea and Coiffet 2003)

Augmented Reality

Augmented Reality is characterized by AZUMA as a variation of Virtual Reality (Azuma 1997). Whereas VR immerses a user in a virtual environment, AR connects the reality with virtual objects which are embedded in it. AR systems are defined by (Azuma 1997):

- The combination of real and virtual world
- Interaction of the user in real world
- Registration in three dimensions

Thus, Augmented Reality is classified between reality and Virtual Reality (Figure 2). Another characterization of AR is "augmenting natural feedback to the operator with simulated cues" (Milgram et al. 1994).

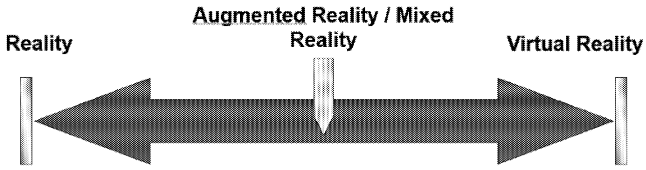


Figure 2: Overview of VR and AR (based on (Milgram et al. 1994))

BROLL categorizes AR in five steps (Dörner et al. 2013):

- Video shot: Streaming the reality with a calibrated camera
- Tracking: Computing the position and orientation of the virtual content
- Registration: Implementing virtual content in reality
- Visualization: Rendering the virtual content
- Output: Illustrating the virtual content on the screen

Following these five steps, an AR scene is implemented.

Virtual Engineering model

For a better organization and structure different approaches of Virtual Engineering were generated.
Virtual Engineering describes the early and continuous support of the product creation process. The support is characterized by coordination, assessment and concretization of development results with the help of virtual prototypes (Ovtcharova 2009).
The approach of OVTCHAROVA is an iterative process-orientated view (Figure 3).
Product Lifecycle Management (PLM) systems deal with the global product program. They include all variants and versions of a product in the definition and feedback phase. Changes of product properties are adjusted during the iterative process (Ovtcharova 2009).
Using CAx-software is important to create the product geometry, mechanical bearing structure and installation space by CAD in the creation phase and simulate functional structures and properties by FEM, CFD and MBS etc. in the analysis phase. The substitution of physical prototypes by deploying these IT solutions for virtual prototypes is called Virtual Prototyping. (Wang 2002)
Virtual Reality takes an important role in the virtualization and validation phase of the Virtual Engineering model.
The technology has variable usage possibilities in these two phases.
In the virtualization phase the main task is the creation of an immersive representation of the virtual prototype.
The virtual design review is a technique to present a product in a three dimensional display.

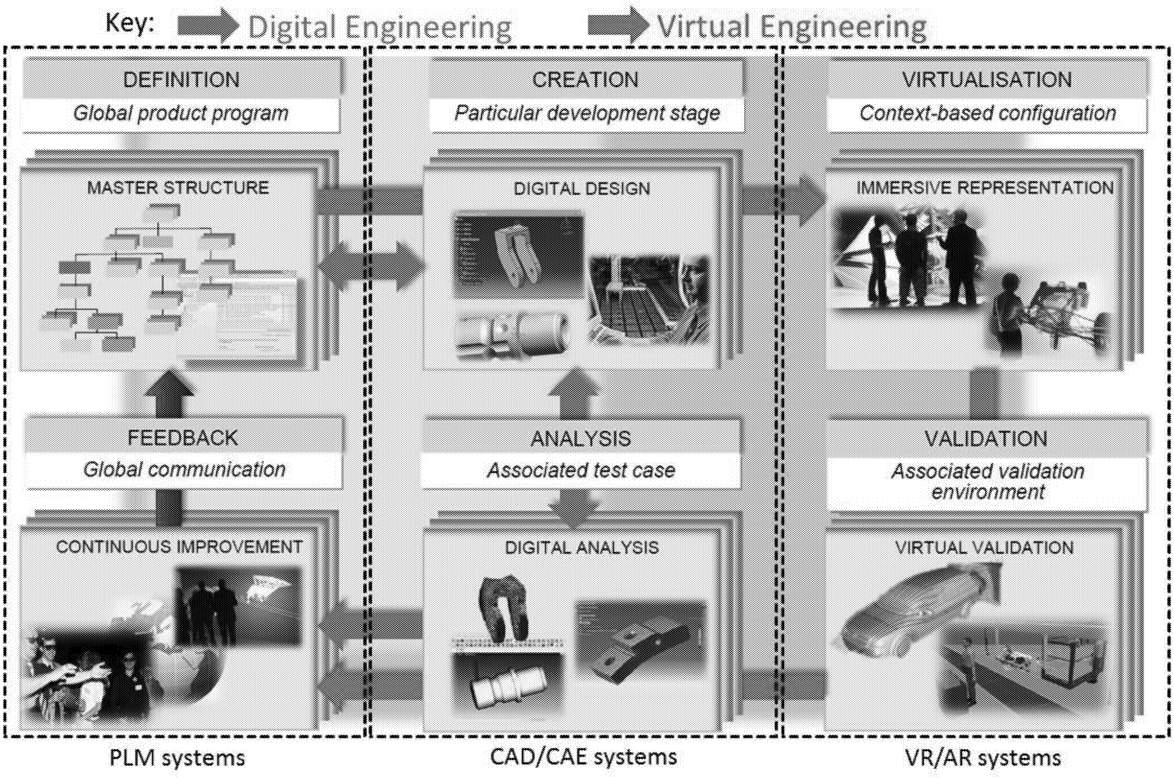


Figure 3: Virtual Engineering model (Ovtcharova 2009)

Objectives are the evaluation of design or the presentation of the different product variant alternatives and versions.
Geometrical properties are only important for a design review. It is useful to investigate surfaces, distances and clearance of the virtual prototype for an accurate design.
Another important technique which is used in the visualization phase is the ergonomics analysis with VR. With the investigation of ergonomics visual obstructions, interferences of comfort or other influencing problems can be identified. The space conditions of a car are affected by virtual human models for example.
In the validation phase simulations of the product and production processes are validated.
The interconnection of simulations like FEM, CFD and visualization technologies VR and AR is a useful possibility to show the results of the simulations (Barbieri et al. 2008). These outputs help to edit the structure of simulations in the analysis phase. After the completion of simulations the product idea has a designed geometry with assigned realistic properties like kinematics, dynamics and strength etc.
The cycle of the approach which includes all stages shows the iterative character. The developed product passes the product creation cycle several times until the maturity phase. Finally the approach displays the implementation of several IT tools like PLM systems, CAD/CAE systems and VR/AR systems in an iterative product creation process.

Virtual Reality visualization systems in the product creation

VR systems are used in the product creation process to visualize prototypes. The paper includes main systems like powerwalls, CAVE Automatic Virtual Environment (CAVE), multi-touch tables, VR and AR glasses. These systems are conventional tools in product creation process.
Powerwalls work with HD projectors which visualize 3D models on a screen. With the application of several HD projectors a stereoscopic illustration can be generated.

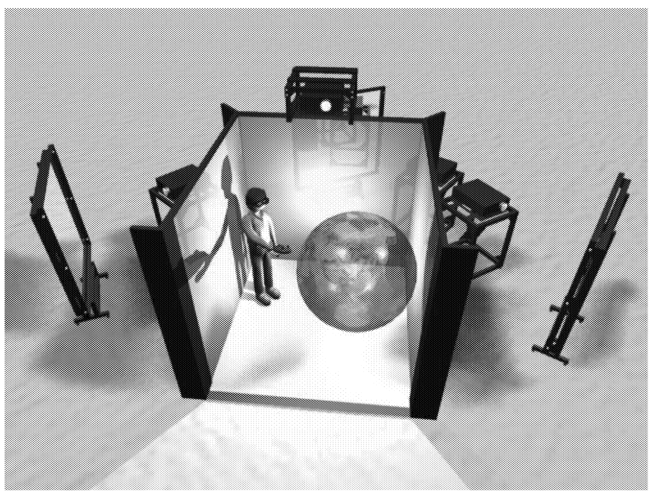


Figure 4: Structure of a CAVE (Jamstec 2007)

The tracking and interaction systems are variable for both VR systems. It can be used with head tracking or natural user interfaces by using a Microsoft Kinect as interaction possibility for example. The Kinect perceives three dimensional point clouds. The exact functionality of both systems are described in following literature (Jung et al. 2011) (Dörner et al. 2013).
Virtual tables are another visualization application. The conjunction of a 2D and 3D view is implemented in the system of virtual tables. The projectors visualize an object on the surface of the table. This application is an ideal possibility to plan a manufacturing process. Another visualization system for the product creation are VR and AR glasses (Figure 5). An established example is the Oculus Rift for VR glasses. It is characterized by a high rate of immersion which is generated by closed glasses within displays. The tracking system is based on sensors which are integrated in the glasses.

Figure 5: Google Glass (Wired 2015)

One of the first industrial application of AR glasses is produced by Google which combines transparent glasses with displayed information on it.

NEW POTENTIALS OF VIRTUAL REALITY TECHNOLOGIES IN THE PRODUCT CREATION

The previous section has shown the temporarily used VR and AR systems in the product creation process.
Certainly new VR/AR systems, which have the potential to support the product creation process, are in progress.
The most promising device for the establishment of Virtual Engineering and the penetration of the product creation process is the Microsoft Hololens which will be put on the market in 2015 (Microsoft 2015). The Hololens is an AR glass which is based on holographic and volumetric displays. After the economic flop of Google Glass the Hololens may provide a new chance of AR technology in engineering.
The technical details are not completely known yet. The hardware consists of a Central Processing Unit (CPU), Graphic Processing Unit (GPU) and a Holographic Processing Unit (HPU). The 3D objects are generated on transparent glasses. It supports stereophonic sound and includes a natural user interface which allows an operating with gestures. Another function is a voice control which permits verbal communication. 3D objects could be generated with the new Microsoft system software Windows 10. Holostudio is the required software of Windows to create 3D objects (Microsoft FAQ 2015).

USAGE OF THE HOLOLENS IN THE PRODUCT CREATION

After the short description of the new AR visualization system Hololens this article deals with different applications of the new AR device. For the classification of the scenarios the mechatronic V-model described in VDI 2206 is taken (VDI 2206 2004).
In Figure 6 the three application scenarios are assigned to the appropriate stages of the Product Lifecycle Management Backbone (Eigner et al. 2014).
The first scenario is a support of the Requirements definition and Product planning stages. It deals with the application in the field of global communication for development project groups. In world-wide operating companies the development departments of domain specific fields must communicate to create an interdisciplinary product. This communication is achieved by tele- and videoconferences. These possibilities are not useful for a visualization of prototypes.
The Microsoft Hololens provides the opportunity of communication for virtual development teams.
The implementation of virtual video displays on the transparent glasses of the Hololens are already shown in a commercial spot of Microsoft (Microsoft 2015). This application can be extended by the visualization of a virtual video display or an avatar for every project team member. With an avatar, which is a virtual projection of a person, the gestures and instructions of a project team member can be visualized clearly. Figure 7 shows an example of interaction between two persons. Additionally, the picture illustrates the visualization of instructions. Conjunctions of different AR glasses all over the world can create a virtual room within which all project team members can interact with each other. With the help of this interconnection of every Hololens everybody will see the same virtual prototype and interact with it. The two possibilities show the potential in the first two stages of the PLM Backbone. Product requirements could be discussed via virtual video displays and avatars with every project member. The product planning would be simplified by the visualization and interaction of the virtual prototype in real-time without being at the same place.
The second application scenario deals with the support of the Design stage in the PLM Backbone.

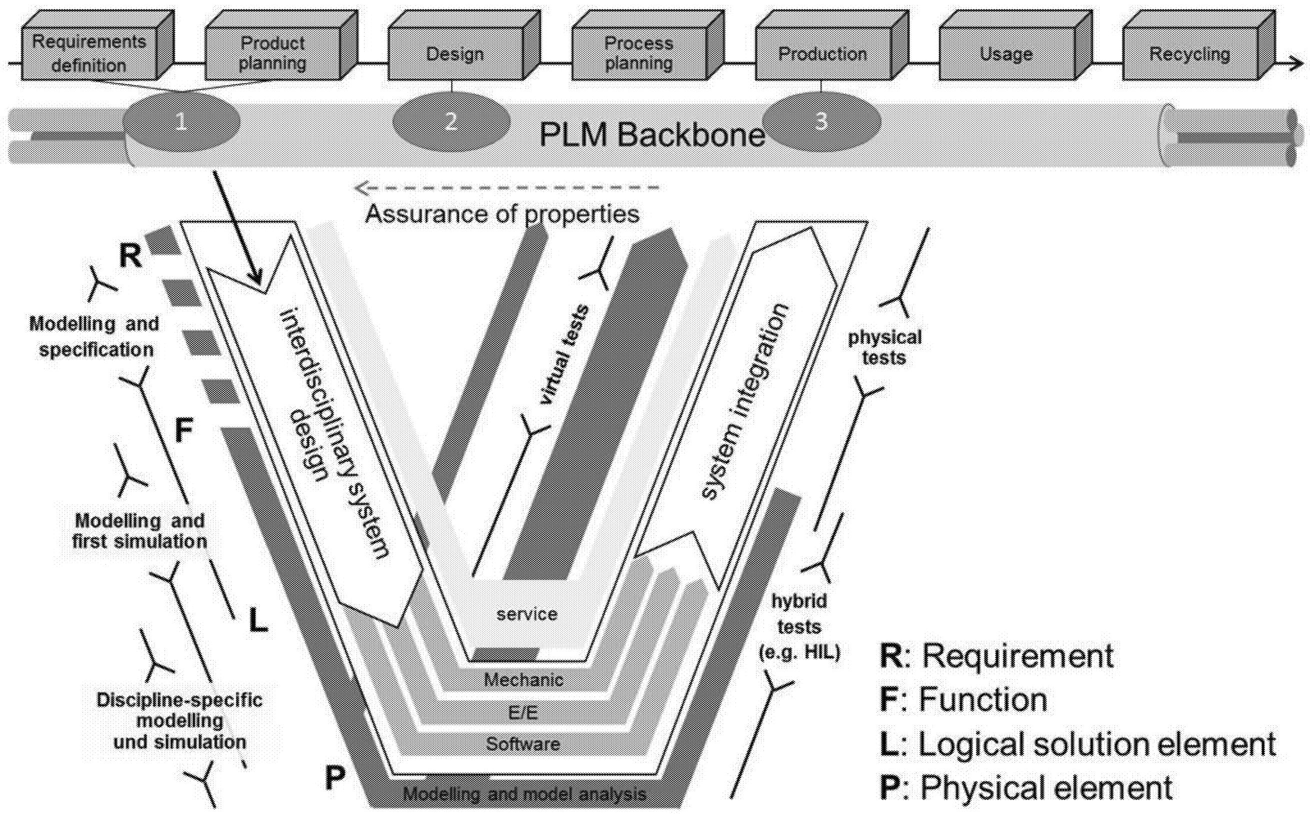


Figure 6: Applications of Hololens in V-model (based on (Eigner et al. 2014))

The design of a complex product is an interdisciplinary development (Peterson et al. 2007). Because of this fact different domain specific departments like mechanics or electronics are involved in the product creation process (VDI 2206 2004). These departments use different software applications. For example a mechanical engineer uses a CAD software to create the geometry of a product.
The results can be visualized on a normal PC screen or virtually on a powerwall or in a CAVE. For a perfect 3D visualization in the real environment an interface between development software and Microsoft´s Hololens can be generated.

These connections between the AR glass and the software applications can visualize a 3D illustration of the prototype while the engineer develops the product.
Figure 8 shows an example which Microsoft uses in a commercial spot (Microsoft 2015). In this picture the product is visualized by the Hololens.
This will be a solution for the presentations of simulation results. Especially CAE data like FEM or CFD can be displayed in a similar kind of way.
A connection of the natural user interface, the voice control of the Hololens and the toolbox of the used development software have high potential to simplify the product creation process.

Figure 7: Communication and interaction (Microsoft 2015)

Figure 8: Visualization of geometries (Microsoft 2015)

With this combination an engineer might change the tool of the CAD software verbally and draw a surface design with his finger in the air.
The third application scenario of the Design stage is a hybrid combination of virtual and physical prototypes. Figure 9 shows a user who adjusts a hybrid prototype of a motorcycle. This possibility has a high potential in automotive industries. Car manufacturers use a platform design for the chassis. For different car models new body works must be designed. A famous example for the combination of standard chassis and individual body works is the Volkswagen Group.

Figure 9: Hybrid of virtual and physical prototype (Microsoft 2015)

With the help of the Hololens the design of a body work could be created virtually on the basis of real standard parts by mechanical engineers. A special aid for the designing process is the natural user interface which is mentioned in the development software integration.
A combination of operating CAD and CAE software and creating hybrid prototypes by the Hololens take an important step to decrease the time-to-market.
The last scenario of the AR glass which is mentioned in this paper finds uses in the production stage.
In times of "Industry 4.0" the visualization of complex and variable production processes becomes more important. Using the AR glasses in the production process provides variable possibilities.
For example a construction of a new manufacturing factory can be visualized by the Hololens. The user regards an empty building and sees the complete production process by watching through the transparent glasses of the AR device. The natural user interface of the Hololens allows it to change the machine positions.
Another application possibility is to visualize the processing of the virtual prototype by real machines in the factory building. The production process of the virtual product could be simulated in this kind of way. This application has high potential to be a help to adjust the production process.
A user of the AR glass can use it to gain important information of the production process while walking through the shop-floor. For example specific values of different production machines can be visualized in the real environment by a created virtual information board.
A connection of production planning and control system (PPC system) and the Hololens might be a new usage possibility. The AR glass has the role to visualize the planned process. The interface of the software application and the AR device would be a presentation platform for people who are not experts for production planning.

CONCLUSION

The article shows the high potential which the new Augmented Reality device Hololens offers for the product creation process. The overview of the V-model, which is used for methodical product development of mechatronic systems, shows the usage of the Hololens in several fields of the product creation process.
It supports the Requirements definition and Product planning with a better interconnection of project team members all over the world.
The usage of Microsoft´s AR glass for the Virtual Prototyping and Engineering process would be interfaces between engineering software and the AR devices. The natural user interface and voice control are useful aids for development processes.
In production there are several usage fields like interfaces of PPC software or virtual constructions of manufacturing halls.
The mentioned application scenarios would decrease the development process of a product and production system. The development expenses would decrease because of the reduction of physical prototypes.
The low cost of the AR glass makes the application in all sizes of enterprises suitable. In comparison to a CAVE the AR glass is much cheaper. Probably the Hololens has a three-digit price, whereas the cost of a CAVE has a five- or six-digit value.
The lower cost of Hololens could accomplish a penetration of the visualization technology in product development departments of every industrial company independently of the enterprise size.

REFERENCES

Azuma Ronald T. 1997(4). "A Survey of Augmented Reality" *Teleoperators and Virtual Environments*, 355–385.

Barbieri L.; Bruno F.; Caruso F.; and Muzzupappa M. 2008. "Innovative integration techniques between Virtual Reality systems and CAx tools". *Int J Adv Manuf Technol* 38 (11-12), 1085–1097. DOI: 10.1007/s00170-007-1160-3.

Burdea G. C. and Coiffet P. 2003. "Virtual reality technology". 2. ed. Hoboken, NJ: Wiley. Online verfügbar unter http://www.loc.gov/catdir/description/wiley039/2002038088.html.

Dörner R.; Broll W.; Grimm P.; and Jung B. 2013. "Virtual und Augmented Reality (VR / AR). Grundlagen und Methoden der Virtuellen und Augmentierten Realität". Berlin, Heidelberg: Imprint: Springer Vieweg (eXamen.press).

Eigner M.; Roubanov D.; and Zafirov R. 2014. "Modellbasierte virtuelle Produktentwicklung". Springer Vieweg.

Jung T.; Krohn S.; and Schmidt P. 2011. "Ein Natural User Interface zur Interaktion in einem Cave Ein Natural User Interface zur Interaktion in einem Cave Automatic Virtual Environment basierend auf optischem Tracking". *3D-NordOst Tagungsband* 2011, 93–102

Mandić V. and Ćosić P. 2011. "Integrated Product and Process Development in Collaborate Virtual Engineering Environment". *Technical Gazette* Vol. 3 (18), 369–378.

Milgram P.; Takemura H.; Utsumi A.; and Kishino F. 1994. "Augmented Reality: A class of displays on the reality-virtuality continuum". *Proceedings of Telemanipulator and Telepresence Technologies* pp. 2351-34.

Ovtcharova J.; Awad R. 2009 "Virtual Engineering I". Course University of Karlsruhe

Peterson C.; Paasch R.K.; Ge P.; and Dietterich T. 2007. "Product Innovation for Interdisciplinary Design Under Changing Requirements". International Conference on Engineering Design, ICED'07. Cite des Sciences et de l'Industrie. Paris, France, 28.08.2007.

Stark R.; Krause F.-L.; Kind C.; Rothenburg, U.; Müller P.; Hayka H.; and Stöckert H. 2010. "Competing in engineering design—The role of Virtual Product Creation". *CIRP Journal of Manufacturing Science and Technology* 3 (3), 175–184. DOI: 10.1016/j.cirpj.2010.08.003.

VDI 2206. 2004. "VDI 2206 - Entwicklungsmethodik für mechatronische Systeme".

Wang G. 2002. "Definition and Review of Virtual Prototyping". *ASME Journal of Computing and Information Science in Engineering* 3, 2002 (Vol. 2), pp. 232–236.

Zorriassatine F.; Wykes C.; Parkin R.; and Gindy N. 2003. "A survey of virtual prototyping techniques for mechanical product development". *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture* 217 (4), 513–530. DOI: 10.1243/095440503321628189.

WEB REFERENCES

Virtual Reality System and Three-dimensional Visualization (18.05.2007) from:
http://www.jamstec.go.jp/esc/research/Perception.old/vr.en.html
Microsoft Hololens from:
http://www.microsoft.com/microsoft-hololens/en-us
http://www.microsoft.com/microsoft-hololens/en-us/get-ready
"Sorry, But Google Glass Isn't Anywhere Close to Dead" (08.02.2015) from:
http://www.wired.com/2015/02/sorry-google-glass-isnt-anywhere-close-dead/

BIOGRAPHIES

Prof. Dr.-Ing. Iris Graessler studied mechanical engineering and graduated in 1999 as a PhD (Dr.-Ing.) at Aachen University of Technology (RWTH Aachen), Germany. In 2003 she qualified as a university lecturer (Habilitation) at RWTH Aachen. For 14 years she has been working at Robert Bosch GmbH in several management functions in the field of Innovation Management, Product Engineering, Lean Manufacturing and Continuous Improvement Process. Since 2013 she is in charge of the Chair for Product Creation at the Heinz Nixdorf Institute at University of Paderborn.

Patrick Taplick was born in Salzkotten, Germany and went to University of Paderborn where he studied industrial engineering and obtained his master degree in 2014. After the study he worked for the Chair for Product Creation with the subject Virtual Engineering ever since.

A V-MODEL BASED COMPARISON OF SYSTEMS ENGINEERING APPROACHES

Iris Gräßler
Julian Hentze
University of Paderborn - Heinz Nixdorf Institute - Chair for Product Creation
Fürstenallee 11, 33102 Paderborn, Germany
E-mail: Julian.Hentze@hni.upb.de

KEYWORDS

Systems Engineering, mechatronic Systems, V-model, development methodology

ABSTRACT

In this paper the two core Systems Engineering approaches of the INCOSE handbook and the guideline of the U.S. Department of Transportation are compared with each other. The different steps of the two approaches are described and analysed up to the stage of implementation. This represents the left side of the mechatronic V-model, which both development approaches refer to in their guidelines. The paper shows the difference of the approaches in terms of granularity, context and points of focus. To have an illustrated model of the whole process, as in the U.S. Department of Transportation's guideline, makes it easier to understand and to follow the approach, but removes flexibility of planning the procedure for using operators.

INTRODUCTION

The methodology of Systems Engineering is an interdisciplinary approach and means to enable the realization of successful systems (INCOSE 2011). It integrates all the disciplines and speciality groups into a team effort forming a structured development process that proceeds from concept to production to operation (INCOSE 2011).
Systems Engineering is done by a variety of approaches and methods. First systematic users of Systems Engineering were the defence and aerospace industry in the 1940s. (Goode and Machol 1957) In 1962, Hall defined Systems Engineering as function with five phases (Hall 1962):

1. System studies or program planning
2. Exploratory planning, which includes problem definition, selecting objectives, systems synthesis, systems analysis, selecting the best system and communicating the results
3. Development planning, which repeats phase 2 in more detail
4. Studies during development, which includes the development of parts of the system and the integration and testing of these parts
5. Current engineering, which is what takes place while the system is operational and being refined

This approach of five phases was one of the first to describe the field of Systems Engineering for a commercial industry, in this case the telecommunication industry, aside from governmental organisations (Buede 2000, Hall 1962). Today there are a lot of different guidelines: First of all the handbook from INCOSE, the International Council of Systems Engineering. Additional to the INCOSE handbook there are other important guidelines, for example the handbook of NASA, the Systems Engineering Fundamentals of the U.S. Department of Defence or the Introduction for Transportation Professionals, published from the U.S. Department of Transportation. All the three examples are specific guidelines for Systems Engineers from the respective organisation, while the INCOSE handbook is a general and cross-industrial guideline (INCOSE 2011, NASA 2007, USDoD 2001, USDoT 2009,).
The goal of this paper is to introduce into two important approaches of Systems Engineering and to analyse the differences and the intersections. The most known and cited the INCOSE handbook and the guideline Introduction for Transportation Professionals from the U.S. Department of Transportation are chosen, because both allude to the V-model to explain the technical development methodology (USDoT 2009).
The paper will refer to the first steps of the development process and will end before the starting stage of the domain-specific design. This represents the left wing of the V-model. Both approaches refer to the V-model as an overview for the whole development process and explain the procedure with the illustration of the V-model (VDI 2206). The relation of both approaches to the V-model predestine them to find out differences and intersections between these two guidelines. The referred part of this paper is illustrated in the mechatronic V-model shown in Figure 1.

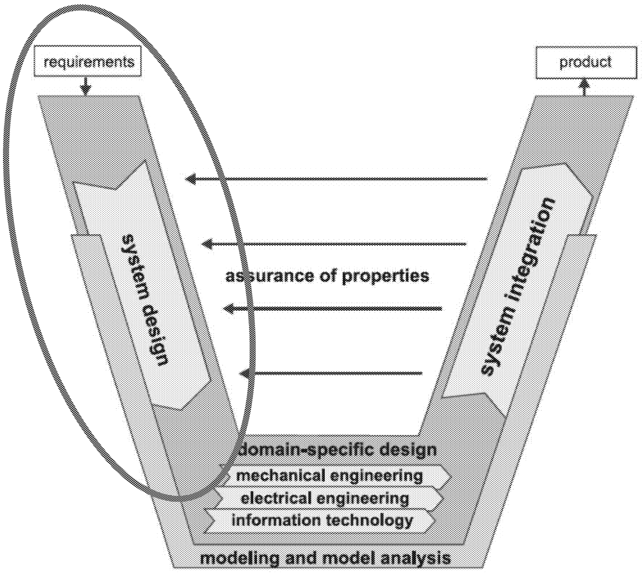


Figure 1: V-model from VDI 2206 with relevant red-marked part for comparison [cf. (VDI 2206)]

DESCRIPTION OF THE APPROACHES

INCOSE Handbook

The INCOSE Handbook v.3.2.2 dated October 2011 proposes the handbook scope, an overview of Systems Engineering and the explanation of the Generic Life-Cycle Stages. The V-model according to VDI 2206 is mentioned in chapter three as "a useful illustration of the Systems Engineering activities during the life-cycle stages" (INCOSE 2011). The following chapters provide a description of Technical-, Project-, Agreement-, Organizational- and Tailoring process activities performed by systems engineers. Chapter four, Technical Processes, starts with the explanation of the different process-stages. The most important processes for analysing and comparison in this paper, which are clear to be placed in the left wing of the V-model, are the first three described technical processes and supplemented to the development process the Project Planning Process, which is explained in chapter five, named Project processes, of the INCOSE handbook.

1. Project Planning Process
2. Stakeholder Requirement Definition Process
3. Requirements Analysis Process
4. Architectural Design Process

The four processes from the INCOSE handbook show no overall image of the left side of the V-model. INCOSE provides no fixed order or describes, which processes are necessary and which are not. Due to the technical based orientation of this paper, the technical processes are focused. The four stages correspond to the marked part in Figure 1. Every stage of the technical or project process is described with a context diagram. Figure 2 shows the structure of this diagram type. It contains the activities of the stage and fields of Input, Controls and Enablers, which have influences on the field of activities in the centre of the stage. Furthermore the Outputs are defined as the results of activities (INCOSE 2001).

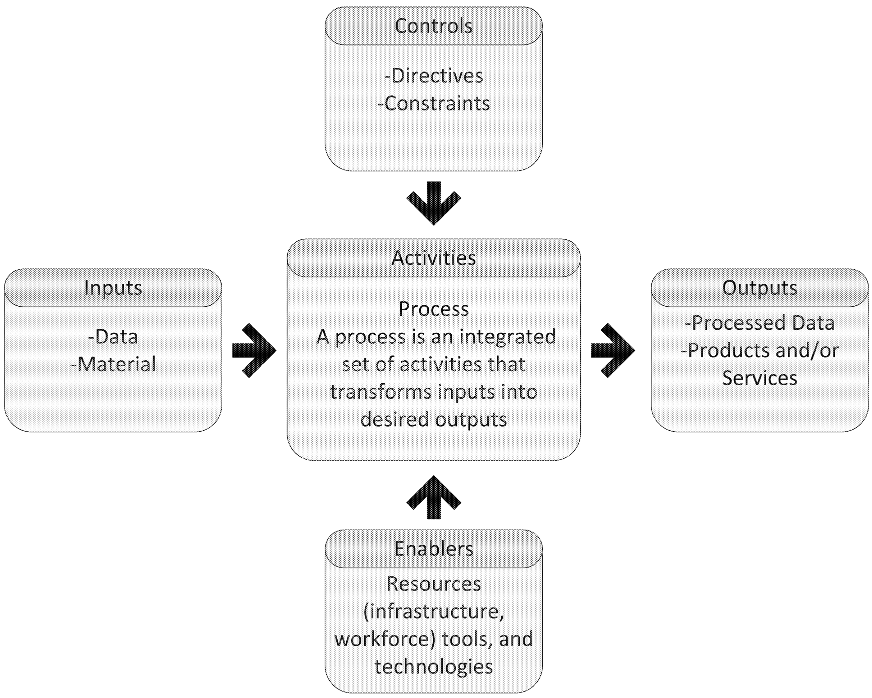


Figure 2: Context Diagram used in the INCOSE Handbook and the U.S. Department of Transportation guideline [cf. (INCOSE 2011, USDoT 2009)]

United State Department of Transportation guideline

In the first chapter of "Systems Engineering for Intelligent Transporting Systems" guideline published in 2009 the following two questions are answered:

- Why is it useful to deploy Systems Engineering?
- How is Systems Engineering defined?

In chapter three the mechatronic V-model is introduced and explained that additional to the original mechatronic V-model according to the VDI 2206 there are wings added "as a part of its adaptation for Intelligent Transportation Systems (ITS) to show how project development fits within the broader ITS project life-cycle" (USDoT 2009).

In chapter three the "ITS Life cycle Processes" are introduced. This procedure and individual V-model were developed especially for the U.S. Department of Transportation, based on the ideas of the V-model-XT (extreme tailoring), which was published in 2005 for software development (RaBr 2005).

In Figure 3 the seven stages are mentioned, which are observed in this paper:

1. Regional Architecture
2. Concept Exploration
3. Systems Engineering Management Plan Framework
4. Concept of Operations
5. System Requirements
6. Subsystem Requirements Project Arch (HLD)
7. Component Level Design (Detailed)

It is important to distinguish between stages and phases, which are both covered in Figure 3: Phases describe the time line of the development process. Stages are the substantive areas of work, which start and end with a decision gate.
Every stage is described in a similar way as in the INCOSE handbook by using context diagrams as shown in Figure 2 above. The seven stages concern the same part of the V-model as shown in Figure 1 and also illustrated in Figure 3.

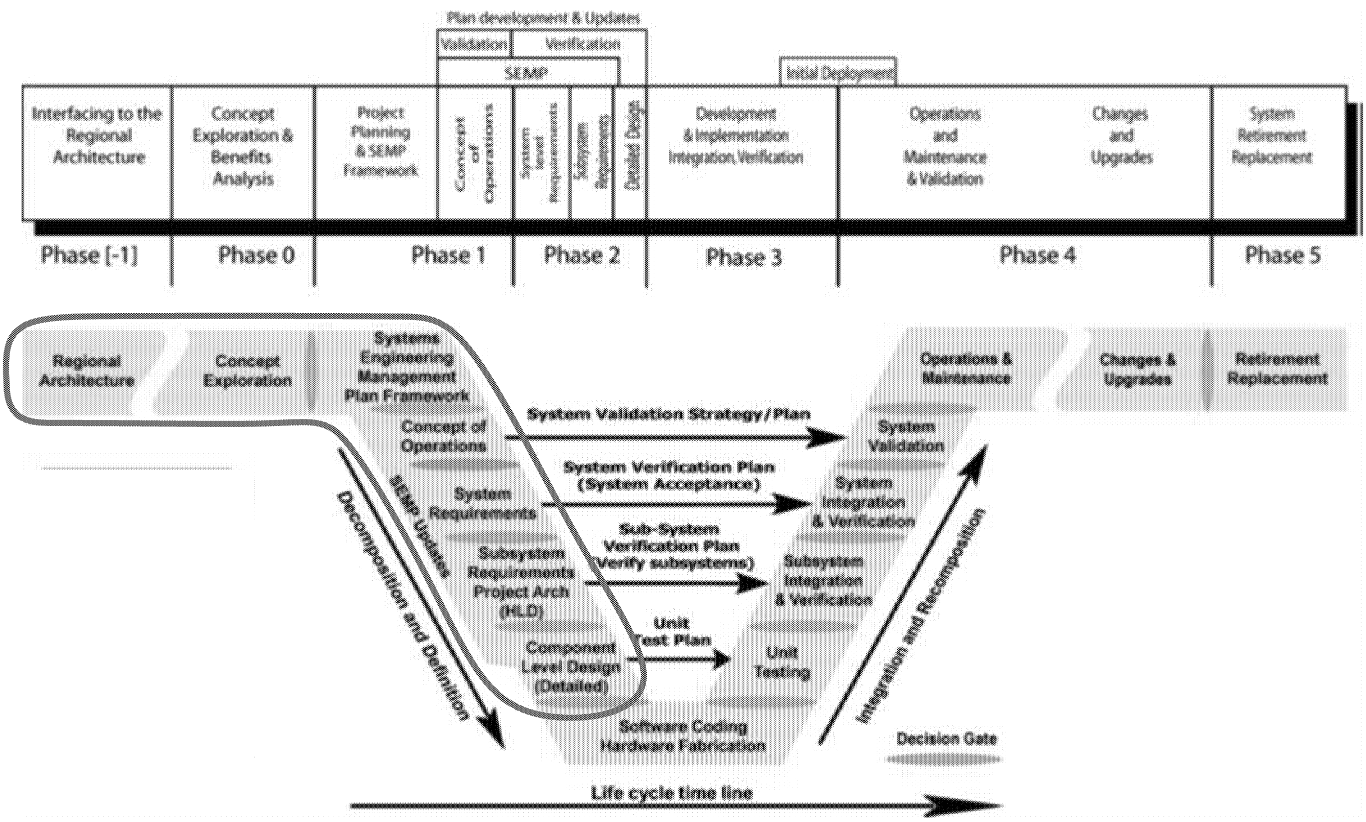


Figure 3: The U.S. Department of Transportations extended V-model with red-marked relevant part for comparison (USDoT 2009)

CONTENTS AND COMPARISON OF THE APPROACHES FROM INCOSE AND THE U.S. DEPARTMENT OF TRANSPORTATION

The first stage which is described in the U.S. Department of Transportation guideline is the regional architecture that deals with the status of regional infrastructure. This means to analyse the area of interest, which ITS-architecture is available, and to identify a portion of regional architecture. After this stage the V-model in Figure 3 shows a gap, which indicates that the first stage does not fit into the following development process. The gap is also described as an interruption in the timeline. Analysis of the regional architecture can take a very long period of time. This pre-development process is marked as phase (-1), that underlines the special early position of this stage. The project starts with the Kick-Off meeting after this phase. INCOSE does not refer to a timeline. The INCOSE handbook includes the analysis of the current status of a product or process to the stakeholder analysis in the development process this paper refers to later on.
The second stage of the U.S. Department of Transportation guideline is Concept Exploration. Important activities are to define goals and objectives of the project, identify constraints and alternative technical- and project concepts. In this "phase 0" the development process starts: The next stage Systems Engineering Management Plan Framework introduces the SEMP (Systems Engineering Master Plan), which contains timelines, schedules and resource planning. The SEMP contains far more than a project plan. SEMP compromises the approach of Systems Engineering which is necessary to manage and control the project in technical-, organizational and process aspects. Stage 4 of the U.S. Department of Transportation guideline, Concept of Operations, has the intention to create the ConOps (Concept of Operation) and the validation plan. Validation originally means checking the validity of a measuring method in empirical social research that is the extent to which test results actually register what is intended to be determined by the test. The ConOps describes the characteristics of a proposed system from the viewpoint of an individual, who will use that system. It is necessary to develop different application scenarios, define the vision of the project and develop and explore the ConOps. This stage concludes Phase 1 of Project Planning and Concept of Operations.
In comparison to the approach of the INCOSE handbook, these four steps are not described in the chapter of technical processes. INCOSE uses an extra chapter for Project Processes: This includes processes for Project Planning, Risk Management, Decision Management and other important management and organisational processes. The prime process for the comparison is the process of project planning, because this is equal to Phase 1 mentioned in the guideline of the U.S. Department of Transportation. One result of this process is the Systems Engineering Plan (SEP) also called Systems Engineering Management Plan (SEMP), which is equal to the Systems Engineering Master Plan (SEMP) from the U.S. Department of Transportation. The Project Planning Process defines the project, plans the resources and the technical- and quality-management.
The U.S. Department of Transportation guideline is much more detailed in the early phases of planning. The approach is clearly structured in one model. The INCOSE approach is divided into different categories. This paper refers to the two categories of Technical- and Project Planning. The extended V-model of the U.S. Department of Transportation shows the development process of the SEMP and the possibility to update the SEMP until the stage of Component Level Design (Detailed). ConOps must be prepared at the decision gate after Phase 1. INCOSE reduces the description to the process and defines no static positions for decision gates or milestones. These must be defined individually while creating the Systems Engineering Plan (SEP) for each individual project. This means, that the Systems Engineer has to define the sequence of the processes and define stages with a decision gate for beginning and ending.
In Phase 2, System Definition and Design, the following three stages of System Requirements, Subsystem Requirements Project Arch (High-Level Design) and Component Level Design (Detailed) are introduced. This part represents the three stages of the INCOSE handbook, which describe the technical processes chosen here for comparison. INCOSE integrates the process of stakeholder analysis, which is discussed in the U.S. Department of Transportation guideline in Phase 0 into the Stakeholder Requirements Definition Process. The INCOSE handbook includes no individual process to detect the stakeholders. Both approaches elicit, analyse and define the stakeholder's requirements. The INCOSE handbook adds the preparation of the validation plan to this stage of Stakeholder Requirement Definition Process, which connects the left- and the right side of the V-model. In the U.S. Department of Transportation guideline this is covered in the earlier stage of creating the ConOps.
The stages of System Requirements in the U.S. Department of Transportation guideline and Requirements Analysis Process at INCOSE define the preparation of the process for verification. Verification means generally demonstrating the truth of statements (VDI 2206). Both define and document the requirements of the system and determine concepts for traceability of requirements. This part of the development process describes the function, the interfaces and the specification of the system. The approach of the U.S. Department of Transportation underlines in this case, that it should be

specified “What the Systems should do”, and not “How the Systems do it”.
The following stage in the U.S. Department of Transportation guideline specifies the requirements of the subsystem. It provides a verification plan for the subsystem requirements in the same way as the systems requirements from stage before. Additionally the high-level Design of the system is developed, based on the requirements from the previous stages. The high-level design equals the overall framework for the system. At this stage the updating process of SEMP must be finished.
In the seventh stage of the U.S. Department of Transportation guideline, Component-Level Design (Detailed), the single components are defined in detail, based on the derived requirements from the overall system, reviewed and scrutinized. Goals of this stage, at the end of the process of decomposition and definition, are detailed design for the components, selection of commercial-off-the-shelf elements (COTS) and the plan to validate and verify every component. In the INCOSE handbook the Architectural Design Process consolidates high-level- and detailed design from the U.S. Department of Transportation in one stage. The activities do not distinguish from the approach of the U.S. Department of Transportation. The analysis, evaluation, definition and documentation of the architecture are most important parts of this stage.

ANALYSIS AND DISCUSSION

The most striking difference between the two approaches is the number of stages for the process until reaching the stage of implementation at the bottom of the V-model. The U.S. Department of Transportation guideline is more detailed and starts much earlier providing guidance in using the development approach of Systems Engineering. The U.S. Department of Transportation guideline illustrates an extended V-model to lead the user through the development process and to define stages with decision gates at the beginning and at the end of every stage. In the INCOSE handbook no overall process model is given. The INCOSE handbook contains the single processes that are due to realize Systems Engineering. The user of Systems Engineering approach according to INCOSE should be able to combine the technical and the project processes and develop an individual concept for each project.
Although the INCOSE handbook contains also project processes, it makes sense to integrate the planning process into an early stage of the development process. This recommendation leads to an oval process model again, that contains an exemplary overview for possible stages and phases, which refer to the aspect of time. This is realized in the guideline of the U.S. Department of Transportation by the illustrated extended V-model and helps to follow and to plan the Systems Engineering Project much more easily. INCOSE renounces an illustration or a clearly structured spreadsheet of the whole process. This would be an important and sensible improvement that helps to follow the described steps of the INCOSE handbook. This does not mean that there must be a model, which raises a generally valid claim. Much more should be attributed to an exemplary procedure or structure, as done in the life cycle stages in the previous chapter 3 of the INCOSE handbook.
Another important missing point as an inherent stage of the INCOSE approach is the analysis of stakeholders. The first process of INCOSE Systems Engineering approach are the technical processes with the definition of requirements. This stage includes the analysis of stakeholders and bases itself on information about stakeholders as an input of this stage. It is pointed out in the INCOSE handbook, that there is a high importance of requirements engineering, by dealing with requirements in two of the three relevant stages of the technical processes. But if the requirements and their description are undoubtedly very important, the detection of the right stakeholders is at least that important to come to the process of eliciting requirements.
Referring to the classic mechatronic V-model, published in VDI 2206, both approaches emphasize the early phases of planning and use the additional stages or processes outside the classic mechatronic V-model, often illustrated as additional wings. The classical idea of requirements as input and soon starting the design of the system, is not that intensive than the methodology of Systems Engineering needs. The phase of Design is only mentioned in two of the seven stages in the U.S. Department of Transportation guideline. At INCOSE one stage describes the design/architectural process. Summing up both approaches can be seen as an extended and modified V-model. In the U.S. Department of Transportation guideline the V-model is clearly illustrated and described by using a possible order of stages.
The idea of interfaces and restrictions from the right wing of the V-model is used in both development approaches. The validation and verification are important parts and are described in the results of the activities of different stages. The only difference is the adaption of additional verification processes, because of the different number of stages in guideline of the U.S. Department of Transportation.
The U.S. Department of Transportation approach illustrates a decision gate by the end of any stage. The INCOSE handbook does not describe such milestone positions, because of no illustrated approach model. The processes, which are described in context diagrams, include a decision management diagram. This is more flexible than in the U.S. Department of Transportation guideline, but does not deal with the positions and points in time, when it is useful or necessary to review the work

and the progress in a decision gate. The handbook does not refer milestones and decision gates to processes. By planning a Systems Engineering Plan (SEP) the important milestones and decision gates are to define in a way that it makes sense for the individual user and project.

CONCLUSION AND FUTURE PERSPECTIVE

The approach of the U.S. Department of Transportation is more detailed and specialised for the usage. Thereby the flexibility of the approach decreases but is simpler to follow and to understand for the operator. At the INCOSE handbook the operator has the options free to combine the technical, project and additional processes in an individual way. This could be useful, but contains additional work in planning the project. An illustration of a possible procedure could be helpful to understand better the INCOSE handbook and its objectives easier. This could be described as an abstract model or as an example. To develop a general concept of Systems Engineering this comparison of the left side from V-model shows many differences in their procedure. To generate an interdisciplinary approach for all users of Systems Engineering, it is necessary to compare the different approaches and to work out facts and discussable processes to implement Systems Engineering as the leading approach for the development of complex mechatronic systems. The comparison shows, that it would be important and helpful to define a degree of adaption to generalize the number of stages in Systems Engineering process.
Next steps of collecting information for a more general methodology of Systems Engineering would be to compare the bottom and the right wing of the V-model. Furthermore it is necessary to introduce additional approaches for example from NASA, the U.S. Department of Defence and guidelines from literature.
It should be a goal of this comparisons to create an overview matrix, exemplary shown for the comparison of this paper in Table 1, for all the approaches, assessed by defined criteria, for example the degree of granularity, the use of a model or generic significance. To create a consistent and global understanding of the ideas and objectives of Systems Engineering, it will be necessary to combine the existent approaches of different organisations and outline the strength and weaknesses of each guideline in order, to help the user to understand and to work with Systems Engineering.

Table 4: Exemplary matrix for comparison of Systems Engineering approaches. In this case the comparison between the INCOSE handbook and the U.S. department of Transportation guideline, representing the left wing of the mechatronic V-model

	INCOSE handbook v3.2.2	**U.S. Department of Transportation guideline v3.0**
Model-based	No	Yes, individual extended mechatronic V-model
Degree of granularity for necessary technical processes	Coarse, but many possible processes are described	Fine, clearly structured procedure and chronological order of processes
Milestone and decision gates	Yes, but individual to define and to plan	Yes, illustrated in V-model after every stage
Proposed timeline	No, individual planning for the project	Yes, phases in the V-model
Generic use	Yes	No, specific for ITS
Description of the methodology	Process-based	Model- and process-based, exemplary chronology

REFERENCES

(Buede 2000) Dennis Buede, "The Engineering Design of Systems: Models and Methods", John J. Wiley &Sons, 2000

(Goode and Machol 1957) Harry H. Goode, Robert E. Machol. "System Engineering". 1957

(Hall 1962) A. Hall. "A Mothodology For Systems Engineering", D. Van Nostrand Company Inc., 1962

(INCOSE 2011) International Council of Systems Engineering. "Systems Engineering Handbook v3.2.2", (Oct) 2011

(NASA 2007) National Aeronautics and Space Administration. "NASA Systems Engineering Handbook", NASA/SP-2007-6105 Rev1, (Dec) 2007

(RaBr 2005) A. Rausch, M. Broy, Das V-Modell XT „Grundlagen, Erfahrungen und Werkzeuge", dpunkt Verlag, 2005

(USDoD 2001) United States Department of Defence. Systems Management College, "Systems Engineering Fundamentals", (Jan) 2011

(USDoT 2009) United States Department of Transportation. "Systems Engineering Guidebook for Intelligent Transportation Systems", (Dec) 2007

(VDI 2206) Verein Deutscher Ingenieure(VDI) „Design methodology for mechatronic systems", Düsseldorf, (June) 2004

BIOGRAPHIES

Prof. Dr.-Ing. Iris Graessler studied mechanical engineering and graduated in 1999 as a PhD (Dr.-Ing.) at Aachen University of Technology (RWTH Aachen), Germany. In 2003 she qualified as a university lecturer (Habilitation) at RWTH Aachen. For 14 years she has been working at Robert Bosch GmbH in several management functions in the field of Innovation Management, Product Engineering, Lean Manufacturing and Continuous Improvement Process. Since 2013 she is in charge of the Chair for Product Creation at the Heinz Nixdorf Institute at University of Paderborn.

Julian Hentze studied mechanical engineering at the University of Paderborn and finished in 2014 with the master's degree. After one year as Junior Project Manager at the chemical company Flint Group Germany GmbH in Ahaus, he returned to Paderborn to work as research assistant at the Heinz Nixdorf Institute at the Chair for Product Creation of Prof. Dr.-Ing. Iris Graessler.

NEW MEDIA FRONTIERS

MEDIATED REALITY IN DENTAL TECHNOLOGY

Christian Steinmetz,
Alexander Christ, Reinhard Heister, Marco Grimm
and Reiner Anderl
Department of Computer Integrated Design
Technische Universität Darmstadt
Otto-Berndt-Str. 2,
D-64287 Darmstadt, Germany
steinmetz@dik.tu-darmstadt.de

Oliver Sandig
and
Christof Ellerbrock

C3System GmbH
Elisabethenstr. 20-22,
D-64283 Darmstadt, Germany

KEYWORDS
Mediated Reality, Dental Technology, Digital Dental Workflow, 3D Dental Models, Computer-Aided Vision and Design

ABSTRACT

Cost pressure, increasing quality requirements of dental products and the effects on the development and manufacturing of dental products demand efficient computer-aided methods and tools. The utilization of mediated reality solutions in dental technology opens up huge potentials for cost savings, automation of dental workflows and improvement of interoperability. The structural integration of electronics, sensors and actuators in the dental work area, digitalization of dental workflows and virtualization of dental products pave the way for a paradigm shift in dental technology and how dental technicians interact with dental products.
In this paper, a concept for the integration of mediated reality in dental technology is introduced. Based on the digital dental workflow, improvements for the context of use of dental technicians are proposed. The concept focuses on the utilization of visual and haptic feedback in the dental product design and manufacturing process. Elements of dental technicians' future work places are analyzed and validated, resulting in a 3D dental Computer-Aided Vision and Design (CAVD) system with haptic feedback and visual motion tracking.

INTRODUCTION

In today's process chain of denture creation, an indirect, plaster model-based procedure is applied. This procedure is characterized by media discontinuities and a low grade of individuality in denture design. This leads to high costs, long processing time, and unpredictable process uncertainties. Due to uncertainties in the dental design and manufacturing process, a consistent quality level can often not be ensured, resulting in dentures with low durability.

Improvements in dental technology and the integration of approved virtual methods enable the optimization of dental products and corresponding processes. Thereby it is essential, that the familiar work environment of dental technicians remains unchanged. Computer-aided approaches enable the fusion of real and virtual environment, resulting in a mediated reality for dental technicians (Barfield and Mann 2003).

In this paper, the concept for a CAVD system is introduced, which allows dental technicians to receive haptic and visual feedback while working on virtual 3D dental models. The CAVD system is characterized by sensomotoric movement patterns based on dental technicians handcraft skills. It enables the fully digital design of dentures with a high grade of individuality and avoids the shortcomings of conventional denture design and manufacturing. The underlying digital dental process chain is described hereafter.

DIGITAL DENTAL PROCESS CHAIN

Three main classes of dental process chains exist to develop patient individual dentures (Heister and Anderl 2013):

- ChairSide
- LabSide
- CenterSide

The *ChairSide* process chain is characterized by inserting the individual denture in just one session. So the patient is not required to leave the dental chair during the treatment. This is achieved by using an intra-oral 3D scanner and an integrated development framework in order to reduce media discontinuities. With this technology only small scaled, low-complex dentures are producible, e.g. crowns, inlays and veneers (Mörmann 2006).

The most common process chain is the *LabSide*, as shown in Figure 1. By using this approach, almost every type of indication, every available material and every complexity grade is producible. The first step is the preparation of the damaged tooth. Afterwards, a dental cast is taken by the dentist and the plaster model is made by the dental technician. In the next step a media discontinuity occurs by the use of a 3D scanner for digitalization of the model. Thereby, the working method changes from physical into virtual. Based on the digitalized patients' jaw, the denture can be designed by a specialized dental CAD (Computer Aided Design) system. As output a CAD file in tessellated form (stored in the Standard Tessellation Language, STL format) of the reconstruction is generated. This file is processed by the use of a Computer Aided Manufacturing (CAM) system for production in a subtractive (e.g. milling) or additive (Selective Laser Melting, SLM) manufacturing process. Subsequently, a further media discontinuity back from digital into physical working methods appears. To finish the denture, a veneering is manufactured by coating the metal with ceramic. This is manually done by using a brush, ceramic substance and furnace. To ensure functionality and accuracy of fit, the denture is validated with an articulator. Finally, the denture is sent back to the dentist. In a dental treatment, the denture can be finalized and inserted (Beuer et al. 2008; Strietzel and Lahl 2007).

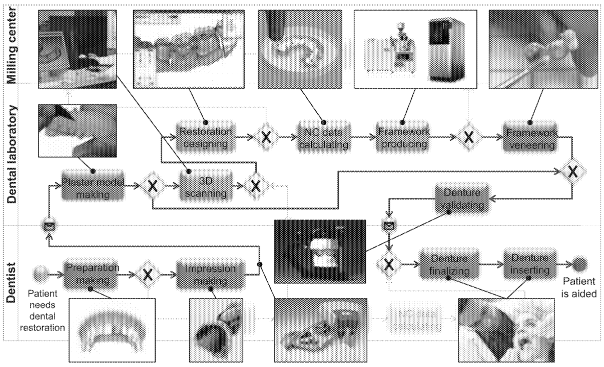


Figure 1: LabSide Process Chain (Heister and Anderl 2013)

The last possibility to walk through the dental process chain is the *CenterSide* method. All steps are equal to the LabSide method, but only CAM processing and manufacturing process are outsourced to a specialized milling center. This paper focuses on the LabSide process chain.

USE CONTEXT OF THE DENTAL TECHNICIAN

Historically, the work of dental technicians is characterized by manual work. The first time digital technologies came up was in the 1970s. In the 1990s, digital dental technology had experienced a breakthrough due to the invention of the CEREC I system. Ever since, the adoption of digital dental work methods has been increased continuously. The advantages of these digital technologies are enhanced productivity, high accuracy as well as improved reliability (Strietzel and Lahl 2007).

Despite all the advances, that have been developed, there are still some disadvantages in the use context of dental technicians:

- No 3D visual experience
- No 3D input interaction
- Lack of physical feedback during interaction
- Lack of digital working methods in education
- Reduced acceptance of non-intuitive digital systems by dental technicians

Today's CAD systems provide the development of 3D parts in a virtual 3D environment. However, the presentation of the 3D representation is rendered on a 2D output device, e.g. a display. Common displays are not able to display virtual 3D parts in an autostereoscopic manner. Therefore, it is challenging to anticipate actual position and orientation of the shape manipulation features during work. Furthermore, accurate assessment of the dental design regarding functional and esthetical requirements is often insufficient (Jobst and Ellerbrock 2013). The human-machine-interface (HMI) of a dental CAD system is shown in Figure 2.

Figure 2: HMI of Dental CAD Systems

Besides the shortcomings in presentation of the virtual model, there is a lack in usability related to the user input. In contrast to intuitive manual movements in the physical denture shaping process, CAD-based denture design is most commonly performed by the use of a 2D computer mouse. The 2D position of the mouse cursor has to be mapped into 3D coordinates. Here, it is difficult to identify the precise 3D position through the 2D cursor, too.

In traditional manual work dental product designers receive direct and immediate feedback. By using digital design technology, designers only receive visual or acoustic feedback. Haptic feedback, comparable to the manual work, is completely missing. Especially for the design of high-accuracy freeform shapes, an immediate and intuitive feedback would contribute to an improved digital design process (Ellerbrock and Kordass 2011).

The German dental technician education is mainly characterized by traditional manual development methods. Digital work methods still play a minor role. As a result, newly qualified dental technicians have to be educated in digital CAD/CAM applications at the dental laboratory. Additionally, sensomotoric movement, which is well-trained during education, is not utilized when using state-of-the-art digital development methods.

CONCEPT FOR THE INTEGRATION OF MEDIATED REALITY IN DENTAL TECHNOLOGY

The integration of approved computer-aided methods and tools offer huge potentials for dental technology. The digitalization and automation of manual processes improve efficiency for the development of dental products and ensure a defined quality level (Christ et al.

2014). A key enabler is the suitable integration of virtual 3D dental models in the digital dental design workflow. 3D dental models allow dental technicians to interact with the virtual representation of real dental products in a virtual 3D environment and pave the way for the further integration of virtual solutions.

For dental technicians the visual and haptic feedback from dental products is very important during the dental manufacturing process. Virtual 3D dental models enable the integration of established solutions like virtual reality applications and haptic devices. The combination of those approaches leads to a mediated reality for dental technicians. There are several approaches for the integration of mediated reality into dental technology. They range from haptic feedback devices to the utilization of camera systems as digital stereomicroscopes and the structural integration of sensors, actuators and electronics in dental input devices.

The innovation of the approach introduced in this paper is the user experience of virtual 3D dental models through the entire development and manufacturing process of dental products. A direct und fully digital procedure is applied to avoid media discontinuities and to increase individuality in denture design. The main objective is the deployment of a Computer-Aided Vison and Design (CAVD) system. Sensomotoric movement patterns based on the dental technicians handcraft skills are used for the interaction between dental technician and CAVD system and for the processing of the 3D dental model.

Architecture of the CAVD System

The proposed concept enables the interaction with 3D dental models and the superposition of virtual and real environments. The concept consists of three modules:

- Haptic feedback system,
- head-mounted display and camera, and
- dental input device.

The *haptic feedback system* enables dental technicians to receive a realistic tactile behavior during the interaction with the 3D dental model. In order to achieve a realistic functioning in the manufacturing of dentures, a pen-based haptic feedback system is used. The pen of the haptic feedback system represents the tools of the dental technician. The pen tip is represented in the CAVD system as cursor position. If the cursor touches the surface of the 3D dental model, a force is applied by the haptic feedback system. Haptic devices are usually intended to interact with 3D models without modification. In this concept, the modification of 3D dental models is one of the main functions of the CAVD system. The modification of the model geometry can be distinguished into operations that add material and remove material.

The *head-mounted display and camera* enables the application of virtual reality and augmented reality methods in the CAVD system. On one hand, scaling of 3D dental models is achieved in the form of a digital stereomicroscope. On the other hand, the display allows a superposition of additional information on 3D dental models by using see-through functionality. Additional information is assigned to the 3D dental model and aligned with the current view. The deployment of the head-mounted display and camera enables dental technicians to interact with dental models in a visual manner. The processing of 3D dental models in a pure virtual environment as well as in an environment that combines virtual and real objects is possible. The system can be used for education and training as well as in daily work for manufacturing and quality control of real dentures.

The *dental input device* is used by dental technicians for modeling of dentures. Usually dental input devices are tongue-shaped plastic plates without further functionality. In the denture manufacturing process, material is added to and removed from these plates. Current dental input devices do not fulfill the requirements of the CAVD system. Therefore, a structure integrated digital dental input device has been developed. Structure integrated sensors enable the continuous determination of position and orientation of the dental input device. For dental technicians the dental input device can be seen as black box: the outer shape remains unchanged, while the digital functionality is achieved with integrated electronics. Thus, the way how dental technicians work stays the same, but is efficiently supported by the virtual modeling system.

Integration of the CAVD System in Dental Workflows

The CAVD system includes the integration of electronic devices, particularly sensors and actuators, into the daily work of dental technicians. In order to integrate the electronic devices to the virtual modeling process, a computer-internal representation of the patient's oral situation is required. Therefore, the oral situation has to be digitalized, as described above in the LabSide process chain. The digitalization is realized by using a 3D scanner, e.g. an intraoral scanner. From the 3D scan a 3D dental model is derived.

The dental technician uses this 3D dental model as basis for the denture modeling process. Standardized 3D teeth models can be loaded from teeth libraries. These teeth models represent relatively coarse geometries that have to be adapted to the individual patient's oral situation by the dental technician. The current procedure is characterized by a variety of manual working tasks on physical objects, time-consuming error corrections and many media discontinuities (Heister and Anderl 2014). This can result in dentures that do not fit correctly and that have a low durability due to overstressing.

The CAVD system enables the digitalization of working procedures of dental technicians, which supports the concept of "cyberizing the physical" (Lee 2010) of Industrie 4.0 (Anderl 2014). Thus, the dental manufacturing process is based mainly on virtual methods. Starting from the 3D dental model, the precise design of dentures can be performed in a pure virtual 3D environment. Scaling of 3D models and overlaying of additional information without interruption of the working process are deployed by the head-mounted display and camera. Besides the visualization, a haptic feedback can be retained by the haptic feedback system. The digitalized dental input device completes the CAVD system as enriched platform for the virtual projection of the 3D dental models, resulting in a mediated reality in form of a smart environment for dental technicians.

Media discontinuities are minimized, while error corrections can be performed more efficiently. Functionality is improved by the individualized design and configuration of the dental technician's tools in the virtual environment. The digitalization leads to time savings in the manufacturing process and to a cost reduction due to material savings compared to the physical dental design process.

Mediated Reality via Multi-sensual Interaction

To bring virtual dental models to a physical user experience, the partial systems

- haptic feedback system,
- head mounted display and camera,
- dental input device, and
- an additional computer with dental modeling software

are integrated to a CAVD system. Therefore, interfaces between these modules have to be harmonized.

The dental input device as well as the haptic feedback input device need to provide high sensor sampling rates for seamless and direct movements of the dental modeling tools. Input delay or lag lead to unnatural movements and modeling errors. To enable fluent and accurate motion, it is important to guarantee low latency transmission with high data rates. To cover the whole 60 full frames per second display output, at least 120 Hz input and output sample rate is recommended.

The modeling core of the CAVD system works with triangulated meshes for virtual constructions. For high accuracy in the dental industry high model resolutions are required. Common virtual dental models contain over 15.000 vertices and 30.000 faces per indication, e.g. a single crown model has a binary file size of about 2 MB (see Figure 3). To prepare the denture model, a 3D scan, stored in the STL file format, is loaded in a conventional dental CAD software and based on the individual situation, a standard denture is generated. Common dental CAD solutions provide a set of denture templates for each required indication and tooth. For modeling of a crown, the specific tooth is selected and the crown is generated by the difference of the existing tooth stump and the tooth template. From the predefined standard denture, the fine modeling is done in the CAVD system. To do so, the pre-designed denture is exported in triangulated form (STL) and loaded in the CAVD software. The data transfer between the systems is done without media discontinuities.

To perfectly fit the patient's dental situation, the pre-designed denture has to be detailed manually. The goal is to fit the occlusion and preparation borders between denture and tooth stump for best functional, hygienic and esthetical results.

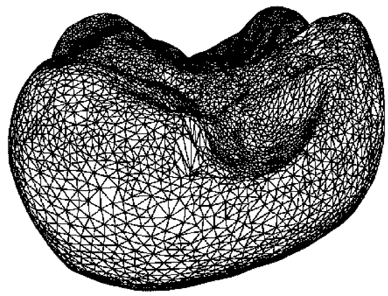


Figure 3: Triangulated Dental CAD Model

The CAVD integrates the linkage-based and pen-type haptic feedback device for both user input and tactile feedback. The input for 3D tool positioning with six degrees of freedom (DOF) is achieved by reading the encoders in the device joints and recalculating the input pen's angular and linear positions. The robotic arm's own weight and introduced momentum is compensated via a closed-loop control integrated in the device to minimize the actual tool weight for the dental technician and make modeling work less straining. The feedback is provided by putting linear force on the end joint of the pen when the virtual modeling tool is touching a surface of the denture model. If the tool is pushed further into the restraining surface, the force of the pen is linearly increased until a maximum value is exceeded. The increase of perceived force is achieved by applying torques to the motors at each joint. With this, forces in all 3D directions can be represented. In practice, a professional haptic feedback device is used for the proposed concept, and its high-level application programming interface (API) is used to cover both input and output streams. See the following section for details on the prototypical implementation.

In accordance to the common manual work the CAVD system is used with both hands. While the haptic feedback system is used with one hand, the dental input device is used with the other hand. The dental input device controls the virtual denture model by acquiring a live spatial position and orientation of the device and maps it to the virtual model in real time. On the model, the virtual presentation of the denture model is projected in a 3D environment. When the input device is moved or tilted, the model is moved and reprojected accordingly. The shape of the device is designed to match the shape of the dental modeling plate that dental technicians are used to. This traditional jaw-like shape provides the best functionality while being ergonomic for the particular dental modeling task. The dental input device does not provide a tactile feedback as the haptic feedback device. The tip of the haptic feedback system modeling pen is not intended to touch the dental input device as the virtual model is projected in between both devices.

The head mounted display and camera is the main module to achieve the mediated reality experience of the overall CAVD system. It is directly connected to the modeling computer, where the CAVD software is running on. The device is held on the head by a head mounting strap, so that the user is looking into the two displays. As displays, high-resolution near-to-eye micro displays are used for vision. The real workplace view is captured by wide angle cameras that are mounted in front of the display. The visual input of the cameras is then projected to the displays and overlaid with the virtual models. In order to detect the dental input device orientation, machine vision algorithms are used to complete the augmentation and gain accurate distance coordinates between viewer and dental input device. The scheme of the overall CAVD system is shown in Figure 4.

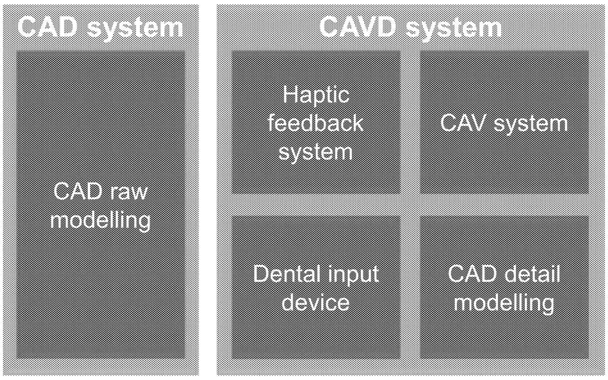


Figure 4: Scheme of the CAVD System

The mediated reality experience is then provided by the combination of the input and output devices to deliver a multi-sensual interaction with the user. The near-to-eye displays provide a natural, first-person visual perspective to the technician on the workplace. The virtual workplace is projected on the real workplace in a way that the augmentation provides a visual feedback to the user about the working process. This visual experience is supported and enriched by the tactile response provided by the haptic feedback system. The dental input device allows natural movements of the model in the virtual environment so that the overall CAVD system provides an immersive and intuitive way of dental work.

PROOF OF CONCEPT

In the proposed operating concept (see Figure 5), the dental technician uses the pen of the haptic feedback device and the dental input device as tools to modify the virtual data that overlay the real scene in the CAVD system (see Figure 6). The real scene is therefore captured by two similar cameras and projected on two similar displays. Both cameras of the CAVD system are connected to the computer via USB 3.0 interfaces and are able to capture images in 1920x1080 pixels (Full HD) resolution at a frame rate of 30 frames per second. The displays are able to show images in 1280x720 pixels (HD) resolution and are connected to the computer via DisplayPort interfaces. The dental input device communicates with the computer via a wireless interface and sends its position and orientation about 120 times per second over this connection. Further, the haptic feedback system with a sampling rate of 1000 Hz is connected over a FireWire (IEEE 1394) interface.

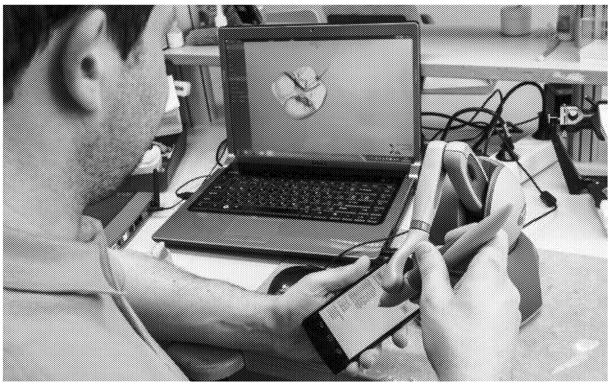


Figure 5: Proposed Operating Concept

The whole optical and mechanical system is optimized for a working distance (distance between the dental technicians head and the working scene) of 300 mm. The computer is thereby the main processing unit of the whole system. While the user or technician can see the live-view of the real scene, a 3D file in STL format is loaded into a 3D viewer based on the Open Graphics Library (OpenGL) and overlays the real scene. Most CAD systems only show the rendered virtual scene in a parallel projection. For a real time manipulation of 3D data, it is advantageous to render a scene with perspective projection. Therefore, a simple and transparent 3D viewer based on OpenGL has been developed in order to overlay the camera images from different points of view, see Figure 6. In addition to these 3D renderings, 3D annotations or other relevant information can be overlaid with the camera images.

The haptic feedback device has six degrees of freedom, a translational resolution of 0.055 mm and a maximum force feedback force of 3.3 N (SensAble 2008). The force feedback is computed and applied by the main processing unit. Therefore the actual position and orientation of the haptic feedback device is compared with the bounding box, nearest point distance and geometry of the loaded CAD model.

Figure 6: Overlay of Digital 3D Scan Data

For various tools of dental technicians different surface properties are applied by calculating a special and individual force feedback. Thereby, it is possible to simulate the haptic behavior of actual und future real and virtual tools. By a basic magnification and almost lossless digital zoom technicians have a huge scope and many possibilities to modify the virtual data. Figure 7 demonstrates the interaction of a dental technician with an early prototype of the CAVD system.

Figure 7: Early Prototype of the CAVD System

In first tests, the cameras have shown a latency/delay of approximately 200 milliseconds between capturing and displaying images which has to be optimized in future studies. The system has also a very good depth of field. The whole jaw model appears sharp in a wide focus range. Further practical tests and developments will show, whether the proposed operating concept is close to traditional manual development methods and can help dental technicians design and manufacture individual dentures.

CONCLUSION AND OUTLOOK

In this paper, the benefits and the necessity of digital dental technology are described. The technology allows vast improvements in quality and development time based on current physical dental modeling techniques. This provides a substantial business potential as well as increased health-care value for patients. However, the complex use context of dental technicians requires comprehensive digital integration to match the flexibility and physical experience of dental technicians work.

As a first step to achieve an integrated virtual environment for dental technicians and to leverage the digital modeling possibilities, a concept for a CAVD system is proposed. This system is based on a mediated reality setup with intuitive input devices and a tactile feedback through the haptic feedback system. The different system elements in combination with the CAVD software provide an immersive modeling experience. Due to the first-person point-of-view, the dental technician's natural work context remains unchanged. To demonstrate and prove the usability and functionality of the proposed concept, a prototypical implementation that combines all modules of the CAVD system has been developed and discussed.

Future research will focus on the extension of the CAVD system as well as on the optimization of the current virtual environment. For example, actuators to provide tactile feedback should be integrated into the dental input device. Furthermore, additional information over the modeling process can be integrated into the user's view through augmentation.

ACKNOWLEDGEMENTS

This project (HA project no. 434/14-24) is funded in the framework of Hessen ModellProjekte, financed with funds of LOEWE – Landes-Offensive zur Entwicklung Wissenschaftlich-ökonomischer Exzellenz, Förderlinie 3: KMU-Verbundvorhaben (State Offensive for the Development of Scientific and Economic Excellence).

LOEWE Exzellente Forschung für Hessens Zukunft

HessenAgentur HA Hessen Agentur GmbH

REFERENCES

Anderl, R. 2014. "Industrie 4.0 - Advanced Engineering of Smart Products and Smart Production". Technological Innovations in the Product Development. 19th International Seminar on High Technology, Piracicaba, Brazil.

Barfield, W. and S. Mann. 2003. "Mediated reality". International journal of human-computer interaction Special issue 15, 2003, 2. Lawrence Erlbaum Assoc, Mahwah, NJ, ISBN: 9780805896046.

Beuer, F.; J. Schweiger; and D. Edelhoff. 2008. "Digital dentistry: an overview of recent developments for CAD/CAM generated restorations". Br Dent J 204, 9, 2008. S. 505–511.

Christ, A.; R. Heister; and R. Anderl. 2014. "Reverse Engineering for the Utilization of Product Model Data for CAx Workflows in Dental Technology". 20th European Concurrent Engineering Conference, ECEC'2014 and the 10th Future Business Technology Conference, FUBUTEC'2014. EUROSIS-ETI, Bruges, Belgium, 2014.

Ellerbrock, C. and B. Kordass. 2011. "Comparison of computer generated occlusal surfaces with functionally waxed-on surfaces". In: Int J Comput Dent., 14(1), 23–31.

Heister, R. and R. Anderl. 2013. "Federative Data Management Based on Unified XML Data Scheme to Support Prosthetic Dentistry Workflows". ASME 2013 International Mechanical Engineering Congress & Exposition (IMECE2013), San Diego, CA, USA.

Heister, R. and R. Anderl. 2014. "Concept for an Integrated Workflow Planning of Dental Products Based on Federative Data Management". ASME 2014 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. Buffalo, NY, USA.

Jobst, T. and C. Ellerbrock. 2013. "Ist Okklusion mit CAD/CAM überhaupt möglich?". 42. Jahrestagung der Arbeitsgemeinschaft Dentale Technologie e. V., Böblingen.

Lee, E. 2010. "CPS foundations". In: Proceedings of the 47th Design Automation Conference (DAC). ACM/IEEE, Anaheim, CA, USA, ISBN: 978-1-4503-0002-5, 737-742.

Mörmann, W. H. 2006. "State of the art of CAD/CAM restorations". 20 years of CEREC. Quintessence, London, Chicago, ISBN: 9781850971641.

SensAble Technologies. 2008. "PHANTOM Omni Specification". Geomagic.

Strietzel, R. and C. Lahl. 2007. "CAD-, CAM-Systeme in Labor und Praxis". Neuer Merkur, München, ISBN: 3937346414.

BIOGRAPHIES

CHRISTIAN STEINMETZ studied Computational Engineering at Technische Universität Darmstadt from 2008 to 2011. In September 2013, he completed the master program Computational Engineering at Technische Universität Darmstadt. Since October 2013 he works as research assistant at the Department of Computer Integrated Design of Technische Universität Darmstadt. His research activities focus on web services and Industrie 4.0.

ALEXANDER CHRIST earned his diploma in mechanical engineering at University of Applied Science Giessen-Friedberg in 2009. As student he worked at GM/Opel in the area vehicle simulation. From 2009 to 2011 Mr. Christ studied mechanical and process engineering at Technische Universität Darmstadt and achieved the Master of Science degree. In June 2011 he joined the Department of Computer Integrated Design of Technische Universität Darmstadt as a research assistant. His research activities focus on the development of methods and tools for CAD and PLM in automotive, aerospace and dental technology.

REINHARD HEISTER studied from 2003 until 2006 Reinhard Heister Mechanical Engineering at the Duale Hochschule Baden-Württemberg in association with the company KARL MAYER Textilmaschinenfabrik GmbH. Subsequently he worked for three years as a project manager in machine development and was responsible for the development of several serial machines. From 2009 until 2011 he studied Mechanical and Process Engineering at TU Darmstadt. His master thesis dealt with the optimization of human-machine interface between the air traffic controller and the air traffic control system at the Deutsche Flugsicherung GmbH. Since October 2011 he is a research assistant at the Department of Computer Integrated Design (DiK).

MARCO GRIMM earned his master's degree in mechanical and process engineering at Technische Universität Darmstadt in 2012. In March 2012, he joined the Department of Computer Integrated Design at Technische Universität Darmstadt. Ever since, he is also engaged at the excellence center CASED (Center of Advanced Security Research Darmstadt). His fields of interest are production systems security, knowledge protection in engineering and connected systems for virtual product development.

OLIVER SANDIG studied mechanical and process engineering at Technische Universität Darmstadt from 2008 to 2014. In September 2014, he completed the master program mechanical and process engineering at Technische Universität Darmstadt. Since September 2014 he works as research assistant at C3System GmbH.

REINER ANDERL earned his diploma in mechanical engineering in 1979 and his Dr.-Ing. in mechanical engineering in 1984, both at Universität Karlsruhe. From 1984 to 1985 he served as technical manager of a medium-sized company. Dr. Anderl received his academic habilitation in 1991 and the venia legendi in 1992, and accepted the call for professorship for Computer Integrated Design at the faculty of mechanical engineering at Technische Universität Darmstadt. Dr. Anderl was dean of the faculty of mechanical engineering from 1999 until 2001, when the bachelor and master programs were implemented. He is member of the Zentrale Evaluierungs- und Akkreditierungsagentur (ZEvA), a national accreditation agency, the PACE manager of Technische Universität Darmstadt (Partners for the Advancement of Collaborative Engineering Education), an adjunct professor at Virginia Tech, and a full member of the Academy of Sciences and Literature, Mainz. Dr. Anderl has authored and co-authored more than 230 publications.

CHRISTOF ELLERBROCK earned his diploma as a dentist in 1999 and the doctorate (Dr. med. dent.) in dentistry at the University of Jena in 2000. Before studying dentistry he studied mechanical engineering at the University of Braunschweig from 1990 -1992. 2001 he became a partner in a dental office in Idar-Oberstein where he specialized in treatment of "craniomadibular diorders". 2007 he started the master course "Treatment of craniomadibular diorders" at the University of Greifswald, where he got his master degree in 2010. 2009 he started another master course at the Danube University of Krems in the field of "Interceptive orthodontics". Here he achieved the master degree in 2012. Since 2012 he owns a dental clinic Darmstadt, where he also founded the dental laboratory "Form for Function GmbH". In 2014 he founded "C3System GmbH". He is holder of several patents, international lecturer and international publications.

CLOUD COMPUTING AND (NEW) MOBILE STORYTELLING IN THE INTERNET OF THINGS

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

KEYWORDS
Cloud Computing, Storytelling, Internet of Things, Digital Humanities, Design Fiction, Human-computer interaction

ABSTRACT

The increasing weight of user-chooser participation in Cloud Computing and the Internet of Things (IoT) urges the making of conventions, terminologies and ontologies that may frame their usage in particular in what concerns tagging the objects of IoT. This labelling will be – is already – surpassing its mere cataloguing mission, being used to convey watever narratives the operator may want to associate them with. New tools give way to novel artistic practices, and a new aesthetics seems to be in the making. The trend towards Mixed Media demands a transdisciplinary dialogue between its performers and Digital Humanities' practitioners. This paper intends to summarize some of the issues arising from the deficiency of conventions, and proposes – to be theoretically explored in future work – the hypotheses of including these novel storytelling practices into the field of Design Fiction.

INTRODUCTION

Artists, scientists and engineers are presently living the same «Desktop Revolution» that affects all fields of knowledge, a Kuhnian epistemological paradigm change already echoing in every human area. The Internet evolution allowed the development of Digital Humanities, embracing the new set of tools available, the new relationships between producers and users, interactivity and networking. The users' number crescendo allows an anticipation of the need for clearer definitions of the utensils that are being used, and the possibilities they offer.

This paper intends to be a speculation over the main issues raised by the lack of clear terminology and ontologies, natural of a period when new expressions, artistic and scientific, are in the making. Giving some practical examples of projects to illustrate some of the issues, it will propose – as a hypothesis to be developed in future work – that these new storytelling strategies do not fit within the existing literary conventions and categories, not even the ones developed by interactive fiction models, demanding to be included into the equally novel idea of Design Fiction.

NEW PRACTICES FOR A NEW PARADIGM

The model of the artist-engineer embodied by Leonardo da Vinci is not confined to the Renaissance. All artists, at all times, used their ingenuity as engineers, exploring the various scientific fields that might interest them to help themselves to the latest discoveries, findings, the most recent advances (even proposing some hypotheses) in the very physical and scientific area of materials, and media used to manufacture their works (Barbas, 2011). In that same way engineers and scientists in their practices are not devoid of aesthetical judgement (MacAllister, 1998). Presently, neurological studies of the human brain activity, through fMRI imaging experiments used to evaluate reactions to beautiful objects (paintings, musical excerpts) are starting to prove that the location of the brain's aesthetical functions and reactions are identical in the human subjects tested (Ishizu 2011).

Artists, scientists and engineers are presently living the same «Desktop Revolution» (Johnson, 1977) that affects all fields of knowledge. This is an epistemological upheaval, attesting the formation of a new scientific paradigm (Kuhn, 1979), which is already echoing in human production, whatever the arena. Nowadays, no one can imagine working without a computer.

However, the tasks and usages are naturally different: «The humanistic turn of mind provides the historical perspective, interpretative skill, critical analysis, and narrative form required to articulate the significance of the scientific discoveries of an era, show how they change our sense of what it means to be human, and demarcate their continuity with or difference from existing ideologies» (Davidson 2008).

In the new Digital Humanities Era, Internet evolution has been carefully followed, and mainstream computer timelines and terminology duly adapted. Tim O'Reilly's distinctions (O'Reilley 2005) between Web 1.0 (around 1991-2000) and Web 2.0 (from 2001 onwards) gave way to Humanities 2.0 as a practice embracing the new set of tools available, the new relationships between producers and users, interactivity and networking. All this echoing also in the artistic expression.

One of the first experiences in Portugal is «Alletsator XPTO – Kosmos.2001» (http://po-ex.net/alletsator) by Pedro Barbosa and Luis Carlos Petry, an opera libretto on electronic text synthesized in computer: «For actors, musicians and other animals».

Figure 1: Alletsator XPTO – Kosmos.2001, P. Barbosa, L. C. Petry, opera libretto - text synthesized in computer

A more recent example - theme of a poster presented at ACM WomEncourage 2014, and embodying a mixed media exhibition in Lisbon City Hall - is the series «Cloud Portraits of 22 Portuguese poets» by Alexandra Antunes-Couts (http://www.alexandraantunes.com/2012/11/herberto-helders-typographic-portrait.html), giving continuity to her MA poetry projects: «Máquina de Emaranhar paisagens» (https://sites.google.com/site/maquinadeemaranharpaisagens) and «Divertimento com Sinais Ortográficos» (https://sites.google.com/site/funnypunctuationmarks/home). The author uses specific software for word frequency which combined, generates word clouds from each poet lexicon, titles and texts, graphically executed with inDesign:

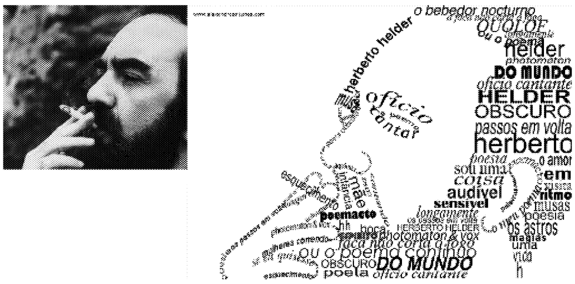


Figure 2: Herberto Helder (2012) A. Antunes Typographic cloud portrait 297×210 mm. Portugal

Consequently, the gradual democratization of access to data combined with a change in the approach from top down to bottom up, was detected. Humanistic users also inherited system administration tasks that «forced them to develop their own IT skills and sensitivities» (Yanosky 2008). This need to learn new tasks, paired with the liberty to choose software and platforms (and to adopt the ones sanctioning a steep learning curve), challenges the idea that users do not know what they want. With the increase of new user-chooser centred participation it has become harder for big IT departments and corporations to control what to offer and to whom. This issue is growing along with the anticipated number of global users:

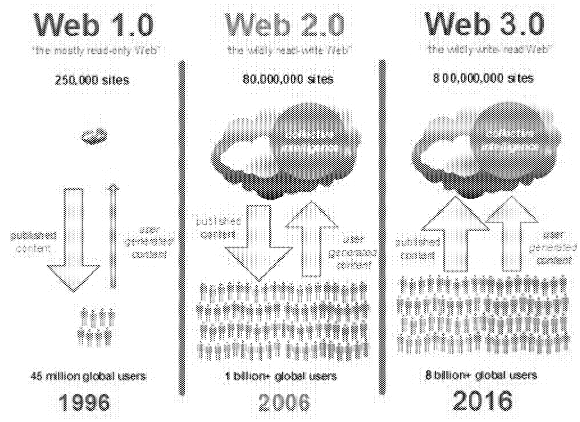


Figure 3: 2016 Expectations for web development

Within these changes and due to their novelty, other questions arise concerning terminology and ontologies - namely the expressions used in the title: Cloud Computing, mobile storytelling, and Internet of Things - for which there are several definitions available, and not everyone agreeing with them.

CLOUD COMPUTING

Cloud Computing is a scientific metaphor for the Internet and a standardized cloud-like graphic shape used to denote a network with the implication that the specifics of how its end points are connected is not relevant for the purposes of understanding the diagram.

For Humanists it immediately recalls the comedy by Aristophanes - where he ridicules the great Greek philosophers. The Clouds (*Nephelai*) were also the nymphs bearing water to the heavens in cloudy pitchers, and nourishing the earth and rivers with their rains. So, only good things can be expected from the present internet-public-clouds with its elasticity of, allegedly, free resources.

Recently, and in accordance with NIST regulations (http://www.nist.gov/itl/cloud/index.cfm), it is generally considered a rich itinerant computing technology that leverages unified flexible assets of wide-ranging Clouds and network structures toward limitless functionality, storage, and mobility. It will serve: «a multitude of mobile devices anywhere, anytime through the channel of Ethernet or Internet regardless of heterogeneous environments and platforms based on the pay-as-you-use principle» (Sanaei 2014).

These enthusiastic definitions are balanced by counter-proposals such as the one presented by the design research IIC - Inhabiting and Interfacing the Clouds (http://www.iiclouds. org), that object to the current expression of Cloud Computing and big proprietary solutions - Cisco, Dropbox, Google Drive, Amazon S3 or CE2 - particularly in forms intended for private individuals and end users. ICC considers that these implementations have chiefly followed the logic of technical development, corporate interests, and are paradoxically envisioned as a purely functional, centralized setup. They (Keller 2015; Nova 2014) defend the concept of Personal Cloud or OpenCloud. Their alternative model implies a «cross-disciplinary approach that links interaction design, the architectural and territorial dimensions as well as ethnographic studies». This matter is still under discussion.

Open, Closed or Hybrid, Cloud allows user access through networked client devices and mobile collaboration via wireless, cellular, broadband technologies, sanctioning working and collaborating independently of space and time.

Users-choosers access Cloud computing via desktop computers, laptops, tablets and mainly smartphones. The latter, as Cloud clients, rely on Cloud computing for a great majority of their applications, becoming essentially useless without them.

MOBILE EMPOWERMENT

The advances in Cloud, wireless and mobile computing, together with the abilities of mobile devices, reached a point where compute-intensive applications can now run on mobile devices with limited computational capabilities. This is attained by using the communication power of mobile devices to establish high-speed connections to vast computational resources located in the Cloud.

This access and potential has originated a boom in the development and creation of Applications (Apps) for handheld devices, which are pre-installed, or delivered via web by request, to perform all imaginable tasks.

As an example, there are two Apps developed by CADA (http://www.cada1.net) a Lisbon-based art group that makes mobile software to explore «the interrelationship between the variable character of human beings and the spatial-temporal dimensions they create and negotiate in everyday life». These Apps are for ambient and personal data visualization.

«Today» (2007) - (http://www.cada1.net/works/today-visualize-your-calls-version-2-for-android/) - converts the amount and type of daily phone calls made and received by the user into a readable graphic. Each call or message generates a specific symbol which appears chronologically in real time. The result is an ever-changing, evolving image.

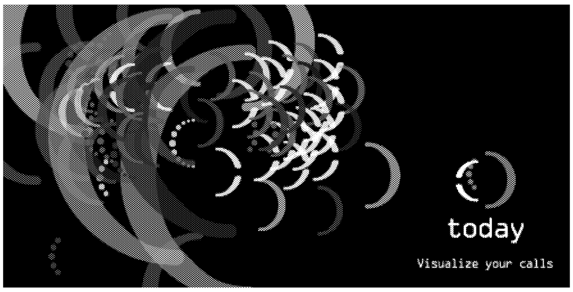


Figure 4: Today – self-tracking data visualisation App (2007) S. Oliveira, J. Hawkey, S. Prudêncio, CADA, Lisbon, Portugal

The second example - «Time Machine» (2007-2009) (http://www.cada1.net/works/timemachine/) - is the outcome of a project developed with Academia participation – CADA and CITI-FCT New University of Lisbon. It converts personal emotional and temporal patterns into a visual form, exploring colour. The output is a data visualization interface displayed on the phone, designed to produce a highly personal topology of time usage.

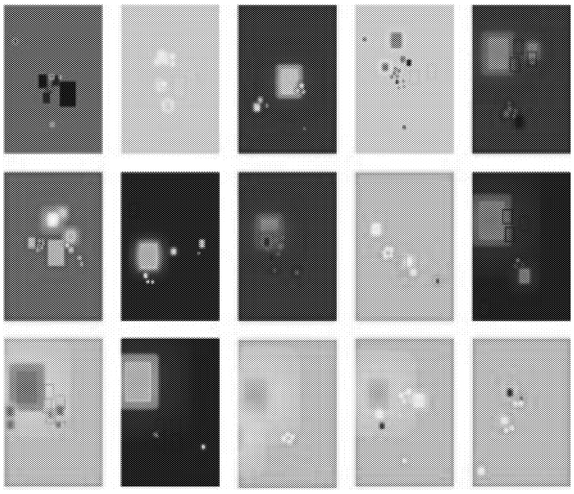


Figure 5: Time Machine – self-tracking data visualisation App (2007-2009) J. Hawkey, S. Oliveira, CADA, Lisbon, N. Correia CITI-FCT UNL, Caparica, Portugal

Apps developers, mainly recruited among the «geeky» users, have to consider the amount and variety of platforms and operating systems available, and are contending in a fierce market. For consumers, competition assures that the most requested mobile phone Apps keep being low-priced or free.

The most novel experiences include crowdsourcing to help people with disabilities. Just launched, Be My EYES (http://www.bemyeyes.org) is an open source non-profit App that aims to help the visually impaired in everyday situations. The user sends a message to the contact. The latter receives a notification. They connect via live video and the volunteer can help by answering the questions asked.

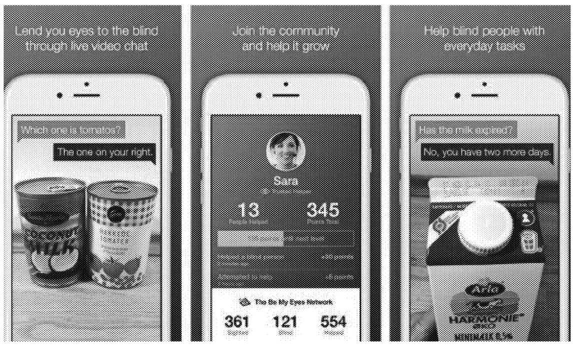


Figure 6 : Be My EYES - video data visualization App for crowdsourcing aid (2012-2015), H. J. Wiberg, Aarhus, Denmark

The outcome is a true empowerment of human users, and all the activities they will want to perform with their mobiles – namely new forms of storytelling in the Internet of Things.

OBJECTS TELLING STORIES BEFORE THE INTERNET OF THINGS

Relying on RFID (Radio Frequency Identification) sensors and a tablet, «Noon - A Secret Told by Objects» (2007) (http://cargocollective.com/tiagomartins/Noon-A-Secret-Told-by-Objects) by Christina Heidecker from Linz, Austria and Tiago Martins from Caparica, Portugal, is a mobile storytelling project, an interactive installation using real objects to convey a narrative, which resorts to a bracer, a wearable interface for ubiquitous gaming based on gesture-driven and tangible interaction.

Embedded with motion, direction, pressure and RFID, this bracer can be connected with mobile devices and allows for freed gaming using real objects. The player has to unveil the origins of the tragic fire that claimed the lives of the Novak family accessing the memories embedded in the remaining objects.

Figure 7 : Noon - A Secret Told by Objects (2007) Interactive installation, Ch. Heidecker Linz, Austria and T. Martins, Caparica, Portugal

In 2013, Tiago Martins, with Vesela Mihaylova and the musician Claudio Pina developed another mobile storytelling project: «Tripo» (http://cargocollective.com/tiagomartins/Tripo).

Three interconnected short stories are delivered through an imaginary device inspired by the tin can radio. Technically, they resort to electronic modules, RFID tags. Participants are led on the exploration of an indoor space, to find hotspots where the device can be used. At each of the hotspots participants may use the object to listen to fragments of stories which evolve and intersect throughout the location, as if they were memories of the space itself.

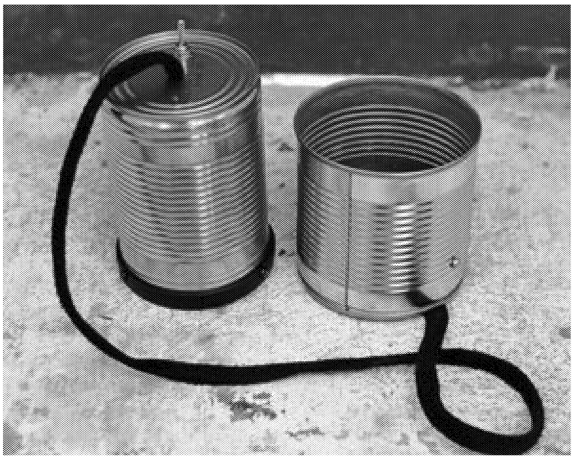


Figure 8: Tripo – Interactive installation (2013) – V. Mihaylova, Linz, Austria, C. Pina, T. Martins, Portugal

The hypotheses of objects talking to each other transferring data between themselves and over a network operating in real time became possible due to the convergence of wireless technologies, micro-electric mechanical systems and the Internet.

INTERNET OF THINGS (IoT)

In general the term refers to a global system of exclusively detectable interconnected objects that can effortlessly interoperate using communication protocols like RFID, Bluetooth, NFC (Near Field Communication), barcodes, embedded sensors and actuator nodes.

The goal of IoT is: «that all daily life objects equipped with unique identifiers (having specific identities and virtual behaviours) can be linked to the Internet and therefore can be managed as well as connect, communicate and interoperate with each other» (Khosrow-Pour 2015). The concept was initially applied in RFID-tags to mark the Electronic Product Code (Auto-IDLab) (Chen 2012).

Alternative definitions are being suggested: IoIT – Internet of Intelligent Things, as it deals with intelligent devices with adequate computing capacities (Chen 2012); Internet of Smart Things (O'Reilly 2005) as distributed intelligence is a part of it; Internet of Stuff (Bell 2015) due to the possible variety of things to be connected; or The Internet of You (Mims 2013) aiming to include a more affectionate relationship with the user. «The Internet of Things extends the Cloud Computing concept beyond computing and communication to include everything, particularly, the physical devices, Internet of intelligent things and robot as a service» (Khosrow-Pour 2015).

The point is that traditional boundaries between hardware and software are falling, and this disruption, this new software-enhanced, networked physical world needs a name. Internet of Things seems as good as any other.

THE THINGS – ENCHANTED OBJECTS

The Things of this IoT seem to be everything, in due course, and the (un)imaginable amount of objects and data are demanding organization, ontologies.

Efforts are being made concerning Machine to Machine (M2M) international standardization, and the three main groups involved are The Industrial Internet Consortium (http://www.iiconsortium.org), The Open Interconnect Consortium (http://openinterconnect.org) and the AllSeen Alliance community (https://allseenalliance.org), the latter offering the AllJoyn software, publicizing itself as: «the first truly collaborative, open ecosystem for interoperable products and services that will communicate engage and delight users in new, exciting and useful ways.» However, akin to what happens with Cloud Computing and Big Data issues, huge proprietary strategies are at stake: «German firms were trying to strike a balance between forming an international software standard for IoT while simultaneously representing German interests» (Geiger 2014).

As to the Things per se, some call them «enchanted objects» (Rose 2015) endowing them with magical qualities; others envision the objects subjectivity and humanization: «These Things have been called "enchanted objects" due to their capacity to contain their own stories, data-enabled to be actuated, connected, and "talk to other objects"» (Lloyd 2013). And these Things, endowed by humans with sensors and communication protocols, are invading everything. And the users, whenever they see it fit, cannot resist the temptation and will resort to anything to tell a story.

Smart-code

The above referred Auto-ID Labs (http://autoidlabs.org/) initial and primary research area was establishing standards for global commerce such as introducing barcodes to the retail industry. These were replaced by Quick Response (QR) codes, whose four information modes, fast readability and greater storage capacity, led them to be artistically used. There are two examples from 2012.

Figure 9: Chicago project - A Poem From Us (2012)
QR code – poetry-tags in objects (barcode)

Felix Jung's The «Chicago project - A Poem From Us» (http://avoision.com/portfolio/a-poem-from-us) encouraged visitors to record themselves reading a favourite poem, and to share their videos and thoughts via the project website. It gave free stickers with customized QR code that would display a random poem.

The first enhanced reality QR code made from Portuguese cobblestone and embedded in the ground was created by BeQRious. It welcomes the user and takes her to a site that offers data about Lisbon's Chiado and nearby places, including touristic and shopping information, local culture, commerce and cuisine.

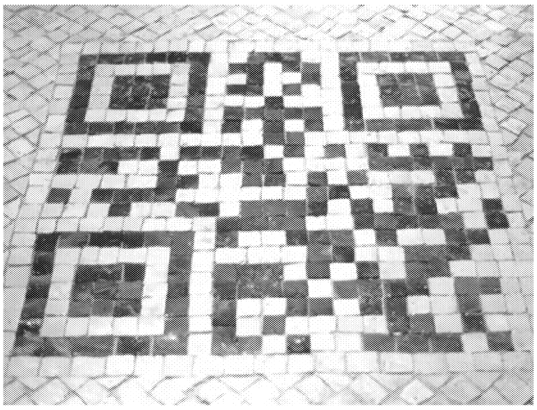


Figure 10: Rua Garrett's QR code (2012)
Urban information Lisbon, Portugal

Smart-homes

The utopian ideal of a smart-home is on the go, and listed in Forbes (Woolf 2014). Home-kit companies are already duelling for a huge future market. A smart-home kit accepts the connection of hundreds of objects among themselves, from the coffee-maker to door security locks.

The main issue is the connection system itself – either the recourse to http via wireless, or the creation of a hub-universal-translator capable of deciphering the multiple protocols and process it all in the Cloud. In all cases the user access is done via smartphone, and the multiple efforts to standardize how multiple connected objects talk to one another is under way.

Smart-wearables

The whole idea of wearable technology - objects that can be used by humans - being smartwatches, smart bracelets, smart glasses or hearing aid devices, is for people to be hands-free, online all the time, with seamless and portable access to the needed data when desired.

Concerning wearables - synonymous with M2M things - some industries, namely e-textiles, are still waiting for truly flexible-circuit boards, strong enough to withstand the demands of clothing and supple enough to contort with skin.

Bendable electronics will allow for a new range of applications, many of which are impossible to achieve using conventional methods.

While waiting for these advances, the hypotheses of some future new referents for «Pillow Talk» have been explored in two 2010 projects.

In the first one by Johanna Montgomery, Scotland, U.K. (http://www.joannamontgomery.com/Pillow-Talk) subject of a CNN news-story among other cases (http://youtu.be/PHrWi7rgxaY), each user wears a ring, and a small speaker is placed inside the pillow case. The ring picks up real-time heartbeat and transmits it via smartphone App to the other pillow. The purpose is to hear the other's heart-beat. The result aims to be an intimate interaction between two lovers, regardless of the distance between them.

Figure 11: Pillow Talk - smart objects (2010)
J. Montgomery, Scotland, U.K.

The second «Pillow-talk» (http://obm.media.mit.edu) comes from M.I.T. and is a project developed by Edwina Portocarrero and David Cranor. It is composed of a seamless voice-recording device embedded in a pillow, intending to capture dreams and that which we normally forget. After the dream is recorded into the pillow, the audio file is transmitted wirelessly to a jar containing shimmering LEDs to display the arrival of a new memory. Electronics in the jar can play back the recordings through a small speaker under its lid when it is opened.

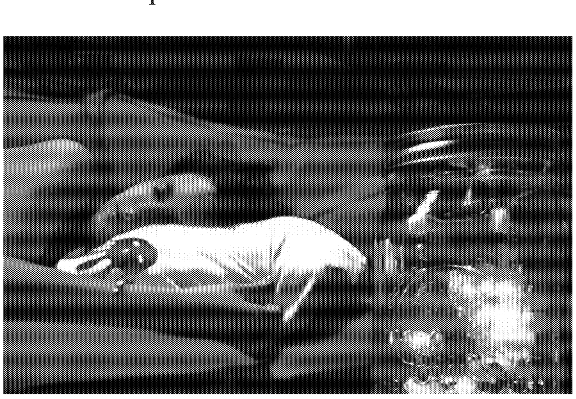


Figure 12: Pillow-talk – smart-objects (2010)
Edwina Portocarrero, David Cranor, MIT, USA

Smart-bodies

Body area networks (BANs) are emergent as enabling technology for many human-centred application domains such as health-care, sport, ergonomics, emergency, security and single assisted living. They consist of wireless wearable sensor nodes coordinated by a static or mobile. The data generated can be processed in real-time by the coordinator, transmitted to a server for processing and storing.

Due to the advance of flexible ultrathin electronics, BANs are being printed directly on the skin and can be worn for an extended period of time, while performing normal daily activities. Bio-stamps, or stick-on electronic Tattoos (http://rogers.matse.illinois.edu/index.php) cling to the skin like a temporary tattoo and can measure electrical activity from the body, allowing medical diagnoses and monitoring health conditions noninvasively.

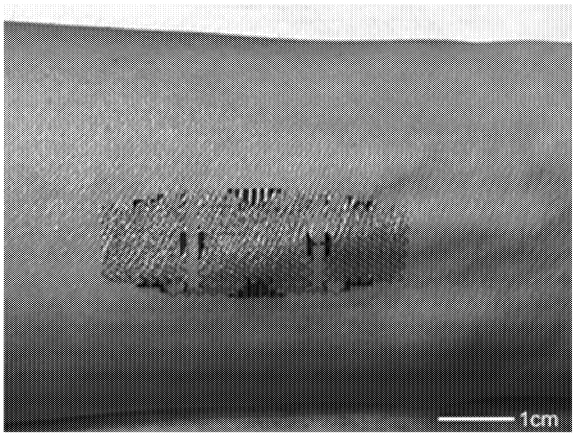


Figure 13: Biostamps - stick-on electronic Tattoos (2011)
smart-objects, epidermal electronics, J. A. Rogers

More intrusive is the aesthetical experience of Tiago Mesquita «The Heart of Living» (http://theartofliving.pt) winner of the Portuguese Young Artists Award (Jovens Criadores) 2014. Tiago implanted a chip on his chest, close to the heart, with a NFC (Near Field Communication) that only requires a smartphone to access the public website where he is recording his memories - images, videos and music at any time. Only he can delete or insert new content.

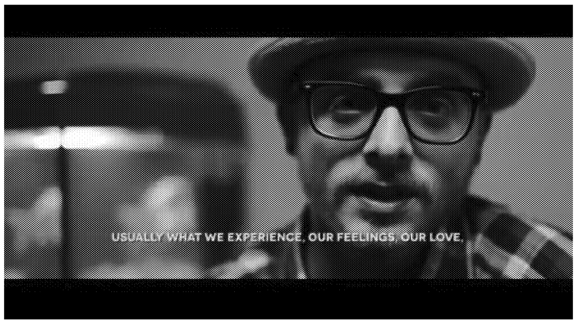


Figure 14: The Heart of Living – smart-body experience, chip implant (2014) T. Mesquita, Portugal – Winner of the Portuguese Award Young Artists (Jovens Criadores) 2014

INTERNET OF THINGS AWARDS

Started in 2011, the IoT Awards (http://postscapes.com/internet-of-things-award/2014/index) is aiming to honour the year's best project, companies, and ideas helping to create the Internet of Things in several areas (including the above referred) - connected: Body, City, DIY, Home, Industry & Environment, Open Source, and Technical. These could be considered the main arrays of public and corporate interests.

This year's 4th. edition reached 22 categories. Design Fiction type was accepted from the first moment: «The results may only be props or prototypes — but the best ones, as recognized by the Design Fiction award, end up helping us navigate our near futures and the stories they contain» receiving storytelling devices. The 2011 third runner up was a «Storytelling and Poetry Reciting Chair» (http://www.mia-kos.com/storytellingchair.html) by Mia Kos, Zagreb/Barcelona. A chair with sensors became enabled to tell stories and recite poetry when pressure was applied to the sitting part. Technically, was used a small system based on simple electronics and Arduino. The aim was to develop the concept of objects with alerted or enhanced behaviours.

Figure 15: – Storytelling and Poetry Reciting Chair (2011)
Mia Kos, Zagreb/Barcelona - IoT Award 2011-2012

IoT AND STORYTELLING

All of the above mentioned projects and examples imply or refer a narrative and/or some kind of storytelling, but none can fit in the normal categories of fiction, even the latest modes of interactive fiction. The idea of things telling stories is not new, and the objects' information content still requires human input and interaction, but new technologies used for storytelling have reached new heights and complexity, invoking the Macluhan equation «medium is the message» (Macluhan 1964).

The Internet of Things will have to shelter «a narrative architecture with the conventions and tools that will allow people to consume and create stories. Like any new media channel, those conventions might borrow from what came before (just like TV borrowed from live theatre) but we'll soon discover the unique affordances of the technology and start to create something entirely new» (Thompson 2013).

If a Thing has the capacity to register information, it will surely be – as it has been – used to tell a story. And Things, even in non-artistic fields, have aesthetical aspirations. The struggle comes from existing conventions, short to embrace this kind of mixed Media. IoT fictions are non-linear, a-chronological and a-synchronous narratives, and could be incorporated in the concept of Design Fiction.

The concept of Design Fiction, itself in the making, also implies several and conflicting interpretations (Blecker 2009; Sterling 2009; Tanenbaum 2012), but they seem open enough to come to include and embed the aesthetic requirements of this new kind of literary practice.

This hypothesis can be substantiated by two things: the fact that Design Fiction admits «narrative elements» having the concept of '*diegesis*' at its heart, or: «Design Fictions are events enabling the coupling of things and words» (Hales 2013). One of the globally accepted models-examples of a Design Fiction object is the Newspad from «2001: A Space Odyssey» (1968) – the novel-film signed by Arthur C. Clark and Stanley Kubrik.

Figure 16: 2001: A Space Odysssey NewsPad (1968)
A. C. Clark, Stanley Kubrik

The main issue would be to unfetter the concept from the field of Science-Fiction allowing it to include/enlarge itself to comprehend all other kinds of fiction, from novels to IoT narratives.

CONCLUSION

The theories and practical examples itemized above allow a glimpse of the speed at which things are changing in the new fields of science invading everyday life.

The increasing weight of user-chooser participation in Cloud Computing and the Internet of Things urges the making of conventions, terminologies and ontologies that may frame their handling in particular in what concerns tagging the objects of IoT – already addressed by some Consortiums and crowdsourcing projects.

As for artistic practices, a new aesthetics seems to be in the making, and the trend towards Mixed Media would require a wider dialogue between its practitioners and Digital

Humanities, Literature (and its branch Science-Fiction) Interactive Fiction and Cyber arts.

In itself, the concept of Design Fiction also implies several contradictory interpretations but they are open and new enough to come to include and embed the aesthetic requirements of this new kind of storytelling practice.

REFERENCES

Bleecker, J. 2009. "Design Fiction: A Short Essay on Design, Science, Fact and Fiction." *Near Future Laboratory*, 49. Retrieved from http://www.nearfuturelaboratory.com/2009/03/17/design-fiction-a-short-essay-on-design-science-fact-and-fiction/

Chen, Y., and H. Hu. 2012. "Internet of intelligent things and robot as a service." *Simulation Modelling Practice and Theory*. doi:10.1016/j.simpat.2012.03.006

Davidson, C. N. 2008. "Humanities 2.0: Promise, Perils, Predictions." *PMLA*, *123* (3), 707–717. Retrieved from http://www.jstor.org/stable/25501892

Hales, D. 2013. "Design fictions an introduction and provisional taxonomy." *Digital Creativity*, *24* (1), 1–10. doi:10.1080/14626268.2013.769453

Ishizu, T., and S. Zeki. 2011. "Toward a brain-based theory of beauty." *PLoS ONE*, *6*. doi:10.1371/journal.pone.0021852

Johnson, S. 1997. *How New Technology Transforms the Way We Create and Communicate.* (1st. ed.,). New York: Basic Books.

Khosrow-Pour, M. 2008. Encyclopedia of Information Science and Technology. In *Encyclopedia of Information Science and Technology* (3rd. ed., p. 5270). doi:10.4018/978-1-60566-026-4

Kuhn, T. S. 1970. *The Structure of Scientific Revolutions*. *University of Chicago Press* (2nd. ed.). London.

McAllister, J. 1998. "Is beauty a sign of truth in scientific theories?" *American Scientist*, *86* (March-April), 174. doi:10.1511/1998.2.174

McLuhan, M., and Q. Fiore. 1964. "The Medium is the Message." In *Understanding Media: The Extensions of Man* (pp. 1–47).

Rose, D. 2014. *Enchanted Objects: Design, Human Desire, and the Internet of Things* (1st. ed.,). New York: Scribner.

Sanaei, Z., S. Abolfazli, A. Gani, and R. Buyya. 2014. "Heterogeneity in Mobile Cloud Computing: Taxonomy and Open Challenges." *IEEE Communications Surveys & Tutorials*, *16* (1), 369–392. doi:10.1109/SURV.2013.050113.00090

Sterling, B. 2009. "Design Fiction." *Interactions*, *16*, 20–24. doi:10.1145/1516016.1516021

Tanenbaum, J., R. Wakkary. 2012. "Steampunk as Design Fiction." *CHI*, 1583–1592. doi:10.1145/2207676.2208279

Yanosky, R. 2008. "From Users to Choosers: The Cloud and the Changing Shape of Enterprise Authority." In R. N. Katz (Ed.), *The Tower and The Cloud* (pp. 126–136). Retrieved from http://net.educause.edu/ir/library/pdf/PUB7202m.pdf

WEB REFERENCES

Barbas H. 2011. "Neo-Modernismos - o artista engenheiro no século XXI." http://www.helenabarbas.net/papers/2011_Artista_Engenheiro_H_Barbas.pdf

Bell, G. 2015. "How to Live with the Internet of Stuff". *MIT Technology Review*. Retrieved March 6, 2015, from http://www.technologyreview.com/news/528126/digital-summit-internet-of-things-think-of-it-as-the-internet-of-stuff/

Geiger, F. And A. Preuschat. 2014. "Germany Moves Away From U.S.-Dominated IoT Standards Groups" Digits. *The Wall Street Journal*. Retrieved March 18, 2015, from http://blogs.wsj.com/digits/2015/03/18/germany-moves-away-from-us-dominated-iot-standards-groups/

Nova, N. 2014. I&IC Workshop #2 at HEAD: output > "Cloudified" Scenarios. Retrieved March 9, 2015, from http://www.iiclouds.org/20141110/iic-workshop-2-at-head-geneve-output/

O'Reilly, T. 2005. What Is Web 2.0 - O'Reilly Media. Retrieved March 7, 2015, from http://www.oreilly.com/pub/a//web2/archive/what-is-web-20.html

Thompson, D. 2013. "A narrative architecture for The Internet of Things | m2m Telefónica." Retrieved March 6, 2015, from https://m2m.telefonica.com/blog/a-narrative-architecture-for-the-internet-of-things

Wolf, M. 2013. "5 Smart Home Companies to Watch in 2014". *Forbes*. Retrieved March 8, 2015, from http://www.forbes.com/sites/michaelwolf/2013/12/31/5-smart-home-companies-to-watch-in-2014/

BIOGRAPHY

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

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

She was a member of the InStory team (2005-2007) – best Portuguese web mobile project 2006. She prepared a project on serious games, PlatoMundi, aiming to introduce e-learning and ethical issues in game playing; she is developing a new project – Numina – that proposes to digitize and study the poetic and dramatic literary estate of José Leite de Vasconcelos (1858-1941), a physician, philologist, ethnologist who founded the Portuguese National Archaeological Museum.

In 2011 she received the (2nd.) SANTANDER Award for the Internationalization of the F.C.S.H. Scientific Production 2010.

Homepage: http://www.helenabarbas.net

TOWARD FREE AND OPEN SOURCE FILM PROJECTION FOR DIGITAL CINEMA

Nicolas Bertrand
Jean-Denis Durou
Vincent Charvillat
IRIT, UMR CNRS 505
Toulouse, France
E-mail: nicolas.bertrand@isf.cc

Carsten Griwodz
Simula Research Laboratory
Oslo, Norway
E-mail: griff@simula.no

KEYWORDS
Image Compression, Digital Cinema, JPEG2000, Open Source, VLC, libavcodec.

ABSTRACT
Cinema industry has chosen Digital Cinema Package (DCP) as encoding format for the distribution of digital films. DCP uses JPEG2000 for video compression. An efficient implementation of coding and decoding for this format is complex, however. Currently deployed equipment is expensive and has high maintenance costs, preventing art-house cinema theaters from acquiring it. Therefore, we conduct this research activity in cooperation with Utopia cinemas, a group of art-house cinemas, whose main requirement (besides functional ones) is to provide Free and Open Source Software (FOSS). This paper presents a solution that achieves real-time JPEG2000 decoding and DCP presentation based on widespread open source multimedia tools, namely VLC and libavcodec library. We present the improvements that were made in VLC to support the DCP packaging format, as well as details on JPEG2000 decoding inside libavcodec (optimization and lossy decoding). We also evaluate the performance of the decoding chain.

I. INTRODUCTION

Cinema theaters switched from 35 mm prints to digital era. The Digital Cinema System Specification (DCSS) [1], provided by the Digital Cinema Initiative (DCI), is now a world-wide standard. The specification describes how to create, distribute and project a Digital Cinema Package (DCP). It requires JPEG2000 as intra-frame codec for video, WAV for audio, and XML for subtitles. Image dimensions are specified for 2K and 4K, with three 12-bit color components sequenced with a frequency of at least 24 fps (frames per second). The DCP size is typically about 80–200 GB according to film duration or compression ratio.

Our research is conducted in collaboration with Utopia cinemas (five independent theaters in France). Those theaters initiated this project because they need to understand the implications of changing to digital. Primarily, they are concerned about becoming dependent on a single company's technology for presentation, and want us to provide Free and Open Source Software (FOSS) for decoding films distributed in the DCP format.

The JPEG2000 [2] format was selected by DCI for video compression because of its compression efficiency. The implementation of a fast JPEG2000 encoder and decoder, however, is complex. Commercial equipment currently used by cinemas relies on VLSI hardware for decompressing DCP at the required frame rate. This kind of equipment is expensive, and not affordable for art-house cinema theaters. Our research goal is to lower the cost for DCP playback by developing a software that runs on today's standard off-the-shelf hardware. There are already software solutions, such as Kakadu [3] and EasyDCP [4] for real-time JPEG2000 decompression, but these are not FOSS solutions. Kakadu is a JPEG2000 coder/decoder that supports all kinds of JPEG2000 profiles, but cannot playback DCP. EasyDCP is dedicated to DCP playback and can play DCP in real-time, but as Kakadu, the software is not FOSS.

Our selection of VLC (Video LAN Client) as the basis for our DCP decoder is due to its position as a FOSS solution with high flexibility, performance and proliferation. We have implemented a DCP module inside VLC, and a JPEG2000 decoder inside the libavcodec multimedia library (used by VLC).

In the next section, we review current fast software implementations for JPEG2000 decoding. Then, we present our projection system called OpenSMS (Screen Management System) in Section III. In Section IV we present the VLC architecture, and the design of our module. The JPEG2000 decoder (named J2K-libavcodec) implementation is detailed in Section V. In Section VI, validation and performance measurements of our solution are presented.

II. STATE OF THE ART – THE PROBLEM

The key point for DCP playback is to reach a frame rate of 24 fps with synchronized playout of audio, video and subtitles. The bottleneck for achieving this is the computational complexity of JPEG2000 decoding. Overcoming it in software requires parallel computation, and there are currently three dominant approaches to this: the use of multi-threading and processor-specific multimedia extensions like Streaming SIMD Extensions (SSE), the use of GPGPU, or a combination thereof. For example, Taubman [5] relies on multi-threading and SSE, the patented EasyDCP [6] relies on GPGPU to present a GOP (Group Of Pictures) approach based on similarity between codeblocks to avoid latency, Le [7] mixes both approaches. Le [8] and Matela [9] explain proposals for parallelizing EBCOT using GPGPU. Taubman [5] shows that a decompression of 24 fps can be achieved with multi-threading and SSE instructions. This solution is implemented in Kakadu. Avoiding GPGPU reduces the decoder's complexity and hardware dependence; the dependency on NVIDIA hardware due to the use of CUDA is a limitation of solutions by Le [8] and Matela [9]. Although OpenCL [7] is meant to be platform-agnostic, detailed hardware-specific tuning is required to achieve a high performance.

Another way to reduce the decompression time is to not completely decode all the compressed bitstream. In Jimenez [10], a method for visually lossless decompression is presented. The method is based on a perceptual model designed by the authors for compression. The proposed strategy is as follows for each codeblock: extract the most significant bitplanes from the codeblock header, compute a Visibility Threshold (VT), based on variance estimation and the author's perceptual model. Then, during codeblock's decompression, at each coding pass an error is compared to VT. If the error is lower than VT, remaining coding passes are skipped. Kakadu also proposes an option called "bitstream truncation" to achieve better decompression performance by stripping away coding passes.

We decided to implement a decoding solution that combines multi-threading, SSE instructions and basic skip of coding passes for least significant bitplanes codeblock's coefficients. This decoding technique is a key component of our alternative projection system.

III. PROJECTION SYSTEM

The standard DCI projection system and our OpenSMS are juxtaposed in Figure 1. Our goal is not to propose a new projection system, but to provide a simpler version of an existing one. Table I presents the features of both systems.

A DCI projection system is composed by a storage device (where the DCPs are ingested), connected to an Integrated Media Block (IMB) via a PCI-Express link. The IMB is in charge of media decryption and decompression. Plain decoded images are sent to the projector, the sound processor receives the plain audio channels.

TABLE I: DCI Projection System Features Compared to OpenSMS.

Feature	DCI	OpenSMS
Picture Size	2K	2K interpolated
DCP 4K Support	Yes	Yes †
Picture Depth	12 bits	10 bits
Sound	PCM 24 bits	PCM 24 bits
Subtitle	Yes	Yes
DCP Decryption	Yes	Yes
Security Manager	Yes	No
Physical Security	Yes	No
Secure Logging	Yes	No
Forensic Watermarking	Yes	No

†To Be Implemented

Our OpenSMS system is composed by an off-the-shelf hardware and an "e-cinema" projector (Barco RLMW8, using tri-DLP as DCI projectors). For our system validation, we have defined two hardwares: a laptop with a quad core Intel i7 and for better performance, a bi-Xeon with 12 cores. The hardware and the projector are connected via a HDMI link, using HDCP encryption.

We have also decided to not implement all the security requirements described in specification [1]. The main reason is to reduce the system complexity and we think that the security constraints do not fit with the projection policy of art-house cinema. Nevertheless we are aware of content protection, and plan to implement playback of encrypted DCP.

Our OpenSMS software is based on VLC and is presented in Section IV.

After video decompression, to preserve correct color display in a monitor or a projector, we implemented a GLSL filter to convert from CIE 1931 XYZ to the required color space, usually sRGB.

IV. DCP PLAYBACK IN VLC

VLC is a widespread cross-platform FOSS media player, downloaded more than 1.4 billion times. VLC is essentially a multimedia framework, where you can dynamically load modules according to the input (files, network streams) and the outputs (audio or video, on screen or network). The framework's core handles low-level operations like threading, timing and synchronization. It is also in charge of pipelining the media streams from input to output by connecting modules, used to do the media processing work. An example of modules loaded for the DCP case is illustrated in Figure 2. Each module has a type according to its purpose, like access, demux, codec, video_output, ...

VLC depends on a large number of FOSS libraries including libavcodec.

Our first contribution is the creation of a DCP module, VLC had not handled it so far. JPEG2000 decoding is done by our decoder (J2K-libavcodec) implemented in libavcodec (cf. Section V). The OpenGL module was slightly modified to implement in GLSL the XYZ to sRGB conversion.

Next we present our module design. A DCP is a set of files stored in a folder. DCP meta-data are stored in XML files, while MXF containers are used to store the media *essence* (video, audio, subtitles). There is at least one MXF file per essence, but several audio and subtitles MXF files can be stored in the folder to handle several languages.

A DCP is a container (the folder) that in turn contains several other containers (the MXF files). Operations for *access* and *demux* are not easily separated, so an *access_demux* module was judged more suitable for DCP. The essence files are accessed via *asdcplib* external library. We chose this library because it supports MXF containers and DCP essence types, and it is popular in cinema tools. Once the essences are extracted, VLC Elementary Streams (ES) are created, and

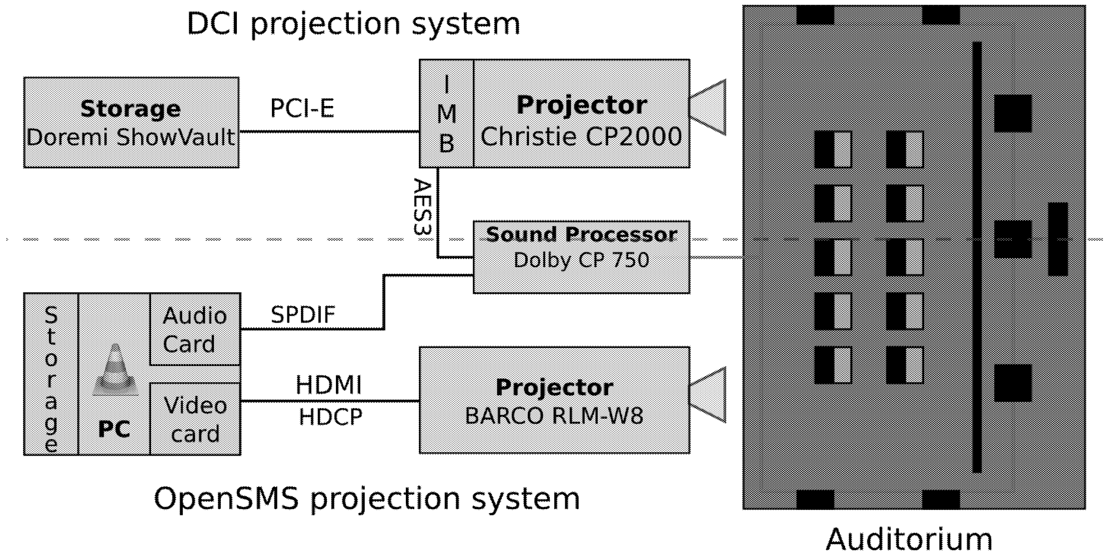


Fig. 1: Digital cinema Projection Systems. The DCI media block, integrated in the projector, performs decoding. Our system (OpenSMS) is not integrated, and can be connected to all kinds of projectors, or display directly on a standard monitor.

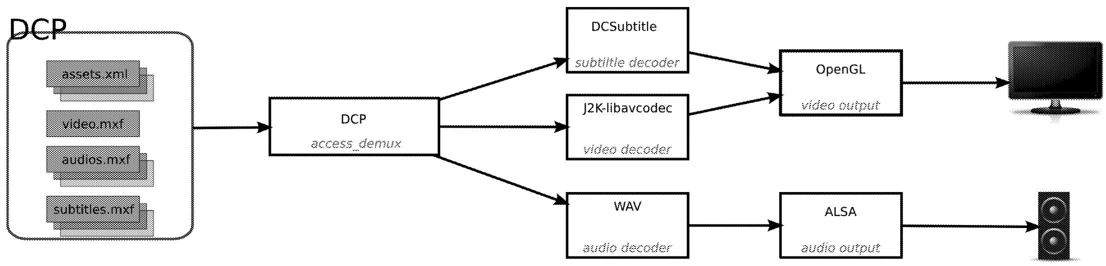


Fig. 2: DCP Playback in VLC and libavcodec. Each box represents a VLC module, and its functionality is written in red. The arrows represent the streams between modules.

sent to the decoders. ES, one per media, form the interface between the demuxers and the decoders.

VLC design is highly multi-threaded, modules are executed in separate threads. Consequently, demultiplexing, decoding and display are executed asynchronously. The synchronization of audio, video and subtitle is performed by a dating mechanism called *Presentation Time Stamps*. The date is set in the ES, through our module and the output modules use this date to play at the right time.

Our DCP module is publicly available in VLC master branch since December 2013.

Code for DCP VLC module is available in git.videolan.org, stored in directory modules/access/dcp (http://git.videolan.org/?p=vlc.git;a=tree;f=modules/access/dcp).

Even if the acceptance of the module required a painstaking work, it was a big step towards a FOSS DCP playback solution.

V. JPEG2000 DECODER IMPLEMENTATION

Our second contribution is a specific JPEG2000 decoder for libavcodec, a coder/decoder multimedia library. Like VLC, libavcodec is a FOSS and cross-platform project. Our JPEG2000 decoder is not the only one. Libavcodec can also use OpenJPEG library for decoding JPEG2000 files. OpenJPEG is a FOSS library hat accepts all JPEG2000 profiles, but is mostly aimed at coding still images. In contrast to this, our codec is aimed at decoding JPEG2000 videos, especially for cinema. By specializing, we can achieve smaller structures and simpler code, avoiding OpenJPEG's complex structure that is required to support all JPEG2000 options.

As the decoder is intra-frame, we decided to multi-thread the decoder at frame level with threading mechanisms provided by libavcodec primitives.

The inverse Discrete Wavelet Transform (DWT) and Main Component Transform (MCT) was optimized with SSE instruction set. The SSE optimization is widely inspired from OpenJPEG, and by focusing on structure size.

The lifting based 2D DWT is optimized by SSE instructions. The six steps of irreversible 1D filtering (cf. chapter F.2.8.2 in JPEG2000 standard [2]) are executed through two SSE functions. One function for the two first steps (linear computes), the other for the remaining steps (nonlinear computes). Implementing SSE instructions for MCT is easy, since it is a matrix multiplication.

Furthermore, for decoding, we increase the performance by

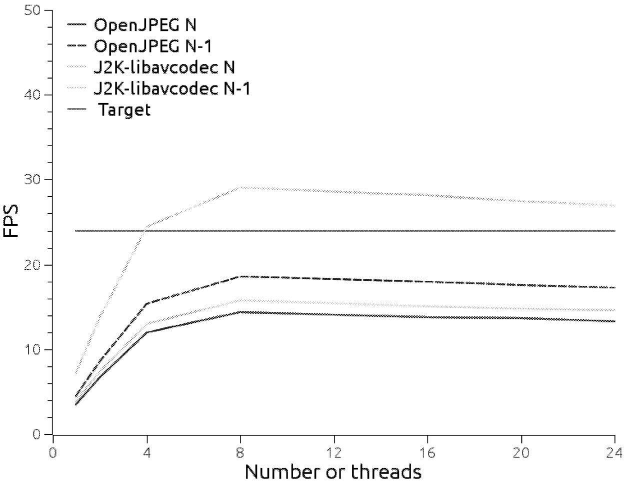


Fig. 3: JPEG2000 Decoders Performance on Machine 1 ($N = 5$ resolution levels).

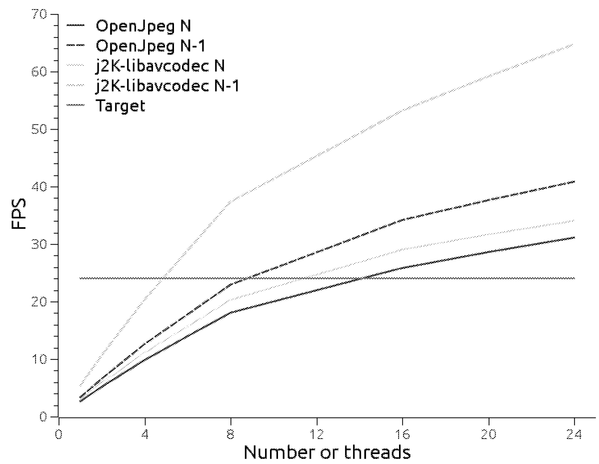


Fig. 4: JPEG2000 Decoders Performances on Machine 2 ($N = 5$ resolution levels).

doing only once some parsing of JPEG2000 headers, as the digital cinema profiles force many parameters.

Our J2K-libavcodec is also publicly available in libavcodec. To get the released code clone git://git.libav.org/libav.git, the JPEG2000 decoder is stored in libavcodec directory, the decoder is stored in jpeg2000* files. All the optimizations are not yet pushed in libavcodec, but are publicly available in our Gitorious account (clone git@gitorious.org:libav/nicoisfs-dondiego-libav.git and checkout expev2 branch for JPEG2000 decoders with all optimizations).

The decoder evaluation and comparison with OpenJPEG is presented in next section.

VI. EXPERIMENTAL VALIDATION

A. JPEG2000 Evaluation on libavcodec

The tests are made on two machines. Machine 1 is a laptop with an Intel Core i7-2820QM operating at 2.30 GHz, the processor has 4 cores for a total of 8 SMT threads (hyper-threads in Intel terminology). Machine 2 is a bi-Xeon E5-2620 operating at 2.0 GHz. Each chip has 6 physical cores, resulting in a total of 24 SMT threads. All tests are conducted with hyperthreading turned on. The tests are performed according to the number of CPU-threads and resolution levels. The tested DCP is the trailer of Moonrise Kingdom, and all numbers refer to decompression of the first 10 seconds (240 frames). The movie was encoded with a bit rate of 128 Mbits/s. Moonrise Kingdom image size is 1998×1080 with 5 resolution levels. The number N of resolution levels tested varies between 5 (2K format, 1998×1080) and 4 (1K format, 999×540).

In Figure 3, the comparison of J2K-libavcodec with OpenJPEG is presented (Machine 1). As the CPU has 8 SMT threads, the decoding performance with both decoders is limited to 8 libavcodec threads. A performance decrease has been observed if we use more libavcodec threads than the CPU has virtual cores. In 2K neither OpenJPEG nor J2K-libavcodec reaches the target (red line) of 24 fps; at least our decoder performs better. In 1K J2K-libavcodec reaches the target with 4 threads.

For Machine 2, the results are presented in Figure 4. All decoders, regardless the resolution level, reach the target of 24 fps: J2K-libavcodec reaches the target frame rate with 12 threads, OpenJPEG with 16 threads.

The fps gain of our decoder at 1K is high. There are two reasons. First, the image to decode is smaller (4 times smaller than 2K), so there are less codeblocks to decode. We cannot expect a gain close to 4 as the non-decoded codeblocks are the high frequency ones: EBCOT encodes high frequency codeblocks with less bits than low frequency. The other reason is memory usage and CPU cache management. With OpenJPEG, we achieve a gain of 1.5 between 2K and 1K, while ours is around 2. The main reason for this difference is memory management: we have smaller structures for decoding data flow, resulting in a reduction of cache misses and page faults.

TABLE II: Fps Comparison of Movies at Several Compression Bit Rates. Run on Machine 2 with J2K-libavcodec in 2K.

Movie	Bit rate (Mbits/s)	12 threads (fps)	24 threads (fps)
Spring Breakers	201	19.05	24.62
Django Unchained	151	19.58	24.64
Moonrise Kingdom	128	27.12	33.76
Jeux d'été	79	30.34	38.28

Table II shows the evolution of the fps according to the compression bit rate. The selected movies are DCP trailers shown in movie theaters. The decoding is tested on the first 240 frames. As expected, lower compression bit rates lead to higher frame rates. For Spring Breakers and Django Unchained we have almost the same fps. Spring Breakers has many black images and completely decoding black images is fast for a JPEG2000 decoder.

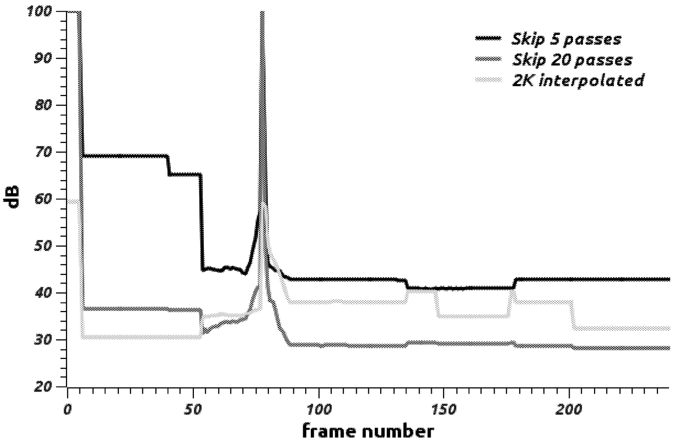


Fig. 5: PSNR Comparison. A peak is present around frame 80; this is due to the display of incrusted subtitles in our sample.

B. Image Quality Measurements

We present here tests of video quality measurements. The used sample is the same as the one used for performance measurements. The reference is the full 2K sample of Moonrise Kingdom. To test performance of lossy decoding, we skip the last decoding passes of EBCOT, by this way we skip the least significant bits of codeblocks coefficients. In Figure 5 we present the PSNR for the 240 frames of the sample. The measures are made by skipping 5 and 20 passes, and also the PSNR of the J2K-libavcodec with decompression at 1K and a bicubic interpolation to reach 2K.

By visual inspection of the video with 5 passes skipped, the image quality is acceptable for projection. With 20 passes skipped the image quality is not acceptable. The 240 frames are decomposed as follows: black image, 2 logos, movie sequence, black image, and 3 movie sequences. The variation of PSNR shows the various decomposition of the video. For 2K interpolated during the same sequences, variations of PSNR are observed. This is due to the incrustation of subtitles in the input video. When PSNR goes down subtitles appear.

TABLE III: Bit Rate (in fps) at Several Resolution Levels and Several Pass Skips for Moonrise Kingdom Sample.

	N	$N-1$
Skip pass 0	33.21	63.16
Skip pass 5	46.08	94.71
Skip pass 20	61.26	204.08

Table III presents fps performance in pass skipping. The test was executed on Machine 2. By skipping 5 passes, for N and $N-1$ resolution levels, a great performance gain is achieved (respectively 40% and 50%). More gain is achieved by skipping 20 passes, but the image quality is not acceptable. As stated by Jimenez [10] for image quality measurement, we confirm that visually lossless decoding is also computationally advantageous for JPEG2000 cinema streams.

C. DCP Playback Evaluation

The full read of DCP with audio and video synchronized, and correct color space display is made at VLC level. There are no tools inside VLC to measure the global performance of full decoding, only warnings or errors can be raised, indicating for instance loss of synchronization between video and audio, or decoding that is too slow. We used those warnings and visual inspection of DCP to evaluate.

TABLE IV: DCP Read at 24 fps, at Several Resolution Levels. 'Yes' indicates that DCP is read without desynchronization.

Machine	2K	2K interpolated (from 1K)
1 (i7 / 4 cores)	No	Yes
2 (bi-Xeon / 12 cores)	No	Yes

The evaluation results are reported in Table IV. We do not succeed to play the DCP at full scale, even with the decoding performances presented before. This is due to remaining work on our VLC module (pre-caching of video and limitation to 16 threads for decoding the video). We can play in interpolated 2K format. The codec provides 1K streams to VLC, the interpolation is made by VLC in the *video_output* module. The execution of the interpolation has no impact in DCP playback (no image freeze, no audio/video desynchronization). The samples used for those evaluations are available in http://utopialab.tetaneutral.net/DCPsamples/ website.

VII. CONCLUSION AND PERSPECTIVES

We have compared our decoder performance with OpenJPEG, the reference and fastest FOSS decoder. We have shown that our decoder is faster, and outperforms OpenJPEG when the highest frequency resolution level is skipped. Skipping some LSB decoding passes in EBCOT provides a great improvement on performances.

Our future work on the decoding part will consist in implementing multi-threading at codeblock level and further reducing data structure size. The purpose of that is to have piece of code that depends only on local data (i.e. no reference to higher level data). In this way, data structure should be kept in L1 and L2 caches during the execution and reduce the data time access.

We need also to validate the pass skipping and the 1K to 2K interpolation at system level to achieve that we will organize visual perceptions tests in cinema projection room. The tests will compare our system with current DCI projection system.

A comparison with closed-source software is more cumbersome and left for future work. Some performance results for Kakadu are presented on the official website, but a direct comparison is impossible because the sample sequences are not provided. The sample image size is 20% smaller, but the compression bit rate (244 Mbits/s) is twice higher than the one from our test sequence Moonrise Kingdom. On a bi-Xeon, Kakadu achieves a framerate of 24.24 fps (35.08 fps with the separate ”speed pack”). In comparison, we achieve 34.04 fps on the same kind of machine. Results for EasyDCP are not publicly available.

As the number of cores increases in current architectures, we expect to avoid the usage of GPGPU at decoding level. In the overall pipeline however, we will use GPGPU, in particular for filtering (color space change and/or image color corrections) after decoding.

At VLC level, we need to optimize the DCP module to handle audio, video and subtitle synchronization with video in full resolution. Asdcplib allows management of encrypted MXF files, so we will implement handling of encrypted DCP. This feature will also be very useful for cinema exhibitors.

We are not far, from the point of view of performance and functionality, from a complete FOSS solution for DCP playback, usable by movie distributors, to preview the DCP, and by exhibitors for DCP playback.

REFERENCES

[1] DCI, *Digital Cinema System Specification*, Digital Cinema Initiatives, 1.2 with errata edition, Mar. 2012.

[2] ITU-T, *Information technology – JPEG2000 image coding system – Part 1: Core coding system*, Aug. 2002.

[3] UNSW, “Kakadu software,” July 2013.

[4] FraunHoffer IIS, “EasyDCP software suite,” July 2013.

[5] D. Taubman, “Multithreaded processing paradigms for JPEG2000,” in *MMSP 2012*, 2012, pp. 164–169.

[6] V. Bruns, H. Sparenberg, and S. Fossel, “Video decoder and methods for decoding a sequence of pictures, patent wo 2012004164 a1,” Jan. 2012.

[7] R. Le, J.L. Mundy, and R.I. Bahar, “High Performance Parallel JPEG2000 Streaming Decoder Using GPGPU-CPU Heterogeneous System,” in *ASAP 2012*, 2012, pp. 16–23.

[8] R. Le, I.R. Bahar, and J.L. Mundy, “A novel parallel Tier-1 coder for JPEG2000 using GPUs,” in *SASP 2011*, 2011, pp. 129–136.

[9] J. Matela, M. Šrom, and P. Holub, “Low GPU Occupancy Approach to Fast Arithmetic Coding in JPEG2000,” in *LNCS 7119*, pp. 136–145. 2012.

[10] L. Jimenez-Rodriguez, F. Auli-Llinas, and M.W. Marcellin, “Visually lossless strategies to decode and transmit JPEG2000 imagery,” *Signal Processing Letters, IEEE*, vol. 21, no. 1, pp. 35–38, 2014.

REMOTE EDUCATION

AUGMENTING DISTANCE EDUCATION SKILLS DEVELOPMENT IN PARAMEDIC SCIENCE THROUGH MIXED MEDIA VISUALISATION

Michael Cowling
School of Engineering & Technology
CQUniversity
160 Ann St
Brisbane QLD AUSTRALIA
E-mail: m.cowling@cqu.edu.au

Emma Moore
School of Medical and Applied Science
CQUniversity
Bruce Highway
Rockhampton QLD AUSTRALIA
E-mail: e.moore@cqu.edu.au

James Birt
Faculty of Society & Design
Bond University
Gold Coast
QLD AUSTRALIA
E-mail: jbirt@bond.edu.au

KEYWORDS
3D printing, augmented reality, mixed media learning, paramedic science.

ABSTRACT

This paper presents a learning intervention using mixed media visualisation (3D printing and Augmented Reality simulation) to enhance skills development for paramedic science students studying through distance education. Presented is a research methodology to evaluate the effectiveness of the mixed media visualisation techniques to provide more hands-on skill practice to paramedic science students studying the course at a distance. The context for this study is the skills acquisition and retention, focusing on Laryngoscopy with foreign body removal.

The project stems from a need, identified through course evaluations, for more opportunity for distance students to practice skills (currently, they can only be practiced in a five day hands on residential school). Selected students will be provided with 3D printed instruments and a mobile phone augmented reality simulation application that they can use to practice skills prior to arrival at the residential school. Additional training will also be provided to those students not selected using the tools during the residential school. Outcomes are expected to be an overall improvement in final skill level for all students.

INTRODUCTION

In 2001, Prensky (Prensky, 2001, 2009) foreshadowed the arrival of a generation of students that he dubbed "Digital Natives", born and bred in a world immersed in digital technologies. Prensky argued that young people born after 1980 differed from previous generations not only in the degree to which they used technology but also in more qualitative ways, such as the way they interact with technologies and even the way they think and learn. He called for a radical overhaul of education systems because "Today's students are no longer the people our education system was designed to teach" (Prensky, 2001, p.1).

Other papers have since argued against the use of the term "Digital Native" to describe a generation of students (Bennett & Maton, 2010), but it's clear that new technology and innovation can now be used in the classroom that has never previously existed. Specifically, tools such as 3D printing and Augmented Reality (AR) are becoming available for use commercially and thus able to be incorporated into the classroom.

The use of visualization as positive learning support tools are well documented and accepted (Mayer, 2005). Numerous academic disciplines incorporate a variety of 2D and 3D visualisations and haptic manipulations including medical anatomy, architecture, geography, chemistry and media/game design (Freitas & Neumann, 2009). In addition, training has been improved in areas such as nuclear power plant operations and astronaut training (Dalgarno and Hedberg, 2001). The fundamental question is therefore not whether visualization affects learning but how to take advantage of the various visualization media, lesson sequencing and student reflection so that instructions and learning can be more effective (Kozma, 1991, 1994).

Meta-analytic studies of 2D and 3D visualization show positive improvements in learning outcomes among low and high spatial learners (Höffler, 2010). However, these studies are to date inconsistent (Huk, 2006). While some learners learn better when provided with non-dynamic media affording them the opportunity to build their own mental model, other learners learn better through provided virtual dynamic models or physical haptic manipulation (Hwang & Hu, 2013).

No particular media is necessary for learning, nor is a particular method, however both media and methods of incorporation into a curriculum influence learning by influencing each other. Choices have to be made as media constrains and enables methods and methods take advantage of media capabilities (Kozma, 1991, 1994). Typical studies examine only single media coding of the visualization (Höffler, 2010) but secondary modality or multi-modal instruction is important (Mayer, 2002, 2005; Moreno & Mayer, 1999).

Visualization for teaching and learning is nearly ubiquitous. In many cases, visualizations represent either reality, or an approximation of a physical reality.

The specific context of this exploratory study is an introductory paramedic science class studied by many students at a distance. In this class, learning is considered to be an active process influenced by prerequisites of the learner (Mayer, 2002) and the class requires numerous "hands-on" exercises in learning to learn the skills required to be a paramedic. Yet, for many distance students, these skills cannot be provided until they attend residential school, often at the end of a semester of study in a location remote to the

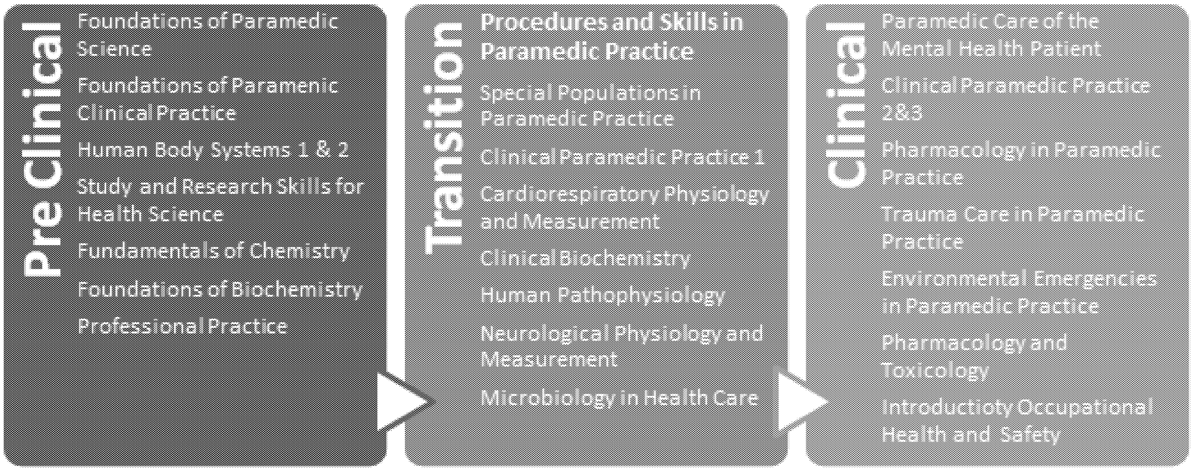


Figure 1: Outline of Paramedic Science Program

university, devoid of the tools needed for skills practice. This type of "hands on" visualization experience is difficult for students studying at a distance.

In this problem domain cutting edge Augmented Reality (AR) could hold the answer as objects can be explored in 3D space but even an object displayed using an augmented reality display can't be touched and held. While haptic technology may hold out the promise of adding the dimension of touch to digital information, this is often difficult for large scale distance education and there is no substitute on the near-term horizon for gaining the knowledge that we gain by holding and manipulating a physical 3D object. 3D printing offers a way to bridge this gap between the virtual and the real (Loy, 2014).

This paper presents an intervention using Mixed Media 3D printing and an AR game to provide a sand box simulation for enhancing skills development for distance paramedic science students and outlines a research methodology to evaluate the effectiveness of 3D printing and augmented reality techniques. This intervention was developed through collaboration of staff from paramedic science, design science and educational technology from two Australian institutions.

BACKGROUND

The CQU Bachelor of Paramedic Science is designed to equip students with the skills to become a Paramedic. They develop foundation knowledge in sciences, human body systems, study and research skills and paramedic practice. The program is three years full time with each year broken down into the following focus areas: Pre Clinical, Transition to Clinical and Clinical (see Figure 1).

As part of this program, one of the first courses looks to develop skills in paramedic science, congruent with the skills required for Queensland Ambulance Service (QAS) paramedics and in line with the QAS procedures. An example is the QAS laryngoscopy procedure and related foreign object removal procedure, as outlined in Figure 2.

In addition, students develop an understanding of paramedic care through investigation of the underpinning theory and practice of procedural applications in the discipline. Knowledge and skills are developed through a series of coursework exercises and practical laboratory sessions as well as a one week residential school for distance students.

Students in the Bachelor of Paramedic Science program are expected to have developed the real world expertise and skills to work as health professionals in emergency medicine and retrieval. Graduates of the program anticipate career options with the government ambulance service, private emergency services or in industry providing paramedic services to mine sites and other areas. Graduates employed in this capacity are eligible for membership with Paramedics Australisia and/or the Australian Registry of Emergency Medical Technicians (AREMT).

Yet despite these very practical requirements, the ability to practice practical skills in the program can be limited for many of the students, who study the program at a distance through a CQUniversity study centre.

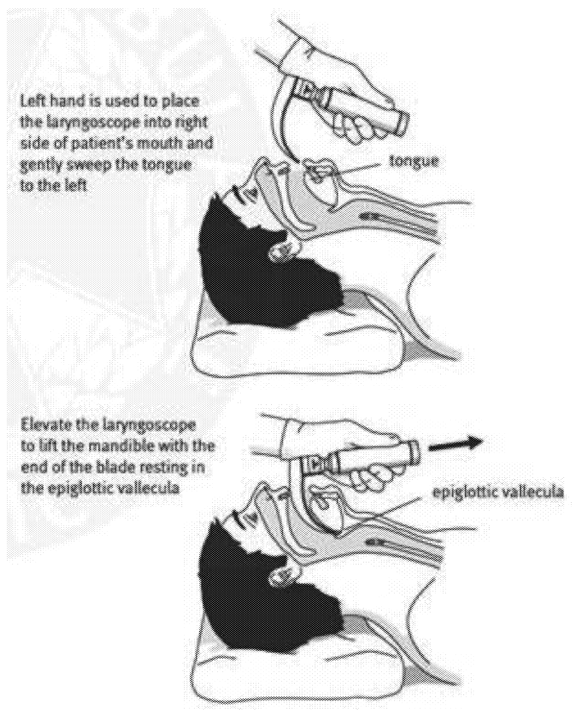


Figure 2: QAS Laryngoscopy Procedure

This is managed to some extent by the program coordinator and course coordinators by intensive residential schools that teach practical skills over the course of a week for various courses, such as the aforementioned skills course, as well as through work integrated clinical courses in the 2nd and 3rd year of the program and the encouragement of students to volunteer for the St John's Ambulance service. However, especially for the second year courses such as the skills development course, more opportunities for practical skill development could be useful for students, and this is discussed further in the next section.

PROBLEM

The CQUniversity paramedic science program seeks to equip students with the practical skills required to be a paramedic, but for distance students the amount of hands on time they have to practice skills can prove challenging. This project stems from a need, identified through course evaluations, for more opportunity for distance students to practice skills (currently, they can only be practiced in the five day residential school). An anonymous sample of the comments made by students in course evaluations is provided in Figure 3 supporting this assertion.

"I believe that because this course is a 'skills' learning course, that there should be a way for us to actually get more time doing skills. I feel that as distance students we are at a severe disadvantage because we spend 5 days doing them in the middle of term and then don't do them again until we hit our placement."

"There is no substitution for experience. Could the school look into either some kind of software or equipment that we could be supplied with so that we can at least go through the motions of doing the skills?"

"I believe that my confidence in performing the procedures and skills could have been improved with a little more 'hands-on' time."

"I feel as an external that I am missing out - they do scenarios every week, I did one or two during res school".

"Studying by distance you can read the skills and kinda do scenarios but it's hard to get feedback and to know if what you're doing is still right."

Figure 3: Examples Student Comments Relating to Hands-On Skills Practice

The proposed intervention seeks to resolve these issues highlighted by students. Selected students will be provided with 3D printed instruments and a mobile application that they can use to practice skills prior to arrival at the residential school. Additional training will also be provided to those students not selected using the tools during the residential school. Outcomes are expected to be an overall improvement in final skill level for all students and immediate attention to the issues raised in the comments in Figure 3. More details on this intervention are provided in the next section.

PURPOSE

The aim of this research is to provide more hands-on skill practice to students, as well as increase overall skill acquisition and retention, focusing on Laryngoscopy with foreign body removal, answering the question "How does the use of 3D printing and augmented reality simulation affect skills development in paramedic science?". The expected outcome of this work is a greater understanding of how 3D printing and augmented reality can assist with skills development and insight into whether these techniques can lead to better learning outcomes.

Whilst previous work has been completed in the field of 3D printing and virtual simulation effect on learner perceptions (for example see Birt and Hovorka (2014)), this project represents a novel combination of the two technologies as a mechanism to test for enhanced skills development. Work in this area will move forward research into the areas of 3D printing and augmented reality for teaching applications and will ultimately benefit the teaching community by providing a greater understanding of how these technologies can be used in classroom practice.

PROPOSED INTERVENTION

Following an action research methodology (Kemmis, 2006), the proposed intervention will involve modifying practice for a subset of students studying the course, creating two learning groups – those with advanced access to the 3D tools and those without. The context of the simulation is a Laryngoscopy procedure with foreign body removal (See Figure 2). To assist in immersion and accuracy a 1:1 scale replication of the actual physical tools is required. In this case we require a 3D printed Laryngoscope with Mac Blade and Magill Forceps (Figure 4). Through the addition of AR markers, these physical models can be tracked and simulated in the virtual game environment (Figure 5).

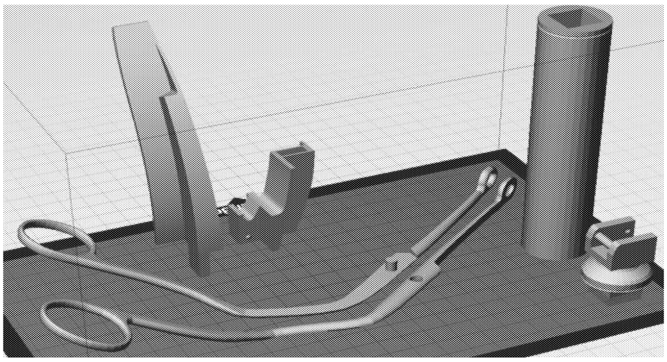


Figure 4: 3D Models representing Forceps and Laryngoscope.

For the augmented reality game simulation the decision was made to use the game engine development platform Unity 3D, (Unity Technologies, San Francisco, CA) and Vuforia AR plugin for Unity 3D (Qualcomm Technologies, Inc). This software is primarily designed for working with object-oriented, multimedia game content and provides a rapid means for deployment to multiple operating systems and mobile device platforms (Figure 6).

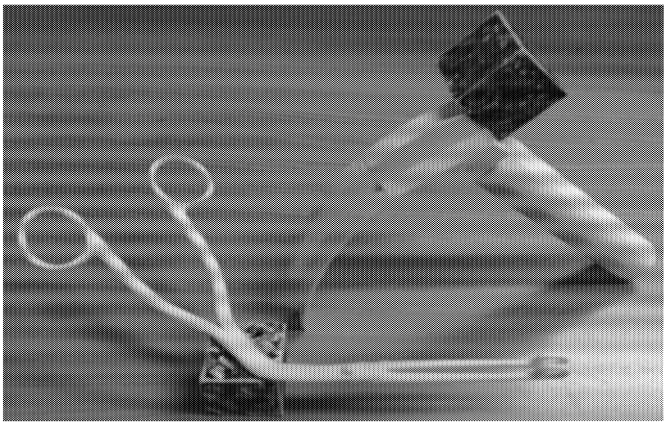


Figure 5: 3D Printed Forceps and Laryngoscope with AR Markers

To allow for the correct view point (looking down the throat of the airway manikin) and keeping the hands free (to use the printed tools with 3D augmented markers) the decision was made to design a 3D printed universal smartphone mount that could be attached to a hat the student wears (Figure 7).

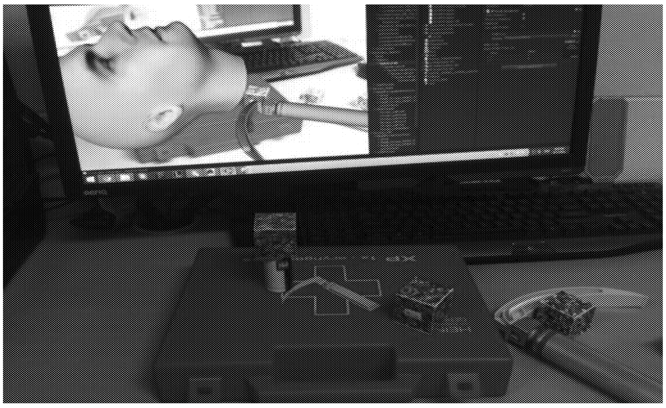


Figure 6: Example of 3D Simulation Objects Visualised in Unity3D with Vuforia Augmented Reality Plugin

The focus of the simulation to be "task appropriate" means only pertinent information relating to the key learning outcomes be included. To this end an audio cue is presented to the user asking them to present the AR tracking markers to the device camera. When all markers have been identified the virtual objects are then intantiated over the camera view of the actual physical tools indicating to the user that the simulation has recognised the markers. An airways maninkin dummy is displayed and a series of steps with audio and visual cues is presented to the user. The aim of the simulation is to follow the steps required to insert the Laryngoscope correctly and then the forceps to remove a foreign body lodged in the patients throat, with cues provided during the simulation to indicate whether the procedure (game) has been successful.

The 3D printed Laryngoscope, Magill Forceps and Phone mount will be provided to Distance Students ahead of the residential school along with the smartphone application. Students will then use the game and printed items to practice foreign object removal on a virtual 3D patient during semester. The game will help students develop correct technique through visual and auditory feedback as the procedure is completed.

RESEARCH METHOD

A stratified sample of 30 students (out of an approximately 120 student cohort) will be selected from the Term 1 2015 cohort to receive the 3D printed instruments (see Figure 5) and access to the augmented reality application (see Figure 7). These students will be given instructions on how to use the tools and will be encouraged to practice prior to the residential school scheduled for late in the term. When students arrive, a pre-test will be conducted with all students to assess skill competency prior to the residential school and to assess the difference between selected students and other students. The traditional "hands on" training will then be provided to all students in the residential school, with an intermediate skills test conducted after this training. Finally, students not selected for the trial will be given extra training using the 3D printed tools and augmented reality application. A final post test will then be conducted with all students and a survey issued to assess how they felt about the intervention methods and use of the tools.

Research data collected will be based on the skills assessment of the students as well as their responses to the survey administered at the end of the skils testing after the residential school. Time commitment from participants will be determined by the students, with the time commitment during the residential school remaining the same as in previous offerings and the only additional time commitment being the time students spent at home using the tools prior to attending the residential school.

Data collected from the skills test will be analysed using SPSS to determine correlations and cross-tabulations between the two groups. Data from the surveys will be analysed using a combination of SPSS for the closed questions and NVIVO for the open questions, with the SPSS data providing correlations between student answers and demographic data, whilst the NVIVO data allows for coding and categoristation and identification of key themes for further research.

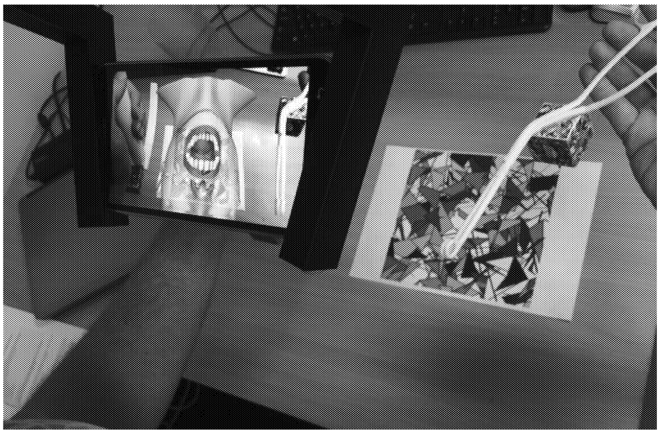


Figure 7: Mounted Mobile Phone with 3D Simulation

EXPECTED RESULTS

It is expected that students using the 3D printed objects and simulation game application will perform better on the pre-test, with the selected students requiring less "time on task" teaching at the residential school. More generally, it is expected that there will be an overall improvement for all students in the post-test for final skills level.

CONCLUSION

This paper has presented a proposed intervention and research design to assess the applicability of 3D printing and an augmented reality game simulation to improve skills development for students studying paramedic science at a distance. Through the use of these interventions, it is theorised that students will have more time on task and therefore perform better with their skill development.

Through the use of an action research paradigm, several tests will be performed at various stages to assess this assertion and student skill levels in performng the task that has been simulated. In addition, a survey will be conducted to assess student attitude towards the intervention methods.

Future work will report on the results of this study and provide correlations of various factors related to student performance, showing whether the use of these interventions have improved skills development and whether the tools were accepted by the student cohort. Through this work, a greater understanding of the use of innovate technology tools and games simulation in the education space will be obtained, providing a foundation for future work in the area.

REFERENCES

Bennett, S., & Maton, K. 2010. Beyond the 'digital natives' debate: Towards a more nuanced understanding of students' technology experiences. *Journal of Computer Assisted Learning, 26*(5), 321-331.

Birt, J. & Hovorka, D.S. 2014. Effect of mixed media visualization on learner perceptions, *Proceedings of the 25th Australasian Conference on Information Systems*, AUT, Auckland, New Zealand, December.

Dalgarno, B. and Hedberg, J. 2001. 3D Learning Environments in Tertiary Education. *Proceedings of the 18th annual conference of the Australasian Society for Computers in Learning in Tertiary Education*. (pp. 253-262). Melbourne, Australia.

Freitas, S. d., and Neumann, T. 2009. "The use of 'exploratory learning'for supporting immersive learning in virtual environments," *Computers & Education,* (52:2), pp 343-352.

Kozma, R. B. 1991. "Learning with media," *Review of educational research*, (61:2), pp 179-211.

Kozma, R. B. 1994. "Will media influence learning? Reframing the debate," *Educational technology research and development*, (42:2), pp 7-19.

Höffler, T. N. 2010. "Spatial ability: Its influence on learning with visualizations—a meta-analytic review," *Educational Psychology Review*, (22:3), pp 245-269.

Huk, T. 2006. "Who benefits from learning with 3D models? The case of spatial ability," *Journal of Computer Assisted Learning*, (22:6), pp 392-404.

Hwang, W.-Y., and Hu, S.-S. 2013. "Analysis of peer learning behaviors using multiple representations in virtual reality and their impacts on geometry problem solving," *Computers & Education,* (62), pp 308-319.

Loy, J. 2014. "eLearning and eMaking: 3D Printing Blurring the Digital and the Physical," *Education Sciences*, (4:1), pp 108-121.

Mayer, R. E. 2002. Multimedia learning. Psychology of Learning and Motivation, (41), pp 85-139.

Mayer, R. E. 2005. *Cognitive theory of multimedia learning*. The Cambridge handbook of multimedia learning, pp 31-48.

Moreno, R., and Mayer, R. E. 1999. "Cognitive principles of multimedia learning: The role of modality and contiguity," *Journal of Educational Psychology*, (91:2), pp 358.

Prensky, M. 2001. Digital natives, digital immigrants. *On the Horizon, 9*(5), 1-6.

Prensky, M. 2009. H. sapiens digital: From digital immigrants and digital natives to digital wisdom.

BIOGRAPHIES

MICHAEL COWLING PhD, MBA, EdD(c) is an information technologist with a keen interest in educational technology and technology ubiquity in the digital age, especially as it relates to International students and those from non-English speaking backgrounds. He is currently a Senior Lecturer in the School of Engineering & Technology at CQUniversity Australia. Dr Cowling is the recipient of 3 CQUniversity Learning and Teaching grants related to teaching technology and was a 2007 recipient of the CQUniversity Award for Excellence in Learning & Teaching (International Campuses). He is also a regular contributor in Australian radio and print media on the topic of Educational Technology and Technology Ubiquity.

EMMA MOORE is a lecturer in the School of Medical and Applied Sciences at CQUniversity. She has come to the University from industry were she has been working as an Advanced Care Paramedic (Level Two) for almost 9 years. Emergency Medicine has always been apart of Emmas life with both her Father and Grandfather working for the Ambulance Service as Paramedics as well. She teaches into the skills development courses of the Bachelor of Paramedic Science. Her research interests are in improving skills development for distance paramedic science students.

JAMES BIRT, B.I.T. (Hons), Ph.D., is an Assistant Professor of Computer games and Multimedia in the Faculty of Society and Design at Bond University Gold Coast, Queensland, Australia. In 2014 James was awarded the prestigious Australian Office of Learning and Teaching citation for outstanding contribution to student learning for his work on improving student learning and engagement through practical juxtaposition of art and science in multimedia education. His research interests are in the areas of mixed media visualization and education technology.

EVALUATION OF THE PROCESS OPERATOR KNOWLEDGE FORMATION RESULTING FROM COMPUTER-BASED TRAINING

Victor Dozortsev
Vladimir Nazin
ZAO "Honeywell"
7 Kievskaya str.
121059, Moscow, Russia
E-mail: victor.dozortsev@honeywell.com
E-mail: vladimir.nazin@honeywell.com

Alexander Oboznov
Svetlana Gutsykova
Institute of Psychology of Russian Academy of Sciences
13 Yaroslavskaya str.
129366, Moscow, Russia
E-mail: aao46@mail.ru
E-mail: gutsykova@rambler.ru

Anastasia Mironova
Moscow Institute of Physics and Technology
9 Institutskiy per.
141700, Dolgoprudny, Moscow Region, Russia
E-mail: mironova@phystech.edu

KEYWORDS

Conceptual model, operator, computer-based training, preliminary training.

ABSTRACT

This paper aims to grasp the changes that the operator experience during computer-based training, and how these changes could be measured. Tools for operator's conceptual model (CM) evolution investigation are discussed. These are 1) certain approaches to training results assessment, 2) preliminary operator training and experimental verification of its efficiency, 3) CM structure and criteria of its assessment, 4) correlation of training success and characteristics of operator's personal features. Results of the experimental research of the CM dynamics are given.

INTRODUCTION

Computer-based training of process operators is an acknowledged element of industrial safety. Dramatic number of accidents (The 100 Largest Losses. 1972-2011 2012), demographic problems and lack of high-skill personnel (Wilkins 2007) require effective methods of training.
Achievements in the area of operator training systems (OTS) are significant: OTS market is huge (1 billion of dollars per year including related hi-tech issues of industrial automation http.arcweb.com) with all OTS components being highly-developed (mathematical models and management environments are complex and accurate, instruction tools are quite advanced (Ortloff 2010)). OTS are good for investments into industrial safety. Average effect of using one training system during its lifecycle is estimated at 18 million dollars (Komulainen, et al. 2012), and annual effect, calculated on the basis of long-term statistics of accidents due to the operator's faults, reaches 0.5 dollar per ton of crude oil per year (Dozortsev 2009). However, despite of these achievements, OTS developers and users suffer shortage of its functionality. Lots of publications are devoted to the experience of development and usage of OTS, this is an area of spirited debates.
Let us focus on the technique of computer-based training. Its influence on the quality of training is not investigated as deeply as the modeling aspects. In this context research efforts focus on the following:
- classification of computer-based training systems,
- structure and amount of training exercises,
- estimation of the training results (Dozortsev 2009).

But the question "*How to measure the impact of training on operators?*" is open.
It is simple to assess the quality of coping with exercises and post-training tests. Based on the results one can formally evaluate operator's "actual" competency level. But this estimation gives no answer to the question stated above, because it does not guarantee operator's success in a different situation. The important questions are: how successfully will the skills learnt during the training process be transferred to the real practice, and will the structure of operator's knowledge change as a result of the training. Apart from that operator's success dependence on the personal features is of interest.
All these issues are important for the people who are responsible for personnel training. They need reliable feedback from training system, which would reflect the objective level of operator's competency apart from the valid training itself.
The research in area of knowledge structure and its characteristics assessments is quite rare. Serious efforts are summarized in the doctoral thesis (Burkolter 2010) with most experimental and practical application in the field of aerospace.
The remainder of this paper is structured as follows. The section below describes different approaches to improve the quality of operator's training. Namely, the preliminary training and CM structure changes assessment (knowledge structure and cluster technique). The next section presents the experimental study of the CM evolution evaluation, the following section contains results discussion and future work outlines. The last section concludes this work.

Let us discuss the background and means of investigation of the CM dynamics.

INVESTIGATION OF THE CM DYNAMICS RESULTING FROM COMPUTER TRAINING

Training results assessment overview

The key principle of OTS quality assessment is the correspondence of operator's activity in training with the real activity. Modern training systems allow achieving the high level of this kind of similarity. They are precise in process modeling and the management environment can be simulated perfectly. Training methods are based on modern concepts of what the operator's cognition needs during the process control. The latter includes hierarchical structure of various skills and mental background (Dozortsev 2009).

High quality training system is prerequisite of operator's success but is not a guarantee. Much depends on training organization. One should be able to estimate a trainee progress and modify the training program on its basis. For instance, one can provide re-training in case of gaps in some of the program sections. Based on formal training results or quality of real process control, it can be regulated by training instructors.

There are known approaches to personnel professional skills assessment using the "responsibilities – competencies – behaviors – proficiency levels" (Pilon, Jonge and Ross 2012) structure. Operator's *responsibilities* can be described by *competencies* which are determined by the professional role (e.g. process control in normal mode, predict and react to the consequences of abnormal situations). Competencies comprise particular operator's *behaviors*. For instance, to control a process in normal state, console operator must read the transmitters, handle the controls, interact with other operators and etc. The latter could be measured by means of training scenarios, tests, etc. related to the behaviors.

Proficiency level (from "beginner" to "expert") is assigned with regard to the exercises performance measurements.

These assessments are significant in the operator's training, but they do not show the changes in trainee's cognitive characteristics as the training outcome.

A rare exception is the publication of (Burkolter 2010) that stated a question of how training can best support process control performance in consideration of individual differences. In particular, one of the goals of this work is to show connection between knowledge structure and training outcome in process control.

The approach presented in our work is based on studies of structural characteristics of operator's conceptual model (Oboznov, Chernetskaya and Bessonova 2013).

About operator preliminary training

How the operator's knowledge evolve during professional growth? After the set of initial trainings, one takes an exam for admission to work and fortifies the skills further in the working practice. At some stage one may face the OTS training. However, complex computer-based training is meaningless for an unprepared person. First, one should develop basic control skills. Primarily, develop the understanding of cause-and-effect links in the system under control, and learn to diagnose causes of system failures.

There are technical means and methods of preliminary training that were implemented in various factory training centers (Dozortsev and Nazin 2011). Following is the "Diagnose" computer training system overview which was used in one of the experimental researches.

The trainee plays step-by-step game with "Diagnose" which imitates diagnostic decision making process in real practice. Game is based on a subset of process failures and measurable process state parameters. System randomly selects one of the failures and keeps it in a secret. The trainee is informed about one of the symptoms (parameter deviations from normal values) relevant to the unknown yet failure. Trainee's task is to identify the chosen failure. One can either choose one failure from the list and validate it (this action "costs" a lot of penalty in case of mistake) or examine the condition of other process parameters (this choice "costs" minimal penalty). Trainee's goal is to identify the hidden failure getting the minimum summary penalty.

In the beginning of the game many failures can be relevant to the initial picture of abnormal process state. In this case it is more profitable to request additional information about symptoms to narrow the candidates list. The key question at that point is "which parameter is better to request?", based on the preliminary understanding of each (candidates list) faults consequences in terms of process state deviation. The good answer to that question leads to a faster solution strategy.

Final score is calculated based on the comparison of obtained penalty points and points that could be gained if the trainee had examined all the process parameters symptoms. Note that the duration of game solution does not affect the score. The stress of time shortage in real life faults handling is replaced by the lack of information in the training game.

Verification of preliminary training efficiency

Objective changes of the operator's skills and professionalism during training and preliminary training can be measured. Experiment that proves reliable improvement of quality of fault diagnostics due to "Diagnose" system is described in paper (Oboznov, Dozortsev, et al. 2013). Here are the brief results.

Two groups of university graduates took part in the experiment (three people in each group). Groups were averaged over gender and work experience.

Experimental group was first trained with "Diagnose" system (training time is X) and then, after an hour break, tackled diagnostic problems using computer-based training system (8 problems were solved, average solution time is O_1). After a two-week break this group tackled new series of tasks (average solution time is O_2). Control group did not use "Diagnose" system and started with computer-based training. They had three series of tasks with an hour and two-week breaks (average solution time in series are O_3, O_4 and O_5).

Table 1:Design of the Experiment

Experimental Group	X	O_1	O_2
Control Group	O_3	O_4	O_5

The following statistically valid results were received (level of significance of Student's t-test was varied from 0.01 to 0.05 and sampling was 24 diagnostic problems for each group and series):
A) $O_2<O_1$ and $O_5<O_4<O_3$ (training effect that remains in both groups even with long breaks);
B) $O_1<<O_3$ (preliminary training improves the results of the first computer-based training series). This result was quite expected because in addition to the hypothetical effect that we wanted to check, preliminary training gave additional acquaintance with the process.
C) $O_1<<O_4$ and $O_1<O_5$ (preliminary training improves the first series of experimental group using OTS in comparison with the second and even the third series in control group). This is the main result that proves the utility of preliminary training.
Hence, improvements in operator's performance in particular tasks achieved by a specific training can be captured. But this result does not prove the fact of objective changes in operator's cognition.

About changes in CM of operator

During self-training and strengthening professional background several types of CM are formed: technological, topological, functional, informational, algorithmic and imaginative. Each of these model types is constructed of unique, semantically consistent units. For instance, technological semantically consistent units (SCU) contain knowledge of certain process elements and engineering equipment (plant equipment, process sections, material and energy flows, etc). Functional SCU include knowledge of process elements conditions (switching equipment modes, execution of technological operations, etc). Different types of relations between SCU can be established: procedural, cause-and-effect, algorithmic and spatial. It is empirically approved that certain relations can correspond to certain types of CM (Galaktionov and Vavilov 1992). Thus, development of CM can be considered as the sequential inclusion of SCU of different types into initially formed CM. The aspect of operator's perception evolution of relations between semantically consistent CM units is especially important. In particular, cause-and-effect relations between SCU (concepts) of CM should be developed.
In one of the empirical studies nuclear plant engineers-operators were asked to estimate the level of interconnection between nuclear unit key-characteristics. 7 discrete values scale was used. Scores from 5 to 7 were used to express strong interconnection, i.e. the change in one characteristic always leads to the change in the other. Mean score (4 points) addressed the conditional connection, i.e. the change in one characteristic may lead to the change in the other. Low scores (1...3 points) regarded to the lowest chances to expect the change in one characteristic as the effect of change in the another. In the table 2 one can see non-linear relationship between operator's job experience and share of one's high values of key concepts interconnections in the whole set. This share is low when the operators have small (≤2 years) and big experience (10-25 years) compared to the operators having the average one (3-9 years) (Oboznov, Chernetskaya and Bessonova 2013).
The non-linear dependence of the high values share of cause-and-effect interconnections and operator job experience was explained through growing tendency to represent a nuclear unit functioning as determined and, therefore, controllable process. It takes a long time to create more detailed and adequate mental representation of nuclear unit functioning as a complex, ergatic (man-machine) system, consisting of many various sub-systems, with numerous connections and feasibility to exhibit unpredictable properties.
Study (Oboznov, Chernetskaya and Bessonova 2013) approach application to training efficiency assessment requires long-term research, because changes are happening in the long term and the factors affecting the training are tricky to capture. In paper (Oboznov, Dozortsev, et al. 2013) this approach was carried over "fast" training with the OTS. In present work which is a sequel to the research the idea of

Table 2: Subjective scoring distribution (%) of interconnection values for nuclear plant key characteristics expressed by operators with different experience (averaged over the group)

Job experience of an operator in the capacity of a lead engineer of reactor control	Subjective scoring of the interconnectedness of key concepts			Total
	High (5...7 pts)	Mean (4 pts)	Low (1...3 pts)	
Last year students (nuclear engineers) that already had a practice, (no job experience)	**20,0**	25,0	55,0	100
Job experience ≤2 years	**46,0**	17,0	37,0	100
Job experience 3…4 years	**53,0**	16,0	31,0	100
Job experience 5…9 years	**58,0**	18,0	24,0	100
Job experience 10…25 years	**37,0**	21,0	42,0	100

High scores of interconnectedness are in bold

CM dynamics measurement was verified in experiment. Therein operator's mental representations about object to control were strengthened by means of computer-based training. A number of applied and experimental problems were tackled for its realization:
— develop a new technique of CM structural changes estimation;
— prepare and conduct a proper psychological experiment;
— research the dependence of task success and CM characteristics dynamics on the trainee's personal features.

About CM assessment (cluster technique)

Estimation of CM structural changes requires construction of formal quantitative measure of models' inequality. In this paper following method is used.
One of the components of CM is operator's concept of cause-and-effect relations between key characteristics (parameters) of the process. Such parameters define influence on abnormal states of certain process units on the functioning of other equipment. In the proposed approach "influence" is expressed in a 6-point scale: from "0" (no influence) to "6" (maximum influence).
Set of influences between all pairs of parameters forms n-by-n matrix $E = (\mathrm{e}_{ij})$, where n is number of parameters.
Note that the essential property of influence is its asymmetry but further, for the estimation of "relation" between parameters, let's use symmetric matrixes:

$$c_{ij} = c_{ji} = \frac{1}{2}(\mathrm{e}_{ij} + \mathrm{e}_{ji}) \tag{1}$$

Row-vectors of matrix C correspond to points in n-space. These points represent a figure with n vertices. Distance between vertices characterizes similarity of corresponding vectors.
Groups of close vertices (clusters) are the structural features of concept of parameters mutual interference.
To analyze the difference between cluster structures let's examine the following measure. Denote $A_1, A_2, ..., A_{m_A}$ - clusters of the first matrix and $B_1, B_2, ..., B_{m_B}$ - clusters of the second matrix, where m_A and m_B - corresponding number of clusters. Let's evaluate a fraction of common elements of set A_i, and B_j relative to content of the set $A_i : \dfrac{n(\mathrm{A}_i \cap \mathrm{B}_j)}{n(\mathrm{A}_i)}$.
Enumerating all the index pairs i and j, we obtain m_A by m_B matrix. Here i-columns present distribution of cluster A_i elements over cluster structure B. Note that sum of the column elements equals to 1, because each element from A_i is contained in one of the sets of B_j. In the same time, the more elements from A_i disturbed over B_j, the more uniform column values spread over the scale [0,1]. For instance, if m_A=m_B=3 and $n(A_1)$=3, then maximum distribution corresponds to the case, when each of three elements from A_1 pass to B_1, B_2 and B_3. On the other hand , identity matrix corresponds to the case when two cluster structures coincide. For quantitative evaluation of the distribution it is convenient to use entropic function, like the following:

$$f_{A\subset B} = \sum_{i=1}^{m_A}\sum_{j=1}^{m_B} \frac{n(\mathrm{A}_i \cap \mathrm{B}_j)}{n(A_i)} \log_{10} \frac{n(A_i \cap B_j)}{n(A_i)} \tag{2}$$

This function takes on a value of zero at complete match of cluster structures and reaches a maximum in case of largest smearing of cluster composition from one structure to another.
In a similar way one can get the measure with respect to B_j sets:

$$f_{B\subset A} = \sum_{i=1}^{m_A}\sum_{j=1}^{m_B} \frac{n(\mathrm{A}_i \cap \mathrm{B}_j)}{n(\mathrm{B}_j)} \log_{10} \frac{n(A_i \cap B_j)}{n(\mathrm{B}_j)} \tag{3}$$

Values of $f_{A\subset B}$ and $f_{B\subset A}$ depend on the number of clusters in A and B structures and their size. Therefore, to compare cluster difference regardless of values m_A, m_B, $n(A_i)$ и $n(B_j)$ one should normalize $(f_{A\subset B}, f_{B\subset A})^{\mathrm{T}}$ by the length of the vector of maximum entropy difference:

$$f^* = 1 - \sqrt{\frac{f_{A\subset B}^2 + f_{B\subset A}^2}{f_{\max(\mathrm{A}\subset\mathrm{B})}^2 + f_{\max(\mathrm{B}\subset\mathrm{A})}^2}}, f \in [0,1] \tag{4}$$

This value evaluates mutual similarity of two cluster structures A and B in the range from 0 to 1. Henceforth, this measure will be used in experimental study of the evolution of the process perception.

EXPERIMENTAL STUDY OF THE CM DYNAMICS

Main hypotheses

The basic hypotheses in the given approach are:
— operator's CM of the technological process evolves during computer-based training;
— mentioned changes can be estimated by investigation of cause-and-effect relations between the system elements.
— the CM evolution depends on the operator's personal features.

Experiment description

To verify hypotheses stated above the following experiment was carried out. 17 final year students from Russian State University of Oil and Gas took part in the experiment. Within the study program they explored the atmospheric distillation unit. The experiment was conducted in two steps, 4 hours each with two-week break in between.
During the first lesson students led by tutor investigated organization and normal state of the unit. During the second lesson they practiced emergency stop procedure. In the beginning of the second lesson testees filled up a form of 16PF questionnaire.

In the end of each lesson students filled the matrix described above. Matrix consisted of 13 key parameters of the atmospheric distillation unit (crude oil flow, overhead temperature, etc). Students estimated impact of each parameter on each another parameter on a 6-point scale from "0" (no influence) to "6" (maximum influence).
Since criterion of "proper" structure is unknown, evolution of the trainees' CM was estimated by using reference cluster structure that belongs to the tutor.

Experimental outcome

Changes in matrices structures

Table 3 shows changes between matrices of the first and the second training days. Strong impacts can be regarded as core in operator's process perception. Therefore, their share among all kind of impacts is an important property of CM.

Table 3: Changes in Matrices Structures

Student №	Average grade	% of strong impacts (5...6)		% of strong differences from reference matrix		% of strong differences between 1st and 2nd day matrices
		1 d	2 d	1 d	2 d	
1	4,89	13	17	8	8	8
2	3,67	7	5	9	6	3
3	4,67	5	6	13	15	3
4	4,11	8	9	8	5	3
5	4,89	7	10	10	8	5
6	4,11	14	12	11	11	14
7	3,33	5	3	3	4	7
8	4,78	11	29	12	13	11
9	5	28	12	13	5	13
10	4,89	6	8	5	5	3
11	3,89	2	4	5	4	1
12	5	17	11	7	6	3
13	4,78	5	11	16	15	5
14	3,89	26	33	19	21	6
15	3,33	14	29	15	13	22
16	5	10	14	10	12	5
17	4,11	12	20	9	13	4
Averag	4,4	11,	13,6	10	10	7
Std.	1	7	9	4	5	6

There are researches reporting that share of strong impacts grows with operator experience, then reaches maximum and decreases to plateau (Oboznov, Chernetskaya and Bessonova 2013).
In a given study changes are captured in a micro-scale (2 weeks) and the average share of strong impact increases. However, these changes are not statistically significant.
Despite of the remarkable positive change in the second day compared to the first one (7% of cases shows strong differences (≥4 points) between first and second training days), strong differences from reference matrix (≥4 points) change are insignificant (10% both days).

Cluster structures evolution

Clustering of the matrix vectors was conducted by hierarchical clustering procedure with complete linkage method.
Values of the cluster similarity f^* (4) are given in table 4.
Average shift turned to be positive and Wilcoxon criterion showed statistical significance (p=0.05). This is a core result that shows that trainee's cluster structures are approaching the reference one after computer training.

Table 4: Cluster Similarity Measure

Student №	Average grade	Measure of cluster similarity, %	
		1 day	2 day
1	4,89	50	80
2	3,67	38	34
3	4,67	42	17
4	4,11	59	79
5	4,89	62	59
6	4,11	36	42
7	3,33	61	79
8	4,78	65	78
9	5	47	62
10	4,89	26	28
11	3,89	13	29
12	5	66	79
13	4,78	31	25
14	3,89	40	40
15	3,33	41	78
16	5	39	44
17	4,11	39	29
Average	4	44	52
Std. Dev.	1	15	23

Personal characteristic and its influence on CM

Studies of contribution of personal characteristics to the construction of CM of sophisticated ergatic system operator are rare and insufficient. Nevertheless, such investigations are essential.
Present study uses 16PF questionnaire procedure (Gutsykova 2011). Sample group of students was divided into three equal subgroups. People of first group have tendency to give ultimative (extreme high or low) assessments, testees of the second group - neutral assessments and the third group testees - mixed ones.
Comparison of these subgroups was fulfilled by means of Mann-Whitney test. According to the results, subgroups that gave ultimative and neutral marks statistically differ in Q2 factor ("self-reliant - self-sufficient") (p=0,016). In contrast, subgroups that gave neutral and mixed marks shows

statically significant difference in Q3 factor ("perfectionistic - self-sentimental") (p=0,023).

Analysis of personal features in these subgroups shows that students inclined to give ultimative assessments are more self-dependent in decision making. Students of this subgroup are less degree focused on group judgments; they enforce the more in decisions made and more responsible for them (Q2 factor). Such people have both feet on the ground (I factor), they behave correctly but detached, they often underestimate themselves and think of consequences (O and N factors).

Students inclined to give neutral marks are less self-containment in decision making and more focus on group judgments (Q2 factor). They are ready to cooperate (A factor) and have positive self-feeling (MD factor). Such people are well-disciplined and care about reputation (Q3 factor).

Therefore, human personality shows influence on the relations assessments between ergatic system parameters even at the initial stages of CM formation. These assessments are determined not only by comprehension of technological process but also by personal features of the operator, in particular, his dependence on group judgments.

DISCUSSION OF THE RESULTS

CM allows operator to act relying on understanding of what happened, what is happening and what will happen with the complex ergatic system. It contains knowledge about cause-and-effect relations in such a system and the share of strong cause-and-effect links increases in the beginning of the career. That is why the focus of operator computer-based training and preliminary training on the development of cause-and-effect understanding is reasonable and is expected to significantly decrease time to reach higher levels of understanding.

Investigation of operator's process perception in complex ergatic systems requires further study. The present work demonstrates only the first outcomes of ambitious research. However, it assures that objective changes in a mental representation do happen during training, and their essence is the technological process conceptual model formation and evolution which are determined by operator's personal features.

CONCLUSIONS

Modern standards of industrial safety demand the high quality operators' training. Computer-based training is the acknowledged instrument to target this demand. To achieve maximum effect it can be complemented by preliminary training. An efficiency of such approach was shown in this paper. Another way to improve computer training methods is to study the explicit changes of CM during learning process. To assess those changes the cluster technique was suggested in this work. Statistically significant results were presented and discussed.

REFERENCES

Burkolter, D.S. *Training in Process Control: Supporting Performance in Consideration of Operator Characteristics.* Ph.D. Thesis, Freiburg University, March 9, 2010.

Dozortsev V.M., Nazin V.A. "Training of Fault Diagnosis Skills: Use of Heuristic Diagnostic Model and Technical System Simulation." *Proceedings of the 9th International Industrial Conference "ISC'2011".* Venice: EUROSIS-ETI, 2011. 273-278.

Dozortsev, V.M. *"Komp'yuternye trenazhery dlya obuchenia operatorov tekhnologicheskikh protsessov (Computer-based Training Systems for Technological Processes Operators)".* Moscow: Sinteg, 2009.

Galaktionov A.I., Vavilov V.A. "Analysis and Organization of Operator Activity." *Psychological Journal*, 1992: No.3, 92-104.

Gutsykova, S.V. "The structure of the personal characteristics of professionals with varying success of activity." *Bullenig of the University (State University of Management)*, 2011: No.19, 40-42.

Oboznov A.A., Chernetskaya E.D., Bessonova Yu.V. "Conceptual Models of the Nuclear Power Plant Operators with Various Job-Experiencel." *Psychological Journal*, 2013: No.4, 47-57.

Oboznov A.A., Dozortsev V.M., Nazin V.A., Gutsykova S.V. "Intellectual System for Process Operator's Mental Representations Formation." *Russian Scientific Journal*, 2013: No.7(38), 133-138.

WEB REFERENCES

Ortloff, Richard C. "Training Approaches for Using Simulators to Teach Process Control Systems." *Control Engineering.* November 30, 2010. Available through http://www.controleng.com/index.php?id=483&cHash=081010&tx_ttnews%5Btt_news%5D=40524.

Pilon S., De Jonge P., Ross M. "Operator Competency." [pdf] *Honeywell Process Web site.* 2012. Available at: https://www.honeywellprocess.com/library/news-and-events/presentations/HUG12-Honeywell-Ross-Operator-Competency.pdf.

"The 100 Largest Losses 1972-2011." [pdf] *Marsh USA Web site.* 2012. Available through https://usa.marsh.com/Portals/9/Documents/100_Largest_Losses2011.pdf.

Tina M. Komulainen, Ronny Sannerud, Bjarne Nordsteien, Havard Nordhus. "Economic Benefits of Training Simulators." [pdf] *Open Digital Archive at Oslo and Akershus University Colledge of Applied Sciences.* December 2012. Available at https://oda.hio.no/jspui/bitstream/10642/1544/1/939013.pdf.

Wilkins, Maurice J. "Changing Workforce Demographics Transform Manufacturing." [pdf] April 19, 2007. Available at: http://www.automation.com/pdf_articles/Changing_Workforce_Demographics_Transform_Manufacturing.pdf.

TEACHING FLUID POWER SYSTEMS BY REMOTE LABORATORIES

Marius D. Bontoș , Daniela Vasiliu, Nicolae Vasiliu, Bogdan Mihalescu
University POLITEHNICA of Bucharest
313, Splaiul Independentei, Sector 6
RO 060042 Bucharest, Romania
E-mail: vasiliu@fluid-power.pub.ro

KEYWORDS

Remote / virtual laboratory, web data acquisition, fluid power systems, remote / online education, open and distance learning.

ABSTRACT

During the last decade, additional teaching strategies have been developed by implementing Internet and Information Technology. One of the most challenging and attractive strategies is remote / online education applied to the traditional hands-on laboratories by two approaches: Remote and Virtual Laboratories (Ertugrul N., 2000, Ma J., Nickerson J.V., 2006, Z Aydogmus, O Aydogmus, 2009). New friendly development environments like graphical programming platforms, distributed and virtual systems became more attractive to promote advanced experiment based teaching. These environments facilitate the simulation of complex, expensive and dangerous equipment, the access to the conducted experiments for a large number of users regardless of their location and the improvement of quality of the experiments. In this paper, the authors describe an original way of using a remote laboratory, built on a LabVIEW platform, in order to teach fluid power systems in power, automotive, and aerospace engineering in the University POLITEHNICA of Bucharest.

INTRODUCTION

Current development of the education system along with the rapid implementation of Internet and Information Technology have generated a trend in developing new methods and techniques used for teaching and learning, leading to an efficient distance education (Ursuț iu D., Cotfas P., Samoilă C., Cotfas D., 2007).

Therefore, the need to develop open and distance learning became an indispensable part of educational systems and has had an enormous impact on engineering, freeing students from the constraints of time and place, and offering flexible learning opportunities to individuals and groups of learners (Ibrahim, W. and Morsi, R., 2005).

The hands-on laboratories are important elements in the technical education. They allow verification of theoretical knowledge and solve practical problems. The laboratory is often a complex structure, so that students cannot safely work directly with laboratory facilities. A high-level technical assistance is always needed.

Easier access to laboratory facilities is an ongoing concern in technical education, some important steps in this area have been made by prestigious universities but also by various software companies.

Additional teaching strategies such as simulation, interactive illustration, demonstration, experimentation, operation, communication and implementation of Virtual Instrumentation concepts have led to the development of special laboratories in the technical field called Remote Controlled Laboratories (Selmer et al. 2007).

New friendly development environments like graphical programming platforms such as LabVIEW (short for Laboratory Virtual Instrumentation Engineering Workbench), Matlab/Simulink, Java Applet, Flash or other software used to simulate the lab environment became more attractive to promote advanced experiment based teaching (Ertugrul N.,2000). All these facilitate the simulation of complex, expensive and dangerous equipment, and the access to the conducted experiments for a large number of users regardless of their location, along with the improvement of quality of experiments (Chen X., Song G. and ZhangY., 2010, Kehtarnavaz, N. 2008)

In traditional fluid power systems laboratory, students have to identify the structure of the testing system, to generate various inputs, and finally to measure certain operational parameters using analog instruments, that requires a considerable amount of laboratory time in order to understand the influence of each parameter on the systems performances.

Modern fluid power systems laboratory implies the use of data acquisition systems implemented on industrial computers with real time control software conducted and controlled remotely through the Internet or a Local Network (Schmid C., 2001).

The development of a unified user-friendly remote laboratory publishing tool for fluid power systems is a great demand because it implies big amount of knowledge in other different fields than the fluid power one: SCADA systems, sensors, transducers and digital control languages. These can be avoided by the aid of a friendly data acquisition interface with simple input entities, which reflects the test strategy only, and the data storage into relational database for a faster numerical and graphical presentation. (Vasiliu D., Vasiliu N. et al. 2014).

The goal of this paper is to imply a new modular education framework based on virtual instrumentation that allows the study of both static and dynamic behavior of the hydraulic systems and enables students to focus only on the main objective of the laboratory. The paper is structured in three major parts: description of proposed laboratory framework, the remote application and experimental results presentation.

MODULAR REMOTE-BASED LABORATORY FRAMEWORK

The block diagram of the Remote Laboratory Framework is shown in Figure 1. Each test bench sensors is connected to a real time machine (PXI) through a NI data acquisition board. The PXI machine also works like real time communication server, web server, database server, media streaming server and content manage server. The camera allows the user to see the system response in real time.

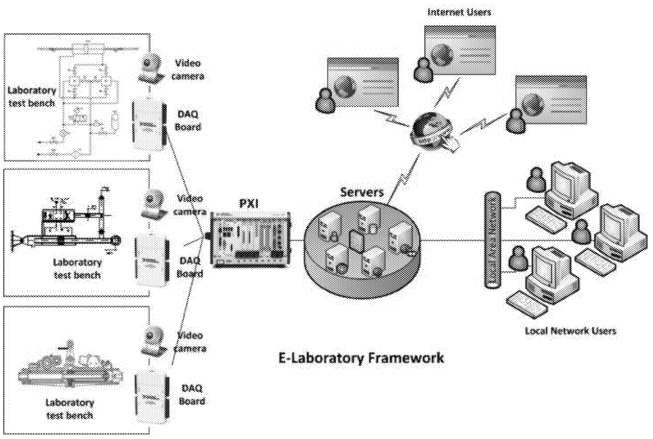


Figure 1: The Remote Laboratory Framework diagram

Hardware

The hardware structure of the remote laboratory for hydraulic servomechanism was developed on a modern test bench. The hydraulic servomechanism receives input signals form an electrohydraulic servomechanism controlled by an analogue interface. The purpose of this arrangement is to obtain an accurate and repeatable input signal. Traditionally a signal generator commands this excitation servomechanism. The output of the servomechanism is applied to a rocker, which replace the load of the mechanism. The displacement of the rocker arm is measured by the data acquisition system with the help of a position transducer. In the remote laboratory arrangement LabVIEW programmable signal processor, controlled locally or remote by the end user supplies the command signal. Figure 2 presents a generic view of the test bench.

Figure 2: Virtual laboratory for hydraulic servos

The PXI computer contains the both the signal generator and the data acquisition and analysis software. Figure 3 presents a generic hydraulic diagram of a hydraulic servomechanism. The free small joint of the error bar is driven by the small electrohydraulic servomechanism controlled by the LabVIEW remote laboratory software. The real view of the tested system is presented in ANNEX.

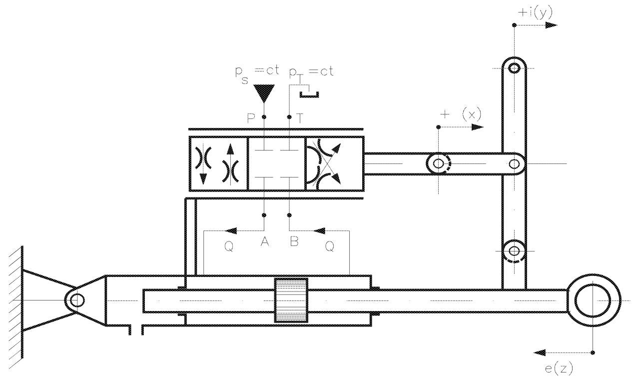


Figure 3: Hydraulic diagram of a hydraulic servomechnaism

Software

The control software of the platform was developed with LabVIEW graphical programming language. It consists of four independent modules: laboratory management, experiment control, data acquisition and data analysis. Out of the four modules, the remote user has access only to the control (signal generation module) and data analysis module. The data acquisition module works with a predetermined set of parameters adequate for the programmed test. Figure 4 presents the block diagram for the data analysis module.

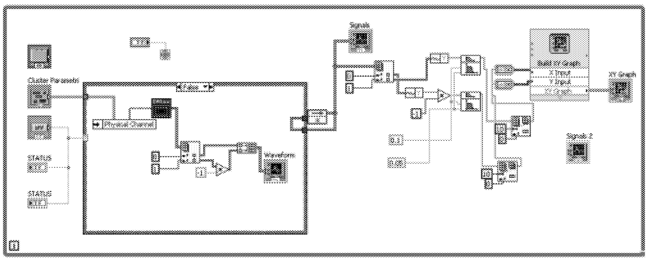


Figure 4: Block diagram of the module for data analysis (partial view)

The data analysis module accepts inputs supplied by the data acquisition module and data read from a previously logged file.

REMOTE LABORATORY APPLICATION

The remote laboratory software application has been developed based on LabVIEW software and MySQL database management system. The associated database contains tables with all the needed information for application's operation, such as:

- Users: name, affiliation, login and contact data;
- Laboratories attributes: name, the necessary time to carry out tests, the initialization period;
- Planning work sessions: users, laboratory name, start date;
- Results after performing samples; each user has a defined area, the access to this area being restricted to other users.

The application manages the information from the database allowing its user's to remote control the fluid power system laboratory stands and access the resulting data. Users can visualize and create defined reports for each completed laboratory. The application works in the same way whether is accessed through the web platform or directly from the laboratory bench and consists from five linked learning steps (figure 5).

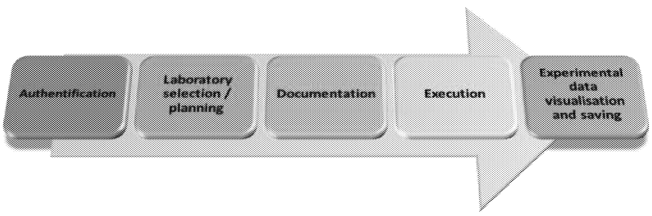


Figure 5: Remote application learning steps

User authentication and documentation

The applications interface between users and remote laboratories module can be accessed through a web browser for each particularly user. Each must provide a valid username and password and after a successful authentication, they can request access to a stand or laboratory if entitled (based on a multi-level user access).

Laboratory selection / planning

The selection / planning interface provides a large amount of information like user login name and affiliation, list of completed laboratories, list of laboratories that can be performed and their planning (Figure 6).
Each logged user can view their performed laboratory list and by selecting one of them, the recorded data will be available for visualization and saving.

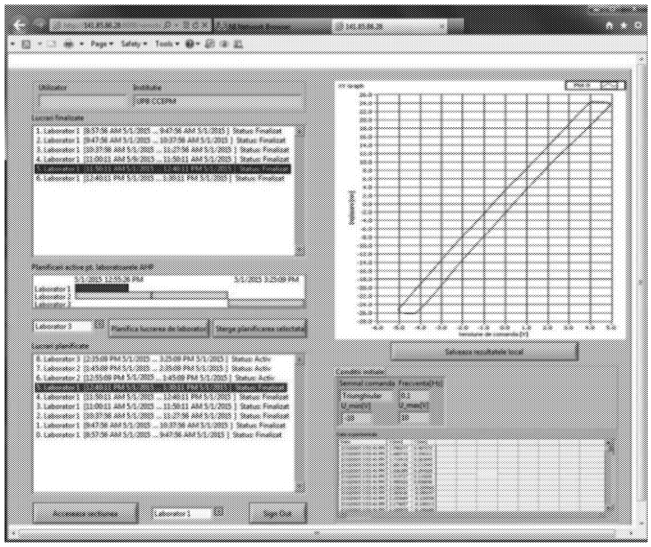


Figure 6: Laboratory selection and planning interface

Also, through the selection of one of the available test bench users can plan the laboratory work. A laboratory cannot be performed without prior planning. This planning is necessary because one user can access only one laboratory at a time; therefore, the planning system does not allow two connections simultaneously by two different users. After laboratory selection, users will go through the mandatory laboratory procedure that contains the necessary theoretical information needed to understand the processes and description of steps performed in laboratory work. All presented materials are available for download.

Laboratory procedure and experimental data visualization

The next step of the learning process is the stand control interface (Figure 7). The test bench can be viewed through a media-streaming server installed on the PXI machine. It provides high-resolution images at 30 frames per second published it the right area of main laboratory interface. Also, the interface contains a necessary parameters setting area such as: form, frequency and amplitude of the input signal.

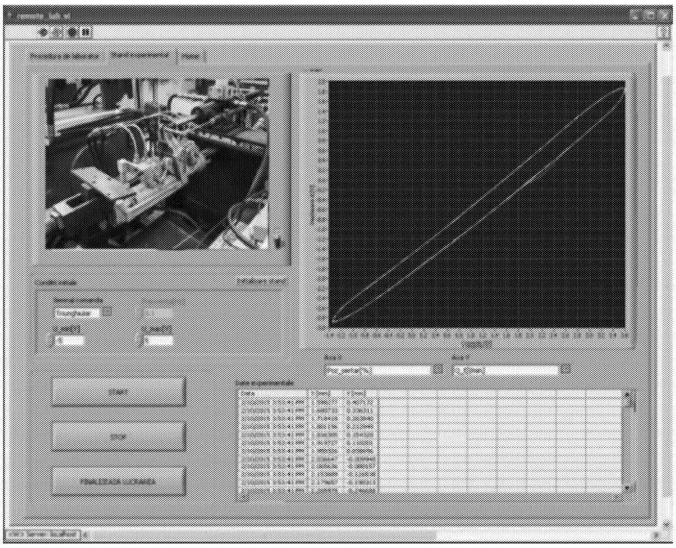


Figure 7: Main laboratory interface.

Start button turns on the acquisition of experimental data. After the acquisition, experimental data is represented graphically in the upper right corner, and numerically in the lower right corner. The users can select the measurement to be represented on each axis.
At the end of the experiment session users can simply stop the program by using the "Stop" button. This action does not save data on the database. In order to save data and return to the laboratory selection and planning interface they must press the "Finalize" button.

EXPERIMENTAL RESULTS PRESENTATION

The experimental results can be visualized on-line or the data files can be downloaded for offline processing.
The figures 8 to 13 presents typical graph for the hydraulic servomechanism obtained form data acquired with the remote laboratory test bench.

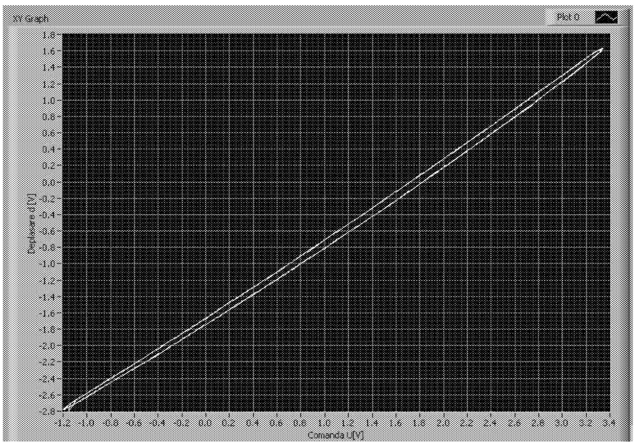


Figure 8: The steady state characteristic of the servomechanism on 0.1 Hz input signal frequency.

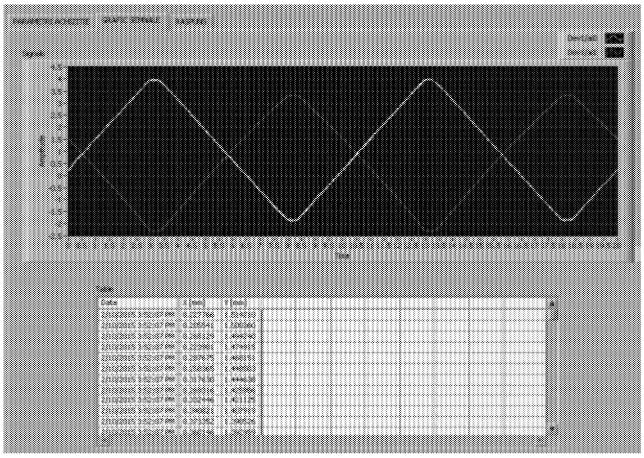


Figure 9: The raw time-domain data of the servomechanism on 0.1 Hz input signal frequency.

Figure 9 presents a typical steady state characteristic obtained with a proper low frequency triangle shaped test signal. The test signal used is presented in figure 10. Improper use of the test bench is exemplified on figure 11 were a higher frequency test signal (figure 12) does not allow a steady state behavior of the servomechanism.

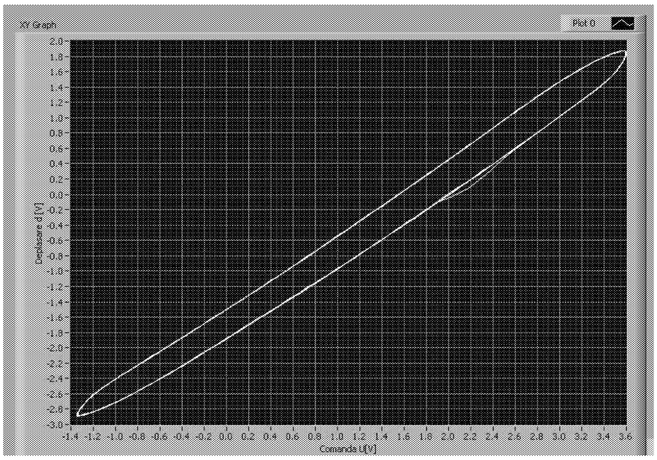


Figure 10: The steady state characteristic of the servomechanism on 0.5 Hz input signal frequency.

Thus, the graph in figure 11 is influenced by the overall response time of the servos, which is rather slow due to the high mass of the inertial load formed by the rocker arm.
The test bench can be used to determine the response time of the servomechanism by applying a square wave input signal. The result of such a test is presented in figure 12. A typical sine input response is presented in figure 13.

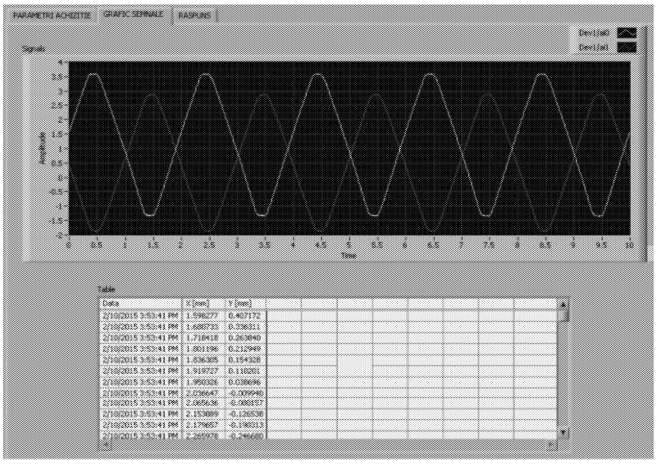


Figure 11: The raw time-domain data of the servomechanism on 0.5 Hz input signal frequency.

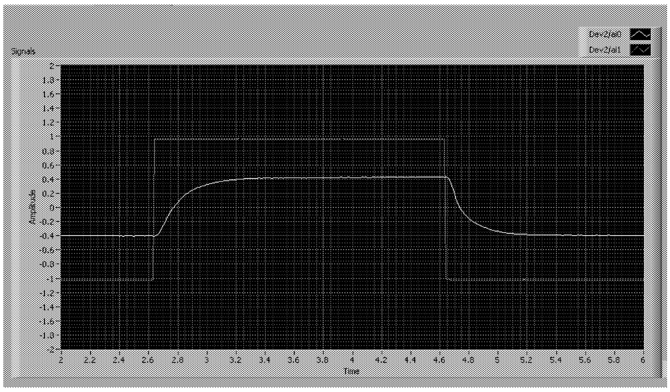


Figure 12: The response of the servomechanism to a step input.

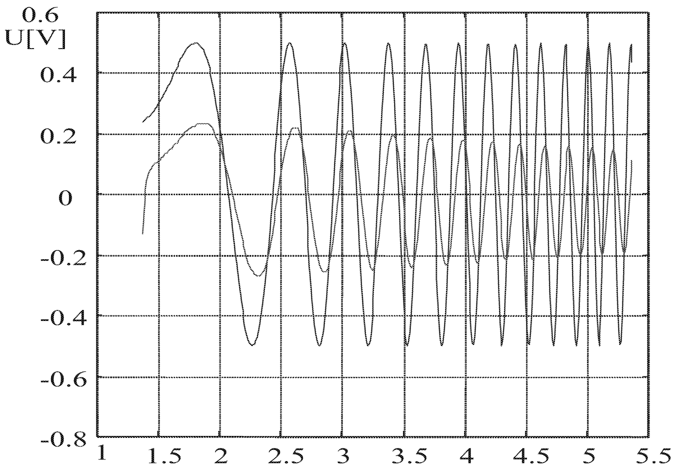


Figure 13: Typical sine input response

CONCLUSIONS

This paper presents the development of remote laboratories by identifying new ways and new areas of applying on-line laboratory technology. The availability of simple and efficient method of result dissemination and remote use of virtual systems becomes attractive to increase the access to students' experimental practice.

Concurrent engineering approach implies a tight collaboration between specialists in many area of expertise. They can rarely find the opportunity to be together in the same laboratory at the same time. A remote or virtual laboratory approach allows for an easier collaboration and shortens the development cycle of a product. Therefore is important to expose students to such a laboratory environment early in their formative years.

ACKNOWLEDGEMENTS

The work has been funded by the Sectorial Operational Program - Human Resources Development 2007-2013 of the Ministry of European Funds through the Financial Agreement POSDRU/159/1.5/S/132397.

NATIONAL INSTRUMENTS CORPORATION made a strong infusion of software in the framework of a permanent co-operation with the Fluid Power Laboratory from the Power Engineering Faculty.

REFERENCES

Aydogmus Z., Aydogmus O. 2009. A web-based remote access laboratory using SCADA, Education, IEEE Transactions on, Volume:52 , Issue: 1, pp.126-132.

Ertugrul N. 2000. Towards virtual laboratories: a survey of LabVIEW-based teaching/learning tools and future trends, International Journal of Engineering Education.

Chen, X., Jiang, L., Shahryar, D., Kehinde, L., and Olowokere, D. 2009. "Technologies for development of virtual and remote laboratories – a case study." Proc., American Society for Engineering Education Annual Conference & Exposition, ASEE, Austin, TX.

Ibrahim, W. and Morsi, R. 2005. "Online engineering education: A comprehensive review." Proc., American Society for Engineering Education Annual Conference & Exposition, ASEE, Portland, Oregon.

Kehtarnavaz, N., 2008. "Digital Signal Processing System Design: LabVIEW-Based Hybrid Programming", Academic Press Elsevier, San Diego, California.

Ma J., Nickerson J.V. 2006. Hands-on, simulated, and remote laboratories: A comparative literature review, ACM Computing Surveys (CSUR), Volume 38, Issue 3.

Schmid, C. 2001. "Virtual control laboratories and remote experimentation in control engineering." Proc., 11th Annual Conference on Innovations in Education for Electrical and Information Engineering. 213–218.

Selmer, A., Kraft, M., Moros, R., and Colton, C. 2007. "Weblabs in chemical engineering education." Education for Chemical Engineers, 2(1), 38–45.

Ursuţiu D., Cotfas P., Samoilă C., Cotfas D. 2007. Tehnologii Utilizate în Laboratoarele Controlate la Distanţă, Conferinţa Naţională de Instrumentaţie Virtuală, Ediţia a IV-a, Bucureşti

Vasiliu, N., Vasiliu, D., Calinoiu, C., Sorohan, Şt., Mihalescu, B., Negoiţă,G.C. 2009. "e-Learning System for Fluid Power Systems", Grant Report CEEX No.162, 2007-2009.

Vasiliu, D., Vasiliu, N., Călinoiu C., Mihalescu B., Bontoş M.D., Puhalschi R. 2014. "Using Virtual Laboratories in Concurrent Engineering", ECEC Conference.

***National Instruments Corporation. 2012. "LabVIEW Course", Manu, Texas.

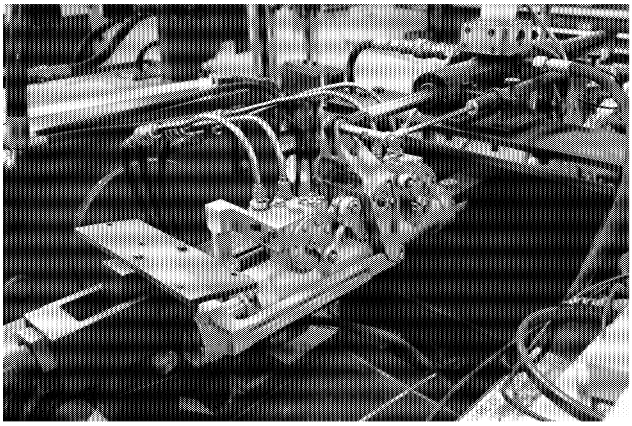
BIOGRAPHIES

MARIUS DANIEL BONTOS graduated in Environmental Engineering from University POLITEHNICA of Bucharest in 2005, and a MD in Environmental Engineering in 2007. After a stage of web designer at Romanian Innovation Agency, he prepare a PhD research work on Using SCADA systems in Environmental Engineering at the Fluid Power Laboratory from the Power Engineering Department of the University POLITEHNICA of Bucharest.

DANIELA VASILIU graduated in mechanical engineering in 1981 and prepared the Ph.D. thesis in the field of the numerical simulation of the hydrostatic transmissions. She is currently professor in the Department of Hydraulics, Hydraulic Machines and Environmental Engineering, head of Fluid Power Laboratory of the University POLITEHNICA of Bucharest with works in the field of modeling, simulation, and experimental identification of the electro hydraulic control systems.

NICOLAE VASILIU graduated in Hydropower Engineering from University POLITEHNICA of Bucharest in 1969. He became a Ph.D. in Fluid Mechanics after a research stages in Ghent State University and Von Karman Institute from Bruxelles. He became state professor in 1994, leading the ENERGY & ENVIRONMENT RESEARCH CENTRE from the University POLITEHNICA of Bucharest. He managed five years the Innovation Romanian Agency. He is member of the Romanian Technical Science Academy, promoting the numerical simulation as engineering tool.

BOGDAN MIHALESCU graduated in Hydropower Engineering from University POLITECHNICA of Bucharest in 2007. He became assistant professor and PhD student in the Fluid Power Laboratory from the same university. His main areas of interest include modeling, simulation and dynamic identification of hydraulic systems.



ANNEX. Front view of the test bench of a section from the automatic pilot of a helicopter

AUTHOR LISTING

AUTHOR LISTING