MATHEMATICS and COMPUTERS in SCIENCES and INDUSTRY

MATHEMATICS and COMPUTERS in SCIENCES and INDUSTRY

Copyright © 2015, by the editors

All the copyright of the present book belongs to the editors. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the editors.

All papers of the present volume were peer reviewed by no less than two independent reviewers. Acceptance was granted when both reviewers' recommendations were positive.

Series: Mathematics and Computers in Science and Engineering Series | 50

ISBN: 978-1-61804-327-6
ISSN: 2227-4588

MATHEMATICS and COMPUTERS in SCIENCES and INDUSTRY

Organizing Committee

Editors:

Prof. Nikos Mastorakis, Visiting Professor University of Salerno, Italy (General Chair)
Prof. Valeri Mladenov, Technical University of Sofia, Bulgaria (Program Chair)
Prof. Imre Rudas, IEEE Felllow, Obuda University, Budapest, Hungary (Special Session Chair)
Prof. Aida Bulucea, Univ. of Craiova, Romania (Finances Chair)
Prof. Branimir Reljin, University of Belgrade, Belgrade, Serbia (International Liaisons)
Prof. George Vachtsevanos, Georgia Institute of Technology, USA (Plenary Speakers Chair)
Prof. Kleanthis Psarris, The City University of New York, USA (Social Part Chair)

Progam Committee:

Prof. Panos M. Pardalos, Univ, of Florida, USA
Prof. Dimitri Bertsekas, IEEE Fellow, MIT, USA
Prof. Ferhan M. Atici, Western Kentucky University, Bowling Green, KY 42101, USA
Prof. Anastassios Venetsanopoulos, IEEE Fellow, University of Toronto, Canada
Prof. Ravi P. Agarwal, Texas A&M University - Kingsville, Kingsville, TX, USA
Prof. Feliz Minhos, Universidade de Evora, Evora, Portugal
Prof. Mihai Mihailescu, University of Craiova, Craiova, Romania
Prof. Aggelos Katsaggelos, IEEE Fellow, Northwestern University, USA
Prof. Alberto Parmeggiani, University of Bologna, Bologna, Italy
Prof. Abraham Bers, IEEE Fellow, MIT, USA
Prof. Lucas Jodar, Universitat Politecnica de Valencia, Valencia, Spain
Prof. Dumitru Baleanu, Cankaya University, Ankara, Turkey
Prof. Martin Bohner, Missouri University of Science and Technology, USA
Prof. Dashan Fan, University of Wisconsin-Milwaukee, Milwaukee, WI, USA
Prof. Luis Castro, University of Aveiro, Aveiro, Portugal
Prof. Kamisetty Rao, IEEE Fellow, Univ. of Texas at Arlington, USA
Prof. Alberto Fiorenza, Universita' di Napoli "Federico II", Napoli (Naples), Italy
Prof. Patricia J. Y. Wong, Nanyang Technological University, Singapore
Prof. Salvatore A. Marano, Universita degli Studi di Catania, Catania, Italy
Prof. Martin Schechter, University of California, Irvine, USA
Prof. Ivan G. Avramidi, New Mexico Tech, Socorro, New Mexico, USA
Prof. Michel Chipot, University of Zurich, Zurich, Switzerland
Prof. Narsingh Deo, IEEE Fellow, ACM Fellow, University of Central Florida, USA
Prof. Xiaodong Yan, University of Connecticut, Connecticut USA
Prof. Ravi P. Agarwal, Texas A&M University - Kingsville, Kingsville, TX, USA
Prof. Yushun Wang, Nanjing Normal university, Nanjing, China
Prof. Detlev Buchholz, Universitaet Goettingen, Goettingen, Germany
Prof. Patricia J. Y. Wong, Nanyang Technological University, Singapore
Prof. Andrei Korobeinikov, Centre de Recerca Matematica, Barcelona, Spain
Prof. Jim Zhu, Western Michigan University, Kalamazoo, MI, USA
Prof. Ferhan M. Atici, Department of Mathematics, Western Kentucky University, USA
Prof. Meirong Zhang, Tsinghua University, Beijing, China
Prof. Lucio Boccardo, Universita degli Studi di Roma "La Sapienza", Roma, Italy
Prof. Shanhe Wu, Longyan University, Longyan, Fujian, China
Prof. Natig M. Atakishiyev, National Autonomous University of Mexico, Mexico
Prof. Jianming Zhan, Hubei University for Nationalities, Enshi, Hubei Province, China
Prof. Narcisa C. Apreutesei, Technical University of Iasi, Iasi, Romania
Prof. Chun-Gang Zhu, Dalian University of Technology, Dalian, China
Prof. Abdelghani Bellouquid, University Cadi Ayyad, Morocco
Prof. Jinde Cao, Southeast University/King Abdulaziz University, China
Prof. Josef Diblik, Brno University of Technology, Brno, Czech Republic
Prof. Jianqing Chen, Fujian Normal University, Fuzhou, Fujian, China

Prof. Naseer Shahzad, King Abdulaziz University, Jeddah, Saudi Arabia
Prof. Sining Zheng, Dalian University of Technology, Dalian, China
Prof. Leszek Gasinski, Uniwersytet Jagiellonski, Krakowie, Poland
Prof. Satit Saejung, Khon Kaen University, Muang District, Khon Kaen, Thailand
Prof. Juan J. Trujillo, Universidad de La Laguna, La Laguna, Tenerife, Spain
Prof. Tiecheng Xia, Department of Mathematics, Shanghai University, China
Prof. Lucas Jodar, Universitat Politecnica de Valencia, Valencia, Spain
Prof. Noemi Wolanski, Universidad de Buenos Aires, Buenos Aires, Argentina
Prof. Zhenya Yan, Chinese Academy of Sciences, Beijing, China
Prof. Juan Carlos Cortes Lopez, Universidad Politecnica de Valencia, Spain
Prof. Wei-Shih Du, National Kaohsiung Normal University, Kaohsiung City, Taiwan
Prof. Kailash C. Patidar, University of the Western Cape, Cape Town, South Africa

Additional Reviewers:

Prof. Abelha Antonio, Universidade do Minho, Portugal
Prof. Alejandro Fuentes-Penna, Universidad Autonoma del Estado de Hidalgo, Mexico
Prof. Ana Maria Tavares Martins, University of Beira Interior, Portugal
Prof. Andrey Dmitriev, Russian Academy of Sciences, Russia
Prof. Angel F. Tenorio, Universidad Pablo de Olavide, Spain
Prof. Athanassios Stavrakoudis, University of Ioannina, Greece
Prof. Audenaert Amaryllis, Universiteit Antwerpen, Belgium
Prof. Bazil Taha Ahmed, Universidad Autonoma de Madrid, Spain
Prof. Bruno Marsigalia, University of Cassino and Southern Lazio, Italy
Prof. Carla Falugi, University of Genova, Italy
Prof. Carlos Gonzalez, University of Castilla-La Mancha, Spain
Prof. Carlos Manuel Travieso-Gonzalez, University of Las Palmas de Gran Canaria, Spain
Prof. Catarina Luisa Camarinhas, Universidade Técnica de Lisboa, Portugal
Prof. Chris Stout, University of Illinois, IL, USA
Prof. Dana Anderson, University of Colorado at Boulder, CO, USA
Prof. Deolinda Rasteiro, Coimbra Institute of Engineering, Portugal
Prof. Dmitrijs Serdjuks, Riga Technical University, Latvia
Prof. Edy Portmann, University of Bern, Switzerland
Prof. Eleazar Jimenez Serrano, Kyushu University, Japan
Prof. F. G. Lupianez, University Complutense, Spain
Prof. Fabio Nappo, University of Cassino and Southern Lazio, Italy
Prof. Francesco Rotondo, Polytechnic of Bari University, Italy
Prof. Francesco Zirilli, Sapienza Universita di Roma, Italy
Prof. Francisco Moya, University of Castilla-La Mancha, Spain
Prof. Frederic Kuznik, National Institute of Applied Sciences, Lyon, France
Prof. Garyfallos Arabatzis, University of Thrace, Greece
Prof. Genqi Xu Tianjin, University, China
Prof. George Barreto Pontificia, Universidad Javeriana, Colombia
Prof. Guido Izuta, Yonezawa Women's College, Japan
Prof. Guoxiang Liu, University of North Dakota, ND, USA
Prof. Heimo Walter, Vienna University of Technology, Austria
Prof. Hessam Ghasemnejad, Kingston University London, UK
Prof. Hirofumi Nagashino, University of Tokushima, Japan
Prof. Hongjun Liu, University of Notre Dame, IN, USA
Prof. Hugo Rodrigues, Universidade Lusofona do Porto, Portugal
Prof. Valeri Mladenov, Technical University of Sofia, Bulgaria
Prof. James Vance, The University of Virginia's College at Wise, VA, USA
Prof. Joao Bastos, Instituto Superior de Engenharia do Porto, Portugal
Prof. John Cater, University of Auckland, New Zealand
Prof. Jon Burley, Michigan State University, MI, USA
Prof. José Carlos Metrôlho, Instituto Politecnico de Castelo Branco, Portugal
Prof. Jose Flores, The University of South Dakota, SD, USA
Prof. Kakuro Amasaka, Aoyama Gakuin University, Japan
Prof. Karel Allegaert, University Hospitals Leuven, Belgium
Prof. Kazuhiko Natori, Toho University, Japan
Prof. Kei Eguchi, Fukuoka Institute of Technology, Japan
Prof. Konstantin Volkov, Kingston University London, UK
Prof. Kun Luo, Zhejiang University, China
Prof. Kyandoghere Kyamakya, University of Klagenfurt, Austria
Prof. Lapo Governi, University of Florence, Italy
Prof. Lesley Farmer, California State University Long Beach, CA, USA
Prof. Luigi Pomante, Università degli Studi dell'Aquila, Italy
Prof. M. Javed Khan, Tuskegee University, AL, USA

Prof. Maling Ebrahimpour, University of South Florida St Petersburg, FL, USA
Prof. Manoj K. Jha, Morgan State University in Baltimore, USA
Prof. Maria Ilaria Lunesu, University of Cagliari, Italy
Prof. Mario Pestarino, University of Genova, Italy
Prof. Masaji Tanaka, Okayama University of Science, Japan
Prof. Mathieu Pétrissans, University of Lorraine, France
Prof. Matteo Nunziati, University of Florence, Italy
Prof. Matteo Palai, University of Florence, Italy
Prof. Matthias Buyle, Artesis Hogeschool Antwerpen, Belgium
Prof. Merzik Kamel, University of New Brunswick, Canada
Prof. Miguel Carriegos, Universidad de Leon, Spain
Prof. Minhui Yan, Shanghai Maritime University, China
Prof. Mokhtari Fouad, University of Quebec at Trois-Rivières, Canada
Prof. Moran Wang, Tsinghua University, China
Prof. Najib Altawell, University of Dundee, UK
Prof. Nicola Simola, University of Cagliari, Italy
Prof. Nikola Vlahovic, University of Zagreb, Croatia
Prof. Ole Christian Boe, Norwegian Military Academy, Norway
Prof. Ottavia Corbi, University of Naples Federico II, Italy
Prof. Pablo Fernandez de Arroyabe, University of Cantabria, Spain
Prof. Pan Agathoklis, University of Victoria, Canada
Prof. Pedro Lorca, University of Oviedo, Spain
Prof. Philippe Dondon, Institut polytechnique de Bordeaux, France
Prof. Philippe Fournier-Viger, University of Moncton, France
Prof. Ricardo Gouveia Rodrigues, University of Beira Interior, Portugal
Prof. Rocco Furferi, University of Florence, Italy
Prof. Rosa Lombardi, University of Cassino and Southern Lazio, Italy
Prof. Santoso Wibowo, CQ University, Australia
Prof. Shinji Osada Gifu, University School of Medicine, Japan
Prof. Sorinel Oprisan College of Charleston, SC, USA
Prof. Stavros Ponis, National Technical University of Athens, Greece
Prof. Sumanth Yenduri, University of Southern Mississippi, MS, USA
Prof. Takuya Yamano, Kanagawa University, Japan
Prof. Tetsuya Shimamura, Saitama University, Japan
Prof. Tetsuya Yoshida, Hokkaido University, Japan
Prof. Thomas Panagopoulos, University of Algarve, Portugal
Prof. Tohru Kawabe, University of Tsukuba, Japan
Prof. Vincenzo Niola, University of Naples Federico II, Italy
Prof. Xiang Bai Huazhong, University of Science and Technology, China
Prof. Xiaoguang Yue, Wuhan University of Technology, China
Prof. Yamagishi Hiromitsu, Ehime University, Japan
Prof. Yary Volpe, University of Florence, Italy
Prof. Yi Liang, Wuhan University, China
Prof. Yuqing Zhou, Wuhan University of Technology, China
Prof. Zhenbi Su, University of Colorado Boulder, CO, USA
Prof. Zhong-Jie, Han Tianjin University, China

Table of Contents

ISBN: 978-1-61804-327-6

ISBN: 978-1-61804-327-6

ISBN: 978-1-61804-327-6

Plenary Lecture

Neurodynamic Optimization Approaches to Parallel Data Selection in the Era of Big Data

Professor Jun Wang
Department of Mechanical & Automation Engineering
The Chinese University of Hong Kong
Shatin, New Territories, Hong Kong
E-mail: jwang@mae.cuhk.edu.hk

Abstract: In the present information era, huge amount of data to be processed daily. In contrast of conventional sequential data processing techniques, parallel data processing approaches can expedite the processes and more efficiently deal with big data. In the last few decades, neural computation emerged as a popular area for parallel and distributed data processing. The data processing applications of neural computation included, but not limited to, data sorting, data selection, data mining, data fusion, and data reconciliation. In this talk, neurodynamic approaches to parallel data processing will be introduced, reviewed, and compared. In particular, my talk will compare several mathematical problem formulations of well-known multiple winners-take-all problem and present several recurrent neural networks with reducing model complexity. Finally, the best one with the simplest model complexity and maximum computational efficiency will be highlighted. Analytical and Monte Carlo simulation results will be shown to demonstrate the computing characteristics and performance of the continuous-time and discrete-time models. The applications to parallel sorting, rank-order filtering, and data retrieval will be also discussed.

Brief Biography of the Speaker: Jun Wang is a Professor and the Director of the Computational Intelligence Laboratory in the Department of Mechanical and Automation Engineering at the Chinese University of Hong Kong. Prior to this position, he held various academic positions at Dalian University of Technology, Case Western Reserve University, and University of North Dakota. He also held various short-term visiting positions at USAF Armstrong Laboratory (1995), RIKEN Brain Science Institute (2001), Universite Catholique de Louvain (2001), Chinese Academy of Sciences (2002), Huazhong University of Science and Technology (2006-2007), and Shanghai Jiao Tong University (2008-2011) as a Changjiang Chair Professor. Since 2011, he is a National Thousand-Talent Chair Professor at Dalian University of Technology on a part-time basis. He received a B.S. degree in electrical engineering and an M.S. degree in systems engineering from Dalian University of Technology, Dalian, China. He received his Ph.D. degree in systems engineering from Case Western Reserve University, Cleveland, Ohio, USA. His current research interests include neural networks and their applications. He published over 170 journal papers, 15 book chapters, 11 edited books, and numerous conference papers in these areas. He is the Editor-in-Chief of the IEEE Transactions on Cybernetics since 2014 and a member of the editorial board of Neural Networks since 2012. He also served as an Associate Editor of the IEEE Transactions on Neural Networks (1999-2009), IEEE Transactions on Cybernetics and its predecessor (2003-2013), and IEEE Transactions on Systems, Man, and Cybernetics - Part C (2002-2005), as a member of the editorial advisory board of International Journal of Neural Systems (2006-2013), as a guest editor of special issues of European Journal of Operational Research (1996), International Journal of Neural Systems (2007), Neurocomputing (2008, 2014), and International Journal of Fuzzy Systems (2010, 2011). He was an organizer of several international conferences such as the General Chair of the 13th International Conference on Neural Information Processing (2006) and the 2008 IEEE World Congress on Computational Intelligence, and a Program Chair of the IEEE International Conference on Systems, Man, and Cybernetics (2012). He has been an IEEE Computational Intelligence Society Distinguished Lecturer (2010-2012, 2014-2016). In addition, he served as President of Asia Pacific Neural Network Assembly (APNNA) in 2006 and many organizations such as IEEE Fellow Committee (2011-2012); IEEE Computational Intelligence Society Awards Committee (2008, 2012, 2014), IEEE

Systems, Man, and Cybernetics Society Board of Directors (2013-2015), He is an IEEE Fellow, IAPR Fellow, and a recipient of an IEEE Transactions on Neural Networks Outstanding Paper Award and APNNA Outstanding Achievement Award in 2011, Natural Science Awards from Shanghai Municipal Government (2009) and Ministry of Education of China (2011), and Neural Networks Pioneer Award from IEEE Computational Intelligence Society (2014), among others.

ISBN: 978-1-61804-327-6

Steganalysis of a pulsed plasma jet ICCD camera image using LabVIEW

Victor J Law and Denis P Dowling
School of Mechanical and Materials Engineering
University College Dublin, Belfield,
Dublin, Ireland
e-mail viclaw66@gmail.com

***Abstract*—A LabVIEW computer program is presented as a steganographic tool to analyze and manipulate, intensified charge-coupled device images of atmospheric pressure plasma jet. The program deselects the red and green color planes and only uses the blue color plane that holds the imperceptible (to the human eye) fluid structure information beyond the visible distal-point of the plasma plume. Low pass pixel filtering of the spatial image followed by Fast Fourier transformation of the image into the complex image is used to access and truncate bit depth information. An inverse Fast Fourier Transformation is then used to convert the complex image back in to the spatial plane where image thresholding after applied to form a binary image of the fluid structure beyond the visible distal-point. Both conceptual and technical programing issues are discussed.**

Keywords—Plasma; fuild structure; image; spatial-domain; frequency-domain; thresholding

I. INTRODUCTION

Today's atmospheric pressure plasmas jet (APPJs) applications range from polymer and composite surface cleaning to medical and dental treatments. To understand and control these plasma applications modern plasma diagnostics embrace an array of electrical, thermal, acoustic, optical and imaging sensors. Associated to these sensors are analytical software tools that are designed to explore the specific signal of the sensor. A specific aspect of interest in these APPJs is the local turbulent flows or instabilities that are aligned with the central gas flow and decaying heat path of the plasma jet. From the early 2000s studies of thermal flame and plasma jets instabilities have used Schlieren imaging [1, 2]; and more recently high-speed Schlieren imaging [3] and direct capture by intensified charge-coupled device (ICCD) camera followed by manipulation of the complete spatial-domain information (red green blue (RGB) colour planes) [4] have been employed. In the latter case, RGB ICCD images have proved to be very successful in identifying and codifying the linear propagation velocities of built-like structures [5], and the non-linear propagation of Snake-like [6] and Kink and Wrinkles structures within a pulsed plasma jet [7, 8]. In this body of work, the majority of the luminous spatial and temporal plasma jet information is contained within the red and green planes while the blue plane that relates to fluid structure beyond the visible distal-point is imperceptible to the human eye. Accessing the hidden information within the blue plane has been recently reported [7, 8]. In these two bodies of work a purpose built National Instruments LabVIEW software program used to deliberately deselect the red and green colour planes and only manipulate the spatial-domain and frequency-domain information of the blue colour plane. The deselection of the base red and green colour planes may be considered to be the digital steganographic (from ancient Greek for 'covered writing') equivalent of the role played by the knife blade edge within the Schlieren imaging process.

Due to the diagrammatic nature of LabVIEW programming it is the aim of this paper to present the design concepts, including panel and block-diagrams of the LabVIEW programs used in [7, 8]. LabVIEW program was chosen for this application because the basic programs have been established in the field of high temperature plasma physics since the 1990s [9] and with continuing software updates [10]. LabVIEW 2011 incorporating vision and motion image acquisition (IMAQ) virtual instruments (VIs) are used in this work and are deployed within a Windows 7 environment. Information on the use of IMAQ VIs within a LabVIEW 7.1 for plasma spraying can be found in reference [11]. It is noted. that similar software, such as MATLAB or Java may be used for this purpose. As regards to the computer used, all LabVIEW programs are installed in a Dell Inspiron laptop computer.

II. PLASMA JET

The plasma jet studied in this work is the kINPen med® it is operated with a square wave pulsed frequency of 2.5 kHz and uses argon as the carrier gas at a flow rate of 5 standard litres per minute. The plasma physics of the plasma jet itself is not discussed here, but can be found in references [1 to 8].

III. ICCD CAMERA

The Andor iStar 334T ICCD camera is used to capture images of the pulse-on period (200μs) of the plasma plume. A 14 cm focal length glass lens focused the region from between 2 mm upstream to 30 mm downstream of the exit nozzle. Using this combination the overall optical chain (between camera and plasma-plume) is of the order of 2 m and the camera spectral range is restricted to 300 to 850 nm by the glass lens. The camera is triggered, via a delay generator, from the rising edge of the photo-diode signal. Within the camera the images are

ISBN: 978-1-61804-327-6

processed using a false-colour scale from blue (low intensity) to yellow (high intensity). For maximum visual differentiation the gain was set to 2817 out of a maximum of 4095, where the final digital images are formatted as a 24-bit depth (8-bit red; 8-bit green; 8-bit blue) Joint Photographic Experts Group (JPEG) image. The colour image is formed as a 2-diamensional array of *N x N* pixels, where *N* = 1024. The RGB image is formed from three overlaying colour filters where the value of each colour pixel location (x, y) has an 8-bit (0 to 256) intensity level value of R(x, y), G(x, y) and B(x, y), respectively. The intensity valve (Z) of each pixel forms the frequency domain location (i.e. $Z_R(x, y)$, $Z_G(x, y)$ and $Z_B(x, y)$). The final value used depends upon the final colour and intensity at each pixel within the spatial-domain image array. Fig. 1 shows the orthogonal relationship of the spatial-domain pixel array and the 8-bit depth frequency-domain for one of the three base colours. In this figure, bit information is placed abruptly in the depth axis, only as an example.

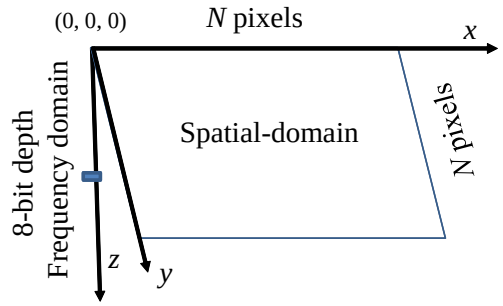


Fig. 1. Orthogonal relationship between the spatial-domain (*N X N* array) and frequency-domain bit depth information.

IV. SOFTWARE DESIGN CONCEPT

The basic design of the processing and analysis of the 24-bit JPEG image processing software follows two distinct data flow pathways. The first pathway has six processing steps. These are: steps 1 and 2 create and read the image; steps 3 and 4 are a user defined line intensity profile that compares the pixel intensities along a chosen line in the *N x N* array of the image. The function of this step in conjunction with step 4 (RGB colour plane separation) allows the operator to visually evaluate (step 5) which colour plane is to be transformed into the frequency-domain; finally step 6 is the automatic save and archive of the line profiles data for later analysis.

This second data flow pathway begins at step 4 of the first data flow pathway. The function of this pathway is to perform conversion between the spatial-domain and frequency-domain and associated low pass filtering and finally thresholding in to the final binary image. This second pathway has six process steps and three temporary image displays that allow the user to make decisions while manipulating the data flow. Step 1 is a user defined colour plane selection; step 2 is an automatic conversion to grey scale for preconditioning for the next step 3; Step 3 is the Fast Fourier Transform (FFT) that generates a complex image in which high frequencies are grouped at the centre, while low frequencies are located at the edges; step 4 is the low pass-filter stage truncation stage; step 5 converts the filtered complex image back in to the spatial-domain using an inverse FFT, and; step 6 is the local threshold algorithm that converts the spatial image into a binary (black and white) image [12]. Finally, the spatial image save option is proved for historical analysis. Shown in Fig. 2 is a simple schematic of the steps of the dual parallel data flow pathways (solid lines) including user-interface (dotted line). The data flow moves from top to bottom.

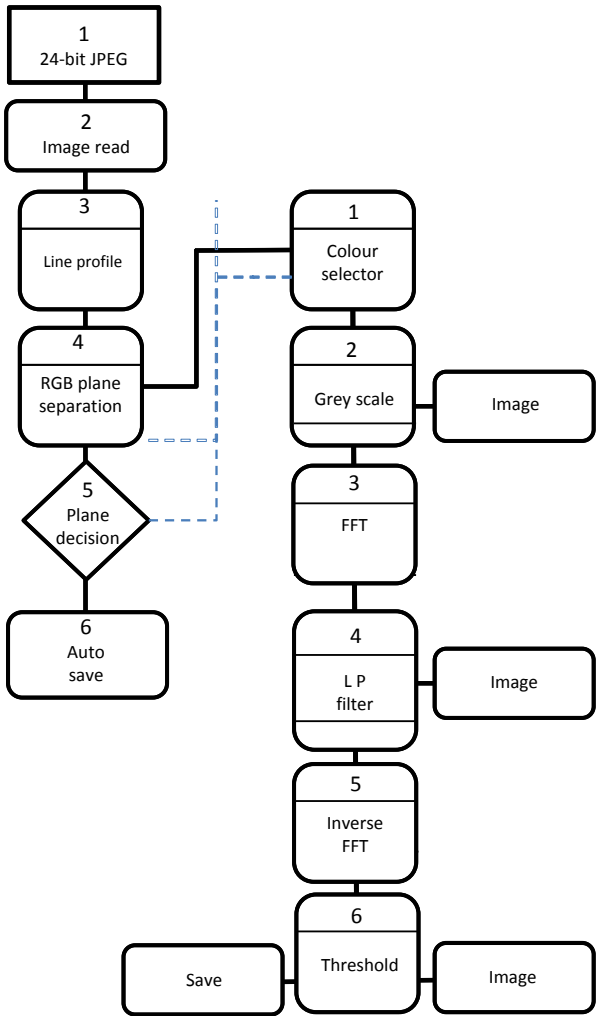


Fig. 2. Schematic of data design showing initial read and dual parallel data flow pathways (solid lines) and user decision making (dashed lines).

V. LABVIEW PROGRAMING

The LabVIEW software packages form a complete programing language based entirely on graphical user interface. The sub-program VIs create graphical constructs equivalent to a while-loop, case-structures and IMAQs and other VIs). These components are placed on a panel (block diagram) which is ultimately hidden from the operator. The data flow through the VIs is controlled by connecting 'wires' between objects, somewhat like constructing an electrical circuit. The graphical user interface (front panel) is created from a flexible set of

ISBN: 978-1-61804-327-6

predefined graphical displays and controls. The image read and display IMAQ VIs are located outside of the while-loop so when the program is operated the image opens and stays open to allow interactive processing of the image within the while-loop. The two data flow pathways are placed within the while-loop. Fig. 3 depicts the complete LabVIEW block diagram. Note it is rotated through 90 degrees and enlarged over 1 column for clarity. The data flow therefore reads from the bottom of the page to the top of the page.

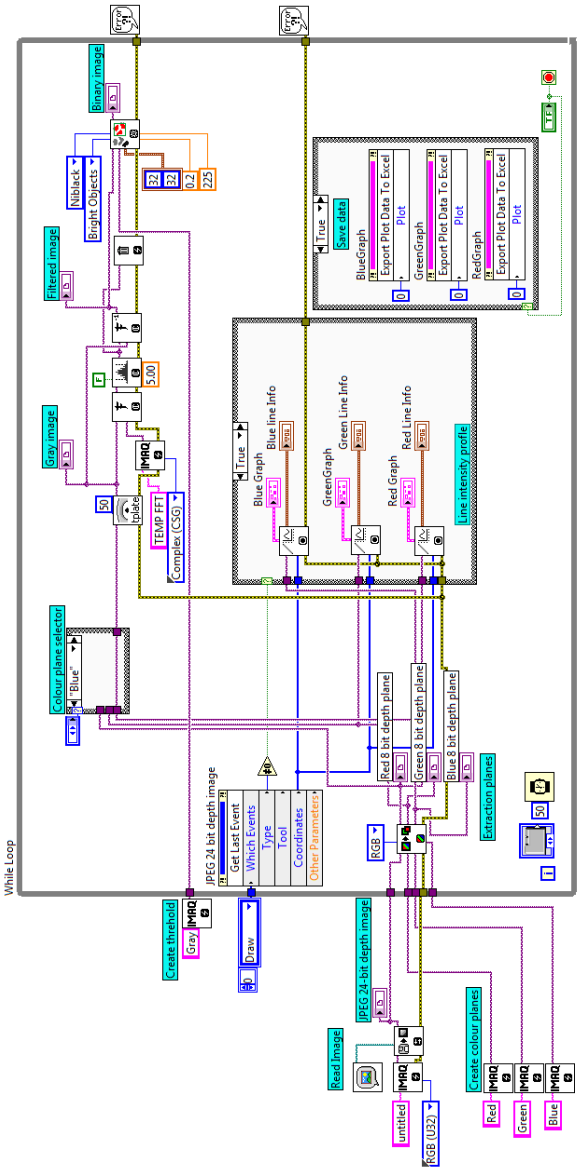


Fig. 3. LabVIEW block diagram.

A. *Data flow path 1*

This section details the construction of data flow pathway 1 within the while-loop, and reference to VIs outside the while-loop. The pathway starts with an IMAQ RGB colour extraction plane VI that has three supporting image displays, each being referenced to their respectively RGB VIs outside the while-loop to maintain the viewing of the original image. For the creation of a user defined line profiles a case-structure (true) containing three line graph IMAQ VIs are used with reference via an event invoke node to a draw command outside the while-loop. Finally three invoke nodes within a separate case-structure (true) is used to automatically save the digital line profile information on closing of the program.

B. *Data flow path 2*

This section details the construction of the data flow pathway 2 within the while-loop and the single reference VI outside the while-loop. The pathway starts by sampling the information generated at the colour extraction VI in pathway 1. From here, a case structure containing the red, green and blue planes is used to select the base colour. The user control dialog box is set to default on blue plane.

The next processing sequence requires spatial-domain gray scale and boarder size control and pixel averaging to clean the noise within the base colour image. This IMAQ VI is not held in the vision VI library, but rather in an external library. Therefore Vision Assistant software is used to both gray scale and perform the pixel average process prior to the conversion into the temporary complex image for frequency filtering, truncation and conversion back into the spatial domain using an inverse FFT sequence. Once this code section was completed the code was inserted into pathway 2, from where the IMAQ NIblack threshold VI is used to make the final binary image.

VI. DISCUSSION

This section presents examples of the software in use. Example ‘A’ shows the original plasma jet image. Example ‘B’ shows the colour extraction and cross- sectional line profile of the plasma plume. Finally, ‘C’ shows pixel averaging and binary image of the frequency-domain truncation.

A. *Origanl plasma jet image*

The original ICCD camera image (JEPEG 24 bit depth) of the plasma jet is shown in Fig. 4. In this image a user yellow cursor (x, y coordinates: 840, 0 to 840, 1025) bisects the plasma plume at 90 degrees. The line thus defines the cross-sectional intensity profile in each base colour plane.

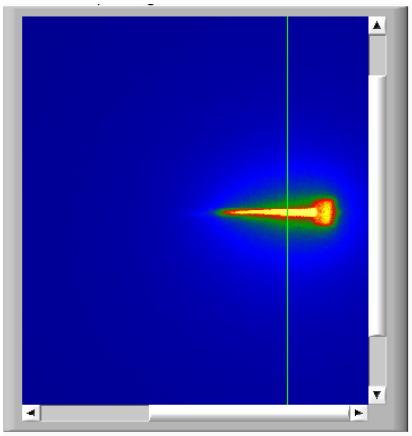


Fig. 4. Original jet image with cursor line.

ISBN: 978-1-61804-327-6

Consider now that the colour changes along the cursor line: starting from the plume intersection region to the outer region blue regions. Here we note white-yellow represent the hottest part of the plasma jet which rapidly cools through red and green, then through a faint light-blue halo and beyond into imperceptible changes (to the human eye) in the blue background.

B. Color extraction and cross-sectional line profile

In this section three base colour planes (red green and blue) of the original JPEG image are separated and displayed along with the cross-sectional profiles of the plasma plume according to the yellow cursor line setting that was established in section V. *A*. Fig. 5 shows all three colour planes, alongside each profile is the cross-sectional line profile. The horizontal axis is in units of pixels of the spatial-domain and the vertical axis is calibrated in pixels: 0 to 256 pixels, or 8-bits, and represents the single colour plane image bit depth information.

Inspection of the red colour plane and its associated plume cross-sectional intensity profile reveals that the plasma plume information is concentrated within 50 pixels along the line of the central gas flow axis and around jet nozzle at bit depth value almost at 256 pixel level.

The green colour plane reveal a similar spatial and bit depth information, but with the addition of halo around the central gas flow axis.

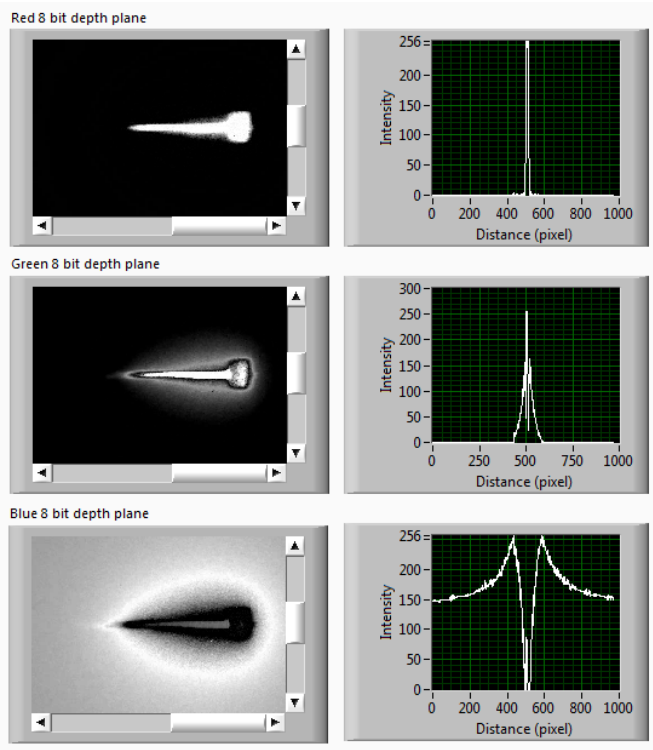


Fig. 5. LabVIEW front panel of color plane extraction and their line profiles.

In contrast to the red and green plane the blue plane contains little pixel information in the region of the central plasma plume and jet nozzle. However, graded pixel line information up to 512 pixels either side of the plasma plume is revealed. This graded line information occurs in the 150 to 200 pixel depth the range and contains packets of noise along the two graded lines. That is the imperceptible information, is revealed to the human eye. Note the characteristic pear-shape information is due to the hot jet nozzle.

C. Pixel averaging, truncation and bianry image

We now consider the spatial-domain low pass average pixel filtering. The filter works by calculating the inter-pixel variation between the pixel under consideration and its neighbouring pixels. For a single pixel this means that 8 neighbouring pixels are compared. When the pixel under consideration has a variation greater than the user defined percentage (here we use 50%) it is set to the average value of neighbouring pixels. Fig. 6 schematically illustrates this process within a 10 x 10 pixel frame.

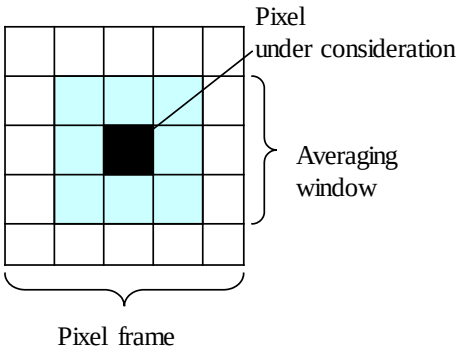


Fig. 6. Spatial-domain pixel average process.

For comparative purposes, Fig. 7 compares the raw pixel line information with averaged line information along the cursor cross-sectional line as defined in Fig. 4 and Fig. 5.

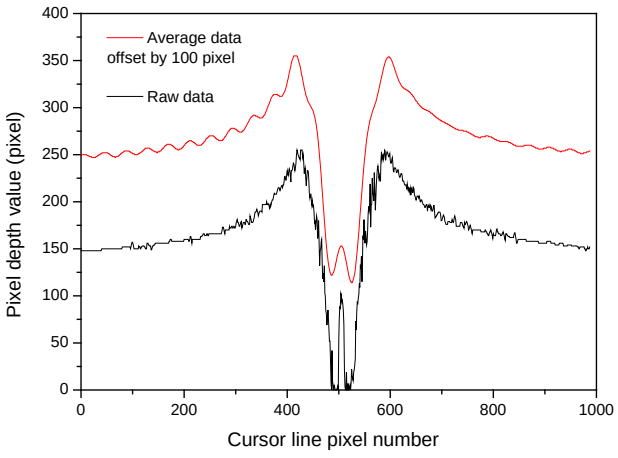


Fig. 7. Raw and averaged pixel data along cursor cross-sectional line.

It can be seen that the line noise in the raw data has been processed into a new line profile that yields a smoothed line containing a repeating periodic profile either side of the central plasma plume; where the bandwidth of the plume and

ISBN: 978-1-61804-327-6

repeating structure is of the order of 40 pixels. Additionally the depth of the periodicity is of the order of 10 pixels. This process is repeated of each pixel in the spatial-domain that starts at the initial x, y pixel coordinates of (0, 0) to yield an image that is ready for truncation (section V. D).

D. Trucation and bianry image

Fig. 8A shows the spatial-domain image of the low pass averaging filtering processes. The image reveals three groups of light gray colour striations against a more uniform and darker gray background: two groups either side of the plasma plume and one group to the front of the plume.

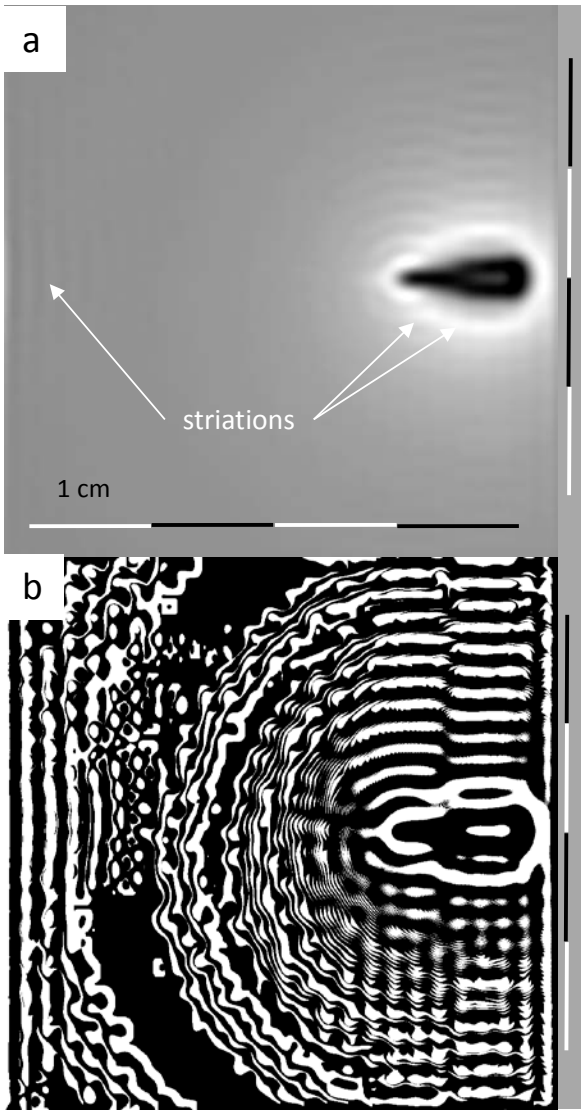


Fig. 8ab. LabVIEW front panel of the spatial averaged filtered image (a) and biany image (b). Scale markers = 1 cm.

An inspection of the parallel striations reveals the same periodic nature as observed and measured in fig 7. Note, also the parallel and frontal striations have similar 40 pixel periodicities. Using the same coordinates as in fig. 4 and fig. 5, A cursor cross-sectional profile of these striations reveals the periodic nature of these striations with only a fine trace of the central plasma plume (40 pixels wide at its broadest). Typically, the parallel and frontal striations have a peak separation of the orders of 65 pixels with a pixel depth of 10 pixels.

A FFT is then applied to the image in Fig 8a to convert the spatial information into the frequency-domain. To smooth the noise a low-pass truncation process is used to remove any remaining high frequency component above a user defined cut-off point. An inverse FFT of the truncated information regains the spatial-domain image.

Fig 8.b shows the results of the local Nibalck thresholding segmentation algorithm when applied to Fig 8a. The operation produces a binary image where the background pixels are set to I = 0 (black) while setting the fluid structure pixel value to I = 1 (white). The result of this process produces a black-and-white binary image that represents the fluid structure within the original blue image.

In fig.8b the three sets of repeating striations (ripples) is again observed to radiate from plasma plume: one set from along the axis of the plasma plume and two sets extending from the point of plume in the direction of effluent flow. The 1 cm scale markers show that these ripples extend up to 4 cm from the nozzle with a local asymmetric complex structure between these ripples structures. This disturbance occurs at around 0.5 cm from the plume distal point. An important feature of notes here is that this break in the ripple continuity is revealed in this thresholding step and not and therefore hidden in the pixel averaging step. Lastly all the ripples, apart from the local disturbance, appear to have a periodic structure (between each white peak) typically of 1 to 2 mm.

VII. CONCULSIONS

In this work, LabVIEW 2011 based software is used as a mean of steganalysis of ICCD camera image of the pulsed-on period (200μs) of the kINPen med® atmospheric plasma jet. To reveal the hidden pixel information within beyond the visible plasma plume distal-point, RGB colour plane separation have been performed with the intention to investigate and reveal the imperceptible changes (to the human eye) in the blue colour plane.

The use of spatial-domain pixel averaging and frequency-domain pixel filtering has revealed hidden pixel information 2 to 3 cm beyond the visible distal-point of the plasma plume i.e. in the cool temperature reigns of the plasma effluent. This far-field fluid structure information may be used to support Schlieren imaging [1-3] investigations, and, or, in the understanding of what has been termed ‘spillover’ [13] and other surface interactions effects when plasma treating thermally sensitive polymers and their biomaterial counter parts. An important point that must be made here this the software is post treatment of ICCD and thus can be performed at any time on historical plasma images and thus extends experiments that did not have the availability of expensive Schlieren image equipment at that original time.

ISBN: 978-1-61804-327-6

Flexibility in the colour plane choice (red, green or blue), not only allows both thermal and non-thermal plasma jets to be analysed, but also may extend the software use to other fields of image analysis.

ACKNOWLEDGMENT

This work is in part supported by the Irish Centre for Composites Research (IComp).

REFERENCES

[1] T. E. Carlsson, R Mattsson, P Gren, M. Elfsberg and J. Tegner. Combination of Schlieren and pulsed TV holography in the study of a high-speed flame jet. Opt, Lasers Eng. 44(6), 2006, pp. 535-554.

[2] Y. Sutton, J. Moore, D. Sharp and N. St. Braithwaite. Looking into a plasma loudspeaker. IEEE Trans, Plasma Sci. 39(11), 2011, pp. 2146-21487.

[3] M. Boselli, V. Colombo, E. Ghedini, M. Gherardi, R. Laurita, A. Ligouri, P. Sanibondi, and A. Stancampiano. Schlieren high-speed imaging of a nanosecond pulsed atmospheric pressure Non-equilibrium plasma jet. Plasma Chem, Plasma Process. 34 (2014), pp853-869.

[4] V. Něničks, J. Hilina and J. Sonský. Identification of dynamic patterns in CCD records of thermal plasma jets. Progress in plasma processing of materials, Begell House, Inc, New York, 2003, pp. 111-116.

[5] J. L. Walsh, F. Iza, N. B. Janson, V. J. Law and M. G. Kong. Three distinct modes in a cold atmospheric pressure plasma jet. J, Phys, D: Appl, Phys 43(7), 2010, pp.75201 (14pp).

[6] S. Wu. Z. Wang, X. Tan, X. Lu and K. Ostrkov. Atmospheric-pressure plasma jets: effect of gas species, and snake-like built propagation. Physics of Plasmas. 20, 2013, pp. 203503.

[7] V. J. Law, A. Chebbi, F. T. O'Neill and D. P. Dowling. Resonances and patterns within the atmospheric pressure plasma jet kINPen-med. CMSIM Journal. 1, 2014, pp. 3-10.

[8] V. J. Law. L. J. Cox. W. Adress, W. G. Graham and D. P. Dowling. Spatial and temporal imaging of a plasma jet plume. CMSIM Journal. 3, 2012, pp. 61-270.

[9] L. Mandrake and W. Gekelman Large scale Windows 95-based data-acquisition system using LabVIEW. Computers in Physics. 11(5). 1997, pp. 498-507.

[10] P. Apte. Using CompactRIO and LabVIEW to monitor and control a compact spherical field Tokamak for plasma Research. National Instruments case study 15630. http://sine.ni.com/cs/app/doc/p/id/cs-15630. Last accessed February 2015.

[11] Z. Zhao, G. Zhou, B. Xiao, and S. Zhang. Plasma jet acquisition and image processing technology. Materials Science Forum. 704-705. Trans Tech Publications, Switzerland, 2012, pp. 1259-1265.

[12] V. J. Law, F. T. O'Neill and D. P. Dowling. Automatic computation of crossing points within orthogonal interpolation line graphs. How Nature Works: Emergence, Complexity Computation 5. chap 9. Springer International Publishing, Switzerland, 2014, pp. 195-216.

[13] J. F. Kolb, A. M. Mattson, C. M. Edelblute, H. Xiaolong, M. A. Malik, and L.C. Heller. Cold dc-operated air plasma jet for the inactivation of infectious microorganisms. IEEE Trans, Plasma Sci. 40(11), 2012, pp. 3007-3026.

Industrial Uses for Authorship Analysis

Patrick Juola
Evaluating Variations in Language Laboratory
Duquesne University
600 Forbes Avenue
Pittsburgh, PA 15282 USA
Email: juola@mathcs.duq.edu
Tel: +1 (412) 396-2276

***Abstract*—Text classification is an important technology to help industry handle millions of words of customer communications. We discuss a less well-known application of text technology, specifically authorship attribution and profiling. By examining not the content but the style of these communications, computers can learn not only what people are writing about, but things about the people writing as well, such as their identity, demographics, and even psychometrics. We provide several applications to illustrate the value of this important emerging technology.**

I. Introduction

Sometimes knowing *who* wrote a document is as important as knowing *what* was written. Sometimes you need to know who your critics are, even when they hide behind anonymous sounding names on Internet forums. Sometimes you need to know who actually sent a letter, a piece of email, or a text message. While the idea of using computers to analyze the contents of a document is well known, the idea of using one to analyze the author is perhaps less well-understood. In this paper, we provide examples both of how this kind of analysis can be done, and more importantly, of why this capacity is important.

II. Background

A. Text Categorization

More text information is available now than ever before in history, and no human can possibly read it all. At the same time, through channels like Internet forums and product commentary, customers have more ability to influence each other than ever before, and suppliers need to be able to read the same documents. But if not humans, then... computers? By reading documents at the speed of electronics, emerging technologies make it possible to keep up with the flood of text.

Two key examples of this technology are topic modeling and sentiment analysis. Topic modeling [1] infers "the main themes that pervade a large and otherwise unstructured collection of documents," in order to "organize the collection according to the discovered themes." Sentiment analysis [2] categorizes documents by the emotions and/or opinions expressed in it, usually based on a positive/negative "polarity," describing whether the author likes/approves of the topic. For example, "awesome" is generally a positive term, as is (less obviously)"zest." "Abdicate" is generally negative. Using these technologies together, a computer can analyze thousands or millions of comments about a product and tell the company both *what* aspects of the product people are discussing as well as *how* they feel about the various aspects.

What this technology will not tell you is *who* is writing these comments. At one level, this is simply useful marketing information; if everyone who likes your product are college-age females, this suggests both new markets that might be opened as well as efficient channels to build the existing market. More subtly, it's hard to tell how many actual people are behind the comments and whether the criticisms are a genuine issue or simply one hard-to-satisfy customer with too much free time. Indeed, review spam [3], [4] where praise and criticisms are written and published on a commercial scale, is a problem of increasing urgency. Another application of text categorization, stylometry (also called stylometrics) provides a potential solution to these problems. By examining the individual style of the individual writers, the computer can tell you not only about the contents of the documents, but also things about the person who wrote it.

B. Theory of Stylometry

So how does this work? The basic theory of traditional stylistics is fairly simple. As McMenamin describes it,

> At any given moment, a writer picks and chooses just those elements of language that will best communicate what he/she wants to say. The writer's "choice" of available alternate forms is often determined by external conditions and then becomes the unconscious result of habitually using one form instead of another. Individuality in writing style results from a given writer's own unique set of habitual linguistic choices.[5]

Coulthard's description is similar:

> The underlying linguistic theory is that all speaker/writers of a given language have their own personal form of that language, technically labeled an idiolect. A speaker/writers idiolect will manifest itself in distinctive and cumulatively unique rule-governed choices for encoding meaning linguistically in the written and spoken communications they produce. For example, in the case of vocabulary, every speaker/writer has a very large learned and stored set of words built up over many years. Such sets may differ slightly or considerably from the word sets that all other speaker/writers have similarly

ISBN: 978-1-61804-327-6

> built up, in terms both of stored individual items in their passive vocabulary and, more importantly, in terms of their preferences for selecting and then combining these individual items in the production of texts. [6]

A non-obvious but key application is to the legal system. For example, a famous dispute over the ownership of a significant part of Facebook (*Ceglia v. Zuckerberg and Facebook*) depended in part upon a set of disputed writings. These writings were email, allegedly written by Mark Zuckerberg, that purported to show that Paul Ceglia owned half of Facebook. Of course, if these writings were not by Zuckerberg, they showed nothing of the sort. McMenamin's report analyzed eleven different and distinct "features" of the writing in both the known (undisputed) email and the disputed email. One feature, for example, hinged on the spelling of the word *cannot*, and in particular whether it was written as one word (*cannot*) or as two (*can not*). Another feature was the use of the single word "sorry" as a sentence opener (as opposed, for example, to "I'm sorry"). [5] submitted a report that showed that the writing style of a set of undisputed email (that Zuckerberg acknowledged having written) differed in a number of important ways from the disputed writings, and concluded that "[i]t is probable that Mr. Zuckerberg is not the author of the QUESTIONED writings." (Capitalization in original.)

Similarly, [6] describes a (redacted) case of authorship of a disputed email leaked from a company under questionable circumstances. Coulthard similarly discussed (among other features) the use of the specific phrase "disgruntled employees." [7] describes a case of potential murder, where the authorship of a set of SMS (text) messages found on a cell phone constituted a key element in establishing both the time of death (when the writing style of these messages shifted radically) and showed strong indications of an attempt to cover up the murder via arson. By examining features including variant spellings such as "wiv" for "with" and "wud" for "would," he was able to show key differences between the writing of the messages and the typical writing of the phone's owner. He was also able to show key similarities between the writings of the (alleged) murderer/arsonist and one of the suspects in the case.

In other examples, [8] describes a case in immigration court, where an applicant for political asylum was able to lay claim to a number of anonymous newspaper columns critical of his home government, and therefore establish a reasonable fear of persecution upon return to his homeland. [9] describes another murder case, one where the crime scene included a suicide note typed on a shared computer, but stylistic analysis was able not only to show that it had not been written by the decedent, but also to identify someone else as the killer.

Computer-based stylometry applies the same general theory, but with a few major differences. The basic assumption that people make individual choices about language still holds, but instead of *ad hoc* features selected by examination of the specific documents, the analysts use more general feature sets that apply across the spectrum of problems. One common feature set is the frequency of common words such as articles and prepositions [10], [11], [12]. Because these words tend both to be common and also not to carry strong semantic associations, their frequencies tend to be stable across documents and genres, but these frequencies can also be shown to vary strongly across individuals. Another commonly used feature set is the frequency of common groups of consecutive words (word n-grams) or consecutive characters (character n-grams) [13], [14], [15]. Using these feature sets or others [16], the features present in a document are automatically identified, gathered into collections of feature representations (such as vector spaces), and then classified using ordinary machine learning algorithms [17], [18], [19] to establish the most likely author.

A particularly good example is Binongo's study of the *Oz* books [11]. The backstory is fairly simple: the series was started with L. Frank Baum's publication of *The Wonderful Wizard of Oz* and continued until his death in 1919. After his death, the publishers asked Ruth Plumly Thompson to finish "notes and a fragmentary draft" of what would become *The Royal Book of Oz*, the 15th in the series, and then Thompson herself continued the series until 1939, writing nearly twenty more books. The underlying question is the degree to which this "fragmentary draft" influenced Thompson's writing; indeed, scholars have no evidence that the draft ever existed. Binongo collected frequency statistics on the fifty most frequent function words across the undisputed samples and analyzed them using principal component analysis (PCA). Reducing these fifty variables down to their first two principal components produced an easily graphable distribution that showed clear visual separation between the two authors. When the *Royal Book* was plotted on the same scale, it was shown clearly to lie on Thompson's side of the graph, confirming that "from a statistical standpoint, [the *Royal Book*] is more likely to have been written in Thompson's hand."

III. Applications

A. Attribution

In 1996, the novel *Primary Colors* was published. A roman-à-cléf purporting to describe Clinton's 1992 presidential campaign, it provided an insightful view into late 20th century American politics. Or did it? If the anonymous author actually had inside knowledge, that was one thing. On the other hand, if it was just a potboiler by an ordinary novelist, it may no more accurately have reflected reality than a Spider-man comic book describes life in contemporary New York City. As part of the discussion surrounding this book, linguist Don Foster [20] showed that the writing was very similar to that of columnist Joe Klein, who later acknowledged authorship. Another recent high-profile example [21], [22] is that of the author of *A Cuckoo's Calling*, by Robert Galbraith. Although Galbraith was a first-time author, numerous critics noted that the authorial voice was unusually polished and confident. Formal analyses of writing style, performed at the behest of the *Sunday Times*, later identified [23] J.K. Rowling, author of the *Harry Potter* books, as the actual author. Literature scholars have been interested in questions of authorship for centuries, as typified by the discussions of authorship of Biblical book of Acts [24], traditionally ascribed to the author of the book of Luke, and of the authorship of the *Illiad* and the *Odyssey*, still an open question [25]. However, identifying the author of a document can be of interest to other parties as well.

ISBN: 978-1-61804-327-6

Another common application is journalism. As with the Rowling case [21], many questions arise from a matter of public interest, driven by journalists. Another recent example is *Newsweek's* analysis of the Bitcoin design documents, officially written by a person named "Satoshi Nakamoto" (which may have been a pseudonym), and attributed by Newsweek to a retired engineer named Dorian Nakamoto. Stylometric analysis of these documents [26] against an appropriate set of known documents showed "that Dorian Nakamoto was not found to be a plausible candidate author, and in fact, one of the distractor authors (Neal J. King) was found to be a better match to Satoshi Nakamoto than any other distractor or than Dorian." [27]

B. Profiling

A related problem is that of authorship profiling [28], the study of other authorial characteristics such as gender, age, education level, native language, personality and so forth. [29], [30], [31], [32], [33], [34], [35], [36], [37], [38], [39] Profiling is in some ways a more important problem than attribution. Profiling can be and is used [28] on a larger scale to infer group properties of a large number of people.

Profiling is done in the same way as attribution, but instead of offering training documents labeled by author, the system is provided with documents representing specific groups, such as essays written by college graduates and by non-college graduates, or by speakers of UK and US English. The same feature selection and classification techniques will infer the appropriate markers for group membership and classify novel documents accordingly. (To illustrate, an obvious feature for distinguishing UK vs. US dialects would be vocabulary, and specifically items like "lorry," "ironmonger," and "tarmac.")

Authorship profiling has obvious commercial potential (what can I learn about the people who post negative reviews of my product?) but is also of significant interest to other fields, such as law enforcement. Among other applications, it forms one of the technologies underlying DARPA's Active Authentication project [40], [15], [41], based on the theory that if I write (or more generally, interact with the computer) like an introvert, but the person actually at the keyboard behaves as an extrovert, then that person is probably not me. In the event of an actual security incident, learning *about* the intruder can provide a useful start for investigation and response. Other applications may include telemedicine, for example, allowing the nonintrusive identification and assessment of risk factors such as bipolar disorder [42], low self-esteem [43], depression [44] or suicidality [45].

The methodological basis of these analyses are very similar to the authorial analysis, and the same software can be used for both applications [32], [46]. Indeed, in many cases [15] very similar feature spaces and classification methods are among the best-performing; the only difference is in the labeling of the training corpus. There are several proofs-of-concept in this space [32], [47], [38], [43], [42], illustrating that it is quite practical to do this kind of profiling for a number of different attributes, including both normal [48], [29], [47], [38], [43] and pathological [44], [45], [42] psychological traits as well as ordinary demographic information [32], and even handedness [15]. This technology has even been used to infer deceptive intentions [49], [50], [51].

IV. Three novel applications

A. Case 1: Identifying commercial sock puppets

We have identified in the previous paragraphs all of the necessary ingredients to begin addressing a key problem of the modern commercial world, that of commercial deceptive social media. Deceptive customer reviews—whether paid-for-positive reviews by shills, or damning reviews placed by agents of the competition—are becoming a major issue in e-commerce and a major problem for businesses whose primary product is review aggregation. Deceptive review spam is used as a marketing tool by corporations [3], [52], political pressure groups [53], and even national governments [4], but can also simply be the acts of a single active person with an axe to grind.

As discussed in the previous section, this deceptive intention can be detected, as can posts by the same author using multiple identifiers and user names. This provides a relatively simple way to allow an analyst to disregard multiple postings or deceptive postings. In fact, it would even be straightforward for an aggregator such as Yelp to eliminate these from consideration in offering "average" customer ratings, or for law enforcement such as the Federal Trade Commission to initiate proceedings as appropriate. By identifying overrepresented (or outright deceptive) comments, this enables the merchant to develop a more representative picture of the customer base and take actions grounded in a better and more realistic understanding of the true situation.

B. Case 2: Identity as a behavioral biometric

Passwords are generally considered to provide weak security. [40] They can be forgotten, guessed, or stolen. More subtly, passwords only provide momentary security up-front, at login time. When the user gets up to get coffee, the computer retains the user's credentials and will continue to provide access to anyone who sits down at the keyboard. Chaski [9] provides an example of a legal dispute hinging on who actually sent inappropriate email from a (shared) corporate computer, but a more common problem might involve insider threats, where someone uses someone else's leftover credentials to access beyond his/her authorization.

In 2012, DARPA [40] proposed to develop "Active Authentication," an alternative approach to computer security where users are continuously and actively reassessed on an ongoing basis to determine whether or not they are still authorized to use the computer. In the event that a user does something that causes the system to question their identity, a security alert can be raised (and appropriate action taken).

One of the technologies assessed for this project is authorship attribution and profiling as a form of linguistic biometric [15], [41]. In simple terms, if the person writing email is not writing the way the authorized user would write, then the person writing may not be the authorized user. Similarly, if a person is drafting a document in the wrong writing style, there may be an issue. Even profiling can be applied to this issue — if the person writing the document writes like a member of the wrong group, a group to which the authorized user does not belong (and which the person normally does not write like), there may again be an issue.

ISBN: 978-1-61804-327-6

Juola et al. [15] have shown this to be feasible. This group collected information about writing style from a group of 80 participants in a simulated work environment. Each person, over a one-week period, was asked to do a long-term blogging task intermixed with smaller, more explicitly-defined writing tasks of a few hours each. Using ordinary stylometric technology, they were able to identify specific participants by their writing style with roughly 60% accuracy based on as few as 500 characters, and to identify personality categories such as introversion/extraversion with approximately 80% accuracy. (By contrast, the chance baseline for identifying specific people is approximately 1%, one in 80, and for identifying personality traits approximately 50%.)

C. Case 3: Psychometrically-informed advertising and customer relations

The idea of using authorship profiling to identify demographic information about actual and potential customers has been discussed in a previous section, and using this technology to infer demographics is well-understood [54], [28]. However, demographics is only half the story, and as the active authentication project has shown, it's possible to infer mental and social traits as well as demographic ones. This makes it both possible and practical to narrowcast messages to specific people based on previous text interactions with them.

The idea of targeting advertisements (or other corporate communications) to a specific person is of course not new; that's one of the fundamental premises behind cookie-based marketing. However, author profiling technology creates new opportunities for analysis, with a new channel providing classification information without needing to gather data from external sources. One specific application for this is in "inside sales," where communications with existing clients can be (re)analyzed to determine both the best approach for maintaining and extending the relationship. This can help, for example, by allowing better matching of successful sales representatives to customers, based on the types of customers and the types of representatives. This assures that customers have representatives that will be able to connect well with them, understand their needs, and create a closer, more beneficial association. This could be done on the basis of demographic and personality data as described in the previous section, or even on the basis of ad-hoc categories for each representative, representing empirically what each representative's strengths and weaknesses are.

V. Conclusions

Text analysis is well understood as a key business technology; it lets companies deal with large sets of documents easily and efficiently. Authorship analysis is not as well known or understood, but provides another key capacity, the ability to deal with large sets of clients and customers easily and efficiently. In this paper, we have described the basics of this technology and outlined several specific, practical applications that can have a major effect on industry.

References

[1] D. M. Blei, "Probabilistic topic models," *Communications of the ACM*, vol. 55, no. 4, pp. 77–84, April 2012.

[2] J. Wiebe, T. Wilson, and C. Cardie, "Annotating expressions of opinions and emotions in language," *Language Resources and Evaluation*, vol. 39, pp. 165–210, 2005.

[3] P. Elmer-DeWitt, "Samsung fined $340,000 for astroturfing in Taiwan," *Fortune*, 2013.

[4] B. Sterling, "The Chinese online 'water army'," *Wired*, vol. June, 2010.

[5] G. McMenamin, "Declaration of Gerald McMenamin," Available online at `http://www.scribd.com/doc/67951469/Expert-Report-Gerald-McMenamin`, 2011.

[6] M. Coulthard, "On admissible linguistic evidence," *Journal of Law and Policy*, vol. XXI, no. 2, pp. 441–466, 2013.

[7] T. Grant, "Txt 4n6: Describing and measuring consistency and distinctiveness in the analysis of SMS text messages," *Journal of Law and Policy*, vol. XXI, no. 2, pp. 467–494, 2013.

[8] P. Juola, "Stylometry and immigration: A case study," *Journal of Law and Policy*, vol. XXI, no. 2, pp. 287–298, 2013.

[9] C. E. Chaski, "Who's at the keyboard: Authorship attribution in digital evidence investigations," *International Journal of Digital Evidence*, vol. 4, no. 1, p. n/a, 2005, electronic-only journal: http://www.ijde.org, accessed 5.31.2007.

[10] J. F. Burrows, "'an ocean where each kind…' : Statistical analysis and some major determinants of literary style," *Computers and the Humanities*, vol. 23, no. 4-5, pp. 309–21, 1989.

[11] J. N. G. Binongo, "Who wrote the 15th book of Oz? an application of multivariate analysis to authorship attribution," *Chance*, vol. 16, no. 2, pp. 9–17, 2003.

[12] D. L. Hoover, "Delta prime?" *Literary and Linguistic Computing*, vol. 19, no. 4, pp. 477–495, 2004.

[13] E. Stamatatos, "On the robustness of authorship attribution based on character n-gram features," *Journal of Law and Policy*, vol. XXI, no. 2, pp. 420–440, 2013.

[14] G. K. Mikros and K. Perifanos, *Authorship attribution in Greek tweets using multilevel author's n-gram profiles*. Palo Alto, California: AAAI Press, 2013, pp. 17–23. [Online]. Available: http://www.aaai.org/ocs/index.php/SSS/SSS13/paper/view/5714/5914

[15] P. Juola, J. I. Noecker Jr, A. Stolerman, M. V. Ryan, P. Brennan, and R. Greenstadt, "Keyboard behavior-based authentication for security," *IT Professional*, vol. 15, pp. 8–11, 2013.

[16] J. Rudman, "On determining a valid text for non-traditional authorship attribution studies : Editing, unediting, and de-editing," in *Proc. 2003 Joint International Conference of the Association for Computers and the Humanities and the Association for Literary and Linguistic Computing (ACH/ALLC 2003)*, Athens, GA, May 2003.

[17] T. Joachims, *Learning to Classify Text Using Support Vector Machines*. Kluwer, 2002.

[18] J. R. Quinlan, *C4.5: Programs for Machine Learning*. Morgan Kauffman, 1993.

[19] M. L. Jockers and D. Witten, "A comparative study of machine learning methods for authorship attribution," *LLC*, vol. 25, no. 2, pp. 215–23, 2010.

[20] D. Foster, *Author Unknown: Tales of a Literary Detective*. New York: Holt Paperbacks, 2001.

[21] R. Brooks and C. Flyn, "JK Rowling: The cuckoo in crime novel nest," *Sunday Times*, vol. 14 July, 2013.

[22] R. Brooks, "Whodunnit? JK Rowling's secret life as wizard crime writer revealed," *Sunday Times*, vol. 14 July, 2013.

[23] P. Juola, "How a computer program helped reveal J. K. Rowling as author of A Cuckoo's Calling," *Scientific American*, vol. August, 2013. [Online]. Available: http://www.scientificamerican.com/article/how-a-computer-program-helped-show-jk-rowling-write-a-cuckoos-calling/

[24] C. A. Evans, *Luke*, ser. Understanding the Bible Commentary Series. Grand Rapids, Michigan: Baker Books, 1990.

[25] R. Garland, *Ancient Greece: Everyday Life in the Birthplace of Western Civilization*. New York: Sterling, 2008.

[26] M. Herper, "Linguist analysis says Newsweek named the wrong man as Bitcoin's creator," *Forbes Magazine*, vol. March 10, 2014. [Online]. Available: http://www.forbes.com/sites/matthewherper/2014/03/10/data-analysis-says-newsweek-named-the-wrong-man-as-bitcoins-creator/

[27] P. Juola, "The Rowling case: A proposed standard protocol for authorship attribution," in *Proceedings of Digital Humanities 2014*, Lausanne, Switzerland, 2014.

[28] S. Argamon, M. Koppel, J. W. Pennebaker, and J. Schler, "Automatically profiling the author of an anonymous text," *Communications of the ACM*, vol. 52, no. 2, pp. 119–123, February 2009.

[29] S. Argamon, S. Dhawle, M. Koppel, and J. W. Pennebaker, "Lexical predictors of personality type," in *Proceedings of the Classification Society of North America Annual Meeting*, 2005. [Online]. Available: citeseer.ist.psu.edu/744868.html

[30] R. H. Baayen, H. van Halteren, A. Neijt, and F. Tweedie, "An experiment in authorship attribution," in *Proceedings of JADT 2002*. St. Malo: Université de Rennes, 2002, pp. 29–37.

[31] M. Corney, O. de Vel, A. Anderson, and G. Mohay, "Gender-preferential text mining of e-mail discourse," in *Proceedings of* Computer Security Applications Conference, 2002, 2002, pp. 282–289.

[32] P. Juola and H. Baayen, "A controlled-corpus experiment in authorship attribution by cross-entropy," *Literary and Linguistic Computing*, vol. 20, pp. 59–67, 2005.

[33] M. Koppel, S. Argamon, and A. R. Shimoni, "Automatically categorizing written texts by author gender," *Literary and Linguistic Computing*, vol. 17, no. 4, pp. 401–412, 2002, doi:10.1093/llc/17.4.401.

[34] T. Kucukyilmaz, B. B. Cambazoglu, C. Aykanat, and F. Can, "Chat mining for gender prediction," *Lecture Notes in Computer Science*, vol. 4243, p. 274283, 2006.

[35] M. Oakes, "Text categorization: Automatic discrimination between US and UK English using the chi-square text and high ratio pairs," *Research in Language*, vol. 1, pp. 143–156, 2003.

[36] H. van Halteren, R. H. Baayen, F. Tweedie, M. Haverkort, and A. Neijt, "New machine learning methods demonstrate the existence of a human stylome," *Journal of Quantitative Linguistics*, vol. 12, no. 1, pp. 65–77, 2005.

[37] B. Yu, Q. Mei, and C. Zhai, "English usage comparison between native and non-native English speakers in academic writing," in *Proceedings of ACH/ALLC 2005*, Victoria, BC, Canada, 2005.

[38] J. Noecker Jr, M. Ryan, and P. Juola, "Psychological profiling through textual analysis," *LLC*, vol. 28, no. 3, pp. 382–387, 2013.

[39] G. K. Mikros, *Authorship Attribution and Gender Identification in Greek Blogs*. Belgrade: Academic Mind, 2013, pp. 21–32.

[40] R. P. Guidorizzi, "Security: Active Authentication," *IT Professional*, vol. 15, pp. 4–7, 2013.

[41] S. Acharya, A. Fridman, P. Brennan, P. Juola, and R. Greenstadt, "User authentication through biometric sensors and decision fusion," in *47th Annual Conference on Information Sciences and Systems (CISS 2013)*, 2013.

[42] J. I. Noecker Jr. and P. Juola, "Stylometric identification of manic-depressive illness," in *Proceedings of DHCS 2014*, 2014.

[43] P. Juola and J. I. Noecker Jr., "Inferring self-esteem from keyboard behavior," in *Proceedings of DHCS 2014*, 2014.

[44] S. Rude, E. Gortner, and J. Pennebaker, "Language use of depressed and depression-vulnerable college students," *Cognition and Emotion*, vol. 18, pp. 1121–1133, 2004.

[45] C. Poulin, B. Shiner, P. Thompson, L. Vepstas, Y. Young-Xu, B. Goertzel, B. Watts, L. Flashman, and T. McAllister, "Predicting the risk of suicide by analyzing the text of clinical notes," *PLoS One*, vol. 9, no. 3, 2014.

[46] P. Juola, "20,000 ways not to do authorship attribution and a few that work," in *Proceedings of 2009 Biennial Conference of the International Association of Forensic Linguists (IAFL-09)*, Amsterdam, 2009.

[47] C. Gray and P. Juola, "Personality identification through on-line text analysis," in *Proceedings of the 2011 Chicago Colloquium on Digital Humanities and Computer Science*, Chicago, IL, 2011.

[48] J. W. Pennebaker and L. A. King, "Linguistic styles: Language use as an individual difference," *Journal of Personality and Social Psychology*, vol. 77, pp. 1296–1312, 1999.

[49] M. Newman, J. Pennebaker, D. Berry, and J. Richards, "Lying words: Predicting deception from linguistic style," *Personality and Social Psychology Bulletin*, vol. 29, pp. 665–675, 2003.

[50] I. Picornell, "The flexible liar – a new approach to detecting deception in written narratives," in *Forensic Linguistics: Bridging the Gap(s) Between Language and the Law: 2012 Regional Meeting of the IAFL*. Porto, Portugual: International Association of Forensic Linguists, 2012.

[51] J. I. Noecker Jr and P. Juola, "Spanish lies : A computational reanalysis of a spanish corpus for detecting deception," in *Proceedings of the 11th Biennial Conference on Forensic Linguistics/Language and Law of the International Association of Forensic Linguists (IAFL 2013)*, Mexico City, MX, June 2013.

[52] D. Folkenflik, *Murdoch's World*. PublicAffairs, 2013.

[53] K. Kleiner, "Bogus grass-roots politics on Twitter," *Technology Review*, 2010.

[54] H. van Halteren, "Author verification by linguistic profiling: An exploration of the parameter space," *ACM Transactions on Speech and Language Processing*, vol. 4, p. n/a, 2007.

Detection Faults for Induction Machine Sensors Based on Fuzzy Logic Techniques

A. Amrane

National School of Technology, National road N 5 Z.I, Rouiba, 16013, Algiers (Algeria) Laboratory of Systems Electric and Industrial (LSEI), Faculty of Electronics and Informatics (FEI), BP 32, EIAlia, Bab-Ezzouar 16111, Algiers, Algeria
e-mail: amraneahmed@gmail.com

A. Larabi

University of Science and Technology (USTHB), Laboratory of Systems Electric and Industrial (LSEI), Faculty of Electronics and Informatics (FEI), BP 32, EIAlia, Bab-Ezzouar 16111, Algiers, Algeria
e-mail: larabiabdelkader@yahoo.fr

***Abstract*—This** paper concerns the development of the detection of faults the sensors of an induction machine by estimation methods for systems modeled in the form of state representation. The sensors monitors in our case are those of speed.

To achieve our objective, many of the techniques of artificial intelligence have been used to fault sensors detection of rotating machinery, where several selection techniques have been explored during the construction of the detection process.

We develop and combine the model reference adaptive system (MRAS) method with fuzzy logic approach to achieve our objectives. This type of estimation is applied to the supervision of the speed sensor in the system; this is done in order to give more robustness of the overall process, the second one is to study and develop an intelligent adaptation method based on the fuzzy logic algorithms, keeping the same performances. The new presented approach improves the performances of our system compared to the usual methods.

Finally, the validity of the proposed scheme is demonstrated by a series of computer simulations and the obtained results show that the designed system can achieve satisfactory performances.

***Keywords* —** Induction machine, model reference adaptive system MRAS, fuzzy detection, fault of sensors

INTRODUCTION

Many industrial applications require fault tolerance and continuity of service. [1] This is due to the growing need to improve the availability of systems. Therefore, strategies are defined in the early stages of design, to facilitate fault detection, localization and reconfiguration of the order. For this, several recent works [2, 3] deal with fault tolerant control of electric drives.

In this article, we focus our study on the induction machine (IM), conventionally designed for constant speed applications, has become, Due to its simple, robustness structure and the evolution of electronics power and control of the control vector machine most commonly used for variable speed drives. This machine has the advantage of being more robust and less expensive, with equal power, as other machines. However, it has drawn backs. This allowed the opening of various lines of research on its control and power [4, 5, 6].

We introduced in this work, an approach widely used for the detection and isolation of faults based on the method of estimator [7, 8], the main components of a tolerant control strategy to defects (CTD) sensor. It is divided into three stages: detection, isolation and reconfiguration.

A novel estimator schemes based on fuzzy logic algorithm for the speed estimation where derived using Lyapunov's stability theorem [6,9,10]. Several strategies have been proposed for the estimation in the vectorial induction motor drives [6,9,11]. Among these techniques, model reference adaptive systems (MRAS) schemes are the most common strategies employed due to their relative simplicity and low computational effort [9,11].

Can be seen, the majority of estimation schemes described in the literature for MRAS observer employ a PI controller to generate the desired value. However, due to the continuous variation in the machine parameters and the operating conditions, in addition to the nonlinearities present in the inverter, the PI controllers may not be able to provide the required performance. Not much attention has been devoted to study other types of estimation scheme [9,11,12].

In this article, the performances of the diagnostic of sensors, the detection of faults and the fuzzy estimation of speed for the induction machine are analyzed by digital simulation. To achieve our objective, this paper is organized as follows. Section I presents the control of the asynchronous machine; the rest of the paper is organized as follows. Sections 2 and 3 present the control of the IM and the MRAS algorithm using to estimate the speed. Section 4 depicts fuzzy logic algorithm for the estimation. Section 5 provides and discusses the simulation results and Section 6 sketches some conclusions.

II. CONTROL OF ASYNCHRONOUS MACHINES

A. Control Scheme of an Asynchronous Machine

Fig.1 shows a simplified block diagram illustrating the essential of our control system. This system consists of a dc power source, a dc link filter, a voltage source inverter, an induction motor, and same circuit of control system. The dc power source converts the constant-frequency ac power to dc power by a three-phase, full wave diode bridge rectifier; the dc

voltage is smothered by a smoothing capacitor dc link filter and then applied to a three-phase bridge inverter witch converts dc power to variable voltage variable frequency ac power supply to the motor. The two control parameters required are frequency and voltage, the frequency command also generates the voltage command through a volts/hertz ratio.

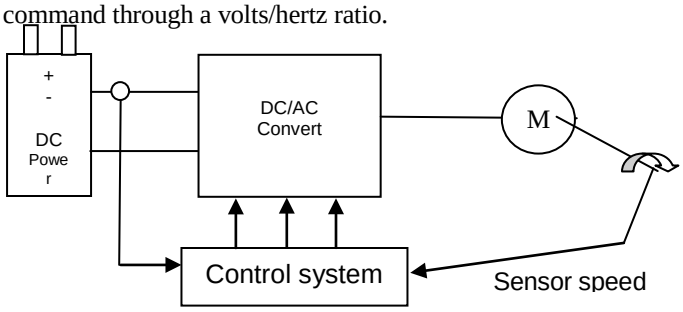


Fig.1: Block diagram of control system

B. INDUCTION MACHINE MODEL

The considered system is an induction machine with a squirrel cage of three phases. The state representation of the mathematical model of this machine can be represented according to usual d-q axes, as follows [6,9,10,11,14]

$$\begin{aligned}
\frac{d\,ids}{dt} &= 1/(\sigma * Ls)\left[-Rsm*ids+\omega s*\sigma*Ls*iqs+\frac{Lm}{(Lr*Tr)}*\phi dr\right]\\
\frac{d\,iqs}{dt} &= 1/(\sigma * Ls)\left[-Rsm*iqs-\omega s*\sigma*Ls*ids-\frac{Lm}{(Lr*Tr)}*\phi qr\right]\\
\frac{d\,\phi dr}{dt} &= \frac{Lm}{Tr}*ids-\frac{1}{Tr}*\phi dr+(\omega s-\omega r)*\phi qr\\
\frac{d\,\phi qr}{dt} &= \frac{Lm}{Tr}*iqs-\frac{1}{Tr}*\phi qr-(\omega s-\omega r)*\phi dr\\
Cem &= \frac{P*Lm}{Tr}(iqs*\phi dr-\phi qr*ids)\\
\frac{d\,\Omega r}{dt} &= \frac{1}{Tr}(Cem-Cr-Kf*\Omega r)
\end{aligned} \qquad (1)$$

C. Indirect Field-Oriented Control of an Induction Machine

This technique consists in assimilating the behaviour of asynchronous machine to the DC machine; this is for applications requiring raised dynamic performances [5,11,12,13].

The aim of such a control technique of the torque and a choice of the (d,q) reference, in order to obtain a decoupling between the torque and the field. From equation system (1); the torque control is made on the components of current iq and is. The electromagnetic torque depends only on component iq. It is a maximum for a given current if we impose is=0. Consequently the obtained torque is then proportional to the current of machine power supply as in case of separately excited DC machine. The model given by the equations (1) is expressed as follows:

$$\begin{aligned}
Vds &= \sigma*Ls*\frac{d\,ids}{dt}-1/(\sigma*Ls)\left[-Rsm*ids+\omega s*\sigma*Ls*iqs+\frac{Lm}{(Lr*Tr)}*\phi dr\right]\\
Vqs &= \sigma*Ls*\frac{d\,iqs}{dt}-1/(\sigma*Ls)\left[-Rsm*iqs-\omega s*\sigma*Ls*ids-\frac{Lm}{(Lr*Tr)}*\phi qr-\omega r*\frac{Lsr}{Lr}*\phi dr\right]\\
\frac{d\,\phi dr}{dt} &= \frac{Lm}{Tr}*ids-\frac{1}{Tr}*\phi dr\\
\omega s &= \frac{Lm}{Tr}*\frac{iqs}{\phi dr}+\omega r\\
Cem &= \frac{P*Lm}{Tr}(iqs*\phi dr)\\
\frac{d\,\Omega r}{dt} &= \frac{1}{Tr}(Cem-Cr-Kf*\Omega r)
\end{aligned} \qquad (2)$$

It is noted that rotor flux depends only on the satatoric current is and iq, that the electromagnetic couple depends only on the quadratic current iq.

III. The Structure of Model Reference Adaptive System MRAS

The aim of this technique of control is to replace the speed sensor by a speed estimator. Into our study, we introduced a speed estimator of type MRAS (Model Reference Adaptive System). The latest replace the mechanical sensor without changing the dynamics of our machine.

The principle of the method MRAS speed estimation rests on the comparison of the sizes obtained in two different ways. One model of such a method is the voltage one (or stator equation) and the other is current model (or rotor equation); because the voltage method doesn't include rotor speed then it does not depend explicitly of speed (model of reference) and the other includes rotor speed (adjustable adaptive model). Fig.2 illustrates the derived MRAS scheme for speed adaptation.

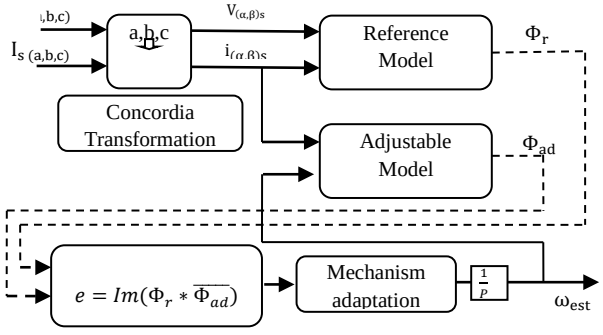


Fig.2: Structure of MRAS estimation of speed.

A. Reference Model

One uses the equations of the currents (2) of the IM, the current expressed in the reference fixed to the stator. The two equations of the model of reference become:

$$\begin{aligned}\frac{d\,\phi\alpha r}{dt} &= \frac{Lr}{Lsr}\left[V\alpha s - \sigma * Ls * \frac{di\alpha s}{dt} - Rs * i\alpha s\right]\\ \frac{d\,\phi\beta r}{dt} &= \frac{Lr}{Lsr}\left[V\beta s - \sigma * Ls * \frac{di\beta s}{dt} - Rs * i\beta s\right]\end{aligned} \quad (3)$$

B. Adjustable Model

To establish the adaptive model, we expressed the current in the reference fixed to the stator, the fluxes expressed with the rotor sizes in a reference αβ are:

$$\begin{aligned}\frac{d\,\phi\alpha ad}{dt} &= -\frac{1}{Tr}\phi\alpha ad + \frac{Lsr}{Tr} * i\alpha s - P * \Omega est * \phi\beta ad\\ \frac{d\,\phi\beta ad}{dt} &= -\frac{1}{Tr}\phi\alpha ad + \frac{Lsr}{Tr} * i\beta s - P * \Omega est * \phi\alpha ad\end{aligned} \quad (4)$$

C. Adaptation Mechanism

The entry of an adaptive mechanism is activated by the error between the reference field and adaptive field. By carrying out the difference between the reference model and the adjustable model, we obtain the following system of equations which govern the adaptive mechanism.

$$\begin{bmatrix}\frac{de\alpha}{dt}\\ \frac{de\beta}{dt}\end{bmatrix} = \begin{bmatrix}-\frac{1}{Lr} & -\omega\\ \omega & -\frac{1}{Lr}\end{bmatrix} + (\omega - \omega ad)\begin{bmatrix}\phi\alpha ad\\ \phi\beta ad\end{bmatrix} \quad (5)$$

The adaptation law chosen to ensure the convergence of ωad towards ω is:

$$\omega ad = Tp * \delta e + Ti\int_0^t \delta e * dx \quad (6)$$

The adaptive mechanism has an integral proportional form:

$$\omega est = {}^{1}\!/_{P} * \left(Tp * e + Ti\int e * dt\right) \quad (7)$$

Where Tp and Ti are positive gain.

IV. THE FUZZY LOGIC ADAPTATION

A) Control by Fuzzy Logic

Fuzzy logic technique makes it possible to control nonlinear systems and complicated models [10,11,12]. In fact, the calculation of the parameters of the system is not necessary to carry out this control [10,11,12].

On opposite of the adaptation of the traditional techniques; the fuzzy logic does not treat a mathematical relations well defined, but uses inferences with a several rules, being based on variables linguistic.

These inferences are treated by operators suitable for fuzzy logic [10,11,12]. Fig.3, shows the structure of a fuzzy regulator with two input (X1 and X2) and one output (Xr).

So, we can note that the calculation of the control is carried out starting from three fundamental stages: an interface of fuzzification; a mechanism of inference (rules); and an interface of defuzzification.

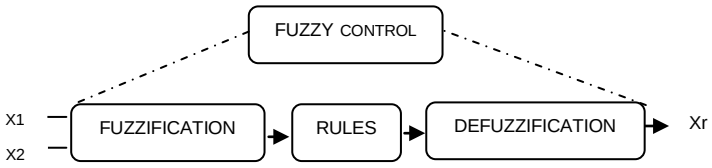


Fig. 3: Functional diagram of the fuzzy control

1. THE FUZZIFICATION

The entries and exits are defined of the fuzzy are defined by membership functions with 7, 5 or 3 sets. The various sets are characterized by standard designations [9,11,12]:

Negative Big NB, Negative Medium NM, and Negative Small NS, Zero Z, Positive Small PS, Positive Medium PM, Positive Big PB.

2. INFERENCE MECHANISM

It is well know that the realizations of the matrix of the rules are deduced by experiment, the experiment of the human operators and rests on the analysis of the system.

This analysis must take into an account the trajectory which one wants to give to the system [9,11,12].

3. THE DEFUZZIFICATION

By this stage; the return to the sets of real exit will be made. It is a question of calculating, from the degrees of membership of all the sets variable of output, the coordinate which corresponds with the value of this exit. Various methods are used [9,11,12].

A. The Fuzzy Logic Adaptation Mechanism Principle

A fuzzy logic controller as shown in (Fig.4) will replace the structure of the proposed PI controller used in the adaptation mechanism.

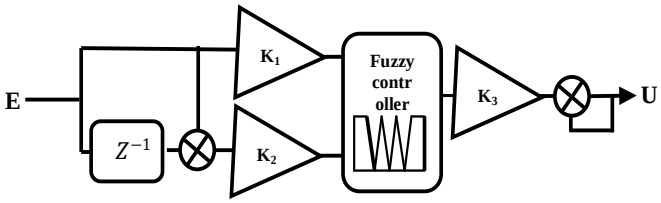


Fig.4: Synoptic diagram of a fuzzy adaptation

In our work, we adopted seven sets for the two variables of input (En and dEn), and a same number of sets for the variable of output dUn. The rules of the controller can be presented in a

matrix with seven sets known as matrix of inference shown in table (1):

Table.1: The Fuzzy Rules (a regulator with 5 sets).

dU_n		E_n				
		NG	NP	EZ	PP	PG
dE_n	NG	NG	NG	NP	NP	EZ
	NP	NG	NP	NP	EZ	PP
	EZ	NP	NP	EZ	PP	PP
	PP	NP	EZ	PP	PP	PG
	PG	EZ	PP	PP	PG	PG

In this article, the memberships function are chosen as the triangular type and trapezoidal, and the method of reasoning is considered as the max-min method, the defuzzification stage is done based on the gravity centre method, as it is frequently quoted in the literature and because they requiring less time computing will be also adopted in our work [10,11,12].

A. The Basic Rules of the Fuzzy Controller

The first step in the design of fuzzy controller is to generate fuzzy rules based on the knowledge of the expert. According to the expert, three situations can be distinguished for the constant time estimation, above, around and below the reference desired.

In our control, in order to determine the rules table that generates a command that will be presented later Based on Fig 5

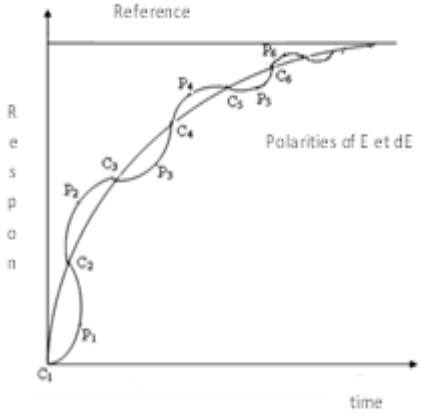


Fig.5: Temporal analysis of the rules of the regulator.

V. THE MECHANISM OF THE DETECTION ON THE SPEED SENSOR

To diagnose the speed sensor, a MRAS estimator is used as an observer. Then, an algorithm for selecting is used to enable the detection of defects and perform the selection between the measured signals and the estimated signals (ξ to predetermine by the user).

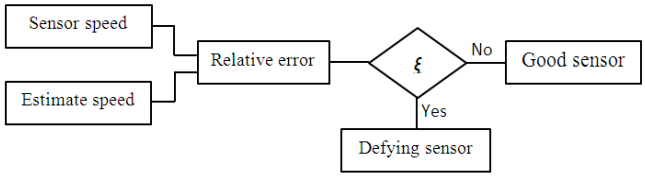


Fig.6: Synoptic diagram of the mechanism of the detection on the speed sensor

VI. SIMULATION RESULTS

To diagnose the speed sensor, a MRAS estimator is used as an observer. Then, an algorithm for selecting is used to enable the detection of defects and perform the selection between the measured signals and the estimated signals.

So, to validate the performances of the MRAS, the proposed simulation of the dynamic behaviour of the machine has been done using the MATLAB/SIMULINK and that for the following conditions.

With, the value of the torque is fixed to zero and the field is fixed to 1Wb. To highlight the influence of speed variations and uncertainties, particularly those of the control process, we gave reference speed of 100 rad/s. Fig.7

Afterwards, constant field and a reference speed Ωref = 100 rad/s, the field rotor field is fixed to 1Wb, the initial values of the torque assumed by the machine is zero, and between 1s and 2.5s, one will apply a nominal torque load (10 N.m). Fig.8

Finally, we will use our estimator to diagnose faults sensor for speed exceeds the rated speed to a value of the torque is fixed to zero, the field is fixed to 1Wb and the initial values of the speed assumed by the machine is zero, and between 1s and 2.5s, we will apply variations. Fig.9

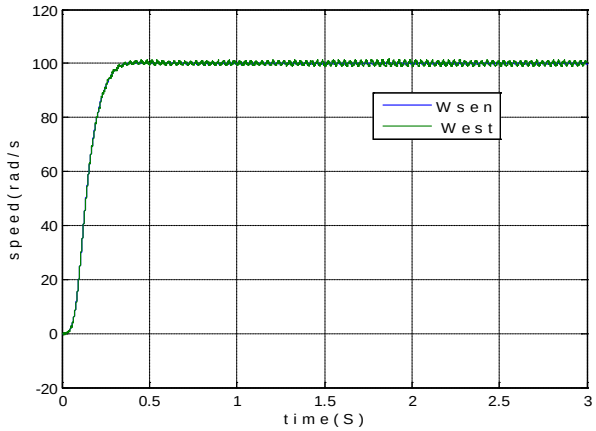


Fig.7: Estimation of speed in without defect estimated speed (West) and speed with sensor (W_{Sen}). With the torque is fixed to zero and the field is fixed to 1Wb and the speed fixed to 100 rad/s

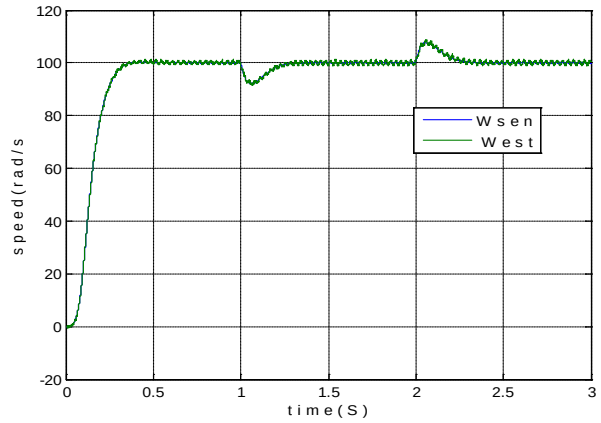


Fig.8: Estimation of speed in without defect, estimated speed (West) and speed with sensor (Wsen). With the torque is fixed to 10 N.m and the field is fixed to 1Wb and the speed fixed to 100 rad/s

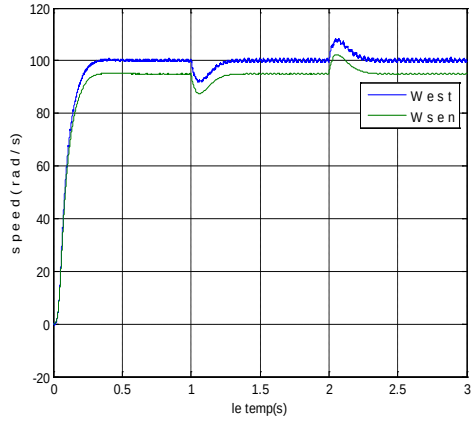


Fig.9: Estimation of speed with defect, estimated speed (West) and speed with sensor (Wsen). With the torque is fixed to 10 N.m and the field is fixed to 1Wb and the speed fixed to 100 rad/s,the tests for a defiant sensor of 0.95 of its value.

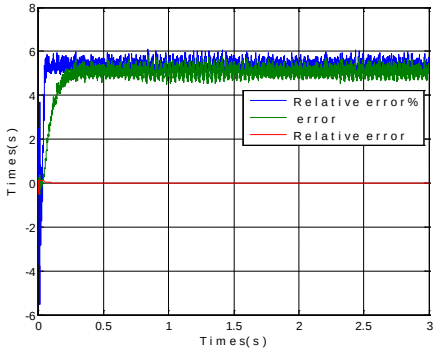


Fig.10: The error of speed with defect. With the torque is fixed to 10 N.m and the field is fixed to 1Wb and the speed fixed to 100 rad/s and The tests for a defiant sensor of 0.95 of its value .

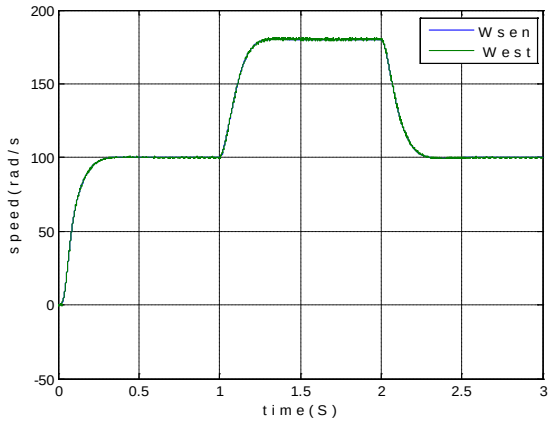


Fig.11 : Estimation of speed with defect, estimated speed (West) and speed with sensor (Wsen). With the torque is fixed to zero and the field is fixed to 1Wb and and between 1s and 2.5s, we will apply variations of speed.

According to the simulation results we can say that the MRAS technique of detection provides an effective solution to the problem of diagnostic. We can say that our objective here has been successfully achieved.

The results presented on the previous figures, show that the dynamics of the flux magnitude are presented can highlight the decoupling role of the flux controller where the flux tracks its nominal value of 1.1 Wb for all speed ranges.

Numerical simulations are performed and shown good results for tricking the faults. Implementation of the proposed algorithm can be used for the control speed issues suitable for IM applications requiring high-performance torque control and faults detection.

The performance of the estimator, based on fuzzy regulators, is verified, We can say that our objective here has been successfully achieved.

VII. CONCLUSION

The work that we presented contributes to the analysis and the synthesis of a robust diagnostic applied to the induction machine. The MRAS is employed for the detection of the faults of sensor speed. The use of the fuzzy logic is a powerful tool in realization of the robust and reliable diagnostic.

We have proposed a method for detection faults of sensor speed using the MRAS and the fuzzy logic algorithm, to ensure a good diagnostic of induction machines. The method proposed in this paper is applicable to a large category of induction motor drives with a gradually varying load torque. The tests of robustness show clearly that the performances of the diagnostic in the presence of estimator, type MRAS, for the tracking of the faults is always fast.

The validity of this method is checked by several tests. The results obtained show that the model suggested for the MRAS adapts perfectly to all the diagnostic of the IM.

The next step of this work is the integration of a real induction motor and estimator MRAS for testing the diagnostic is required for the practical case.

ISBN: 978-1-61804-327-6

VIII. ANNEXES

The parameters of the induction machine cage used are shown below:

Rated power:	1.5 Kw
Nominal voltage:	220/380 V.
Rated power factor:	0.8.Rated
Speed:	1420 rev/ min.
Nominal frequency:	50 Hz.
Stator resistance:	4.85Ω.
Rotor resistance:	3.805Ω.
Stator cyclic inductance:	0.274 H.
Cyclic inductance of Rotor:	0.274 H.
Cyclic mutual inductance:	0.25 8 H.
Number of pole pairs:	2.
Moment of Inertia:	0,031Nm-s^2 / rad
Friction:	. 0.008 Nms /r

IX. NOMENCLATURE

Ls	Stator inductance [H].
Lr	Rotor inductance [H].
Lm	Mutual magnetizing inductance
Lsr	Mutual Inductance between the stator and rotor [H].
Kf	Friction coefficient [N.s/rad].
J	Total inertia [kg.m2].
P	Number of pole pairs.
ωs	Synchronous Pulsation [rad/s].
ωr	Electrical angular Pulsation [rad/s].
Cem	Electromagnetic torque [N.m].
Cr	Resistive torque [N.m].
Tr	Rotor time constant [s].

X. REFERENCES

[1] Benbouzid, M.E.H., Diallo, D., Zeraoulia, M., March 2007; "Advanced Fault-Tolerant Control of Induction-Motor Drives for EV/HEV Traction Applications: From Conventional to Modern and Intelligent Control Techniques," Vehicular Technology, IEEE Transactions on, vol.56, n° 2, pp.519-528.

[2] Green, S.; Atkinson, D.J.; Jack, A.G.; Mecrow, B.C.; King, A. , 2003; "Sensorless operation of a fault tolerant PM drive," Electric Power Applications, IEE Proceedings, vol.150, n° 2, pp. 117- 125

[3] Abolhassani, M.T., Toliyat, H.A., May 2009, "Fault tolerant permanent magnet motor drives for electric vehicles," Electric Machines and Drives Conference, 2009. IEMDC '09. IEEE International, pp.1146-1152, 3-6

[4] H.Shoudao, W.Yaonan, G.Jian, L.Jiantao and Q.Sihai, 2004, "The vector control based on MRAS speed sensorlesse induction machine drive, intelligent control and automation", WCICA, 5th World congress, 4550-4553.

[5] M. Montanari, S. Peresada, A. Tilli, & A. Tonielli, , 2000," Speed sensorless control of induction motor based on indirect field- orientation", 35th Annual Meeting, IEEE Industrial Application Society, Rome, 1858–1865.

[6] M. Ouhrouche, 2002, "Estimation of speed, rotor flux and rotor resistance in cage induction motor using the EKF algorithm", International Journal of Power and Energy Systems, 103–109.

[7] Chen, J., Patton, R., 1999, "Robust model-based fault diagnosis for dynamic systems", Kluwer Academic, Boston.

[8] J. Holtz, 1996 ," Methods for speed sensorless control of AC drives ", in Sensorless Control of AC Motor Drives, Piscataway, NJ, IEEE Press

[9] A.Amrane, A.Larabi, 2012," Speed control sensorless of an induction machine powered by photovoltaic generator ", 10th Global Conference on Sustainable Manufacturing (GCSM'12), 31 October au 02 November 2012. Istanbul, Turkey.

[10] A. Amrane, A. Larabi, A. Hamzaoui, 2014, "Robust sensorles control of an induction machine based on a fuzzy MRAS" International Conference on Electrical and Electronics Engineering (ICEEE 2014) du 21-23 April, Antalya, Turquie.

[11] A. Amrane, M.Louri, A. Larabi, A. Hamzaoui, 2013, "A Fuzzy Model Reference Adaptive System Control for Induction », 3rd International Conference on Systems and Control, October 29-31, Algiers, Algeria

[12] A. Larabi, A. Amrane , M.S.Boucherit,2013, "Robust Speed Sensorless Fuzzy Control of the Induction Machine Drive with Adaptation of the Rotor Time Constant Supplied by Photovoltaic Solar Energy », International Conference on Power Engineering, Energy and Electrical Drives, POWERENG-2013.13-17,May, Istanbul, Turkey.

[13] M. Louri, A. Amrane, L. Barazane, 2013, Comparison between the performances of variables structures Control and the theory of synergetic on applied to the squirrel motor drives », 3rd International Conference on Systems and Control, October 29-31, Algiers, Algeria

[14] T.-J. Ho and L.-Y Yeh, 2010, " Design of a Hybrid PID Plus Fuzzy Controller for Speed Control of Induction Motors", 5th IEEE Conference on Industrial Electronics and Applicationsis, pp 1352-1357, in KINTEX, Gyeonggi-do, Oct. 27-30, Korea.

[15] Youssef agrebi, Moncef Triki, Yassine Koubaa et Mohamed Boussak, 2006, "Commande avec et sans capteur mécanique d'une machine asynchrone". 7ème conférence internationale des sciences et techniques de l'automatique STA'2006,17-19 décembre, Hammamet Tunisie, pp. 573-584.

[16] F. Peng and T. Fukao, 1994, "Robust speed identification for speed-sensorless vector control of induction motors". IEEE Trans. Ind. Applicat., vol. 30, Sept. /Oct, pp. 1234–1240.

[17] A.Hazzab, 2006, "Commande des systèmes par logique floue, Réseau de neurones et Algorithmes géniques ", Thèse de doctorat, Université des sciences et de la Technologie Mohamed Boudiaf d'Oran, USTO.

Shape Matching Method Based on Spatial Features of Multi-scaled Contours

Min Han, Yafei Yang, Danchen Zheng and Jun Wang
Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology
Dalian, China
minhan@dlut.edu.cn

***Abstract*—In order to improve the robustness and discrimination power of the triangle-area representation, a shape matching method based on multi-scaled contour space relationship is proposed in this paper. The angle information in multi-scales, which used to represent shape contours, can be obtained by analyzing the spatial relationship of different points, and a new shape descriptor is further constructed. With the proposed descriptor, the multi-scaled information of contours can be expressed accurately, and the cost between descriptors of the corresponding points from similar shapes can be effectively decreased. As the orders of all the contour points are already known, dynamic programming method is used to calculate shape matching results during the feature matching step. The proposed algorithm has been tested on different shape databases, and the performances are superior to many other methods.**

Keywords—shape matching; multi-scale; contour space relationship; dynamic programming

I. INTRODUCTION

Shape is an important visual basis to recognize objects, and shape matching is a very critical problem in computer vision [1]. The similarities between different shapes can be obtained by using some certain shape matching methods, which have been applied to image retrieval, medical image analysis, face recognition and many other fields [2].

Shape matching methods mainly consist of two steps: shape feature extraction and shape feature matching [3]. The accuracy of shape matching results generally depends on the construction of the shape descriptors. And to some extent, better descriptors can develop higher accuracy of feature matching. Shape matching algorithms can be classified into two categories by the differences of shape feature extraction methods: the contour-based shape matching method and the skeleton-based shape matching method [4]. Since the contour-based method can achieve comparable recognition results with low computation complexity, numerous scholars have paid attention to this method. As a representative type of method in contour theories, the methods based on multi-scale theories are commonly used to extract the features of shapes [5]. The local and global information of shapes can be captured by the multi-scaled methods using a continuously changed scale parameter.

Mokhtarian et al. [6] proposed the curvature scale space image (CSS image), which constitutes a descriptor which is one of the standard descriptor in the MPEG-7 database. Latecki et al. [7] used a descriptor based on vison part to replace the continuous segments between two points with a straight line. Adamek et al. [8] proposed using Multi-scale convexity concavity (MCC) as shape descriptor and judged the concavity and convexity of contours through the Euclidean distance between a point and the same point on an adjacent scale. During the processes of generating the features in the above methods, the contours should be evolved at first, and this step makes the methods complex and time-consuming. These kinds of methods employed different shape descriptors to describe shapes on multi-scales. While using the above methods to extract features, the contours need to be evolved at first, this step will increase the computing efforts. In order to simplify the procedures of feature extraction, Donoser et al. [9] proposed a shape matching method based on multi-scaled chord: choose two points p, q and a point r between p and q randomly; the corresponding angle of q is the shape descriptor; the size of scales is adjustable by the chord length between p and q. All the methods above provide satisfied results in shape retrieval and recognition.

Alajlan et al. [10] investigated the triangle-area representation (TAR) method to represent the shape. The contour is convex, concave or straight, while the triangle-area made up of three continuous points is positive, negative or zero. And as the distance between the points changes, the local and global information of shapes can be obtained. Alajlan et al. [11] further proposed a dynamic programming method called dynamic space warping (DSW) to match the extracted features.

During the recognizing of some similar shapes, the triangle-area representation method couldn't operate robustly, and the sampling points at corresponding position in similar shapes may not be exactly matched. And when dealing with dissimilar points on different contours, the triangle-area representation method shows low discrimination power, which leads to mismatching of dissimilar points. To solve the above-mentioned problem, this paper presents a shape matching method based on spatial features of multi-scaled contours (SFMC), which can get contour features with better robustness and more powerful discrimination. As the orders of all the contour points are already known, dynamic programming (DP) algorithm is introduced to match the features and the accuracy of recognition and retrieval can be further ensured.

This work was supported by the National Natural Science Foundation of China (No. 61374154), the National Basic Research Program of China (973 Program) (No. 2013CB430403) and the Fundamental Research Funds for the Central Universities (No. DUT14RC(3)128).

II. Shape Matching Method Based on Triangle-Area Representation

In triangle-area representation method, each contour is represented by N sampling points, and the nth point is expressed as (x_n, y_n). Then the area of triangle which is made up of the points (x_n, y_n), (x_{n-t_s}, y_{n-t_s}) and (x_{n+t_s}, y_{n+t_s}) can be defined as

$$TAR(n,t_s)=\frac{1}{2}\begin{vmatrix} x_{n-t_s} & y_{n-t_s} & 1 \\ x_n & y_n & 1 \\ x_{n+t_s} & y_{n+t_s} & 1 \end{vmatrix} \quad (1)$$

where $n\in[1, N]$, $t_s\in[1, T_s]$, $T_s= N/2-1$. And watching from clockwise direction, when the values of (1) are positive, negative or zero, the contours are convex, concave or straight, respectively. For the convenience of processing the data, the values of (1) are normalized in the same scale

$$\overline{TAR}(n,t_s)=\frac{1}{\max TAR_{t_s}}TAR(n,t_s) \quad (2)$$

$$\max TAR_{t_s}=\max_{1\le n\le N} TAR(n,t_s) \quad (3)$$

Finally, the triangle-area features are represented as $\overline{TAR}(n,t_s)$. Given two contours A and B, for the nth point in A and the mth point in B, the loss function of the two points is

$$d(n,m)=\frac{1}{T_s}\sum_{t_s=1}^{T_s}\left|\overline{TAR}_A(n,t_s)-\overline{TAR}_B(m,t_s)\right| \quad (4)$$

where $\overline{TAR}_A(n,t_s)$ and $\overline{TAR}_B(m,t_s)$ are the triangle-area features of A and B, respectively. $n, m\in[1, N]$.

After extracting the shape features, dynamic space warping (DSW) algorithm is used to solve the problem of feature matching. Fig. 1 illustrates the three kinds of triangle-area, the number of sampling points is N, and three sampling points n_1, n_2, n_3 represent convex, concave and straight-line area, respectively. Fig. 2 shows the triangle-area features of all the points in multi-scales. In Fig. 2, n is the index value of the points, t_s is the size of the scales, and $\overline{TAR}(n,t_s)$ is the feature value.

III. Spatial Features of Multi-Scaled Contours

With the triangle-area representation, shape information can be effectively obtained. However, given two similar shapes, the distances between descriptors of corresponding points, in some cases, are extremely large. Using the triangle-area representation method with the two similar shapes shown in Fig. 3(a), the point n_a on the left shape is regarded as a convex point and the point m_a on the right shape is regarded as a concave point. Given a small t_s, $\overline{TAR}(n_a,t_s)$ of sampling point n_a provides a big positive value, and $\overline{TAR}(m_a,t_s)$ of the corresponding sampling point m_a provides a negative number with big absolute value. Therefore, the relationships between the corresponding points on similar shapes cannot be exactly reflected.

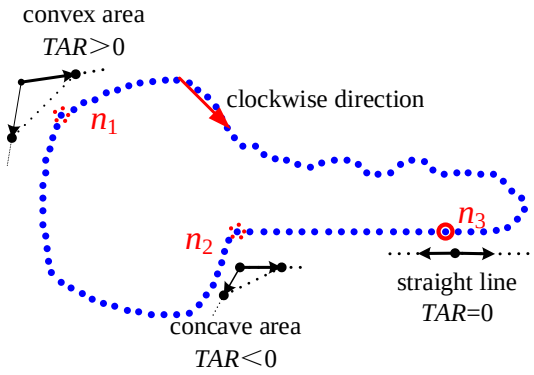


Fig. 1 The sketch of triangle-area representation

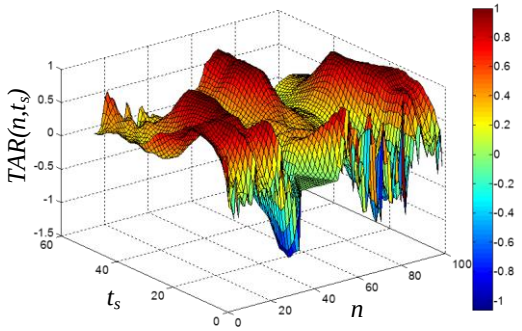
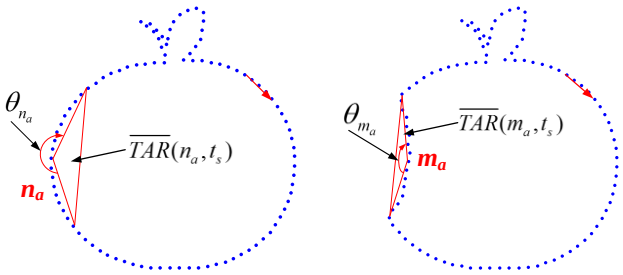
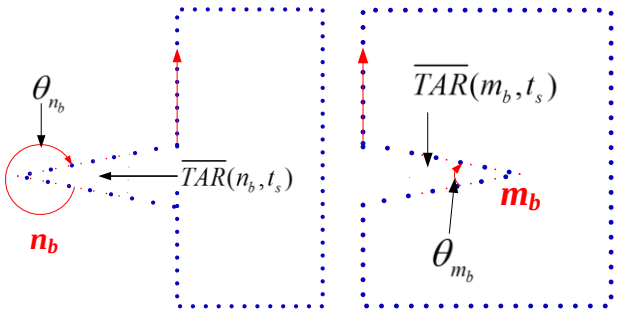


Fig. 2 The triangle-area features of the shape in Fig. 1



(a) Comparison of points at corresponding position on similar shapes



(b) Comparison of points at corresponding position on dissimilar shapes

Fig. 3 Comparison of triangle-area representation method and the method proposed in this paper

Meanwhile, sometimes the matching cost between shape features from different shapes is low, which may lead to the mismatching of sampling points. Fig. 3(b) shows parts of shapes with sharply convex contour or sharply concave contour. $\overline{TAR}(n_b,t_s)$ provides a small positive value, and $\overline{TAR}(m_b,t_s)$ provides a negative number with small absolute value. Though the configurations of the two shapes in Fig. 3(b) are very

different, the distance between the descriptors of n_b and m_b is small making the points mismatched.

In order to solve the above problems, this paper proposes a method based on spatial features of multi-scaled contours. Each contour is represented by N sampling points, with the nth point expressed as (x_n, y_n). Then the angle which is made up of (x_n, y_n), (x_{n-t_s}, y_{n-t_s}) and (x_{n+t_s}, y_{n+t_s}) is expressed as $\theta(n,t_s)$. Watching from clockwise direction, when the value of the angle is no more than π, it is defined as

$$\theta_1(n,t_s) = \arccos \left(\frac{(x_{n-t_s}-x_n, y_{n-t_s}-y_n)\cdot(x_{n+t_s}-x_n, y_{n+t_s}-y_n)}{\left\|(x_{n-t_s}-x_n, y_{n-t_s}-y_n)\right\|_2 \cdot \left\|(x_{n+t_s}-x_n, y_{n+t_s}-y_n)\right\|_2} \right) \quad (5)$$

when the value of the angle is larger than π, it is defined as

$$\theta_2(n,t_s) = 2\pi - \theta_1(n,t_s) \quad (6)$$

Besides, The discriminative power of shape descriptors depends more on the nearby samples than the points far away. Therefore, the weight of local features should be increased. In order to increase the weight of local features and avoid the influence of global features, the exponential function of e is used to weight the feature values in various scales. Besides, using angle features to describe shapes in the same scale, the loss function of different shapes is very small when the value of $\theta(n,t_s)$ is near 0, π or 2π. Then the sine function is brought in to map the features generating better shape descriptors.

Furthermore, the features can be represented as

$$SFMC(n,t_s) = \begin{cases} e^{-\alpha t_s} \cdot \sin(\theta_1(n,t_s)) \\ e^{-\alpha t_s} \cdot \sin(\theta_2(n,t_s)) \end{cases} \quad (7)$$

For the nth point in A and the mth point in B, the loss function of the two points is

$$d(n,m) = \frac{1}{T_s} \sum_{t_s=1}^{T_s} \left| SFMC_A(n,t_s) - SFMC_B(m,t_s) \right| \quad (8)$$

where $SFMC_A(n,t_s)$ and $SFMC_B(m,t_s)$ are the shape features of A and B. $\alpha \in [0.2, 0.3]$ is a weighting coefficient. $t_s \in [1, T_s]$ and T_s is an integer less than $N/2$. n, $m \in [1, N]$. For many shapes, when the number of scale value t_s is large, the value of $SFMC(n,t_s)$ tends to be zero, which can result in the mismatching of shapes. So the upper bound of scales must be restricted, which means only the features in low scales are used to describe the shapes. Thus, the local and global information of shapes can be expressed, and the computation complexity is decreased at the same time. In this paper, T_s and α are set as T_s=25 and $\alpha = 0.25$ via cross validation.

Fig. 4 demonstrates an example of shape feature construction using the proposed method, and the points n_1, n_2, n_3 represent convex, concave and straight-line area, respectively. Using the proposed feature extraction method, the shape features of Fig. 4 can be obtained , as shown in Fig. 5. Moreover, Fig. 5 gives the features of all the points in multi-scales, in which n is the index value of the points, t_s is the size of the scales, and $SFMC(n, t_s)$ is the feature value.

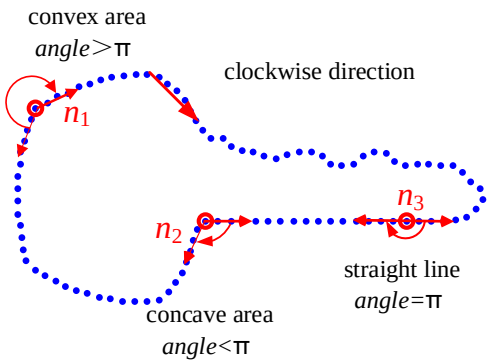


Fig. 4 The sketch of the proposed feature extraction method

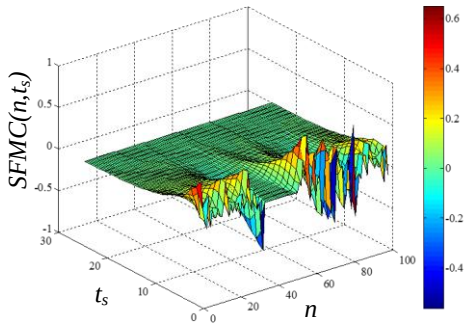


Fig. 5 The values of the features getting through the method in this paper

IV. Matching the Features of the Shapes

Previously, a novel shape extraction method is proposed, and this section introduces the corresponding feature matching process. With the orders of all the contour points already known, dynamic programming algorithm is used to solve the feature matching problem.

A. Feature Matching Method Based on Dynamic Programming Algorithm

The principles of using dynamic programming algorithm to match the features is illustrated in Fig. 6. As shown in the figure, the shadow parts at the top and the left side represent the features of different shapes. Starting from arbitrary points on A and B, the features are matched from left to right, top to bottom, through a diagonal window with width w starting from top left. By limiting the width of the window to w ,which is an integer less than $N/2$, there is no need to compute the distances between all the points of A and B. The computational complexity is largely reduced by only computing the distances between the points which are closer. The best matching path of A and B is presented by the arrows.

For A and B, a distance matrix D with size $N \times N$ can be obtained by using (8) to compute the distances of all the corresponding points. Then a desirable matching path between A and B can be obtained by combining dynamic programming algorithm. The minimum distance of the

corresponding points is expressed as $DT(n, m)$, and the initial value of the elements in DT is

$$DT_{initial}(n,m) = \infty \tag{9}$$

Then the elements of DT are updated by (10)

$$DT(n,m) = \min \begin{cases} DT(n-1,m-1) + D(n,m) \\ DT(n-1,m) + \Delta \\ DT(n,m-1) + \Delta \end{cases} \tag{10}$$

where $DT(1,1) = D(1,1)$, $\Delta \in [0.9, 1.3]$ is a threshold value which can avoid one point of a contour matching several times with different points of another contour. And n, $m \in [1, N]$, $\max(1, n-w+1) \le m \le \min(N, n+w-1)$. The last element of DT is $DT(N, N)$, which represents the minimum distance between A and B.

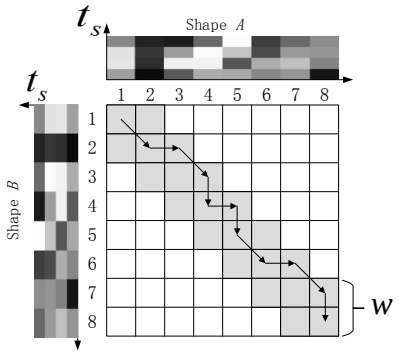


Fig. 6 The sketch map of dynamic programming algorithm

Besides, to avoid the influences caused by the rotating of shapes, we fix the query target A and rotate the test target B several times, and choose the minimum distance from all the results. The specific operation is: for the N points of query target A, fix the first point as the initial sampling points; for the N points of test target B, the $(1+\lfloor N/g \rfloor * (i-1))$th point is set as the initial sampling points in turn; after matching the two shapes for g times, record the minimum distance in all the results. $\lfloor \ \rfloor$ is a ceiling function, $1 \le i \le g$, and g is an integral[12] chosen from 8~12, and in this paper g is set as 10.

Finally the minimum distance between A and B is

$$DT_{\min} = \min\{DT_i(N,N), 1 \le i \le g\} \tag{11}$$

where $DT_i(N, N)$ is the minimum distance between A and B after the two shapes are matched for i times.

B. Basic Steps of Shape Matching

There are four main steps to match the query target A and the test target B using the proposed method:

1) Extract the contours of shapes A and B, and represent contours A and B with N sampling points, respectively;

2) Get the feature sets of all the N sampling points from A and B using (7), and the feature sets of A is $\{SFMC_A(n,t_s)\}$, $1 \le n \le N, 1 \le t_s \le T_s$, the feature sets of B is $\{SFMC_B(m,t_s)\}$, $1 \le m \le N$;

3) Use the feature sets $\{SFMC_A(n,t_s)\}$ and $\{SFMC_B(m,t_s)\}$ and equation (8) to compute the distances between A and B, $D(A,B) = [d(n,m)]$;

4) Use dynamic programming algorithm to search for the best matching path between shapes A and B, and use (11) to compute the minimum distance $DT_{\min}(A, B)$.

Furthermore, the flipping of shapes can influence the matching results, so when it is needed, we should fix the query target A, and match A with the test target B and B^*, where B^* is the flipped shape of B. The minimum result from the twice matching is taken as the final result.

V. Experiments

To evaluate the effectiveness and accuracy of the proposed method, experiments are carried out with different databases in this section. In all the following experiments, each shape in a certain data set is a query shape, and the rest of the shapes are target shapes. If the query shape and the target shape are classified correctly, the matching result is right; if not, they are mismatched.

A. Kimia-25 Data Set and Kimia-99 Data Set

Kimia-25 data set [13] and Kimia-99 data set [14] are firstly used to test the retrieval accuracy. Kimia-25 data set is composed of images from 6 different categories with a total of 25, as shown in Fig. 7 and Kimia-99 data set is composed of images from 9 different categories with a total of 99, as shown in Fig. 8. To insure the fairness of comparison, the number of sampling points is set as N=100 during all the experiments. In the experiments of Kimia-25 dataset and Kimia-99 data set, the width of window is set as w=18.

Fig. 7 Images in Kimia-25 database

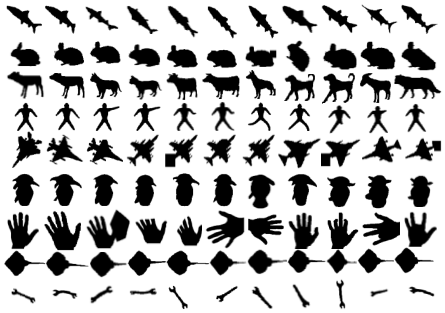


Fig. 8 Images in Kimia-99 database

In the experiment of Kimia-25 database, Δ is set to 1, and the retrieval results are shown in Table I. The results represent the number of the first three shapes that are correctly classified, and the proposed method achieves comparable accuracy.

ISBN: 978-1-61804-327-6

TABLE I. COMPARISON OF RETRIEVAL RATES FOR DIFFERENT ALGORITHMS TESTED ON KIMIA-25 DATABASE

Method	1st	2nd	3rd
Sharvit et. al[13]	23	21	20
Gdalyahu et. al[15]	25	21	19
SC[16]	25	24	22
MDS+SC+DP[12]	23	20	19
IDSC+DP[12]	25	24	25
IS-Match[9]	25	25	24
Our method	**25**	**24**	**24**

In the experiment of Kimia-99 database, Δ is set to 1.25, and the retrieval results are shown in Table II. The results summarize the number of the top 10 shapes that are correctly matched. Compared with the other methods, the proposed method achieves higher classification rates. Synthesizing the results of the two data sets, we get the conclusion that the proposed method performs well when retrieving and recognizing on databases with small size.

TABLE II. COMPARISON OF RETRIEVAL RATES FOR DIFFERENT ALGORITHMS TESTED ON KIMIA-99 DATABASE

Method	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	Total
SC[16]	97	91	88	85	84	77	75	66	56	37	756
Generative Model[17]	99	97	99	98	96	96	94	83	75	48	885
Path Similarity[18]	99	99	99	99	96	97	95	93	89	73	939
TAR[11]	99	99	99	98	98	97	98	95	96	80	956
IDSC+DP[12]	99	99	99	98	98	97	97	98	94	79	958
MDS+SC+DP[12]	99	98	98	98	97	99	97	96	97	85	964
Our method	**99**	**99**	**99**	**99**	**99**	**98**	**96**	**96**	**93**	**87**	**965**

B. MPEG-7 Data Set

To prove the effectiveness of our method on databases with large size, MPEG-7 database [19] is used. The database consists of 1400 images into 70 categories, and each category has 20 different images. The bull's eye score is used in the retrieval test: each shape is compared to all the other shapes and the number of correctly matched shapes among the 40 most similar shapes is recorded, then the corresponding recall ratio is computed.

In this experiment, the width of window is w=14 and the threshold value is Δ=1.1. Besides, the flipping of shapes is considered. The retrieval and recognition results of MPEG-7 database of different methods are listed in Table III. For retrieving and recognizing, our method performs comparably to the listed methods. Through the experiment on three different databases containing both large and small samples, the effectiveness and applicability of the proposed method are demonstrated.

TABLE III. COMPARISON OF RESULTS FOR DIFFERENT ALGORITHMS TESTED ON THE MPEG-7 DATABASE

Method	Retrieval Accuracy (%)
SC[16]	76.51
CSS[6]	75.44
MCC[8]	84.93
Generative Model[17]	80.03
IS-Match[9]	84.79
MDS+SC+DP[12]	84.35
IDSC+DP[12]	85.40
Polygonal Multi-resolution[5]	84.33
Symbolic Representation[20]	85.92
Our method	**86.12**

In recent years, many methods have achieved the retrieval accuracy of over 95% [21, 22] or even 100% [23] on MPEG-7 database. But all these methods use the processing step of metric learning after pair-wise matching. While, the proposed method and all the other contrast methods in this paper mainly focus on the pair-wise matching research, so the results of methods relevant with metric learning are not listed in Table III.

VI. CONCLUSION

In this paper, a novel shape matching method based on spatial features of multi-scaled contours is proposed, which shows better robustness and recognition power in describing the shapes, compared with the triangle-area representation method. Besides, when combined with the dynamic programming method, the feature matching results become more comparable. The merits of our method have been validated by the results of the experiments on different databases.

REFERENCES

[1] D. Zheng and M. Han, "Shape retrieval and recognition based on fuzzy histogram," Journal of Visual Communication and Image Representation, vol. 24, pp. 1009-1019, 2013.

[2] V. Premachandran and R. Kakarala, "Perceptually motivated shape context which uses shape interiors," Pattern Recognition, vol. 46, pp. 2092-2102, 2013.

[3] I. C. Paula, F. N. Medeiros, and F. N. Bezerra, "Shape retrieval by corners and Dynamic Space Warping," 18th International Conference on Digital Signal Processing (DSP), pp. 1-6, 2013.

[4] R.-X. Hu, W. Jia, H. Ling, Y. Zhao, and J. Gui, "Angular pattern and binary angular pattern for shape retrieval," IEEE Transactions on Image Processing, vol. 23, pp. 1118-1127, 2014.

[5] E. Attalla and P. Siy, "Robust shape similarity retrieval based on contour segmentation polygonal multiresolution and elastic matching," Pattern Recognition, vol. 38, pp. 2229-2241, 2005.

[6] F. Mokhtarian, S. Abbasi, and J. Kittler, "Efficient and robust retrieval by shape content through curvature scale space," Series on Software Engineering and Knowledge Engineering, vol. 8, pp. 51-58, 1997.

[7] L. J. Latecki and R. Lakamper, "Shape similarity measure based on correspondence of visual parts," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, pp. 1185-1190, 2000.

[8] T. Adamek and N. E. O'Connor, "A multiscale representation method for nonrigid shapes with a single closed contour," IEEE Transactions on Circuits and Systems for Video Technology, vol. 14, pp. 742-753, 2004.

[9] A. Temlyakov, B. C. Munsell, J. W. Waggoner, and S. Wang, "Two perceptually motivated strategies for shape classification," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2289-2296, 2010.

[10] N. Alajlan, I. El Rube, M. S. Kamel, and G. Freeman, "Shape retrieval using triangle-area representation and dynamic space warping," Pattern Recognition, vol. 40, pp. 1911-1920, 2007.

[11] N. Alajlan, M. S. Kamel, and G. H. Freeman, "Geometry-Based Image Retrieval in Binary Image Databases," IEEE Transactions

on Pattern Analysis and Machine Intelligence, vol. 30, pp. 1003-1013, 2008.

[12] L. Haibin and D. W. Jacobs, "Shape Classification Using the Inner-Distance," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, pp. 286-299, 2007.

[13] D. Sharvit, J. Chan, H. u. Tek, and B. B. Kimia, "Symmetry-Based Indexing of Image Databases," Journal of Visual Communication and Image Representation, vol. 9, pp. 366-380, 1998.

[14] T. B. Sebastian, P. N. Klein, and B. B. Kimia, "Recognition of shapes by editing their shock graphs," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, pp. 550-571, 2004.

[15] Y. Gdalyahu and D. Weinshall, "Flexible syntactic matching of curves and its application to automatic hierarchical classification of silhouettes," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 21, pp. 1312-1328, 1999.

[16] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, pp. 509-522, 2002.

[17] Z. Tu and A. Yuille, "Shape Matching and Recognition – Using Generative Models and Informative Features," T. Pajdla and J. Matas. Computer Vision - ECCV 2004. vol. 3023, Springer Berlin Heidelberg, pp. 195-209, 2004.

[18] B. Xiang and L. J. Latecki, "Path Similarity Skeleton Graph Matching," IEEE Transactions on Pattern Analysis and Machine Intelligence (CVPR), vol. 30, pp. 1282-1292, 2008.

[19] L. J. Latecki, R. Lakamper, and T. Eckhardt, "Shape descriptors for non-rigid shapes with a single closed contour," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), vol.1, pp. 424-429, 2000.

[20] M. R. Daliri and V. Torre, "Robust symbolic representation for shape recognition and retrieval," Pattern Recognition, vol. 41, pp. 1782-1798, 2008.

[21] X. Yang, L. Prasad, and L. J. Latecki, "Affinity learning with diffusion on tensor product graph," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, pp. 28-38, 2013.

[22] R.-X. Hu, W. Jia, Y. Zhao, and J. Gui, "Perceptually motivated morphological strategies for shape retrieval," Pattern Recognition, vol. 45, pp. 3222-3230, 2012.

[23] M. Donoser and H. Bischof, "Diffusion processes for retrieval revisited," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1320-1327, 2013.

Comparative Analysis of Algorithms for Communication Encryption

Milena Karova
"Computer Science and Technology"
Technical University of Varna
Varna, Bulgaria
e-mail: mkarova@ieee.bg

Mariana Todorova
"Automation of manufacturing"
Technical University of Varna
Varna, Bulgaria
e-mail: mgtodorova@tu-varna.bg

Gergana Todorova
master degree student
major "Computer systems and technology (SE)"
Technical University of Varna
Varna, Bulgaria
e-mail: gergana.todorova.ji.90@gmail.com

Ivailo Penev
"Computer Science and Technology"
Technical University of Varna
Varna, Bulgaria
e-mail: ivailo.penev@tu-varna.bg

Ventsislav Nikolov
"Computer Science and Technology"
Technical University of Varna
Varna, Bulgaria
e-mail: v.nikolov@tu-varna.bg

***Abstract*—** A comparative analysis of the algorithms DES, Triple DES-128, Triple DES-192, AES-128, AES-192 and AES-256 is done. For this purpose an application for research of algorithms for cryptographic secure data transmission is created. The application is developed in C#. It offers to the user access to the transmitted data via password. The used resources and time required for encryption of groups of files with different length are studied with the help of the application. The obtained results were compared and analyzed and corresponding conclusions are made.

***Keywords*—** AES, Comparative analysis, Cryptography, Decryption, DES, Encryption, TripleDES.

I. INTRODUCTION

With the fast growth of universal electronic connectivity, people from around the world have a wide access to the information available from different sources. The need for confidential and security information increases. The data and resources must be protected from disclosure and security computer communication systems must be protected from network based attacks.

From ancient times different cryptography algorithms and methods have been employed to encrypt and hide valuable information from eavesdroppers [1]. The algorithms in cryptography are categorized into two classes: Symmetric and Asymmetric algorithms. In symmetric encryption both the sender and receiver share the same secret key [3]. The presented study compares symmetric algorithms. The security level of an encryption algorithm is measured by the size of key space [4]. Large key size means greater security but may decreases encryption or decryption speed. The larger the key space the more possible keys can be constructed. The strength of encryption schemes relies on the key secrecy, length of the key, the initialization string and how they all work together [2].

The main objective of presented work is evaluating the suitability of 6 symmetric popular algorithms. The implementation of algorithms is developed in the software application. It compares performance and resource requirements of encryption of communication and study of different secret key cryptographic algorithms.

The application is designed in C # for use cryptographic secure transmission of data using algorithms DES, TripleDES-128, TripleDES-192, AES-128, AES-192 and AES-256. It offers the user access to password information transmitted. To ensure trouble-free operation is carried out testing of the application and its components. As resources and time required for encryption algorithms groups of files with different length are used.

It needs to define the experimental topics on which it can evaluate and compare encryption schemes. The essential performance parameters are: average time for encrypting/decrypting, length of source file encrypting, size of secret key and resources used during file encryption [5, 6].

The developed application allows comparative analysis of various algorithms for communication encryption on various parameters. The analysis is depicted with a large number of figures, summarized data and tables.

ISBN: 978-1-61804-327-6

II. Development Of Software Application For Research Of Algorithms For Communication Encryption

Two programs – server (Fig. 1) and client (Fig.2) are developed. After starting the server passphrase and "salt" are set. They have to be strings with a length of at least 8 characters. From a security perspective, it is desirable for "salt" to contain uppercase and lowercase letters, numbers and symbols. There are built-in passphrase and "salt" that can be used for quick test of the application.

On the basis of the passphrase and the "salt" are derived a key and an initialization vector to be used for decryption of the received data.

There is a possibility of using algorithms AES (128-bit, 192-bit and 256-bit key), DES and Triple DES (128-bit and 192-bit key). After starting the server expects to receive files with encrypted data, decrypts and stores them in a folder on the disk. The file names, their length and the time for decryption are written in the server's log.

Upon receipt of a connection request from the client, the server starts a thread in which the receiving and decryption of data is done. After completion of the exchange of data, or after a period of idle time (15 seconds), the connection with the client is closed.

The client requires to be entered a passphrase and "salt" or to be used the built-in (for quick test). The method of encryption has to be the same as the one selected in the server.

A. *Encrypted TCP Server*

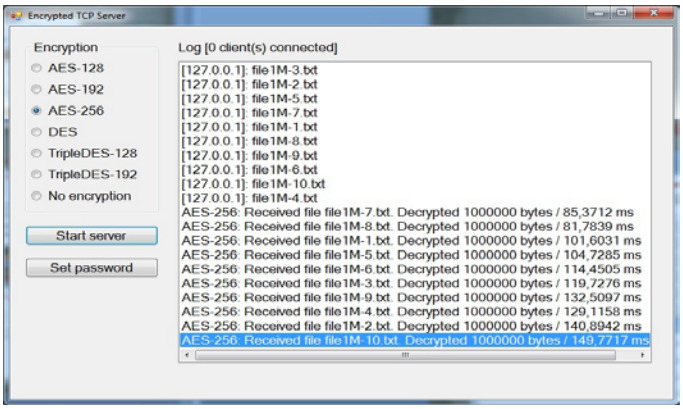


Fig. 1. Server's main window

B. *Encrypted TCP Client*

Pressing the button "Send file" opens a dialog box that allows you to choose one or several files. For each of the selected files a separate thread is started that connects to the server. Then the client sends to the server a block of 256 randomly generated bytes encrypted with the selected encryption algorithm and starts the encryption and transmission of the file. The purpose of the block of 256 randomly generated bytes is to increase security. Thus, even if the same file is transmitted repeatedly, every time the encrypted data will look different. After the transfer of all files, the screen displays a graph showing the time to encrypt each file and the average time for encryption.

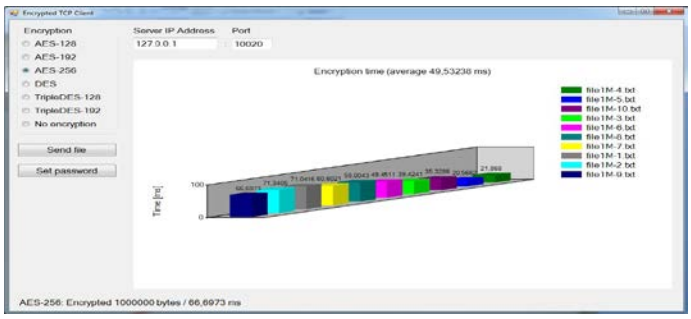


Fig. 2. Client's main window

III. Research Of Algorithms For Communication Encryption

To compare the algorithms groups of files with four different lengths (10,000 bytes; 100,000 bytes; 1,000,000 bytes and 10,000,000 bytes) is done. The tests were performed with minimal use of computer resources from other processes. The encryption is done by using each of the selected algorithms. Some of the obtained results are shown in the figures below.

A. *Encryption time testing*

A comparison of the encryption time of files of the same length using different algorithms is made. The obtained results are shown in Fig. 3 – Fig. 6.

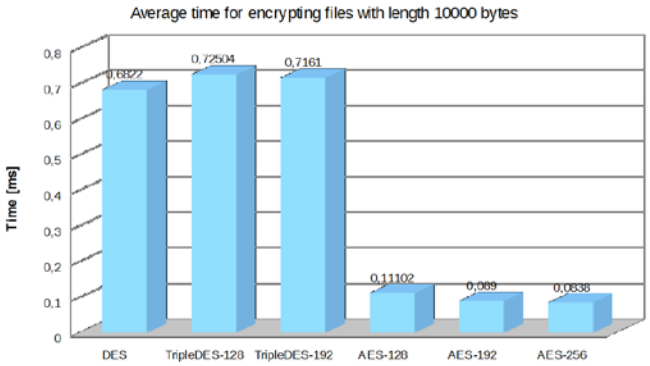


Fig. 3. Average time for encrypting files with length 10,000 bytes

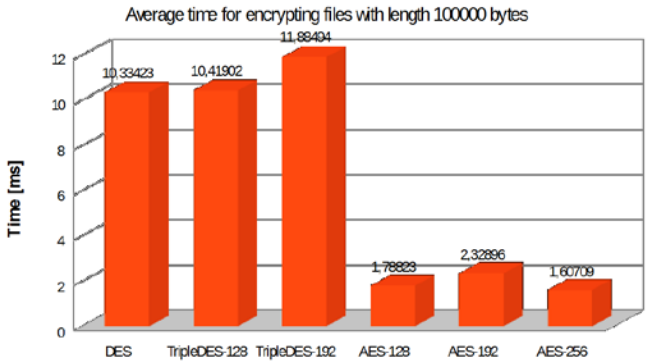


Fig. 4. Average time for encrypting files with length 100,000 bytes

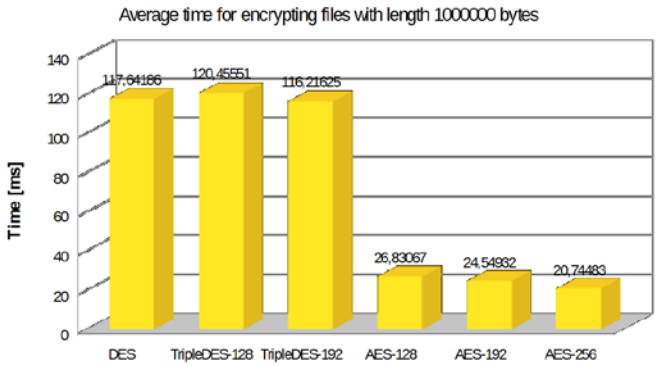


Fig. 5. Average time for encrypting files with length 1,000,000 bytes

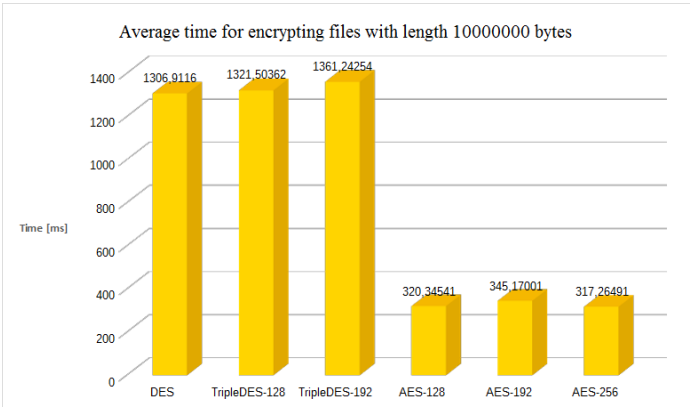


Fig. 6. Average time for encrypting files with length 10,000,000 bytes

The obtained results of encryption of 10K, 100K, 1M and 10M files using the considered algorithms are given on Table I. The ratio of the encryption time $K=t_{max}/t_{min}$ is shown in the last column, where t_{max} – maximum time for encryption, t_{min} – minimum time for encryption.

TABLE I. AVERAGE TIME FOR ENCRYPTION [MS]

	DES	Triple DES-128	Triple DES-192	AES-128	AES-192	AES-256	K
10K	0,6822	0,7250	0,7161	0,1110	0,0890	0,0838	8,6
100K	10,3342	10,4190	11,8849	1,7882	2,3289	1,6071	7,3
1M	117,6419	120,4555	116,2163	26,8307	24,5493	20,7448	5,6
10M	1306,91	1321,50	1396,24	320,34	345,17	317,26	4,4

Note: The least time for encryption is highlighted in red and the highest – in blue.

B. *Used resources*

System resources are monitored during the tests of groups of files with size 10 M. Summarized data achieved by Resource monitor are given below (Fig. 7 – Fig. 14).

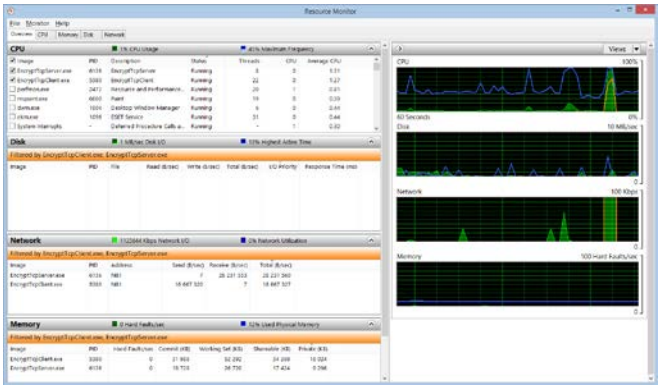


Fig. 7. Summarized data for DES achieved by Resource monitor

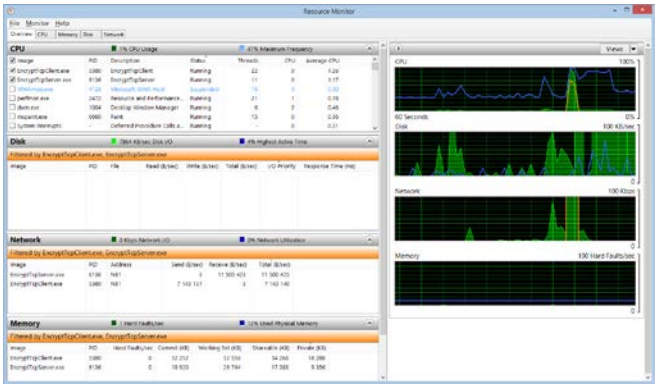


Fig. 8. Summarized data for TripleDES-128 achieved by Resource monitor

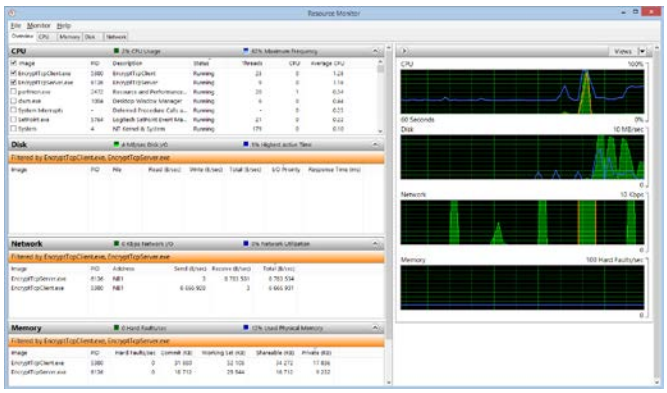


Fig. 9. Summarized data for TripleDES-192 achieved by Resource monitor

ISBN: 978-1-61804-327-6

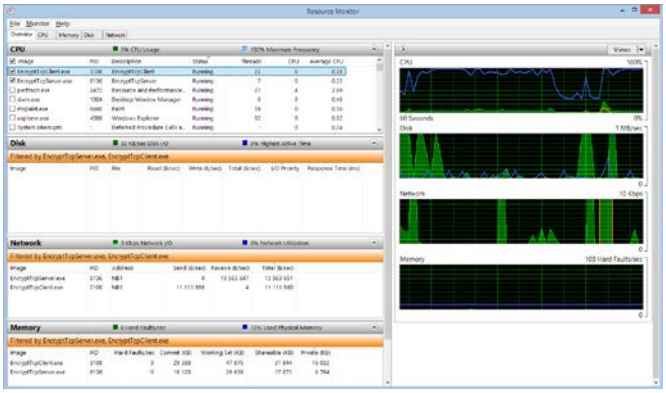


Fig. 10. Summarized data for AES-128 achieved by Resource monitor

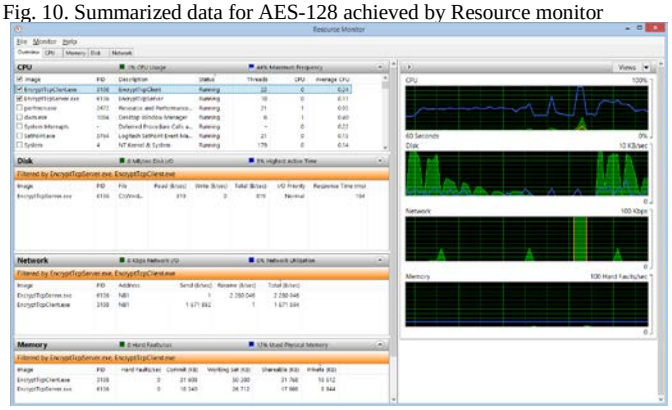


Fig. 11. Summarized data for AES-192 achieved by Resource monitor

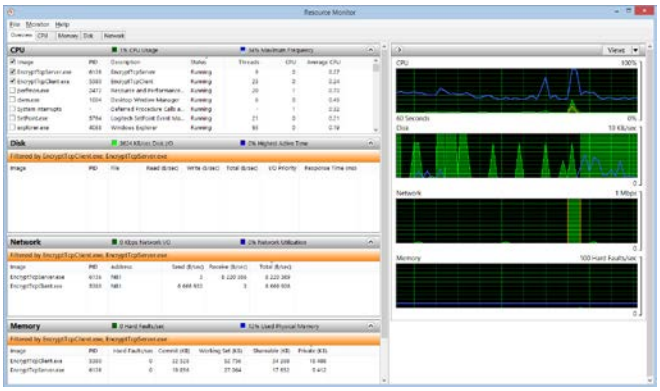


Fig. 12. Summarized data for AES-256 achieved by Resource monitor

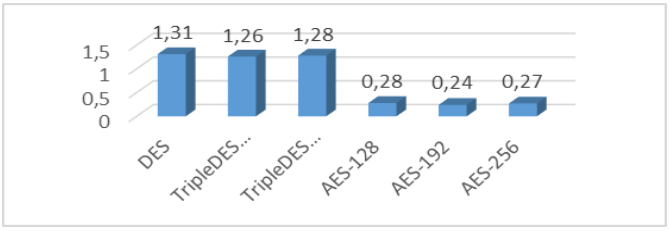


Fig. 13. Average server CPU load

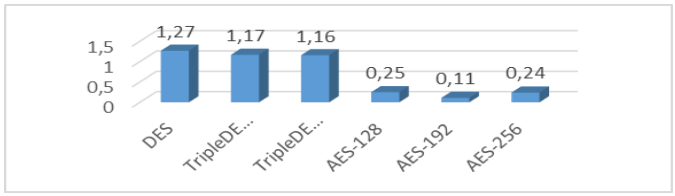


Fig. 14. Average client CPU load

IV. Conclusions And Future Work

The developed application allows for testing and analyzing the considered algorithms. The flexible architecture of the application allows easily add new features that make it even more attractive to users.

In the paper, we aim to practically compare the work of the algorithms mentioned above. We do so by comparing the used resources and time required for encryption of groups of files with different length.

As a result of the tests, the following conclusions can be drawn:

- AES-256 is the fastest encryption method, regardless of the file size. Slightly slower are AES-192 and AES-128.
- Triple DES has low encryption speed. Encryption of files is faster using DES than using Triple DES.
- It makes an impression that with increasing the file size, the speed of AES-256 relatively decreases comparing to the slowest algorithm (Triple DES). Table I shows that:
  - Encryption of small files (10,000 bytes) is performed about 9 times faster using AES-256 than using Triple DES-128;
  - Encryption of files of size 100,000 bytes is carried out about 7 times faster using AES-256 than using Triple DES-192;
  - Encryption of files of size 1,000,000 bytes is performed about 5 times faster using AES-256 than using Triple DES-128;
  - Encryption of large files (10M) is performed about 4 times faster using AES-256 than using Triple DES-192;
  - The average CPU load of the server and the client is significantly lower (5-6 times) when using the AES algorithm as compared to that when using DES and TripleDES. Differences in CPU load when using DES and TripleDES are minor.

The considered problem offers a future scope on work related to:

- Testing with real time voice encryption;
- Data encryption using all methods in parallel;
- Using the application for test of other encryption algorithms.

ISBN: 978-1-61804-327-6

References

[1] Meneses A.J et al., Handbook of Applied Cryptography, CRC Press, 1996.

[2] Manzoor H., Mittal P., Kumar V., A Comparative Study of Cryptographic Algorithms, International Journal of Computer Science and Network, Volume 3, Issue 3, ISSN:2277-5420, 2014.

[3] Christof Paar, Jan Pelzl, Understanding Cryptography, A Textbook for Students and Practitioners, Springer, ISBN 978-3-642-04100-6, 2010.

[4] Henk C. A. van Tilborg, Encyclopedia of cryptography and security, Eindhoven University of Technology The Netherlands, Springer, ISBN-13: (HB) 978-0387-23473-1, 2005.

[5] Pfleeger Charles, Security in computing, ISBN 0-13-185794-0, Prentice Hall PTR Prentice - Hall, Inc, 1997.

[6] Schneier Bruce, Applied Cryptography: Protocols, Algorithms and Source Code in C, Printed in the USA, ISBN 0-471-11709-9, p. 657, 1996.

Studies Regarding the Specificity of the Abrasive Processes

Badea LEPADATESCU, Anisor NEDELCU, Adela-Eliza DUMITRASCU
Department of Manufacturing Engineering
Transilvania University of Brasov
Brasov, Romania
lepadatescu@unitbv.ro, a.nedelcu@unitbv.ro, dumitrascu_a@unitbv.ro

***Abstract*—** In the paper are presented some aspects regarding the specificity of the abrasive processes compared to other machining processes and taking into account the main characteristics of superfinishing work parameters for radial ball bearing parts. Also, in this study are showed some test results regarding the performances obtained in increasing the surface finish, deviations from circularity and waviness after machining the ball bearings through superfinishing process for two types of workpieces.

***Keywords*—**surface finish, abrasive processes, quality of machining, cost reduction

I. INTRODUCTION

In his "Abrasives for engineering purposes" D.Roşca [1] states that "regardless of the abrasive process considered (grinding, honing, superfinishing, etc.), the abrasive grain fixed in a body abrasive tool must have a relative movement with respect to the workpiece surface to be processed". As a result, we find the characteristic elements of abrasion process:

- a large number of cutting tools with undefined geometry consisting of a lot of abrasive grains acting simultaneously or successively;
- relative movement between the abrasive grains and the workpiece.
- Researchers have tried to study the abrasion process comparing it to the machining process with defined cutting tool geometry:
- for grinding were tried to find similarities to milling, each abrasive grain being considered as cutter tooth [2] (grinding wheel is a rotating body as milling, but possessing a small number of edges bordering abrasive grains embedded in the body);
- for superfinishing was assimilated broaching, each abrasive grain corresponding to a tooth of the broach [3], abrasive bar having a similar movement to broach, each abrasive grain moving on the trajectory on a previous grain.

But today, specialists in the field believe that abrasive process differs fundamentally from the other machining processes with defined tool cutting edges. One of the differences between abrasive process and other machining technology is the phenomenon of self-sharpening of abrasive body.

As the abrasive grit blades become dull and blunt, cutting forces grow properly and the grains are torn out from the body tool. If the cutting tool (grinding body) has a low hardness, abrasive grains are easily plucked from the body tool and gradually and continually fresh grains appears with sharp edges and abrasive process continues normally. If the cutting tool has a hardness too high, dull abrasive grains are retained for a longer time so that the abrasive grain gradually lose its capacity of cutting or complete clogging and the workpiece surface finish will have a poor quality.

Between other limitations from abrasive processes and milling or broaching process we can mention:

- Cutting edges of the abrasive grains located on the periphery of the abrasive body have an irregular layout and are not included, in general, to the same initial surface;
- The geometry of the abrasive grain is variable and varies from grain to grain, being possible cutting rake angles both positive and negative, within a very broad field;
- Cutting edges have irregular shapes and sizes.
- Grain sizes range in a great domain, from 10 - 1400μm;
- Cutting fluid penetration is difficult in the area of contact between the abrasive grain and workpiece surface because a small size of the contact surface between the abrasive body and workpiece ,a high speeds under which the cutting process and centrifugal forces "throws out" of the cutting fluid, an "obstacle" which has to overcome the cutting fluid to one point at the periphery of the workpiece.

In superfinishing, it can be stated [5] that the rake angle of the abrasive grain can turn into a relief angle after a reciprocate movement of the tool during machining process (Fig.1).

ISBN: 978-1-61804-327-6

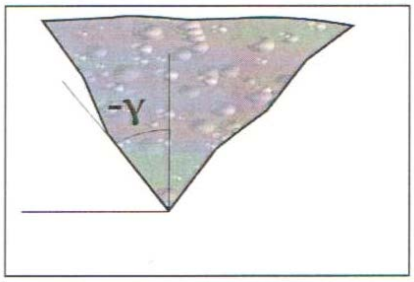


Fig. 1. Transition of the rake angle into relief angle

In the early processing, abrasive grain has sharp edges, but shortly grain radius increases due to wear and wear facet occurs which results in decreased of cutting productivity (Fig.2).

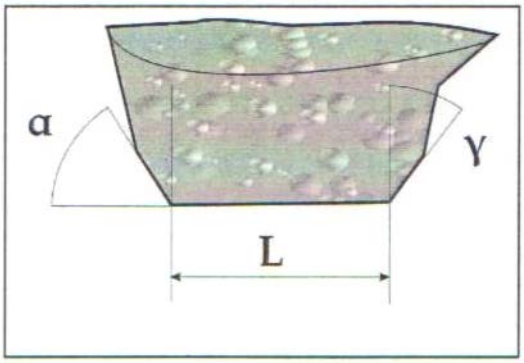


Fig. 2. The wear facet of an abrasive grain

During abrasion process is observed a phenomenon of lateral deformation of the workpiece surface, due to the penetration of the abrasive grain in the part body (Fig.3). These bumps may be expelled by the action of next abrasive grains.

Studying contact between the abrasive grains and the workpiece surface (Fig.4) is observed that there are three areas: the elastic zone I, the plastic zone II, and the cutting zone III. Research [3] has shown that the intensity of the lateral deformation has different values in the three areas, and the maximum value was obtained for the area II.

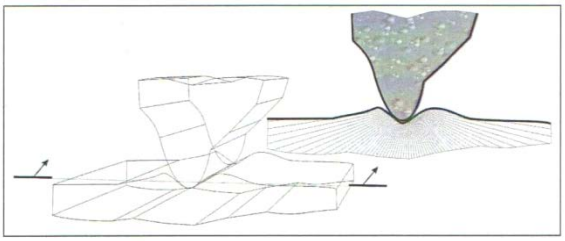


Fig. 3. The phenomenon of lateral deformation

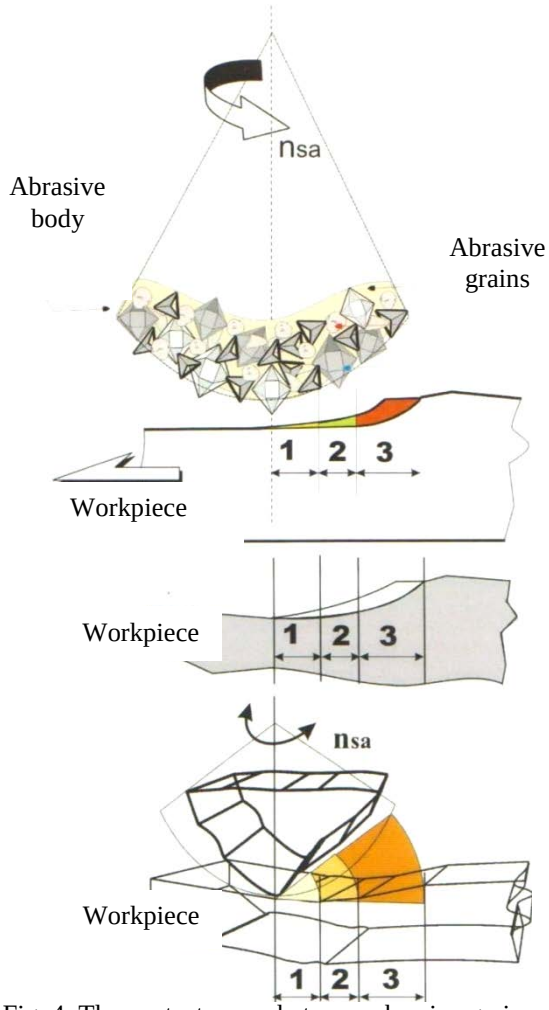


Fig. 4. The contact zones between abrasive grains and workpiece surface.

Cutting heat diffused into the workpiece during abrasion process has, in addition to dimensional changes that should not be neglected, the following effects on the workpiece surface: internal stresses that remain, changing the surface hardness, the formation of cracks on the surface (Fig.5), changing the structure of the surface layer piece rectified (Fig.6).

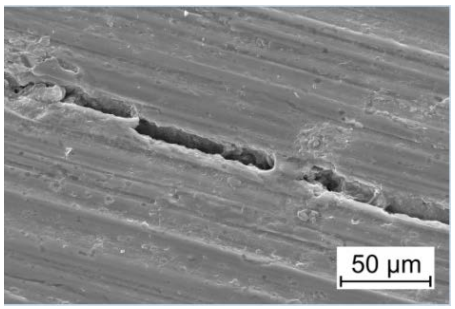


Fig. 5. Cracks that appear during abrasion process.

ISBN: 978-1-61804-327-6

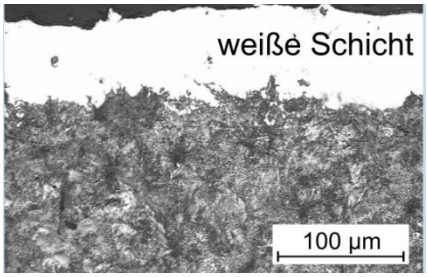


Fig. 6. Changing the structure after abrasion process.

II. Test Results Regarding the Superfinishing of the Ball Bearing Ring Groove

During the tests was used a superfinishing machine KM 150 which has two working stations for roughing and finishing machining, an oscillation frequency of 1200 double strokes per minute and a force of abrasive stone on the workpiece surface of 50 daN.

The principle of the superfinishing process of the ball bearing ring groove is shown in Fig.7.

To highlight the relationship between surface finish of the radial ball bearing ring grooves after superfinishing process, were measured the deviations of circularity, waviness and roughness for two groups of workpieces, 6209-10 and 6312-10 respectively.

It were used two abrasive tool, for roughing EK1 600-08-100 VKH S-Atlantic and for finishing SC 9 1000-1-65- VUB – Atlantic.

To exclude the influence of the blank material on the quality of superfinished part, measured workpieces were from a lot of parts processed on the same machine tool under the same conditions, previous operations (simultaneous surface grinding, centerless grinding, and grinding ring groove).

Measurement deviation from roundness and waviness was done with a device type Talyrond and roughness measurement was performed with a Talysurf device type.

Dispersions values deviation from circularity, waviness and roughness Ra were performed using SPSS program and are shown in the Figs. 8-10.

Fig. 7. The principle of the superfinishing process.

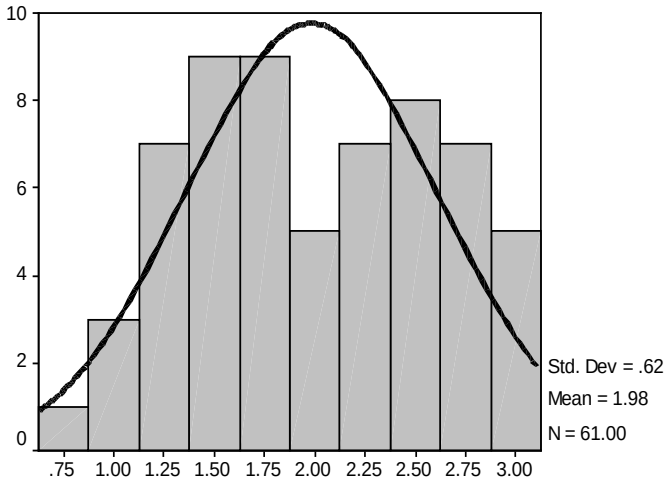


Fig. 8. Deviation from circularity (Piece nr. 6209-10).

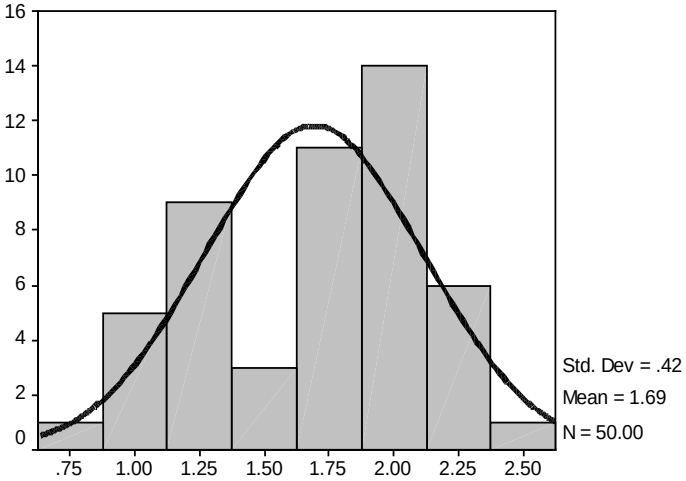


Fig. 9. Deviations from circularity (Piece nr. 6312-10).

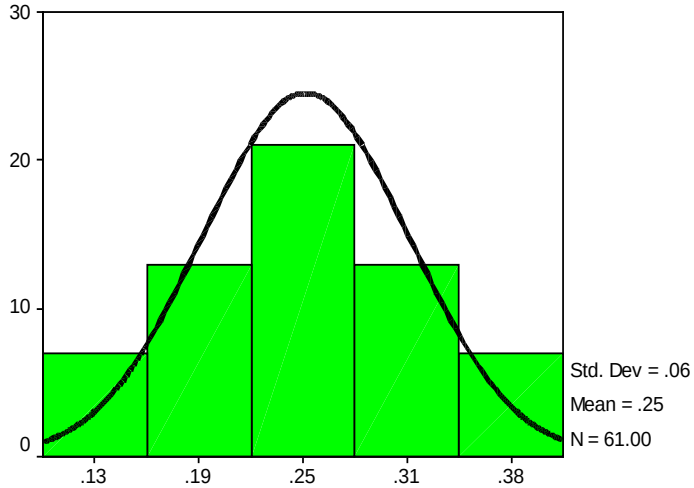


Fig. 10. deviations from waviness (piece nr. 6209-10)

ISBN: 978-1-61804-327-6

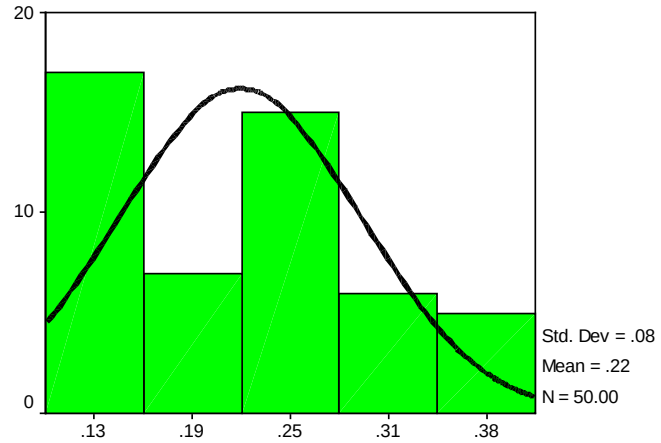


Fig. 11. Deviation from waviness (piece nr. 6312-10).

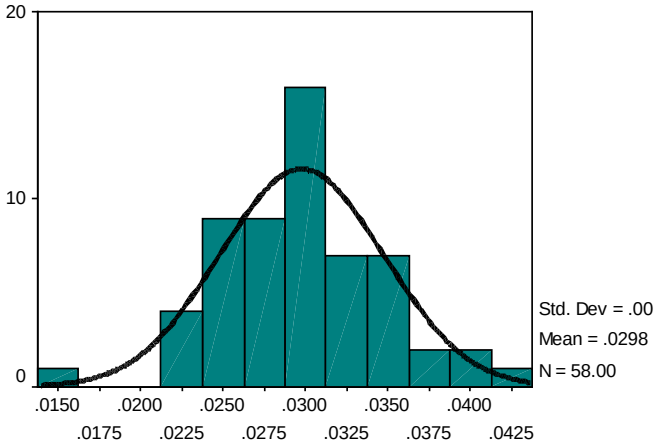


Fig. 12. Deviation from roughness (piece nr. 6209-10).

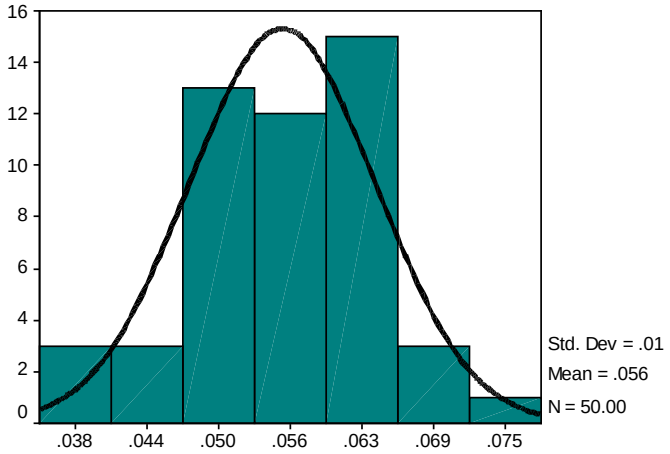


Fig. 13. Deviations from roughness (piece nr. 6312-10)

Analysis of correlation between the workpieces characteristics after superfinishing process show that there is a significant correlation between waviness and deviation of roundness, but there is no correlation between roughness and parameters characterizing the shape of the piece (Tables 1-2).

TABLE I. CORRELATIONS BETWEEN WORKPIECE PARAMETERS AFTER SUPERFINISHING (PIECE NR. 6209-10).

		N	Correlation	Sig.
Pair 1	AB.CIRC2 & OND2	61	0295	0.021
Pair 2	AB.CIRC2 & RA2	58	0.040	0.763
Pair 3	OND2 & RA2	58	0.027	0.838

TABLE II. CORRELATIONS BETWEEN WORKPIECE PARAMETERS AFTER SUPERFINISHING (PIECE NR. 6312-10).

		N	Correlation	Sig.
Pair 1	AB.CIRC2 & OND2	50	0.1	0.992
Pair 2	AB.CIRC2 & RA2	50	0.135	0.348
Pair 3	OND2 & RA2	50	-0.245	0.086

TABLE III. DESCRIPTIVE STATISTICS

Parameter	Characteristics' parameters	6209-10	6312-10
Deviation from circularity of ball bearing ring groove	N	61	50
	The arithmetic mean	1.98	1.69
	Standard deviation	0.62	0.42
Waviness of ball bearing ring groove	N	61	50
	The arithmetic mean	0.06	0.08
	Standard deviation	0.25	0.22
Roughness average of ball bearing ring groove	N	58	50
	The arithmetic mean	0.0298	0.056
	Standard deviation	0	0.01

In Table 3 were presented characteristic parameters of dispersions deviation from circularity, waviness and roughness for the two groups of pieces with numbers 6209-10 and 6312-10 with the outer diameter of 95 mm and 120 mm respectively.

It appears that there are not big differences between arithmetic mean and standard deviations for circularity deviation, waviness and roughness Ra average of the two groups of rings.

III. CONCLUSIONS

- Using the data processing program of data SPSS were drawn frequency diagrams for surfaces of the bearing rings after superfinishing process – deviation from circularity, waviness and roughness average of two groups of workpieces 6209-10 and 6312-10
- The characteristic sizes of descriptive statistics (arithmetic mean, standard deviation) do not differ in a much from those two groups of workpieces.
- The correlation tables show that there is a significant correlation between the deviation of circularity and waviness. One explanation would be that the waviness and deviation from circularity are obtained by

ISBN: 978-1-61804-327-6

"smoothing" of measurement results forming the real profile of the surface by applying different filters.

- Tables of correlations showed that there is not a significant correlation between the waviness and deviation of circularity. This shows that, in the circumstances of superfinishing, the rings are scratched of loose abrasive grains of abrasive bars and unfiltered properly by individual system of machine tool used. Scratches cause an increase in roughness average. Another limitation of the technological process is that there is no perfect control of radius of the ring groove obtained through grinding. The deviations from the correct form of the radius of the ring groove lead to the formation of "bumps" on this surface. Superfinishing stone cannot penetrate into all these areas, and therefore the area is insufficient superfinished and thus will have a higher waviness.

References

[1] D. M. Rosca, Abrasive tools used in machining, LUX LIBRIS Publishing House, Brasov, 1995.

[2] M. Cozmanca, S. Panait, and C. Constantinescu, Basics of cutting, „Gh. Asachi" Publishing House, Iasi, 1995.

[3] C. Buzatu, A. Nedelcu, I. Piukovici, and B. Lepădătescu, Working smoothing surfaces in mechanical engineering, LUX LIBRIS Publishing House, Brasov, 1998.

[4] G. Draghici, Manufacturing engineering, Technical Publishing House, Bucuresti, 1983.

[5] I. Sarbu, O. Pruteanu, and C. Carausu, "Static and dynamic stability study in superfinishing process," Paper presented to TMCR ´99, Iasi, 1999.

[6] C. Picos, Manufacturing engineering, Didactic and Pedagogic Publishing House, Bucuresti, 1974.

[7] J. Koehler, Randzonenbeeinflussung und Prozessueberwachung beim Aussenrundschleifen, Schaeffler Technologieplattform Schleifen, May 20, 2008, Bucuresti.

Software-Hardware Complex For Drill Core Scanning

Dolgy K., Belashev B., Gorkovets V.

***Abstract*—**A software-hardware complex, which forms a core scan from the images of individual sites, had been developed. The structure, circuit design and data processing algorithms, which provide the functioning of the complex are described. Designed complex within a reasonable time with a high space resolution forms the lateral surface of the core samples, has a small size, weight and consumption of energy. Scan of the core sample contains all the visual information about the core sample, reduces the risk of loss, mixing, chemical decomposition of the samples. It may duplicated and does not require the cost of storage and shipment. 3D-model of the scan of core sample allows submit a sample in a conventional three-dimensional form. The database designed for storage of scans of core samples uses the traditional core samples storage base scheme with fast sampling for comparison of data.

***Keywords* —** core samples, 3D-model, scan, software-hardware complex.

I. Introduction

Data on the character, composition, texture, fracturing, mode of occurrence and relative position of rocks are obtained by studying a core – a monolithic cylindrical column extracted from a borehole while drilling [6]. To study the core matter, some core samples are cut, one of its portions is crushed and its substance is studied. The rest of the core is put into boxes and placed in a core storage, where it is described and analyzed. As the core occupies a large volume and the physical condition and informative value of individual samples are lost with time, the core is reduced later. If a core is stored as described, loss can be avoided by obtaining a core scan – a flat image of its lateral surface. As a digital model of a core, a scan stores information, increases the speed and quality of analysis, reduces the risks of loss, mixing and chemical decomposition of samples. It can be duplicated and requires no storage and mailing expenses. The goal of the project is to develop a device for obtaining a core scan. A software-hardware complex, which forms a core scan from the images of individual sites, is presented. The structure, circuit design and data processing algorithms, which provide the functioning of the complex, are described.

II. STRUCTURE AND FUNCTIONS OF COMPLEX

The complex (Fig.1) consists of an executive device, control and image processing units and the database of scan storage and additional information. The executive organ of the complex was designed by selecting immobile and mobile constituents. A rotating sample and a video camera, which moves in step-by-step manner, were used and a core was photographed the way it is commonly done in geological prospecting [6]. The complex, composed of an immobile sample and video cameras mounted on its sides, is simpler in design. The models of the devices designed for both approaches are shown in Figure 2. A sample is supported and rotated (Fig.2a) by guides equipped with friction linings. One of the guides is immobile, while another is rotated by a reducer with gears and belt transmission controlled by a step motor.

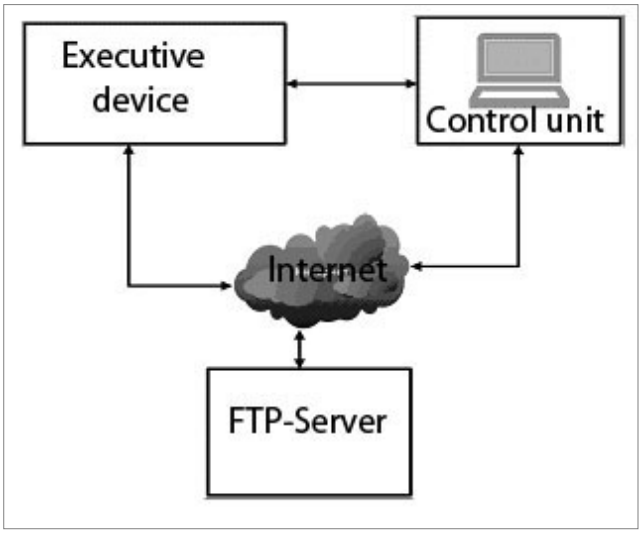


Fig.1. Structural scheme of a complex for scanning the lateral core surface.

The video camera, controlled by another step motor, moves intermittently along the guides located above the sample and gives the images of the sites below it. In the model shown in Figure 2b, one end of the core sample is rigidly fixed. A ring with four video cameras located on its sides is a mobile element coaxial to the sample. The operation of the executive organ is controlled by a mini-computer which works according to the program or is remote-controlled by the user based on SSH-protocol. In the model shown in Figure 2a, the camera, which moves along the sample, scans it and on reaching the end, returns to its original position. The sample rotates by the preset angle, and the process is repeated until the core makes a complete revolution. If the sample is immobile, the cameras scan the lateral surface sites of the core simultaneously, and the controller controls only the movement of the ring movement. On completion of the work, the images are sent to either a FTP-server or a computer, where they are “glued together” are processed and are

This work was supported in part by the U.S. Department of Commerce under Grant BS123456.

K. A. Dolgy is with the Petrozavodsk State University 185910 Petrozavodsk, Russia (corresponding author to provide phone: ; fax: ; e-mail: dolkons@ gmail.com).

B. Z. Belashev is with Institute of geology of Karelian Research Center of Russian Academy of Sciences and with the Petrozavodsk State University 185910 Petrozavodsk, Russia (e-mail: belashev@ krc.karelia.ru).

V. Ya. Gorkovets is with the Institute of geology of Karelian Research Center of Russian Academy of Sciences (e-mail: gorkovets@krc.karelia.ru).

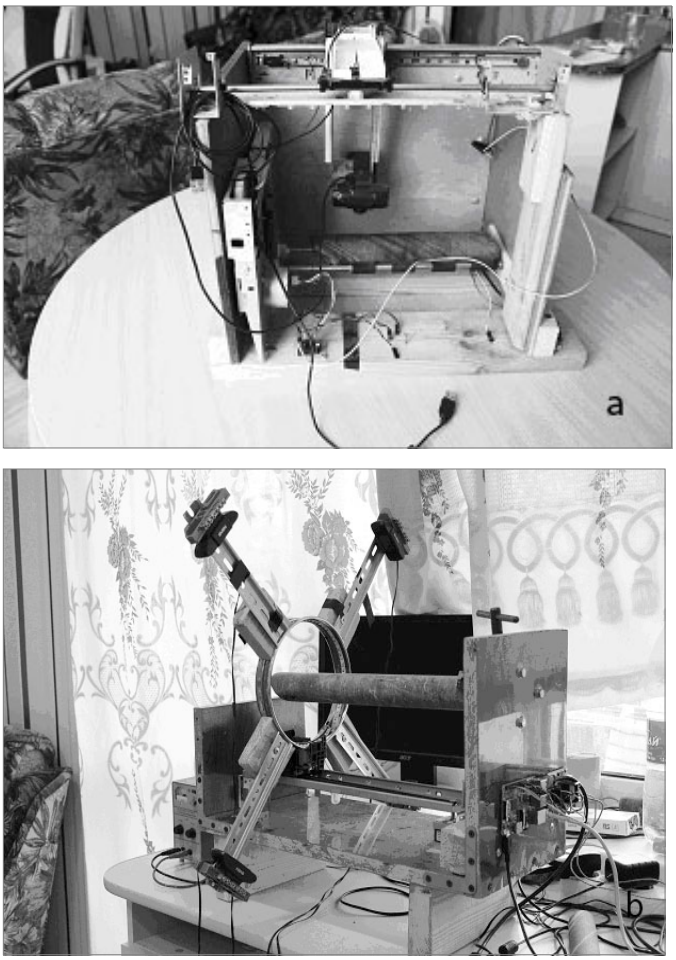


Fig.2 Models for scanning a core with rotating "(a)" and immobile sample "(b)".

used to construct a 3D – model of core scan. In the course of processing, the video images were corrected to form a scan and a 3D-model showing the core in the normal form, the mineral composition of the samples was estimated, and the analytical results were either kept in the computer database or mailed at preset addresses. Distortions, caused by photographing cylindrical samples and varying illumination, were removed by correction. A scan of the lateral surface of the core was obtained by "gluing together" site images in certain order. A 3D-model of the lateral surface of the sample was formed by transferring the scan to the virtual cylindrical surface. The mineral composition of the samples was assessed from the areas covered by the minerals on the core scan. A relational database, imitating the traditional core storage system, which has divisions corresponding to site numbers, borehole numbers and the depth of occurrence of the samples, was used for storing samples scans. Data regarding the date of scanning and the assess of the mineral composition of the sample are entered together with the scan of sample.

III. OPERATIONAL CHARACTERISTICS OF COMPLEX

The models of the complex were tested to assess their characteristics and technical parameters. The advantage of the device with a rotating core is stable illumination and other factors that simplify the obtaining and processing of images. The power and time consumption of the model depends on the rotation of massive samples, the use of several motors and the duration of transition processes. The power consumed by the model is 27 W, and the time taken by the production of a scan of a 0.5 m long sample is about 20 min. A model with an immobile sample and a moving ring with cameras controlled by one step motor is more simple and economic. In this version, power and time consumption is reduced because the core does not rotate and the cameras scan its lateral surface sites simultaneously. The consumed power of the device so designed is 16 W, and the time taken by scanning a sample, about 1 m in length, is one minute. One disadvantage of this method is that samples are changed manually. A Raspberry PI mini-computer (model B+ manufactured by Raspberry PI Foundation) with 4 USB 2.0 connectors, an Ethernet connector, 512Mб OZU and 40 user GPIO outputs was used as a controller of the control unit. The minicomputer was supplied with voltage from the USB – port of the computer or from the supply unit with a rating of 5V. The maximum allowable consumption current is 2.5A. The mini-computer is connected via the GPIO ports to the drivers which control the quantized step motors L293D [4]. Figure.3 shows the printed-circuit board of the computer (a) and the microelectronic circuit of the quantized step motor driver L293D with the connection circuit (b). As Raspberry PI is as small as a banking card, the control unit was made smaller than its counterparts [7]. Connection between Raspberry PI and the analytical unit and the aposterior processing of the image is maintained via Ethernet.

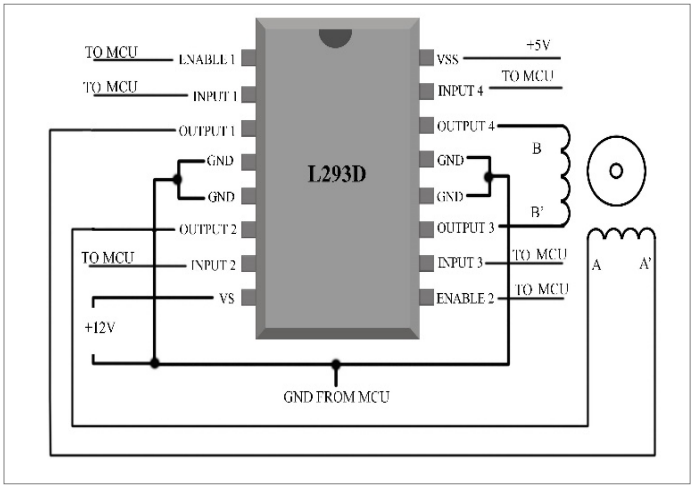


Fig.3. The microelectronic circuit of the quantized step motor driver L293D with the connection circuit.

The blocks of the complex are connected within the Ethernet network and have access to the Internet. The models are controlled by scripts which use the procedures of the

процедуры Raspbian operation system on the basis of the official distributor Debian Linux and a Python language interpreter, version 2.7 [3]. To form a scan of samples from their images, the utility fgwebcam, called from command line bash of the Linux shell, was used. To construct a 3D-model of the core, the algorithm [5] was modified. To control the model, the free database control system MySQL by the Oracle Company was employed [1].

IV. RESULTS

The core samples of Kostomuksha iron ore deposite were used for testing the functions of the designed complex. Core samples rhyodacites, gneiss with quartz veins, ferruginous quartzite, their scans and 3D - models of scans are shown in Fig. 4. The fact that the core scan of ferruginous quartzites consist from many separate images, reflects the way it was received using the variant of executive device with rotating sample. The mineral compositions of the samples were assessed by segmenting scan pixels in color space and estimating the area of the segments. The processing of a rapakivi granite image by the K-mean method [2], used to identify the pixel clusters of K-feldspar, quartz and oligoclase is shown in Fig. 5. From figure it seen that oligoclase crystallized after K-feldspar and quartz occupies intermediate space between their grains.

V. CONCLUSIONS

The software - hardware complex is intended for automatical registration of the scans of core samples. Storing all the visual information about the core sample, scanning removes the disadvantages associated with its storage and carriage of core samples. 3D-model of the core scan allows us to describe and explore the core sample in the usual way. Distinctive features of the complex are low power consumption, simplicity, small size and weight, effective control algorithms and data processing. The ability to evaluate the mineral composition of the rocks on scans of core samples allows you to obtain detailed information on the results of drilling. However, it is clear that for geologists working with a core samples, visual information is incomplete. For reliable identification of rocks and minerals they need a variety of characteristics of rocks, such as for example hardness, magnetic and others features. For this reason it is not yet possible to abandon the cutting, crushing and other ways destroy the core samples. Nevertheless, we believe that computer's methods have a big potential and could improve the efficiency of technologies of mining and mineral exploration.

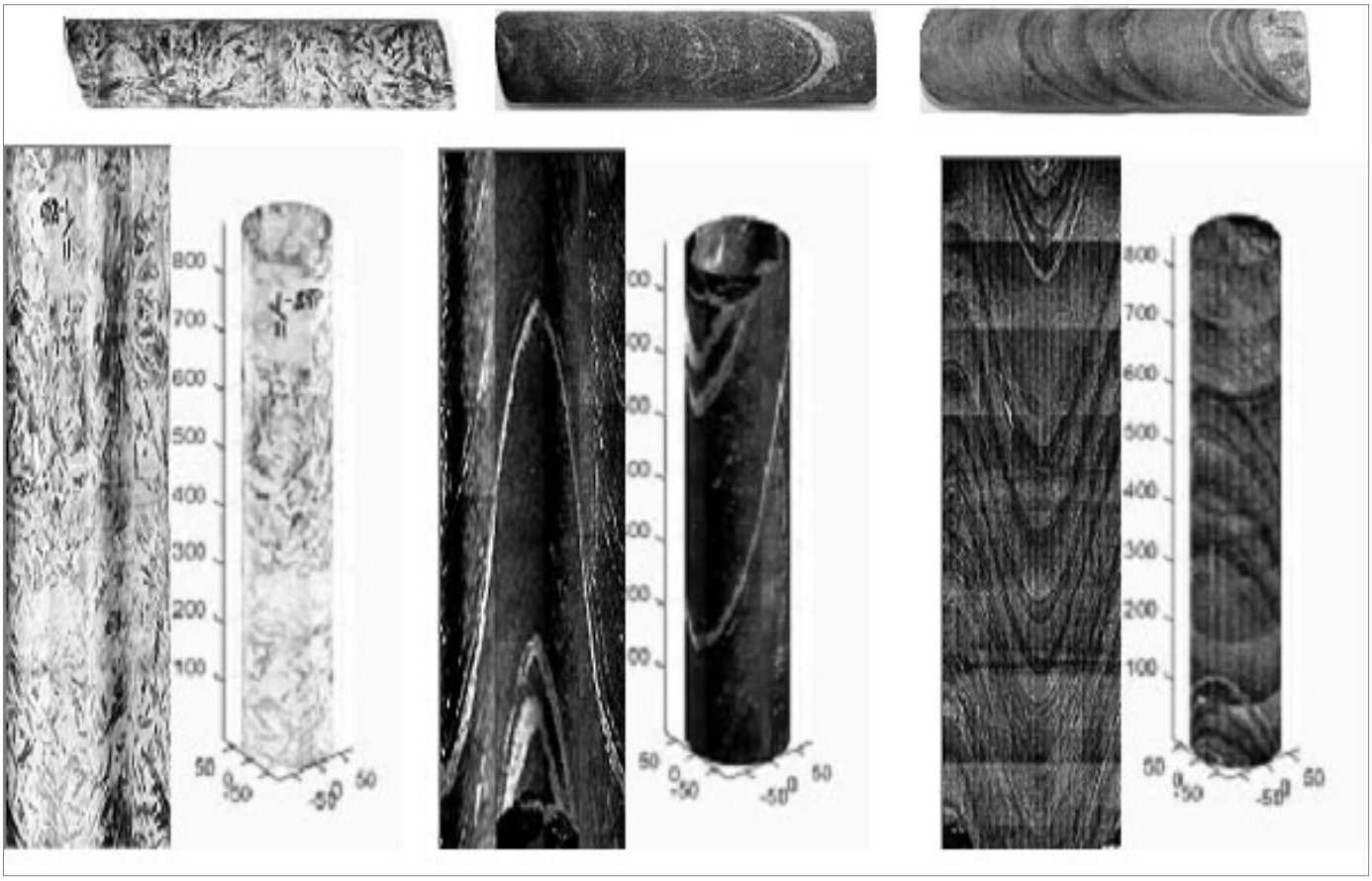


Fig.4. Core samples of rhyodacite, gneiss with quartz veinlets, ferruginous quartzites, their scans and 3D-models of scans

ISBN: 978-1-61804-327-6

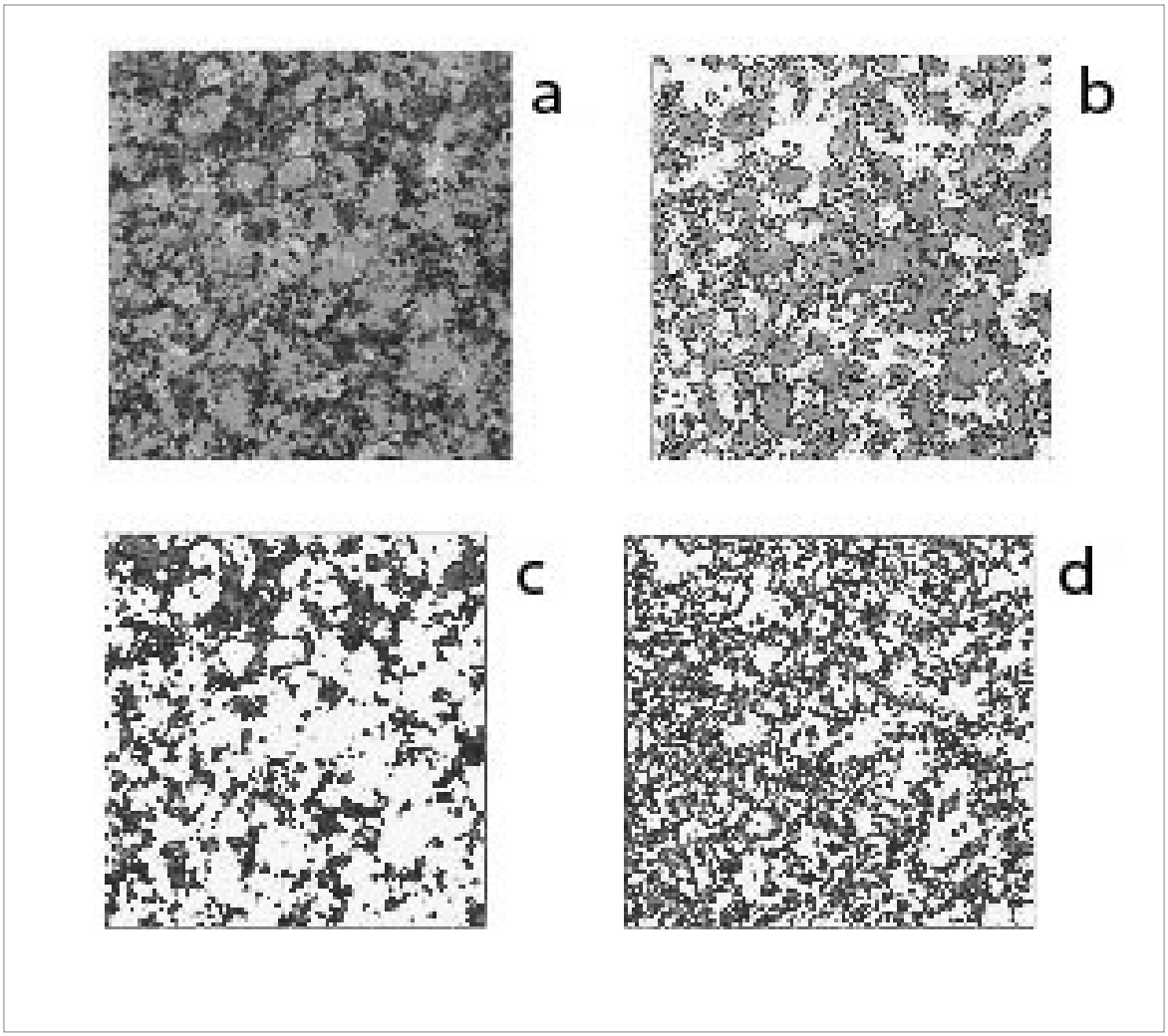


Fig. 5 Rapakivi granite image “(a)” and their clusters: K-feldspar “(b)”, quartz “(c)” and oligoclase “(d)” which cover the areas of 46, 26 and 28 %, respectively.

ACKNOWLEDGMENT

The authors thanks the U.S. Department of Commerce under Grant BS123456 for sponsors and financial support.

REFERENCES:

[1] MySQL Administrator’s Guide. M. Williams Publishers, 2005, pp. 624

[2] Shtovba S.D. Designing indistinct systems by MATLAB. M. Goryachaya liniya-Telecom, 2007, 288 pp.

[3] Marc Lutz. Programming on Python/ Translated from English, 4th edition. SPb: Symbol-Plus, 2011, V.1, 992 pp.

[4] Emelyanov A.V., Shilin A.N. Step motors. Volgograd:VolgGTU, 2005, pp.48

[5] Anufriev I.E., Smirnov A.B., Smirnova E.N. MATLAB 7. SPb: BHV-Petersburg, 2005, 1104 p.

[6] Selling and application of advanced technologies used in geological prospecting in Russia and CIS-countries. AGT-systems and service //agtsys.ru

[7] Raspberry PI Products https://www.raspberrypi.org/

.

Inductive Transmission of Electromagnetic Energy – From M. Faraday to XXI Century

Ph.D eng. ARDELEANU Mircea-Emilian
Electrical Engineering Faculty,
University of Craiova
B-dul Decebal nr.107, 200440 Craiova, Romania
mircea_emilian@yahoo.com

eng. RĂSCĂCEA Bogdan
Electrical Engineering Faculty,
University of Craiova
B-dul Decebal nr.107, 200440 Craiova, Romania

***Abstract*—The discovery of the phenomenon of electromagnetic induction and its scientific justification are cues in the history of physics and electrical engineering in particular. Its practical application through the development of mass production of electrical transformers and motors and generators led to one of the most important economic sectors namely electrotechnical industry.**

Some important biographical highlights of English physicist Michael Faraday's life and his first experiments and conclusions on electromagnetic induction are presented.

The principle of electromagnetic induction was taken over some decades later by the Serbian engineer Nikola Tesla who made a prototype of the induction motor. Also, this article refers to Tesla's brilliant idea on the possible development of wireless transmission of electricity over long distances.

In the end are reminded a few solutions on the inductive transmission of electricity to an electric vehicle recharging or pacemaker

Keywords-electromagnetic induction, resonant circuit, wireless transmision,

I. INTRODUCTION

The discovery of the phenomenon of electromagnetic induction and its scientific justification are milestones in the history of physics and electrical engineering in particular.

This article aims to present a brief history of the discovery of this phenomenon and its use in the service of scientific and technical progress of humanity (making transformers, motors and electric generators, etc.) and current and future uses in the wireless transmission of electricity.

First experiments and conclusions of M. Faraday on electrical induction are presented.

Particular attention is shown to brilliant personality of engineer N. Tesla and his tenacious, ambitious work which by the spectacular achievements and experiments conducted at Wardenclyffe promoted the idea of the possibility of wireless transmission of electricity over long distances.

Some current applications of wireless transmission of electrical energy and important future uses such as wireless power of laptops, charging batteries in electric cars and mobile phones etc. are presented.

In conclusion are presented perspectives and technical and economic advantages of future uses of wireless transmission of electricity both for industrial purposes as well as household uses.

II. ELECTROMAGNETICAL INDUCTION

A. Michael Faraday- father of electrotechnical engineering

On September 22, 1791, in the house of a poor blacksmith from Newington, near London, Michael Faraday saw daylight. He had a difficult childhood, full of indigence. At the age of twelve he is forced to abandon primary school, where he had managed to learn to read and count at all, and works as a bookbinder apprentice. In addition, he read books to the bound, being interested in physics and chemistry books.

After completing his apprenticeship he works for a traveling bookbinder.

Thus, by unknown destiny games, he gets to know the famous Professor of Chemistry, Humphry Davy at the Royal Institution. This impressed with young Faraday hires him in 1813, as laboratory – assistant at the Royal Institution. Perseverance, dedication, thoroughness, diligence, desire to know as many things as possible made young Faraday to be noted not only by Davy but also by other teachers who needed his services.

During 1813-1815 he accompanies the famous professor in a course of lectures in the great European capitals.

Due to its scientific and experimental merit and qualities faculty in 1815 at the Royal Institution unanimously agreed that Michael Faraday to be employed as an university assistant

ISBN: 978-1-61804-327-6

In his brilliant scientific and teaching career that has lasted over thirty years, he carried out many, varied and original works. He developed extensive studies, meticulously prepared, accompanied by convincing experiments in support of the conclusions set [1].

Fig.1 Michael Faraday [2]

On August 29, 1831 Michael Faraday recorded in his workbook notes on the experience that says "that a circuit bathing in a variable magnetic flux, as long as the flow variation lasts, became the seat of electromotive force. If the circuit is closed in this phase of the magnetic flux change, it is crossed by an electric current. If the circuit is open at its terminals there is a potential difference." He called this phenomenon that produces an ***electromotive force*** and an ***induction current, induction***. [1] This day can be recorded as a historic day, a memorable page of physics, can be considered the birth date of discovery of the law of electromagnetic induction and of electromagnetic induction.

Discovering induction is regarded by the entire world as the highest academic achievement of Faraday, the phenomenon that gave a new course to electromagnetism and is the cornerstone of electrical engineering. Therefore, Michael Faraday is considered the ***father of electrotechnical engineering***.

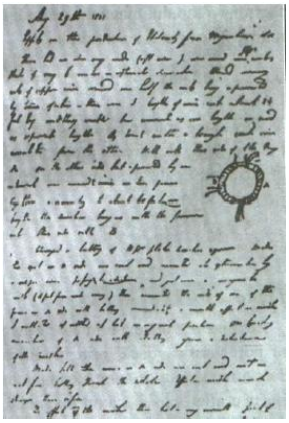


Figure.2 1831 manuscript page in which are shown the results of the "conversion of magnetism into electricity [1]

Scientific and laboratory experiments were doubled by a distinguished teaching activity.

On 25 August 1867 after a simple and modest life dedicated with passion and a strong spiritual dedication to scientific discoveries, simple, clear conclusive experiments, understood by its auditor, although ill health forced at times to stay away from his laboratory, Michael Faraday died.

B. Faraday's experiments

With tenacity, dedication and interest in the phenomena of physics (with preference to electricity and magnetism) coupled with a soul kindness and modesty recognized by colleagues and collaborators, he managed through simple, clear, persuasive experiments, understandable to everyone to present and define electromagnetic induction law, the fundamental law of electrical engineering.

Next we will briefly present Faraday's experiments.

Experiment 1

Consider a circular coil of conductive material on which is mounted a galvanometer (fig.3).

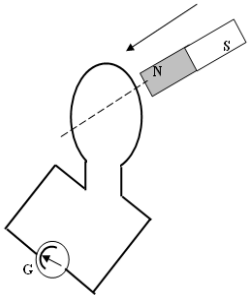


Figure3 Figure explanatory for the first experiment

The coil is approaching a bar shaped magnet with the North Pole toward the coil. As the magnet approaches the coil, galvanometer needle deflects off indicating that an electrical current is flowing through the coil. If the magnet does not move relative to the coil galvanometer does not deviate.

When moving in reverse the magnet (the magnet remove coil) galvanometer needle deflects off but in reverse than before. If you repeat the experience but bringing close and then removing the magnet with South Pole toward coil galvanometer needle deflects off, but the direction of deviation is opposite than in the first experiment.

The conclusion that emerged from this experience extremely simple is that *what matters is the relative displacement of the magnet and the coil.*

It doesn't matter which is the element that is moving, namely the magnet toward the coil or vice versa.

The current that occurs in the coil is called *induction current* being determined by *induced electromotive voltage* [3].

Experiment 2

Another simple demonstration made by Faraday to explain the phenomenon of induction, he used a device similar to that shown in fig.4 .

The experimental device is made up of two circuits. A circuit is made up of a coil with a galvanometer mounted at the ends (G). The second circuit is made up of a coil where

there is a DC voltage source (E), a switch (K) and a resistor (R).

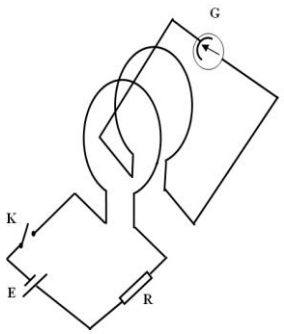


Figure 4 Figure explanatory for the second experiment

The two coils are arranged side by side (face to face) and in the rest one against the other, as shown in fig.4 When the switch K is closed, in the circuit of the other coil is established an electrical current notified by the galvanometer G through its indicator needle deflection. Deflection is short as the needle returns to zero.

If the switch K is open also it can be seen, for a short period of time, the deflection of the galvanometer needle, but in the opposite direction than before.

The experiment showed that in the coil with passive circuit (who does not have power supply) occurs an induced electromotive voltage whenever the current in the coil in the circuit powered by the power supply E ranges (closes or opens the switch K).

The conclusion that was drawn from this experiment is that what is important in this phenomenon is *the variation speed of current and not its intensity* [3].

C. Law of electromagnetic induction

Law of electromagnetic induction is an important and fundamental law to electrical engineering showing how to produce voltage always stating that the presence of a time-varying magnetic field is accompanied by an electric field [4].

The statement of the law of electromagnetic induction states that "u_e electromotive voltage induced in a circuit is equal to variation speed of magnetic flux taken with the opposite sign."

$$u_e = -\frac{d\Phi}{dt} \tag{1}$$

Integral form of the law, both for bodies at rest and bodies in motion has the form:

$$u_e = \oint_{\Gamma} \vec{E} \cdot d\vec{l} = -\int_{\Sigma_\Gamma} \frac{\partial \vec{B}}{\partial t} d\vec{S} + \int_{\Sigma_\Gamma} rot\left(\vec{v} \times \vec{B}\right) d\vec{S} \tag{2}$$

And local form is:

$$rot\vec{E} = -\frac{\partial \vec{B}}{\partial t} + rot\left(\vec{v} \times \vec{B}\right) \tag{3}$$

In relation (2) :

$-\int_{\Sigma_\Gamma} \frac{\partial \vec{B}}{\partial t} d\vec{S}$ - represents the variation in magnetic flux due to local variation of magnetic induction, the contour is assumed stationary and is called electromotive voltage induced by transformation;

$\int_{\Sigma_\Gamma} rot\left(\vec{v} \times \vec{B}\right) d\vec{S}$ - corresponds to the variation of the magnetic flux due to the movement of contours with the body, the magnetic induction is assumed invariable in time and represents electromotive voltage induced by motion.

III. Nikola Tesla and wireless transmission of electricity

A. Nikola Tesla – engineering genius

Nikola Tesla saw the light of day on July 10, 1856 in the small village Smilijan in the province of Lika in Croatia. He was the fourth child of Milutin and Djouka Tesla. He was born and raised in a family educated and devoted to the principles of family and Orthodox faith, his father being an appreciated priest.

Primary School classes begin in Gospic, where his father received a new parish, proving to be a good student who likes to read a lot, even obtaining a job in the school library.

Destiny makes that in Karlovac (Carlstad) where he continued his studies at the village high school he knows physics Martin Sekulic who stimulated and impressed him with the experiments he made for the young students, opening the way to the sciences.

Wanting to become an electrical engineer he enrolls at the renowned Ecole Polytechnique Graz, where he meets famous teachers such as: Rogner, Poeschl, who initiated him into the mysteries of physics and superior mathematics. From financial reasons is forced to abandon studies. He never graduated from Ecole Polytechnique Graz and received no grade for last semester spent there [5].

Fig.5 The young Tesla [2]

Later, in 1880, he manages to participate in some summer courses of the University of Prague, one of the most famous and important university in Europe with famous rector Professor Ernst Mach.

After a spell in Budapest, where he works in the telephone company, and then to Paris where he was employed at the

ISBN: 978-1-61804-327-6

newly founded company Continental Edison Company, in 1884 decides to go to America.

United States of America, a country growing economic, industrial and financial, country where each newcomer saw the Promised Land proved a fertile area for the development and flourishing of the technical ideas ingenious and courageous of young Nikola Tesla.

With the recommendation of the Director of Continental Edison Company in Paris he gets to work for the company of famous Thomas Edison, a personality that will mark much his life and activity.

At one point Tesla said: "The meeting with Edison was a memorable moment of my life. I was amazed by this wonderful man who has achieved so much with no scientific training." [6]

In 1887, he built the first induction motor, brushless, AC power, which he presented at the *American Institute of Electrical Engineers* (now IEEE, Institute of Electrical and Electronics Engineers) in 1888 [2].

As a child he was fascinated by the native water mills and water power to move the mill wheel. Since then cornered in small Nikola mind the idea of using flowing water energy for the production of electricity.

Making the great project for the power plant at Niagara Falls has raised many specialized issues constructively, finding many ingenious technical solutions, but also created great controversy, much publicized at the time, between Thomas Alva Edison and Nikola Tesla, the American and world's electrical engineering titans at the time.

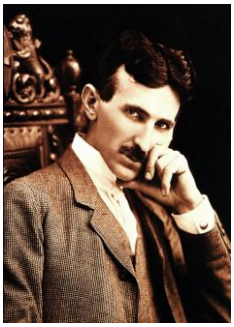


Fig.6 Nikola Tesla at maturity [2]

After five years of great human and technical endeavor, points of renunciation and distrust, financial crisis and millions of dollars invested in November 1886 Niagara Falls hydropower plant project was completed. The first hydropower plant in the world was now functional, the author of this project, greeted by many with distrust and suspicion, considered likely to fail, was Nikola Tesla.

Making the large hydropower plant from Niagara in the last decade of the nineteenth century brings Tesla at no forty years, a great victory and enhances the reputation and respect in the world of engineering. He managed to defeat the favorable trend of DC supported by Edison which was financially supported by tycoon J.P.Morgan requiring alternating current generators.

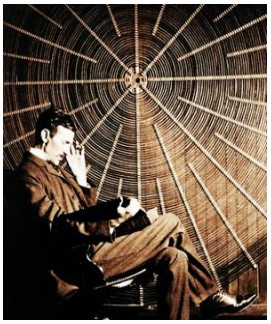


Fig.7 Tesla in his laboratory [2]

B. Tesla's vision on wireless transmission of electricity

For instance, he understood that wireless power transmission is possible, but did not know that there are limits to the amount of energy transmitted in this way.

Towards the end of 1898 he begins a systematic research, undertaken over several years in order to perfect a method for transmitting electricity through the natural environment.

Thus, in 1899 Tesla moved to Colorado Springs. Colorado Springs a plateau located at 2,000 m altitude above sea level and covers an area of 1,000 km^2. Here he designed the new laboratory to achieve his goals.

Atmosphere of Colorado Springs allowed him to make some important scientific observations noting "Consequently, lightning in the atmosphere are very frequent and sometimes incredibly violent. Once occurred about twelve thousand lightning in a span of two hours, all on distance less than fifty miles around the laboratory. Many of them resembled giant fire trees with trunks facing when upwards when downwards. I have never seen ball lightning, but as compensation for my disappointment I succeeded later to determine how they are formed and I could produce them artificially" [6].

Publication of scientific observations made during the period when he worked at his laboratory in Colorado has attracted the attention of many scientists and interests of businessmen. Morgan, initially taken by Tesla's ideas, accepted such a giant project financing by providing the scientist and his team a significant amount of money, namely $ 150,000 to build a transmission tower and a power plant.

Full of energy, thoughts and brilliant ideas, surrounded by a team of collaborators Tesla ventured may be into the most deep, known and publicized scientific adventure – achievement of wireless transmission of electricity.

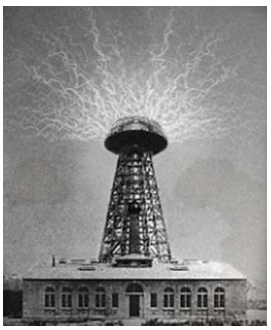


Fig.8 Tesla Wardenclyffe tower

ISBN: 978-1-61804-327-6

To implement his plan Tesla begins the construction of the famous tower Wardenclyffe in Long Island. In September 1904 the tower had reached a maximum height of 55m. Remaining with little money he failed to complete the dome atop the tower. With reduced funding Tesla continued his work and experiments.

J.P.Morgan was concerned, more obsessed with the possibility that Tesla's achievements can provide the transmission of "unlimited amounts of energy" wireless, although Tesla ensured him that the plant at Wardenclyffe can send only "small amounts" of energy.

A close friend of Morgan suggested him "Look, this man has gone crazy. What he does is that he wants to give them all free electricity and you do not have to put counters. We'll go bankrupt if we support this man."

This was the straw that broke the camel's back and prompted tycoon J.P.Morgan to remain deaf to all Tesla's attempts to continue its research and investments.

Lacking funds, preset by the lenders, Nikola Tesla was forced to stop in 1904 the work on Wardenclyffe complex.

Thus due to more pragmatic than visionary thinking on long term of potent American businessmen and bankers was interrupted a brilliant idea that could have well changed the lifestyle of mankind.

Tesla has lost a battle but happily continued to have a brilliant mind and ideas that today amaze the scientific world and are waiting to be applied to peaceful purposes.

IV. Current and Future Uses of Wireless Transmission of Electricity

A. Electric vehicle battery charging system

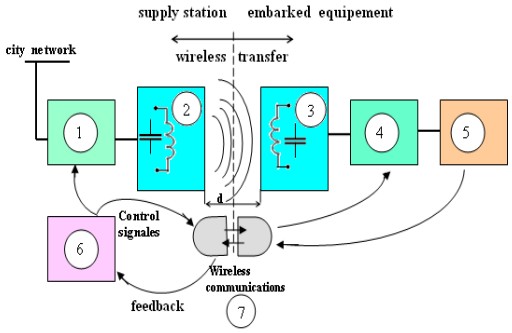


Fig. 9 Block diagram of a system for wireless transfer of energy and data for batterry charging of power electric vehicles

1.) AC power converter (50Hz)-medium/high frequency AC; 2.)transmitter circuit; 3.) receiver circuit; 4.) medium/high frequency AC converter–c.c.; 5.) storage battery; 6.) monitoring and control block;7.) wireless communication

The schematic diagram in fig.9 [7] the energy transfer takes place between the emitter circuit (2) or base station and receiver circuit (3) or mobile station (located on the vehicle) separated by a variable distance (d) of ground clearance time. The transfer is one of the inductive type, such as that described above, the power transmitted to the load (5) can be of the order of kW and provides charging the battery in a time equivalent to the duration of the charging process using a galvanic coupling type.

Variable frequency static converter (1) used in this scheme is adaptive type to ensure, on the one hand, the resonance conditions of the transmitter and receiver circuits, and on the other hand can automatically compensate for variable load caused disagreement. Converter (4) is designed to provide battery charging regardless of its status (degree of discharge) in an optimal regime. Status and battery charging system is monitored and information about them is transmitted through a wireless communication channel (7) to monitor and control block (6) who has command over converters (1) and (4) so that yield global transfer of the system to be maximum regardless of battery status and power factor at the point of power supply to the equal [7].

B.Pacemaker battery charging system.

A heart stimulator (pacemaker) is a medical device of small size (3-4 cm) emitting electrical impulses, transmitted by means of electrodes which are in contact with the heart muscle in order to regulate the heartbeat. The miniaturized electronic device delivers regular rhythm excitations to a heart with physiological centers unable to provide normal heart rhythm.

Implantation of a pacemaker is a minimally invasive surgery under local anesthesia by which the boxy of pacemaker is buried in a place specially prepared between the chest skin and pectoralis major muscle. For power of electronic circuits of this pacemaker, it is fitted with lithium batteries which ensure a smooth operation for several years (5-7 years) [8].

In fig.10 there is a block diagram of a wireless inductive system of pacemakers power battery.

The system consists of an outer and an inner coil. Outer coil is supplied by an external source of AC. Internal coil takes over inductive power transmitted by the external coil and adapts it to the needs imposed by electronic implants.

The system shown in fig.10 [9] is a series-parallel resonator type. The outer member is composed of an inductor L_{ext} and a capacitor connected in series C_s. The inner member (mounted in the body with the pacemaker) is made up of a coil L_{int} and a capacitor mounted in parallel C_p.

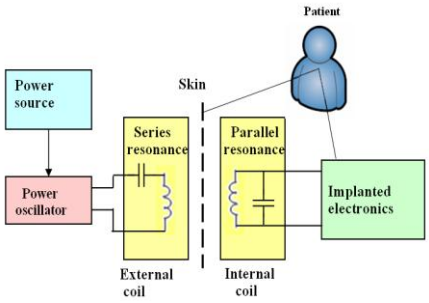


Fig.10 Block diagram of a wireless inductive system of pacemakers power battery.

At resonance is satisfied the relationship known for pulsation:

$$\omega_0 = \frac{1}{\sqrt{L \cdot C}} \quad (4)$$

Optimum working frequency for such a resonator used for medical purposes for cardiac implants is between 5 ÷ 10 MHz for both coil size reduction and compliance with rules imposed by international medical bodies. For use in medical purposes of such implants, an electronic device based on inductive power transmission must release a power of around 10 mW [9].

C. Supply of electronic devices equipped with rechargeable batteries

The explosive development of electronics by implementing new technologies that ensure miniaturization increasingly sharper of circuits and electronic devices has led to a technological revolution. Thus, both production and market of "devices" portable is in permanent mobile content diversification and full expansion. These products such as mobile phones, laptops, tablet, camera, etc. heavily penetrated everyday life, so for many people they have become virtually indispensable.

Use of wired chargers, although widely used and available today becomes thick, unsightly (wires across your office or home, a possible danger of accidents, etc.).

Modern solutions, which are expected in the not too distant future, to supply such low power consumers is the use of wireless transfer of electricity

In fig.11 [10,11] are some aspects which appear to be detached from a science fiction movie, on the use of wireless transmission of electricity to power the equipment of an apartment or office

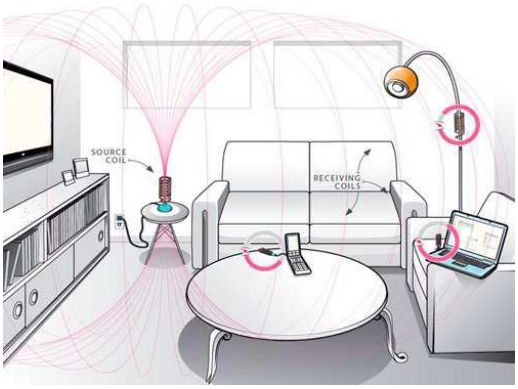


Fig.11 Supply of electronic devices equipped with rechargeable batteries

Modern solutions, which are expected in the not too distant future, to supply such low power consumers is the use of inductive transfer (wireless transfer) of electricity.

V CONCLUSIONS

Discovery and presentation of simple, clear, persuasive experiments and understandable to anyone of the law of electromagnetic induction by Michael Faraday was a moment of crossroad in the evolution of the electromagnetic field theory. Making electrical transformers and motors and generators, whose operation is based on the phenomenon of electromagnetic induction started a major industrial branch, Electrotechnical Engineering Industry.

Nikola Tesla, a brilliant and visionary engineer, took the principle of electromagnetic induction not only to improve design and functionality of electric motors or generators but also to achieve a larger project that foresees the possible transmission of electricity over long distances wirelessly. From financial reasons his experiments in the wireless transmission of electricity could not be completed.

Today, over a century after this attempt of Tesla, the idea acquires new meanings and more and more engineers and researchers lean forward towards realizing it.

Modern technologies allow for electrical equipment to perform wireless voltage supply of portable devices such as mobile phones, tablets, laptops or electric road vehicles.

This paper presents several such proposals.

Future achievements in the field of wireless transmission of electricity will materialize Tesla's unfulfilled dream.

REFERENCES

[1] Florea,I.C., *Trei savanți iluștri*, Editura științifică și enciclopedică, București, 1978
[2] www.wikipedia
[3] Holliday,D.,Resnic,R., *Fizica (vol.II)*,Editura didactică și pedagogică, București,1975
[4] Sora,C., *Bazele electrotehnicii*, Editura didactică și pedagogică, București, 1982
[5] Seifer,M., *Tesla, bigrafia unui geniu*, Editura ProEditură și Tipografie, Bucuresti,2008
[6] Tesla,N.,Childress,D.H., *Fantasticele inventii ale lui NIKOLA TESLA*, Editura Vidia, Bucuresti,2011
[7] Marinescu,A., Iordache¸M., Mandache,L., Transferul fără contact al energiei electrice, a VIII-a Conferință anuală a ASTR, Bucuresti, 2013
[8] www.sfatul medicului.ro
[9] Qingyun,M.,s.a., Power-oscillator Based High Efficiency Inductive Power-Link for transcutanous Transmission
[10] Georgiana Nicoleta ZAINEA, Contribuții privind transferul wireless a energiei electromagnetice,teza de doctorat, Universitatea politehnica Bucuresti,Bucuresti, 2013
[11] http://phys.org/news182608923.html

Molecular modeling of interaction between ribavirin and nucleic acids

L.E. Vîjan, C. M. Topală
Department of Natural Sciences
Faculty of Science, University of Pitesti
Pitesti, Arges, Romania
loredana.vijan@upit.ro, carmen.topala@gmail.com

***Abstract*—**The ribonucleoside analog ribavirin shows the antiviral activity against a variety of DNA and RNA viruses. Ribavirin, in combination with interferon, has predominantly been applied in the treatment of hepatitis C virus infection and its potential antitumor efficacy has recently become a point of interest. The molecular structure of ribavirin was investigated by the semiempirical AM1 method, which triggered two polymorphic modifications of the antiviral drug also reported in the literature. The interactions of two polymorphic modifications of ribavirin (V_1 and V_2) with nucleic acids by the molecular mechanic and semiempirical AM1 methods were analysed. Previous experimental data pointed out that in the ribavirin – nucleic acid complexes, the 1,2,4-triazole-3-carboxamide chromophore is intercalated between the bases of the nucleic acid helix, the carboxamidic group is set outside of the helix toward the major groove and the 4-hydroxymethyl-tetrahydrofuran-2,3-diol fragment is located in the minor groove. In order to evidence the sequence specificity of the drug, some model mono- and double-stranded nucleic acid containing the bases: adenine (A), thymine (T), cytosine (C) and guanine (G) in AAAAAA, TTTTTT, CCCCCC, GGGGGG, ATATAT, CGCGCG, ATCGAT and CGATCG sequences were used. The results outline the differences in the contributions of the electrostatic and van der Waals interactions to the total binding energy and the preference of ribavirin for the binding at the sequences of nucleic acids containing adenine and thymine bases.

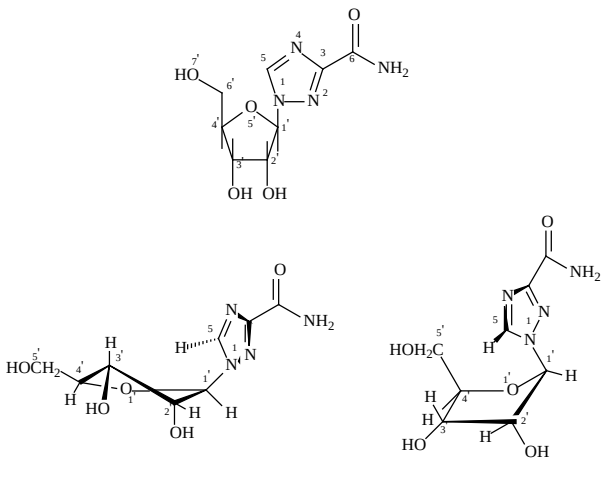
Keywords*—ribavirin; nucleic acids; molecular modeling*

I. INTRODUCTION

Ribavirin is a purine nucleoside analogue that is active against a number of DNA and RNA viruses [1]. There are numbers of proposed mechanisms of action for ribavirin. These include indirect effects such as inhibition of inosine monophosphate and immunomodulatory effects and direct effects such as polymerase inhibition and interference with viral RNA capping. Recent studies use double or triple combinations of ribavirin with other antiviral drugs, such as oseltamivir or/and amantadine in order to increase the activity against multiple virus strains in vitro [2, 3]. In addition, the antiviral drugs, such as ribavirin, are used by our research group in vitro chemotherapy to obtain of grapevine virus-free and potato virus-free plants [3].

Ribavirin has a complex structure (figure 1), comprising an 1,2,4-triazole-3-carboxamide chromophore and a ribose moiety, 2-hydroxymethyl-tetrahydro-furan-3,4-diol.

Ribavirin crystallises in two polymorphic forms V_1 and V_2. The forms V_1 and V_2 were found in slightly different conformations concerning the glycosyl bond, V_1 in „normal *anti*" and V_2 in "high *anti*" [4]. The V_1 form exhibits a glycosyl torsional angle χ (O_1'-C_1'-N_1-C_5) of 10.4° denoted as high *syn* and the ribose conformation is 3'-*endo*-2'-*exo*. The V_2 form has a χ value of 119.0° referred to as high *anti* and a 2'-*exo*-1'-*endo* ribose conformation (Fig. 1) [5]. Some inactive derivatives with substituents in 5-position (methyl, chloro) which also exist in the high *syn* forms, an lack energy minimum in the high *anti* region. Ribavirin does have a second minimum energy corresponding to the high *anti* conformation and it has been suggested that the active conformation of ribavirin at the enzyme site is the high *anti* form.



V_1 **V_2**

Fig. 1. Chemical structure for ribavirin and their two conformations V_1 ("high *syn*") and V_2 ("high *anti*")

Our experimental studies for ribavirin–DNA system [6] have pointed out a complex nature of the binding process. Two binding processes were highlighted: a process of the internal binding, that involve the drug intercalation between the bases from nucleic acid and a process of the external binding, that involve the drug binding to the grooves from the nucleic acid structure. It was found that the external binding prevails, the binding constant of this process being with an order of magnitude greater than the binding constant of the second process.

This work was supported by a grant of the Romanian National Authority for Scientific Research, CNDI-UEFISCDI, project number 104/2012 (PN-II-PT-PCCA-2011-3.1).

In addition, the dependence of the binding constants on the ionic strength of the medium allowed the dissection of the binding free energy in electrostatic and non-electrostatic contributions. It was found that the non-electrostatic contribution prevails.

The purpose of this paper is to perform a theoretical modeling of the interaction of ribavirin with some model nucleic acids. We will focus on the following aspects:
i) The study of the electronic structure of the antiviral drug, which involves the analysis of the possible conformers in order to find the optimal conformations for the interaction with the nucleic acids, the calculation of the charge distribution and the electrostatic potential;
ii) The theoretical modeling of the drug - nucleic acid complexes in order to estimate the relative contributions to the interaction energy (van der Waals and electrostatic terms) and to get an insight on the sequence selectivity of the drug.

II. COMPUTATIONAL DETAILS

The structures of the ribavirin conformers and the sequences of nucleic acids were built within the HyperChem Release 7.5 program and optimized by the semiempirical AM1 method (parameters: SCF control of 0.01, RHF spin pairing, Polak - Ribiere optimizer, RMS gradient of 0.01 kcal/mol·Å for the ribavirin conformers and the sequences of mono-stranded nucleic acids, RMS gradient of 0.1 kcal/mol·Å for the sequences of double-stranded nucleic acids).

The calculations on the complexes of two ribavirin conformers (noted V_1 and V_2) with the sequences of mono- and double-stranded nucleic acids were performed in vacuo by both the Molecular Mechanics (MM+ force field) and the semiempirical AM1 methods. The optimization criteria for the drug – nucleic acids complexes were 0.01 kcal/mol·Å for the MM method and 0.1 kcal/mol·Å for the AM1 method.

III. RESULTS AND DISCUSSION

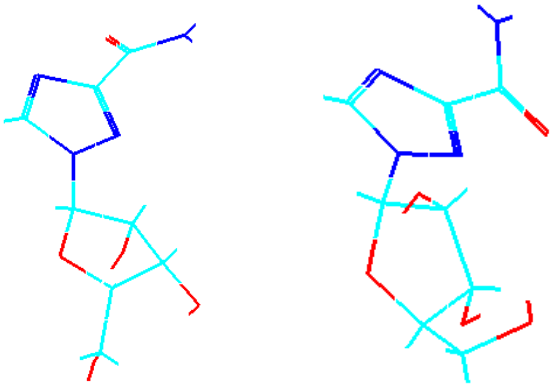
A. *Conformers of ribavirin*

After the conformational analysis, we have obtained a series of conformers presenting the features indicated in literature for the polymorphic forms V_1 and V_2. From all of the conformers obtained, they were selected two optimum conformations, with the minimum energy and corresponding to the crystallographic data presented in literature.

Table 1 shows the values of torsions angles between the binding atoms of the two rings from the ribavirin structure and figure 2 shows the molecular structures of the ribavirin conformers obtained by AM1 method.

TABLE 1. Values of torsion angles in ribavirin conformers

Torsion angles	Ribavirin V_1	Ribavirin V_2
N_1-C_1'-O_1'-C_4'	101.95°	113.08°
C_1'-O_1'-C_4'-C_5'	130.78°	-115.76°
O_1'-C_4'-C_5'-O_5'	-67.83°	102.83°

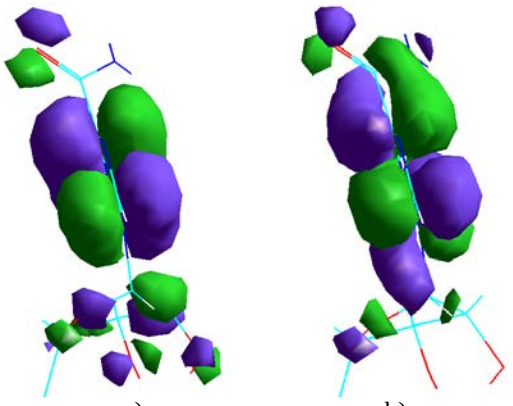


(a) Conformer V_1 (b) Conformer V_2

Fig. 2. The molecular structures of the ribavirin conformers obtained by AM1 method

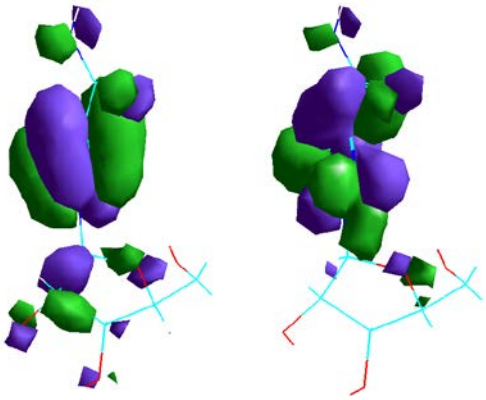
The charge distribution in the two conformers of ribavirin is similar, excepting the oxygen and nitrogen atoms from carboxamide group, the N_2 and N_4 atoms from 1,2,4-triazole group and the O_5' atom from ribose moiety. In the V_1 conformer, the smallest charge density was found at the nitrogen atom from carboxamide group and the O_5' atom from ribose moiety while in the V_2 conformer, the nitrogen and oxygen atoms on carboxamide group bears the smallest charge density.

The frontier molecular orbitals (Figs. 3 and 4) are π orbitals.



a) b)

Fig. 3. The HOMO (a) and LUMO (b) orbitals of ribavirin V_1



a) b)

Fig. 4. The HOMO (a) and LUMO (b) orbitals of ribavirin V_2

ISBN: 978-1-61804-327-6

In both conformers of ribavirin, the highest occupied molecular orbital (ε_{homo}~-10.45eV) and the first vacant molecular orbital (ε_{lemo}~-0.1eV) are preferentially localized on the heterocyclic oxygen and 1,2,4-triazole-3-carboxamide chromophore, ensuring a high superposition with the π-system of the base pairs from the nucleic acids and explaining thus the strong intercalating tendency of ribavirin conformers.

The electrostatic potential is a useful analytical tool in the analysis of chemical reactive behaviour involving the electrophilic and nucleophilic processes, as well as recognition and hydrogen bonding interactions. The electrostatic potential was calculated for each conformer of ribavirin and the minimum and maximum values for these conformers are presented in table 2.

TABLE 2. Results of AM1 calculations

Parameter	Ribavirin	
	V_1	V_2
$\Delta H_{formation}$, Kcal/mol	-139.47	-141.25
ε_{homo}, eV	-10.42	-10.47
ε_{lumo}, eV	-0.05	-0.13
V_{min}, Kcal/mol	-41.67	-36.39
V_{max}, Kcal/mol	266.53	246.01

Due to the presence of the lone pair electrons, the oxygen and nitrogen atoms are characterized by negative regions of the electrostatic potential while the rings are characterized by positive electrostatic potential. There are not significant differences considering the extension of the positive and negative regions of the electrostatic potential between the conformers of ribavirin. However, the V_2 conformer is more stable than the V_1 conformer, although the differences between them are small.

The both ribavirin conformers were optimized by both MM and AM1 methods. The optimized values of the ribavirin conformers energies calculated by the MM method were used to calculate the drug – nucleic acid interaction energy.

B. Sequences of nucleic acids

Nucleic acids are complex organic molecules that contain the genetic code for the organism. Nucleic acids act as drugs by different mechanisms, they may bind with the synthesized proteins, and they can hybridize to a messenger RNA leading to translation arrest or may induce degradation to target RNA. In this way the nucleic acids act as drug for inhibiting gene expression or protein synthesis.

A remarkable feature of the nucleic acids is that in these macromolecules there are several reactive sites, uniquely displayed on the surface of the helix, depending on the nitrogenous bases succession in the nucleic acids sequences. For instance, in the minor groove of deoxyribonucleic acid, the exocyclic N2 amino group of guanine and the N3 atom of both guanine and adenine bases are particularly susceptible to the drugs action. In the major groove, the N7 atom of both guanine and adenine bases is particularly susceptible to drug action. Finally, the C4', C5', and C1' atoms of the deoxyribose in the backbone of nucleic acid double-helix are other reactive sites from the nucleic acids sequences [7-9].

Because the literature data pointed out that the intercalating drugs have a sequence selectivity for nucleic acids that does not extend beyond two or three nitrogenous base level [10-14], we have chosen some model mono- and double-stranded nucleic acids containing the AAAAAA, TTTTTT, CCCCCC, GGGGGG, ATATAT, CGCGCG, ATCGAT and CGATCG sequences. The nucleic acids sequences were constructed by the charge neutralization of phosphate groups with hydrogen atoms. The nucleic acids sequences were optimized by both MM and AM1 methods.

The results obtained by the optimization the nucleic acids sequences using the AM1 method indicate a low net charge on the nitrogen atoms and negative area for the electrostatic potential on the oxygen and nitrogen atoms, which have lone pair electrons.

The optimized values of the nucleic acids sequences energies calculated by the MM method were used to calculate the drug – nucleic acid interaction energy.

C. Complexes of ribavirin with sequences of nucleic acids

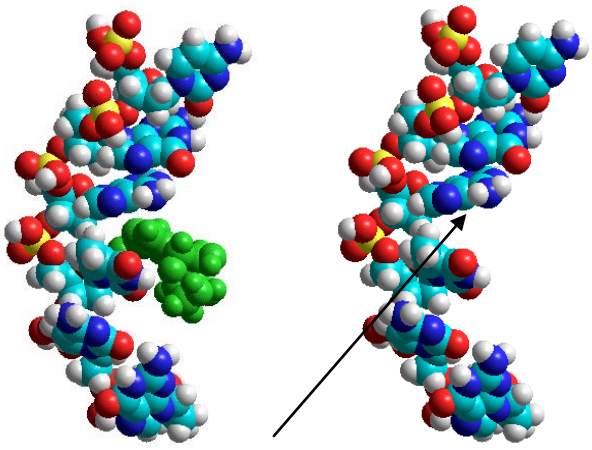
The optimized conformers of ribavirin and the optimized sequences of mono- and double-stranded nucleic acids were utilized in the optimization of the drug - nucleic acid complexes. For the optimization of the drug - nucleic acid complexes by both MM and AM1 methods, the solvent effect was not considered.

The starting structures of the drug - nucleic acid complexes were built by the docking procedure. Initially, several restraints were imposed, so that the 1,2,4-triazole-3-carboxamide chromophore to be oriented parallel to the nitrogenous bases from the nucleic acid helix, the carboxamidic group to be set outside of the helix toward the major groove and the 4-hydroxymethyl-tetrahydrofuran-2,3-diol moiety to be located in the minor groove of nucleic acids structure. After optimization of the drug - nucleic acid complexes until the required gradient, the restraints were eliminated and the complexes were optimized again.

In Figs. 5 and 6 are presented the optimized geometries for two complexes of the drug with the mono-stranded nucleic acids sequences and in Figs. 7 and 8 are presented the optimized geometries for two complexes of the drug with the double-stranded nucleic acids sequences.

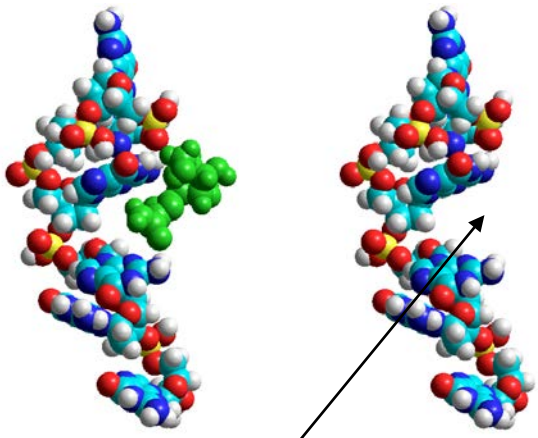
In all drug - nucleic acid complexes, the formation of some intercalation complexes was observed. Initially, the nucleic acid undergoes a conformational change that leads to the obtaining of an intercalation site. In this step, the nitrogenous bases from the nucleic acid structure were separated to form the cavity in which the drug will intercalate. Then, in the second step occurs an external binding of drug at the nucleic acid sequence and in the third step occurs the drug intercalating between the nitrogenous bases from the nucleic acid structure.

ISBN: 978-1-61804-327-6



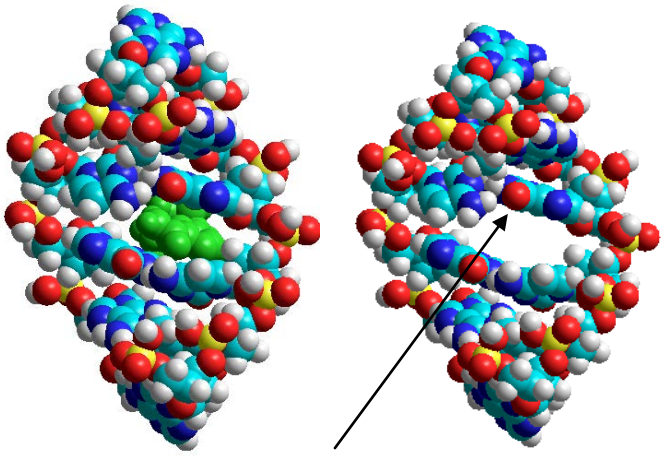
a) The intercalation site b)

Fig. 5. The optimized geometry of ribavirin V_1-CGATCG complex (a). The intercalation site of the ribavirin V_1 to mono-stranded nucleic acid sequence (b)



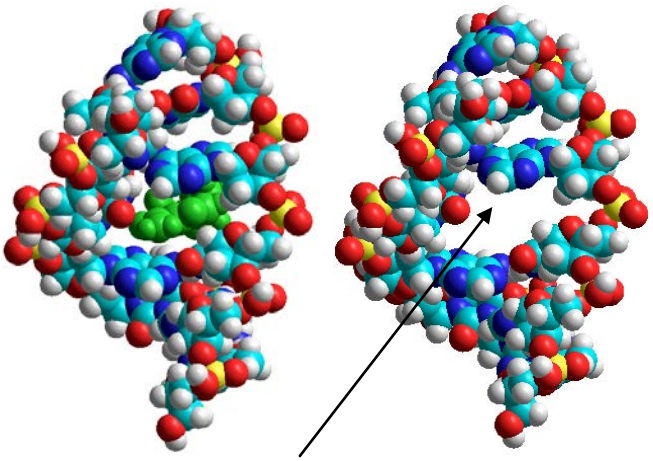
a) The intercalation site b)

Fig. 6. The optimized geometries of ribavirin V_2-GGGGGG complex (a). The intercalation site of the ribavirin V_2 to mono-stranded nucleic acid sequence (b)



a) The intercalation site b)

Fig. 7. The optimized geometries of ribavirin V_1-ATCGAT-TAGCTA complex (a). The intercalation site of the ribavirin V_1 to double-stranded nucleic acid sequence (b)



a) The intercalation site b)

Fig. 8. The optimized geometries of ribavirin V_2-ATATAT-TATATA complex (a). The intercalation site of the ribavirin V_2 to double-stranded nucleic acid sequence (b)

The energies of the drug – nucleic acid complexes were used for the evaluation of the following quantities [10, 12, 15]:

- the interaction energy:

$$E_{\text{int}\,eraction} = E_{complex} - (E_{drug} + E_{DNA})_{optimized} \quad (1)$$

- the binding energy:

$$E_{binding} = E_{complex} - (E_{drug} + E_{DNA})_{frozen\ in\ complex} \quad (2)$$

- the perturbation energy:

$$E_{perturbation} = E_{i\,nteraction} - E_{binding} \quad (3)$$

The values of the binding energies of the ribavirin conformers at the mono- and double-stranded nucleic acids sequences calculated by MM method are presented in table 3. The van der Waals (VdW) contribution to the binding energy are also included in table 3.

TABLE 3. MM results of drug – nucleic acid interaction

Nucleic acid	Ribavirin V_1		Ribavirin V_2	
	$E_{binding}$, Kcal/mol	% VdW	$E_{binding}$, Kcal/mol	% VdW
AAAAAA	-19,05	86,31	-17,85	69,47
TTTTTT	-21,24	79,81	-24,37	78,05
ATATAT	-23,75	71,98	-20,02	78,12
CCCCCC	-20,26	77,84	-18,77	73,73
GGGGGG	-15,25	93,38	-16,39	88,41
CGCGCG	-18,92	81,61	-17,85	88,63
ATCGAT	-14,24	80,76	-15,12	74,67
CGATCG	-14,23	83,91	-12,95	69,88
AAAAAA-TTTTTT	-27,97	78,48	-26,68	86,58
ATATAT-TATATA	-23,91	82,93	-24,39	83,81
CCCCCC-GGGGGG	-25,14	70,59	-28,27	72,56
CGCGCG-GCGCGC	-30,13	80,95	-27,11	86,09
ATCGAT-TAGCTA	-22,77	96,25	-23,24	89,51
CGATCG-GCTAGC	-20,46	79,62	-22,74	75,45

In all cases, the binding energies have negative values reflecting the drug - nucleic acid interaction. The results underline the significant van der Waals contribution (>70%) to the binding energy and, consequently, the low percentage of the electrostatic interactions, in agreement with our previous

ISBN: 978-1-61804-327-6

experimental data [6]. A slight preference for the sequences containing adenine and/or thymine bases can be noticed for both ribavirin conformers with both mono- and double-stranded nucleic acids sequences.

In table 4 are presented the values for the interaction and perturbation energies, characteristic for the inclusion processes of the two ribavirin conformers in the nucleic acids structures. It is noted that the intercalation of ribavirin conformers in the nucleic acids helix causes a small disturbance in the drug structure and a big disturbance in the structure of nucleic acids.

TABLE 4. MM results of drug – nucleic acid interaction

Ribavirin – nucleic acid complexes		$E_{interaction}$, kcal/mol	$E_{perturbation}$, kcal/mol	
			total	nucleic acid
RIBAVIRIN V_1	AAAAAA	-3,96	15,09	14,74
	TTTTTT	-9,67	11,57	10,76
	ATATAT	-16,19	7,56	6,79
	CCCCCC	-13,61	6,63	6,32
	GGGGGG	-2,71	12,54	12,12
	CGCGCG	-7,79	11,13	10,84
	ATCGAT	-12,76	1,48	0,64
	CGATCG	-13,41	0,82	0,73
	AAAAAA-TTTTTT	-25,75	2,22	1,06
	ATATAT-TATATA	-22,54	1,37	0,22
	CCCCCC-GGGGGG	-20,32	4,82	3,59
	CGCGCG-GCGCGC	-11,87	18,26	17,52
	ATCGAT-TAGCTA	-11,81	10,96	10,21
	CGATCG-GCTAGC	-15,99	4,47	2,95
RIBAVIRIN V_2	AAAAAA	-3,78	14,07	13,59
	TTTTTT	-15,88	8,49	7,81
	ATATAT	-11,31	8,71	8,11
	CCCCCC	-5,9	12,87	9,73
	GGGGGG	-2,7	13,69	12,84
	CGCGCG	-9,96	7,89	7,33
	ATCGAT	-9,67	5,45	4,61
	CGATCG	-12,04	0,91	0,44
	AAAAAA-TTTTTT	-24,39	2,29	0,55
	ATATAT-TATATA	-25,55	2,22	1,06
	CCCCCC-GGGGGG	-16,05	12,22	10,91
	CGCGCG-GCGCGC	-23,13	3,98	3,3
	ATCGAT-TAGCTA	-13,48	9,76	8,83
	CGATCG-GCTAGC	-11,38	11,36	10,34

The binding energies obtained by the AM1 method are presented in table 5.

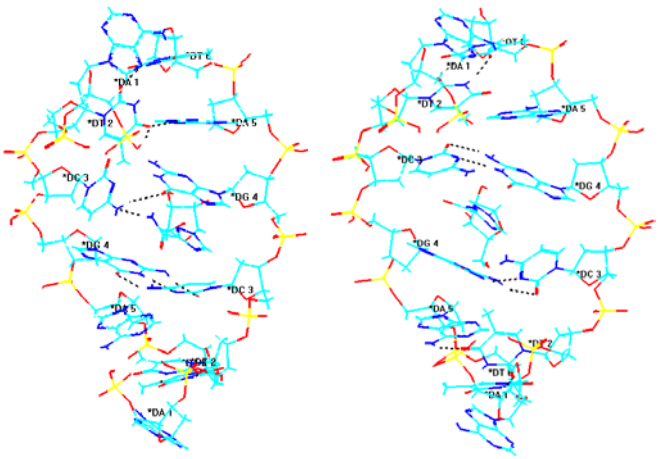
TABLE 5. Results of drug – nucleic acid interaction obtained by AM1 method

Nucleic acid	$E_{binding}$, kcal/mol	
	Ribavirin V_1	Ribavirin V_2
AAAAAA	-119,15	-102,68
TTTTTT	-21,89	-20,81
ATATAT	-20,54	-24,96
CCCCCC	-26,31	-8,31
GGGGGG	-22,95	-23,64
CGCGCG	-27,07	-13,54
ATCGAT	-16,69	-10,01
CGATCG	0,94	2,44
AAAAAA-TTTTTT	-98,73	-97,87
ATATAT-TATATA	-18,03	-9,96
CCCCCC-GGGGGG	-20,48	-16,66
CGCGCG-GCGCGC	7,96	-9,37
ATCGAT-TAGCTA	0,74	3,76
CGATCG-GCTAGC	16,11	-19,13

A slight preference for the nucleic acids sequences containing adenine and/or thymine bases can be noticed for both ribavirin conformers.

They were found a lot of features of the dyes and drugs intercalation [10-17] between the purine and pyrimidine bases from the nucleic acids structure, namely: the distortion in nucleic acid structure by the angle opening of the phosphate groups for allowing the dyes/drug intercalation, the lengthening of the helix by approximately 3.4 Å which causes a conformational change of some sugar moieties involved, the increase in the distance between nitrogenous bases at the intercalation site level.

Although some features of the intercalation are found in the ribavirin - nucleic acid system, however there is a remarkable difference, determined by the "accordion type" motion (specified in the model Lerman [16]) that occurs with the breaking of some hydrogen bonds between the purine and pyrimidine bases from the nucleic acids structure. In addition, the formation of new hydrogen bonds between ribavirin and the nitrogenous bases from the nucleic acid sequence at the intercalation site level was observed.



a) b)

Fig. 9. Hydrogen bonds in ribavirin V_1 (a), respectively ribavirin V_2 (b) - ATCGAT-TAGCTA complexes

CONCLUSIONS

The results of the molecular modeling points out that the complexes of ribavirin with nucleic acids are stabilized mainly by van der Waals forces involving the 1,2,4-triazole-3-carboxamide chromophore and the nitrogenous bases from the nucleic acids structure and that the electrostatic term brings a minimal contribution (<20%) to the binding energy. For both ribavirin conformers, a slight preference for nucleic acids sequences containing adenine and/or thymine bases was found. As a result of the ribavirin – nucleic acid interaction, only the nucleic acids structure is significantly perturbed, the structure of the drug being practically unchanged.

The theoretical calculations in the ribavirin – nucleic acid system predict an increase of the distance between the adjacent nitrogenous bases at the intercalation site level.

ISBN: 978-1-61804-327-6

The turning of the polynucleotidic helix is produced and the "accordion type" motion takes place determining a breaking in the hydrogen bonds between base pairs from double-stranded nucleic acid structure.

REFERENCES

[1] E. De Clercq, "Toward Improved Anti-HIV Chemotherapy: Therapeutic Strategies for Intervention with HIV Infections", Adv. Virus Res., vol. 42, pp.1–55, 1993.

[2] J.T. Nguyen, J.D. Hoopes, D.F. Smee, M.N. Prichard, E.M. Driebe, D.M. Engelthaler, M.H. Le, P.S. Keim, R.P. Spence, and G.T. Went, "Triple combination of oseltamivir, amantadine, and ribavirin displays synergistic activity against multiple influenza virus strains in vitro", vol. 53, pp. 4115–4126, 2009, doi:10.1128/AAC.00476-09

[3] I.C. Guţă, E.C. Buciumeanu, E.M. Vişoiu, and C.M. Topală, Modifications of Some Biochemical Compounds of Virus Infected Grapevine *in vitro* Cultured on Media with Chemotherapeutic Agents, Rev. Chim. (Bucharest), vol. 65(8), 884-887, 2014.

[4] P. Prusiner, and M. Sundaralingam, "The Crystal and Molecular Structures of Two Polymorphic Crystalline Forms of Virazole (1-β-D-Ribofuranosyl-l,2,4-triazole-3-carboxamide). A New Synthetic Broad Spectrum Antiviral Agent", Acta Cryst., B32, pp. 419-426, 1976.

[5] D.L. Miles, D. W. Miles, P. Redington, and H. Eyring, "Theoretical studies of the conformational properties of ribavirin", Proc. Natl. Acad. Sci. U. S. A., vol. 73(12), pp. 4257-4260, Dec.1976

[6] L.E. Vîjan, C.M. Topală, "UV-VIS absorption study of ribavirin – DNA interaction", unpublished

[7] X.L. Yang, and A.H. Wang, "Structural studies of atom-specific anticancer drugs acting on DNA", Pharmacol. Ther., vol. 83, pp. 181-215, 1999.

[8] F. Gago, "Stacking interactions and intercalative DNA binding", Methods, vol. 14, pp. 277-292, 1998.

[9] Y.M. Chang, C.M. Chen, and M.H. Hou, "Conformational changes in DNA upon ligand binding monitored by circular dichroism", Int. J. Mol. Sci., vol. 13, pp. 3394-3413, 2012.

[10] L. E. Vîjan, E. Volanschi, and M. Hillebrand, "Molecular modeling of anthracycline-DNA interaction", Progress in Colloid and Polymer Science, vol. 122, pp. 67-72, 2003.

[11] J. Ren, and J.B. Chaires, "Sequence and structural selectivity of nucleic acid binding ligands", Biochemistry, vol.38, pp. 16067-16075, 1999.

[12] L.E. Vîjan, "The theoretical study of the interaction of some anthracycline drugs with nucleic acids", Analele Universităţii din Bucureşti, vol. I-II, pp. 377-382, 2005.

[13] W.D. Wilson, S. Mizan, F.A. Tanious, S. Yao, and G. Zon, "The interaction of intercalators and groove-binding agents with DNA triple-helical structures: the influence of ligand structure, DNA backbone modifications and sequence", J. Mol. Recognit., vol. 7, pp. 89-98, 1994.

[14] K. Sandström, S. Wärmländer, J. Bergman, R. Engqvist, M. Leijon, and A. Gräslund, "The interaction of intercalators and groove-binding agents with DNA triple-helical structures: the influence of ligand structure, DNA backbone modifications and sequence", J. Mol. Recognit., vol. 14, pp. 277-285, 2004.

[15] M. Cervero, and F. Mendicuti, "Inclusion complexes of dimethyl 2,6-naphthalenedicarboxylate with alpha- and beta-cyclodextrins in aqueous medium: Thermodynamics and molecular mechanics studies", J. Phys. Chem. B, vol. 104, pp. 1572-1580, 2000.

[16] L.S. Lerman, "Structural considerations in the interaction of DNA and acridines", J. Mol. Biol., vol. 3, pp. 18-30, 1961.

[17] B.A. Neto, and A.A. Lapis, "Recent developments in the chemistry of deoxyribonucleic acid (DNA) intercalators: principles, design, synthesis, applications and trends", Molecules, vol. 14, pp. 1725-1746, 2009.

Rheological Modeling for Shape-Memory Thermoplastic Polymers

Hossein Hosseini
Dept. Chemical Engineering
Abadan Branch, Islamic Azad University
Abadan, Iran
Pedram465@yahoo.com

Boris V. Berdyshev
Dept. Chemical Equipment
Moscow State University of Mechanical Engineering
Moscow, Russia

***Abstract*—** This paper presents a rheological model for producing shape-memory thermoplastic polymers. Heat-shrinking occurs as a result of internal rearrangement of the structural elements of a polymer. A non-linear viscoelastic model was developed that allows qualitative and quantitative prediction of the stress-strain behavior of shape-memory polymers during heating. This research was done to develop a technique to determine the maximum possible change in size of heat-shrinkable products during heating. The rheological model used in this work was particularly suitable for defining process parameters and constructive parameters of the processing equipment.

***Keywords*—**shape-memory polymers; rheology; modeling; Elastic deformation.

I. INTRODUCTION

Shape-memory occurs as a result of internal rearrangement of the structural elements of a polymer. The ability of a polymer to shrink when it is subjected to heat is due to elastic de-formation of the polymer. Heat-shrinkable polymeric articles have wide-ranging industrial application such as in heat-shrinkable tubing, packaging industry, insulating wires, provision of abrasion resistance and environmental protection for stranded and solid wire conductors, connections and joints and terminals in electrical engineering. Heat shrinking can also be used to repair the insulation on wire or to form bundles of wires, to protect wires or small parts from minor abrasion and to create cable entry seals for environmental protection [1-4].

Existing multi-stage manufacturing technology of polymeric heat-shrinkable products requires several steps and uses specialist kinds of equipment.

Disadvantages of this technology are that it requires several steps, and specialist equipment and implementation of radiation exposure or chemical crosslinking of the polymer, which makes the technology low in efficiency, costly and environmentally unsafe. Furthermore, the process of heat-shrinking often uses the technique of applying various types of adhesive to the surface of insulated or connected products, which is undesirable because radiation or chemically modified polymers usually lose their adhesive properties. As melted polymer materials exhibit not only a viscous property, but also elastic properties, this raises the following question: Is it possible when blow molding hollow tubular articles to suppress viscous properties of the deformed polymer and translate it into a state of forced elasticity [5-6].

Most of the earlier modeling research has introduced rheological models consisting of spring, dash-pot, and frictional elements in one-dimensional models, in order to quantitatively describe the shrinkage behavior in amorphous polymers. However, despite their simplicity, such models usually lead to predictions agreeing only qualitatively with experiments. This study was done because there is currently a lack of appropriate constitutive rheological relations for the production of hollow heat-shrinkable polymeric articles.

II. RHEOLOGICAL MODELING

To describe the stress-strain behavior of the polymeric tubular blank, inflated by compressed gas, using the following rheological model.

$$\begin{cases} \bar{\sigma} + p\bar{\delta} = 2\bar{c}\,W_1 - 2\bar{c}^{-1}W_2 \\ \bar{e}_f = \dfrac{1}{\theta_0(T)G_0(T)}\exp\left\{\alpha\psi\dfrac{I_1-3}{I_1-1} - \beta\dfrac{W^S}{G_0(T)}\right\}\left[\left(\bar{c} - \dfrac{I_1}{3}\bar{\delta}\right)W_1^S - \left(\bar{c}^{-1} - \dfrac{I_2}{3}\bar{\delta}\right)W_2^S\right] \\ \dfrac{d\bar{c}}{dt} + \bar{\omega}\bar{c} - \bar{c}\bar{\omega} - \bar{c}\left(\bar{e} - \bar{e}_f\right) - \left(\bar{e} - \bar{e}_f\right)\bar{c} = 0, \end{cases} \quad (1)$$

where $\bar{\sigma}$ is the stress tensor, p is the Lagrange multiplier, determined by the boundary condition, $\bar{\delta}$ is the identity tensor, $\bar{c}$ is the Cauchy strain tensor, $\bar{e}_f$ is the flow strain rate tensor, $\bar{\omega}$ is the vortex tensor, $\bar{e}$ is the strain rate tensor, ψ is dimensionless parameter ($\psi = 0$ at $\bar{\omega} = 0$ and $\psi = 1$ at $\bar{\omega} \neq 0$), α is dimensionless parameter ($\alpha = 1$ at $\omega \neq 0$ and $\alpha = 0$ at $\omega = 0$) that characterize the presence or absence of reversible destruction of the structure of the polymer during deformation,

β is the flexibility of macromolecular chains, $\theta_0(T)$ is the relaxation time , $G_0(T)$ is the tensile modulus, W is the strain energy function $W = W(I_1, I_2)$, I_1 and I_2 are the primary and the secondary strain tensor invariants, t is the time, $f(I_1, I_2)$ is the dimensionless function that defines relaxation time, and $2W^S = W(I_1, I_2) + W(I_2, I_1)$is the symmetric function of W. The strain energy function parameters can be shown by:

$$W_1 = \frac{\partial W}{\partial I_1},\ W_2 = \frac{\partial W}{\partial I_2},\ W_1^S = \frac{\partial W^S}{\partial I_1},\ W_2^S = \frac{\partial W^S}{\partial I_2}.$$

To describe the elastic properties of the polymer material in the rheological model (1), the following elastic potential which gives a fairly adequate results in a variety of kinematic types of polymer media under loading [2].

$$W = 0{,}25G_0(I_1 + I_2 - 6) \quad (2)$$

Given the fact that during inflation of the tubular blank only increase its diameter (Figure 1), and its length is virtually unchanged. Due to the fact that the upper and lower ends of the blank clamped in the mold, we can assume that the deformation of the blank is carried out by mechanism of pure shear. In this case, the kinematic tensors and tensor of elastic deformation in the rheological model (1) will have the following form:

$$\bar{e} = \dot{\varepsilon} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}; \bar{\omega} = 0; \qquad \bar{c} = \begin{pmatrix} c & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & c^{-1} \end{pmatrix};$$

$$\bar{c}^{-1} = \frac{1}{\det \bar{c}} \begin{pmatrix} c^{-1} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & c \end{pmatrix} \quad (3)$$

where $\dot{\varepsilon} \equiv \frac{1}{\lambda}\frac{d\lambda}{dt}$ is the rate of deformation of the tubular blank in its circumferential direction, $c \equiv \lambda_e^2$; λ_e, λ are the elastic and total stretch ratio in the polymer, respectively. $\lambda_e = exp(\varepsilon_e^H)$, $\lambda = exp(\varepsilon^H)$ where ε^H is the Hencky strain, and ε_e^H is the elastic Hencky strain.

The primary and secondary invariants of tensor C are resulted from Equation (3) as:

$$I_1 = I_2 = c + 1 + c^{-1} \quad (4)$$

By utilizing Equations (2), (3), and (4), the following form of Equation (1) can be developed.

$$\bar{\sigma} + p\bar{\delta} = 0{,}5G_0(T) \cdot \begin{pmatrix} c - c^{-1} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & c^{-1} - c \end{pmatrix} \quad (5)$$

$$e_f \equiv \frac{d\bar{\varepsilon}_f^H}{dt} = \frac{1}{4\theta_0(T)} \exp\left[-\beta(c + c^{-1} - 2)\right] \cdot \begin{pmatrix} c - c^{-1} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & c^{-1} - c \end{pmatrix}$$

(6)

$$\frac{d\bar{c}}{dt} = 2\bar{c}(\bar{e} - \bar{e}_f) \quad (7)$$

where $\bar{\varepsilon}_f^H$ is the flow strain tensor as defined in Hencky.

III. RESULTS AND DISCUSSION

The following physical considerations were used to solve the problem.

Firstly, it was noted that the normal stresses acting on the thickness of a deformable blank can be neglected because of the small gas overpressure that is installed in the cavity of the inflated blank. This makes it possible from Equation (5), to determine expression for the Lagrange multiplier:

$$p = 0.5G_0(T) \cdot (c^{-1} - c) \quad (8)$$

Secondly, by taking into account that the stretch ratio is defined as $\lambda = r(t)/r_0$, the rate of deformation can be determined from Equation (3) as follows:

$$\dot{\varepsilon}(t) \equiv \frac{1}{\lambda(t)}\frac{d\lambda(t)}{dt} = \frac{1}{r(t)}\frac{dr(t)}{dt} \quad (9)$$

Relationship between the dimensionless current radius of the deformable blank ($r(t)/r_0$) and the process parameters of inflating will be as follows:

$$\frac{r(t)}{r_0} = \sqrt{1 + \left(\frac{P_0 + \Delta P}{P_u}\right)^{-1/k} \frac{G_u}{V_0} t}, \quad (10)$$

where P_0 is the initial gas pressure within the blank, P_u is gas pressure, ΔP is gas overpressure in cavity of blank, V_0 is the volume of the cavity of initial tubular blank, k is the adiabatic sign of the blowing gas.

The value of the volumetric flow rate of gas (G_u) through an inflatable nipple can be determined from the following equation [19]:

$$G_u = \mu_P S_{нип} \sqrt{\frac{2k}{k-1} R_P T_P} \quad (11)$$

where μ_P is the coefficient of pneumatic consumption, providing a supply of pressurized gas from the receiver, $S_{нип}$ is the cross sectional area of openings of nipple, R_P, T_P are the universal gas constant and temperature of the compressed gas in the receiver, respectively.

Considering expressions (10) and (11) the rate of deformation of the inflated blank is determined as follows:

$$\dot{\varepsilon}(t)=\frac{\dot{\varepsilon}(0)}{1+2\dot{\varepsilon}(0)t} \tag{12}$$

where $\dot{\varepsilon}(0)\equiv\dot{\varepsilon}(t=0)=\frac{1}{2}\left(\frac{P_0}{P_u}\right)^{-1/k}\frac{G_u}{V_0}$.

Expression (12) allows determination of the values of all components included in the first equation of the kinematic tensors (3). Now, using this tensor and Equations (6-7) the following scalar differential equation can be obtained that describes kinetics of the process of elastic deformation in the material in its circumferential direction:

$$\frac{dc}{d\tilde{t}}=2c\left\{\frac{E(0)}{1+2E(0)\tilde{t}}-\frac{\left(c-c^{-1}\right)}{4}\cdot\exp\left[-\beta\cdot\left(c+c^{-1}-2\right)\right]\right\} \tag{13}$$

Since, in initial extruded tubular blanks, any missing deformation including elastic deformation allows formulation of the following initial conditions for solutions of the differential equation (13):

$$\text{at } \tilde{t}=0 \quad\Longrightarrow\quad c\equiv 1 \tag{14}$$

Solution of Equation (13) with initial condition (14) for different values of the dimensionless parameter of $E(0)$ is presented in figure 1.

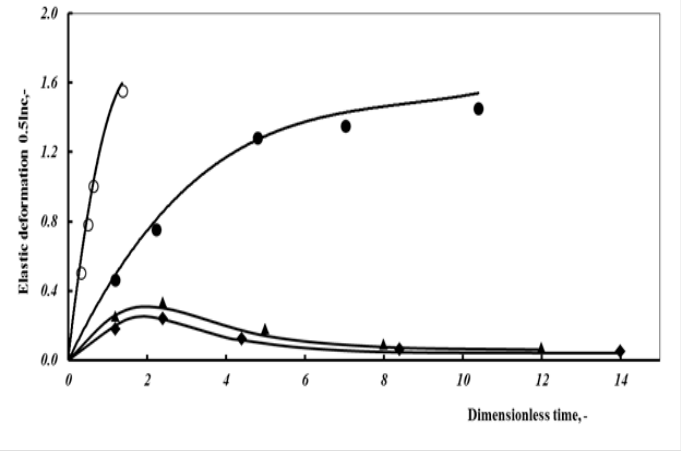


Figure 1. The kinetics of elastic deformations of LDPE in the inflated tubular blank at various modes of its deformation: ■ - *E(0)=0.41;* ▲ - *E(0)=0.6;* ● - *E(0)=1.25;* ○ - *E(0)=8.6;* points: experimental data; curves: theoretical model; rheological parameters: *β=0.38 and* θ_0*=0.05 sec; T=423K.*

Analysis of solutions of Equation (13) is confirmed by the experimental data, the comparison of which is represented in figure 1. Figure 1 shows that under certain conditions the deformation process of polymeric material goes into a state of "forced" elasticity and practically ceases to flow, like a deformed elastic medium. Data presented in Fig. 1 clearly demonstrates realization modes corresponding to curves 1 and 2 and that elastic deformation of the material was developed only in the initial moment of deformation of the blank, and then it relaxed to zero. Implementation of the same modes corresponding to curves 3 and 4, shows that elastic deformation of the material developed during the deformation time, and reached hundreds of percent or more. These results indicate that under practical conditions, the level of accumulated elastic deformation can be determined by dimensionless initial rate of deformation of the material, $E(0)$. Consequently, for the production of heat-shrinkable articles from conventional unmodified polymer, it is necessary that the initial rate of deformation exceeds a certain critical value, which is formalized as the following conditions:

$$E(0)\equiv\dot{\varepsilon}(\tau=0)\cdot\theta_0>E_{кр} \tag{15}$$

Where *E(0)* is the initial rate of deformation, and $E_{кр}$ is the critical value of rate of deformation.

IV. CONCLUSIONS

From the obtained equations it is obvious that identically sized heat shrinkable products may have completely different heat-shrinking characteristics that will define the value accumulated in the polymer material at the end of the formation process. These results demonstrate that for the practical implementation of the one-step production technology of shape-memory products, we can use well-known equipment commonly used for the extrusion blow-molding of hollow articles subject to adjustment of certain technological parameters of the formation process.

REFERENCES

[1] F. Khan, J. Koo, D. Monk, E Eisbrenner *Polym. Test.* vol.**27,** pp.498-502, 2008

[2] J. Diani, C. Frédy, P. Gilormini, Y. Merckel, G. Régnier, I. Rousseau *Polym. Test.* vol. **30,** pp.335-342, 2011.

[3] J. Morshedian, H. Khonakdar, S. Rasouli, *Macromol. Theory Simul.*, vol. 14, pp. 428-432, 2005.

[4] R. Rado, P. Zelenak, *Int. Polym. Sci. Technol.* vol.**19,** pp.72-82, 1992.

[5] S. R. Chowdhury, C. K. Das, *J. Appl. Polym. Sci.*, vol.**87,** pp.1414-1420, 2003.

[6] A. Landlein, *Shape-Memory Polymers.* Springer, New York, 2010.

Independence tests for financial variables

Sergio Ortobelli Lozza, Tommaso Lando

***Abstract*— T**his paper proposes an alternative method to evaluate the independence between random variables. The new method is particularly useful when the tested random variables are continuous, because the most used tests for independence are not able to give precise evaluations. In particular, we analyze and compare two different methods to test the independence among financial variables. The first is the classical chi-squared test generally used to evaluate the independence of historical observations in the portfolio risk valuation. The new alternative method is based on a conditional expectation estimator. Thus, we can compare the results of the two methods by evaluating the performance in terms of goodness-of-fit tests.

***Keywords*—**test of independence, conditional expectation, Kernel, Non Parametric test.

I. Introduction

This paper discusses two different methods to test the independence among random variables. On the one hand, several well known methods test independence between random variables by evaluating the independence between their realizations. Clearly, if the events are not independent, this criterion is sufficient to guarantee that the random variables are not independent. Thus, these methods can be properly used for discrete random variables. However, when random variables are continuous, we cannot guarantee that the random variables are independent only if a few events are independent. Moreover, in several financial applications, tests of these kinds are generally used although the financial random variables are assumed to be continuous. For example, when we evaluate the risk interval forecasts, with reference to the information available at each time, we use the tests proposed by [1], [2], and with a chi-squared test we also evaluate the time independence. In this paper, we propose an alternative method to test the independence among random variables, based on the conditional expectation between random variables. As observed by [3], the conditional expectation between two random variables $E(Y|X)$ can be estimated using different methodologies: the Kernel method and the OLP method. On the one hand, the kernel non-parametric regression (see [4] and [5]) allows to estimate $E(Y|X=x)$ as a locally weighted average, based on the choice of an appropriate kernel function: the method yields consistent estimators, provided that the kernel functions and the random variable Y satisfy some conditions, described in Section II. On the other hand, an alternative methodology was recently introduced by [6] for estimating the random variable $E(Y|X)$: this method has been proved to be consistent without requiring any regularity assumption. In this paper we use both methods to evaluate the difference between tests based on the conditional expectation and the classic chi squared test for the independence. In order to compare the effects of the two tests we discuss and examine the case of some financial variables using both alternative methodologies for estimating the random variable $E(Y|X)$. Then, we can perform a simulation analysis, drawing a bivariate random sample from (X,Y), and finally investigate which test better fits to the true case.

The paper is organized as follows: in Section II we present the different methodologies and their properties; in Section III we examine a method to compare the two tests; in Section IV we briefly illustrate the financial interpretation and possible application of the tests of independence.

II. Tests of Independence

In this section, we describe two different procedures to evaluate the independence among random variables. First, we present the well know Pearson chi-square test, used to test independence between random variables and events. Then, the second alternative test is based on the conditional expectation and thereby differs from several other tests which have been proposed for continuous random variables (see [7],[8] and [9]):

The chi-squared independence test

Two random variables X and Y are independent if for any couple of Borel sets A and B then

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B).$$

Therefore, if $X = \sum_{i=1}^{n} a_i I_{[X \in A_i]}$ and $Y = \sum_{j=1}^{m} b_j I_{[Y \in B_j]}$ are discrete random variables (where the collections $\{A_i\}_{i=1,\dots,n}$ and $\{B_j\}_{j=1,\dots,m}$ are partitions of the real line) we can easily test independence using the chi squared test. As a matter of fact, in order to prove the independence of X and Y it is sufficient to show that, for any i=1,…,n and j=1,…,m, we have that

$$p_{i,j} = P(X \in A_i, Y \in B_j) = P(X \in A_i)P(Y \in B_j) = p_i p_j.$$

Hence, we can use the statistic

This paper has been supported by the Italian funds ex MURST 60% 2014, 2015 and MIUR PRIN MISURA Project, 2013–2015, and ITALY project (Italian Talented Young researchers). The research was also supported through the Czech Science Foundation (GACR) under project 13-13142S and through SP2013/3, an SGS research project of VSB-TU Ostrava, and furthermore by the European Regional Development Fund in the IT4Innovations Centre of Excellence, including the access to the supercomputing capacity, and the European Social Fund in the framework of CZ.1.07/2.3.00/20.0296 (to S.O.) and CZ.1.07/2.3.00/30.0016 (to T.L.).

S.O. L. Author is with University of Bergamo, via dei Caniana, 2, Bergamo, Italy; and VŠB -TU Ostrava, Sokolská třída 33, Ostrava, Czech republic; e-mail: sergio.ortobelli@unibg.it.

T.L. Author is with University of Bergamo, via dei Caniana, 2, Bergamo, Italy; and VŠB -TU Ostrava, Sokolská třída 33, Ostrava, Czech republic; e-mail: tommaso.lando@unibg.it.

$$\chi^2 = \sum_{i=1}^{n}\sum_{j=1}^{m}\frac{(f_{ij}-e_{ij})^2}{e_{ij}} = N\sum_{i=1}^{n}\sum_{j=1}^{m}\frac{(p_{ij}-p_i p_j)^2}{p_i p_j} \quad (1)$$

where N is the sample size, p_{ij}, p_i, p_j are the estimated probabilities and similarly f_{ij} is the observed frequency count of the events belonging to both the *i*-th category of *X* and *j*-th category of *Y*, while e_{ij} is the expected count when *X* and *Y* are independent. Thus, the null hypothesis of the independence assumption must be rejected when the p-value of the chi squared statistic (1) (that is chi squared distributed with (m-1)(n-1) degrees of freedom) is less than a given significance level α. Observe that this statistic can also be used to test the independence of continuous random variables. However, in this case, the statistic cannot be applied in order to evaluate whether the random variables are independent, indeed we can only guarantee that the random variables are not independent if the null hypothesis is rejected.

Independence test based on the conditional expected value

Let $X:\Omega\to\mathbb{R}$ and $Y:\Omega\to\mathbb{R}$ be integrable random variables in the probability space $(\Omega,\Im,P)$. As observed by [3] when two integrable random variables *X* and *Y* are independent, then $E(Y|X)$=$E(Y)$ and generally the converse is not true, except in the case that *Y* is positive (negative). Thus, given a positive non constant measurable function g such that $E(g(Y)) < \infty$ we can easily test the independence of two integrable random variables *X* and *Y* by considering the variance of $E(g(Y)|X)$. As a matter of fact, the variance of $E(g(Y)|X)$ is equal to zero if and only if *Y* is independent from *X*. Hence, assume that "*Y* is independent from *X*" represents the null hypothesis of the test. We reject the null hypothesis anytime the variance of $E(g(Y)|X)$ is significantly greater than a given positive benchmark value. We call this test *conditional test*. Typically, we consider the function g(x)=|x|. Let $(x_1,y_1),(x_2,y_2),\dots,(x_n,y_n)$ be a random sample of independent observations from the bi-dimensional variable (X,Y). Next, we need an estimator of $E(|Y||X)$. In particular we recall that [3] proposed two alternative estimators of the conditional expected value: the first one is based on the Kernel non-parametric regression, and the other is based on the approximation of the sigma algebra generated by *X*. Thus the first procedure is aimed at estimating the conditional expectation of $|Y|$ given $X = x$, which is a mathematical function of X; the second method yields an unbiased and consistent estimator of the random variable $E(|Y||X)$.

The kernel non-parametric regression. It is well known that, if we know the form of the function $f(x) = E(|Y||X = x)$ (e.g. polynomial, exponential, etc.), then we can estimate the unknown parameters of $f(x)$ with several methods (e.g. least squares). In particular, if we do not know the general form of $f(x)$, except that it is a continuous and smooth function, then we can approximate it with a non-parametric method, as proposed by [4] and [5]. Thus, $f(x)$ can be estimated by:

$$\hat{f}_n(x) = \frac{\sum_{i=1}^{n}|y_i|K\left(\frac{x-x_i}{h(n)}\right)}{\sum_{i=1}^{n}K\left(\frac{x-x_i}{h(n)}\right)}, \quad (2)$$

where $K(x)$ is a density function such that i) $K(x) < C < \infty$; ii) $\lim_{x\to\pm\infty}|xK(x)| = 0$; iii) $h(n)\to 0$ when $n\to\infty$. The function $K(x)$ is denoted by *kernel*, observe that kernel functions are generally used for estimating probability densities non-parametrically (see [10]). It was proved in [10] that if $|Y|$ is quadratically integrable then $\hat{f}_n(x)$ is a consistent estimator for $f(x)$.

The OLP method. We now describe an alternative non-parametric approach [6] for approximating the conditional expectation, the method is denoted by "*OLP*", which is an acronym of the authors' names. Define by $\Im_X$ the σ-algebra generated by X (that is, $\Im_X = \sigma(X) = X^{-1}(\mathcal{B}) = \{X^{-1}(B): B\in\mathcal{B}\}$, where $\mathcal{B}$ is the Borel σ-algebra on $\mathbb{R}$). Observe that the regression function is just a "pointwise" realization of the random variable $E(|Y||\Im_X)$, which can equivalently be denoted by $E(|Y||X)$. $\Im_X$ can be approximated by a σ-algebra generated by a suitable partition of Ω. In particular, for any $k\in\mathbb{N}$, we consider the partition $\{A_j\}_{j=1}^{b^k} = \{A_1,\dots,A_{b^k}\}$ of Ω in b^k subsets, where b is an integer number greater than 1 and:

- $A_1 = \left\{\omega: X(\omega)\le F_X^{-1}\left(\frac{1}{b^k}\right)\right\}$,
- $A_h = \left\{\omega: F_X^{-1}\left(\frac{h-1}{b^k}\right) < X(\omega)\le F_X^{-1}\left(\frac{h}{b^k}\right)\right\}$, $for\ h = 2,\dots,b^k-1$
- $A_{b^k} = \Omega - \bigcup_{j=1}^{b^k-1}A_j = \{\omega: X(\omega) > F_X^{-1}\left(\frac{b^k-1}{b^k}\right)\}$.

Starting with the trivial sigma algebra $\Im_0 = \{\emptyset,\Omega\}$, we can obtain a sequence of sigma algebras generated by these partitions, for different values of *k* (*k*=1,…,*m*,…). For instance, $\Im_1 = \sigma\{\emptyset,\Omega,A_1,\dots,A_b\}$ is the sigma algebra generated by $A_1 = \{\omega: X(\omega)\le F_X^{-1}(1/b)\}$, $A_s = \left\{\omega: F_X^{-1}\left(\frac{s-1}{b}\right) < X(\omega)\le F_X^{-1}\left(\frac{s}{b}\right)\right\}$, *s*=1,...,*b*-1 and $A_b = \{\omega: X(\omega) > F_X^{-1}((b-1)/b)\}$. Generally:

$$\Im_k = \sigma\left(\{A_j\}_{j=1}^{b^k}\right), k\in\mathbb{N}. \quad (3)$$

Hence, it is possible to estimate the random variable $E(Y|\Im_X)$ by

$$E(|Y||\Im_k)(\omega) = \sum_{j=1}^{b^k}\frac{1_{A_j}(\omega)}{P(A_j)}\int_{A_j}|Y|dP = \sum_{j=1}^{b^k}E(|Y||A_j)1_{A_j}(\omega), \quad (4)$$

where $1_A(\omega) = \begin{cases}1 & \omega\in A\\ 0 & \omega\notin A\end{cases}$. It is proved in [6] that $E(|Y||\Im_k)$ is a consistent estimator of the random variable $E(|Y||X)$, that is, $\lim_{k\to\infty}E(|Y||\Im_k) = E(|Y||X)$ a.s.

From a practical point of view, given *n* i.i.d. observations of Y, if we know the probability p_i corresponding to the *i*-th outcome y_i, we obtain:

$$E(|Y||A_j) = \sum_{y_i\in A_j}|y_i|p_i/P(A_j). \quad (5)$$

Otherwise, we can give uniform weight to each observation, which yields the following consistent estimator of $E(|Y||A_j)$:

$$\frac{1}{n_{A_j}}\sum_{y_i\in A_j}|y_i|, \quad (6)$$

where n_{A_j} is the number of elements of A_j. Therefore, we are always able to estimate $E(|Y||\Im_k)$, which in turn is a consistent estimator of the conditional expected value $E(|Y||X)$.

A simple proof of the potentiality of the test can be given when we compare uncorrelated but dependent random variables as in the following section.

III. A COMPARISON AMONG TWO PORTFOLIOS

Let us consider two portfolios of daily returns X and Y, taken from the NYSE, which are empirically uncorrelated.[1] Consider that we have about three years of historical daily joint observations (750 trading days). First of all, we want to test if the losses and gains of the two portfolios are independent. Using the chi square test with one degree of freedom we could not reject the independence of the two portfolios at 95% significance level. Secondly we want to test if the two portfolios are independent. Thus, we apply the conditional test to the standardized random variables $\tilde{X}$ and $\tilde{Y}$ of X and Y. We get a variance of $E(|\tilde{Y}|/\tilde{X})$. equal to 0.0512 with the OLP estimator and 0.0445 with the Kernel estimator. We observe that the joint distribution of the two standardized portfolios can be well approximated by a bivariate t-student with 5 degrees of freedom. Thus, with a bootstrap technique based on bivariate t-student, we estimated the variance obtained for a sample of the same dimension (750 observations) under two different hypotheses: X and Y are independent t- distributed or X and Y are dependent but uncorrelated. For independent t distributed random variables we get an average variance of $E(|\tilde{Y}|/\tilde{X})$ equal to 0.0082, while for uncorrelated dependent t distributed random variables we get an average variance of $E(|\tilde{Y}|/\tilde{X})$ equal to 0.0431. This simple observation suggests to reject the independence hypothesis even if the two portfolio are uncorrelated.

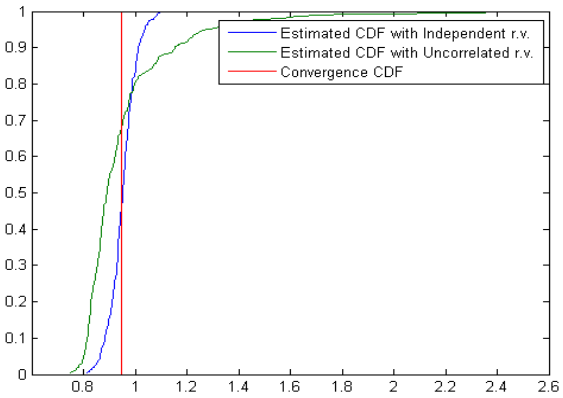


FIG. 1 Distributions of $E(|Y|/X)$ for uncorrelated or independent t-student random variables.

One further example of this analysis is given in Fig. 1 where we report the distributions of $E(|\tilde{Y}|)$ and of $E(|\tilde{Y}|/\tilde{X})$ (estimated with the OLP method) assuming $\tilde{X}$ and $\tilde{Y}$ to be uncorrelated or independent t distributed with 5 degrees of freedom.

[1] The procedure to get portfolio uncorrelated is very simple and can be useful in several practices for example in the PCA to reduce the dimensionality of the problem.

IV. CONCLUSION

In this paper, we deal with tests of independence among random variables. In particular, we show that the well known chi squared test for independence is not always able to evaluate correctly the independence between random variables. On the other hand, a newly proposed test is able to capture the dependence of random variables even when they are uncorrelated. In particular, we show that the new test could be based on two different methodologies for estimating the conditional expectation, namely the kernel method and the OLP method recently proposed by [6].

REFERENCES

[1] Christoffersen P. Evaluating interval forecasts. International Economic Review 1998; 39; 841-862.
[2] Kupiec P. Techniques for verifying the accuracy of risk measurement models, Journal of Derivatives 1995; 3; 73-84
[3] Lando T., Ortobelli S. (2015) "On the Approximation of a Conditional Expectation" WSEAS Transactions on Mathematics, Volume 14, pp. 237-247
[4] E. A. Nadaraya, "On estimating regression," *Theory of Probability and its Applications*, vol. 9, no. 1, pp. 141-142, 1964.
[5] G. S. Watson, "Smooth regression analysis," *Sankhya, Series A*, vol. 26, no. 4, pp. 359-372, 1964.
[6] S. Ortobelli, F. Petronio, T. Lando, "A portfolio return definition coherent with the investors preferences," under revision in IMA-Journal of Management Mathematics.
[7] K. Sricharan,R. Raich, and A. Hero. 2012 Empirical estimation of entropy functionals with confidence. *Technical report* arXiv:1012.4188
[8] A. Gretton, O. Bousqet, A. Smola, and B. Schölkopf, (2005) Measuring Statistical Independence with Hilbert-Schmidt Norms, *Algorithmic Learning Theory Lecture Notes in Computer Science* 3734, pp 63-77 .
[9] B. Schweizer and E.F. Wolff, (1981) On Nonparametric Measures of Dependence for Random Variables, *Annals of Statistics* 9(4), 879-885.
[10] V. A. Epanechnikov, "Non-parametric estimation of a multivariate probability density," *Theory of Probability and its Applications*, vol. 14, no. 1, pp. 153-158, 1965.

Time Difference Calculation based on Signal Starting Point Detection

Wan-zhen Zhou, Yu Ling, Yong-qiang Zhang, Wei-dong Wu

***Abstract*—The research of the high-precision sound source localization system which is based on time difference of arrival (TDOA) of different sensors positioning system has increasingly become popular research. The outcome of the traditional generalized cross-correlation function (GCC) time delay estimation algorithm which is under the condition of the intense noise, has an apparent error, therefore, we can't guarantee the accuracy of the positioning system. In this paper, we put forward a spectral entropy-double threshold endpoint detection algorithm, through obtaining the exactly starting time of source sound to get a time difference, then improving it. Experiments show that the algorithm has the characteristic of simple process, small amount of computation and accuracy. It is much suitable for DSP and small system time difference estimation module.**

***Index Terms*—**sound source localization; TDOA; GCC; endpoint detection algorithm.

I. INTRODUCTION

TDOA (Time Difference of Arrival) is a kind of positioning technology [1][2] which is based on the time delay difference of the simultaneous reception of the signal time parameters of each sensor. Traditional time difference algorithm [3][4][5] which delay difference is obtained by signal correlation operation has the large amount of computation and obvious delay difference errors because of the overlay operation. Since the traditional dual threshold endpoint detection algorithm [6][7] has been turned out inaccurate in the case of intense noise, we propose an improved spectral entropy-double threshold endpoint detection method.

Wan-zhen Zhou is from the School of Information Science & Engineering, Hebei University of Science and Technology, China. He received his doctor's degree from China people's Liberation Army Ordnance Engineering college in 2006. His main research interests include the search on Artificial intelligence, data mining and IOT. Email: zhou_wanzhen@163.com.

Yu Ling Zhang is from the School of Information Science & Engineering, Hebei University of Science and Technology, China. His main research interests include IOT and cloud computing. Email: lpingaoxy@163.com.

Yong-qiang Zhang is from the School of Information Science & Engineering, Hebei University of Science and Technology, China. His main research areas include the research on IOT and distributed computing. Email: zyqwork@gmail.com.

Wei-dong Wu is from the from the School of Information Science & Engineering, Hebei University of Science and Technology, China. His main research interests include the research on IOT and the embedded system. Email: wwd@hebust.edu.cn

II. DOUBLE THRESHOLD ENDPOINT DETECTION

A. The traditional double threshold endpoint detection

Energy always as the parameters for audio process. Put short-time energy as the judgement which is mean to subframe the signal first, and then calculate the the energy for each frame respectively. Average short-time energy of the signal {x(n)} at the time of n is defined as:

$$E_n = \sum_{m=-\infty}^{\infty} x^2(m)h(n-m) \quad (1)$$

Where h(n) is the window frame, used to implement subframe.

Short-time zero crossing rate means the count of the a frame of signalwaveform through the horizontal axis ,defined as:

$$Z_n = \frac{1}{2}\sum_{m=0}^{N-1} \left| \operatorname{sgn}\left[x_n(m)\right] - \operatorname{sgn}\left[x_n(m-1)\right] \right| \quad (2)$$

Traditional double threshold endpoint detection method through the product of short-time energy and zero-crossing rate for detecting. The basic idea is to determine two thresholds for the short-time energy and zero-crossing rate, i.e. make the short-time energy as the primary judge, and choose a higher threshold, and set the average zero-crossing rate as a secondary judgment, on this basis, select a lower threshold. It may not be the beginning of a valid signal when the lower threshold is exceeded which is likely caused by subtle noise. When the higher threshold is exceeded and the lower threshold is exceeded during a period of time, it means the beginning of effective signal. Although the calculation of the traditional dual threshold endpoint detection algorithm is simple, it usually can't detect the starting point of a sound signal accurately when the noise is large. Therefore, it can't get an accurate arrival time.

B. spectrum entropy of endpoint detection

Entropy is a measure of the degree of system chaos, which is used to represent homogeneous degree of energy distribution in the space. It has a certain robustness feature cause of the small change of spectral entropy compared with the energy characteristics. Spectral entropy sound endpoint detection algorithm is mainly to detect flatness of spectrum. The energy of the information section of sound signal which

ISBN: 978-1-61804-327-6

is mainly concentrated on certain frequency band compared with the background noise changes clearly and it's entropy value is small. While, the distribution of noise is relatively flat over the entire frequency band, with a large entropy value. So we can use this difference to get the accurate signal arrival time by distinguish the valid information segment with noise segment.

Spectral entropy is obtained by first calculating a Fast Fourier Transform (FFT) to get the frequency spectrum of each frame signal, whose coefficients shows the distribution in size of the frame signal at the frequency points, calculate the proportion of each spectral component in the total energy of each frame, which would be set as a signal energy probability in a certain spectrum point. Its probability density function defined as:

$$p_i = s(f_i) \Big/ \sum_{k=1}^{N} S(f_k), \quad i = 1, \ldots, M \tag{3}$$

Where, $s(f_i)$ is the energy of f_i, p_i is the corresponding probability density, and M is the points of the FFT frequency components. The corresponding spectral entropy of each frame is defined as:

$$H = -\sum_{k=1}^{M} p_k \cdot \log P_k \tag{4}$$

C. Improved spectral entropy - double threshold endpoint detection

In order to get a more accurate time difference of the signal, we need to reduce the signal noise by spectral subtraction before detecting the endpoint. By using the spectral subtraction, which needs to cut the noise spectrum from the sound we can obtain a relatively pure signal spectrum. In order to further increase the accuracy of the time difference, it is necessary to divide the sound signal into several frequency bands, the energy spectrum of frequency band is defined as:

$$G(n, m) = \sum_{k=1+5(n-1)+3}^{1+5(n-m)} | X(k, m) |^2, 1 \le n \le N \tag{5}$$

$$G(m) = \sum_{k=1}^{N} E(k, m) \tag{6}$$

Where N is the number of bands, every five points form a band, then the probability density of G(n, m) is:

$$P(n, m) = \frac{G(n, m)}{G(m)} \tag{7}$$

The multi-band spectral entropy is:

$$H(m) = -\sum_{n=1}^{N} P(n, m) \log P(n, m) \tag{8}$$

Finally, we determine the double threshold T1 and T2 according to multi-band energy, and T1 > T2. During detecting, we set that when status = 0 it represents the present situation, status = 1 represents probably in an effective state, status = 2 represents the beginning to enter the effective information, status = 3 represents ending of the signal, count is the timing length count according to the change of the status, specific detection process is as follows:

Step 1. Initially, when it is in the silent period, we set the status and count both as 0, if the spectral entropy value exceeds the low threshold T2, we begin to mark the starting point, and then enter the transition stage, the status become the value 1, and the count also has to add 1; when the spectrum entropy is under the lower threshold T2 in this section, we consider that the current situation is back to the silent period, the status becomes 0, and then the count also is back to the value 0.

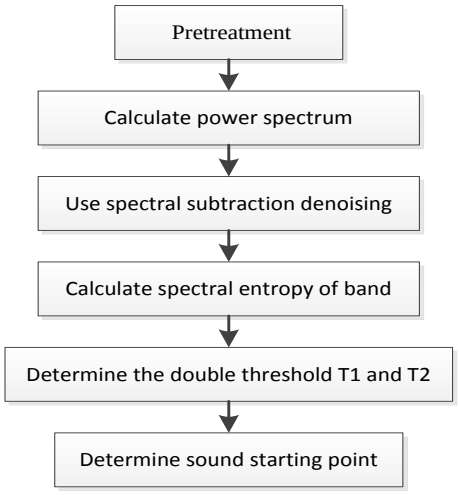


Fig. 1. flowchart of the improved spectrum entropy-double algorithm

Step 2. If in the transition section spectral entropy exceeds the high threshold T1, it enters into the valid segment, status becomes the value 2.

Step 3. When in effective information period, in case that the spectrum entropy value is less than the lower threshold T2 and the total time length is less than the shortest length of time, it is considered to be the noise.

To sum up, the flow chart of the improved spectral entropy-double endpoint algorithm as shown in Fig. 1.

III. THE EXPERIMENT SIMULATION

A. The experimental environment and parameter settings

In order to verify the performance of the algorithm, we use the MATLAB software to do the simulation experiment. Sound signal which is considered as a particular music is recorded respectively in a quiet laboratory and outdoor with large noise. We set the sampling frequency as 24KHz, sampling accuracy for 16bit. By framing, we take the hamming window, whose length is 256 points and shift is 90 points, and pre-emphasis by the filter with (1-0.95Z-1) filter. Then we detect the endpoint by the traditional and the improved double threshold algorithm respectively,

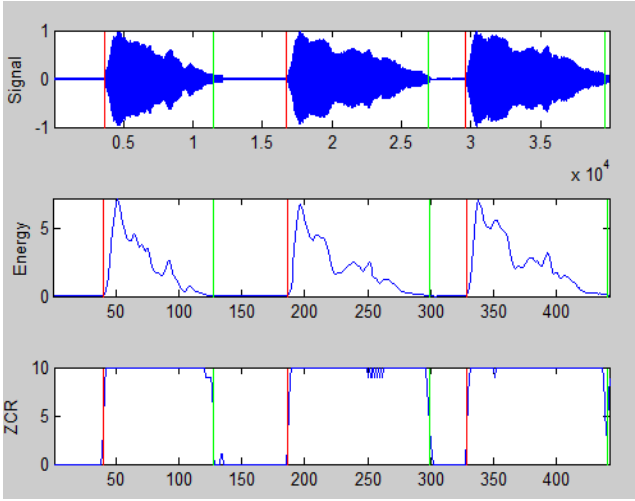


Fig. 2. The traditional double threshold endpoint detection Map when noise is small

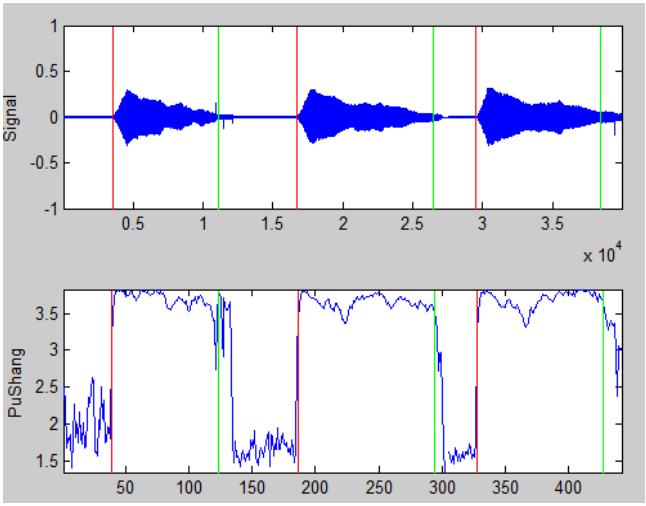


Fig. 3. The improved double threshold endpoint detection map when noise is small

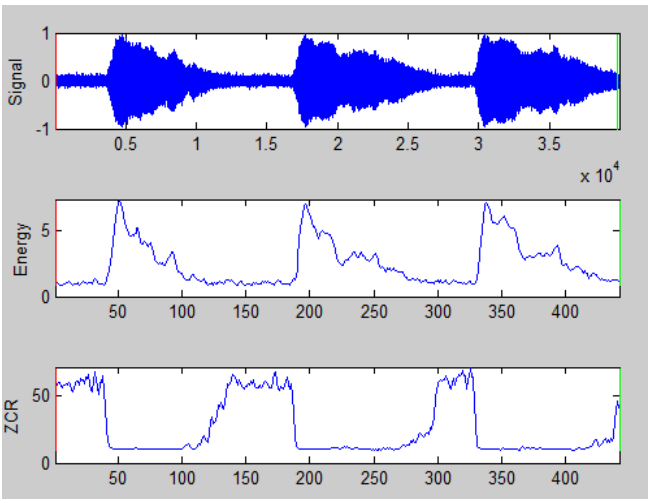


Fig. 4. The traditional double threshold endpoint detection map when noise is bigger

comparing the results made from the two kinds of algorithm. The contrast of the positioning results from the endpoint detection between the results of the traditional cross-correlation function, in [2][3][4] and the operation time

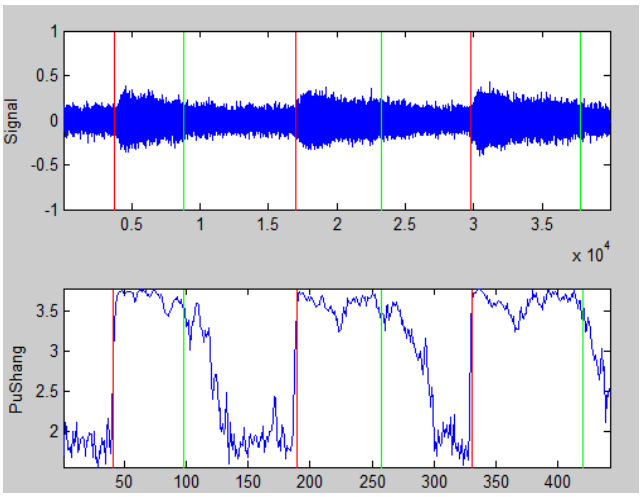


Fig. 5. The improved double threshold endpoint detection map when noise is bigger

of each algorithms.

B. *Performance Analysis*

This study adopts the traditional dual threshold endpoint detection algorithm and the improved endpoint detection algorithm to test the sound signal respectively, by which the same special music signal under the different noise conditions is detected. The results are shown in the figure below: the red vertical lines represent sound starting point, and the green line shows the end of the signal. From Fig. 2 and Fig. 3 it can be seen that when the noise is small, the results of the traditional double threshold endpoint detection algorithm and the improved spectral entropy - double threshold detection results are equally accurate. However, from the Fig. 4, with the increase of noise, the accuracy of traditional double threshold endpoint detection decreased, while the improved spectral entropy - double threshold endpoint detection algorithm can still be accurately detected under the same conditions, as shown in Fig. 5.

C. *Further Experiments*

Further, we recorded 10 different sets of special music signal, and respectively applied the traditional cross-correlation function method in the literature [3][4][5], and traditional endpoint detection algorithm in the literature [6][7] and the improved spectral entropy algorithm to calculate the time difference. The sound source TDOA positioning system as shown in Fig. 6.

The average operation time of each arithmetic, as shown in TABLE 1. The average position obtained from time difference results and system average operation time as shown in TABLE 2.

Where r, R respectively represent the condition of small and large noise, the actual two-dimensional coordinates of sound is (7m, 3m). As we can see from the table, when the noise is large we often can't get the accurate time difference data, which make the accuracy of the system positioning failed, from the traditional cross-correlation method and the traditional dual threshold endpoint detection method while

ISBN: 978-1-61804-327-6

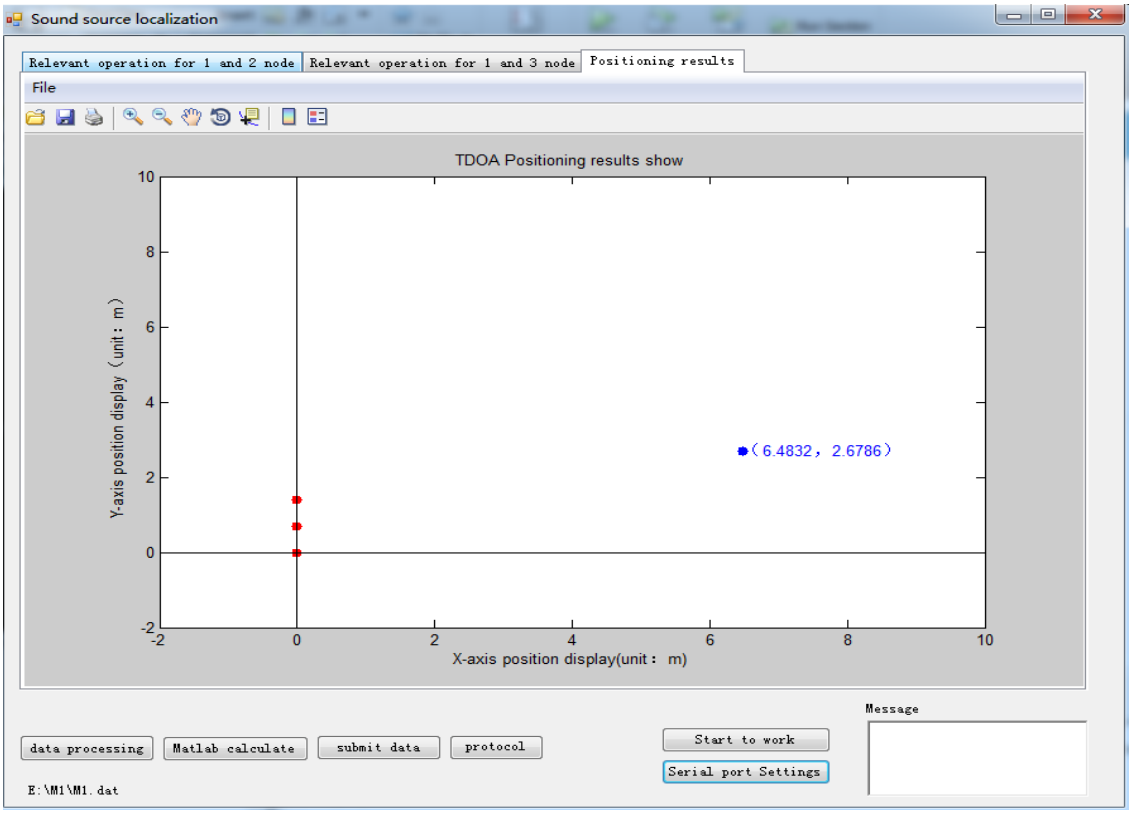


Fig. 6. TDOA positioning system

TABLE 1. AVERAGE OPERATION TIME OF EACH ARITHMETIC

Traditional cross-correlation function algorithm	traditional double threshold detection algorithm	the improved spectral entropy detection algorithm
560ms	280ms	410ms

TABLE 2. SYSTEM AVERAGE OPERATION TIME

position results / algorithm	X coordinate (m)	Y coordinate (m)	Operation time (ms)
traditional cross-correlation (r)	5.8	2.4	760
traditional cross-correlation (R)	13.2	-3.7	
traditional double threshold (r)	5.5	2.0	480
traditional double threshold (R)	fail	fail	
the improved spectral entropy (r)	5.6	2.3	610
the improved spectral entropy (R)	5.3	2.0	

the improved spectral entropy can still get a more accurate data in a larger noise condition with the character of small amount of computation and short operation time.

IV. CONCLUSION

This paper puts forward an improved spectrum entropy – double threshold endpoint detection algorithm. And the time difference between the sensors is obtained by the proposed algorithm when it is applied to the TDOA localization system. The experimental results show that the proposed algorithm is not only steady and strong anti-noise performance, but the amount of computation is small, compared with the traditional cross-correlation function algorithm and the double threshold endpoint algorithm. Thus it is better to meet

ISBN: 978-1-61804-327-6

real-time requirements for DSP and other small positioning system.

REFERENCES

[1] Dirk Bechler, Kristian Krosechel. "Considering the second Peak in the GCC Function for Muti-source TDOA Estimation with a Microphone Array". International Workshop on Acoustic Echo and Noise Control(IWAENC2003), Kyoto, 2003.

[2] Yong Ren. FGPA Implementation of sound source localization based on microphone array [master's degree thesis]. XiAn: Xian university of electronic science and technology, 2011.

[3] David Lindgren, Olof Wilsson. Shooter Localization in Wireless Microphone Networks. EURASIP Journal on Advances in Signal Processing. 2010: 10-30.

[4] Peng Li, The sniper positioning algorithm based on Mach wave research and simulation. Beijing jiaotong university master's degree thesis. 2011: 5-20.

[5] ZHAO X.F, ZHANG W, YANG T.L, JIANG C.S. A DBF Method of Localizing Accuracy Improvement and Speech Enhancement Based on A New Microphone Arrays International Conference on Signal Processing 2012.10.

[6] Hang Hu. Speech signal processing [M]. HarBin: Harbin industrial university press, 2000.

[7] XueYing Zhang. Speech signal processing and MATLAB simulation [M].Beijing: Electronic industry press, 2010.

Discrete Event Simulation Robotic Technology of Mining

Vasily V. Sinoviev, Alexy N. Starodubov,
Mihail U. Dorofeev
Siberian Branch of the Russian Academy of Sciences
Institute of coal
Kuzbass State Technical University
Kemerovo, Russia
zinovyev@kemsc.sbras.ru

Victor V. Okolnishnikov
Siberian Branch of the Russian Academy of Sciences
Design Technological Institute of Digital Techniques
Novosibirsk, Russia
okoln@mail.ru

***Abstract*—**The experience of discrete event simulation of robotic technology of mining for comparison of conducting works on the selected criteria is presented.

***Keywords*—**discrete event simulation, queuing system, robotic mining technology

Mining is carried out by drilling and blasting or by a combine. The main works are carried out in the mine when blasting and drilling are charging, rock mass loading and support fixing.

The drill rig 1 first enters the heading face (Fig. 1.) and bores a specified number of holes. After charging of holes and blasting cargo transport vehicle 2 exports mountain range to the distance depending on the distance of the face. Then anchors installer 3 enters the face and fixes support by anchors. Simultaneously ancillary works are performed (installation of drainage, building ventilation pipes and others).

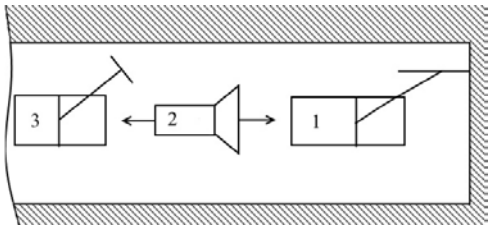


Fig. 1. Schemes of gallery driving by drilling and blasting.

The combine, simultaneously, (Fig. 2) drives in gallery executing unit 1 for a given program, and loads the rock mass when unit 2 is loaded with its transportation conveyer bridge 3 to haulage unit 4. After gallery driving to a predetermined depth the combine is off and the support is installed by robotic anchors installer 5, located on the frame of the combine. Then gallery-driving cycle repeats.

A variety of proposals for robotics technology of mining and its high cost, the difference between the conditions at coalfield and mining, multivariate methods of work organization are available at the design stage while comparing the options of mining on computer models [1].

This work is supported by the Russian Foundation for Basic Research (Project 13-07-98023 r_ siberia_a).

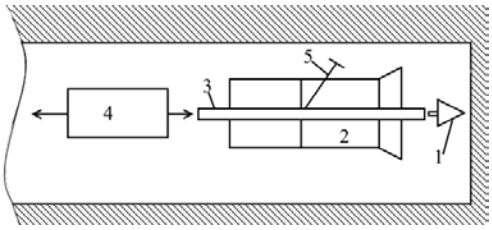


Fig. 2. Schemes of gallery driving by combine.

Most of the operations during mining are discrete with a finite number of variables. Such operations include the beginning and ending of drilling, loading and unloading of coal, the beginning and the end of a combine cycle.

It is exigent to use the mathematical apparatus of queuing systems (QS) for the simulation of such works [2].

The models of robotic technology of mining in the form of multi-channel multi-phase closed QS, where requests were the moments the equipment was ready for the next cycle, were developed.

Requests service implies their delay for the time of duration of processes of the gallery driving execution in devices that simulate the appropriate equipment. The duration of the process of technological cycle is displayed by entering random time delays in the devices QS. The length of the ongoing gallery driving is specified in the QS by the amount of the requests at the input of system (1)

$$m = L/l \quad (1)$$

where L - length of tunnel; l - face advance per cycle.

In actual practice equipment may start execution of the next gallery driving after the previous cycle, the time interval between the beginning and the end of the gallery driving depends on the random cycle time process operations. This feature is displayed in the model by the feedback input through another request, which is input after the QS application has been handled at the system output. Thus, feedback forms the input stream requests. Speed of the applications receipt is the speed of their services, therefore the

ISBN: 978-1-61804-327-6

queue on the entrance is not formed and the objectives of the study are to assess the overall service time of the application and utilization of the equipment.

The average duration of the gallery driving is defined in the model as the sum of random variables (2)

$$t_{\mu} = (\sum_{i=1}^{m} t_{di} + \sum_{i=1}^{m} t_{li} + \sum_{i=1}^{m} t_{si})/m \qquad (2)$$

where t_{di}, t_{li}, t_{si} - random values of the simulating equipment service time for a request, where the equipment simulates the devices for the destruction of the rock mass, rock mass load and support at the *i*-th cycle.

For example, the robotic technology of gallery driving by drilling and blasting method is presented in the form of a two-channel polyphase closed queuing system (Fig. 3).

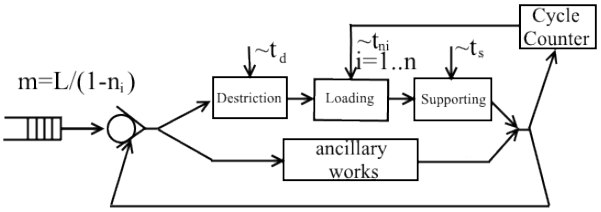


Fig. 3. Model of the robotic technology of production drilling and blasting as QS.

Request is the time when the equipment is ready for the next driving cycle, and devices are tunnelling machines, execution of requests for a random time; the rate of receipt of requests in the system is determined by the speed of their service. The request is served by robotic drilling rigs, loaders and support installer using one channel. Another channel is to perform ancillary works. At the end of the gallery driving the serviced request changes the value in cycle counter and allows unserved request to enter the system as an input. Service time requests devices are presented as functional dependencies (3)

$$\begin{aligned} t_d &= f(n_h, l_h, P^d, f_s, S, n^d) \\ t_l &= f(S, f_s, l_h, \eta, P^l, n^l, k_l, V) \\ t_s &= f(S, l_s, n_m) \end{aligned} \qquad (3)$$

where t_d, t_l, t_s - random variables of the time of destruction of the rock mass, rock mass loading and supporting; S - cross-section of heading; f_s- coefficient of hardness of rocks; n_h- number of holes per cycle; l_h - the length of the drill hole; η - utilization of holes; k_l - coefficient of loosening rocks; P^d, P^l - performance drilling and loading machines; n^d, n^l - number of drilling and loading machines; V - volume haulage unit; l_s - the length of the supported gallery driving during one cycle; n_m - number of miners busy in supporting.

To account for changes in trip times of haulage unit cycle counter that increases the delivery time, depending on the number of driving cycle (4) is mounted

$$t_{li} = f(l_d), i = (1, m) \qquad (4)$$

where l_d - the length of the path for delivery of the rock mass; m - the number of cycles required for gallery driving.

The current version of the simulation language GPSS - GPSS World best fits as the means of software implementation models of mining technology. GPSS language is one of the most effective and common software tools for modeling complex discrete systems on the computer and is successfully used for the simulation of various branches of industry [3, 4], including mining operations [5], formalized as queuing systems.

There are two basic types of objects: transacts and blocks, - in GPSS World. Blocks define the operation logic of the model and determine the path of transacts on it. Blocks analogous are QS devices showing the combine, drilling machines, loaders, and support installer. Transacts, moving from block to block, simulate mining works: snapping-distilled drilling rigs, load haulage unit, support etc. Transacts are analogous of QS requests to perform operations of the gallery-driving cycle. The modeling system transacts interact with blocks, resulting in a change of their attributes, as well as converted arithmetic or logical values. Such transformations are called events.

Using GPSS World model, we have developed software modules in order to build models of the excavation technology, "cut with loading", "drilling", "charging", "loading", "supporting". Fig. 4 shows an example of the module "loading."

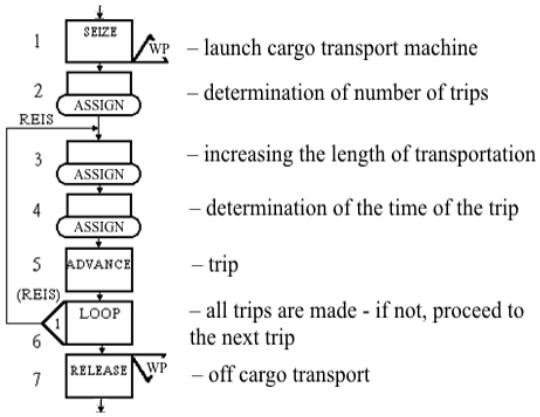


Fig. 4. Block diagram of module “Loading”.

In block 2 (ASSIGN) is determined by the desired number of trips for removal of separated rock mass (5).

$$\begin{aligned} Q &= Sl_h k_s \\ n &= Q/Vk_f \end{aligned} \qquad (5)$$

where Q - separate volume of the rock mass; S - cross-section of heading; l_h - the length of the drill hole; k_s - coefficient of softening; n - the number of required trips of haulage unit for removal of the separated rock mass; V – capacity of haulage unit; k_f - fill factor.

Duration of cargo transporting is modeled by delay in ADVANCE block (6).

ISBN: 978-1-61804-327-6

$$t_{Li} = 2\left(\frac{l_i}{v}\right) + t_l + t_{un} \quad (6)$$

where t_{Li} – duration of the i-th cycle of loading; l_i – length of gallery driven at i-th cycle v – velocity of haulage unit; t_l - the time of filling; t_{un} - unloading time.

Increasing the distance of transportation of rock mass l_i is determined in block 4 (ASSIGN) the expression (7)

$$l_i = l_{i-1} + l \quad (7)$$

where l_i - length of gallery driven at i-th cycle; l - face advance per cycle.

Turning off the cargo transport vehicle by block RELEASE occurs when all the trips that are defined by a block LOOP, which reduces the required number of flights per unit and checks the condition n_i=0, are performed. If not, it repeats the cycle of loading.

The developed model can be used to select the optimal volume of the loading and transport vehicles buckets; the maximum length of the rock mass transportation when the performance of the gallery driving is limited; the desired volume of haulage vehicles; the effectiveness of robotic technology and others evaluations.

For example, operating time for cargo transport machine depends on the distance from the place of unloading to the face. During the development, this distance increases with each cycle the value of face advance. The problem arises of assessing the impact of transport on the length of the key indicators of the gallery driving.

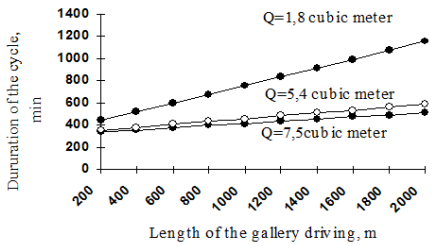


Fig. 5. Effect of the length of output per cycle at different bucket capacity of cargo transport machine.

In the mines several sizes of cargo transport vehicles with different volumes of the bucket (1 to 7.5 m^3) are used. To select a machine it is necessary to determine the dependence of the duration, the gallery driving on the capacity of cargo transport machine. Fig. 5 is a graph of the cycle time of the length of production for different bucket capacity of cargo transport machine obtained from the results of simulation experiments.

It was found during the development of robotic technology of individual machines that average cycle time increases linearly with the length of the production from 200 to 1800m. A reduction of capacity bucket loading and transport vehicles from 7.5 to 1,8m^3 increases the gallery driving to 24-56% depending on the length of the driven working.

CONCLUSION

Discrete event simulation displaying the mining technology as the QS, followed by software implementation on a specialized simulation language GPSS World allows comparing various robotic technologies and justifying their effective structure and parameters through the computer experiments. This opens up new choices of technological options and allows avoiding significant loss of investment during the design stage of the mine.

References

[1] V. Konyukh, "Simulation of mining in the future", in Proc. of the IASTED International Conference on Automation, Control, and Information Technology - Control, Diagnostics, and Automation, ACIT-CDA, Novosibirsk, Russia, June 15-18, 2010, pp. 1-6.

[2] Kleinrock L. "Queuing theory", Moscow, Russia, Mashinostroenie, 1979 (in Russian).

[3] V.A. Poletayev, V.V. Sinoviev, A.N. Starodubov, I.V. Chicherin, "Design of the computer integrated production systems" / under the editorship of V. A. Poletayev, Moscow, Russia, Mechanical engineering, 2011 (in Russian).

[4] A.N. Starodubov, V.V. Sinoviev, and M.U. Dorofeev "Simulation of the technological complex for deep coal processing", Ugol, 2010, № 2, pp. 8-12 (in Russian).

[5] V. V. Sinoviev, “Modeling multi face gallery tunneling using a simulation approach”, Mining Informational and Analytical Bulletin, Moscow, Russia, 2013, vol. 6, pp. 138-144.

Simulation of Coal Mining in Flat-Lying Coal Seam

Victor V. Okolnishnikov, Sergey V. Rudometov, Sergey S. Zhuravlev
Siberian Branch of the Russian Academy of Sciences
Design Technological Institute of Digital Techniques
Novosibirsk, Russia
okoln@mail.ru

Vasily V. Sinoviev
Siberian Branch of the Russian Academy of Sciences
Institute of coal
Kuzbass State Technical University
Kemerovo, Russia
zinovyev@kemsc.sbras.ru

***Abstract*—**The article describes a new specialized library of simulation models of mining machinery for coal mining in flat-lying coal seam. This library is based on the visual interactive discrete simulation system of technological processes. The article contains a brief description of this simulation system. A number of models of coal-mining in flat-lying coal seam developed with using of the specialized library are presented.

Keywords—visual interactive simulation; coal mining; longwall mining; highwall mining

I. INTRODUCTION

It is specified in different investigations, that situation in different areas of underground mining that requires automatic control (e.g. coal mining) is that resources become less available, and harder to be extracted.

The technologies of coal mining are well-known. Today mine uses mining machinery for mining, transportation, roof support, etc. How exactly will these machines function? What additional machines will they require? What is the cost of machines installed in mine? Will big universal machines be more appropriate rather than a lot of small, specialized mechanisms? What will the cost of a final product be? These tasks can be solved by methods of computer-based simulation of coal mining process.

The problem of the particular mining installation is that it is effective in terms of minimum costs and maximum productivity only if it is correctly planned. Usually big layout has many components from different vendors interconnected. That makes it hard or, even, impossible to predict the exact effectiveness. This situation gets worse if there is also a requirement to create new components for management of a part of such layout. In all these cases, computer-based simulation can be used to solve these problems.

Because of the importance of these problems, there are a large number of papers on the use of simulation in the development and optimization of coal mining systems [1, 2, 3, 4, 5, 6]. There are also a large number of simulation tools both universal simulation systems and specialized systems and packages for simulation of coal mining systems.

A number of models for various technologies of coal mining were developed with the help of own simulation system. Technologies of coal mining are presented in section 2. Brief description of the simulation system is presented in section 3. Sections 4, 5 contain description of a specialized library and examples of the models.

These models are used for developing of process control systems for underground coal mines in Kuznetsk Coal Basin (Russia, Western Siberia).

This work is supported by the Russian Foundation for Basic Research (Project 13-07-98023 r_ siberia_a).

II. TECHNOLOGIES OF COAL MINING

There are several well-known technologies of coal mining.

A. Longwall mining

Longwall mining system is a highly-automated, very powerful and productive way to mine a product. It is most-widely applied around the world. Its main advantage is that it leaves almost no product inside mines. But it is limited with the depth of the mine (measured from the surface). Also it is applied in relatively flat areas of coal, from 0.8 up to 10 meters high, from 150 to 450 meters face, and up to 4 kilometers in depth.

Longwall system consists at least of armored face conveyor (AFC), a shearer, and roof support sections. The AFC is connected to outbound belt conveyor. The shearer cuts the product from coal seam face, in a series of passes along the AFC. The AFC delivers product to the belt conveyor connected. The roof support moves itself and the AFC, pushing it (and itself) forward with hydraulics. One of the problems of longwall mining is roof caving. This can lead to serious environmental problems. The other problem is that longwall requires significant amount of work to be done before its massive equipment will be installed in production. This work also must be planned.

B. Highwall mining

In case when it is impossible to mine a product with longwall or other systems, a relatively new technology can be applied, named Highwall. This is a shearer tool, mounted on the top of the chain of special sections. These sections can be updated to each other, making a long (up to 300 meters) support chain for a shearer. Each section can transport the product developed by shearer to the end of a sections chain.

The shearer and the chain of sections cut the product (coal) from the very thin and curved seams. This is the main advantage of the Highwall technology.

C. Coal mining of flat-laying coal seam

This approach uses a number of front-cutting mining machines that cut a coal on a special scheme, and a number of self-moving wagons that move a coal from mining machines to the storage area. It solves the same problem as highwall mining, but requires no specialized equipment. Also this approach can be used in deep mining.

III. The Simulation System

A visual interactive Manufacturing and Transportation Simulation System (MTSS) is developed at Design Technological Institute of Digital Techniques of Siberian Branch of the Russian Academy of Sciences (DTIDT) [7, 8]. It is a process-oriented discrete simulation system intended to the development and execution of models of technological processes.

MTSS is a set of program interfaces for creating *elementary* models and for forming complex models from them. The elementary model is a ready-to-use submodel of an equipment unit with capability of low-level control for it.

The elementary model consists of the following parts:

- Two-dimensional and three-dimensional graphic images.
- Input and output parameters.
- Functionality algorithm describing dependence between parameters.
- States which the elementary model can reach during the simulation process.
- Control commands defining switching process between elementary models states.

A model in MTSS is created by graphical connection of images of elementary models.

MTSS is also a tool for running of complex models built from elementary models. The running model performs the movement of the model time and visualization. Statistics is collected as well. Statistics are available as a short overview when model runs, and more statistics are available after model completion.

This simulation system is effective in solving the task for the rapid creation of correct simulation model by mining engineers. Usually engineers have not enough qualification to create a simulation models in details, but they know how to connect correctly elementary models to create the required topology.

MTSS uses the 2D as the graphical editor and 2D, 3D for the visualization of model running. Such approach seems more natural for mining engineers, when all installations and machines appear first on 2D plans. 3D is more useful for visualizing complex vertical movement.

Process control systems often have two levels: the low level of equipment and simple control logic and the upper level of complex control of production. Therefore one of the distinguishing features of MTSS is a separation of the logic of simulation model into two parts: low-level logic and an upper level logic.

Such separation allows us not only to correspond to the usual structure of the process control systems but to use such models for embedding them into actual process control systems in the following ways: to emulate equipment, to simulate upper level logic, and to send commands to actual process control system for debugging and testing. This separation into upper and lower logics allows also organizing a switch between various implementations of the decomposition. It allows coexisting simulation of upper level logic and a proxy that allows communicating with the upper level logic of actual process control system.

The model of coal mining can communicate with a new process control system developed in DTIDT, to be a source of input signals, emulate equipment, test actual control program with simultaneous visualization of overall process of mining.

This allows debugging and tuning of a new process control system in accordance with behavior of simulated system, even allows simulating various accidents. This allows minimizing time and costs on site for commissioning.

IV. Simulation of Coal Mining in Flat-Lying Coal Seam

A specialized library of simulation models of mining machines for coal mining was developed. This library is a part of MTSS and its prime goal is to simulate interactively and visualize various aspects of coal mining in flat-lying coal seam. The library consists of new elementary models of:

- Highwall mining system.
- Longwall mining system.
- Coal seam.
- Mining machines.
- Self-moving coal wagon.
- Storage area.

The library contains also a simulation model of a flat-lying coal seam. This model is a source of the product in a simulation model, while storage area is a consumer of a product. The product itself is coal.

These new components can communicate with existing libraries of MTSS [9, 10, 11, 12] which simulate mines subsystems like:

- Belt conveyor subsystem.
- Power supply subsystem.
- Ventilation subsystem.

A simulation model of coal mining subsystem in a flat-lying coal seam was developed using the library of mining machines. Fig. 1 contains a sample layout in the simulation model of the flat-lying coal seam (2D and 3D view combined in different views).

ISBN: 978-1-61804-327-6

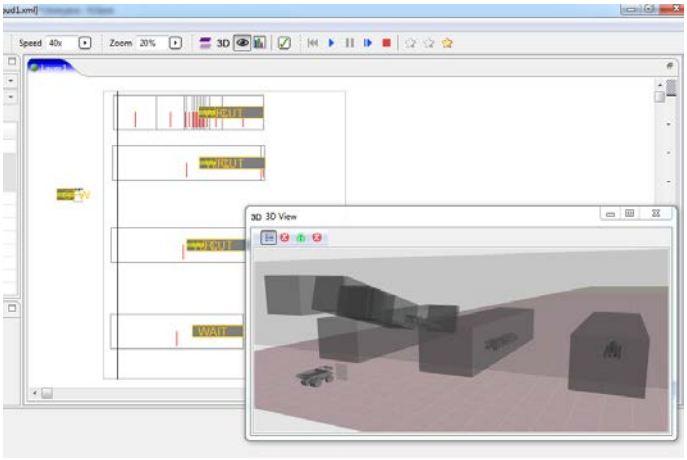


Fig. 1. Simulation model of coal mining in flat-lying coal seam (4 mining machines and 4 self-moving coal wagon).

The simulation model built from the content of the library will simulate movement of all mobile objects of the model. Both 2D (top view) and 3D visualization are available. Statistical data is also collected.

Fig. 2 shows a sample layout in another simulation model of the flat-lying coal seam. Main window consists of 6 areas:

1) 2D top view. It contains: main mine, side mines. The origin coordinates is at the left top corner of this view.
2) 3D view. On Fig. 2 there is a view from point *3)*.
3) Point of view for the 3D.
4) Parameters of a simulation model.
5) Specialized view for fast navigation in simulation model.
6) Settings for the time start, time end and current model time.

The model was created to visualize new technology of coal-mining in flat-lying coal seam [13]. This simulation model also allows receiving statistical results for usage of this technology.

As patent describes, mining is done by frontal winning machine, paired with self-moving coal wagons. Patent describes the directions of cut of a flat-lying coal seam for the frontal winning machine. Simulation model of this process allows visualizing the process described.

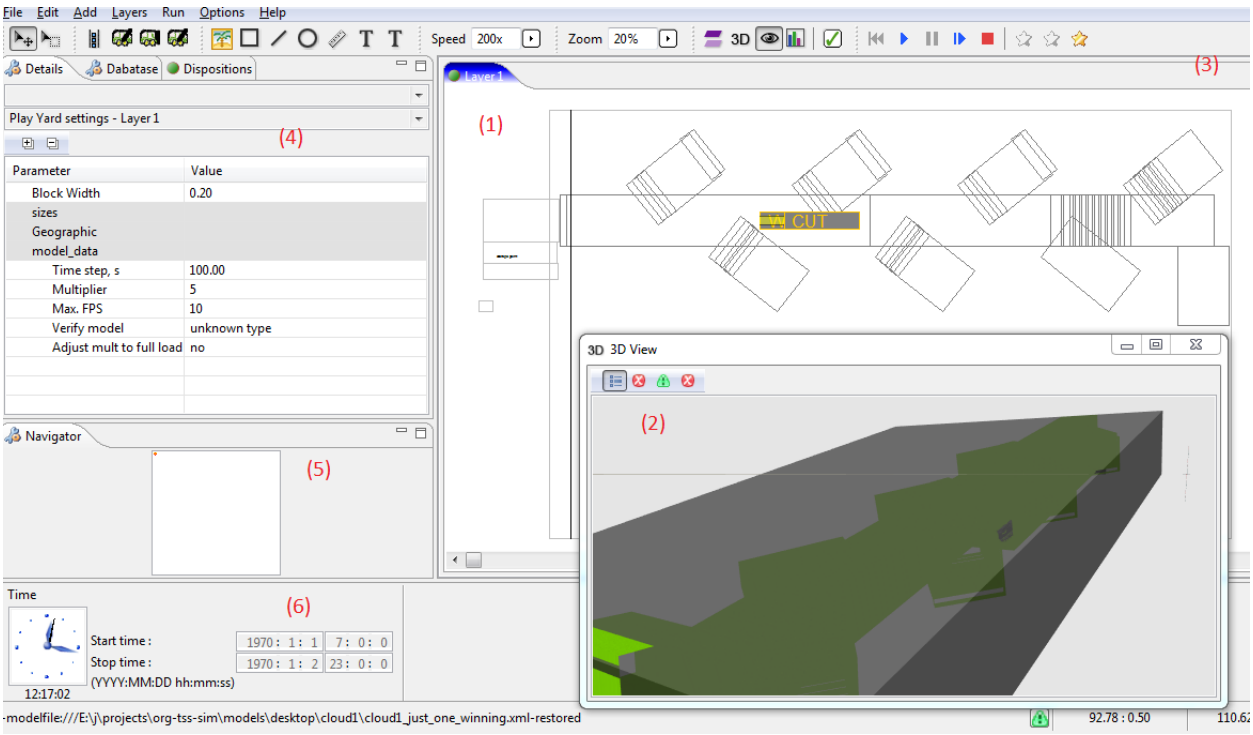


Fig. 2. Main window of MTSS system while running of simulation model of coal mining in flat-lying coal seam.

V. Simulation of Longwall Mining

For detailed simulation of longwall mining system we finished with next decomposition:

- Armored Face Conveyor (AFC).
- Shearer.
- Roof support sections.

Simulation model for longwall mining system can function only if it is connected (in terms of MTSS) with belt conveyor simulation model.

The algorithm for longwall mining system simulation is done closer to the control algorithms of the real longwall mining system. The task will define the amount of time for the

ISBN: 978-1-61804-327-6

shearer to move to the state while it will allow one roof support section to be free to move, and an amount of a product to be mined. Then it will advance the simulation to the time defined, and then simulate the roof support sections movement. The amount of a product will be moved to the AFC that will deliver it to the belt conveyor. Task will repeat these steps until the “done” or “postpone” conditions will be achieved.

Task is “done” when shearer reaches the end of an AFC line. Task is “postponed” when belt conveyor is overloaded and cannot accept the next portion of product or gas level is not safe.

During any of these steps, the simulation model of flat-lying coal seam will simulate roof fall (behind the roof support) and gas level increasing. Also ventilation simulated (remove gas from working area).

In most cases (as it is shown in our previous works), there is no need to simulate in details any technology like longwall or flat-lying coal mining, if it is used just as a source of a product for belt conveyor system, for example. All that is really needed in such cases are:

- To define that longwall or flat-lying coal mining is a source of a product for a big mining system like conveyor.
- To know the performance of the longwall coal mining during some time period (working day, 8-hours time interval, 1-hour time interval). Note that the emergency stops (like gas or coal dust) are already included in this statistical data.

But if the goals of simulation are:

- To predict how longwall or flat-lying coal seam mining automation will behave in details (i.e. movement of its parts depending on various situations in mines).
- To make a detailed visualization of mining process.
- To define how this mining will impact to the overall performance of the mines.
- To define scenarios of broken or temporarily inaccessible parts interactively.

Then achieving these goals will require detailed decomposition of common longwall (or highwall) mining system and a detailed visualization of all its parts.

In our work both detailed and statistical approaches are presented.

VI. FUTURE WORK

Detailed simulation of longwall system, connected with detailed simulation of flat-lying coal seam (or multiple flat-lying coal seams), will allow creating of simulation that will not only predict the behavior of big underground mining system, but also simulation of land subsidence while using longwall, especially with very heavy longwall systems that can cut 10-meters-high coal seams.

Simulation system MTSS can be used not only for simulation of existing coal mining techniques but also for perspective robotized techniques.

References

[1] J.R. Sturgul, "Modelling and Simulation in Mining - Its Time Has Finally Arrived", Simulation, 76(5), 2001, pp. 286-288.

[2] M. S. Kizil, A. McAllister and R. Pascoe, "Simulation of Development in Longwall Coal Mines", in Proc. of the 11th Underground Coal Operators' Conference, University of Wollongong & the Australasian Institute of Mining and Metallurgy, 2011, pp. 91-98.

[3] J. Greberg, F. Sundqvist, "Simulation as a tool for mine planning", in Proc. of Second International Future Mining Conference, 2011, pp 273-278.

[4] D. Cai, E. Baafi & I. Porter, "Modelling a longwall production system using flexsim 3D simulation software", in R. Singhal, E. Topal, K. Fytas, M. Yellishetty & A. Mehrotra (Eds.), Mine planning and equipment selection, 2012, pp. 107-114.

[5] C. Meng, S.S. Nageshwaraniyer, A. Maghsoudi, Y. Son, and S. Dessureault. “Data-driven Modeling and Simulation Framework for Material Handling Systems in Coal Mines.” Computers & Industrial Engineering, 2013, pp. 766-779.

[6] M. Fioroni, L. Santos, L. Franzese, I. Santana, G. Telles, J. Seixas, B. Penna and G. Alkmim, "Logistic Evaluation of an Underground mine Using Simulation", in Proceedings of the 2014 Winter Simulation Conference, 2014, pp. 1855-1865.

[7] V. Okolnishnikov, S. Rudometov, and S. Zhuravlev, “Simulation environment for industrial and transportation systems”, in Proc. of the International Conference on Modelling and Simulation, Prague, Czech Republic, 2010, pp. 161-165.

[8] S. Rudometov, “MTSS simulation system (Patent style),” FAP SB RAS, Novosibirsk, Russia, 2011, URL: http://fap.sbras.ru/node/2325 (In Russian).

[9] V. Okolnishnikov, S. Rudometov, and S. Zhuravlev, “Monitoring System Development Using Simulation”, in Proc. of the 2010 IEEE Region 8 International Conference, Irkutsk, Russia, July 11-15, 2010, vol. II, pp. 736–739.

[10] Victor Okolnishnikov, "Development of Process Control Systems with the Use of Emulation Models", International Journal of Mathematics and Computers in Simulation, issue 6, vol. 5, 2011, pp. 553 – 560.

[11] V. Okolnishnikov, S. Rudometov, and S. Zhuravlev, “Simulation Environment for Development of Automated Process Control System in Coal Mining”, in Proc. of the 2013 International Conference on Systems, Control, Signal Processing and Informatics (SCSI 2013), Rhodes Island, Greece, July 16-19, 2013, pp. 285–288.

[12] Victor Okolnishnikov, Sergey Rudometov, and Sergey Zhuravlev. Simulation Environment for Development of Automated Process Control System in Coal Mining // International Journal of Systems Applications, Engineering & Development. Issue 5, vol. 7, 2013. pp. 255–262.

[13] V. Fedorin, V. Shachmatov, B. Anferov, L. Kuznetsova, "The method of open-underground mining of thick flat coal seam", Patent for invention of Russin Federation № 2490456, 2013 (in Russian).

Comparison of ACO and GA Techniques to Generate Neural Network Based Bezier-PARSEC Parameterized Airfoil

Waqas Saleem
School of Mechanical and Manufacturing Engineering
National University of Sciences and Technology
Islamabad, Pakistan
waqas_jeral@hotmail.com

Riaz Ahmad
CIE Building, Research Directorate
National University of Sciences and Technology
Islamabad, Pakistan
dresearch@nust.edu.pk

Athar Kharal
College of Aeronautical Engineering
National University of Sciences and Technology
Islamabad, Pakistan
atharkharal@gmail.com

Ayman Saleem
College of Aeronautical Engineering
National University of Sciences and Technology
Islamabad, Pakistan
aimen1173@hotmail.com

***Abstract*—** This research uses Neural Networks to determine two dimensional airfoil geometry using Bezier-PARSEC parameterization. Earlier, Ant Colony Optimization (ACO) techniques have been used to solve combinatorial optimization problems like TSP. This work extends ACO method from TSP problem to design parameters for estimating unknown Bezier-PARSEC parameters that define upper and lower curves of the airfoil. The efficiency and the performance of ACO technique was compared to that of GA. The work established that ACO exhibited improved performance than the GA in terms of optimization time and level of precision achieved. In the next phase, Neural Network is implemented using Cp as input in terms of C_l, C_d and C_m for learning and targeting the corresponding Bezier-PARSEC parameters. Neural Networks including Feed-forward back propagation, Generalized Regression and Radial Basis were implemented and were compared to evaluate their performance. Similar to earlier work with GA and Neural Nets, this work also established Feed-forward back propagation Neural Network as a preferred method for determining the design of airfoil since the technique presented better approximation results than other neural nets.

***Keywords*—** Airfoil Optimization; Ant Colony Optimization; Bezier-PARSEC; C_p; Neural Network

I. Introduction

Airfoil design is one of the most challenging processes [1] in development of aircraft aerodynamic surfaces as it affects various aircraft performance parameters like lift, drag, spin-stall, cruise and turning radius [2]. Studies indicate that selecting the right design of airfoil with required characteristics reduces overall cost and improves the performance of air vehicle. Airfoil design largely depends on desire for high lift to drag ratio that is in conflict with the performance requirements [3].

There are two major techniques for designing an airfoil; direct and inverse [4]. First method involves designing a new or modifying an existing airfoil (UIUC Airfoil Database [5] and computing pressure distribution across the surface to achieve desired set of parameters. This approach may limit the approximation for desired specifications due to inherent limitations in airfoil's aerodynamics. For faster approximations, reduced degrees of freedoms are required but such reduction results in computational errors like round off, truncation and discretization error. In fact, determining the airfoil geometry should be based on requirements for aircraft's performance. Thus later method involves using desired operational characteristics and performance parameters unless the airfoil geometry so generated meets the desired criteria. To reduce the computational time and meet the required design criteria various techniques including CFD, fuzzy logic, neural networks [6] and heuristics based algorithms like PSO [7] and GA [8] have been implemented to advantage the aerodynamic design process.

This research, largely inspired by Saleem and Kharal [9], uses neural network based approach for airfoil generation exploiting Bezier-PARSEC 3434 parameterization rather than full coordinates for a given Cp. However, this research implements ACO to optimize Bezier-PARSEC unknown parameters instead of GA as in earlier work.

II. Artificial Neural Network

In machine learning and data mining, Artificial Neural Network is a set of learning algorithms modeled after neural network structure of the cerebral cortex and is used to approximate functions involving a larger number of the unknown input variables [10] Each neuron receives input from external sources or neighbors in the network, computes output and propagates to other neurons. Another function is the weight adjustments in the connections between neurons. Incremental learning is the technique by gathering information on cumulative error and consequently adjusting weight coefficients, w_{ij}. Mathematically, a Neural Network can be defined as a triple (N, C, w) where N is the set of neurons, C $\{(i, j)|i, j \in N\}$ is a set of connections, and function w((i, j)), shortened as w_{ij} is called weights between neurons i and j. For every neuron, there is an external input ϑ_j and an activation function F_j to establish the new activation level based on effective input of a neuron S_j and is determined by following propagation rule in "(1)".

$$S_j = \sum_i w_{ij}\ (t)\ y_i\ (t) + \theta_j\ (t) \quad (1)$$

Besides, a threshold is also introduced as linear, non-linear or sigmoidal function [11] that helps avoid the situation when training is not successful at ||σ||>0. A threshold function for each neuron is given by "(2)"

$$F_{S_j} = \frac{1}{1 + e^{-S_j}} \quad (2)$$

A. Feed-forward Back propagation Neural Network

A feed forward Back Propagation Neural Network (FFBP) contains a multi layered interconnected feed forward structure where every layer gets input from below and gives output to layer above it. Back propagation is a learning technique where output values are compared to a desired value to calculate the error using a pre-determined error function. This value of error is then fed back through the network repeatedly for minimizing through neural network algorithm by adjusting weights for each network connection until the network converges to a bare minimum acceptable level of error [12] Generally, a non-linear optimization method called gradient descent is implemented where derivative of the error function is determined w.r.t. weights, that are adjusted till the reduction of error.

B. Radial Basis Function Neural Network

A Radial Basis Function Neural Network (RBF) consists of an input layer, a hidden layer with non-linear Radial Basis activation function and an output layer. For Radial Basis Neural Network, the input is modeled as vector of real numbers (R^n) while output is a scalar function φ, given in "(3)" by [13]

$$\boldsymbol{\varphi(x)} = \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{a_i p(||x - c_i||)} \quad (3)$$

where n is number of neurons, a_i is weight of neuron and c_i is center vector.

In Radial Basis Neural Networks, neurons respond to inputs close to their center in contrast with other neural networks. Although Radial Basis Neural Network requires more neurons for high dimensional input spaces, it can be trained faster than standard multi layered neural networks and have proven efficiency in regression and classification problems.

C. Generalized Regression Neural Network

A Generalized Regression Neural Network (GRNN) consists of one each input layer, pattern layer, summation layer and output layer. Training patterns are presented by neurons in pattern layer. In GRNN, pattern layer is connected to summation layer. Sum of weighted responses and un-weighted responses of pattern neurons are computed by two neurons in summation layers [9] The summation layer consists of both summation and single division units. Normalization of output is performed together both by summation and output layers. GRNN exhibit single pass learning algorithm with high parallel structure for estimating continuous variables and do not require iterative process as in multi-layered networks. GRNN converges to optimal regression even in noisy environments given a large number of sample data is available. Generalized Regression Neural Network is particularly advantageous with sparse data but as the training data increase, the error converges to zero.

III. PARSEC PARAMETERIZATION & BEZIER CURVES

PARSEC parameterization has the capability to describe the airfoil shape and its flow using engineering parameters [10] On the other hand, a Bezier curve is a parametric curve of degree n defined by polygon of n+1 vertex points called control points of nth order Bezier curve and is given by "(4)"

$$P(t) = \sum_{k=0}^{n} P_k \binom{n}{k} t^k (1-t)^{n-k} \quad (4)$$

where P_k is the kth control point while parameter t ranges from 0 to 1 with 0 at the zeroth control point and 1 at the nth control point. Eq. (5) gives Third order Bezier Curve

$$\begin{cases} x = x_a\ (1-t)^3 + 3x_b(1-t)^2 t + 3x_c(1-t)t^2 + x_d t^3 \\ y = y_a\ (1-t)^3 + 3y_b(1-t)^2 t + 3y(1-t)t^2 + y_d t^3 \end{cases} \quad (5)$$

Eq. (6) present fourth order Bezier Curve

$$\begin{cases} x = x_a\ (1-t)^4 + 4x_b(1-t)^3 t + 6x_c(1-t)^2 t^2 + 4x_d(1-t)t^3 \\ y = y_a\ (1-t)^4 + 4y_b(1-t)^3 t + 6y_c(1-t)^2 t^2 + 4y_d(1-t)t^3 \end{cases} \quad (6)$$

IV. BEZIER-PARSEC PARAMETERIZATION

Bezier-PARSEC parameterization is a technique in which Bezier Curves are described using PARSEC parameterizations [14] and is further subdivided into BP3333 and BP3434.

A. BP3333 Parameterization

BP3333 Parameterization employs third order Bezier Curves for camber shape and thickness of airfoil [15] Twelve PARSEC parameters represent Bezier control points as shown in Fig 1.

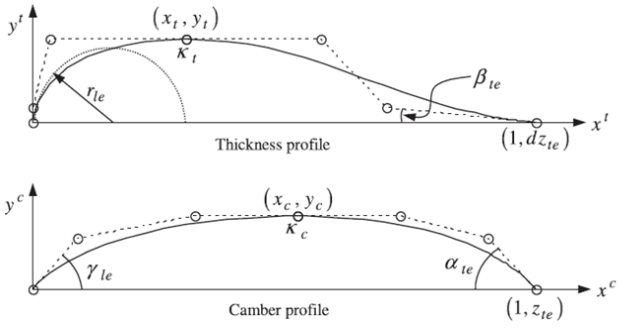


Fig 1 : BP3333 Bezier PARSEC Control Points and Respective Airfoil Geometry

Main advantages of BP3333 include close relevance to airfoil aerodynamics parameters, faster optimization, continuity characteristics, reduced deviation of design process and avoidance of sharp leading edges. Disadvantage of this technique is reduced degree of freedom resulting in failure to parameterize airfoils having radical camber trailing edge

* Corresponding author. Tel.: +92 322-4362442
E-mail addresses: waqas_jeral@hotmail.com (W.Saleem), dresearch@nust.edu.pk (R. Ahmad), atharkharal@gmail.com (A. Kharal), aimen1173@hotmail.com (A. Saleem)

B. BP3434 Parameterization

BP3434 Parameterization depends on 10 PARSEC parameters and 5 Bezier parameters for airfoil shape representation. Here, camber and thickness leading edge of airfoil is defined by third order Bezier Curves while fourth order Bezier Curves are used to define camber and thickness trailing edge of airfoil shape [15] This allows increased degree of freedom for trailing edge parameterization of airfoil as shown in Fig 2.

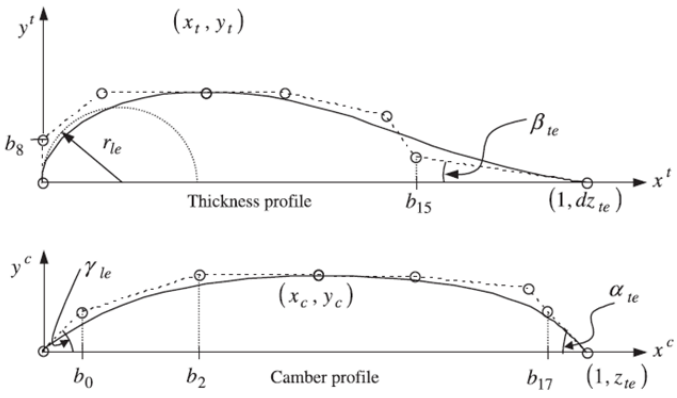


Fig 2 : BP3434 Bezier PARSEC Control Points and Respective Airfoil Geometry

BP3434 proves to be efficient than BP3333 when airfoil camber becomes negative along any part of the chord length. However, the convergence speed for this method reduces due to greater number of variables as compared to BP3333. In presence of high computing numerical computers, the convergence speed of BP3434 can be compensated for effective application of the method.

V. Ant Colony Optimization

Ant Colony Optimization (ACO) is the meta-heuristic motivated from the working of natural ants that suggests that ants follow different paths to reach food source initially. Thus the ants with shortest path would reach the source in least time than the longer paths [16] Ant Colony Optimization is an algorithm where artificial ants are used to probabilistically construct solutions guided by higher pheromone trails and promising heuristic information [17] In actual, ants implement a randomized construction heuristics that differ from greedy heuristics by adding a probabilistic component to partial solution than a deterministic one. Generally, ACO algorithm consists of two phases. In first phase, artificial ants construct a solution where in second phase, pheromone trail is updated by first reducing by an evaporation factor to avoid unlimited accumulation followed by adding pheromone proportionate to quality of their solutions [18]. Thus most important is to update pheromone for generating quality solutions in future iterations of algorithm. ACO algorithms can be considered as competitive solution technique where previous solutions known to be part of good solutions are used to generate even better solutions in future cycles [19].

VI. Methodology

In this research work, our methodology was quite similar to earlier work; however, ACO was preferred as a choice for optimization technique instead of GA to determine unknown Bezier-PARSEC parameters.

A. Airfoil Representation

A vector of 71 points is used to represent x-y coordinates of an airfoil where x_i ranges from 1 to 0 for upper airfoil curve and lower airfoil curve, thus only values for y change which determine the shape of both curves.

Mean Camber Line is a line at equal distance from both upper and lower surfaces of airfoil. Therefore, camber curve y points were obtained by taking average of upper and lower coordinates corresponding to the same x coordinate. These upper and lower coordinates were divided by chord length for non-dimensionalizing. The camber profile of an airfoil is calculated by "(7)", "(8)", "(9)" and "(10)"

c-$|x_1$-$x_{36}|$ for i=1 to 36 and j=36 to 71

$$y_i^u = \frac{y_i}{c} \quad \text{and} \quad y_j^l = \frac{y_j}{c} \qquad (7)\ \&\ (8)$$

$$x_i^c = \frac{x_i}{c} \quad \text{and} \quad y_i^c = \frac{y_i^u + y_j^l}{2} \qquad (9)\ \&\ (10)$$

Thickness curve used to define the airfoil thickness is the difference between the camber curve and upper curve of the airfoil i.e.

$$y_i^t = y_i^u - y_i^c \qquad (11)$$

Next a two dimensional analysis of airfoil was carried out using Panel Method to obtain values for lift coefficient C_l quarter-chord pitching moment coefficient C_m and drag coefficient C_d at ten angles of attack α. Thus the airfoil would be represented by x_i^c, y_i^c, C_l, C_d, C_m *and* α.

B. Calculating Bezier-PARSEC Parameterization

Table 1 presents the required parameters for Bezier-PARSEC

TABLE I : Known Bezier-Parsec Parameters

Parameters	Calculations
Maximum Thickness Point	$y_t = C^t(\{x_t \left\|\frac{dC^t}{dx_i}\right\|_{x_i - x_t} = 0\})$
Maximum Camber Point	$y_c = C^c(\{x_c \left\|\frac{dC^c}{dx_i}\right\|_{x_i - x_c} = 0\})$
Trailing Edge Vertical Displacement	$Z_{te} = C^c(x)\|_{x=1}$ and $dZ_{te} = C^t(x)\|_{x=1}$
Trailing Camber Line Angle	$\propto_{te} = -tan^{-1}(\frac{dC^c}{dx}\|_{x=1})$
Trailing Wedge Angle	$\beta_{te} = -tan^{-1}(\frac{dC^t}{dx}\|_{x=1})$
Leading Edge Direction	$\gamma_{le} = tan^{-1}(\frac{dC^c}{dx}\|_{x=0})$
Leading Edge Radius	r_{le}

While ten parameters are calculated using Bezier-PARSEC equations, there is no specific mathematical expression for finding remaining five parameters i.e., b_0, b_2, b_8, b_{15} and b_{17} and therefore are calculated by curve fitting. Since actual airfoil is known, Bezier Curves with correct five

control points would suffice given a smallest Sum-of-Least-Square Error. Table II shows the four curves and corresponding unknown Bezier points.

TABLE II : UNKNOWN BEZIER-PARSEC PARAMETERS

Curve	Bezier Curve	Order	Unknown Bezier Control Points
Camber	Leading Edge	3rd	b_0, b_2
	Trailing Edge	4th	b_{17}
Thickness	Leading Edge	3rd	b_8
	Trailing Edge	4th	b_8, b_{15}

C. *Optimization of Unknown Bezier Control Points Using ACO*

To determine optimal value of these unknown parameters, Ant Colony Optimization was implemented requiring fitness functions for each Bezier Curve that was equal to the difference between Bezier generated and actual airfoil. For this a Simple ACO code was written to determine each of these parameters i.e., b_0, b_2, b_8, b_{15} and b_{17}. In ACO, 6 ants were used to determine the optimal path to the destination and since the destination point was unknown; therefore, SSE for each curve was calculated for each generated point. Thus, a decrease in SSE over the path indicates that the ant is close to the destination point and vice versa. The pheromone is inversely proportional to the distance so the path with least distance or least SSE would have maximum pheromone. For each value of b_0, a corresponding value of b_2 is calculated through ACO. Thus a number of combinations (pair of b_0 and b_2 values) are made where pair with the least SSE is finally chosen. Same approach was used for b_8 and b_{15} while value of b_{17} was calculated separately. Fig 3 present flow charts for the method used.

D. *Error Calculation*

All 15 BP3434 parameters determined are used for generation of airfoil geometry. The error is calculated by comparing Bezier generated airfoil with actual airfoil. To calculate this error, at a certain x-value, y-value from parameterized and actual airfoils should relate to this x-value. The main challenge was to determine y-values of Bezier parameterized airfoil corresponding to these x-values. After generating x and y values of trailing and leading edge of thickness curve, these are arranged into a single set of x-y array in which first element corresponds to leading edge followed by trailing edge. Then cubic spline interpolation is used to fit a curve in the vector of x and y values which is then evaluated for 36 x-values of actual airfoil. Same procedure was followed for camber curve. These thickness and camber curves can be used to determine the shape of airfoil. The airfoil geometries of parameterized and actual airfoils are then plotted against same axis for comparison.

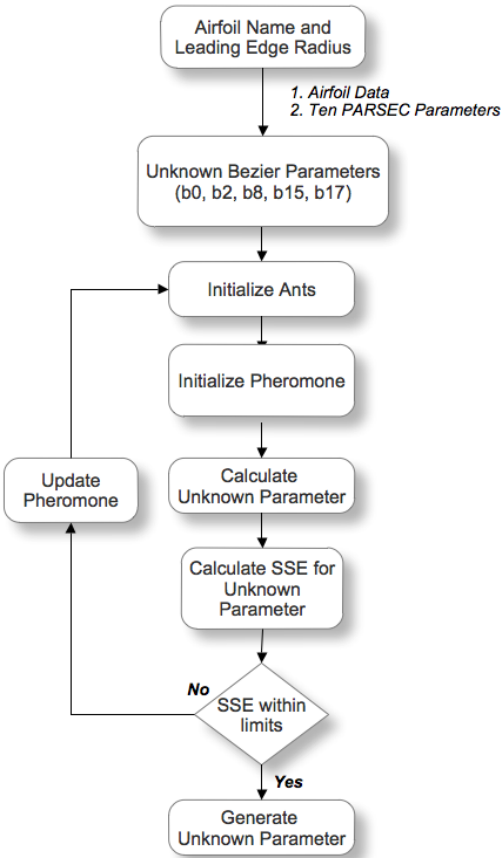


Fig 3 : Ant Colony Optimization Methodology

Fig 4 shows flow chart for SSE calculations while Fig 5 presents results for Eppler 433 sailplane parameterized airfoil.

E. *Neural Networks Estimations*

Neural Networks of three types as discussed in Section 2 were implemented in this research work. A 10X4 matrix of Cl, Cd, Cm at ten angles of attack for 500 heterogeneous airfoils was input to neural network while target was 15 Bezier-PARSEC parameters for airfoil generation.

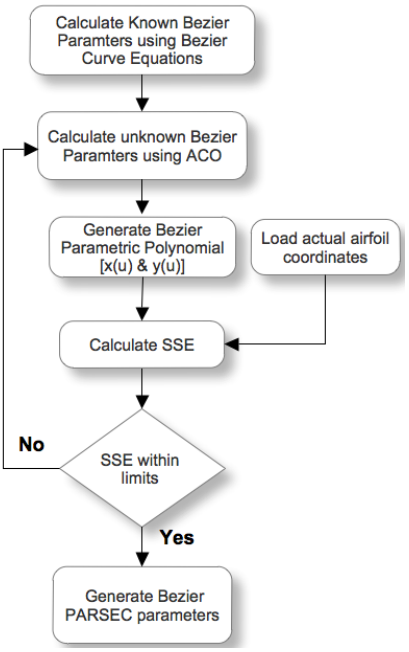


Fig 4 : SSE Calculations for Bezier-PARSEC Parameters

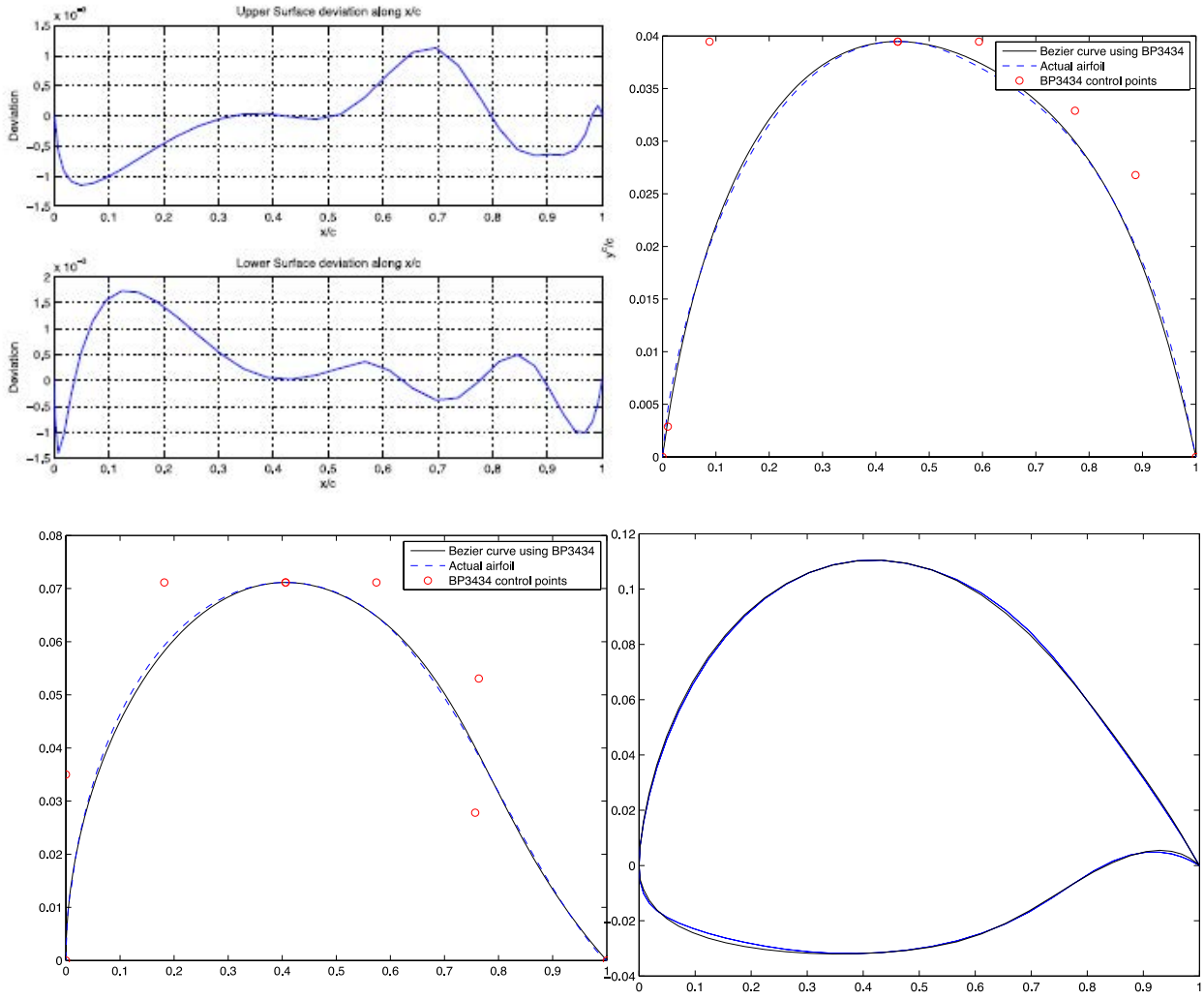


Fig 5 : Bezier Parameterization Results for Eppler 433 Sailplane Airfoil

VII. Results And Discussions

A. *Comparison of ACO Results with GA*

Implementation of ACO for finding unknown Bezier Curve parameters proved to be more efficient than Genetic Algorithm. We were able to achieve a precision level of $\leq 1X10^{-5}$ as compared to GA based version of the program. Also time to optimize the missing BP3434 parameters was greatly reduced. For example, Eppler 433 Sailplane airfoil took 30.905144 seconds to optimize BP3434 missing parameters using ACO as compared to GA that took 87.869966 seconds for optimization of said airfoil using 2.7GHz Processor and 4GB RAM. Table III gives a comparison of ACO and GA optimizations for few airfoils for reference.

TABLE III : Comparison of ACO and GA Optimization Results

Airfoil	Ant Colony Optimization		Genetic Algorithm	
	Time (Seconds)	Level	Time (Seconds)	Level
Eppler E433	30.905144	$\leq 1X10^{-5}$	87.869966	$\leq 1X10^{-4}$
NACA 65(4)-421	55.187357	$\leq 1X10^{-5}$	90.952194	$>1X10^{-4}$
Eppler E335	65.389595	$\leq 1X10^{-5}$	109.694796	$>1X10^{-4}$
Gottingen GOE426	44.489090	$\leq 1X10^{-5}$	82.259980	$\leq 1X10^{-4}$
Eppler E399	55.089536	$\leq 1X10^{-5}$	94.445729	$\leq 1X10^{-4}$

From Table III, we see that optimization time has remarkably been reduced to almost half for above airfoils.

B. *Results of Neural Networks*

As discussed above, three types of neural networks were implemented and tested against 500 airfoils for training and 200 airfoils unknown to the neural nets. Consolidated results for these airfoils is shown in Table IV.

The results from Table IV show that Feed Forward and Back Propagation has proved to be more promising in terms of better performance as indicated by increased fraction of both known and unknown airfoils within acceptable MSE values. On the other hand, GRNN and RBF showed improved efficiency with known airfoils than for the unknown airfoils. Comparison of Results for a known to network airfoil (Eppler 399 airfoil) and an unknown to network airfoil (Gottingen 426 airfoil) to the three types of neural networks is shown in Fig 6

ISBN: 978-1-61804-327-6

The plots for Gottingen 426 airfoil and Eppler 399 airfoil support application of Feed Forward Back Propagation Neural Network for solving this problem. However, results from RBF and GRNN largely favour known to network airfoils than unknown airfoils as is evident from RBF and GRNN plots for Gottingen 426 airfoil. Results for 200 airfoils unknown to network also support similar findings. MSE for GRNN and RBF is higher than FFBP with RBF performing the worst with a high MSE.

TABLE IV : COMPARISON OF TEST RESULTS FOR THREE NEURAL NETS

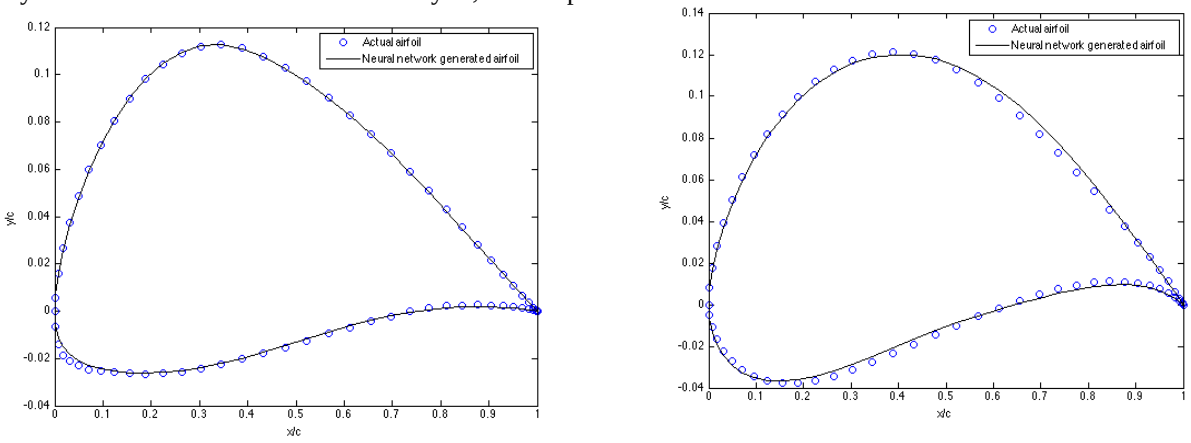
Artificial Neural Network	$\leq 1X10^{-5}$		$\geq 1X10^{-5}$	
	Count	%age	Count	%age
Feed Forward and Back Propagation				
500 Known Airfoils	273	54.6	227	45.4
200 Unknown Airfoils	113	56.5	87	43.4
Radial Basis Neural Network				
500 Known Airfoils	394	78.8	106	21.2
200 Unknown Airfoils	47	23.5	153	76.5
Generalized Regression Neural Network				
500 Known Airfoils	363	72.6	137	27.4
200 Unknown Airfoils	78	39.0	122	61.0

C. *Regression Analysis*

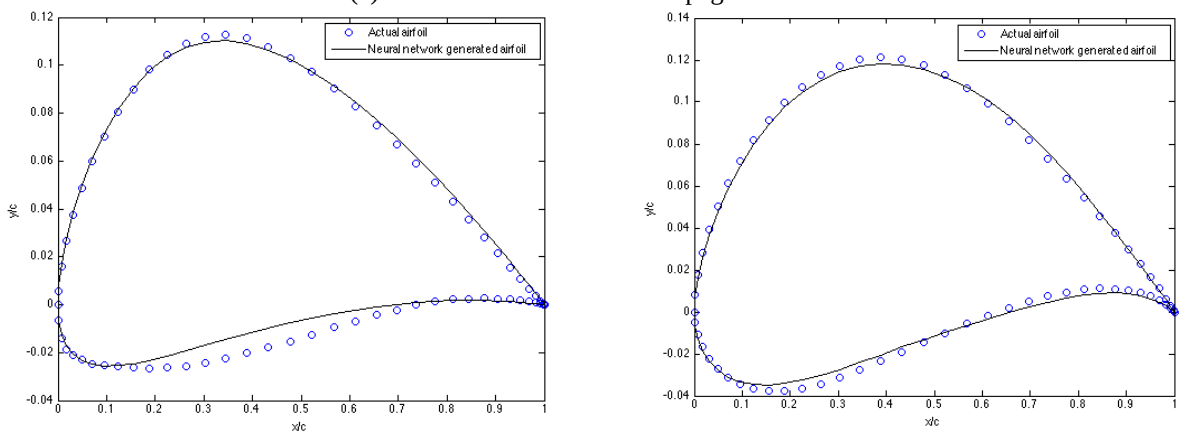
A post training regression analysis was performed to analyze the neural networks. In this analysis, the output of neural networks for known targets was compared. Thus neural network output would match the target values and would ideally be a straight line with 45° slope passing through the origin as shown in Fig 7. Fig 7 shows that performance of Feed Forward and Back Propagation is better than other two types of neural nets as indicated by the high regression values and low training R-values. On the other hand, both Generalized Regression Neural Network and Radial Basis Neural Network have higher R-values but shown poor results with test and validation data. The main reason is their architecture as both determine distance between input and weight vectors, which are incrementally multiplied by biased vectors. This would lead an input close to weight vector, produce an output close to unity while output would be close to zero if input is different from weight vector.

VIII. CONCLUSION

This work determines airfoil geometry for a given C_p using Neural Network and Bezier-PARSEC parameters. The main consideration of this paper is to use Ant Colony Optimization technique to optimize missing BP3434 parameters instead of complete set of airfoil coordinates. Further, three types of Neural Networks; Feed Forward and Back Propagation, Radial Basis and Generalized Regression were employed. Similar to earlier findings with GA based code, we proved that Feed-forward and Back Propagation exhibited greater efficiency than the other two types of Neural Networks. However, we were able to achieve higher precision with reduced time for optimization using ACO to determine missing BP3434 parameters. Besides, percentage of known and unknown airfoils with precision $\leq 1X10^{-5}$ has shown a slight increase.

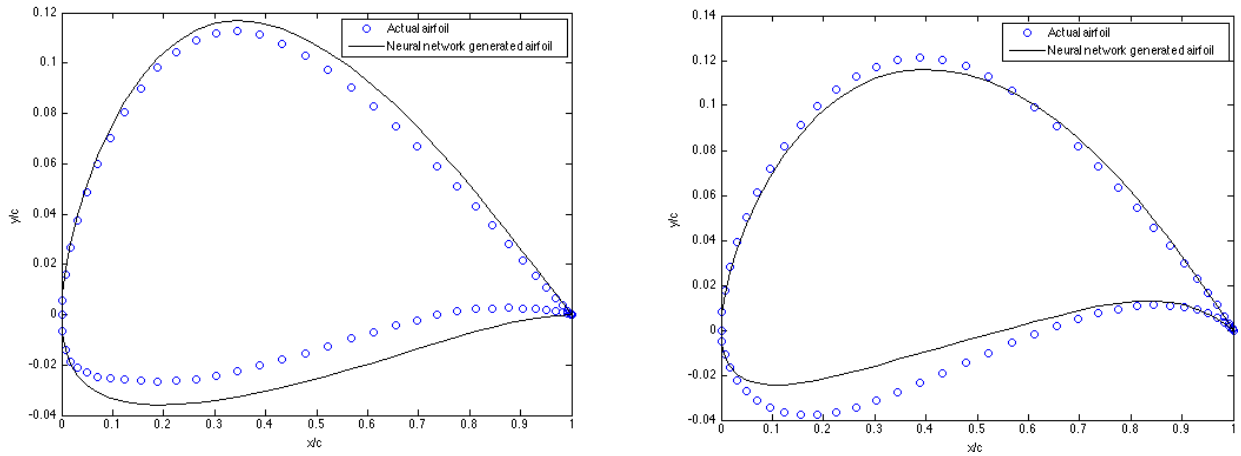


(a) Feed Forward and Back Propagation Neural Network



(b) Generalized Regression Neural Network

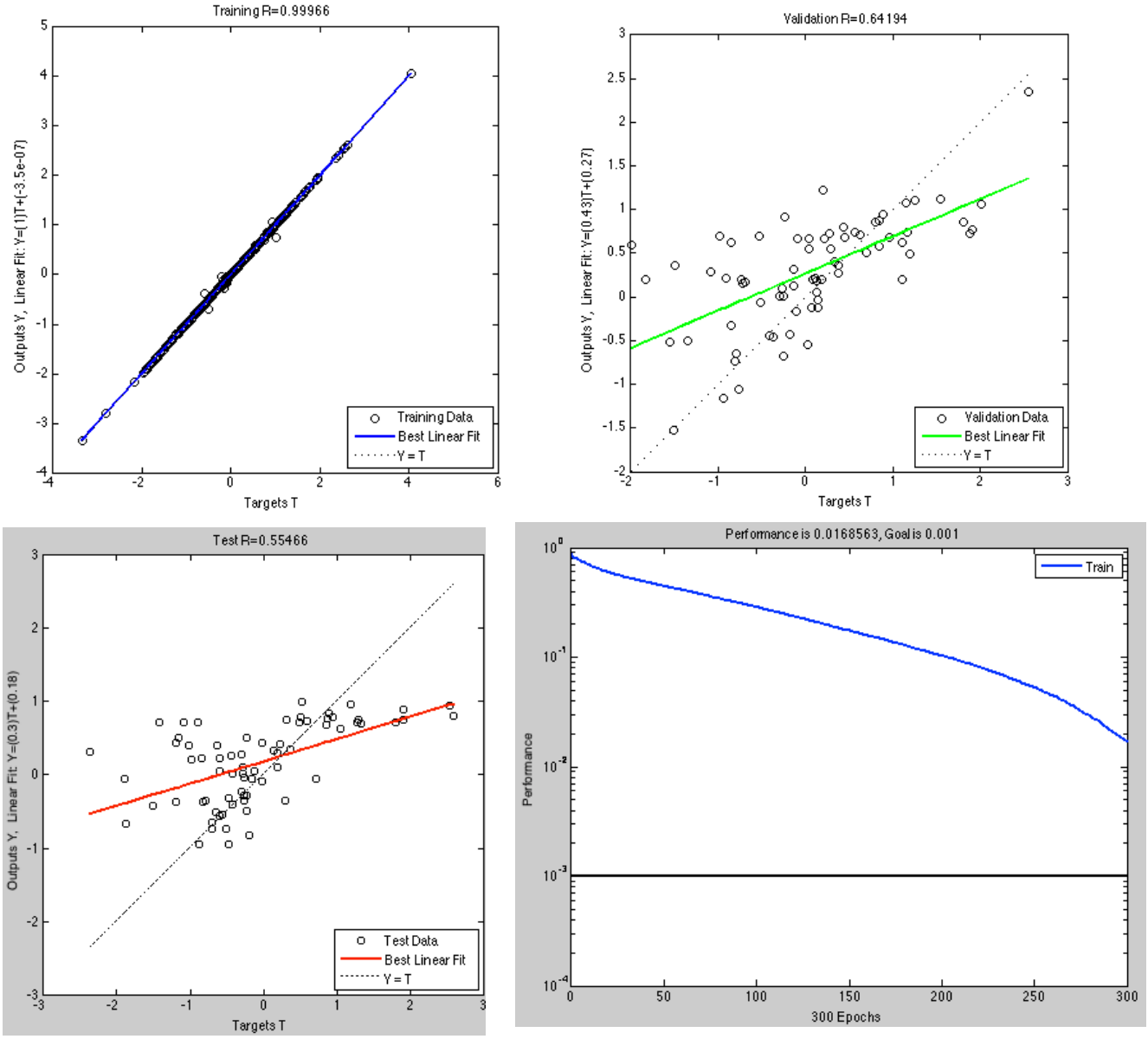
ISBN: 978-1-61804-327-6



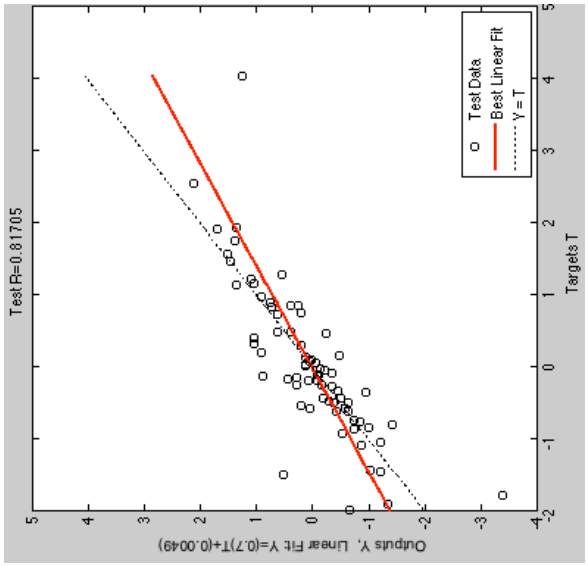
(c) Radial Basis Neural Network

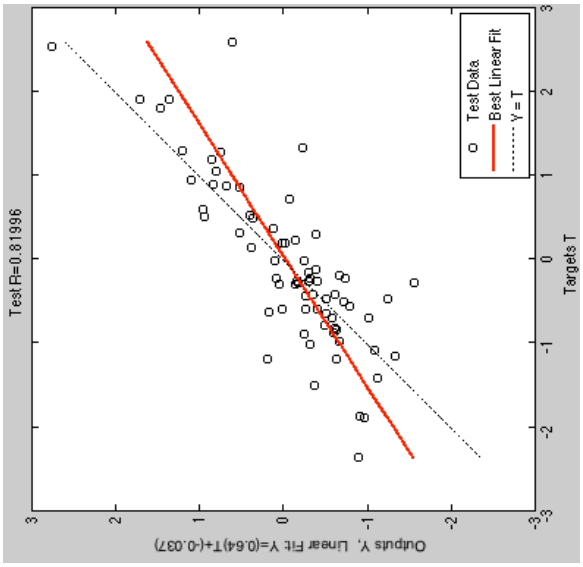
Gottingen GOE426 Airfoil Eppler E399 Airfoil

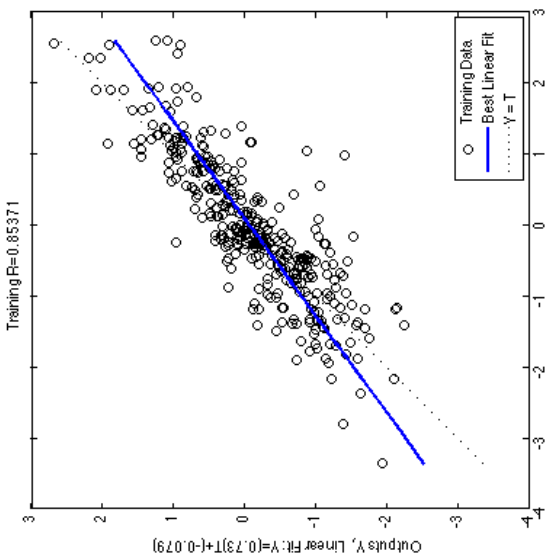
Fig 6 : Comparison of Results of Known to Unknown Airfoil to Neural Network

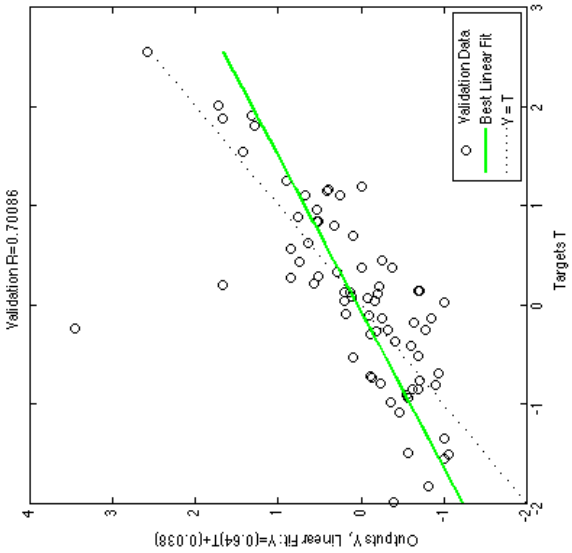


(a) Radial Basis Neural Network

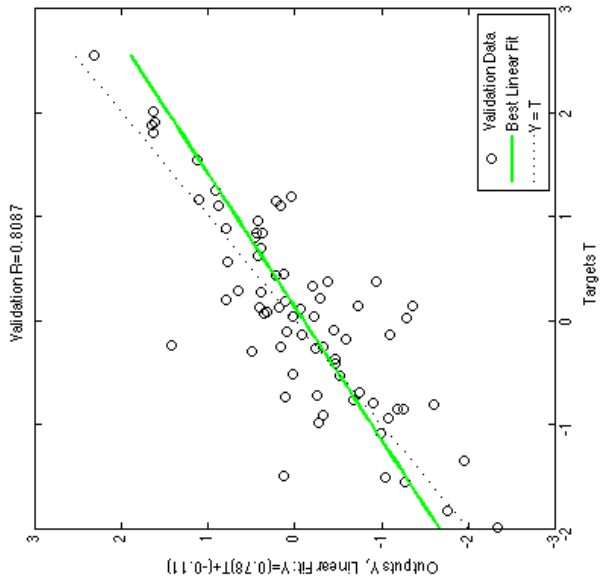


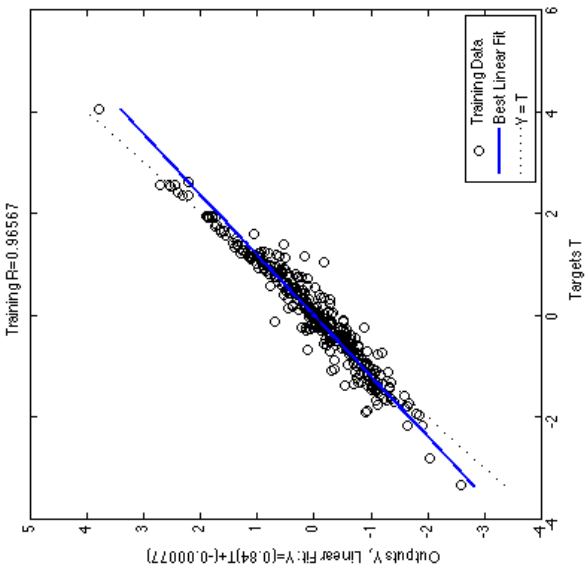






(b) Feed Forward and Back Propagation Neural Network





(c) Generalized Regression Neural Network

Fig 7 : Regression Analysis

ISBN: 978-1-61804-327-6

IX. Future Work

We have implemented Simple ACO in this research work. Future works may consider implementation of other extensions of ACO techniques like Elitist AS, Ant-Q, Max-Min As, Hyper-cube AS and etc to achieve high performance in order to further reduce the optimization level and attain higher level of precision.

References

[1] G. Trapani, T. Kipouros and M. Savill, 'Computational Aerodynamic Design for 2D High-Lift Airfoil Configurations', 2010.

[2] K. Lane, 'Novel Inverse Airfoil Design Utilising Parametric Equations', Graduate, California Polytechnic State University, 2010.

[3] Z. Zhu, F. Hongyan, Y. Rixin and L. Jie, 'Multi-objective Optimization Design of Airfoil and Wing', *Science in China Series E Technological Sciences*, vol. 47, no. 1, pp. 15-25, 2004.

[4] P. Hewitt and S. Marques, 'Aerofoil Optimisation Using CST Parameterisation in SU2', in *Royal Aeronautical Society Applied Aerodynamics Group Conference*, Bristol, 2014.

[5] M-selig.ae.illinois.edu, 'UIUC Airfoil Data Site', 2015. [Online]. Available: http://m-selig.ae.illinois.edu/ads/coord_database.html. [Accessed: 11- Jan- 2014].

[6] A. Hacioglu, 'Fast Evolutionary Algorithm for Airfoil Design via Neural Network', *AIAA Journal*, vol. 45, no. 9, pp. 2196-2203, 2007.

[7] U. Wickramasinghe, R. Carrese and X. Li, 'Designing Airfoils Using a Reference Point Based Evolutionary Many-objective Particle Swarm Optimization Algorithm', *IEEE Congress on Evolutionary Computation*, 2010.

[8] M. Ebrahimi and A. Jahangirian, 'Aerodynamic Optimization of Airfoils Using Adaptive Parameterization and Genetic Algorithm', *Journal of Optimization Theory and Applications*, vol. 162, no. 1, pp. 257-271, 2013.

[9] A. Kharal and A. Saleem, 'Neural Networks Based Airfoil Generation for a Given Cp Using Bezier-PARSEC Parameterization', *Aerospace Science and Technology*, vol. 23, no. 1, pp. 330-344, 2012.

[10] G. Sun, Y. Sun and S. Wang, 'Artificial neural network based inverse design: Airfoils and wings', *Aerospace Science and Technology*, vol. 42, pp. 415-428, 2015.

[11] K. Thinakaran and R. Rajasekar, 'Design of Airfoil Using Inverse Procedure and Neural Network', *An International Journal of Advanced Computer Technology*, vol. 2, no. 9, pp. 285-290, 2013.

[12] L. Prabhu and J. Srinivas, 'Flutter Prediction Using Neural Networks-based Hybrid Optimization Scheme', *International Journal of Research in Aeronautical and Mechanical Engineering*, vol. 2, no. 4, 2014.

[13] M. Ahmed and N. Qin, 'Surrogate-Based Aerodynamic Design Optimization: Use of Surrogates in Aerodynamic Design Optimization', in *13th Conference on Aerospace Sciences & Aviation Technology*, Cairo, 2009.

[14] R. Derksen and T. Rogalsky, 'Bezier-PARSEC: An Optimized Aerofoil Parameterization for Design', *Advances in Engineering Software*, vol. 41, no. 7-8, pp. 923-930, 2010.

[15] N. Salunke, R. Juned and S. Channiwala, 'Airfoil Parameterization Techniques: A Review', *American Journal of Mechanical Engineering*, vol. 2, no. 4, pp. 99-102, 2014.

[16] M. Darigo and L. Gambardella, 'Ant Colonies for the Travelling Salesman Problem', *BioSystems*, vol. 43, pp. 73-81, 1997.

[17] M. Darigo and T. Stutzle, *Ant Colony Optimization*. Cambridge: The MIT Press, 2004.

[18] J. Bell and P. McMullen, 'Ant Colony Optimization techniques for The Vehicle Routing Problem', *Advanced Engineering Informatics*, vol. 18, no. 1, pp. 41-48, 2004.

[19] G. Fainekos and K. Giannakoglou, 'Inverse Design of Airfoils Based on a Novel Formulation of the Ant Colony Optimization Method', *Inverse Problems in Engineering*, vol. 11, no. 1, pp. 21-38, 2003.

Multiple Choice Question Tests – Advantages and Disadvantages

Jindrich Klufa
Dept. of Mathematics
University of Economics
Prague, Czech Republic
klufa@vse.cz

***Abstract*—** In this paper we shall describe situations in which the multiple choice question tests are better than the standard tests. On the other hand, there are situations in which classical tests are more convenient. As an example we shall use tests in mathematics at the Faculty of Finance and Accounting at University of Economics in Prague. We shall consider the tests in mathematics in entrance examinations and the tests in the basic course of mathematics for the first year. For this analysis we shall use some probability methods.

***Keywords*—** Multiple choice question tests, tests in mathematics, probability methods.

I. INTRODUCTION

MULTIPLE QUESTION TESTS are widely used in testing knowledge of students. One of the advantages of such type of test is that the results can be evaluated quite easily even for large number of students. On the other hand, a student can obtain certain number of points in the test purely by guessing the right answers and this fact affects reliability of the test and should be considered in interpretation of test scores. This problem is addressed in education research – see Premadasa (1993), Klufa (2012).

An analysis of a multiple choice question test from probability point of view is provided in this paper. This test is for example used for entrance examinations at University of Economics – see Klufa (2011). Note that standard (no multiple choice questions) tests are used for checking knowledge of students in mathematics courses at University of Economics – for analysis of such test see (Otavova and Sykorova, 2014), but regarding entrance examination, multiple choice questions are preferred so that the results of tests can be obtained quickly and there is clearly no impact of any subjective factor in evaluation. Similar problems are in Moravec, Štěpánek and Valenta (2014), Brozova and Rydval (2013).

Entrance examinations at the Faculty of Finance and Accounting at University of Economics in Prague include mathematics and English. Test in mathematics has 10 questions for 5 points and 5 questions for 10 points (100 points total). Each question has 5 answers. Test in English has 50 questions for 2 points (100 points total). Each question has 4 answers. Questions are independent (one answer is correct), wrong answer is not penalized. We provide an answer to the following questions (under assumption of random choice of answers): what is probability that number of right answers exceeds given number, what is expected number of right answers, what is standard deviation, and finally what is a risk of success of students with lower performance levels.

II. METHODS

Multiple choice question tests (the test has n questions, each question has m answers) are applied for the entrance examinations at the Faculty of Finance and Accounting at University of Economics in Prague. Therefore a model of binomial distribution can be used for the entrance examinations. From probability point of view a multiple choice question test means:

Let us consider n independent random trials having two possible outcomes, say "success" (right answer) and "failure" (wrong answer) with probabilities p and $(1-p)$ respectively. Probability of correctly answered question p (under assumption that each of m answers in particular question has the same probability and one answer just is correct) is $p=1/m$.

Let us denote X as number of successes (right answers) that occur in n independent random trials. X is random variable distributed according to the binomial law with parameters n and p. Probability that number of successes is k ($k=0, 1, 2, \ldots, n$) is (see e.g. Marek (2012))

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k} \quad (1)$$

ISBN: 978-1-61804-327-6

The expected value and the standard deviation of random variable *X* distributed according the binomial law with parameters *n* and *p* is

$$E(X) = np, \quad \sigma(X) = \sqrt{D(X)} = \sqrt{np(1-p)} \qquad (2)$$

where *D(X)* is dispersion of random variable *X*.

The distribution function of random variable *X* distributed according to the binomial law with parameters *n* and *p* is

$$F(x) = 0, x < 0, \qquad F(x) = \sum_{k=0}^{[x]} \binom{n}{k} p^k (1-p)^{n-k}, x \geq 0 \qquad (3)$$

where *[x]* is integer part of *x*.

III. Entrance Examinations in Mathematics

Entrance examinations in mathematics have 10 questions for 5 points and 5 questions for 10 points (100 points total). Questions are independent. Each question has 5 answers; the wrong answer is not penalized. Under assumption that each answer has a same probability, probability of right answer is *p=1/5*.

Example 1. Under assumption of random choice of answers we shall find probability that number of points in the test in mathematics is 15.

Let us denote

Y = number of points in the test in mathematics
X_1 = number of right answers in the first 10 issues
X_2 = number of right answers in 10-points tasks

It holds

$$P(Y=15) = P[\,(X_1=1 \cap X_2=1)\ U\ (X_1=3 \cap X_2=0)\,] =$$
$$= P[\,(X_1=1 \cap X_2=1)\,] + P[\,(X_1=3 \cap X_2=0)\,]$$

Random variables X_1, X_2 are independent, therefore we have - see e.g. Rényi (1972)

$$P(Y=15) = P(X_1=1)\,P(X_2=1) + P(X_1=3)\,P(X_2=0)$$

Random variable X_1 has binomial distribution with parameters *n=10* and *p=0,2*. Random variable X_2 has binomial distribution with parameters *n=5* and *p=0,2*. According to (1) we obtain

$$P(Y=15) = \binom{10}{1} 0{,}2\ 0{,}8^9 \binom{5}{1} 0{,}2\ 0{,}8^4 + \binom{10}{3} 0{,}2^3\ 0{,}8^7 \binom{5}{0} 0{,}8^5 = 0{,}175922$$

Analogously, we can calculate the probability *P(Y=k)* for other *k=0, 5, 10, 15, ..., 95, 100* (see Table 1 and Figure 1). For this calculation we used software Mathematica (Statistics 'Discrete Distributions') – see Wolfram (1996).

Points in the test	Probability	Points in the test	Probability
0	0,035184	55	0,002890
5	0,087961	60	0,000957
10	0,142937	65	0,000275
15	0,175922	70	0,000067
20	0,174547	75	0,000014
25	0,146098	80	0,000002
30	0,105227	85	3×10^{-7}
35	0,066057	90	2×10^{-8}
40	0,036467	95	1×10^{-9}
45	0,017761	100	3×10^{-11}
50	0,007634	Sum	1,000000

Tab. 1 Distribution of number of points in the test in mathematics

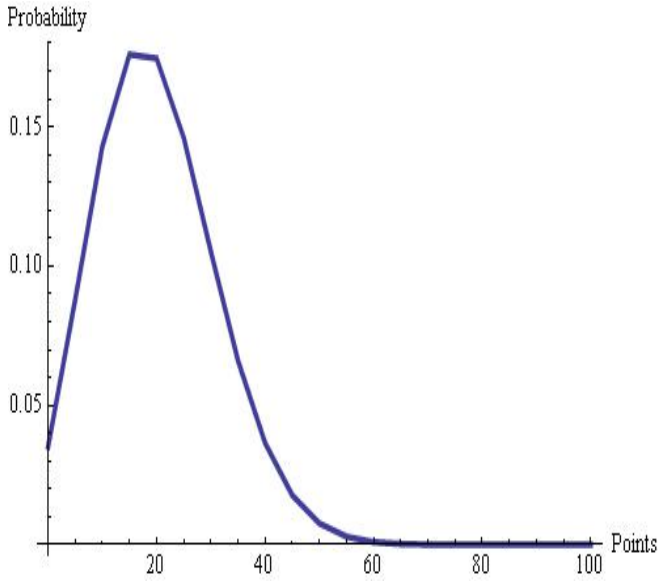


Fig. 1 Distribution of number of points in the test in mathematics (polygon)

Example 2. Under assumption of random choice of answers we shall find probability that number of points in the test in mathematics is

(a) 30 and more,
(b) 40 and more,
(c) 50 and more.

(a) Using notation from example 1 we have - see e.g. Rao (1973)

$$P(Y \geq 30) = 1 - P(Y<30) = 1 -$$

$$P[(Y=0) \cup (Y=5) \cup (Y=10) \cup (Y=15) \cup (Y=20) \cup (Y=25)] =$$

$$= 1 - [\, P(Y=0) + P(Y=5) + P(Y=10) + P(Y=15) + P(Y=20) + P(Y=25) \,]$$

Finally from Tab.1 we obtain

$$P(Y \geq 30) = 1 - 0{,}762649 = 0{,}237351.$$

Under assumption of random choice of answers almost a quarter of students get the test score 30 or more points.

(b) Analogously, we obtain

$$P(Y \geq 40) = 1 - P(Y<40) =$$

$$= 1 - [\, P(Y=0) + P(Y=5) + P(Y=10) + P(Y=15) + P(Y=20) + P(Y=25) + P(Y=30) + P(Y=35) \,]$$

Finally from Tab.1

$$P(Y \geq 40) = 1 - 0{,}933933 = 0{,}066067.$$

Under assumption of random choice of answers approximately 6,6% of students get the test score 40 or more points.

(c) Finally

$$P(Y \geq 50) = 1 - 0{,}988161 = 0{,}011839.$$

Under assumption of random choice of answers approximately 1,2% of students get the test score 50 or more points.

Example 3. Under assumption of random choice of answers we shall find expected number of points in the test in mathematics and mode.

Using notation from example 1 we have

$$Y = 5\,X_1 + 10\,X_2$$

Therefore - see e.g. Anděl (1978)

$$E(Y) = E(5X_1 + 10X_2) = 5\,E(X_1) + 10\,E(X_2)$$

According to (2) we obtain (e.g. $E(X_1) = 10 \,.\, 0{,}2 = 2$)

$$E(Y) = 5 \,.\, 2 + 10 \,.\, 1 = 20.$$

Expected number of points in the test is 20. The mode is the most probable number of points. From Tab.1 is

$$\hat{y} = 15.$$

I. IV. Conclusion

Entrance examinations at the Faculty of Finance and Accounting at University of Economics in Prague include mathematics and English. Probability that number of points from test in mathematics is 50 and more is 0,011839 (see example 2). Analogously, we can calculate this probability for test in English. We obtain 0,000122. That means (both tests are independent: 0,011839 x 0,000122 = 0,000001) that approximately one student from million (under assumption of random choice of answers and using 50 points as a cut-off value for successful completion in each test) successfully makes the entrance examinations at the Faculty of Finance and Accounting at University of Economics by pure guessing the answers.

Multiple choice question tests are optimal for entrance examinations at University of Economics. These tests are objective (there is clearly no impact of any subjective factor in evaluation). Moreover, results can be evaluated quite easily for large number of students. From results of this paper follows that risk of acceptance students with lower performance levels is negligible.

On the other hand, number of students in the basic course of mathematics is not large. In this case is better use the standard tests. These tests (the solution of concrete examples) examine the students' knowledge of mathematics better than the multiple choice question tests.

References

[1] Anděl, J. (1978) *Matematická statistika.* Prague: SNTL/ALFA.

[2] Brozova, H., Rydval, J. (2013) 'Analysis of the exam test quality', *Efficiency and Responsibility in Education*, Proceedings of the 10th International Conference, Prague, pp. 47–54.

[3] Klufa, J. (2011) Mathematics for entrance examinations at University of Economics, Prague: Ekopress

[4] Klufa, J. (2012) 'Tests from probability point of view', Efficiency and Responsibility in Education 2012, Proceedings of the 9th International Conference, Prague, pp. 229–233.

[5] Marek, L. (2012) Pravděpodobnost, Prague: Professional Publishing,

[6] Moravec, T., Štěpánek, P., Valenta, P. (2014) The Impact of Progress Testing of Students On Their Results At Final Exam. *International Conference on New Horizons in Education*. Paris, pp. 829–834.

[7] Otavova, M., Sykorova, I. (2014) 'Analysis of scores from mid-term and final test by a contingency table' , *Efficiency and Responsibility in Education, Proceedings of the 11th International Conference*, Prague, pp. 527–533.

[8] Premadasa, I. (1993) 'A reappraisal of the use of multiple-choice questions', Medical Teacher, vol. 15, no. 2-3, pp. 237-242.

[9] Rao, C.R. (1973) Linear Statistical Inference and Its Applications, New York: John Wiley.

[10] Rényi, A. (1972) *Teorie pravděpodobnosti.* Prague:Academia.

[11] Wolfram, S. (1996) Mathematica. Addison-Wesley.

Two-dimensional Finite Elements Thermal Analysis of a Switched Reluctance motor

Gholam Reza ZANDESH Javad Shokrolahi MOGHANI
Power Engineering Center of Excellence, Faculty of Electrical Eng.
AmirKabir Unviersity of Tech. (Tehran PolyTechnic)
Tehran,Iran
Zandesh77@yahoo.com

Mina GHOORCHIAN
Research Department of
Hirbodan Ariana Professional Engineering Training
Tehran,Iran
m.ghoorchian@gmail.com

***Abstract*—this paper report on the thermal analysis carried out on switched reluctance motor (SRM) by finite element (FE) methods. Thermal calculation, especially in non-conventional machine in which there is only limited information about heat dissipation, is an important stage in machine design. SRM is such a machine hence necessitating a defined methodology to carry out thermal analysis; this paper intensely attempts and provides simulation method for thermal analysis. The thermal flux plot, the isothermal distribution, thermal gradient in different part of SRM at its different rotor position and the respective governing equation are presented.**

Keywords— Switched Reluctance Motor; finite Element Analysis;Thermal analysis.

I. INTRODUCTION

Electric motors play an important role in consumer and manufacturing industries. Among all different kinds of electric motors, Switched Reluctance Motors (SRMs) have a special place. That is because of their simple construction, high speeds and their very low cost comparing with the other Motors . In these motors, the rotor does not have any permanent magnet or windings, in contrast with other motors .This reduces the rotor weight and consequently, with a precise commutation, makes it possible to reach much higher speeds. Variable reluctance in the air gap between rotor and stator poles provides torque in SRMs. In other words, the motion system tends to retain its energy in the minimum state, so the rotor poles, rotate in such a way to face the stator poles. However, the different angles between the rotor and stator poles cause some of them to be out of phase and with a proper commutation a smooth motion can be achieved. Another advantage of SRMs is their reliability, which is the result of a high independency in the physical and electrical phases. On the other hand, in SRMs, not only do we need a commutator, but also a complex control system is required, Because of the nonlinearity of the magnetic characteristics.

II. SWITCHED RELUCTANCE MOTOR

The cross-section of SRM under study [2] is shown in Fig 1. There are six salient stator poles; the winding-less simple solid laminated lot or has four salient poles (“doubly-salient”). The DC excitable windings on the stator are of particularly simple form and concentrated type (“singly excited”);

Two diametrically opposite stator windings are connected in series and exited together providing Ns/2 ‘pinases’ for an SRM of Ns number of stator poles. The production of torque lies with the rotor to get aligned with the magnetic field produced by the excitation current (DC) [1] in the windings on the stator.

Currents in the subsequent stator windings are processes switched ON and OFF in accordance with the rotor-positions and, with this simplest form of control, SRM develops a wide range of torque-speed characteristics of a best DC series rotor. By proper design of switching the phases [3] a wide variety of applications can derive from SRM.

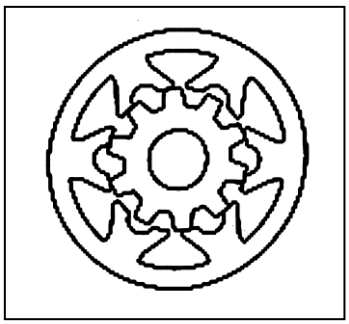


Fig. 1. CAD model of SRM 12/10

III. ESTIMATION OF LOSSES

Losses of SR motor are divided into three parts:

A. Copper losses

current passing through the stator windings generates the heat and temperature rise of the motor . If the current waveform of each phase is taken to be a rectangular wave , then the copper loss can be obtained as follows:

$$I \approx i_p / \sqrt{m} \ , P_{cu}=mRI^2 \quad (1)$$

B. Core losses

Complicated magnetic circuit of SRM , makes it difficult to calculate the core losses analytically. These losses depend on the shape of current waveform , angle and period of switching

ISBN: 978-1-61804-327-6

and the operation of drive. An experimental equation for core losses in SR machines is expressed as:

$$Pc=afb\ (v/f)c \qquad (2)$$

in this case , values of a,b,c are:

$a=5.3*10^{-3}$,b=1.466,c=2.159

where 84% of these losses are in the stator yoke ,8% in the stator teeth, 3.3%in the rotor teeth and 4.7% in the rotor core.

C. Friction losses

They are results of relative motion in bearings as well as the viscous flow over the rotor because of its rotation. These losses can be written as:

$$P_f=aN+bN^2 \qquad (3)$$

$a=0.0118,b=0.1434*10^{-4}$

These losses will be shared between different parts of motors as sources of heat flow.

IV. THERMAL ANALYSIS

FE analysis is a more detailed. Structured and equipped modeling tool for numerical study of SRM, A detailed FE analysis enhances the performance prediction as well as provides a conventional platform for further detailed analysis [4].

The governing equation in the usual format as given in (1),

$$d(u)\frac{\partial u}{\partial t}-\nabla.(c(u)\nabla u)+\partial(u)u=f(u) \qquad (4)$$

Where a,c,d,f and u are the unknown complex time-varying functions on boundary problem domain , u, on the boundary behave as the boundary condition (Dirtchlet)and also as the

$$\rho\alpha\frac{\partial T}{\partial t}-\nabla.(k\nabla T)+h_{cv}.s(T-T)=Q \qquad (5)$$

Normal derivative *of* u on the boundary (Neumann). Equation (4) for thermal analysis becomes equation (5).

In this equation T is the unknown temperature distribution, hcv is the specific heat (heat transfer co-efficient), K is the heat conductivity, Q is the heat source, s is the outer surface of dissipation and T is the ambient temperature ; $\alpha and \rho$ the material properties.

In (5) the individual three terms are the heat dissipation terms, respectively by radiation conduction and convection processes. The heat production (hot sport) is dissipation by radiation in such a closed case as it's and [6]. Equation (5), to be solved by FE method for thermal analysis thus reduces to the form of equation (6); however, the dissipation of heat considerably takes place by other two

$$-\nabla.(k\nabla T)+h_{cv}.s(t-t)=Q \qquad (6)$$

Had the heat been transferred by thermal conduction it is enough to consider the boundary for analysis; but heat dissipation by convection too plays a vital role. It can be taken that thermal regions on the boundary can be considered as made up of several thermal nodes. These nodes on the boundary are connected to the thermal nodes in the several layers of the inner heat region and in this way the heat is dissipated amongst several layers. Thermal flow takes place between the nodes in the layers as there will certainly be temperature difference between the regions. This flow occurs against the thermal resistance which is a material property. Thus there is thermal energy exchange ('thermal gradient') which has convective heat transfer with surroundings. Heat dissipation by this convective heat transfer is given by second term of the (6), which is hcvs.(t-t).For naturally convicted cooling which is the east with SRM, the hcv value is emphatically given by,

$$hcv=1.08(T-T)1/3 \qquad (7)$$

The above discussion is in analogy with electric circuits where an electric current flows when a potential difference exists between two nodes, against the resistance of the path. The thermal analysis is carried out exactly adopting these basic electrical fundamentals; in order to estimate the temperature difference the surface and the surroundings the which motor is considered as a single volume element from which the heat dissipates into the surroundings. One current source is considered for each node injecting current ('heat') into the node, considering the ambient temperature as the source voltage and the node voltages as the different potential points (Fig2)

Such formed non-linear equation (6) is solved by FE methods by the numerical R-K approach.

The operating case for thermal simulation is shown in Fig3. A finite element heat run simulated whose results are presented in the next section. Note that the value of Q (watts/m3) which is the main heat source is to be calculated for different parts like stator core, rotor core, coils etc., form the core and copper losses of the machine.

The working conditions of the c considered SRM [2] is simulated rigorously though the finite element analysis; the power of 2D-Maxwell electromagnetic field simulator (ANSOFT) was used by the authors. Improved stages to fine-time the simulated results, originally dictated by the authors [4], were carefully implemented. For this simulated working conditions as appearing in Fig .3. I, finite element thermal analysis using ANSYS and FAST. THERMAL toll was carried out.

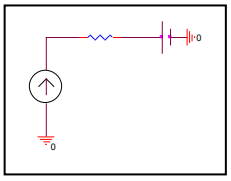


Fig.2. Electric equivalent thermal network

ISBN: 978-1-61804-327-6

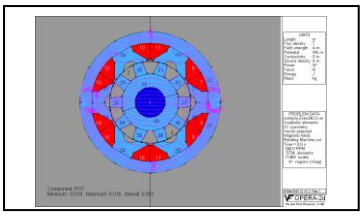


Fig.3. Nominal flux plots for full load

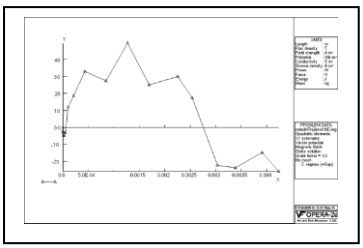
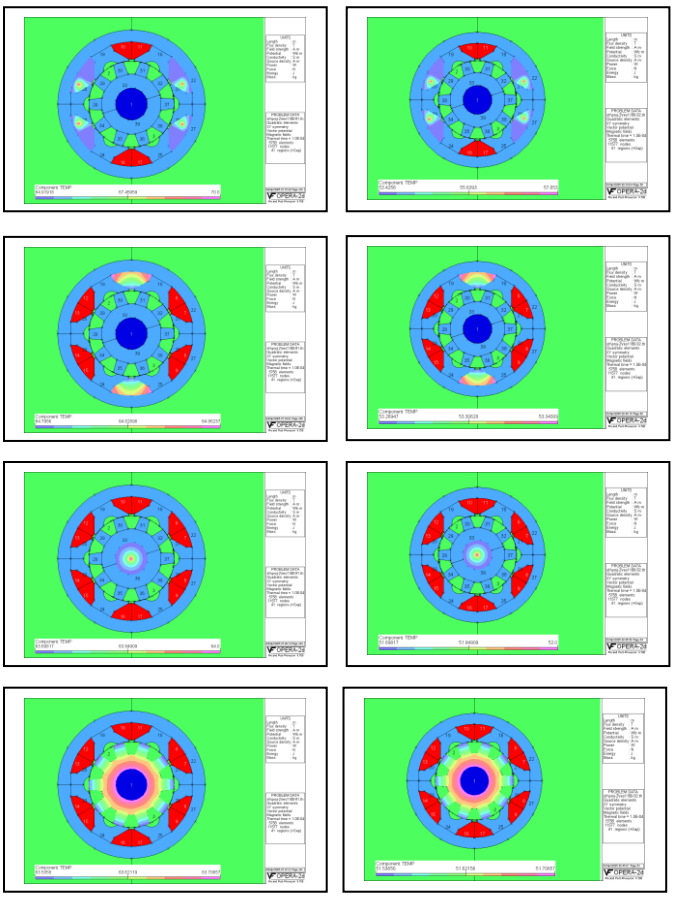


Fig.4. Torque profile

V. RESULTS OF SIMULASION

The results obtained by the simulated thermal analysis are presented in this section. The foot-notes in the figs5 are (color plots shown monochromatically) are the simulation descriptions of the respective plots.



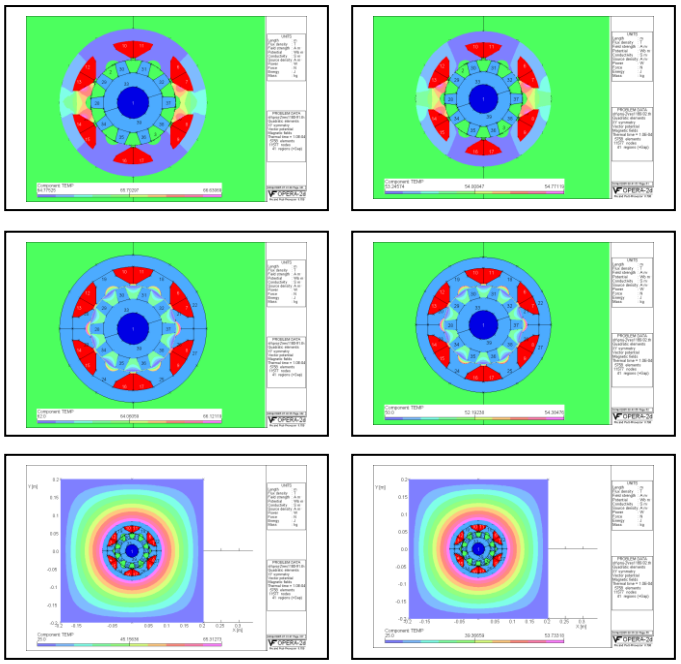
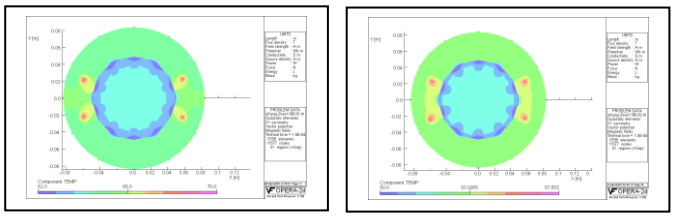


Fig.5. Simulated thermal analysis results-color plot

VI. CONCLUSIONS

A useful simulation method used to analyze the thermal characteristics of switched reluctance motor based on finite element method has been presented. The key points in determining the thermal gradients at the different parts of the machine.

The hot spot temperature is : 70c.



VII. VII.REFERENCES

[1] P.J.Lawrenson, 'Switched reluctance motor motor drives' , LEE proceeding, fed 1983, pp:144-147

[2] T.J.E. Miller, 'Brushless permanent magnet and reluctance motor drives' ,calerendon press, Oxford, 1989.

[3] Mike Barnes, power electronic circuits for witched reluctancedrives' , IEEE Transactions on power Electronics, November 1991, pp;1100-1111.

[4] K.N. srinivas and R. Arumumgam, 'Improved finite element analysis for switched reluctance motor' , Proceedings on SASESC-2000 National conference, Dayalbagh educational institute,Agra ,pp:128-133.

[5] M. Jakob, 'Elements of heat transfr' ,John Wiley, 1957.

[6] J.P.Holman, 'Heat Transfer' , Mc-Graw Hill,USA,1976.

[7] Yifan Tang, 'characterization and numerical analysis of Switche reluctance motor' , LEEE Transactions on Industrial Applications, Vol:33 pp:1545-1552

ISBN: 978-1-61804-327-6

Introduction to the IDL application in the Weather

Wan-zhen Zhou, Quan-bing Hou

***Abstract*—**It mainly introduce a ideal tool that the IDL as the meteorological data analysis, visualization expression of meteorological data set and weather data analysis platform for cross-platform application in the field of climate science data(measured rainfall isoline,4D meteorological data platform, two and three-dimensional isoline, doppler radar).It has an advanced and integrated development environment which involves multitudinous weather data analysis toolkits and application of high-speed meteorological three-dimensional graphics display technology, large sets of visualization, interactive data analysis. In a word, the meteorological department scientific research becomes fast and effective in meteorology data processing because of IDL's application.

***Keywords*—** Data component; IDL;Meteorological data visualization; Data interaction

I. INTRODUCTION

IDL [1] is a programming language of data analysis and visualization application which invented by the American company ITT. At first the original intention is to help climate scientists analysis the data from meteorological satellite. It can converts the vast meteorological data to graphics in the late 1970s , then IDL is widely used in weather forecast and complex weather statistics. As the fourth generation of scientific computing visualization language [2] , it not only help the user to establish the IDE development environment of visual programming environment but also to provide plenty of programming tools and the built-in math library meteorological forecast [3] analysis function, moreover it greatly reduce the workload of meteorological data image visualization development. the program wrote by IDL can run in other platforms with change infrequently, which makes meteorological data monitoring system flexible involuntary.

II. IDL'S ADVANTAGE IN THE FIELD OF WEATHER SCIENCE DATA

A. Support luxurious meteorological data format

A large number of data in different formats will be used in meteorological science research, IDL can storage these data effectively and express corresponding meteorological data graphical visualization to developers exactly.IDL provides a mass of meteorological data tools which can be readed and wrote, supports common data format reading and writing directly, such as general image data format (DCM, JPEG2000, TIFF/Geo , TIFF , etc.).It also support some scientific meteorological data format(HDF, CDF, NCDF, etc) that widely used in the US national oceanic and atmospheric administration institutions

B. Meteorological data analysis, the weather chart data statistics

IDL integrates perfect meteorological data analysis, data statistics, image processing software package. it provides the intelligent tool--iTools which combine a series of interactive graphics display high quality tools with data analysis and visualization function. Futhermore using the iTools intelligence tool can interactively display image and contour data of meteorological information. The project using the traditional C or Fortran language needed for days or weeks can be done in a matter of hours if we use IDL.

C.Capabilities of advanced meteorological contour 2D and 3D data interactive image visualization

From simple two-dimensional chart (drawing from multi-dimensional surface,contour image display) to the use of OpenGL [4] hardware acceleration for interactive meteorological satellite cloud three-dimensional image browsing. Moreover, it can support the multiprocessor render body data rapidly. Luxuriant meteorological data visualization effect got easily because of using IDL.Considering the heavy and complicated of meteorological data, IDL adopt process design and give full play to the multiple processing functions on the processor system.In addition its built-in meteorological data processing support powerful scientific computing with very few lines of IDL code, however other language do barely about the meteorological data of image visual function [5] . So as a powerful tool,it has been used for analysising the meteo rological satellitedata for a long time. And it becomes the most preferred language used by climate scientists ,which is known as the "milestones NASA" in recent 40 years .Meteorological satellite is equipped with multi-channel high resolution radiometer, infrared spectrometer and microwave radiometer meteorological remote sensors. It can measure the global cloud cover, wind, parameters such as pressure, temperature and humidity. What's more, it obtain global meteorological data for providing a global meteorological information.The second generation of polar orbit meteorological satellite—FY – 3 weather satellite [6] make a great step on the basis of the FY- 1 meteorological satellite. The specific requirement is to solve three-dimensional

Wan-zhen Zhou is from the School of Information Science & Engineering, Hebei University of Science and Technology, China. He received his doctor's degree from China people's Liberation Army Ordnance Engineering college in 2006. His main research interests include the search on Artificial intelligence, data mining and IOT. Email: zhou_wanzhen@163.com.

Quan-bing Hou is from the School of Information Science & Engineering, Hebei University of Science and Technology, China. His main research interests include the weather three dimensional image processing . Email: hou596678352@163.com.

atmospheric detection, improving the global information obtaining ability greatly and further improving the cloud area remote sensing ability thus they can obtain global, all-weather, three-dimensional, quantitative and multispectral atmospheric parameters. The application of FY - 3 meteorological satellite mainly includes three aspects:

1)、Providing global uniform resolution of meteorological parameters for medium-term numerical weather prediction

2)、Providing various meteorological and geophysical parameters for climate prediction through researching the relation between global change and climate change

3)、Providing meteorological information in any areas of the world For a variety of professional activities

III.IDL APPLICATION IN THE WEATHER

A. The IDL application in scientists measured the global precipitation

Global precipitation observation is an international satellite network which cooperated by the Japan aerospace exploration agency and NASA together. it provides the next generation of rain and snow observed. The concept of global precipitation observation center deployed on a core satellite which carry advanced satellite/radiometer system. The system can measure space and also can be used as a reference standard. The purpose is to merge all of space and ground-based measurement information available to develop high resolution, near real-time global precipitation data of the product with a series of research and the business of the satellite rainfall observation as a standard. Through these improved measurement, GPM tasks will help us to advance our understanding of the earth's water and energy cycle, Through improving the extreme events of natural risk and hazard prediction, we can use accurate and timely rainfall information ability to benefit society.

GPM will provide a complete coverage of higher precision and the dynamic range of global precipitation measurement for studying the rainfall, at the same time, the IDL module build tools such as the structure of language libraries, file access, data analysis and visualization toolkit library which could help scientists to study it .GPM improve weather and rainfall forecast by the instantaneous rainfall data assimilation.GPM will provide more advanced scientific contribution and social benefits compared to the TRMM. It can improve the understanding of the earth's hydrologic cycle and related to climate change as well as the full understanding of climate sensitivity, At the same time its expanded the monitoring and the ability to predict extreme weather events such as hurricane.

B. 4 D meteorological data visualization platform

1)、The development purpose of 4D meteorological data visualization platform

With the increase of the meteorological data redundancy and data explosion,today the existing tools rarely to fusing, processing and interpreting meteorological data. The amount of data collected by satellites and other means is huge including meteorological data [7] in any fields of science. NOAA successfully launched a polar orbit satellite on October 28, 2011.It continue to increase the meteorological data sources. Related meteorological researchers worry about the weather forecasters and their forecast tool can carry such a large amount of data especially when the major weather events occurs. In order to applicate IDL more effectively and mitigate large pressure of meteorological data derived from meteorological data explosion ,4 D meteorological data visualization platform is developed by NWS.

2)、The formation of 4 D meteorological data sets and applications

A study is conducting from the national weather service to the federal aviation administration government offices in the United States. They hope to merge various meteorological data to set up an integrated meteorological data warehouse management platform in order to strengthen the cooperation of distributed meteorological information based on network, moreover they also make meteorological monitoring dynamic decisions. Visibility cause disorder of some factors (cloud, rain and snow) and meteorological characteristics (convection, ice, snow, cloud screen visibility, wind speed and direction, etc.) [8] will be described in 3 D space. Time is joined the meteorological data in the model as the fourth dimension which forms the 4D meteorological data sets, meteorological data demonstrate diagram as shown in figure 1, it makes the original characteristic parameters with time characteristics.4D data set [9] which be stored in the server can set out through a service oriented architecture distributed system. In this way, the meteorology experts provide technical support for federal aviation administration through determinating each cycle of meteorological data.

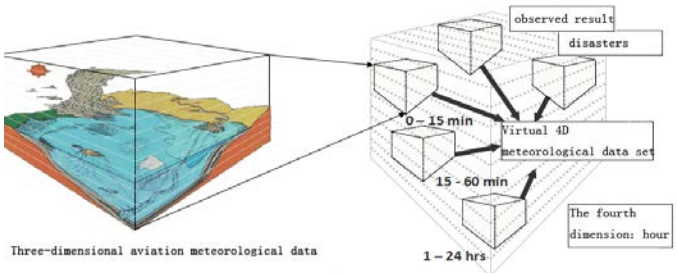


Fig. 1.4D weather simulation data show

C. IDL meteorological contour data of 2- 3D visualization

The platform get the weather data (mainly isoline under different height) multidimensional visual image display, the figure 2 is provided by the meteorological department the contour of the two-dimensional figure which is 500 pa and 1000 pa on July 30, 2003, 8 pm. We can from the top, low, before, after, left, right, six view direction to watch visual images at the control modules of this platform, the figure 3 is different height field contour of two 3 D display which equips any highly and angle off the meteorological data of 3D simulation at the same time. We can use the mouse control

ISBN: 978-1-61804-327-6

casual observation images from any angle, so that we could analyze weather data visualization and make an accurate prediction of weather through the integration of two 3 D GIS technology. The program language of IDL realize the GIS 3D drawing of contour map, and then put the existing meteorological data (3D contour) loaded into the meteorological and display in the form of graphical visualization. Finally it achieved the integration of two 3 D GIS meteorological data [10].

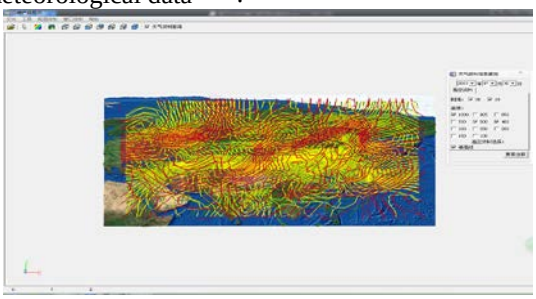


Fig.2.IDLtwo-dimensional display

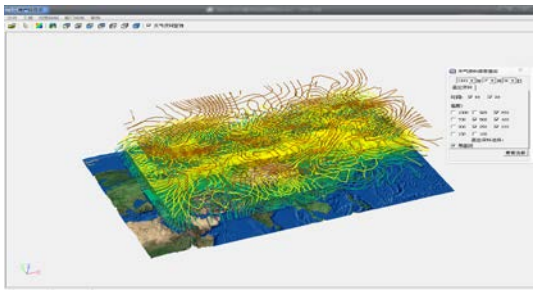


Fig.3. The meteorological data of 3d simulation meteorological data

D.IDL read date and visualization of the new generation doppler weather radar data

CINRAD/CC weather radar is a new generation of C band coherent doppler weather radar that China meteorological administration use it for domestic meteorological operational network detection .It monitor the scope of the heavy rain and tropical cyclones within 400 km. And it can also monitor dragon volume, strong hail cloud, hail and other medium and small scale weather phenomena in the distance of greater than 200 km. The scope of the radial velocity measurement to achieve + 36 m/s. CINRAD/CC type weather radar body scan data finished nine times nine elevation scanning in 5 minutes. Each layer including the echo intensity (Z) and radial velocity (V), velocity spectrum width (W), there are 512 radial, each radial store 500 Z, 500 V, 500 W when radar scanning each layer.IDL include data display and analysis function through reading echo intensity 9 layer, and then 3D interpolation.
radar three-dimensional data display as shown in figure 5.IDL read and display radar data simply and quickly what's more it also analysis radar data. the body of data display as shown in figure 4, it uses function(OPENR READU) read binary files 2013081416.02 V, the first reads a header file which stored in a binary file for acquiring CINRAD/CC type weather radar data description information [11], such as radar's longitude and latitude, observation starting time, echo types, the antenna elevation Angle, whether correction, etc**.**

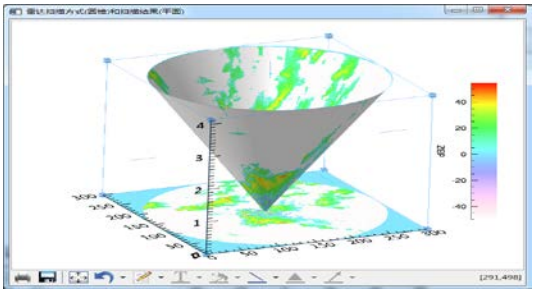


Fig.4. Radar scanning way (cone)

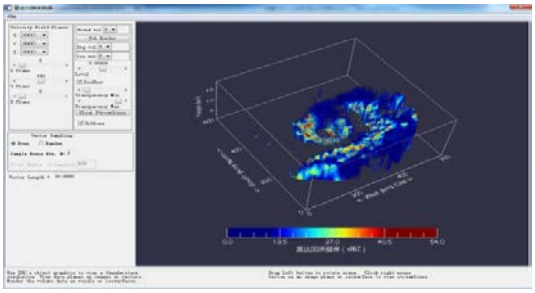


Fig. 5. Radar three-dimensional data show the graphic results

Part of the code as follows:

```
PRO  coor_trans,array, vdata, e,hdata,volumndata
  r_numda = lon*!PI/180.0;  r_phy = lat*!PI/180.0;
  r_h = h*0.001;
  res=300;//lon lat h Radar station of longitude and latitude and
high elevation
  dNdz=40;// linear gradient is 40
  y0=0
  EarthRadius=6371.1
  r_res = res*0.001 ;
  ry = y0 ;
  Km = 1/(1-EarthRadius*dNdz*0.000001) ;Km=4.0/3.0
  e=0.5;//elevation degree
```

IV .CONCLUSION

This article mainly introduce the IDL application in the weather especially in scientists measured the global precipitation, the application of 4D meteorological data visualization platform, doppler weather radar and 3D visualization.IDL can be used in the field of 2D meteorological and multidimensional meteorological data visualization,3D cloud meteorological graphical modeling, scientific data read meteorological cloud data because of the IDL powerful meteorological data analysis,which makes tens of thousands of meteorological data researchers, engineers and programmers quicken the process of meteorological research.

ISBN: 978-1-61804-327-6

REFERENCES

[1] Dian Wu Yan. Introduction to IDL visualization tools and increase [M]. China Machine Press,2003.06

[2] Pei You Han.IDL visualization analysis and application [M]. Northwestern Polytechnical University Press,2006.10

[3] Qiao Ren Wang,Li Xin Wang.A generation of weather forecast system and its technical development trend of the business.Wu Han center station of meteorology,2011.02

[4] Zhong Ying Pan.The 3D visualization method based on OpenGL. Computer Science and Engineering,2014.03

[5] Wen Hao Tu.3 D study of meteorological data.Journal of computer graphics,2008.06

[6] He Qun Yang,Hong Mei Zhou. The realization of FY - 3 meteorological satellite MERSI data preprocessing quickly.Remote sensing technology and application,2012.08

[7] Xian Feng Li.The research of meteorological observation data graphics base on Web and IDL .Meteorological hydrological Marine instruments,2014.09

[8] Yan Li,Feng Huang. The object of field visualization and the applications based on the meteorological element scalar.Meteorological science,2001.03

[9] Yi Yu Lu,Xin Min Tang. 4D track forecast based on dynamic weather forecast data.Aerospace computing technology,2013.06

[10] Bradley Zavodsky.Development and Implementation of Dynamic Scripts to Support Local Model Verification at National Weather Service Weather Forecast Offices,2014.02.03

[11] Wei Li. The analysis and application of doppler radar data on weather modification.Nanjing information engineering university,2005.03

Virtual Reality Contents using the OculusLift and Kinect

Dongik Lee, Giyeol Baek, Yangwon Lim and Hankyu Lim(corresponding author)
Department of Multimedia Engineering
Andong National University
Andong-si, Republic of Korea
yoha3567@naver.com, qkkndq@naver.com, limyw@anu.ac.kr, hklim@anu.ac.kr

***Abstract*—**Recently, in the game market, new kinds of peripheral devices in the controller and display area are being developed and games using them are being produced. Controllers diversely utilized in recent times are motion recognition controllers such as Kinect and Leap Motion, and as a display there is head mounted display (HMD). There have been various games and contents utilizing such peripheral devices but development of games and contents using them in combination is low. This paper designed and embodied virtual reality contents using Kinect, a motion recognition controller most actively used recently, and Oculus Lift, an HMD.

Keywords*—VR contents; HMD; oculus lift; kinect*

I. INTRODUCTION

Recently, motion recognition controllers and relevant controllers are being developed and sold through many home console game devices and personal computers (PCs) [1]. Among them, Kinect supports both Xbox and PCs and is utilized in diverse areas such as medicine and education as well as the game area. However, they are not being actively developed yet in other than the game area owing to the low recognition rate and level of awareness.

In the case of HMDs, at an early stage, wearing them was uncomfortable because of their size and weight and they had many problems such as low resolution and expensive price of displays, but thanks to the development of displays and improvement in their performance, diverse HMDs such as Sony's HMD, Oculus' Oculus Lift, and Samsung's Gear VR have been developed and the growth of the relevant market has been continuing [2].

This paper designed and embodied virtual reality contents using together Kinect, a motion recognition controller, and Oculus Lift, an HMD. The contents were to embody prototypes which moved characters utilizing Kinect and made display with Oculus Lift using C++ and DirectX11. Through this, developing the virtual reality game and contents proposed in the paper is expected to be possible.

II. THE RELATED RESEARCH

When the motion of the user is recognized using Kinect, the location information values of the user's skeleton may be derived. Information on the skeleton is composed of a total of 20 joints including humans' hands, feet, and head [3]. We utilized location information values of these joints and used them in order to move characters.

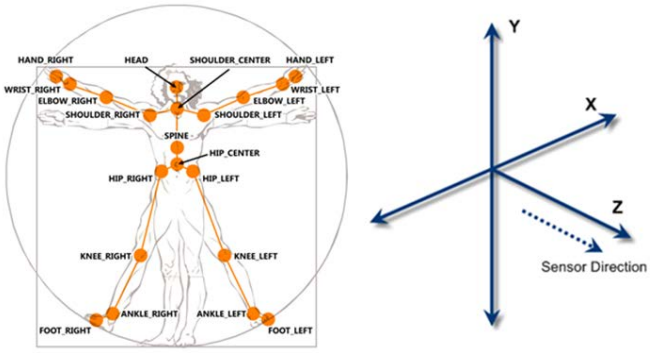


Fig. 1. Skeleton and Joint of Kinect

Although countless HMDs have been developed since the 1990s until now, the unique characteristics of Lift are wide field of view and rapid response speed [4]. There are two methods to create a game supporting Oculus Lift. When developing a game, the rendering screen of framework is set as Oculus Lift, and View Matrix and Projection Matrix, corresponding with the left and right eyes, are calculated, and fish eye rendering is made for each of the left and right screens, and there is a method of rendering by porting the existing game to be suitable for the Oculus Lift. vorpX is not supported for all games and in order to utilize it, the relevant program should be purchased separately. Here, fish eye rendering is to distort the screen and perform rendering in order to express as if one looks with the eyes, and wide field of view may be obtained through the convex lens. However, owning to the problem of pixel density by fish eye rendering, the pixel of the center area has relatively small image quality decline but the periphery has degraded image quality and during rendering the calculation ability to make fish eye rendering of the screen is needed [5].

ISBN: 978-1-61804-327-6

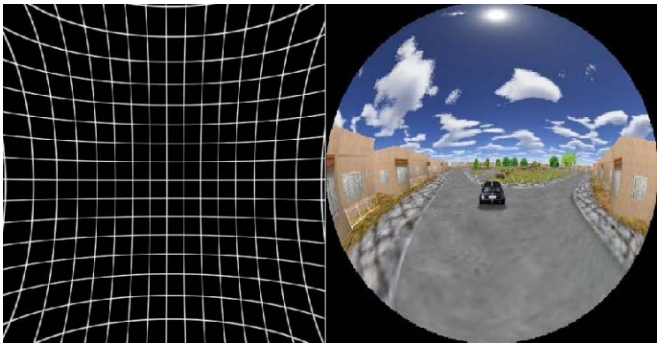


Fig. 2. Fish Eye Rendering

Another function of the Oculus Lift is head tracking, which is a function of applying changes in field of vision within the game and contents by sensing the head's rotation and forward and backward and up and down movements by the Oculus Lift.

III. Function and Implementation

A. Function Analysis

The functions of the program are divided into three in order to perceive a user's motions through this Kinect and utilize the perceived motions as data to move characters. First, Kinect has a function of perceiving a user's motions and manipulating characters of the game and is divided into head tracking of Oculus Lift, display, and game and contents. The figure 3 below is the perspective plan of virtual reality contents utilizing the Oculus Lift and Kinect.

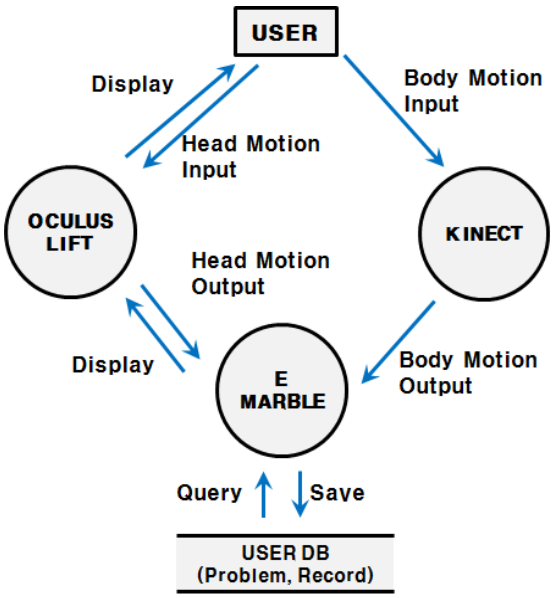


Fig. 3. DFD Structure_perspective plan

In the perspective plan of Figure 3, E Marble is the name of contents this paper designed and embodied. It is a main function of the game which is to bring information on the skeleton of the Kinect, manipulate characters and the game, bring information on head tracking from the Oculus Lift, calculate View Matrix and Projection Matrix, and perform rendering them to the Oculus Lift. Owing to the characteristics of the controller and the display, each function has high degree of mutual coupling.

B. Kinect Implementation

In order to move characters in the same way as the movements of the user, information on skeleton joints read from the Kinect should be applied to the character mesh's bones by calculating rotational matrix of each joint based on the information on the skeleton's joints. The Figure 4 below represents composition to correspond the character's bones with the Kinect's joints, and Kinect Tutorial and Avaterring Sample Source were referred to in order to derive the formula to calculate each joint's rotational matrix with information on joints.

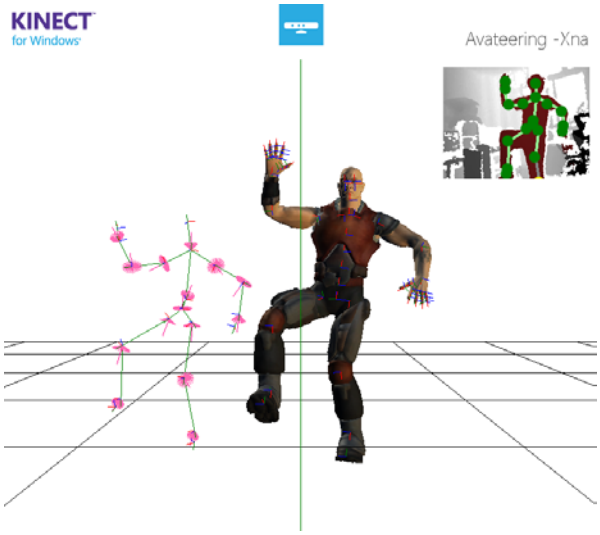


Fig. 4. Screen of Avateering C# Sample
(Left) Skeleton & Joint, (Middle) Character, (Right) Player

The Figure 5 below represents Kinect Skeleton Joint and Implementation Character Bone. Bone matrix calculated in such a way may be applied to character mesh through Skinning Animations.

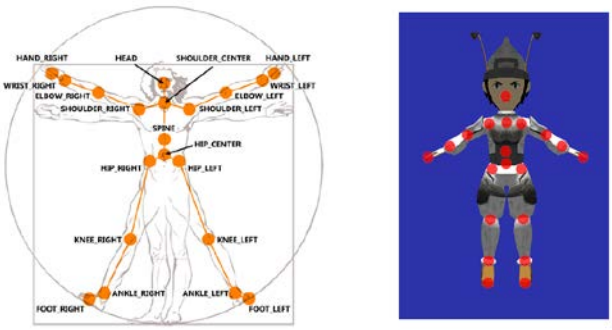


Fig. 5. Kinect Skeleton Joint and Implementation Character Bone

ISBN: 978-1-61804-327-6

C. Event Handling

The game embodied in this paper enabled virtual reality experience by using Kinect and Oculus Lift, but it is impossible to employ a mouse or keyboard for game manipulation. Therefore, manipulation should be made using Kinect as well. Such event was processed by setting a collider on the hands in the character's mesh and composing an object for manipulation and then detecting mutual collision.

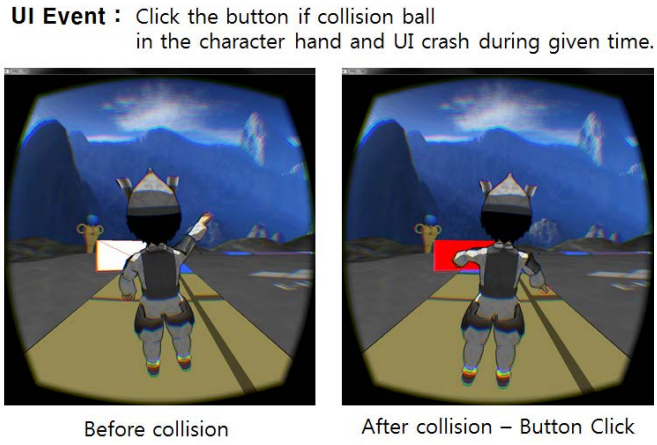


Fig. 6. UI Event Handling

Figure 6 above shows establishment of OBB(Oriented Bounding Box) aimed at event processing to the plan for UI output and event processing by testing collision with the sphere in the hands of the characters. The relevant character moves to the movement of a user and it was embodied for the user to click the button by extending the hand in order to trigger an event.

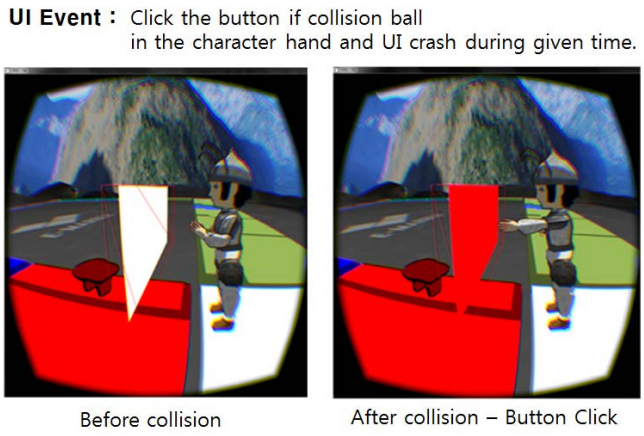


Fig. 7. UI Event Handling2 : View from Side

IV. Conclusion

At present, HMD and motion recognition controller have grown enough to produce games and contents, but still they have problems of high prices, low performance, and lack of key contents. Nonetheless, research and development on the area of virtual reality is continuously being made and performance of peripheral devices is being improved day by day and therefore they are good enough to be established as future new contents. In addition to the combination of Kinect and Oculus Lift proposed in this paper, utilization of other controllers will be possible to produce various virtual reality games and contents.

References

[1] Daegeun Kwon, Hyeongeun Shin, Yongje Kim, Jinyeong Jung, Seola Kim, Suncheon Jang, Daejun Lim, Yangwon Lim and Hankyu Lim, “Action simulation game design using the leap motion controller”, Proceeding of ICMEM 2014, Turkey, 2014, pp.381-391.

[2] Davis, Simon, Keith Nesbitt, and Eugene Nalivaiko. “Comparing the onset of cybersickness using the Oculus Rift and two virtual roller coasters.” Proceedings of the 11th Australasian Conference on Interactive Entertainment (IE 2015). Vol. 27. 2015, pp.30-33.

[3] Guo, B., Sun, J., Wei, Y. C., Meekhof, C., amd Leyvand, T. “Kinect identity: Technology and experience”, Computer, (4), 2011, pp. 94-96.

[4] Oculus Lift, https://en.wikipedia.org/wiki/Oculus_Rift

[5] Avateering C# Sample, https://msdn.microsoft.com/en-us/library/jj1 3104 1.aspx

e-Education VS traditional education: perspectives and challenges

Deniss Sceulovs
Riga Technical University
Faculty of Engineering Economics and Management
Riga, Latvia
deniss.sceulovs@rtu.lv

Elina Gaile-Sarkane
Riga Technical University
Faculty of Engineering Economics and Management
Riga, Latvia
elina.gaile-sarkane@rtu.lv

Elina Miezite
Riga Technical University
Faculty of Engineering Economics and Management
Riga, Latvia
elina.miezite@yahoo.com

***Abstract*—**E-education at higher education establishments helps cut the costs and its importance will continue growing in the future. Study methods for data analyses have been employed in the article. The monographic method has been used to obtain and collect theoretical information, whereas the qualitative study method — interviewing — has been used to find out the stakeholders' views. The quantitative study method — surveying — has been used to obtain information from stakeholders. The authors of the article distinguished between eight of the possible types of communications of stakeholders involved in e-education, as well as identified the tools used in the virtual environment in each of these types. The conceptual scheme of interaction between key players involved in the communication within e-education has been developed. Based on expert interviews, the most suitable areas of the e-education training process have been identified. Fundamental scientific research, laws of the Republic of Latvia and government decisions, general and special literature, national and EU statistical data, periodicals, public materials of organisations, and internet resources have been used to form the theoretical basis of the article.

***Keywords*—**e-education, stakeholders, information technology, communication, business.

I. Introduction

Education is a tool for gaining success, as it offers opportunities of not only self-development, but also a more successful career growth. Increasingly, knowledge and skills acquired specifically thanks to the offer of education establishments are what make an employee more competitive and consequently better appreciated on the labour market.

The labour market directly depends on the graduates and their competencies; therefore there should be long-term plans to ensure a sufficient number of specialists, bearing in mind that many professions take many years to master. Nowadays, when not only the competition on the labour market increases, but also the demand for qualified specialists rises, the education and acquired knowledge of an individual play a decisive part.

The education system is undergoing continuous changes, because it is under the direct influence of multiple factors, such as the overall global tendencies, development of technologies, changes in the national economy, falling demographic tendencies, increasing immigration, and accordingly changes in the labour market, which undeniably affect also the education system and force education establishments to introduce changes in the educational work. Overall, the number of students in Latvia over the last ten years has decreased from 130,706 to 89,671 [1], which is a reduction of more than 31 %. Higher education establishments are organisations the main aim of which is not to gain profit, however, the strategic plan is of great importance as there is competition also amongst universities — private and public universities alike are competing — because they must be able to present themselves to attract as many aspiring students as possible, who would appreciate the gained education, which in turn would be appreciated by employers and society on the whole. As the education system becomes ever more international, there is a range of new challenges that must be faced in the daily operations of universities; the simplest example would be a university's ability to ensure free communication of the teaching staff with international students, therefore it is very important to understand, how much the university can do. Owing to the rapid development of internet technologies over the last years, the e-education study methods at higher education establishments have gained an ever increasing popularity. One of the broadly recognised definitions has it that e-education is the use of information and communication technologies in the learning and teaching process [2]. Seeing as there is a lot of talk about specialists, who are in demand on the labour market, then the employer's opinions and evaluation must be taken into account, because occasionally stereotypes have been voiced claiming that e-education cannot replace

ISBN: 978-1-61804-327-6

traditional study methods. E-education provides access to education irrespective of the location, which is important for the modern global society, and it is also a way of implementing the currently very topical lifelong learning. "The development of e-training has resulted in increased importance of lifelong learning. Accordingly, the necessity originates from this movement to update knowledge and skills, requalify, because one job for a lifetime is no longer a relevant concept" [3]. Distance learning offered by universities, employing various e-education tools, is gaining ground around the world, ensuring studies in an international environment, which additionally ensures an exchange of experience and knowledge of students from various countries.

The authors of the article draw attention specifically to e-education trends in the higher education, in particular to the methods employed and the views of stakeholders. *The objective* of the article is, based on a study of scientific literature and practical research, to assess the e-education methods used by universities and their development trends in the view of the stakeholders.

II. The Most Recent Higher Education System Trends in E-Education

A. Trends in the application of e-education in the higher education

Influenced by globalisation processes, the higher education is undergoing major changes triggered by new technologies, especially by internet. According to statistics, during the period from 2000 until 2014, the number of internet users around the globe has increased by 741 % [4]. The development of communications and IT allow continuous education establishments to consolidate their positions in the area of education. It also ensures a future focus on lifelong learning for all, as it simultaneously spreads in conventional universities, which are increasingly employing distance learning methods in its work, thus, the differences between two types of establishments are almost irrelevant [5]. The availability of electronic education is gaining importance nowadays, and it is also facilitated by the development of various e-education tools and, accordingly, the environment, in which knowledge can be acquired. E-education, according to Oblinger and Hawkins, is defined as "course teaching and communication between teachers and students online" [6]. E-education is electronic delivery of learning, practice, or an education programme, which always involves the use of a computer or another electronic device [7]. On the European and international scope, education in a knowledge-based society becomes a strategic element for sustainable economic growth. The education environment must actively respond to external challenges, by demonstrating flexibility and openness to the new [8]. Electronic education has given an opportunity to the educators to reach out to an increasing number of students from all countries of the world, which otherwise would not even be possible [9].

The conventional higher education model has not changed for centuries. Establishment costs keep increasing due to the expensive investments in technologies, teaching staff's salaries and ever growing administrative costs. At the same time, the technological revolution is an important part of the higher education model. Online training is often free of charge, which implies that knowledge can be acquired by anybody, who has a smart phone or a computer. Universities are responsible for staff training and requalification throughout their career. However, will universities manage the ever increasing popularity of e-education and, if not, who will replace them? The demand for requalification and continuous education increases among the employed in all age groups around the globe. Globalisation and automatization have reduced the number of jobs, which require an average education level. Employees are investing every effort to top up their education thus trying to outpace the labour force demand curve. In the USA, higher education enrolment for students aged 35 and older increased by 314,000 in 1990-ties, whereas in 2000's, by 899,000. It is expected that 47 % of professions can be automatized over the upcoming decades [10]. According to *The Economist*, The future of universities is obtaining a degree in a digital environment [10]. However, this judgment can be disputed, because obtaining a degree in a digital environment has many cons. As the demand for education increases, online courses known as massive open online courses (MOOC) are becoming available as competition to universities. Clayton Christensen from the Harvard Business School believes that the strong technologies of MOOC can eliminate ineffective universities. "Within fifteen years, more than a half of universities [USA] will be on the brink of bankruptcy", he forecast in 2013. The first MOOC originated in Canada in 2008 as an online computing course. Online education has its shortcomings: only about 10 % of first time MOOC applicants complete the course. It is a problem when students involve somebody else to perform the tests online in their stead. Most universities and employers still regard online education as an additional option besides the conventional study courses and not as a replacement [10].

There are also several myths related to electronised education, such as that it bears practically no cost and is available to all who wish to have it. It requires careful planning for the part of the student and the university alike, because it can be checked in the system at any time whether the student has performed the specific task or not [11]. There is no single principle of operations for all universities — it should be developed. In an article published in *The Economist*, e-education is analysed and criticised, as well as referred to as the destructor of creativity. The revolution has started owing to three forces: the increasing cost, variable demand, and high technologies. MOOC is trying to bring down universities in various ways. However, the universities of Oxford and Harvard might even benefit. The most popular universities will be able to sell their MOOC around the globe, but average universities might experience a collapse. It is estimated that revenues of universities might decrease by more than one half, and employment in the sphere might reduce by nearly 30 %. In some ways MOOC will increase inequality among students (it will be more convenient for a more talented student rather than a weaker student outside the structured university environment) and among the faculty alike [12]. Last year, the Harvard Business School introduced HBX — an online training platform [13]. Since June 2014, within the framework of HBX, the following programmes are offered: financial accounting, business analysis and economics. The target group of these courses are bachelor level students and initially the training is

available to students of Massachusetts. It is planned to enrol 500–1000 students, and the learning duration is generally 60–90 hours. Before admission, the student's knowledge level is evaluated as there can be situations when a student has been working for 4 years at a bank, but does not know the difference between revenues and profits. In these situations, it is offered to choose the training programme online. They emphasise that, irrespective of our basic knowledge, these courses will help us understand business terminology and feel confident on the labour market [13].

The teaching approach is based on 2 conditions: firstly, not only hearing lectures, but also active learning and discussion. At the beginning of each lesson, the teacher is seen. First, the topic of the lecture is announced. During the learning process, videos shot by various professors are used. After the teacher's speech, students are urged to answer questions, a minute is given to explain the answer, and later it is commented and evaluated by other students. This course costs 1500 $, and it is currently coordinated by a team of 32 people [13]. It must be concluded that electronised education gives universities an opportunity to develop, and Harvard is aware of the fact that under the modern circumstances of competition something new must be created and it must be better and more competitive. E-online education has great future prospects in acquiring the humanities, whereas in exact sciences this type of training will not be as effective. The introduction of this type of training also requires sufficiently extensive financial investments, especially upon commencing a new course of lectures. Only recently, greater attention was paid to the impact of IT on learning and teaching at the higher education level, therefore the European University Association came to a decision to conduct a broad survey to better understand the current situation and to be able to support its members in their attempts to improve and foster innovation when learning and teaching. The survey results showed that, apart from several exceptions, nearly all institutions are involved in various e-learning forms [14].

The aim of the study was to find out the opinion of broader circles and to encourage discussions on further development in e-learning matters. Altogether, 249 answers were received from universities in 38 (the EU and elsewhere in Europe) countries. In surveys, it was asked about the application of e-learning in the functioning of establishments, about the experience in this area and future points of view. The most important study data indicate that mostly, out of the surveyed institutions, 91 % use mixed learning, by integrating e-learning in conventional teaching, 82 % of institutions also mentioned that they offer online learning courses. Less often, however with an increasing tendency, interinstitutional co-operation and online degree courses were mentioned. It was concluded that, besides the educational and economic motive, institutions refer to increasing needs for flexible timing and space, by using resources, which offer benefits to students, professional and other lifelong learning trainees alike [14]. In the EU university survey, institutions were also asked about the possible motivation of students to participate in online learning. As a result, 69 % of respondents said that combining work with education was an important motivational aspect. Furthermore, 63 % emphasised the remoteness of residence and 55% emphasised continuing education. Familial and social-economic situation has been mentioned by nearly three quarters of respondents [14]. Thus, the modern dynamic lifestyle does not always permit full dedication to the study process and the student's life.

B. Advantages and challenges of e-education

E-education offers extensive opportunities for aspiring students, however, to be able to evaluate what it can give, advantages must be identified, along with challenges. Having conducted studies, when senior level managers were surveyed, it is suggested that organisations classify benefits of e-education according to different features.

TABLE I. E-EDUCATION BENEFITS [15]

Feature	E-education benefit
By improving resource use	savings on travel/commute time and accommodation costs when the student has to attend the training in person; cost-effective training courses; less time needed to introduce improvements, updates; better use of available technologies; possible bulk financial savings when the number of trainees is large; 24-hour availability
Geographic reach	flexibility — the course can be mastered anywhere at any time; people from remote areas are ensured access; faster transfer of knowledge; several simultaneous training courses.
Organisational benefits	achieving best performance offered by technologies; support to employees at places of employment — (just-in-time) learning improves problem-solving abilities; demonstrate the actual employee development; profitable benefit within the framework of continuing education.
Improved learning	self-instruction; increasing individual management abilities; options of individual choice, ensuring alternative approaches to learning and development.

Mostly, all of the aforementioned benefits can also be referred to higher education organisations, however, there are also properties aimed specifically at the academic environment.

Scientists have identified several strengths of e-education: it ensures student-focused learning; learning at any time and place; facilitates interaction between the student and contents of the study subject; facilitates and encourages communication and co-operation; facilitates course administration; helps keep track of time that the student spends to perform the task; helps keep track of the time that the student spends to complete a task; can reduce costs related to delivering instructions; adds a global dimension to the course [16]; students can have an option to choose the learning materials according to their level of knowledge and of interest; self-instruction allows a student to learn at his/her own pace; flexible approach at any time to discussion forums to communicate with course mates and teachers; additionally, skills will be improved when learning with the intermediation of internet and computer; an option to choose the educational establishment, accordingly, if only few are nearby geographically, then e-education ensures access to any establishment, and only the desired area is then chosen; e-education is also mentioned as a solution for specific national problems: too many students in a class, excessive conventional education fee [18].

ISBN: 978-1-61804-327-6

Many organisations are using e-education learning as they can be as effective as conventional learning, at lower costs. To create e-education learning is significantly more expensive than preparing material for a classroom and to train lecturers, especially if multimedia or high interactivity methods are employed. Nevertheless, the overall e-learning costs, including servers and technical support, are much lower than to purchase equipment for classrooms, the time of lecturers, the students' travel time and missing work to attend classes [19]. Even though there are many positive aspects to acquiring e-education, a range of problems can still arise. For instance, dependence on technologies — students will need access with minimum specifications as determined by the creator of e-education training; material incompatibility — possibly, what works in one browser, does not in another; not applicable in all educational sciences — what is learnt in classroom, can be supplemented with e-education learning; not suitable for all students, learning in the electronic environment demands a high level of self-discipline, as well as must be very self-motivated [20], students having low motivation or poor habits can lag behind; without a structured schedule and coming to lectures, the student can be confused as regards the course activities and deadlines of final assignments; students cannot ask questions right away, answers are given only after some time, when the teacher has received it in their e-mail or elsewhere; not possible to identify the person, who is mastering the course, and it is possible that the tasks are performed by somebody else; the student can feel isolated from the course teacher/instructor and course mates, insufficient contact with the teacher [21]; slow internet or older computers can encumber access to course materials; laboratory and practical work is difficult to simulate in a virtual environment [17]; limited English and IT knowledge can hinder engagement in e-learning courses [22]; not all employers recognise degrees obtained in an electronic environment [23]. Likewise, the course language is of importance, since not all programmes are available in a languages we understand, and e-training can simply be unsuitable for various psychological types of people.

III. E-education methods and their classification

The beginnings of e-learning are linked to creating such approaches, which are intended to adapt distance learning [24]. E-learning means that learning is no longer a passive experience, where students are sitting opposite the teacher and learning by telling; e-training is an active acquisition of experience. E-training has the possibility of using a custom and individual learning approach of sorts, because we all learn at different paces: when reading, watching, investigating, discovering, doing, discussing, co-operating, sharing our knowledge, accordingly the students involved in e-training have an access to extensive learning resources. The co-operation tools that are used in e-training can be classified depending on whether or not synchronous or asynchronous co-operation is ensured. Synchronous communication, also referred to as a real-time conference, is possible when all participants are online simultaneously. Means to communicate synchronously: chatting, sharing software, using an interactive board, audio conferences, video conferences. Asynchronous communication does not require all participants to be online simultaneously. By using asynchronous communication, a participant sends a message at a time convenient for him/her. E-mails and discussion forums are used in asynchronous communication [24]. Synchronous communication ensures quick support to the student by other students or teachers, whereas asynchronous communication allows time flexibility, however time must be planned and it must be taken into account that the addressee also works by the same principle and, accordingly, a longer time can pass until a reply is received. Classification of e-education learning methods according to the presence of a teacher: ***physical presence*** (face-to-face, use of additional visual aids in front of the classroom — presentations, images, videos, etc.); ***virtual presence*** *(*e-learning — online learning, computer-based learning, internet-based learning, audio learning, video learning, distance learning, online courses MOOC;) ***Physical, virtual presence*** *(blended learning).*Not one and the same training model can be applied to all, hoping to achieve equally good results. Therefore, it is important for everybody to know, which type of perception we represent, as well as for educational establishments to take that into account and adapt the acquisition of the learning process insofar as possible to all types of perception.

Authors have distinguished between 8 types of stakeholder communication, based on four units, and have identified them — an individual, a robot, a group, a teacher — involved in e-learning **An individual** — a single person, who wishes to acquire e-education at a time and place suitable for them; **peer (artificial intelligence) —** a set of information prepared in advance, which is mastered with the intermediation of technologies by a group or an individual, or, in other words, a university, which ensures communication with the interaction of the teaching staff and artificial intelligence; **a group —** a community of several people, who wish to acquire e-education, they are not restricted by geographical conditions, but are restricted by time limitations, the dates of starting and completing e-learning; **the lecturer —** a person, with whom synchronous or asynchronous communication besides the robot is possible.

Eight such types of stakeholder communication were identified: **1. Person to Peer** — the person, who wishes to acquire e-education, has access to the necessary technologies to master the topic, and he/she can access the contents of the study programme from any place at any time convenient for them. Acquisition of e-education by employing this type does not restrict the individual to time or space. **2. Group to Peer** — several people with computers and an internet connection start the acquisition of e-education in a specific period, and a scope of contents is determined, or deadlines that are integrated in the programme, which must be completed at a specific time. In a group, restrictions in time, but not space, apply. **3. Person to Teacher** — the individual has the necessary technologies, the material prepared in advance to master the topic, however, synchronous or asynchronous communication with the teacher is possible to ensure a more in-depth understanding of the topic to be mastered. **4. Group to Teacher** — several people are ensured with material prepared in advance and with technologies, as well as additionally synchronous or asynchronous communication process is occurring with the lecturer to improve the knowledge, in which not all group members must participate. **5. Person to Person —** an individual learning the material at a

ISBN: 978-1-61804-327-6

time convenient for him/her, is given an opportunity to communicate with another individual, who already has acquired the topic or is still in the process of acquisition. **6. Person to Group —** an individual learning the material at a time convenient for him/her, an opportunity is given to communicate with a group of individuals, who already have acquired the topic or are still in the process of acquisition. **7. Person to Person in the group —** one person, who is learning the topic together with other group members, communicates with another individual, who is also a group member, who is acquiring the same topic and who is subject to a performance deadline. **8. Person to group —** a single individual, who is mastering the topic within a group, communicates with other group members, who are united by the acquisition of the same topic or an opinion, as well as are subject to deadlines.

TABLE II. TYPES OF PARTIES INVOLVED IN THE PROCESS OF E-LEARNING COMMUNICATION AND THE TYPES OF USED E-LEARNING TOOLS [AUTHOR'S ORIGINAL]

	1	2	3	4	5	6	7	8
Tests with immediate results	x	x						
Tests with the option of inserting a textual or visual material for assessment	x	x						
Web tour — introduce with a web (A)*	x	x	x	x				
Use of software to show activities (A)	x	x	x	x				
3D game simulation and model-based learning	x	x						
Study materials placed on intranet	x	x						
Online surveys and quizzes	x	x						
Submitting assignments online	x	x						
E-books	x	x	x	x				
MOOCS		x		x				x
Open-ended questions in assessment			x	x				
E-mail (A)			x	x	x	x	x	x
Chatrooms (S)**			x	x	x		x	
Video conferences (S)			x	x	x	x	x	x
Audio conferences (S)			x	x	x	x	x	x
Use of an interactive board (S)			x	x				
Discussion forums				x		x		X
Blog and wiki sites				x		x		x

Notes: A – asynchronous, S – synchronous, numbers 1–8 correspond to the classification of types of communication of stakeholders.

It must be emphasised that in the individual-peer and group-peer types of communication, initially, the teacher is also involved, who prepares in advance and places the learning contents in the relevant environment, which the students must later present to peers, if the students do not have an access to the individual-teacher or group-teacher types of communication, then they completely lack communication with the teacher. On rare occasions, the teacher is not involved at all in the course development — they are prepared by companies specifically trained for that. The uneven distribution seen in Table II between the use of methods points to that different methods are used in the communication between the peer and teacher with another student. The classification also portrays the fact that many e-learning tools overlap in a number of communication types. Out of the mentioned 8 types, the authors propose the first option as the most suitable for e-training — when an individual and a peer are involved in the communication, as this situation most accurately corresponds to when a student himself/herself makes a decision as to when, where and what he/she will be learning. It must be mentioned that discipline and motivation are required to organise the daily schedule in this way. The group-peer type of communication restricts the student in time, however other types of communication are rather a supplement to the already existing e-learning. In order for the student not only acquire the prepared content, but also to gain an in-depth understanding and to communicate with others, for whom the specific matters are important, communication with the teacher and group mates, who have already started or start learning the specific topic, is possible. Practical applicability of the classification: allows the higher education establishments to clearly determine which types of communication are implemented and which are not, and such classification offers diversification of the methods used in the learning process, as different experiences are ensured for the student. The uneven distribution of communication tools points to the need of using various tools so that the learning content is acquired in the virtual environment and it would enable discussions with other students and the teacher.

IV. STAKEHOLDERS IN THE HIGHER EDUCATION AND DETERMINATION OF THEIR EXPECTATIONS

Nowadays, companies and society determine that the success of an organisation depends on the stakeholder evaluation. The organisation's approach to human resources management plays an important role in meeting the needs of all stakeholders [26]. This also applies to higher education establishments. Stakeholders are benefiting from the organisation's success or, on the contrary, suffer from failures and errors. The organisation is generally interested in maintaining good relations with the key stakeholders. If some of the stakeholders end relations, the organisation might suffer [26]. It is important for the organisation to always be aware of the stakeholders' interests to maintain good relations, which means further co-operation and development. It is concluded that it is important to identify all stakeholders and their concerns in various matters.

A university is a complex structure consisting of several stakeholders: the administration, faculty, students, graduates, parents and the organisation itself. Each of them promotes university culture and can affect the consumption decision-making [27]. According to the Law on Institutions of Higher Education [28], the identified stakeholders are the academic staff, scientific staff, and the student. The activity of the scientific staff is regulated according to the Law on Scientific Activity [29]. Within the article, the stakeholders of the higher education will be considered from the market perspective, taking into account the perspective of the law. Classification of stakeholders of the higher education entails two categories: primary and secondary. The basic group consists of stakeholders, which ensure the performance of key tasks: the government, the Ministry of Education and Science, study departments, society. The second group: students, staff, competition, agency for the education and training of faculty [30]. Another option of distinguishing between various stakeholder categories could be: internal and external stakeholders of higher education establishments. Internal stakeholders are those, who are participating in the daily

ISBN: 978-1-61804-327-6

activities of an establishment: academic and administrative staff, students. External stakeholders or involved persons are those, who are interested in higher education, even though they are not members of the higher education community. They are parents, tax payers, employers, the State and even international organisations. It is important to define two different categories of stakeholders because changes in the higher education point to reduced influence of internal stakeholders, thus external stakeholders' role is ever increasing [31]. The increased importance and role of external stakeholders is mentioned by many authors [32]. Academics are losing their former dominant power over universities [33]. This is directly linked to the increasing control of the State and the market in the higher education. Stakeholders of higher education establishments can also be classified according to four dimensions: firstly, by determining whether the stakeholders are internal or external; a second aspect is to what extent the stakeholder group is active or passive with regard to the establishment or a specific project, such as, integration of e-learning in a university; thirdly, to what extent the stakeholder is co-operating with the establishment and supporting the initiative or a project; and, finally, matters on the extent of threats that the stakeholder can cause to the initiative or the project [34]. Based on the stakeholder classification, it is easier to prepare further development models, being aware of the influence, needs, and expectations of each stakeholder. To achieve success, it is not only sufficient to introduce technologies — changes are necessary in the educational approach. In this process, universities are subject to changing economic circumstances and university budgets are affected by economic fluctuations. Even though there are institutions which have introduced e-learning as a new source of income, it is, nevertheless, difficult for them to successfully gain revenues due to the high competition and the global market [35]. Although there are some important similarities between companies and universities, there are also important differences. Companies are trying to increase profits, while universities, with some exceptions, are trying to improve something completely different — prestige. Everybody, who is affiliated to the university either as a student, a faculty member, an administrator, or a graduate, is important as regards the prestige of the university [36]. Universities, just like companies, are making a range of products that are in demand in various market segments. It is clear that the product of universities is education for its graduates. In case of companies, value is added to raw materials, however, in case of universities; a significant difference is that they add value for its clients.

In the last decade, an important point of discussion has been the co-operation between the higher education and the market, by stimulating the development of new sectors, as well as maintaining a qualified people/social capital [37]. This situation has originated due to overproduction of students in one sector, but a significant deficiency in another. Most stakeholders agree that the teacher education must change in the face of students and challenges in a digital era in technology-based society. However, obstacles are presented by the lack of a different policy or vision in co-operation with the stakeholders [38].

Students' expectations with regard to the higher education have increased over the past few years. The academic staff expects a certain level of student engagement in the studies and a high academic standard in the students' course work performance. Nevertheless, the university staff must understand that conventional evaluation, such as A levels in the UK, SATS and ACTS in the USA, etc., is no longer the primary qualification indicator when enrolling in a university, thus, there might be students with differing academic abilities in a single course [39]. The lack of skills in important labour force areas (contrary to unemployment in others), as well as the employers' expectations concerning the graduates' skills makes governments (and employers) set higher requirements to the higher education establishments. Furthermore, while the willingness to be employed in the future does not always serve as the grounds for acquiring university education, it is the desirable educational outcome among most students. It is expected that the graduates will acquire the necessary knowledge and skills to be able to adapt to the labour market thus taking into account the students' wishes, labour market demands, government plans, as well as the employers' and general public's expectations [40]. Students' motivation to choose e-learning is most often based on convenience and availability [41]. Motivation regarding the e-learning will increase once they feel that efforts are invested to meet their expectations and the added value as individuals [42]. Students appreciate access and support from the e-learning instructor or teacher [43]. A suitable and pleasant learning environment is expected from the study department [44]. Employers need graduates, who can contribute by immediately starting to work. However, employers appreciate the university graduates' ability to quickly acquire new information [45]. In various ways, the employers' influence on higher education programmes is ever increasing: dissatisfaction is voiced with graduates as 'a product'; there are specific expectations concerning the skills and attitude; they become the higher education providers by spotting differences in the regular higher education system [46].

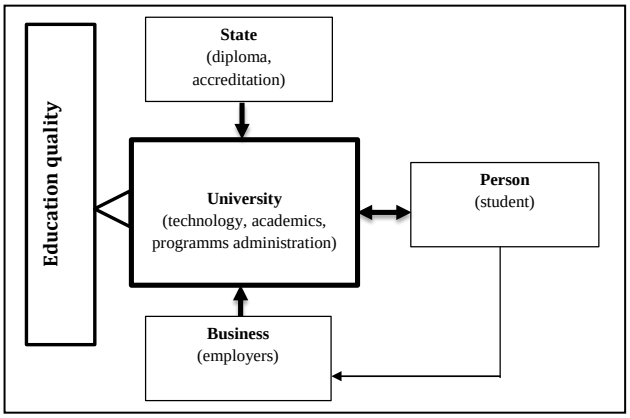


Fig.1. A conceptual scheme of interaction between e-education key stakeholders [authors' original]

The conceptual scheme is based on interaction between the stakeholders. The most important aspect is that the communication process for an individual can be only virtual, whereas, in the types of communication 'State-university-employers', physical and virtual communication exists.

Therefore, it is important to assess the students' opinion of the received educational content, the process, acquired knowledge and other aspects, because he/she is the main consumer of studies.

Entrepreneurs are interested to hire specialists who are as qualified as possible and who could, within the job duties, use their theoretical and practical knowledge, therefore entrepreneurs communicate with educational establishments on the programme contents and their correspondence to the actual labour environment, to increase the quality of the labour process, which can be done, by engaging professionals who better meet the requirements. *The task of the university* is to hear out the employers, the state, and the students, to maintain its prestige and gain profits. *The State*, for its part, is responsible for ensuring that the university is operating in line with all educational requirements and in compliance with all regulatory enactments, so that the State could accredit the university and recognise its issued diploma.

V. ASSESSMENT OF E-EDUCATION RELEVANCE IN LATVIA

To find out the opinion about e-education, a questionnaire survey was conducted for two respondent groups: 1. individuals, 2. employers (managers, company employees related to staff matters). Survey period: April 2015. Respondent base: 50 companies and 70 individuals (students). The questionnaire consisted of 10 questions for employers and 9 questions for individuals.

78 % of employers and 85 % individuals recognised that e-education is very relevant or relevant. For the most part, employers and individuals alike recognised that e-learning in a virtual environment in future could not fully replace conventional on-site education (only 14 % of employers and 21 % of individuals answered rather yes than no), which could be explained with companies being a part of different sectors. The most important advantages: learning at any time and place, lower costs; student-focused learning and the possibility of learning at your own pace; e-learning can be better combined with other educational programmes and work. If comparing the assessment by employers and individuals concerning the e-education advantages, it was concluded that the time and place is of the essence, and that can be ensured by learning in a virtual environment. It is important for an employer, who hires an educated employee, as well as for an individual, who, without losing a job, can continue studying at a high level; likewise, the financial aspect is important, because the acquisition of conventional education demands extensive resources throughout the study process, whereas in the case of e-education, most resources are spent during the programme development stage. Among shortcomings pointed out by employers, there are the requirement for a high self-discipline of the student and insufficient communication with the teacher. The most important shortcoming indicated in the survey of individuals was the language barrier. The second most important shortcoming is the unsuitability for all educational sciences, which leads to a conclusion that, no matter how motivated the student, there are sectors, in which e-education acquisition opportunities simply will not be available. Among other important shortcomings, there are the need of a high student self-discipline and lack of communication with the teacher, no immediate feedback from the teacher/lecturer, who could help correct errors, which could constitute problems of the lack of development of socialisation skills, which prevents developing such skills as team work, a shared project, public presentation, as well as improving practical skills. For examples, how could doctors learn to perform a surgery without professional supervision? There are rather many respondents, who believe that unsuitability for all students or all sciences is a minor deficiency. 70 % of employers recognised that there is an opinion of e-education as incapable of ensuring the same as conventional education. This proves that these methods of acquiring education due to stereotypes are assessed negatively or there has been negative experience with educational establishments, which have provided this educational service, accordingly not meeting the quality or requirements. Job duties are different when learning about science and practically working in the sphere. Therefore, the graduate receives training in the practical work and specifics at the company, which could be applicable to 20 % of respondents, who answered not, or rather not than yes. Most managers admit that they are rather appreciated or rather yes than not. However a large share of respondents, 28 % of all, are not informed, possibly, about the acquired employees' skills or are not informed due to other reasons. General tendencies show that e-education is topical among employers and individuals, that it has equal number of advantages and disadvantages, which suggest that in the nearest future it will still serve as a supplement to the conventional education approach.

CONCLUSION

There is no single system regarding the models of presenting e-learning. In line with the authors' developed classification, by combining several types of stakeholders involved in the e-learning process, a very extensive use of tools used in the virtual environment can be achieved, ensuring versatility of the learning process and allowing the student to choose study contents that are more convenient to grasp (video, audio, text, etc.). The conceptual scheme of ensuring the quality of e-education which is based on interaction between stakeholders proves that for an individual the communication process can be purely virtual, however in the communication types 'the State-university-employers', physical and virtual communication exists, which points to the need for the higher education establishment to assess the student's opinion of the received study contents, process, acquired knowledge and other aspects, because it is the student, who is the main consumer of studies, whereas the university is the main provider of education quality.

REFERENCES

[1] Educational organizations and number of students [online]. Rīga: LR CSB, 2014, [viewed 10 Feb 2015] http://data.csb.gov.lv/pxweb/lv/Sociala/Sociala__ikgad__izgl/IZ0010.px

[2] T. Mayisela. Academic Staff and Student Adoption Rate of Blended Learning at Walter Sisulu University, The Proceedings of the 6th International Conference on e-Learning. Balcaen, P. ed. UK: Academic Publishing Limited, 2011. pp.219-226.

[3] R. Mason, F. Rennie. Elearning: The Key Concepts [online]. USA: Routledge, 2006. 200 p. [viewed 10 Feb 2015]. e-ISBN 978-0-203-09948-3. https://play.google.com/books/reader?id=gZx-AgAAQBAJ&printsec=frontcover&output=reader&hl=lv&pg=GBS.PT15.w.10.0.5>

[4] Internet Users in the World Distribution by World Regions – 2014 Q2 [online]. Miniwatts Marketing Group, 2014. [viewed 18 Jan 2015]. http://www.internetworldstats.com/stats.htm>

[5] T. Varis. e-Learning and Higher education. In: eLearning Conference "Toward a learning society". May 19-20, 2005.University of Tampera, 2005. p.1.

[6] D.G. Oblinger, B.L. Hawkins. The Myth about E-Learning [online]. Educause, 2005 [viewed 25 Nov 2014]. https://net.educause.edu/ir/library/pdf/erm05411.pdf>

[7] D. Stockley. E-learning Definition and Explanation (Elearning, Online Training, Online Learning) [online]. Australia: Derek Stockley, 2003 [viewed 25 Nov 2014]. http://www.derekstockley.com.au/elearning-definition.html>

[8] E.C. Raluca, E. Avram, A. Tufan. Educational inovation and consumer behaviour. A study of student` perceptions on the use of e-learning in class. In Annals of the University of Oradea, Economic Science Series, 2011, Vol. 20 Issue 1. pp. 736-740.

[9] A. Edmundson. Globalized E-Learning Cultural Challenges. United States of America: Idea Group Inc, 2006. p.2.

[10] The future of universities, The digital degree [online]. UK: The Economist, 2014 [viewed 23 Nov 2014]. http://www.economist.com/news/briefing/21605899-staid-higher-education-business-about-experience-welcome-earthquake-digital.

[11] K. Shelton, G. Saltsman. An Administrator`s Guide to Online Education. USA: Information Age Publishing, Inc, 2005. p.4.

[12] M. Hart. Proceedings of the 5th European Conference on elearning. University of Winchester, UK, September 11-12, 2006. p. 422.

[13] R. Knight. Harvard takes a step into the world of Moocs [online]. The Financial Times LTD, 2014 [viewed 23 Nov 2014]. http://www.ft.com/cms/s/2/674f02f8-ade8-11e3-9ddc-00144feab7de.html.

[14] M. Gaebel, V. Kupriyanova, R. Morais, E. Colucci. E-learning in European Higher education institutions. Results of a mapping survey conducted in October-December 2013. [online]. Belgium: European University Associatio, 2014, p.6-7, 29.

[15] D. Birchall, D.Woolfall, Corporate e-learning: Delivering business benefits. UK: Grist Ltd, 2003. p.6.

[16] S. Waterhouse. The Power of e-Learning: The Essential Guide for Teaching in the Digital Age. New York: Pearson Education, 2005. p.10.

[17] Advantages and Disadvantages of eLearning [online]. ISU, 2001. [viewed 7 Feb 2015]. http://www.dso.iastate.edu/asc/academic/elearner/advantage.html.

[18] M.F.E Zayat. A Strategy to Improve E-Learning Adoption, Implementation and Development in Higher Education in Egypt [online]. Academia.edu, December 2008. p.5-6. [viewed 23 Feb 2015]. https://www.academia.edu/4028842/A_Strategy_to_Improve_e-Learning_Adoption_Implementation_and_Development_in_Higher_Education_in_Egypt.

[19] E-learning methodologies. A guide for designing and developing courses [online]. Rome: FAO, 2011. p.8-9. [viewed 13 Feb 2015] http://www.fao.org/docrep/015/i2516e/i2516e.pdf.

[20] R. Mobbs. Disadvantages of eLearning [online]. December, 2003. [viewed 13 Feb 2015]. http://www.le.ac.uk/users/rjm1/etutor/elearning/disadvofelearning.html.

[21] N. Prakashan, "Current Trends in Management. PUNE, 2007.

[22] J. Bolton, A. Braddell, S. Dunn. Learning technologies in social care: A guide for employers. 1 edition. Skills for Care, 2014. p. 60-61.

[23] Advantages and Disadvantages of Distance Learning [online]. Distance Learning Net, 2015 [viewed 21 Feb 2015]. http://www.distancelearningnet.com/advantages-and-disadvantages-of-distance-learning.

[24] J. Gold, R. Thorpe, A. Mumford. Gower Handbook of Leadership and Management Development. Fifth edition. England: Gower Publishing Limited, 2010. p.442.

[25] S. Nangia. S., B. Global Trends in Teacher Education. New Delhi: APH Publishing Corporation, 2007. p. 136.

[26] S. Jackson, R. Schuler, S. Werner, S. Managing Human Resources. USA: Cengage Learning, 2011. p.4.

[27] C. Gilde. Higher Education: Open for Business. UK: Lexington Books, 2007. p. 7.

[28] LR Highschool law [online]. Riga: Legislation of the Republic of Latvia vortal, 1995 [viewed 15 Mar 2015] http://likumi.lv/doc.php?id=37967

[29] Scientific activity law [online]. Rīga: Legislation of the Republic of Latvia, 2005 [viewed 30 Mar 2015]. http://likumi.lv/doc.php?id=107337

[30] I. Maric. Stakeholder analisys of higher education institutions. Interdisciplinary Description of Complex Systems, 2013, vol.11, iss.2, pp. 222-223.

[31] A. Amaral, G. Jones G., B. Karseth. Govering Higher Education: National Perspectives on Institutional Governance. The Netherlands: Kluwer Academic Publisher, 2002. p.11.

[32] A. Bernhard. Quality Assurance in an International Higher Education Area: A Case Study Approach and Comparative Analysis. Germany: Spinger Science Business Media, 2011. p.22.

[33] Study directions accreditation [online]. Riga: Ministry of Education and Science, 2015. [viewed 20 Apr 2015]. http://www.izm.gov.lv/lv/izglitiba/augstaka-izglitiba/studiju-virzienu-akreditacija.

[34] J. Burrows. Going Beyond Labels: A Framework for Profiling Institutional Stakeholders. USA: Contemporary Education, 1999, vol. 70, iss. 4, pp. 5-8.

[35] F.Z. Moser. The Strategic Management of E-Learning Support. Munster: Waxmann Verlag, 2007. p.87.

[36] W. Bienkowski, C.J. Brada, G. Stanley. The University in the Age of Globalisation: Rankings, Resources and Reforms. UK: Palgrave Macmillan, 2012. p.66-67.

[37] B. Galbraith. Higher Education, Stakeholders and Collaborative Work for Entrepreneurial Learning. In: 9th European Conference on Innovation and Entrepreneurship: ECIE2014. September 18-19, 2014. UK: Academic Conferences Limited, 2014. p.322.

[38] R. Mohan. Teacher Education. New Delhi:PHI Learning Ltd, 2011. p.8.

[39] M. Morgan. Supporting Student Diversity in Higher Education: A practical guide [online]. [viewed 2015 2015]. https://books.google.lv/books?id=YdgcAAAAQBAJ&pg=PT49&dq=government+expectation+of+supporting+higher+education&hl=lv&sa=X&ei=fbcZVZTwBoWPsAHE1oOQAQ&ved=0CD0Q6AEwBA#v=onepage&q=government%20expectation%20of%20supporting%20higher%20education&f=false.

[40] S. Marshall, R. Henry, P. Ramburuth. Improving Assessment in higher Education: A Whole of Institution Approach. UNSW Press, 2013. p.6.

[41] A. Edmundson. Globalized E-Learning Cultural Challenges. United States of America: Idea Group Inc, 2006. p.204.

[42] V.P. Dennen, A. Darabi, L.J. Smith. Instructor-learner interaction in online courses: The relative perceived importance of particular instructor actions on performance and satisfaction. Distance Education, 2007. Vol. 28, iss. 1, 65-79. pp. doi:10.1080/01587910701305319

[43] M. Simonson, C. Schlosser, C. Quarterly Review of Distance Education. IAP: 2009, vol. 10, Nr. 4. 352.p.

[44] P. Kumarawadu P. Motivation of online learners: Review of practices and emerging trends [online]. [viewed 15 Mar 2015]. Sri Lanka: Sri Lanka Institute of Information Technology, 2001. http://www2.uca.es/ordenacion/formacion/docs/jifpev5-doc5.pdf.

[45] F. Stephen, C. Steven. Integrating Key Skills in Higher Education: Employability, Transferable Skills and Learning for Life. Routledge, 2013. p.128.

[46] K. Cox, W.B. Imrie, Student Assessment in Higher Education: A handbook for Assessing Performance, London: Kogan Page Limited, 2014,p.13.

Quasi-nilpotent equivalence of weakly decomposable operators

Cristina Serbanescu
Faculty of Applied Sciences
University Politehnica of Bucharest
E-mail: mserbanescuc@yahoo.com

Ioan Bacalu
Faculty of Applied Sciences
University Politehnica of Bucharest
Email: dragosx@yahoo.com

***Abstract*—In this paper we study several properties of weakly decomposable operators, which are defined as a generalization of decomposable operators. We open the way for defining other categories of operators with similar properties to decomposable operators to be used in Differential Equations Theory and in Fractal Theory.**

AMS 2000 Mathematics Subject Classification: 47B47, 47B40. **Keywords:** decomposable; weakly decomposable; spectral equivalence; commutator.

I. Introduction

Let X be a Banach space, let $\mathbf{B}(X)$ be the algebra of all linear bounded operators on X and let $\mathbb{C}$ be the field of complex numbers. In what follows, by subspace of X we understand a closed linear manifold of X. If $T \in \mathbf{B}(X)$ and Y is a linear subspace of X, let us denote by $T|Y$ the restriction of T to Y. An operator $T \in \mathbf{B}(X)$ is said to have the *single-valued extension property* if for any analytic function $g\colon G \to X$ (where $G \subset \mathbb{C}$ open) with $(\lambda I - T)g(\lambda) \equiv 0$, it results that $g(\lambda) \equiv 0$ ([2] [8]). For an operator $T \in \mathbf{B}(X)$ having the single-valued extension property and for $x \in X$, we consider the set $\rho_T(x)$ of all elements $\lambda_0 \in \mathbb{C}$ such that there is a X-valued analytic function $\lambda \to x_T(\lambda)$ defined on a neighborhood of λ_0 which verifies $(\lambda I - T)x_T(\lambda) \equiv x$; $x_T(\lambda)$ is unique, $\rho_T(x)$ is open and $\rho(T) \subset \rho_T(x)$. We also denote $\sigma_T(x) = \complement\rho_T(x) = \mathbb{C} \setminus \rho_T(x)$ and

$$X_T(F) = \{x \in X; \sigma_T(x) \subset F, F \subset \mathbb{C}\}.$$

$\rho_T(x)$ is called the *local resolvent set of x with respect to T* and $\sigma_T(x)$ is the *local spectrum of x with respect to T*.
A subspace $Y \subset X$ is called *spectral maximal space* of $T \in \mathbf{B}(X)$ if it is invariant to T and for any other subspace $Z \subset X$, also invariant to T, the inclusion $\sigma(T|Z) \subset \sigma(T|Y)$ implies the inclusion $Z \subset Y$ ([2], [8]).

Definition I.1. An operator $T \in \mathbf{B}(X)$ is called *weakly decomposable* if

(i) T has the single-valued extension property and $X_T(F)$ is closed, for any $F \subset \mathbb{C}$ closed;
(ii) for any finite open covering $\{G_i\}_{i=1}^n$ of $\sigma(T)$, there is a system $\{Y_i\}_{i=1}^n$ of spectral maximal spaces of T such that
  1) $\sigma(T|Y_i) \subset G_i \ (1 \leq i \leq n)$
  2) $X = \overline{\sum_{i=1}^n Y_i}$.

Examples of weakly decomposable operators are weakly contractions ("almost unitary contractions") (see [3], [8], Cap. VIII) and the adjoins of decomposable operators (see [6], [10]).

Remark I.1. From Proposition I.3.8. and Theorem II.1.5. ([2], [10]), it results that if $T \in \mathbf{B}(X)$ is weakly decomposable, then $X_T(F)$ is spectral maximal space of T, for any $F \subset \mathbb{C}$ closed and $\sigma(T|X_T(F)) \subset F \cap \sigma(T)$; moreover, we have that any spectral maximal space Y of T can be defined by the relation $Y = X_T(\sigma(T|Y))$.

II. Weakly decomposable operators and analytic functions

In this section we present several results of the theory of weakly decomposable operators which are related to functional calculus with analytic functions on spectrum.

Theorem II.1. *Let $T \in \boldsymbol{B}(X)$ be weakly decomposable and let $f\colon G \to \mathbb{C}$ ($G \supset \sigma(T)$ open and connected) be an analytic function. Then $f(T)$ is also weakly decomposable.*

Proof: Obviously, if f is constant, then $f(T)$ is weakly decomposable. Let us suppose that f is not constant and let $F \subset \sigma(T)$ be a closed set.

It is easily shown that $X_{f(T)}(f(F))$ is spectral maximal space of $f(T)$ (see [11]) and

$$\sigma(f(T)|X_{f(T)}(f(F))) \subset \sigma(f(T)) \cap f(F). \tag{II.1}$$

Let $\{G_i\}_{i=1}^n$ be a finite open covering of $\sigma(f(T)) = f(\sigma(T))$. Then

$$\bigcup_{i=1}^n f^{-1}(G_i) = f^{-1}\left(\bigcup_{i=1}^n G_i\right) \supset f^{-1}(f(\sigma(T))) \supset \sigma(T) \tag{II.2}$$

and hence $\{f^{-1}(G_i)\}_{i=1}^n$ is a covering of $\sigma(T)$. T being weakly decomposable, there is a system $\{Y_i\}_{i=1}^n$ of spectral maximal spaces of T defined by the relation $Y_i = X_T(\sigma(T|Y_i))$ such that

$$\sigma(T|Y_i) \subset f^{-1}(G_i) \ \ (1 \leq i \leq n) \tag{II.3}$$

and

$$X = \overline{\sum_{i=1}^n Y_i}. \tag{II.4}$$

If we denote

$$X_i = X_{f(T)}(\sigma(f(T)|Y_i)),$$

ISBN: 978-1-61804-327-6

then X_i is spectral maximal space of $f(T)$ (because $X_i = X_T(f^{-1}(\sigma(f(T)|Y_i)))$ is closed) and we obtain

$$X_i = X_{f(T)}(f(\sigma(T|Y_i))) \supset X_T(\sigma(T|Y_i)) = Y_i. \qquad \text{(II.5)}$$

From the relations (II.2)-(II.5), it follows that

$$X = \overline{\sum_{i=1}^{n} X_i}$$

and

$$\sigma(f(T)|X_i) = \sigma(f(T)|X_{f(T)}(\sigma(f(T)|Y_i))) \subset f(f^{-1}(G_i)) \subset G_i$$

hence the operator $f(T)$ is weakly decomposable. ■

Corollary II.1. *Let $T \in \boldsymbol{B}(X)$ be weakly decomposable and let $f\colon G \to \mathbb{C}$ ($G \supset \sigma(T)$ arbitrary and open) be an analytic function. Then $f(T)$ is also weakly decomposable.*

Proof: $\sigma(T)$ being compact, there is only a finite number of connected components of G which intersecting $\sigma(T)$. Let $G_1, G_2, \ldots, G_n$ be these connected components and let us denote $f_i = f|G_i$ $(i = 1, 2, ..., n)$. The sets $\sigma_i = G_i \cap \sigma(T)$ are separated parts of $\sigma(T)$. From $\sigma(T) = \bigcup_{i=1}^{n} \sigma_i$, it follows that

$$X = \bigoplus_{i=1}^{n} E(\sigma_i, T)X$$

and

$$T = \bigoplus_{i=1}^{n} (T|E(\sigma_i, T)X).$$

According to previous theorem and from Proposition **??** and Corollary **??** , it results that the operators $T|E(\sigma_i, T)X$, $f(T)|E(\sigma_i, T)X = f_i(T|E(\sigma_i, T)X)$ and $f(T) = \bigoplus_{i=1}^{n} (f(T)|E(\sigma_i, T)X)$ are weakly decomposable. ■

Proposition II.1. *Let $T \in \boldsymbol{B}(X)$ and let $f\colon G \to \mathbb{C}$ ($G \supset \sigma(T)$ open) be an analytic function such that $f(\lambda_1) \neq f(\lambda_2)$, if $\lambda_1 \neq \lambda_2$, $\lambda_1, \lambda_2 \in \sigma(T)$ (i.e. f injective). If $f(T)$ is weakly decomposable, then T is also weakly decomposable.*

Proof: Let F be a closed set of $\sigma(T)$. Then $f(F) = F'$ is a closed subset of $f(\sigma(T)) = \sigma(f(T))$ and $f^{-1}(F') \cap \sigma(T) = F$.

From Theorem 1.1.6. ([2]), we deduce that

$$X_{f(T)}(F') = X_T(f^{-1}(F')) = X_T(f^{-1}(F') \cap \sigma(T)) = X_T(F)$$

hence $X_T(F)$ is spectral maximal space of T.

Let now $\{G_i\}_{i=1}^n$ be a finite open covering of $\sigma(T)$. We may assume that $G_i \subset G$. The function f being analytic, then it is an open map, thus $H_i = f(G_i)$ is open and $\{H_i\}_{i=1}^n$ is an open covering of $\sigma(f(T))$. $f(T)$ being weakly decomposable, there is a system $\{Y_i\}_{i=1}^n$ of spectral maximal spaces of $f(T)$ such that $\sigma(f(T)|Y_i) = F_i' \subset H_i$ and $X = \overline{\sum_{i=1}^{n} Y_i}$.

But

$$Y_i = X_{f(T)}(F_i') = X_T(f^{-1}(F_i')) = X_T(F_i)$$

(where $F_i = f^{-1}(F_i') \cap \sigma(T)$), hence Y_i is also spectral maximal space of T. From $\sigma(T|Y_i) \subset F_i \subset G_i$, it results that T is weakly decomposable. ■

III. Quasi-nilpotent equivalence of weakly decomposable operators

Definition III.1. We say that the operator T_1, $T_2 \in \mathbf{B}(X)$ (not necessarily commutable) are *quasi-nilpotent equivalent* (or *spectral equivalent*) and write $T_1 \sim T_2$, if

$$\lim_{n\to\infty} \left\|(T_1 - T_2)^{[n]}\right\|^{\frac{1}{n}} = 0 \text{ and } \lim_{n\to\infty} \left\|(T_2 - T_1)^{[n]}\right\|^{\frac{1}{n}} = 0,$$

where

$$(T_1 - T_2)^{[n]} = \sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} T_1^k T_2^{n-k}.$$

Definition III.2. We consider the following notations

$$\rho_{sp}(T_1, T_2) = \limsup_{n\to\infty} \|(T_1 - T_2)^{[n]}\|^{\frac{1}{n}} \text{ and}$$

$$p(T_1, T_2) = \max\{\rho_{sp}(T_1, T_2), \rho_{sp}(T_2, T_1)\}.$$

p defines on $\mathbf{B}(X)$ a topology called *the disturbing topology* (or *p-topology*) ([25], [26]). Moreover, we denote

$$D(F, \rho) = \{\mu; \inf|\lambda - \mu| \leq \rho\} \text{ and } D(\{\lambda\}, \rho) = D(\lambda, \rho).$$

where $F \subset \mathbb{C}$ closed and $\rho > 0$.
In particular, $D(\{\lambda\}, \rho) = D(\lambda, \rho)$ is the disc with center λ and radius ρ.

Theorem III.1. *Let T_1, $T_2 \in \boldsymbol{B}(X)$. If T_1 is weakly decomposable and $T_1 \sim T_2$, then T_2 weakly decomposable and*

$$X_{T_1}(F) = X_{T_2}(F)$$

for any $F \subset \mathbb{C}$ closed.

Proof: Because T_1 and T_2 are weakly decomposable, then $X_{T_1}(F)$ and $X_{T_2}(F)$ are spectral maximal spaces of T_1, respectively T_2, for any $F \subset \mathbb{C}$ closed and any spectral maximal space of T_1 (respectively T_2) is defined by the identity $Y_1 = X_{T_1}(\sigma(T_1|Y_1))$ (respectively $Y_2 = X_{T_2}(\sigma(T_2|Y_2))$). It is easily seen that the spectral maximal spaces of T_1 are also spectral maximal spaces of T_2 (this assertion follows from the equalities $Y = X_{T_1}(\sigma(T_1|Y)) = X_{T_2}(\sigma(T_1|Y))$, because $\sigma_{T_1}(x) = \sigma_{T_2}(x)$ if $T_1 \sim T_2$).

Let $\{G_i\}_{i=1}^n$ be a finite open covering of $\sigma(T_2)$, which is also a covering of $\sigma(T_1) = \sigma(T_2)$ (I.3.8., [2]). T_1 being weakly decomposable, there is a system $\{Y_i\}_{i=1}^n$ of spectral maximal spaces of T_1 such that $\sigma(T_1|Y_i) \subset G_i$ and $X = \overline{\sum_{i=1}^{n} Y_i}$. But Y_i $(1 \leq i \leq n)$ are also spectral maximal spaces of T_2 and by relation

$$\sigma(T_2|Y_i) = \sigma(T_2|X_{T_2}(\sigma(T_1|Y_i))) \subset \sigma(T_1|Y_i) \cap \sigma(T_2) \subset G_i,$$

it results that T_2 is weakly decomposable.
The equality $X_{T_1}(F) = X_{T_2}(F)$, $F \subset \mathbb{C}$, is obvious. ■

ISBN: 978-1-61804-327-6

Proposition III.1. *The set of all weakly decomposable operators is closed in p-topology.*

Proof: The set of all operators of $\mathbf{B}(X)$ having the single-valued extension property is closed in p-topology of $\mathbf{B}(X)$. ([14])

Let $\{T\}_{n=1}^{\infty} \subset \mathbf{B}(X)$ be a sequence of weakly decomposable operators and $T \in \mathbf{B}(X)$ such that $\lim_{n\to\infty} p(T_n, T) = 0$. Then T has the single-valued extension property. From Corollary 2.5. ([10]), it results that

$$X_T(F) \subset X_{T_n}(D(F, p(T_n, T))) \subset X_T(D(F, 2p(T_n, T)))$$

and since $X_{T_n}(D(F, p(T_n, T)))$ is closed, we have

$$\overline{X_T(F)} \subset X_T(D(F, 2p(T_n, T))).$$

Let $x \in \overline{X_T(F)}$; then $\sigma_T(x) \subset D(F, 2p(T_n, T))$, for every $n \geq 1$, hence $\sigma_T(x) \subset F$, i.e. $X_T(F)$ is closed.

Let $\{G_i\}_{i=1}^{n}$ be an open covering of $\sigma(T)$. There are the compact sets σ_i, $\sigma_i \subset G_i$ $(1 \leq i \leq n)$ such that $\bigcup_{i=1}^{n} Int\sigma_i \supset \sigma(T)$. According to Proposition 2.2. ([10]), we can take n_0 so that $\bigcup_{i=1}^{n} Int\sigma_i \supset \sigma(T_n)$, for $n \geq n_0$. We take $n \geq 0$, with $G_i \supset D(\sigma_i, p(T_n, T))$ and according to Corollary 2.5. ([10]) we have

$$X_{T_n}(\sigma_i) \subset X_T(D(\sigma_i, p(T_n, T))).$$

But T_n is weakly decomposable and $\underline{\{Int\sigma_i\}_{i=1}^{n}}$ being a covering of $\sigma(T_n)$, we deduce that $X = \sum_{i=1}^{n} X_{T_n}(\sigma_i)$. If we denote

$$X_i = X_T(D(\sigma_i, p(T_n, T)))$$

we obtain $X = \overline{\sum_{i=1}^{n} X_i}$, with X_i spectral maximal spaces of T and $\sigma(T|X_i) \subset D(\sigma_i, p(T_n, T)) \subset G_i$, hence T is weakly decomposable. ■

Corollary III.1. *Let $\{T\}_{n=1}^{\infty} \subset \boldsymbol{B}(X)$ be a sequence of weakly decomposable operators which commute with $T \in \boldsymbol{B}(X)$. If $\lim_{n\to\infty} \|T_n - T\|_{sp} = 0$, then T is weakly decomposable operator.*

Proof: Since $\rho_{sp}(T_n, T) = p(T_n, T) = \|T_n - T\|_{sp}$, by previous proposition, it results that T is weakly decomposable operator. ■

IV. Conclusions

This article generalizes properties of decomposable operators emphasizing on quasi-nilpotent equivalence of weakly decomposable operators. It represents a starting point for defining the commutator of two weakly decomposable operators and studying its properties.

References

[1] ALBRECHT, E.J., *Funktionalkalküle in mehreren Veränderlichen*, Dissertation, Mainz, 1972.

[2] ALBRECHT, E.J., *Funktionalkalküle in mehreren Veränderlichen für stetige lineare Operatoren in Banachräumen*, Manuscripta Math., **14**, 1-40, 1974.

[3] ALBRECHT, E.J., VASILESCU, F.H., *On spectral capacities*, Revue Roumaine Math. Pures et Appl., **18**, 701-705, 1974.

[4] APOSTOL, C., *Some properties of spectral maximal spaces and decomposable operators*, Revue Roumaine Math. Pures et Appl., **12**, 607-610, 1965.

[5] APOSTOL, C., *Spectral decompositions functional calculus*, Revue Roumaine Math. Pures et Appl., **13**, 1483-1530, 1968.

[6] APOSTOL, C., *Restrictions and quotients of decomposable operators in Banach Spaces*, Revue Roumaine Math. Pures et Appl., **13**, 147-150, 1968.

[7] APOSTOL, C., *Remarks on the perturbation and a topology of operators*, J. Funct. Anal., **2**, 395-408, 1968.

[8] BACALU, I., *On restrictions and quotients of decomposable operators*, Revue Roumaine Math. Pures et Appl., **18**, 809-813, 1973.

[9] BACALU, I., *S-decomposable operators in Banach spaces*, Revue Roumaine Math. Pures et Appl., **20**, 1101-1107, 1975.

[10] BACALU, I., *Some properties of decomposable operators*, Rev. Roum. Math. Pures et Appl., **21**, 177-194, 1976.

[11] BACALU, I., *S-Spectral Decompositions*, Ed. Politehnica Press, Bucharest, 2008.

[12] BISHOP, E., *Spectral theory for operators on a Banach space*, Trans. Amer. Math. Soc., **86**, 414-445, 1957.

[13] COLOJOARĂ, I., FOIAŞ, C., *Theory of generalized spectral operators*, Gordon Breach, Science Publ., New York-London-Paris, 1968.

[14] COLOJOARĂ, I., FOIAŞ, C., *Quasinilpotent equivalence of not necessarily commuting operators*, J. Mat. Mech., **15**, 521-540, 1965.

[15] COLOJOARĂ, I., FOIAŞ, C., *Commutators of decomposable operators*, Revue Roumaine, Math. Pures et Appl., **12**, 807-815, 1967.

[16] DOWSON, H.R., *Restriction of spectral operator*, Proc. London Math. Soc., **15**, 437-457, 1965.

[17] DUNFORD, N., SCHWARTZ, J.T., *Linear Operators*, Interscience Publishers, New-York, Part I (1958), Part II (1963), Part III (1971).

[18] FOIAŞ, C., *Spectral maximal spaces and decomposable operators in Banach Spaces*, Archiv der Math., **14**, 341-349, 1963.

[19] FOIAŞ, C., *On the spectral maximal spaces of a decomposable operator*, Revue Roumaine, Math. Pures et Appl., **15**, 10, 1970.

[20] FRUNZĂ, ŞT., *An axiomatic theory of spectral decompositions for systems of operators* (*Romanian*), I, St. cerc. mat., **27**, 656-711, 1975.

[21] FRUNZĂ, ŞT., *An axiomatic theory of spectral decompositions for systems of operators* (*Romanian*), II, St. cerc. mat., **29**, 329-376, 1977.

[22] HILLE E., PHILLPS R., *Functional analysis and semi-groups*, Amer. Math. Soc. Colloq. Publ., Providence, **31**, 1975.

[23] RADJABALIPOUR, M., *Equivalence of decomposable and 2-decomposable operators*, Pacific J. Math., **77**, 243-247, 1978.

[24] VASILESCU, F.H., *Residually decomposable operators in Banach spaces*, Tôhoku Mat. Jurn., **21**, 509-522, 1969.

[25] VASILESCU F.H., *Residually properties for closed operators on Fréchet spaces*, Illinois J. Math., 377-386, 1971.

[26] VASILESCU, F.H., *Analytic Functional Calculus and Spectral Decompositions*, D. Reidel Publishing Company, Dordrecht; Editura Academiei, Bucharest, 1982.

[27] VASILESCU, F.H., *On the residual decomposability in dual spaces*, Revue Roumaine, Math. Pures et Appl., **16**, 1573-1578, 1971.

Cross-disciplinary methodology for development of entrepreneurial skills: the case of Riga Technical University

Anita Straujuma
Institute of Business Engineering and Management,
Riga Technical University
Riga, Latvia

Elīna Gaile-Sarkane
Institute of Business Engineering and Management,
Riga Technical University
Riga, Latvia

***Abstract* —** The impact of introducing and testing of cross-disciplinary methodology for development of entrepreneurial skills within study process becomes more and more important. This paper describes cross-disciplinary methodology and main results of the testing phase.

***Keywords* —** entrepreneurship education, cross-disciplinary methodology, skill development

I. INTRODUCTION

Many scientists have identified the importance of cross-disciplinary approach in education. Some of scientists stress that we even do live in a cross-disciplinary age [1]. A lot of researches are devoted to cross-disciplinary collaboration, culture, learning, experiments, programs etc., but just a few of them towards cross-disciplinary methodology of the entrepreneurial education.

When higher education institutions deal with entrepreneurship education often in reality it becomes cross-disciplinary. There are a lot of different definitions of entrepreneurship, but most of them are about development of a business from the ground up, developing of an idea and turning it into business, about getting profit from this idea.

Entrepreneurship education usually is oriented towards ensuring entrepreneurial success by providing students with motivation, knowledge and skills needed for particular field of business, but today's innovative and technologically oriented society requires new approaches to cross-disciplinary education.

II. ON CROSS-DISCIPLINARY METHODOLOGY IN EDUCATION AND HOLISTIC APPROACH

Cross-disciplinary approach in education is described by many authors like Koke, Lace, Oganisjana, [2] who in their researches stress that entrepreneurship is a dynamic system of the components determined in mixed method research. In this article elaborated holistic functional-structural model of entrepreneurship illustrates that entrepreneurship can be developed holistically as a system when students identify, generate and realize opportunities into new personal, social or economic values not only in the context of specialized entrepreneurial disciplines, but within a wide range of disciplines in the unity of theory and practice [2].

Previous research by Caune, Lapiņa, Gaile-Sarkane, Ozoliņš, Borkus [3] on managers competences and capabilities confirmed specific competencies that are the most important for successful management of a company. These competences were grouped in four groups: professional, social and communication, personal and responsibility, leadership and innovative and learning competencies, with detailed description of each [4]. Very good analysis of competences that are essential for new entrepreneurs is described by E. Frank, in "The UK's Management Charter Initiative: the first three years" [6].

In order to achieve better results in development of methodology, there are many other factors that are important, for example – strategic orientation, strategic background, voice of customer, possibility to apply integrated management systems etc. [4,5,6].

Cross-disciplinary methodology was based on:

- Open innovation concept developed by Chesbrouh in 2003;
- the value model;
- holistic functional-structural model;
- skills and competences that should be developed during the study process.

For testing purposes of developed methodology, step by step approach to testing process form successfully completed Leonardo Da Vinci Innovation Transfer project Nr. LLP-LdV-ToI-12-CY-1671210 „YOUNG LEADERS: Developing and enhancing leadership skills for young managers in times of crisis: an innovative training package for European young professionals, 2012.-2014. were taken. Project results assured that methodology can be tested within study process.

A. Open innovation concept

Open innovation concept has been discussed since beginning of the 21st century and becomes very attractive in many areas. Firstly it was announced by Henry W. Chesbrough in 2003 and up to now many scientists have discussed open

ISBN: 978-1-61804-327-6

innovation paradigm and concepts in their researches. Eelko K.R.E. Huizingh in his article „Open innovation: State of the art and future perspectives" has mentioned „Many open innovation studies focus on specific industries, such as consumer electronics (Christensen et al., 2005), food (Sarkar and Costa, 2008), financial services (Fasnacht, 2009), automotive (Ili et al., 2010), and biotechnology (Fetterhoff and Voelkel, 2006; Bianchi et al., this issue). Other studies confirmed that there are minor differences in adoption rate between industries (e.g., Chesbrough and Crowther, 2006; Keupp and Gassmann, 2009; Lichtenthaler, 2008; Lichtenthaler and Ernst, 2009; Van de Vrande et al., 2009), whereas Gassman (2006) suggests the nuclear and military industries as typical examples of closed innovation industries" [9]. There are a lot of different researches published on this topic in manufacturing, service industries, education and other areas.

Basically open innovation is defined as the use of knowledge that exists inside and outside of organizations and can increase or create significant value for innovation, usually by exploiting ideas, intellectual property and technologies that drive a company towards better business model than competitors have [8].

In his book "Open innovation: the new imperative for creating and profiting from technology" Henry W. Chesbrough claims: "If you were trying to develop mechanisms to access useful knowledge today, you would start by surveying the surrounding knowledge landscape. The factors that promote knowledge diffusion create new opportunities. Knowledge diffusion rewards focused execution: You need not invent the most new knowledge or the best new knowledge to win. Instead, you win by making the best use of internal and external knowledge in a timely way, creatively combining that knowledge in new and different ways to create new products or services." [8].

Trott and Hartmann claim: „If open innovation is in essence nothing new, why then has this concept been so readily embraced by firms and the R&D community? Much of this surely is due to its simplicity (it is appealing because it is simple and retains the linear notion of science to marketplace) and the partial deception which was created by describing something which is undoubtedly true in itself (the limitations of closed innovation principles), but false in conveying the wrong impression that firms today still follow these principles." [10]

B. The value model

The value model – value for customers' creation approach described by Lindstedt, Burenius [7] is a holistic model on theory of how to create customer value in products (including systems and services). The model explains how customer value works and how customer value is made. The model describes entrepreneurial approach of developing an idea and turning it into value that can be turned into a profitable business. The model deals with agile product development approach in a multi project environment [7].

Value is a fundamental term in management and is used in several meanings. One of them, immaterial values, Schwartz defines as "desirable transsituational goals, varying in importance, that serve as guiding principles in the life of a person or other social entity… Values are the criteria people use to evaluate actions, people, and events" [11], [12]. Stakeholders' values play important role in different stages of product development. Each group of stakeholders have dominant values that influence their decisions, interests and actions. Also perception of products' material value. At each stage of the process, it is essential for all stakeholders to be aware of the value divergence and have an internal agreement on it to ensure most effective results.

C. Holistic functional-structural model

Holistic functional-structural model describes entrepreneurship as a holistic system where participants identify, generate and realize opportunities into new personal, social or economic values. Model emphasizes the unity of theory and practice. The successful results are ensured by personal interest and involvement in the process. The model suggests active involvement of students in the process of product development in all phases of the course where professor performs as a coach [2].

D. Methodology for cross-disciplinary education

Inspired by thoughts of Chesbrough, simplicity of open innovation concept, holistic functional-structural model of entrepreneurship, and value creation model that is essential for delivering value to customers, group or Riga Technical university researchers have developed methodology for cross-disciplinary education.

Methodology consists of 8 different steps and matrix-type structure for competence development. Main steps are following:

1) Idea/ problem development
2) Identification of customer needs;
3) Identification of potential functions of the product (system or service);
4) Development of concepts;
5) Analysing of concepts;
6) Prototyping and testing;
7) Development of sales plan;
8) Economic justification of the project.

Ideas form Lindstedt and Burenius value model have been adapted for graphical description of cross-disciplinary methodology for development of entrepreneurial and innovation skills.

ISBN: 978-1-61804-327-6

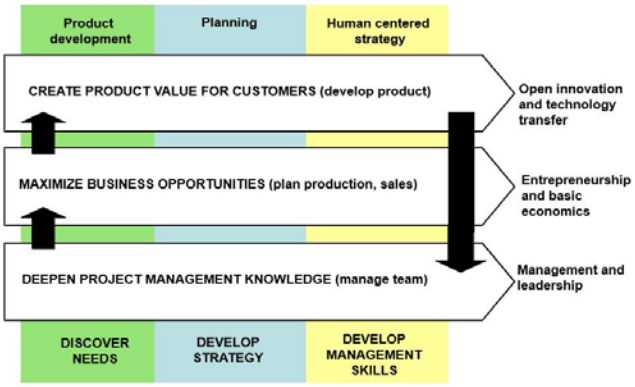


Figure 1 Methodological Model of Cross-Disciplinary Methodology in Education and Holistic Approach in study process.

Methodology was tested in academic year 2014./2015. in Riga Technical University, faculty of Engineering Economics. For testing purposes two groups of 30 students in each were involved. Both groups were part time postgraduate students, typically employed. Testing phase was done within one study semester.

III. TESTING OF THE METHODOLOGY

Survey was performed on a representative sample – 57 students who studied Entrepreneurial and Innovation Skill Development course (thereinafter – ESID). Test results have good internal consistency – Cronbach's alpha 0,83.

Students were tested with 360 degree test before starting ESID course. Students had to rate their own skills, scale 1 to 5 (self-assessment). Scores above 4 are considered strengths, 3-3.5 are considered moderate and can be further improved. Scores below 2.5 are considered weaknesses and need improvement. Test uses [4] classification of management skills:

- Self-management skills (self-confidence, optimism, trustworthiness);
- Business management skills (making analytical decision, strategic thinking and planning, creative problem solving);
- People management skills (collaboration and team work, interpersonal communication, building networks and connections, motivating others).

Same test was given to student's friend, manager and colleague to rate the skills of particular student – i.e. each person has three external evaluations and as external test result is used mean of all three results.

Test results show that for all students self-evaluation is always lower than external evaluation except male students' self-evaluation of interpersonal communication. There are certain tendencies depending on the gender of the student. Men rate themselves quite close to external evaluation, women are more critical of themselves. Men rate themselves higher than women in all questions. Women rate themselves lower than 4 in all questions. Also employers and friends evaluate men higher than women in all positions except building networks and connections. Though external rating of women is close to the one of men except strategic thinking and planning where women are rated significantly lower.

Figure 2 Variety of self-management skills rating before ESID course

Both, men and women, give highest rating to self-confidence. And all students rate highest following skills: self-confidence, interpersonal communication, creative problem solving, motivating others. External evaluation gives highest rating of following skills: self-confidence, motivating others, trustworthiness, collaboration and teamwork. Self-evaluation does not match with external evaluation except for self-confidence.

Figure 3 Business management skills before ESID course

Both genders rate low following skills (these skills will be trained in ESID course): building networks and connections (team management), strategic thinking and planning (enterprise

economics), making analytical decisions (enterprise economics.), optimism. Colleagues and friends rate low following skills: building networks and connections, strategic thinking and planning, making analytical decisions, interpersonal communication. Self-evaluation coincides with external evaluation.

Largest gap between self-evaluation and external evaluation (external evaluation is higher) is for following skills: trustworthiness, strategic thinking and planning, building networks and connections.

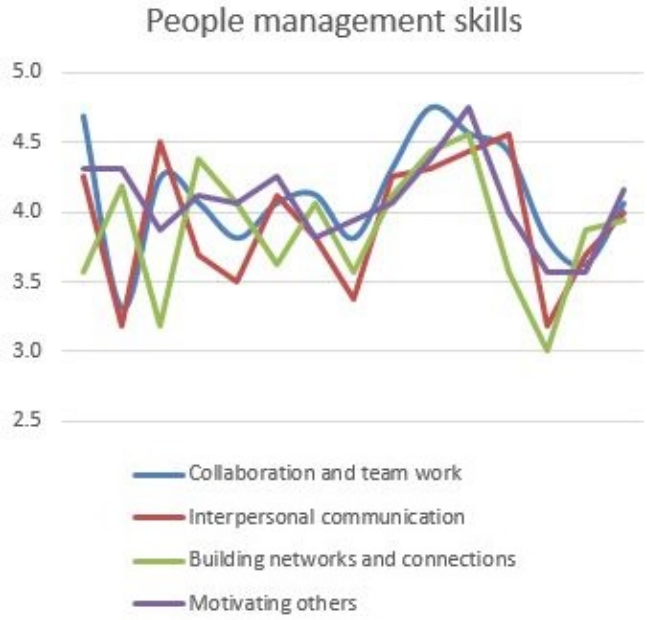


Figure 4 Evaluation of people management skills before ESD course.

After finishing ESID course students were tested again by asking them to evaluate themselves by answering same questions as before starting the course. From 57 students after finishing this testing module we received 52 filled questionnaires.

External evaluation of students was not performed since the course lasts for one semester. Students' skills and behavior may change at the end of semester, close to the moment of testing. Even if students' friends, managers and colleagues would have noticed changes in students' skills and behavior, they might not recognize it or might not perceive it as stable results. Thus, as a future work, 360 degree test will be performed one year after students' finished the course and the results will contain also external evaluation and describe how permanent are the new skills.

Self-evaluation showed improvement of results in most of the areas. Out of all tested self-management skills most improved were self-confidence and optimism. Trustworthiness had not changed.

Business management skills need long-term practice and experience. Test results show that skills "making analytical decision" and "strategic thinking and planning" have minor improvement. On contrary self-evaluation of "creative problem solving" improves significantly since methodology gives both – theoretical basis and practice of different creativity techniques throughout the course.

Most improving people management skills are "collaboration and teamwork" and "interpersonal communication" since both are practiced constantly within the group of course participants. Skill "building networks and connections" improve although less than the other two. This skill improves because students apply it within the group but for significant improvement they need long-term practice outside the group boundaries. Skill "motivating others" improves for few students since it depends on the person's role in the group or the company. Team leaders train the skill while others stay on the same level.

Feedback from participants in testing methodology was positive. Most of the participants agreed that participation in ESID course has broadened their knowledge on the topics covered, improved their skills, increased their competitiveness in the business environment, helped them to adopt new and modern ways of communication, increased their self-confidence and stimulated to recognize their strength and weaknesses.

IV. CONCLUSIONS

The Cross-Disciplinary Methodology for Entrepreneurial and Innovation Skill Development, based on open innovation concept developed, the value model and holistic functional-structural model has good success in first stages of its' testing. Results approved that main skills and competences can be developed by applying this methodology in education process of higher education institutions. At the same time there are a lot of fundamental points for improving, testing and verifying in the future.

First test results confirmed that methodology should be modified for graduate students since they have no work experience and thus need to improve basic skills. Differences in skill development also can be tested in multicultural environment.

Holistic approach requires change of didactic approach starting from teaching theory issues and content of teaching up to organizational forms and methods, that should be changed towards coaching.

ACKNOWLEDGMENT

The paper was supported by the project «The Development of Innovation and Entrepreneurship in Latvia in Compliance with the Smart Specialization Strategy» within the National Research Program 5.2. «Economic Transformation, Smart Growth, Governance and Legal Framework for the State and Society for Sustainable Development - a New Approach to the Creation of a Sustainable Learning Community (EKOSOC-LV)».

REFERENCES

[1] Fontaine P. Introduction: The Social Sciences in a Cross-Disciplinary Age. Journal of the History of the Behavioral Sciences, Volume 51, Issue 1, 1 December 2015, Pages 1-9

[2] Koke T., Lace N., Oganisjana K. Understanding entrepreneurship development in Latvia: A cross-disciplinary approach. International Journal of Etrepreneurship and Small Business, Volume 13, Issue 3, July 2011, Pages 293-312

[3] Caune, J., Lapiņa, I., Gaile-Sarkane, E., Ozoliņš, M., Borkus, I. Managers Capabilities and Competences in Changing Environment: Empirical Study in Latvia. Proceedings of the 18th World Multi-Conference on Systemics, Cybernetics and Informatics (WMSCI 2014). Vol.1, United States of America, Orlando, 15-18 July, 2014. Florida: International Institute of Informatics and Systemics, 2014, pp.7-12.

[4] Mežinska, I., Lapiņa, I., Mazais, J. Integrated Management Systems towards Sustainable and Socially Responsible Organisation. Total Quality Management & Business Excellence, 2015, Vol.26, Issue 5-6, pp.469-481

[5] Tamošiuniene, R., Jasilioniene, R. Customer relationship management as business strategy appliance: Theoretical and practical dimensions. Journal of Business Economics and Management, Volume 8, Issue 1, 2007, pp. 69-78

[6] Frank E., The UK's Management Charter Initiative: the first three years, Journal of European Industrial Training, Vol. 17, No. 1, 1991, pp. 9–11.

[7] Lindstedt, P., Burenius, J. "The Value Model: How to Master Product Development and Create Unrivalled Customer Value" – textbook, 2003, ISBN 91-630-6349-2

[8] Chesbrough, Henry W. „Open innovation : the new imperative for creating and profiting from technology" – 2003, Harvard Business School Publishing Corporation, 216 pp. (p.50).

[9] Huizingh, Eelko K.R.E. „Open innovation: State of the art and future perspectives" - Technovation 31 (2011) 2–9, www.elsevier.com/locate/technovation,

[10] Trott P., Hartmann D."Why 'Open Innovation' is Old Wine in New Bottles" - Imperial College Press, International Journal of Innovation Management. Vol. 13, No. 4 (Dec. 2009) pp. 715–736

[11] Schwartz, S. H. (1994). Are There Universal Aspects in the Structure and Contents of Human Values? Journal of Social Issues, 50(4), 19–45. doi:10.1111/j.1540-4560.1994.tb01196.x

[12] Schwartz, S. H. (2006). Basic Human Values : An Overview Basic Human Values : Theory , Methods , and Applications Introduction to the Values Theory. Jerusalem Hebrew University, 48, 49–71. doi:10.1080/026999499377664

Some kinds of use of i-textbooks

Eva Zmazek
Institut Logik, Ptuj and
Faculty of Mathematics and Physics,
University of Ljubljana
Slovenia
eva.zmazek@gmail.com

Jan Zmazek
Institut Logik, Ptuj and
Faculty of Mathematics and Physics
University of Ljubljana
Slovenia
jan.zmazek@gmail.com

Blaz Zmazek
Institut Logik, Ptuj and
Faculty of Natural Sciences and Mathematics
University of Maribor
Slovenia
blaz.zmazek@gmail.com

***Abstract*—**ICT is a part of modern life. It provides us with an instant access to information, materials and services and provides a new opportunity for equality in education. The structure and didactic guidelines of i-textbooks are based on the knowledge of many subject-specific didactics and didactics of technology use in teaching. This paper highlights the role of higher level of interactivity in i-textbooks. We believe that i-textbooks will contribute to the deepening of knowledge and we hope we set an example of good practice in ICT driven didactics, which is still developing in Slovenia.

***Keywords*—** i-textbook, interactivity, ICT

I. INTRODUCTION

The use of information and communication technology (ICT) for teaching in Slovenia has been a constant for twenty years. A teacher was predominantly a direct user of this technology in the initial period, but over the last few years, students have also turned into direct ICT users. Students actively apply all available options of use:

a) simplest, such as electronic communications between teachers and students (e.g., e-mail, social networking sites, online classrooms),

b) simple, the use of software tools (e.g., Geogebra, Mathematica, Wolfram Alpha in mathematics, etc.),

c) use of didactic-substantive quality collections of e-materials (e.g., http://www.e-um.si) and i-textbooks (http://iucbeniki.si/) through various electronic media.

The subject of e-materials in Slovenia has been evolving through several projects since the introduction of the Internet, but a real revolution, from the viewpoint of use in Slovenia, was achieved with the project "E-um", which has followed (later set) guidelines for i-textbooks development since its concept design. In the last period, the growing trend of ICT use in education has seen more and more e-textbooks in digitised form of conventional (printed) textbooks, which exploit the new media only as compensation. This is why we introduced the concept of i-textbook for interactive e-textbooks and started a so-called ivolution of textbooks into i-textbooks (e-textbooks, which exploit new media to upgrade interaction with users, enabled by the new technology).

Today, the use of ICT in teaching and learning has become indispensable. Not only does it help us optimize time and spice up learning and teaching, it allows us to reach a common objective of teachers and students – to achieve an in-depth understanding and useful knowledge in many subjects.

II. I-TEXTBOOK AS A RESOURCE FOR LEARNING

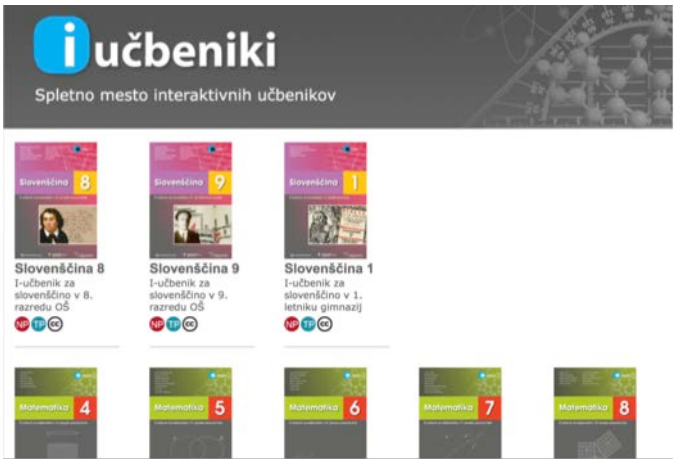
I-textbook brings new experience in the Slovenian school environment for a teacher and a student. We can leave aside some of the electronic learning materials, which are a welcome support in the teaching environment, but do not have a wide technical background as collections of worksheets, texts and examples and they cannot be equated with the quality of an i-textbook. We define i-textbook as an e-textbook, which is dominated by (reasonably included) i-learning building blocks with high degree of interactivity. The prevalence means that the core concepts are mainly introduced by highly interactive learning building blocks. As an example of such an interactive widget, interactive conceptual applets are proposed. The applet is relatively small software application that is built around pre-designed graphic representation. Conceptual applets are applets that allow the detection of relationships of key concept or relations between several concepts of the topic. The ability of visualization of mathematical concepts and processes, and interactive nature of the interactive applet can be a powerful conceptual teaching tool.

In preparation stage of e-textbook and confirmation of its suitability, following basic axioms define content-didactical

ISBN: 978-1-61804-327-6

requirements: axiom of curriculum coverage, axiom of interactivity, axiom of multimedia, axiom of inductive approach, axiom of developing profound knowledge, axiom of developmental suitability (for more details see [1]). I-textbooks encourage students to be more active and discourage from learning without understanding the topic. In the i-textbook, the content is presented with audio-visual and interactive elements, which enable impressive acquisition of knowledge. It is recommended that each unit contain summary that allows the user to repeat the acquired knowledge with its integrated content. The content is split into content sections and is accessible in tree structure of menu and. Basic elements present closed learning units, combined in chapters and content sections. Each web page (screen image) includes closed activity, acquisition or repetition of knowledge without the need to navigate between other web pages and may contain content in whole or just a part of it. The content is freely accessible without required registration the content can be simply modified and adjusted.

Slovenian i-textbooks that can be found on http://iucbeniki.si/ and http://itextbooks.eu/ offer a lot of options for use inside and outside of a classroom is easily accessible and can function with or without a web access. The purpose of application dictates the manner of i-textbook use. An i-textbook for a specific class is designed linearly, but the curriculum in a spiral: each next class builds new contents based on known topics from previous classes. Students revise and consolidate known facts, concepts, properties and procedures and then upgrade an existing knowledge network with new experiences and new knowledge. The basic elements of an i-textbook are student-centered activities. Students implement these activities directly with an i-textbook (applets, demonstrations, spreadsheets, images) or in their own environment (working with material, pictures and sketches in a notebook, observing phenomena, monitoring of events, browsing the Internet, use of ICT, etc.).



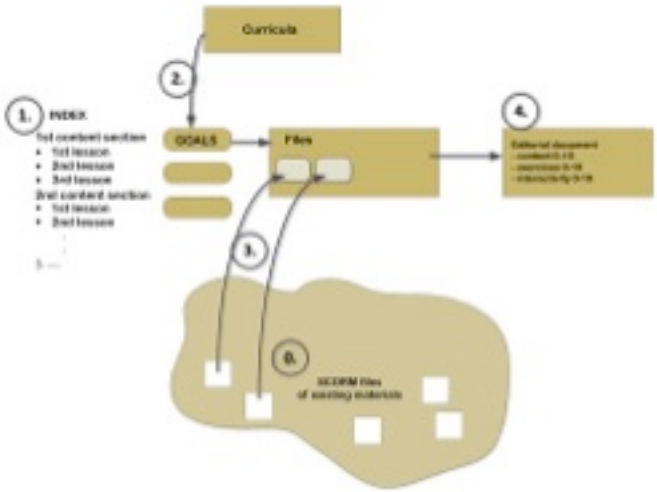
III. SOME KINDS OF USE OF I-TEXBOOKS

Interactive elements from i-textbooks are very suitable for independent learning and research. Interactive elements with high degree of interactivity allow students to repeatedly do the activity but on the different sets of input data. With that student could find and learn the new relations between the known objects. New i-textbooks contain a lot of interactive elements, much more than existing collections of the old e-learning materials. These elements are a didactical challenge for the teacher and his main task is to properly focus and train students to correctly work with them. The main advantages and great didactical value of interactive elements could be lost if a teacher used them just for his own interpretations.

Lessons could be very interesting and very effective if the teacher motivates students to use content in the i-textbooks in an appropriate way, but using i-textbooks also hides some potential traps. Students' motivation may be lost because using computers and tablets provides access to other online non-educational content. This risk decreases when the teacher clearly defines the work instructions and expectations at the start of lesson. Another problem is related to the student's individual speed of learning. To overcome this gap, teachers must continuously use effective differentiation in the classroom, more precisely; teachers must build various learning pathways for different groups with the aim of the best acquisition of new content.

Different levels of interactivity, parameterized exercises and various content lmake i-textbooks revolutionary. Students now have much more options for learning at home, content plays a very active role in learning and it encourages the student's learning activities. I-textbook is not just a textbook - it is some kind of virtual tutor, who helps both weaker and better students. I-textbooks are of great help to students who miss a lot of lessons in schools, but they also could be a very helpful tool for students with special needs, mainly due to the visualization of the content.

I-textbook has an index structured content covering each learning set that contains learning units with meaningful content covering one and up to three teaching lessons. Development of an i-textbook included an incorporated targeted approach, with a pre-formatted index table, which included the objectives of the curriculum



E-learning units encourage active self-learning, which is optimally achieved by solving problems with reactive cooperation instead of passive reading. That is why they must offer a sufficient amount of interactive and dynamic elements provided by the modern technology (graphics, audio sequences, video projections, animations, slides simulations

ISBN: 978-1-61804-327-6

and applets). They also include drawings with a high level of educational usability that introduce the essence of a problem, which enables understanding or provides a solution to a problem. A hyperlink is not considered as an interactive element.

A. Independent learning of students at home

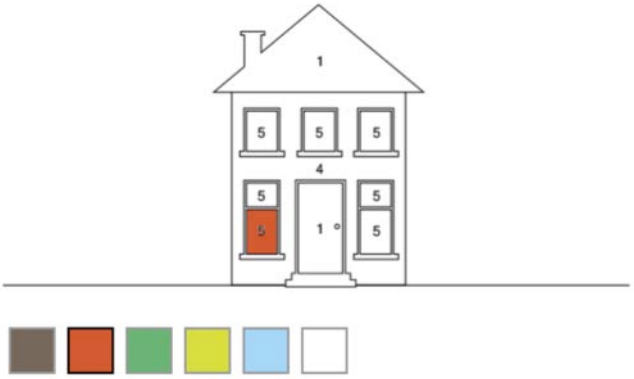
Classic home learning has, just a few years ago, been carried out mostly by the use of printed textbooks. Textbooks are otherwise substantially correct, but dull in form and especially difficult to understand for weaker students. Different interactivity levels, generated tasks and text placed under the buttons have made many improvements in new i-textbooks. High school students have much more opportunities when studying at home, to take on a more active learner role when dealing with content, because the interactive units encourage them to be more active. On the introduction page of every unit, which encompasses motivation or content contextualisation and repetition of needed prior knowledge, the student already has to engage in active learning.

B. Use of i-textbooks in class

For independent research in class, the most appropriate applets are those with high levels of interactivity, which are defined and described in [2]. With these applets student may learn independently by repeating activities and diverse data sets and discover new connections between already known concepts. New i-textbooks contain significantly more of these applets than existing collections of e-materials. These applets are an educational challenge for a teacher, because they must motivate their students to work with them. However, they can lose their value, if a teacher uses them only as enrichment to their own interpretations. A teacher can use an applet with frontal instruction, but it would be even better to encourage students to investigate the applet by themselves, especially in classes equipped with computers or tablets.

Lessons, where a teacher directs students to reading already developed e-materials and individual units of an i-textbook, are interesting for students however, this method has a few traps. Students' motivation can deteriorate rapidly and the use of computers or tablets offers many opportunities for other forms of entertainment (non-educational). This risk is reduced if a teacher clearly defines work instructions and expectations at completion of work. We must also be aware of great advantages, particularly in developing effective differentiation in classroom, as students may adjust the speed of learning and a teacher may offer various paths to adoption of new content. Differentiation is a problem due to different levels of students' prior knowledge. Teachers have been working to solve this problem for a long time, but it seems that it is extremely difficult to solve [3]. Technology can most likely offer great help in solving this problem.

Modern interactive and dynamic elements can play a constructive as well as a destructive role in the learning process.



They can be a useful tool in ensuring active participation of a student, with a better presentation of the facts and a deeper understanding of the material. Misused elements can cause student to lose focus of the essential objective. Dynamism and interactivity therefore should not be only a means by itself. They must be used to achieve constructive goals [4,5,6].

Even though a teacher will be able to use certain parts of an e-learning unit, it is in its entirety designed to encourage individual work. It should therefore be designed to lead a user throughout the whole unit. It must be unambiguous, motivational, encourage curiosity and creativity and should enable clear understanding of its contents. It should also encourage activities like reading, writing, designing, troubleshooting, research, cooperation, etc. and ensure feedback. E-learning units are not by any means theoretical didactic preparations for a teacher, but a stimulating and productive environment for learning. That is why hyperlinks must be avoided, except at the end of a unit, where we can cite and link literature for broadening and deepening of knowledge. This is one of the ways to ensure that students focus on the unit and do not lose themselves on the Internet, where one click quickly leads to another.

IV. Conclusion

We want to point out that modern interactive and dynamic elements in i-textbooks can play a constructive or a destructive role in the learning process. They can be a useful tool in ensuring active participation of a student, with a better presentation of the facts and a deeper understanding of the material. Misused elements can cause a student to loose focus of the essential objective. Dynamism and interactivity should therefore not only be a means by itself [5].

Advantages of i-textbooks in learning at home are numerous. With a more active learner's role and a new textbook design we developed a virtual instructor who helps weaker and also more curious students. I-textbooks also help students with special needs, especially because of the content visualisations, and students with a specific status, who are often absent form school. Because i-textbooks are freely available online they can accompany us everywhere with no additional luggage.

ISBN: 978-1-61804-327-6

REFERENCES

[1] Zmazek, B., Lipovec, A., Pesek, I., Zmazek, V., Šenveter, S., Regvat, J., Prnaver, K. (2012) What is an e-textbook?. Metodički obzori,. 7 (15), 127-139)

[2] Repolusk, S. Zmazek, B., Hvala, B. in Ivanuš Grmek, M. (2010). Interaktivnost e-učnih gradiv pri pouku matematike. Pedagoška obzorja, 25(¾), 110-129.

[3] Žakelj, A. (2012). The impact of level education (ability grouping) on pupils' learning results. V: Uzunboylu, H. in Demirok, M. (Ur.). The 3rd World conference on psychology, counseling and guidance, WCPCG 2012, May 9-12, 2012, Izmir, Turkey, str. 383-389. New York: Elsevier, 2012.

[4] Zmazek, B., Kobal, D., Zmazek, V. and Hvala, B. (2007). The challenge of E-learning. V: 11th World Multiconference on Systemics, Cybernetics and Informatics jointly with the 13th International Conference on Information Systems Analysis and Synthesis, July 8-11, 2007, Orlando, Florida, USA. proceedings. Vol. 1, p. 5.

[5] Zmazek, B., Kobal, D. and Zmazek, V. (2007). E-um learning in e-society. V: 2nd International Conference on e-Learning, New York, 28-29 June 2007. REMENYI, Dan (ur.). ICEL 2007. Reading: Academic Conferences, 2007, str. 521-544.

[6] Zmazek, B., Kobal, D. and Zmazek, V. (2007). E-um learning in e-society. V: 2nd International Conference on e-Learning, New York, 28-29 June 2007. REMENYI, Dan (ur.). ICEL 2007. Reading: Academic Conferences, 2007, str. 521-54.

Internet of Things as a Framework for E-recruitment's Business Model?

Deniss Sceulovs
Faculty of Engineering Economics and Management
Riga Technical University
Riga, Latvia
deniss.sceulovs@rtu.lv

Vladimir Shatrevich
Faculty of Engineering Economics and Management
Riga Technical University
Riga, Latvia
vladimirs.satrevics@rtu.lv

***Abstract*—** Modern e-businesses are developing rapidly as new modern enterprises, e-business management is an important topic across contemporary management and modern information technology. This paper investigates e-recruitment based on Business Model Ontology framework, to provide useful implication of e-recruitment as a business model.

***Keywords*—** e-recruitment, Internet of Things, e-business, Business Model Ontology, value creation, information and communication technology.

I. INTRODUCTION

Under conditions of globalisation the boundaries of labour market continue to expand generating new opportunities and challenges. Workforce became more geographically mobile searching for well paid job and better life circumstances. The increasing migration of workforce creates a surplus of available labour in developed countries, and the shortages of labour in less prosperous countries. Also, organisations look for developmental possibilities expanding their market. These organisations' activities include the expansion or relocation of business aboard, the utilisation of cost efficient forms of labour (e.g. flexible types of employment) etc. In majority of cases, organisations encounter the matter of personnel recruitment. As Deloitte's survey "Global Human Capital Trends 2015" demonstrates, the organisations' need for talent and contingent workers will continue to grow [1], thereby recruitment process must be extensive and high calibrated simultaneously. The Boston Consulting Group's survey "Creating people advantage 2014-2015" reveals that the future importance, urgency and invested effort of recruiting processes and strategy are higher in high performance organisations than in low performance organisations [2].

The rapid electronic environment development over the last decade has fostered the e-recruitment growth and has provided companies with opportunities that they previously did not have. By employing advantages offered by the e-recruitment, entrepreneurs can ensure expedient and effective communication with the target audience, by promoting their services on the global market. The performed scientific studies show that proper and skilful use of modern technologies can contribute to significant development of companies. Up to now, no unequivocal studies have been performed about the use of the electronic environment in ensuring development of micro, small, and medium enterprises. Ph. Kotler, D. Tapscott, P. Drucker, and J. A. Pearce [3] maintain that two parallel markets exist and are developing – the traditional and the electronic environment. The electronic environment is used for various needs – for trade, marketing, advertisement, studies, communication, training, etc. Simultaneously, there is an opinion claiming that in future, the majority of businesses will be performed on the electronic market, hence advancing the dominant position of the e-environment in achieving entrepreneurship competitiveness. In recent years, companies' intellectual capital (IC) has gained increased attention due to globalisation and integration of capital markets, greater mobility of monetary and actual goods, tougher competition, new dominating industries, and developments in information and communication technology (ICT).

Scientists [4; 5; 6; 7] have argued that demand for information (external communication) on knowledge-based resources is growing as companies increasingly base their competitive strength in the value of know-how, patents, skilled employees and other intangibles. The electronic environment already now offers companies practically all the necessary marketing and communication tools for ensuring company development by creating competitive advantages, nevertheless, not all companies can employ the opportunities rendered by the e-environment, in order to increase company competitiveness and productivity. These trends promotes e-recruitment as a new form of business that has changed conventional re-cruiting to a more efficient "continuous mode" [8] and has reduced hiring costs compared to traditional recruiting through newspapers and magazines [9]. Competitive advantages provided by e-recruitment methods and value creation process principles should be better explained in order to create effective business model. Recent findings stated that more than 20% of job seekers have rejected job opportunities simply based on poorly designed websites [10] and that company-designed websites are so complicated that about three-quarters of all job seekers are unable to use them successfully [11]. Conventional management studies of employee e-recruitment methods have failed to provide managers with a theory-based understanding of how e-recruitment contribute to recruiting success [12, 13] or explain "not only what happens, but why it happens" [14].

Considering challenges coming from using e-recruitment methods, this article develops an e-recruitment evaluation

ISBN: 978-1-61804-327-6

system framework based on Business Model Ontology (BMO). New model is created to identify crucial e-recruitment factors. This model is based on statement that conventional evaluation system is not suitable for the recruiting process and should be developed. From this perspective, the model presents e-recruitment as an e-business and evaluates in BMO context. The model is aimed to create sustainable e-business by identifying value creation process and significant factors.

II. E – Recruitment Business Requirements

In general, e-recruitment (or online recruitment, internet recruitment, web-based recruitment) is the process of human resource (HR) recruitment exploiting electronic resources. The majority of the definitions of e-recruitment are derived from view of human resource management theory and practice, so focused to e-recruitment as instrument or process by which organisation's needs for workforce is ensured. For example, Armstrong defines e-recruitment as the use of the internet to advertise or 'post' vacancies, provide information about jobs and the organization and enable e-mail communication to take place between employers and candidates; the latter can apply for jobs online and can e-mail application forms and their CVs to employers or agencies [15]. Some HR specialists interpret e-recruitment not only as the using internet for hiring, but also emphasize the application of HR software [16]. The academic works examining e-recruitment are increasing, but many studies analyse this subject from human resource management, psychological or information and communication technology perspectives. Searching relevant articles in database Scopus, applying keywords "e-recruitment", or "e-recruiting", "online recruitment", "internet recruitment", and "business" and "model" for document's title, abstract or keywords, only fifteen results were obtained. Approximately half of this search results could be referred to the theme of e-business.

The typical forms of e-recruitment are corporate websites, commercial job boards and recruitment agencies' sites [15]. The last two represent e-business. According to usual sequences of staffing process, commercial job board is relatively narrow form of e-business, where main source of revenue is advertisement of vacancies. In contrast, recruitment agencies offer much more services and its completions – from investigation of pool of potential candidates to support for hired employee. The vacancies market handled by job boards and recruitment agencies is divided by location, economy sector, job types and level [17].

There are different types of recruitment agencies that provide external recruiting services for organisations including retained search, contingency search, full-scale recruitment process outsourcing (RPO), on-demand RPO, and staff augmentation/placing consultants [18]. Retained search agencies, or executive search firms, provide search services for senior, executive, or other highly compensated positions. Retained search agencies work exclusively with clients, require an upfront retainer, and typically charge 30 to 35 percent of the salary of the position. Payments are made according to milestones in the recruitment process, so at least some fees will be paid regardless of whether a hire is actually made. Contingency search agencies search for candidates for their clients and get paid when a candidate they present is hired. Their search fees are typically 20 percent of the candidate salary when hired. Full-scale RPO agencies acts as a company's internal recruitment function for a portion or all of its jobs. RPO is utilized when a company experiences high volume staffing needs that internal HR can't cost-effectively handle along with their core responsibilities, or when there is no HR function in the company. On-demand RPO agencies provides recruiting, sourcing, and coordination on as requested basis rather than with long-term contracts. The agencies can also provide companies with own consultants for project or high-demand business objective.

To provide effective and efficient services, all types of recruitment agencies have to use e-recruitment advantages, which are associated with wider access, faster processes, reduced costs, corporate image promotion and reinforcement [17]. As Kelly Outsourcing and Consulting Group's survey "Global Trends in RPO and Talent Recruitment 2014" demonstrates, organisations are very interested in recruitment service outsourcing [19], so there are favourable conditions to develop commercial side of e-recruitment.

III. Value Creation Process in a Business

The most important aspect of value creation process is e-recruitment revenue streams that the source for business model. (e.g., advantages over other products, website design, attitude elevant information, user friendly interface and etc.). In practice, these main advantages in e-recruitment environment are promoting intensive knowledge (information) interactions between employers and applicants (i.e., its interactive characteristics), and the degree to which it provides necessary information. The goal of e-recruitment business is to influence job seeker attitudes toward job application and that the ability to accomplish this goal is affected by ability to manage key indicators. Management decisions should promote these key indicators that affecting job seeker decisions.

Various theories were developed many years ago, when the electronic market was not yet developed, and hence are suitable for the conventional market. Due to this reason, the authors of the article suggest that companies use the Alexander Osterwalder's value proposition concept [20] or the approach that is a constituent element of the author's developed business model canvas. (see Fig. 2). The Osterwalder's business model was formed based on Freeman's stakeholder theory [21]. The model is adapted to today's market needs and conditions, and the importance of the electronic environment, i.e. of the electronic market, in entrepreneurship is taken into account. Osterwalder distinguishes between "value proposition" and "elementary value proposition", which is an element of value proposition. The authors wish to draw attention to Osterwalder's "value life cycle" consisting of five stages: value creation, appropriation, consumption, renewal, and transfer [20].

All life cycle stages are linked to value consumption, using the electronic environment: value creation process (based on ICT) – adaptation of various products for the needs of an individual consumer. Value appropriation – "a one click purchase" at an internet shop. Value consumption – listening to music, watching a movie and etc. Value renewal – various

software updates, value transfer – disposal of old computers and other machinery, giving away unnecessary books and equipment for further use, etc.

Upon combining analysed models, it can be seen that the information and communication technologies (in the Osterwalder's model) or the information communication technology bear great importance in creating value for consumers and that they undoubtedly affect the company's image. The value concept is broadly used in various business models, including e-business models. The value forms the basis of several business models. The e-business model is based on mutual integration of key flows and values and implementation thereof between e-market participants, through the use of the e-environment. Three main e-business model elements can be distinguished: flows, participants, value. The term e-business model describes a broad spectrum of informal and formal models, which may be used in companies to depict various business aspects, such as operational processes, organisational structures, and financial forecasts [22].

The conceptual business models enable companies to analyse the current condition more broadly and to evaluate the already existing business. By employing this analysis, companies can develop new business development directions or improve the existing ones, because a modern market demands that companies change and are aware of their global condition. Entering the global market allows companies to reduce their dependency on local market fluctuations. The use of ICT promotes communication (Fig. 1); moreover, ICT is at the basis of the first stage "value creation" of the value life cycle.

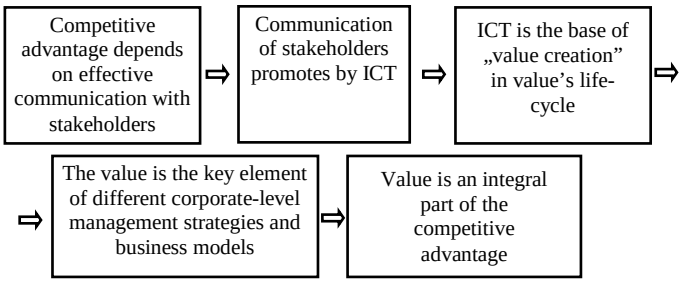


Fig. 1. Competitive advantage, ICT and value intermediation [23]

Based on the authors' performed study about the use of e-environment in e-business companies [24], having studied value formation theories, having analysed the types and theories of business models, the authors have drawn a conclusion that the most suitable course of action would be to base further development on the Osterwalder's Business Model Canvas [11]. Forbes has referred to this business model canvas as a simple instrument for creating innovative business models [3]. The model is based on active use of the e-environment in entrepreneurship. There are nine stakeholder groups at the basis of the model. Meanwhile, reciprocal and effective interaction and communication between the stakeholders promotes a company's competitiveness [25].

At the same time, value is an intrinsic part of a competitive advantage. It can be concluded that a competitive advantage depends on effective communication with stakeholders and customers. The previous study done by the authors about competitiveness of companies' shows that it is the use of communications networks, being a constituent element of competitiveness of companies, that the companies are using the least [3]. Thus, the authors of the paper assume that by increasing e-environment element as part of IC system, the competitiveness companies will also increase.

For the practical use of the quantitative evaluation model in e-business authors of the paper develop BMO [20]. The BMO's roots are found in management science and information systems research. Its four basic areas of preoccupation of a business model, the value proposition, the customer interface, the infrastructure management and the financial aspects stem from management literature [26; 27; 28]. The proposed business model elements providing practical contribution for business users. Its scientific roots originate in so-called design science [29] and its recent upsurge in Information Systems research [30; 31; 32; 33].

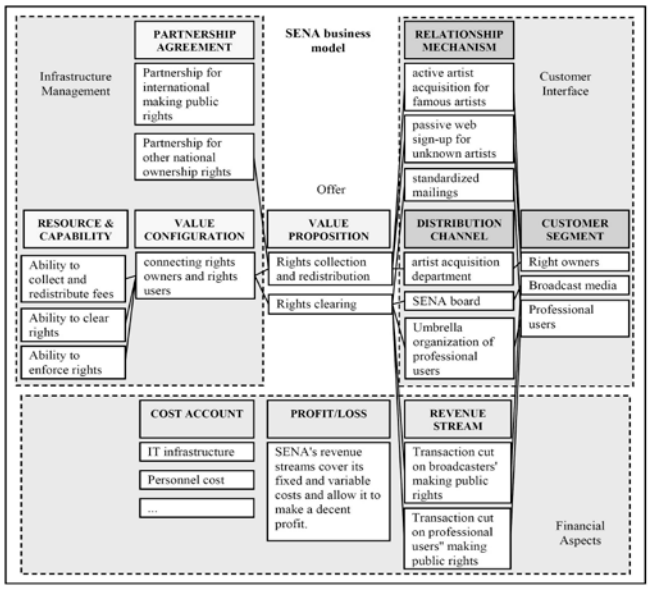


Fig. 2. SENA business model [34]

Authors of the paper see business model as a conceptual tool that contains a set of elements and their relationships that allows expressing the business logic of a company. It is a description of the what, the who, the how and the how much in a company [26; 27; 28] In other words it describes the value a company offers (what?) to one or several segments of customers (who?) and the architecture of the firm and its network of partners for creating, marketing and delivering this value and relationship capital (how?), in order to generate profitable and sustainable revenue streams (how much?). This business model has a good visualization, allowing understanding value creation logic.

In human resource management sources, many indicators for recruitment evaluation can be found allowing to measure some quantitative and qualitative aspects of this process (Table 1).

ISBN: 978-1-61804-327-6

TABLE I. EXAMPLES OF TRADITIONAL INDICATORS FOR RECRUITMENT [35]

Quantitative indicators	Qualitative indicators
Number of applicants attracted per method	Ratio of qualified to unqualified applicants attracted
Number of candidates interviewed	Job performance of employee attracted by method
Costs per applicant attracted	Tenure of employee attracted by method
Total recruiting cost per employee hired	Proportion of those interviewed who receive invitations to visit
Time from start to hiring of applicant	Organisation's or Applicants' satisfaction with recruitment process

Obviously, the traditional indicators for recruitment are not sufficient to evaluate e-recruitment as business. Using discussed above approach authors propose an additional set of indicators to evaluate e-recruitment (Table 2).

TABLE II. NON-FINANCIAL INDICATORS FOR E-RECRUITMENT COMPANY'S DEVELOPMENT DETERMINATION [3]

Indicator name	Explanation
Market share by purchased units	Market share by purchased units (%) = (Purchased units (%)) / (Total units of Purchased units (%))
Market share by revenue	Market share by revenue (%) = (Revenue from sales) / (Total revenue from market sales) x 100%
Relative market share	Relative market share (%) = (Brand market share) / (Biggest competitors market share) x 100%
Market concentration	Shows which a relatively small number of companies account for a large market share.
Brand development index	Brand development index = ((Brand sales for a group)/(Household in a group)) / ((Total brand sales)/(Total household))
Penetration (market or brand)	Market penetration (%) = (Customers who bought product) / (Total population) x 100%
Penetration share	Penetration share (%) = (Brand penetration) / (Market penetration) x 100%
Awareness	Awareness scale with point grading system.
The total number of active consumers	Percentage of consumers who at least once certain periods of time have bought a brand or product.
Desire to search	Percentage of the number of consumers who want to postpone purchase, changes stores or reduce purchases volume, focuses on other brands
Trial rate	Trial rate (%) = (Applied first time in period t) / (Total population (number of customers)) x 100%
Penetration t	Penetration = (Penetration t x Replicates rate (%)) x first purchased in period t
Sales forecast	Sales forecast = Penetration x The average purchase frequency x Average number of sold units
Repeated purchases	Number of repeated number of buyers = Trial number x Repetitions rate (%)
Sales forecast	Sales forecast = Penetration x The average purchase frequency x Average number of sold units
Repeated purchases	Number of repeated number of buyers = Trial number x Repetitions rate (%)
Trial volume	Trial volume = Trial number x Number of appliances
Repeated purchases volume	Repeated appliances volume = Repeated buyers number x Number of appliances made by one customer x Repeat times
Numerical distribution	Numerical distribution (%) = (Number of brand banners) / (Total number of banners) x 100%
All products distribution	All products distribution (%) = (Total sales volume of all brand's sales places) / (Total sales volume of sales places (banners)) x 100%
Distribution of particular type of product (PTP)	Distribution of PTP (%) = (Total PTP brand's sales places sales volume) / (Total sales volume of sales places (banners)) x 100%
Premium price	Premium price (%) = (Revenue market share)/(Product market share) x 100%
Impressions, Opportunities-to-See, Exposures	Impressions = Network Reach x Frequency Network Reach – the percentage of reach of the certain audience through the media; Frequency – certain ad or others activity views number, which done by one user.
Clickthrough Rate	Clickthrough rate = Clicks / Effect
Visits indicators	Visits, Sessions –a particular company's website first-time attendance of users. Visitors, Unique Visitors – the number of users who visit a particular website of the company for a given period. Clickstream – the way, how user find website. Abandonment Rate – the percentage of abandoned number of websites. Cookie – the small visitor's file, which recorded by website and helps identify user next on visiting time.
Website traffic statistics dynamics	How many internet users visited a given site during a given period.
Web site visit duration	Average time which users spent on the site.
Site visitors characterization	Behaviour: new and repeated visitors, frequency etc. Demographic data: language, location, gender, etc.
Technologies	Technologies used in site attendance: device, from which the attendance made; browser and operating system, with which help made attendance; provider used for site visiting; visitors flow (what content were visited on the site); in what way was visited site - directly or via link and/or divert from other sites.

Business Model focuses on the design of a company's value creation model, visualization of value creation in BMO is highly relevant, and such visualisations are used to explain a model to stakeholders. Additionally, it proposes specific diagrams, for instance for distribution channel strategies or activity configurations.

IV. THE CONCEPT OF E-RECRUITMENT AS A BUSINESS MODEL

Authors see e-recruitment business model advantages in e-recruitment methods, transferring knowledge for job seeker

through automated processes creating the ability to accomplish these processes in a shorter time. These advantages are the main technical feature of the e-recruitment business. Particularly it is interaction with job seeker in e-environment, thus reducing cost associated with first phase (Socialisation, Fig. 3).

The impact of information was noted long ago by Behling, Labovitz, and Gainer (1968), who observed that job choice decisions are based on thoughtful assessment of key information concerning objectively measurable job attributes such as pay and working conditions [36].

Information interaction is suggested by authors to be central point as soon as job seeker's interaction process is very important. The importance of knowledge is found in information processing studies, which have demonstrated that prior knowledge of product characteristics greatly affects the way in which consumers investigate, process, and organize product related information [37].

This phase (Externalisation, Fig. 3) is analysed in management studies showing interaction experience create substantially different variations in the ways that job seekers gather and use labour market information [38]. Experienced customers are better able to extract and analyse important central information [39].

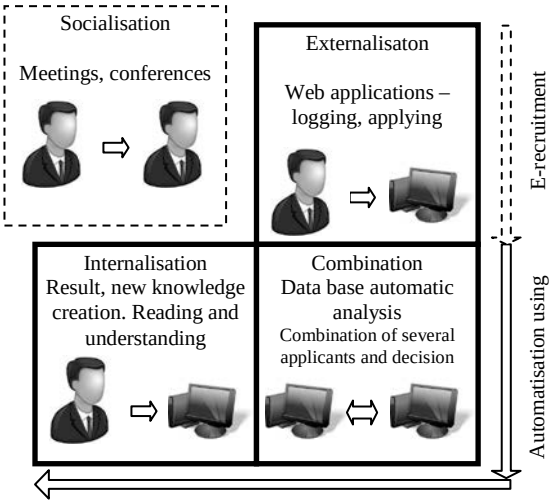


Fig. 3. E-recruitment as a business model (framework) [40]

The implication of these findings for e-recruitment is that e-recruitment business model potentially could neglect these shortcomings by using only last two phases (Combination and Internalisation). The advance of modern ICT has launched the Industry 4.0, to take up a leader role in industrial IT which is currently revolutionizing the manufacturing engineering sector [41].

Technology breakthrough is allowing to increase the level of automation for interaction with job-seekers and labour cost decreased. These trends will is more focused on intangible assets (associated with IC) managing company data flow, plantspecific software and the "hardware" of manufacturing technology.

Since ICT is only one part of the Industry 4.0, the other is its use in the industrial sector and the utilization of the benefits that it brings to the value chain (Fig. 4).

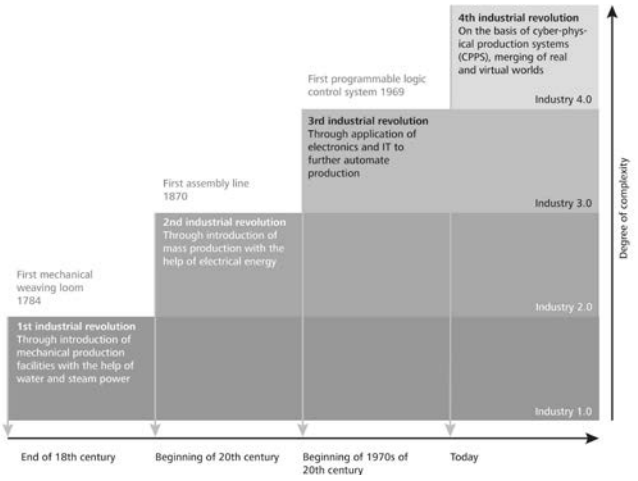


Fig. 4. The evolution of embedded systems into the internet of things, data and services [42]

"Industry 4.0" (sometime referred as Smart industry) advantages are coming from the technological evolution from embedded systems to cyber-physical systems. Industry 4.0 connects embedded system production technologies and smart production processes associated with the new technological age advantages (Fig. 4). Decentralized intelligence helps create intelligent object networking and independent process management, with the interaction of the real and virtual worlds representing a significant new aspect of the manufacturing and production process. Industry 4.0 creates the vision (Fig. 4) of an entirely networked production, in which orders managed automatically throughout entire value chains, order processing machines and material and organize their delivery to the customer [43].

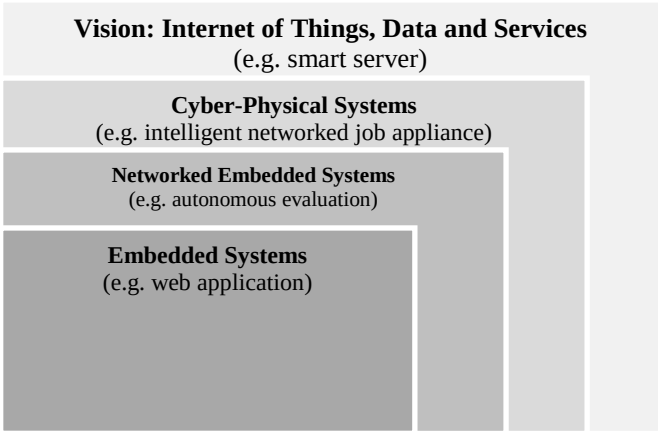


Fig. 5. The evolution of embedded systems into the internet of things, data and services [44]

Using these data efficiently provides a considerable competitive advantage (reducing downtimes, accurate planning, reducing unit costs and etc.).

ISBN: 978-1-61804-327-6

New Industrial revolution (Industry 4.0) is also called Internet of Things, Data and Services (Fig. 5). Cyber-physical systems provide the basis for the creation of an Internet of Things, which combines with the Internet of Services to make Industry 4.0 possible.

New Industrial revolution (Industry 4.0) is also called Internet of Things, Data and Services (Fig. 5). Cyber-physical systems provide the basis for the creation of an Internet of Things, which combines with the Internet of Services to make Industry 4.0 possible.

The widespread adoption by e-recruitment automatic operations of ICT is increasingly blurring the boundaries between the real world and the virtual world in what are known as cyber-physical production systems (CPPSs) [45].

In contrary to e-recruitment, studies of interviewers as recruiting sources have found that such factors as interviewer personableness, competence, empathy, interest in the applicant, communication skills, and enthusiasm often play significant roles in applicant interest in a job and intention to accept a job offer [46; 47; 48]. E-recruitment lacks these advantages and these findings show that motivation of job seeker will be affected [49] and decreasing their motivation [50]. These findings emphasize that search motivation is a key element of a job-seeker, and that e-recruitment must carefully consider negative effects.

Nonetheless, the authors see E-recruitment as a business model is concentrated in “Combination” phase efficiency due to business requirements and new trends.

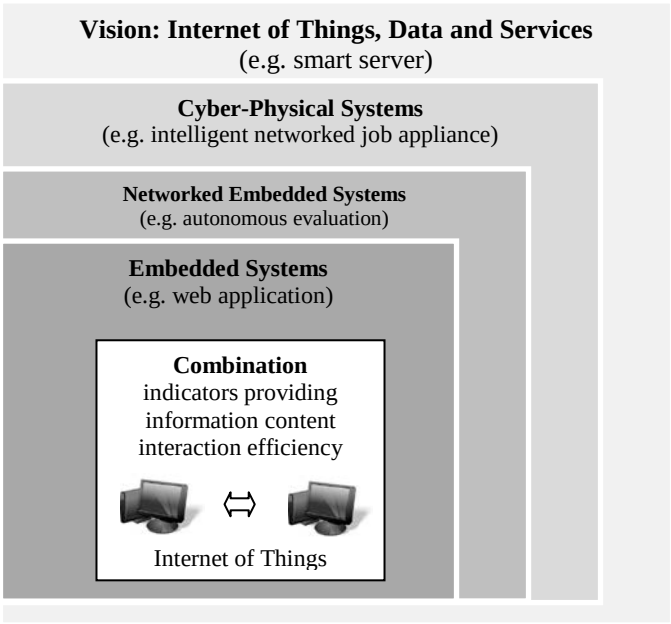


Fig. 6. E-recruitment as a business model (framework) [created by the authors based on Federal Ministry of Education and Research [44]]

V. Conclusion

The contemporary cornerstone advantage of e-recruitment methods lies in labour cost. E-recruitment methods represent a growing and high potential opportunity for business to reduce recruiting costs [51]. Hence the goal of e-recruitment is to satisfy job-seekers needs by providing competitive virtual environment to traditional one. The ability to manage value creation process as interaction effect`s efficiency is of the main goals, meeting job-seeker needs and web applications capabilities.

Controversially the problems discussed here about socialisation affect toward motivation and the effort needed to attract a job offer should be taken in consideration. E-recruitment focus only on outcomes such as job acceptance decisions or application attractiveness should be carefully evaluated. Thus e-recruitment`s effect on initial job-seeker interest is limited, decreasing the potential possibility to attract a job-seeker and receive positive feedback. Information interaction playing a certain role for a job seeker's attitude and job acceptance decisions, but the motivation enhancing possibilities are likely to be less effective than traditional ones.

Such our findings provide e-recruitment with a contemporary approach in value creation.

Based on our business model approach, e-recruitment should create a virtual recruiting environment that effectively interacts with job-seeker partially motivating his decision process. E-recruitment as a business model should be based on indicators providing effective (user friendly) information content required to affect job seekers` positive decision, besides labour cost efficiency advantages. Offering modern “Internet of Things” concept to e-recruitment provides new business models with efficient framework for indicators identification. To contribute to this process, authors present their indicators to evaluate important issues associated with information interaction in order to develop effective e-recruitment business model.

References

[1] Deloitte, “Global Human Capital Trends 2015“, 2015. Retrieved from http://www2.deloitte.com/us/en/pages/human-capital/articles/introduction-human-capital-trends.html

[2] Boston Consulting Group, “Creating people advantage 2014-2015“, WFPMA. Retrieved 2015 from https://www.bcgperspectives.com/content/articles/human_resources_creating_people_advantage_2014_how_to_set_up_great_hr_functions/

[3] D. Sceulovs, “Use of Electronic Environment in Entrepreneurship Development”. (Doctoral Thesis). Riga Technical University, LV, 2013.

[4] V. Beattie, K.Pratt, “Business reporting: Harnessing the power of the internet for users”. Edinburgh, UK: Institute of Chartered Accountants of Scotland, 2001.

[5] C. Eustace, “The intangible economy-impact and policy issues“, Bruselas, Comisión Europea, 2000.

[6] B. Lev,”Communicating knowledge capabilities”, Stern School of Business, New York, 2000.

[7] W.S. Upton, “Business and financial reporting, challenges from the new economy”, In Financial Accounting Series (FASB), Special report (219-A), 2001. Retrieved from http://www.fasb.org/new_economy.shtml

[8] I. Lee, “The evolution of e-recruiting: A content analysis of Fortune 100 career web sites“ Journal of Electronic Commerce in Organizations, 3, 2005, pp. 57−68.

[9] J. Gill, “Now hiring“, July 18, 2001,. Apply on-line. Businessweek Online. Retrieved from http://www.businessweek.com/careers/content/jul2001/ca20010718_003.htm

[10] M. Pastore, "Web expands role in corporate recruiting". ClickZ Network. 2000, Retrieved from http://www.clickz.com/showPage.html?page=330331

[11] Brown, D. Unwanted online job seekers swamp HR staff. Canadian HR Reporter, 17, 2004, pp. 1−2. Business Model Foundry. The business model canvas. 2014. Retrieved from http://www.businessmodelgeneration.com/canvas

[12] D.G. Allen, J.R. Van Scotter, R.E. Otondo, "Recruitment communication media: Impact on pre-hire outcomes", Personnel Psychology, 57, 2004, pp. 143−171.

[13] J.A. Breaugh, M. Starke, "Research on employee recruitment: So many studies, so many remaining questions", Journal of Management, 26, 2000, pp. 405−424.

[14] A.E. Barber, "Recruiting employees: Individual and organizational perspectives". Thousand Oaks, CA: Sage Publications, 1998.

[15] M. Armstrong, "Armstrong's handbook of human resource management practice", 11th ed., 2009.

[16] HRZone. "What is E-Recruitment?", 2015, Retrieved March 16, 2015, from http://www.hrzone.com/hr-glossary/what-is-e-recruitment

[17] L. Barber, "E-recruitment developments". Brighton, UK: Institute for Employment Studies, 2006. Retrieved from http://www.employment-studies.co.uk/resource/erecruitment-developments

[18] A. Greenberg, "5 types of recruiting companies", 2013. Retrieved from http://www.recruitingdivision.com/5-types-of-recruiting-companies/

[19] P. Berklich, "Global trends in RPO and talent recruitment", 2014. Retrieved from http://www.kellyocg.com/Workforce_Trends/Global_Trends_in_RPO_and_Talent_Recruitment_2014/#.VRmsDo5G2-0

[20] A. Osterwalder, "The Business Model Ontology - a proposition in a design science approach". (Ph.D. thesis), 2003. Retrieved from http://www.uniempre.org.br/user-files/files/TheBusiness-Model-Ontology.pdf

[21] R. Freeman, Edward "Strategic management: a stakeholder approach". Boston, MA: Pitman, 1984.

[22] Laudon, K., Traver, C. "E-commerce", (6th ed.), Upper Saddle River, NJ: Prentice Hall, 2010.

[23] V. Shatrevich, D. Sceulovs, "Using of e-environment as a tool for value creation in a company". WMSCI 2014 - 18th World Multi-Conference on Systemics, Cybernetics and Informatics, Proceedings, Vol. 1, 2014, pp.62-67.

[24] D. Sceulovs, E. Gaile-Sarkane, "Identification of factors affecting consumer habits in the e-environment". Faculty of Business Management at Vilnius Gediminas Technical University 6th International Scientific Conference "Business and Management–2010 Proceedings, Vol. 13, No. 14.05, 2010.

[25] A. Osterwalder, Y. Pigneur, "Business model generation", Hoboken, NJ: John Wiley & Sons, 2009.

[26] J. Hagel III, M. Singer, "Unbundling the Corporation", McKinsey Quarterly, 3, 2000, pp. 148-161.

[27] R.S. Kaplan, D. P. Norton, "The balanced scorecard--measures that drive performance"', Harvard Business Review, 70, 1992.

[28] C. Markides, "All the Right Moves". Boston, MA: Harvard Business School Press, 1999.

[29] C. Owen, "Design research: Building the knowledge base". Journal of the Japanese Society for the Science of Design, 5, 1997, pp. 36-45.

[30] Y.A. Au, "Design science I: The role of design science in electronic commerce research". Communications of the Association for Information Systems, 7 , 2001.

[31] N.L. Ball, "Design science II: The impact of design science on e-commerce research and practice". Communications of the Association for Information Systems, 7, 2001.

[32] A.R. Hevner, S.T. March, J. Park, S. Ram, "Design science in information systems", MIS Quarterly 28, 2004, pp. 75-105.

[33] S.T. March, G.F. Smith, "Design and natural science research on information technology". Decision Support Systems, 15, 1995, pp. 251-266.

[34] J. Gordijn, A. Osterwalder, Y. Pigneur, Comparing two business model ontologies for designing e-business models and value constellations. BLED 2005 Proceedings, 15, 2005.

[35] J. J. Phillips, "Accountability in human resource management". Routledge, 1996.

[36] O. Behling, G.Labovitz, M. Gainer, "College recruiting: A theoretical basis. Personnel Journal, 47, 1968, pp. 13−19.

[37] J.W. Alba, J.W. Hutchinson, "Dimensions of consumer expertise". Journal of Consumer Research, 13, 1987, pp. 411−454.

[38] S.L. Rynes, M.O. Orlitzky, R.D. Bretz, "Experienced hiring vs. college recruiting: Practices and trends". Personnel Psychology, 50, 1997, pp. 309−340.

[39] B.M. Meglino, A.S. Denisi, E.C. Ravlin, "Effects of previous exposure and subsequent job status on the functioning of realistic job previews". Personnel Psychology, 46, 1993, pp. 803−822.

[40] I. Nonaka, "Managing innovation as an organizational knowledge creation process". In J. Allouche & G. Pogorel (Eds.), Technology management and corporate strategies: A tricontinental perspective, Amsterdam, NE: Elsevier, 1995, pp. 73–109.

[41] Germany Trade and Invest. "Future markets. Industrie 4.0 Smart manufacturing for the future", 2014, Retrieved from http://www.gtai.de/GTAI/Content/EN/Invest/_SharedDocs/Downloads/GTAI/Brochures/Industries/industrie4.0-smart-manufacturing-for-the-future-en.pdf

[42] J. Schlick, P. Stephan, D. Zühlke, "Produktion 2020: Auf dem Weg zur 4. industriellen Revolution". IM -- Fachzeitschrift für Information Management und Consulting, 27, 2012, pp. 26-33.

[43] R. Berger, "Industry 4.0: A driver of innovation for Europe", 2014, Retrieved September 17. from http://www.think-act.com/blog/2014/industry-4-0-a-driver-of-innovation-for-europe/

[44] Federal Ministry of Education and Research. "The evolution of embedded systems into the internet of things, data and services" ("Industry 4.0" Working Group report). 2012. Retrieved from http://www.bmbf.de/en/19955.php

[45] Federal Ministry of Education and Research. "Securing the future of German manufacturing industry". Recommendations for implementing the strategic initiative INDUSTRIE 4.0 (Final report of the "Industrie 4.0" Working Group), 2013, Retrieved from http://www.plattform-i40.de/finalreport2013

[46] M.M. Harris, L.S. Fink, "A field study of applicant reactions to employment opportunities: Does the recruiter make a difference?" Personnel Psychology, 40, 1987, pp. 765−784.

[47] S.D. Maurer, V. Howe, "Comparing personnel representatives and line managers/engineers as engineering recruiters: An application of the elaboration likelihood model of persuasive communication". Journal of High Technology Management Research, 6, 1995, pp. 243−259.

[48] G.N. Powell, "Applicant reactions to the initial employment interview: Exploring theoretical and methodological issues". Personnel Psychology, 44, 1991, pp. 67−83.

[49] R. E. Petty, J. T. Cacioppo, D. Schumann, "Central and peripheral routes to advertising effectiveness: The moderating role of involvement". Journal of Consumer Research, 10, 1983, pp. 35−146.

[50] S.B. MacKenzie, R. A. Spreng, "How does motivation moderate the impact of central and peripheral processing on brand attitudes and intentions? " Journal of Consumer Research, 18, 1992, pp. 519−529.

[51] P. Cappelli, "Making the most of on-line recruiting". Harvard Business Review, 79, 2001, pp.139−146.

Design and Implementation Unified Model for Testing Object-Oriented Application Development Tools

Pavel P. Oleynik, Russia, Rostov-on-Don, xsl@list.ru

***Abstract*—** The paper presents a unified model for testing tools for object-oriented application development. Based the available papers were identified shortcomings of existing work and identified the following optimal criteria, which shall comply the resulting model:

1. To deep inheritance hierarchies
2. To presents of multiple inheritance hierarchies
3. To presents of abstract classes in the hierarchy
4. To presents of multiple (n-ary) associations
5. To presents of associations with attributes
6. To presents of a composition between classes
7. To presents of recursive associations
8. To presents of associations between classes belonging to the same inheritance hierarchy
9. To presents of association classes
10. To presents between the association class and other classes
11 To presents enumerations in model

With a unified graphical language UML class diagram unified model testing. The paper we verified compliance with the resulting implementation of the selected criteria was presented.

Currentlythe implementation of applications using object-oriented programming languages and relational databases. To overcome the object-relational mismatch it is necessary to implement object-related mapping patterns presents. The paper presents three methods used to represent the class hierarchy highlighted the advantages and disadvantages of each method.

For test the feasibility a unified model chosen development environment SharpArchitect RAD Studio which is designed object applications in C# and are implementing a relational database. The paper presents the developed object model in the form a class diagram showing the interfaces and inheritance relations diagram containing all the tables and columns the resulting database.

In the conclusion recommendations on the areas for further development work and identified the need of implement a unified model with other approaches proposed by the authors was used.

***Keywords* —** UML; Object modeling; Design of Information Systems; Databases; Object-oriented design; Object-Relational Mapping Patterns; Impedance Mismatch

I. INTRODUCTION

At the moment there are many tools provide object approach to application development. Despite the existence of their own advantages and disadvantages the main goal is provide the advantages of the developer of object-oriented paradigm. The paper are describes in detail the unified model test tools development of object-oriented applications for demonstration, graphical Unified Modeling Language which used. The practical implementation of the model is demonstrated by the use of classical methods (patterns) object-relational mapping (ORM) in the tool, developed the author. The object model is put into a relational database environment. This approach is most justified from the point of view the author, because the RDBMS is the most popular type of database management systems now.

II. DESIGN UNIFIED MODEL TESTING

When designing a unified testing model used the same approach as in the description of the design patterns in [1]. This approach is involves the description of reusable solutions widespread problems in software development without reference to particular domain. The main task of this section – is a description of the model and the structural elements (classes and associations), and not the correctness of the model and the accuracy of its fitness for a particular domain area.

Standard graphical language modeling various aspects of object systems is the language UML. This language is namely structural class diagrams will be discussed in this paper. As a result under the unified model test tools development of object-oriented applications we mean a class diagram, consisting of classes and attributes and containing common practice relationship classes.

The idea of the article is not new and there are works of similar subjects. In [2] has attempted to construct a unified model testing. However, there were no multiple (n-ary) associations and association with attributes that are an integral part of any complex information system.

In [3] presented test model to study the design of object-oriented databases. But the model is relatively simple, which is justified by its purpose. This article used dignity previously existing works and corrected drawbacks of them.

Before designing a unified model testing were nominated optimality criteria (OC) is representing the requirement of a certain structural elements in the class diagram, and which must comply with the finished implementation. Have been put forward the following requirements for the unified model test tools development of object-oriented applications:

1. Must have deep inheritance hierarchies. In realworld applications, very often there are deep hierarchy, is the

relational of inheritance and combining transitive least three classes.

2. To presents of multiple inheritance hierarchies. This will show a variety of options and modes available in the development tool.
3. To presents of abstract classes in the hierarchy. Abstract classes cannot have instances in the system and described as a container for attributes and methods used in the inherited (instantiated) classes.
4. To presents of multiple (n-ary) associations. In applications that automate realworld domains, often an association involving three or more classes. Such a relationship is called multiple or n-ary associations.
5. To presents of associations with attributes. Many domains contain attributes that do not belong to certain entities (classes), and their values appear only in the organization of associations between instances of classes. The designing unified model should have associations with attributes.
6. To presents of a composition between classes. Composition - an association between the classes which are Part and Whole. The peculiarity is that the class represents a Part can belong to only one instance of the class that represents the Whole. In this class represents the Whole manages the life cycle is a class represents a Part. When removing the Whole all Parts also deleted. This peculiarity of behavior is very important for many application domains.
7. To presents of recursive associations. Recursive call the association, the ends of which bind the same class. These relationships allow you to implement a hierarchy of subordination.
8. To presents of associations between classes belonging to the same inheritance hierarchy. In terms of implementation is necessary to provide the implementation of the association, the edges of which are associated classes belonging to the same inheritance hierarchy, are represents the base class and the child together.
9. To presents of association classes. Association class - an association which at the same time a class. Especially the use of that class association represents a unique association, i.e. combination of instances of classes in this association is unique.
10. To associationed between the association class and other classes. From a theoretical point of view, the association class is a class, so it can participate in other associations. From the point of view of the implementation of the class association presents a class that contains the attributes (fields or properties of the programming language) that refer to other classes. In turn, for the organization of the association with the class association necessary depending class to create an attribute whose type supports class association.
11. To presents enumerations in model. From a theoretical point of view, enumeration is a set of predefined constants, and the user can not extend this set by adding new values.

In accordance with the selected criteria was implemented hierarchy shown in Fig. 1.

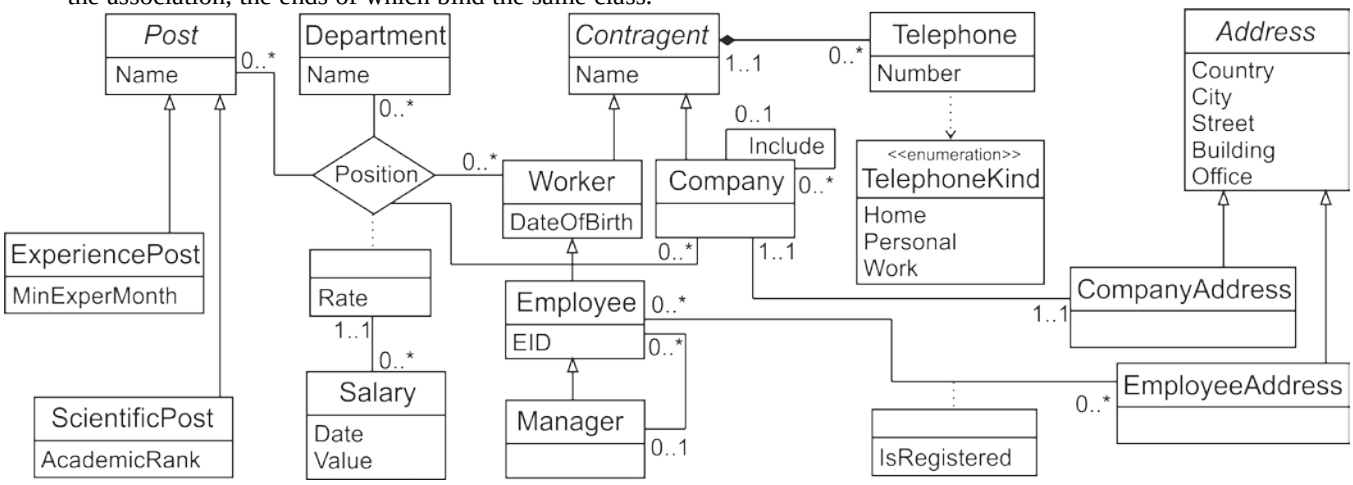


Fig. 1 - Unified model for testing object-oriented applications development tools

Consider the appointment of the main classes of diagrams are presented. As mentioned earlier this class diagram is a fictional and is not intended to describe a particular domain therefore contains some illogical (fictional) classes and associations.

For representation of employees and organizations assigned to the base abstract class Contragent. Inherited Company class is present organizations and the class Worker is the base for the employee of organization. Inherited Employee class is an employee and an attribute EID, representing the employee unique number. Class Manager is the staff who are heads of other workers.

Post an abstract class is a position that can be occupied by staff. Inherited class ExperiencePost is a position that requires a minimum amount of experience of the applicant, expressed as number of months (attribute MinExperMonth). The second class is implemented ScientificRank describes the position of

the applicant, which requires the presence of a scientific degree, whose name is value in the attribute AcademicRank.

For presentation departments of organizations and entering into an n-ary association a class of Department was introduced. Salary class is paid wages, accrued to employees occupying positions represented by a complex association which called Position.

Class Telephone allows saving the number of phone of company. Phone type (like Home, Personal, Work) represented by enumeration TelephoneKind. For presentation address used by the base abstract class Address. Two derived class CompanyAddress and EmployeeAddress used to represent the address of the organization and address of the employee, respectively.

Check the conformity of the model presented previously selected criteria of optimality. The need for a deep class hierarchy, represented by at least three transitive inherited classes, described OC_1 and implement a class Contragent, Worker, Employee, Manager. In addition to this, there are two hierarchies: 1) Post, ExperiencePost (ScientificPost); 2) Address, CompanyAddress (EmployeeAddress). I.e. the model contains multiple inheritance hierarchies, therefore, the condition OC_2. The presence of abstract classes in the hierarchy due OC_3 and holds classes Post, Contragent and Address.

OC_4 requirements are also performed as there are n-ary association Position, combining classes Post, Department, Worker, Company. Described association has an attribute Rate, which implemented class association and binary association between Employee and EmployeeAddress classes also contains an attribute (IsRegistered) it can be argued that the requirement OC_5 fulfilled.

Each contractor represented derived from Contragent classes, a list of telephone numbers represented instances of Telephone, and both classes related with composition, OC_6 requirement is satisfied. Unified model allows you to store information about a group of companies, organize the tree structure using a recursive association connects Company class with a same. The presence of recursive association dictated OC_7.

In OC_8 written requirement for associations between classes belonging to the same inheritance hierarchy. Figure 1 between classes Employee and Manager provides this association satisfying OC_8. As previously noted, the models have a association class Position, which corresponds OC_9. Described association class is linked with addition association with Salary class. This is a consequence of the implementation OC_{10}. The presence of the models listed due to the implementation of OC_{11}. Of the present disclosure can be seen that the unified model is fully consistent with all previously selected criteria of optimality. Therefore we can move on to the implementation of the unified model.

III. The Classical Object-Relational Mapping Patterns

To implement of this model development environment software systems based on the organization of the metamodel object system presented in [4-5] was used. This development environment is called SharpArchitect RAD Studio and as storage of information uses a relational DBMS. Because information system is designed in terms of object-oriented paradigm, and implemented in a relational database environment, there is a so-called "object-relational impedance mismatch" to overcome the consequences of which object-relational mapping patterns are used. The most commonly used patterns for represent the class hierarchy.

In SharpArchitect RAD Studio implemented three classic patterns for implementing object-oriented inheritance relationships of classes in a relational structure (relational tables), presented in Fig. 2 [2, 4].

Consider the basic patterns is presented in more detail. Single Table Inheritance pattern physically represents an inheritance hierarchy of classes in a single relational database table whose columns correspond to the attributes of all classes within the hierarchy and allows you to display the structure of inheritance and to minimize the number of joins that must be performed to extract information. In this pattern each instance of the class represented by one row of the table. When you create the object values are entered only in the columns of the table that match the attributes of the class, and all the rest are empty (have a null-value).

The pattern has advantages:

- In the structure of the database contains only one table are representing all classes of whole hierarchy.
- To selection of instances of classes hierarchy do not need to make the joins of tables.
- Move fields from a base class to a derived (as well from the derivative in the base) does not require changes to the structure of the tables.

The pattern has disadvantages:

- In the study of the structure of the database tables can cause problems, because not all the columns in the table are intended to describe each domain class. This complicates the process of refining the system in the future.
- If you have a deep inheritance hierarchy with a large number of attributes, many columns can have empty values (null-values). This leads to inefficient use of the available space in the database. However, modern DBMS can compress strings containing a large number of null-values.
- Table may be too large and contain a huge number of columns. The main way to optimize the query (to reduce the execution time) is created a covering index. However, the index set and a large number of queries to a single table can lead to frequent blockages that will have a negative impact on the performance of software applications.

ISBN: 978-1-61804-327-6

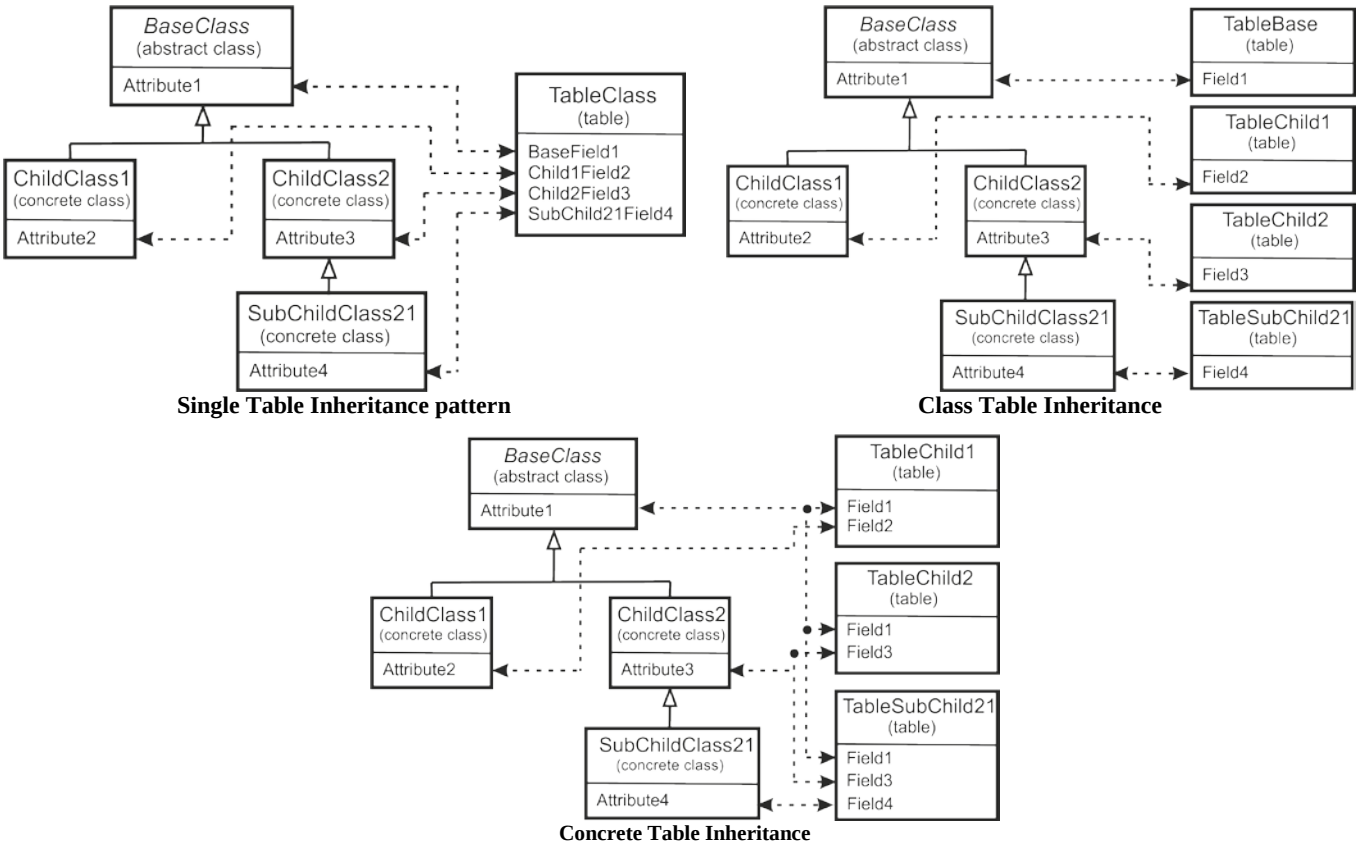


Fig. 2 - Classical object-relational mapping patterns which used to represent the class inheritance in the form of a relational structure (relational tables)

An alternative pattern is called Class Table Inheritance, representing a hierarchy of classes for one table for each class (as an abstract and concrete). Class attributes are mapped directly on the columns of the corresponding table. With this method, the key is the task of joins the respective rows of several database tables that represent a single object of domain.

The pattern has the following advantages:

- Each table contains a field, the corresponding attribute of a certain class. The therefore tables are easy to understand and take up little space on your hard drive.
- The relationship between the object model and relational database schema is simple and clear.

However, there are disadvantages:

- When you are create an instance of a particular class you want to upload data from several tables, which requires either their natural joins or a plurality of database calls followed by join results in memory.
- Move the fields in the derived class or base class requires changes in the structure of several relational tables.
- Base class table can become weaknesses in performance, since access to such tables will be carried out too often, leading to a variety of locks.
- High degree of normalization can be an obstacle to the implementation of unplanned advance queries.

The Concrete Table Inheritance pattern present is an inheritance hierarchy of classes using one table for each concrete (non-abstract) class of the hierarchy. From a practical perspective, this pattern assumes that each instance of the class (object), which is in memory, will be shown on a separate row in the table. In addition, each table in our case contains columns corresponding to attributes as a particular class, so all of his ancestors.

The advantages are that:

- Each table not contains extra fields, so that it is convenient to use in other applications that do not use object-relational mapping tools.
- When creating objects of a certain class in the application memory and retrieve data from a relational database sample is made of a single table, i.e. is not required to perform relational joins.
- Access to the table is carried out only in the case of access to a particular class, thus reducing the number of locks imposed on the table and spread the load on the system.

There are disadvantages:

- Primary keys can be inconvenient by handling.

ISBN: 978-1-61804-327-6

- There is no ability to model relationships (association) between abstract classes.
- If the class attributes are moved between base classes and derived classes needed to change the structure of several tables. These changes are not as often as in the case of Class Table Inheritance pattern, but they cannot be ignored (as opposed Single Table Inheritance pattern in which these changes are absent).
- If in base class to change the definition of at least one attribute (for example, change the data type), it will require to change the structure of each table representing a derived class because a superclass fields are duplicated in all tables of its derived classes.
- In implementing the method of searching for data in the abstract class is required to view all the tables represents an instance of the derived classes. This requires a large number of database calls.

Selection of an required ORM-pattern depends on the initial logical model, i.e. from the class hierarchy of the domain. At the same time can be used two or more ORM-patterns, which is associated with the need to optimize the structure of a relational database and reduce the number of tables used, which will increase the speed of data retrieval queries.

After describing SharpArchitect RAD Studio object-relational mapping patterns which are available to the developer we can start implementing the unified model for testing tools.

IV. Implementation of the Unified Testing Model

In order to simplify the implementation of the three existing class hierarchies in Figure 1 will separate in available ORM-patterns. The result is shown in Fig. 3.

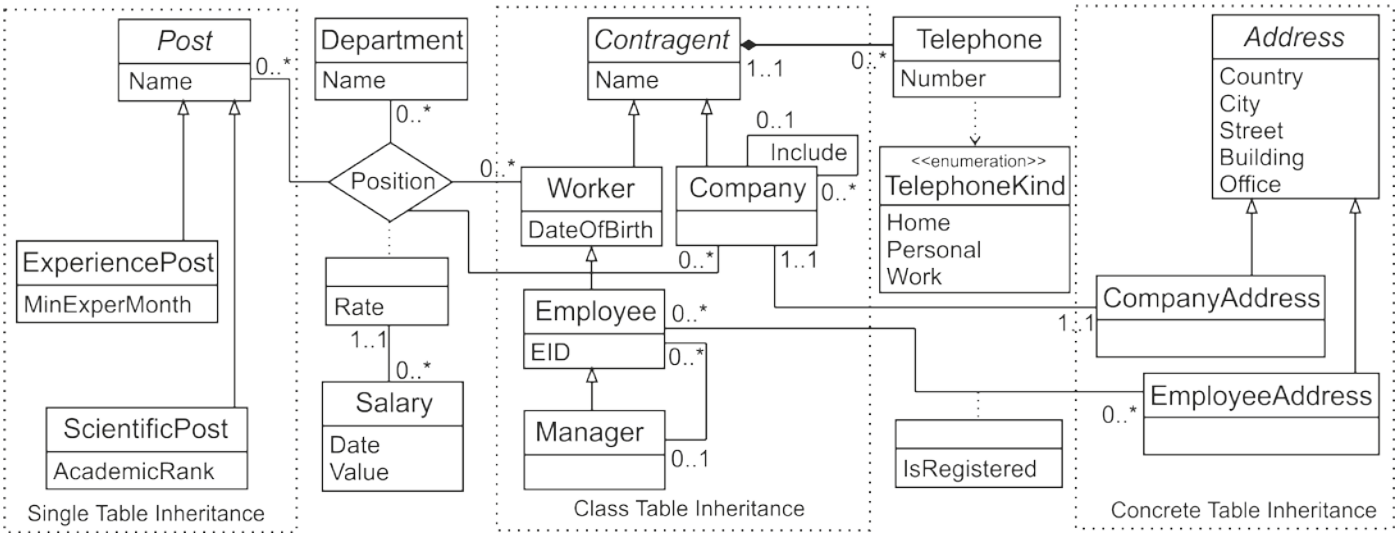


Fig. 3 - The use of the classical ORM-patterns for the implementation of the unified model for testing object-oriented applications development tools

The Single Table Inheritance for the class hierarchy Post, ExperiencePost (ScientificPost) was used. As a result, it is assumed that in the RDB will create one single table (relational table), which will be retained instances of all listed non-abstract classes. For the class hierarchy with classes Contragent, Worker (Company), Employee, Manager uses the Class Table Inheritance pattern. I.e. for all classes regardless of whether he or abstract concrete will create a separate table in RDB. Address class is abstract and has no association with other classes in model, so it will not create a separate table in the RDB. And for child classes will be created two tables (one for each heir). I.e. in hierarchy Address, CompanyAddress (EmployeeAddress) was used Concrete Table Inheritance. For other classes outside the hierarchy described, will be created on a separate relation table.

One of the main features of SharpArchitect RAD Studio support multiple inheritance is implemented by means of interfaces C# language construction, as described in detail in [4]. Used C# language does not support this syntax as an association. To represent the binary associations, regardless of the multiplicity was used properties (property construction), containing a single value or collection of values.

Multiple n-ary association are represents a separate class, the attributes of these associations (as well as the attributes of binary associations) are converted into property of classes. To simplify information searching and extraction of all the associations are bidirectional both ends of the relevant classes there are properties whose type corresponds to the opposite end of the class association. All of the above arguments are presented graphically in Fig. 4.

In implementing the interfaces used language C#, so it is impossible italics abstract classes. Bidirectional associations are shown corresponding arrows connecting classes. In implementing the association used the following approach. From the "one" was declared property, which is a type of list (C# type IList<>), containing the elements, which is a type of class, located on the side "to-many". From the "to-many" is declared in the class property whose type is a class, located on the side "one". Association of the "many-to-many" (without attributes) can be represented by two lists is declared in class

antagonisms. In a SharpArchitect RAD Studio development environment has a number of base classes that implement the most common functionality. For example, the class IBaseRunTimeDomainClass is the root of all domain classes. To implement the tree structure will enough inherited from IBaseRunTimeTreeNodeDomainClass. At the time code generation will automatically generate additional attributes Nodes and Owner, allow you to save a reference to the parent and subnodes, respectively. It is implemented in such a way recursive association. For submission to the transfers and sets used syntax construction "enum".

Applying the classical ORM-patterns was obtained relational database schema of the unified model now. Fig. 5 is depicts the result.

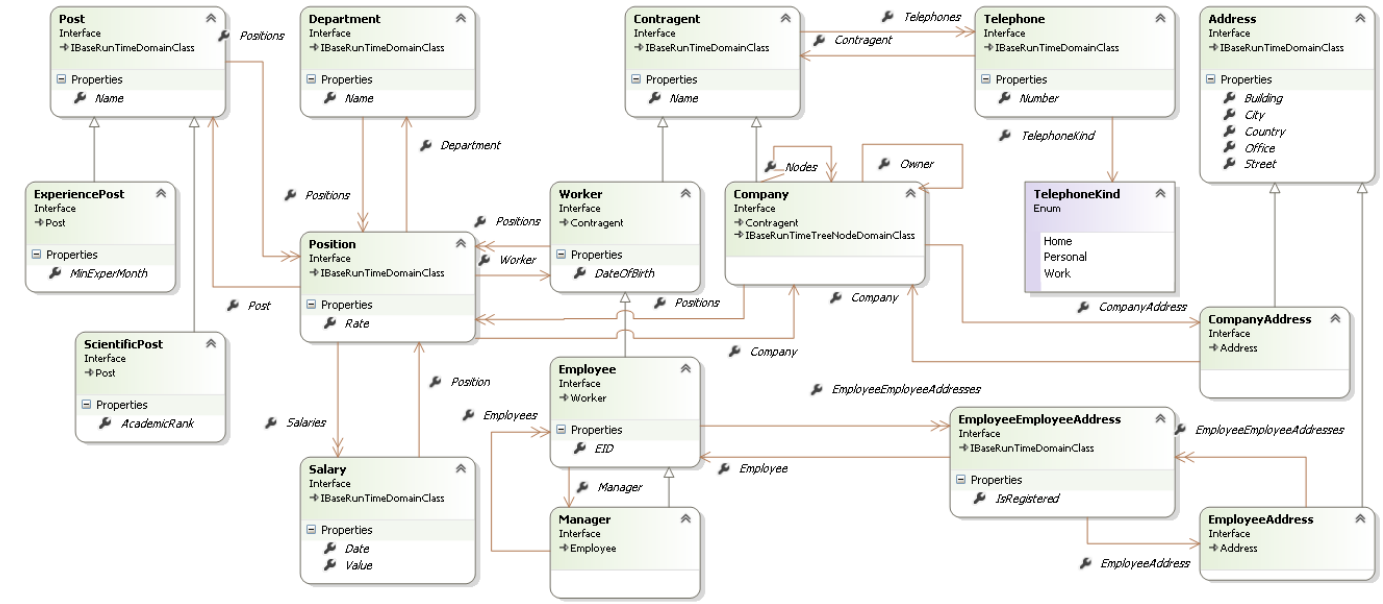


Fig. 4 - Unified model for testing object-oriented application development tools, implemented in SharpArchitect RAD Studio in C#

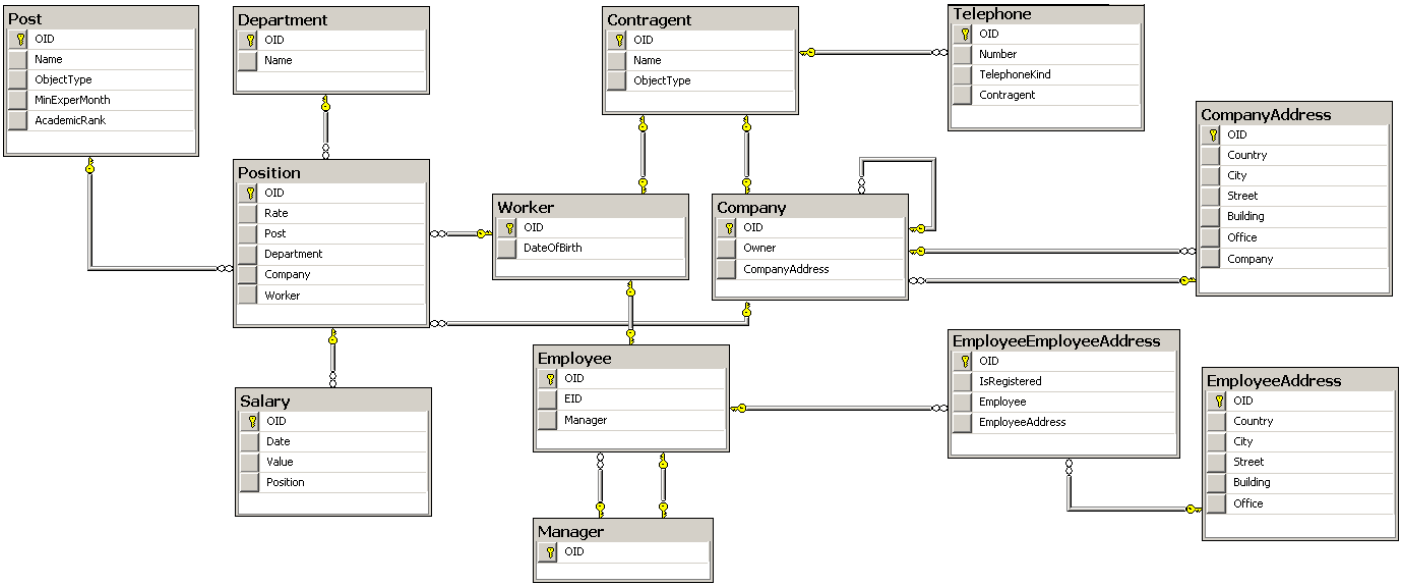


Fig. 5 - A relational database schema of the implementation of the unified model testing in SharpArchitect RAD Studio

Figure requires is explanation. For all posts submitted by three classes of Post, ExperiencePost and ScientificPost, created one single table Post, which has all the attributes of classes. Additionally, there is a column in the table OID, representing an object identifier (primary key in a relational model). ObjectType column contains the identifier of the class whose objects are stored in the form of table rows. This value by the application to create a class of object-oriented programming language and to load the attribute values is used.

In implementing Class Table Inheritance pattern have been created for the table Contragent for abstract class and table Worker, Company, Employee, Manager for the concrete classes. Instances of classes are physically stored in multiple database tables. A copy of the Manager class is stored in all tables.

ISBN: 978-1-61804-327-6

In implementing the Concrete Table Inheritance pattern is applicable for classes Address, CompanyAddress and EmployeeAddress, was created two tables: CompanyAddress and EmployeeAddress, because CompanyAddress class is abstract. All abstract class attributes stored in tables physically specific classes.

For an n-ary association Position create a separate table as well as for the binary association linking the Employee class and EmployeeAddress, for that created the table EmployeeEmployeeAddress, containing foreign keys.

Note that for the enumeration Telephone-Kind separate table is not created. An approach representations enumeration values as a bit mask and store it in the form of an integer value, where appropriate attributes are used. So the table has a column Telephone TelephoneKind, SQL-type is Integer.

After analyzing of the above it can be argued that shown in Fig. 5 implementation, created in a development environment SharpArchitect RAD Studio, fully consistent with the unified model for testing object-oriented application development tools, presented in Fig. 1.

V. CONCLUSION

Further development of the unified model is to test the feasibility of a variety of application development environments. In this alternative implementation is planned and using the approach presented by other authors dealing with similar scientific problems.

REFERENCES

[1] Erich Gamma, Richard Helm, Ralph Johnson, John Vlissides. Design Patterns: Elements of Reusable Object-Oriented Software, Addison-Wesley, USA, 1994, 395 pp.

[2] Oleynik P.P. A unified model for testing object-relational mapping tools // Object Systems – 2011: Proceedings of the Third International Theoretical and Practical Conference. Rostov-on-Don, Russia, 10-12 May, 2011. Edited by Pavel P. Oleynik. - 65-69 pp. (In Russian), http://objectsystems.ru/files/Object_Systems_2011_Proceedings.pdf

[3] Oleynik P.P. Test model for training in design of object-oriented databases // Object Systems – 2014: Proceedings of the Eighth International Theoretical and Practical Conference (Rostov-on-Don, 10-12 May, 2014) / Edited by Pavel P. Oleynik. – Russia, Rostov-on-Don: SI (b) SRSPU (NPI), 2014. – pp 86-89. (In Russian), http://objectsystems.ru/files/2014/Object_Systems_2014_Proceedings.pdf

[4] Oleynik P.P. The Elements of Development Environment for Information Systems Based on Metamodel of Object System // Business Informatics. 2013. №4(26). – pp. 69-76. (In Russian), http://bijournal.hse.ru/data/2014/01/16/1326593606/1BI%204(26)%202013.pdf

[5] Oleynik P.P., Computer program "The Unified Environment of Rapid Development of Corporate Information Systems SharpArchitect RAD Studio", the certificate on the state registration № 2013618212/ 04 september 2013 (In Russian).

Reducing Employee Turnover in Small Business

An Application of Employee Turnover Models

Iveta Ozolina-Ozola
Department of Innovation and Business Management
Riga Technical University
Riga, Latvia
iveta.ozolina-ozola@rtu.lv

***Abstract*—**The purpose of this paper is to investigate the role of human resource management practices for employee turnover reducing in small business organizations. This investigation is based on the analysis of employee turnover models. The results of content analysis show that job and organizational factors are frequently mentioned in employee turnover models, and among these factors the issues of job content, rewards, control and trust are highlighted. Taking into account the human resource management context of small business organizations, the appropriate practices are proposed for employee turnover reducing.

***Keywords*—**human resource management; small business; voluntary employee turnover; employee turnover model; content analysis

I. INTRODUCTION

Employee turnover reflects the employees voluntary or involuntary leaving the organization. As recent reviews of employee turnover studies show [1-3], these studies mostly are focused on voluntary employee turnover, i.e. on movement initiated by the employee. The involuntary turnover as layoffs and discharges are initiated by the employer, therefore this process is more controllable and forecastable. There are many empirical studies confirmed the negative effect of the high rate of voluntary employee turnover on organization's productivity and profitability [4, 5], workforce performance [6], instrumental communication and behavioral commitment [7], social capital [8] and organizational capital [9]. Since high rate of voluntary employee turnover has negative impact to organization's economic and social processes, the issues of employee turnover management are important for both researchers and practitioners in human resource management, organizational behavior and labor economics field.

The strategy, practices and climate of organization's human resource management have the essential impact on employee turnover. In contemporary scientific studies, the effect of human resource management on employee turnover is investigated by various ways:

- measuring the correlations between organizational performance, including the variables characterized the employee turnover, and human resource management effectiveness or human resource management systems, expressed as a set of specific practices or techniques, e.g. [10-12];
- similarly to previously mentioned but focusing directly to the connections of employee turnover with human resource management systems, e.g. [13];
- measuring the employee turnover correlations with the specific human resource management practice, e.g. [14], or with the specific technique of some practice, e.g. [15];
- presenting the results of review or meta-analysis of studies on employee turnover, e.g. [2];
- developing the employee turnover models, e.g. [16];
- discovering the non-managerial causes of employee turnover that could be avoided by the human resource management measures, e.g. [17].

Considering the abundance of employee turnover studies, nevertheless there is a lack of the narrow review of human resource management function in employee turnover models. Such review could be useful to create a framework for analyzing of human resource management impact on employee turnover in systematic way. In this regard, the following research question is posed: what practices of human resource management are viewed as essential in employee turnover models?

The next focus of author's interest is the features of human resource management to influence employee turnover in small business organizations. This theme is not sufficiently explored in academic literature. For example, the recent reviews of studies on employee turnover [2, 3] do not present details concerning the employee turnover factors dependence from the size of organization. However, it is obvious that some human resource management practices recommended as effective tools to reduce or prevent high rate of employee turnover are unusual or hardly implemented in small business. Therefore, the second research question is what practices of human resource management are applicable to influence employee turnover in small business organizations?

Consequently, the purpose of paper is to investigate the role of human resource management practices for employee turnover reducing in small business organizations. The principal tasks of this study were:

1) to find, review and select for further exploration employee turnover models;

ISBN: 978-1-61804-327-6

2) to explore selected models and extract the bands of employee turnover's factors;

3) to identify the role of human resource management in employee turnover models;

4) based on the previous analysis, to propose the certain human resource management practices for employee turnover reducing in small business organizations.

The deliberate review of employee turnover models in the context mentioned above could contributes to better theoretical understanding of human resource management impact on employee turnover in small organizations.

II. Previous research

Since the early 20th century, a lot of studies had been conducted on employee turnover in psychology, sociology, management and economics. The recent reviews of studies on employee turnover are presented by Allen et al. [1], Steel and Lounsbury [3], and Holtom et al. [2]. According to Steel and Lounsbury, there are at least 24 conceptual models of employee turnover including variations and refinements. Despite of the existence of such models variety, a few attempts were realized to compare these models in detail. The well known academic works devoted to the synthesis of employee turnover models are performed by Maertz and Campion [18], Steel and Lounsbury [3].

Currently, there are not certain reports exploring the role of human resource management in employee turnover models. Implicitly this issue is highlighted in Steel and Lounsbury's review [3], where conceptual analysis of 16 employee turnover models is presented. According to authors' opinion, all analyzed models are constructed around the same three core turnover mechanisms – attitudinal variables (job satisfaction and organizational commitment), job-search mechanisms (whether perceptual or market-based), and turnover intentions or stay-quit intentions. The similarities were also observed in the secondary mechanisms, or additional dimensions, of employee turnover models. Respectively, the authors sorted these dimensions into five broad topic areas:

- personal factors (personal traits, values, age, tenure, skill, training, professionalism, family responsibilities);
- job and organizational factors (job-related perceptions, expectations of present job, job rewards, costs of quitting, job stress, role conflict, organizational size);
- mechanisms of external and internal change (desirability of movement, expectations of future job, efforts to change situation, intra-organizational transfer possibility, promotion possibility, demotion possibility, alternative forms of withdrawal);
- consequences of quitting or staying (non-work consequences, job performance);
- decision process mechanisms (system shocks, thoughts of quitting).

The closest topic area mentioned above to human resource management practices is "job and organizational factors". As it is seen, Steel and Lounsbury define job and organizational factors as a set of objective and subjective variables. Nevertheless, in their "integrative map of the turnover-theory concept map" the job and organizational factors are placed with strict dependence on personal factors and attitudinal variables. Such explanation of job and organizational factors function is discussable because of some personnel characteristics, e.g. skill, training, job satisfaction etc., can be controlled or influenced by purposive human resource management. In other words, there is an interaction between job or organizational factors and employee's characteristics.

Among the objective variables of job and organizational factors, was organizational size. However, this variable was observed only in two employee turnover models [3]. It could be explained not so much by ignorance of organizational size effect, but rather by anticipation of this effect in other, more proximal to employee turnover process, variables.

The concepts of human resource management are also noted in topic area "mechanisms of external and internal change" such as possibility of intra-organizational transfer, promotion or demotion. In overall, Steel and Lounsbury's description of employee turnover secondary mechanisms is quit general, and there is a necessity to conduct an in-depth analysis for accurate understanding of the role of human resource management in employee turnover models.

III. Methods

The study was formed of three stages. In the first stage, the models of employee turnover were found, reviewed and selected for further exploration. As main sources of information on employee turnover models were peer-reviewed articles, including its references, obtained from electronic databases as *Scopus, EBSCOhost Web, Google Scholar, ScienceDirect, Springer Link*. For searching of relevant articles the keywords "employee turnover", "personnel turnover", "labor turnover", "job turnover" were applied in electronic databases. The models for further exploration were selected by four criteria:

1) comprehension (does the conceptual description of employee turnover comprehends the whole process of employee turnover?);

2) level of analysis (does the conceptual description of employee turnover concentrates on the individual level of analysis?);

3) originality (does the conceptual description of employee turnover contains the principal novelty?);

4) significance for development of employee turnover research.

The articles proposed some model of employee turnover were analyzed primary in qualitative mode. To be sure about significance of each model, the ranking of corresponding articles was additionally made on the base of articles' citation data acquired from *Scopus* and *Google Scholar*.

ISBN: 978-1-61804-327-6

In the second stage of study, applying content analysis method, the variables of employee turnover models were grouped into thematic bands, or categories, and the frequencies (*f*) of these categories were calculated. The scale of category occurrence was dichotomous assuming "1", if description of employee turnover model contained at least one code of category, and "0", if description of model has not any relevant code. The codes of categories were composed without preliminary list, but concurrently reviewing descriptions of employee turnover models. The content analysis was also applied to determine presence and frequency of the subcategories of job and organizational factors. For this procedure the created list of codes was exploited. The technique of subcategories' frequencies calculation was similar to previously mentioned.

In the third stage of study, revealed job and organizational factors were linked with appropriate human resource management practices in small business organizations.

IV. Results and Discussions

Based on the search results of publications, at least 50 articles, proposed some theory or model of employee turnover, were found. Evaluating these articles by four criteria, stated above, 16 employee turnover models were selected. It should be noted that the list of selected models slightly differs from Steel and Lounsbury's list. All selected models describe voluntary turnover process, except Jackofsky's employee turnover model [19], because it integrates involuntary turnover as well.

The analysis of employee turnover models shown that variables of these models, excluding "employee turnover", or "quitting" and similar terms, can be grouped into ten thematic bands, or categories (Tab. I). Commonly, employee turnover is explained by employee attitude towards job (by job satisfaction or organizational commitment or by both these attitudes). In the majority of models, these attitudes towards work are linked with formation of employee intention to stay or leave. The employee attitudes towards job, in turn, are often explained as consequences of other individual factors (demographic, professional, socio-economic and psychological factors) interaction with job and organizational factors. The role of external alternatives is also accepted in employee turnover models, in which employee perceived opportunities are more emphasized than such objective variables of labor market as unemployment rate.

The content analysis of employee turnover models revealed that variables related to job and organizational factors are included as often as the employee job satisfaction (Tab. I). Both categories are the most frequent categories presented in employee turnover models. The least presented category is employee's demographic factors. It could be explained by assumption that impact of these factors on individual decision to leave or stay is distal. For example, the more proximal factors to such decision is employee attitude towards job or work situation than employee age [23].

TABLE I. Frequencies of Categories in Employee Turnover Models

Employee turnover models / **Categories**	Organizational equilibrium theory [20]	Met expectations model [21]	Intermediate linkages model [22]	Structural model [23]	Expanded process model [24]	Multidisciplinary model [25]	Multi-route model [16]	Cusp-catastrophe model [26]	Investment model [27]	Integrated process model [19]	Labor economic model [28]	Alternative linkage model [29]	Unfolding model [30]	Job embeddedness theory [31]	Evolutionary search model [32]	Forces model [33]	Frequency	Relative frequency
Intention to leave or stay	1	1	1	1	1	0	1	0	0	1	1	1	1	1	1	1	13	0.81
Job satisfaction	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	15	0.94
Organizational commitment	0	0	0	1	0	1	1	1	1	0	0	1	1	1	1	1	10	0.63
Job search activities	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	14	0.88
Labor market	1	0	0	0	1	1	1	0	0	1	1	1	1	0	0	0	8	0.50
Job and organizational factors	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	15	0.94
Employee's demographic factors	1	1	0	0	1	1	1	0	0	1	0	0	0	0	0	0	6	0.38
Employee's professional factors	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	0	13	0.81
Employee's social-economic factors	1	1	1	1	1	1	1	0	1	0	0	0	1	1	1	1	12	0.75
Employee's psychological factors	1	1	1	1	1	1	1	0	0	1	0	1	1	1	1	1	13	0.81

ISBN: 978-1-61804-327-6

Among job and organizational factors the pay, peer group interaction and supervision are the most pointed out in employee turnover models (Fig. 1). The size of organization is mentioned in four models, i.e. in 25% of cases.

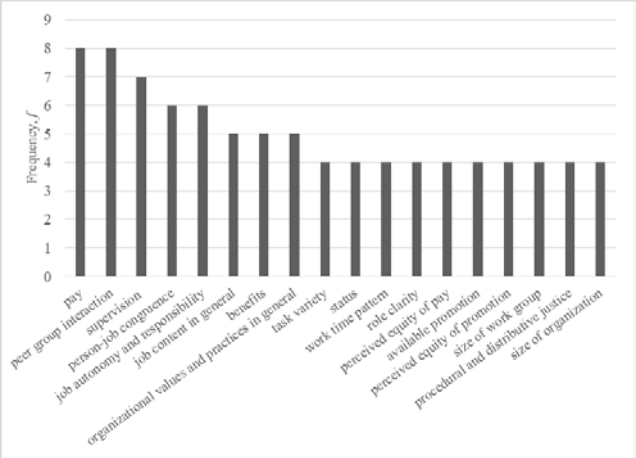


Fig. 1. The frequencies of job and organizational factors' subcategories (with *f* at least four)

The detected subcategories were consolidated into thematic groups, i.e. into higher level of subcategories:

1) job content (*f*=24): job content in general, person-job congruence, task variety, workload, status, job stress in general;
2) pay and benefits (*f*=17): pay, benefits, perceived equity of pay;
3) control and trust (*f*=17): supervision, job autonomy and responsibility, opportunities for participation, feedback and recognition;
4) work group (*f*=12): size of work group, peer group interaction;
5) organizational culture and climate (*f*=9): organizational values and practices in general, procedural and distributive justice;
6) promotion (*f*=8): available promotion, perceived equity of promotion;
7) job related information (*f*=7): available information about the job and organization, role clarity.

The frequencies of the rest thematic groups were lower: hardly controlled organizational variables (f=6), work time (f=4), employee expectation for keeping job (f=3), organizational critical events (f=1), and organizational reputation (f=1).

Summarizing the results of presented above thematic groups' analysis, it could be concluded that the greater focus is put on the adjustment of job content, rewards system and management-subordinate relationships. The improvement of peer-group interaction, the employee involvement in decision-making process, and adherence to principles of justice and fairness are also highlighted in employee turnover models.

The majority of job and organizational factors, described in employee turnover models, are realizable in small business organizations. However, there are some restrictions for ensuring the employee promotion or transfer, task variation and employment security. The restrictions covered these and other possibilities to reduce employee turnover could be determined not only and so much by small size of organization, but by insufficiency of professional expertise in human resource management. Usually, small organizations do not have human resource units or specialists, therefore the quality of human resource management is often depended on owner competencies.

TABLE II. EXAMPLES OF HUMAN RESOURCE MANAGEMENT PRACTICES FOR EMPLOYEE TURNOVER REDUCING IN SMALL BUSINESS ORGANIZATIONS

Subcategories	Examples of human resource management practices
Job content	Task assignment considering employee motivation and abilities; preventing high-level stress by workload optimization.
Pay and benefits	Clear and fair system of reward; performance or tenure related reward contingencies; competitive level of wage; fringe benefits.
Control and trust	Control practices confirming to employee's independence level; supportive supervisory style; recognition expressing; employee involvement in decision-making.
Work group	Regular collective events; constructive conflict solving; feedback about performance of firm and unit.
Organizational culture and climate	Suitable candidates selection; distributive and procedural justice keeping; clear communication; work conditions improvement.
Promotion	Employee assignment to higher position or status (if it is possible).
Job related information	Realistic job preview; new employee induction; training; clear and timely feedback on performance; regular meetings.
Work time	Holidays and vacation scheduling with respect to employee wishes; flexible work schedule.
Employee expectation for keeping job	Clear and fair rules of sanctions and dismissals; clear and timely informing about essential changes.
Organizational critical events	Clear and timely informing about organizational changes in future; employee involvement in decision-making determining their future employment.
Organizational reputation	Fair employment practices.

As results of content analysis show, the job and organizational factors of employee turnover are formulated in quit general terms that do not communicate to employer what specific practices or tools is effective to manage employee turnover. Therefore, it is necessary to operationalize these factors into specific human resource management practices (Tab. II).

To reduce employee turnover in effective way, the bundles of human resource management practices should be selected. Besides objective activities, initiating by owner or manager to reduce employee turnover, employee perception of these activities is crucial. For that reason, owners and managers should pay attention to communication with subordinates receiving their opinions and forming their understandings.

V. Conclusions

The high rate of employee turnover is costly for all organizations. The resources and methods for employee turnover reducing in small organizations differ from those in larger organizations. As critical factor in small organization's employee turnover processes is owner competencies and attitude towards subordinates.

There are few investigations on small organization's employee turnover management in academic literature. The analysis of employee turnover models allows to provide a general comprehension of human resource management impact on employee turnover in small organizations.

Factors described in employee turnover models could be grouped into ten thematic bands: employee intention to leave, job satisfaction, organizational commitment, job search activities, labor market, job and organizational factors, employee's demographic, professional, socio-economical and psychological factors. As results of content analysis show, job and organizational factors are one of the most frequent category presented in employee turnover models. These factors are often mentioned as determinants of job satisfaction.

Content analysis also revealed that among job and organizational factors the issues of job content, rewards, control and trust are highlighted. Operationalization of these factors into specific practices of human resource management could contributes to development of employee turnover theory and its practical application.

References

[1] D.G. Allen, J.I. Hancock, and J.M. Vardaman, "Analytical mindsets in turnover research," Journal of Organizational Behavior, vol. 35, pp. 61-86, 2014.

[2] B.C. Holtom, T.R. Mitchell, T.W. Lee, and M.B. Eberly, "Turnover and retention research: A glance at the past, a closer review of the present, and a venture into the future," The Academy of Management Annals, vol. 2, pp. 231-271, 2008.

[3] R.P. Steel and J.W. Lounsbury, "Turnover process models: Review and synthesis of a conceptual literature," Human Resource Management Review, vol. 19, pp. 271–282, 2009.

[4] S. Brown, G. Garino, and C. Martin, "Firm performance and labour turnover: Evidence from the 2004 workplace employee relations survey," Economic Modelling, vol. 26, pp. 689-695, 2009.

[5] J.C. McElroy, P.C. Morrow, and S.N. Rude, "Turnover and organizational performance: A comparative analysis of the effects of voluntary, involuntary, and reduction-in-force turnover," Journal of Applied Psychology, vol. 86, pp. 1294-1299, 2001.

[6] J.D. Shaw, N. Gupta, and J.E. Delery, "Alternative conceptualizations of the relationship between voluntary turnover and organizational performance," Academy of Management Journal, vol. 48, pp. 50-68, 2005.

[7] C.W. Mueller and J.L. Price, "Some consequences of turnover: A work unit analysis," Human Relations, vol. 42, pp. 389-402, 1989.

[8] J.D. Shaw, M.K. Duffy, J.L. Johnson, and D.E. Lockhart, "Turnover, social capital losses, and performance," Academy of Management Journal, vol. 48, pp. 594-606, 2005.

[9] R. Eckardt, B.C. Skaggs, and M. Youndt, "Turnover and knowledge loss: An examination of the differential impact of production manager and worker turnover in service and manufacturing firms," Journal of Management Studies, vol. 51, pp. 1025-1057, 2014.

[10] J.B. Arthur, "Effects of human resource systems on manufacturing performance and turnover," Academy of Management Journal, vol. 37, pp. 670-687, 1994.

[11] M.A. Huselid, "The impact of human resource management practices on turnover, productivity, and corporate financial performance," Academy of Management Journal, vol. 38, pp. 635-672, 1995.

[12] O.C. Richard and N.B. Johnson, "Strategic human resource management effectiveness and firm performance," International Journal of Human Resource Management, vol. 12, pp. 299-310, 2001.

[13] T.M. Gardner, P.M. Wright, and L.M. Moynihan, "The impact of motivation, empowerment, and skill-enhancing practices on aggregate voluntary turnover: The mediating effect of collective affective commitment," Personnel Psychology, vol. 64, pp. 315-350, 2011.

[14] M.R. Barrick and R.D. Zimmerman, R. D. "Hiring for retention and performance," Human Resource Management, vol. 48, pp. 183-206, 2009.

[15] C.R. Williams, "Reward contingency, unemployment, and functional turnover," Human Resource Management Review, vol. 9, pp. 549-576, 2000.

[16] R.M. Steers, R.T. Mowday, and L.W. Porter, "Employee turnover and post decision accommodation processes," Graduate School Of Management and Business, University of Oregon, Eugene, Technical Report No. 22, 1979.

[17] R.D. Zimmerman, "Understanding the impact of personality traits on individuals'turnover decisions: A meta-analytic path model," Personnel psychology, vol. 61, pp. 309-348, 2008.

[18] C.P. Maertz and M.A. Campion, "25 years of voluntary turnover research: A review and critique," International review of industrial and organizational psychology, vol. 13, pp. 49–81, 1998.

[19] E.F. Jackofsky, "Turnover and job performance: An integrated process model," Academy of Management Review, vol. 9, pp. 74–83, 1984.

[20] J.G. March and H.A. Simon, Organizations. New York: John Wiley, 1958.

[21] L.W. Porter and R.M. Steers, "Organizational, work, and personal factors in employee turnover and absenteeism," Psychological Bulletin, vol. 80, pp. 151-176, 1973.

[22] W.H. Mobley, "Intermediate linkages in the relationship between job satisfaction and employee turnover," Journal of Applied Psychology, vol. 62, pp. 237–240, 1977.

[23] J.L. Price, The study of turnover. Ames, IA: Iowa State University Press, 1977.

[24] W.H. Mobley, R.W. Griffeth, H.H. Hand, and B.M. Meglino, "Review and conceptual analysis of the employee turnover process," Psychological Bulletin, vol. 86, pp. 493–522, 1979.

[25] P.M. Muchinsky and P.C. Morrow, "A multidisciplinary model of voluntary employee turnover," Journal of Vocational Behavior, vol. 17, pp. 263–290, 1980.

[26] J.E. Sheridan and M.A. Abelson, "Cusp catastrophe model of employee turnover," Academy of Management Journal, vol. 26, pp. 418–436, 1983.

[27] C.E. Rusbult and D. Farrell, "A longitudinal test of the investment model: The impact on job satisfaction, job commitment, and turnover of variations in rewards, costs, alternatives, and investments," Journal of Applied Psychology, vol. 68, pp. 429–438, 1983.

[28] C.L. Hulin, M. Roznowski, and D. Hachiya, D. "Alternative opportunities and withdrawal decisions: Empirical and theoretical discrepancies and an integration," Psychological Bulletin, vol. 97, pp. 233–250, 1985.

[29] P.W. Hom and R.W. Griffeth, "Structural equations modelling test of a turnover theory: Cross-sectional and longitudinal analyses," Journal of Applied Psychology, vol. 76, pp. 350–366, 1991.

[30] T.W. Lee and T.R. Mitchell, "An alternative approach: The unfolding model of voluntary employee turnover," Academy of Management Review, vol. 19, pp. 51–89, 1994.

[31] T.R. Mitchell, B.C. Holtom, T.W. Lee, C.J. Sablynski, and M. Erez, "Why people stay: Using job embeddedness to predict voluntary turnover," Academy of Management Journal, vol. 44, pp. 1102–1121, 2001.

[32] R.P. Steel, "Turnover theory at the empirical interface: Problems of fit and function," Academy of Management Review, vol. 27, pp. 346–360, 2002.

[33] C.P. Maertz and M.A. Campion, "Profiles in quitting: Integrating process and content turnover theory," Academy of Management Journal, vol. 47, pp. 566–582, 2004.

Measuring the industrial processes performance by simulation

Florina-Cristina Filip*
Department of Engineering and Industrial Management
Transilvania University of Brasov, Mihai Viteazul 5,
Brasov, 500174, Romania
E-mail: florina-cristina.filip@unitbv.ro
*Corresponding author

Vladimir Mărăscu-Klein
Department of Engineering and Industrial Management
Transilvania University of Brasov, Mihai Viteazul 5,
Brasov, 500174, Romania

***Abstract*—** This paper describe the development of a reliable product that enable to measuring the industrial processes performance by framework management methods. It briefly describes the Witness simulation system and three Excel tools built and used to modeling and simulating a complex assembly line known as AML (Assembly Main Line). Also, it describes typical problems that call for use of modeling and simulation process and then explains how the relevant data are collected. Validation of the AML model is carried out through a significant case study that describes and analyzes the impact and effect of rejection rate (measured by first time through, FTT) on throughput of the assembly line, and the use of simulation as a method and supporting tool used to improve the industrial processes performance. By validating the AML model, it's confirming that it was properly built, provided that the input data are correct and correspond with reality.

***Keywords*—**assembly line; overall throughput; performance evaluation; rejection rate

I. INTRODUCTION

Many papers in the specialty literature highlights the fact that, to succeed with high-quality products in today's highly competitive world markets, companies have a great demand to highly adaptive management and plays a significant role in improving their industrial processes performance in order to fulfill the customers' needs and sustain their competitive advantage [1, 2]. Due to the customer's needs and expectations, the modern companies are facing a higher pressure to identify and implement management methods to improve their responsiveness in 21st century market dynamics [3]. One of the most important advances of implemented management methods in industrial processes have opened up the research possibility of optimizing the modeling and simulation process [4]. The global trends on software development allow processes to be developed in a large area which require several changes and reliable software's [5-7].

Due to the emergence and improvement of electronic computers, simulation is now widely used in scientific research and design. The simulation process is a technique that involves the use of mathematical and logical models that describe the behavior of a real system over a period of time and requires the use of electronic computer. Modeling complex systems such as industrial process is a difficult task and as such, requires a simple and effective way to facilitate the simulation process [6]. Simulation has an important role in guiding the processes understanding and development without requiring costly manufacturing trials [8] and accordingly, the computer-aided techniques are found mainly in modeling and simulation process [9].

To characterize the impact of changing the values and parameters to measuring the industrial processes performance, simulation has been widely used as a support decision in modeling and analysis the processes. Thus, when the direct measurement of parameters isn't easy and trusted, modeling and simulation process provides efficient methods of observing the behavior over time of a process [6], [10]. Simulation is a rewarding tool for industrial processes performance, especially in the system design and launching phases and a model is a conceptual or mathematical representation of a system or process [11]. Mainly, the modeling and simulation process includes specific techniques and methods of computer-aided applications [9].

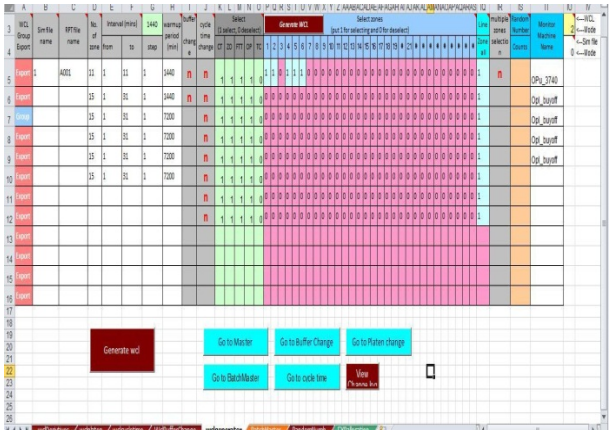
II. THE ASSEMBLY MAIN LINE (AML) MODEL AND SIMULATION TOOLS

An assembly line has different sub assembly lines or process flows, defining a process flow as any set of operations that feed another line that would stop the line from working, because of slow or varying cycles, breakdowns, quality issues etc. Simulation is one of the techniques available to study large and complex systems, is a collection of methods and applications designed to mimic the actual behavior of industrial processes. Today, there are many simulation tools available that can model all kinds of systems, regardless of their complexity [6].

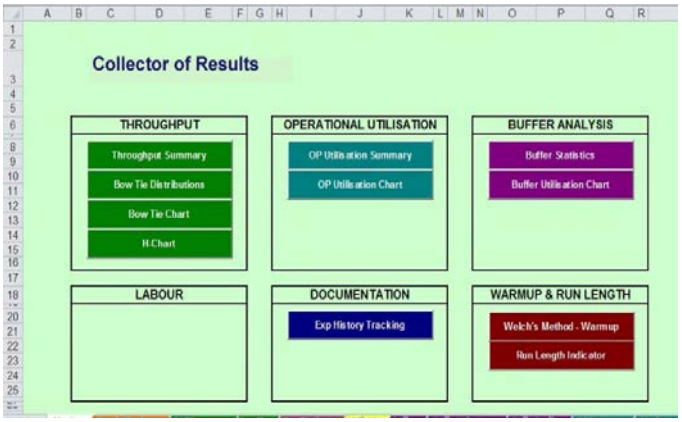
To build the AML model as a complex line composed by eleven production zones (areas) was used the Witness simulation system. Although Witness is quite efficient and allows modeling and simulation of complex production systems, however, in certain circumstances, it cannot cover their versatility. In this case, to simplify and ease the simulation experiments of the AML model was developed three Excel tools (Fig. 1), named as Database Centralization (DC), Orders Generator (OG) and Collector of Results (CR):

ISBN: 978-1-61804-327-6

a. The Database Centralization



b. The Orders Generator



c. The Collector of Results

Fig. 1. Tools used to simulate the AML Line [12]

- Database Centralization (DC) – contains all the needed data and worksheets with information on the functioning line, input and output data, areas description of that make up the assembly line, transport means used in the production process, processed parts, the check points, data on the production capacity of machines, their cycle time, areas of repairs or improve parts, data on quality rules and standards, storage areas, waiting times etc. (Fig.1.a);
- Orders Generator (OG) – allow the possibility to generate a number of „n" simulation experiments and sets up the experimental scenarios using the Witness command lines (Fig.1.b);
- Collector of Results (CR) – is responsible to display and analyze the simulation experiments results (Fig.1.c).

The AML model built and simulated in Witness system and a detail of its operation and components, are shown in Fig. 2. Studies and data presented are part of the PhD experimental research performed by the first author of this paper.

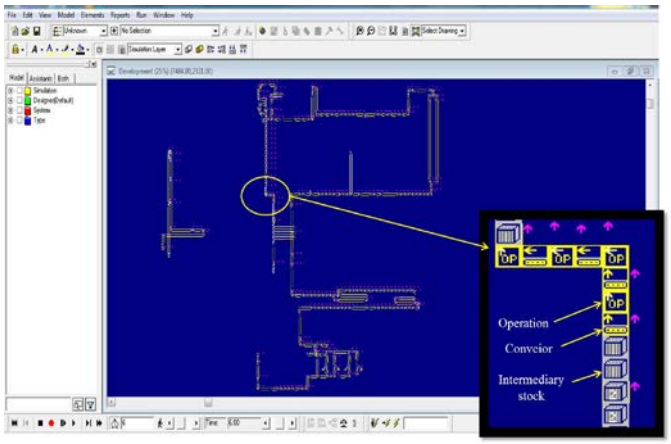


Fig. 2. The AML modeled in Witness

Connection between tools created in Excel and Witness simulation system is possible because, inside of Excel were created macros which automatically generated the „simulation codes" of assembly line (Fig. 3). These codes make the connection with a so-called „library modules" designed for Witness and this library is made up of elements and existing equipment in an automated assembly line (manual or

ISBN: 978-1-61804-327-6

automated workstations, output materials, data quality, conveyors, inventory etc.).

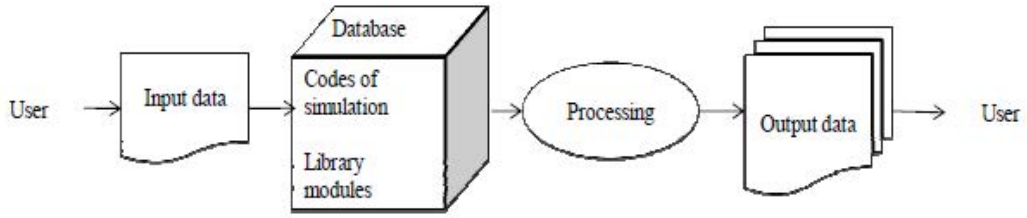


Fig. 3. Connection between DC and Witness [12]

III. Validation of the AML Model

The issue of testing and validation of a model is an essential element of all credible modeling and simulation processes [13-15]. The AML model built using the Excel tools (DC, OG and CR) and Witness simulation system, is accurately represented provided that the input data from Excel are accurate and the conversion is carried out automatically with help of the verified modules. Any building error can also be visually identified from the way that the elements are constructed in Witness.

The main verification and validation process, therefore, is the running of the AML model for a specified period of time (plus the warming up time) and checking that the simulation throughput is consistent with the expected cyclic throughput. This is done by first disabling all causes of losses including breakdowns, over cycles and first time through (FTT) prior to running. The simulation is then run and the output in JPH compared with the cyclic JPH which is calculated as 60 divided by the longest cycle time (in minutes) of the AML line operations. If the two values are the same, then it can safely be said that there are no logical errors in the model. If the result between theoretical JPH and real JPH is the same, the model is valid. Otherwise, the model is check. In this case, the theoretical JPH is equal to real JPH.

Once the model has been verified, it is reset to include all the relevant data (including the losses) and is it rerun. The simulation output is then compared with the real data from the manufacturing floor. The acceptable variation to validate the model is generally 1 JPH or 5% on the average JPH.

The simulation outputs are discussed with manufacturing personnel and experimenting decisions set out. The fundamental consideration is how the causes of the bottleneck could be alleviated. The normalized average JPH profile analysis is valid only for a model with at least 100 data points. By data points is understood the provided data by a single transfer (change) in one day.

IV. The AML Simulation and Rejection Rate Analysis

For the AML model a lot of experiments have been done on production losses caused by machine breakdown, overcome, lack of balancing, but not much attention was given to the quality issues (rejection rate) that directly affect the first time through (FTT) and the job per hour (JPH). The effect of FTT (rejection rate) on individual, zonal and overall throughput is more obvious when rework is involved and the rework items have different entry points, as in the assembly line this case study is based on. The experiments consist in performing a number of „n" different simulations of the AML model by making several changes on the values of product quality, in order to determine the JPH percentage in order to achieve the rejects rate analysis.

For the first experiment, in the Database Centralization (DC) were made changes on the rejection percentage of operations from zones 1, 2 and 4. After saving the changes and running the model in Witness, it opens the Orders Generator (OG) program and it activates with „1" the input factor of the quality problems FTT (first time trough or quality). After conducting the first simulation experiments of AML model, it has found that the most significant differences on the rejection rate, is recorded at the operation 1000 (OP1000) from zone 4 (Z4), the largest cycle time is recorded to the operation 860 (OP860) from zone 4 (Z4) and the bottleneck is recorded also in zone 4 (Z4). Based on these first results, was insisted on making changes in zones 1, 2 and 4, zone 3 is outside of the line.

Another experiments consisted in making changes on the rejection rate of OP1000 and OP860 from Z4, by activating in OG all the input factors: CT, FTT, ZO, OP and TC. Based on the simulation results, was done another experiment where were made changes on the rejection rate of OP1000 only, by activating again all the input factors: CT, FTT, ZO, OP and TC.

Another set of simulation experiments was consisted in making changes on the rejection rate of OP1000 and OP860 from Z4. After making changes in DC and running the AML model in Witness simulation system, is activated in OG the CT only. The last simulation experiments consist in changing the rejection rate of OP1000 and OP860 from Z4 and activation in OG the input factors CT and FTT.

As the first step to analyses the effect of FTT, the AML model has already been verified and validated. Variation of the FTT values of OP1000 from the existing 2.68% to as low as 0.5% and as high as 4.0% has the results shown in Fig. 4. The throughput improved when the rejection rate was reduced from 2.68% to 0.5%, and when the rejection rate deteriorates to 4.0% the throughput dropped.

Experiments were also carried out with changes on the FTT of OP860, combinations of OP860 and OP1000, and changes in zone 3 and zone 5. The results did not significantly improve the overall throughput when compared to the output of FTT changes on OP1000. For instance, a combined reduction on FTT of OP860 from 1.3% to 0.5% and that of

OP1000 from 2.68% to 1.5% resulted in an overall throughput of 106 JPH. This is the same as the throughput obtained by changing the FTT of OP1000 only (Fig. 4).

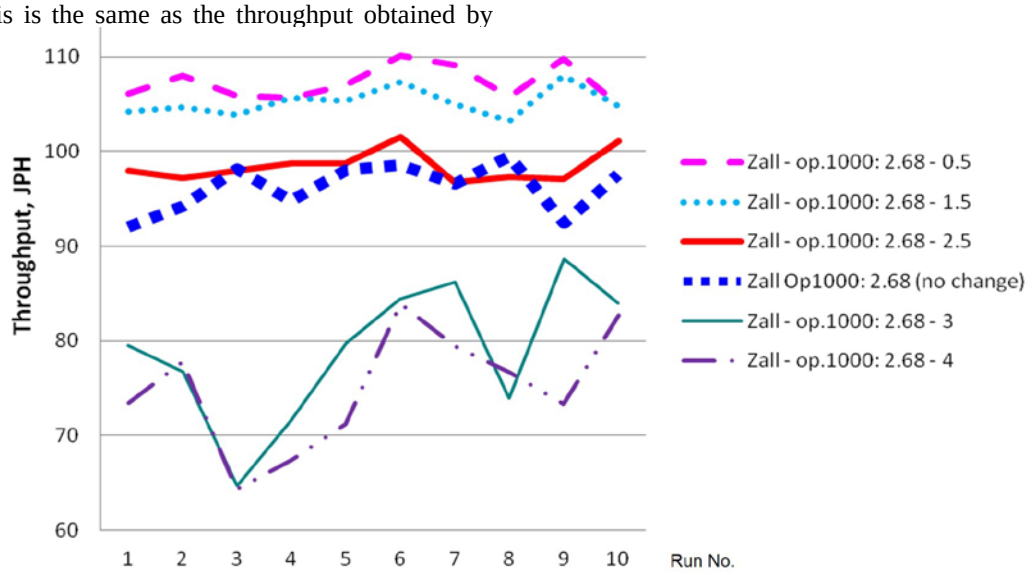


Fig. 4. Effect of the rejection rate on the overall throughput [12]

V. Conclusion

To describing and observing the behavior of a real assembly line in a certain period of time, is resorted to computer-aided simulation which allows identifying, analyzing and implementing the management methods to industrial processes performance. The main objective followed by modeling an assembly line submissive to the simulation experiments, is to actually restore and reproduce the essential elements and basic activities of its. A model subjected to the simulation process is able to cope with a system containing more than one hundred sequential operations and adopt the philosophy of continuous improvement and change after each improved version. These models help managers in making decisions on improving and extending the existing lines or to build new assembly lines.

In any simulation process, the quality of provided data has a paramount importance. For the AML simulation process, data to model building are mainly gathered from the assembly line layout plan (layout drawing), the work standard, and the real-time computer monitoring system. The layout gives a scaled drawing of the assembly plant layout and includes information such as loading/ unloading area, operation numbers, relative sizes of stations, repair areas, team sections and, in some cases, line cycle times.

Modeling and simulation of the AML assembly lines was carried out by using the Database Centralization (DC) which is an Excel based on front-end user interface to the Witness simulation system, outputs are generated using the Orders Generator (OG), and the results analysis is carried out by using the Collector of Results (CR) tool. The AML model building was also enhanced by the use of specialized sub modules (Witness modules). The DC is a component-based simulation system that enables engineers and non-simulation experts to develop models easily and quickly. It is designed to make model construction of assembly lines easier, with an interface that is readily understood by end users. The OG is used to setup experimental scenarios, run the model based on the scenarios and generate results. The CR tool then collects the simulation outputs and generated tables and charts, as appropriate.

The simulation experiments of the AML model led to following results:

- Development the Excel tools used to simulation process of the AML assembly line;
- Verification that the AML model is correctly built and corresponds with reality;
- Generate the simulation experiments of AML model;
- Identify the bottleneck and the largest cycle time of assembly line;
- Change the rejection rate (FTT) in certain zones of AML model;
- Analysis the experimental results of AML model simulation in order to identify and implement management methods to increasing the overall throughput (JPH) of assembly line;
- Determination and selection of the most relevant reported data of AML model simulation experiments;
- Analyzing of the data reports regarding the rejection rate (FTT) on the overall throughput (JPH);
- Validation of the AML model.

The results from the study performed in this paper, have led to identification and analysis of the rejection rate influence, also known as first time through (FTT) on the throughput rate, termed as jobs per hour (JPH) of assembled parts. To investigate the effect of rejection rate on throughput,

ISBN: 978-1-61804-327-6

the bottleneck zone (zone 4) and the zones feeding to and receiving from the bottleneck zone were identified and experiments carried out by varying the FTT values of selected operations. The OP1000 and OP860 were identified as those with the highest and the next highest rejection rates in Z4.

The impact of quality issues especially focused to FTT has been demonstrated to show that its effect on throughput could be significant. As complete as the simulation process is, it also has areas where improvement could be made such as in the debugging process where the user has to be familiar with the error messages to understand where the causes of the problem could be.

On the whole, the Witness simulation system and the Excel tools have proven to be effective in building the AML model and do the simulation experiments, and in collecting and analysis the relevant results of rejection rate on overall throughput. These tools are used not only in the simulation experiments of the AML model, but it can be used to simulate any real industrial process in order to measuring its performance. The component-based nature of the simulation system has made building of accurate assembly line models very easy even to end users with limited simulation expertise. This helps managers to concentrate more in analysis and evaluation of simulation results on implementing the framework management methods. It is hardly surprising therefore that modeling and simulation is used by managers to take the risk out of business change.

The future research directions imply analysis the possibilities of developing a informatics product that can facilitate the linking mechanism and access between the Excel tools and Witness simulation system, in order to simplify selection of managerial decisions on AML line performance. Analysis of a model is the key of any simulation study and the models are built in order to meet the experiments requirements. Accurate modeling but an improper analyze, can lead to incorrect decisions and may even be harmful, especially when financial investments are involved.

Acknowledgment

This paper is supported by the Sectoral Operational Programme Human Resources Development (SOP HRD), financed from the European Social Fund and by the Romanian Government under the project number POSDRU/159/1.5/S/134378.

References

[1] T. C. Poon, K.L. Choy, F.T.S. Chan, G.T.S. Ho, A. Gunasekaran, H.C. Lau, and H.K.H. Chow, A real-time warehouse operations planning system for small batch replenishment problems in production environment, Expert Systems with Applications, Vol. 38: 8524-8537, 2011.

[2] C. Schlick, R. Reuth, and H. Luczak, A Comparative Simulation Study of Work Processes in Autonomous Production Cells, Human Factors and Ergonomics in Manufacturing, Vol. 12(Issue 1): 31-54, 2002.

[3] S.C.L. Koh, MRP-Controlled Batch-Manufacturing Environment under Uncertainty, Journal of the Operational Research Society, Vol. 55(Issue 3): 219-232, March 2004.

[4] M.C. Fu, F.W. Glover, and J. April, Simulation Optimization: A Review, New Developments, and Applications, Proceedings of the Winter Simulation Conference, pp. 85-95, 2005.

[5] X. Bai, L. Huang, H. Zhang, and S. Koolmanojwong, Hybrid modeling and simulation for trustworthy software process management: a stakeholder-oriented approach, Journal of Software Maintenance and Evolution: Research and Practice, Published online in Wiley InterScience, May 2010.

[6] F.C. Filip, H. Mebrahtu, V. Mărăscu-Klein, and B. Deaky, Method for Optimization Production Systems by Computer Aided Modeling and Simulation, Journal of Electrical and Electronics Engineering, Vol. 5(Issue 1): 67-72, May 2012.

[7] S. Setamanit, W. Wakeland, and D. Raffo, Planning and improving global software development process using simulation, Proceedings of the 2006 International Workshop on Global Software Development for the Practitioner, pp. 8-14, 2006.

[8] F. Zhang, A. Mawardi, E. Santos Jr., R. Pitchumani, and L.E.K. Achenie, Examination of load-balancing methods to improve efficiency of a composite materials manufacturing process simulation under uncertainty using distributed computing, Future Generation Computer Systems, Vol. 22: 571-587, 2006.

[9] L. Iribarne, R. Ayala, and J.A. Torres, A DPS-based system modelling method for 3D-structures simulation in manufacturing processes, Simulation Modelling Practice and Theory, Vol. 17: 935-954, 2009.

[10] S. Cho, A distributed time-driven simulation method for enabling real-time manufacturing shop floor control, Computers & Industrial Engineering, Vol. 49: 572-590, 2005.

[11] D. Riseborough, N. Shiklomanov, B. Etzelmüller, S. Gruber, and S. Marchenko, Recent Advances in Permafrost Modeling, Permafrost and Periglac. Process, Vol. 19: 137-156, 2008.

[12] F.C. Filip, Contributions on implementing the efficient management methods specific for production systems, PhD thesis, October 2012.

[13] A.M. Law, and W.D. Kelton, Simulation Modelling and Analysis, Third Edition, McGraw Hill: International Series, 2000.

[14] S. Robinson, Simulation: The Practice of Model Development and Use, John Wiley & Sons Ltd. 2004.

[15] H. Mebrahtu, and J. Ladbrook, Simulation and Performance Improvement Process at Ford Powertrain Operations, The European Multidisciplinary Society for Modelling and Simulation Technology, Universite de LyonI, Lyon, France. 2008.

Proposal of knowledge discovery platform for big data processing in manufacturing

Lukas Spendla, Lukas Hrcka, Pavol Tanuska
Faculty of Materials Science and Technology
Slovak University of Technology
Trnava, Slovakia
lukas.spendla@stuba.sk, lukas.hrcka@stuba.sk, pavol.tanuska@stuba.sk

***Abstract*—In the proposed paper, we described the approach to building Data Lake based knowledge discovery platform. The proposal is focused on integrating Data Lake based storage, built on Hadoop framework and NoSQL systems, into traditional data warehouse discovery platform, preserving the well proven and robust data warehouse decision support and analytic tools. The proposed knowledge discovery platform processes data from all hierarchical control levels in manufacturing and can be used to address the main manufacturing issues in the knowledge discovery domain**

Keywords—knowledge discovery; data warehouse; data lake; hadoop; manufacturing; hierarchical control

I. INTRODUCTION

The current trend in manufacturing is marked by the large increase in amount of data, originating from the field level of hierarchical control. This increase is mainly due to implementation of new automation technologies and machines based on internet of things concept, a part of Industry 4.0, enabling direct communication with upper control levels.

Each parameter of manufacturing process is represented by a large amount of production data applicable in information or control systems at various levels. Despite the fact that most of manufacturing companies gather these data, they are not further used as information or knowledge in decision support process.

This was one of the reasons resulting in the urgent need for storing and processing large quantities of data and yet, it will be possible to work with them flexibly. These needs are reflected by current big data technologies based on NoSQL systems and Hadoop framework. However, integrating these new technologies into a company structure disrupts the well-established architecture based on data warehouses. This structure represents proven and robust solution from the company decision support point of view. Therefore, these new technologies must be integrated into manufacturing companies in a way allowing users to preserve the currently used solutions based on the data warehouse concept, while exploiting the advantages of the deployed NoSQL or Hadoop solution.

II. HIERARCHICAL CONTROL MODEL

Current information and control systems primarily employ hierarchical (pyramid) architecture integrated as a whole with elements of physical and logical distribution thus providing open and scalable solutions. Many of hierarchical control systems are built as multiprocessor control systems enabling both horizontal and vertical communication. Intelligent features arising from deploying sensors and actuators have been intensely utilised recently with direct hierarchical relations being transformed into network relations. Emerging tendencies such as connecting previously independent systems leading to new behaviour attributes are strongly reflected in current systems.[1] [2]

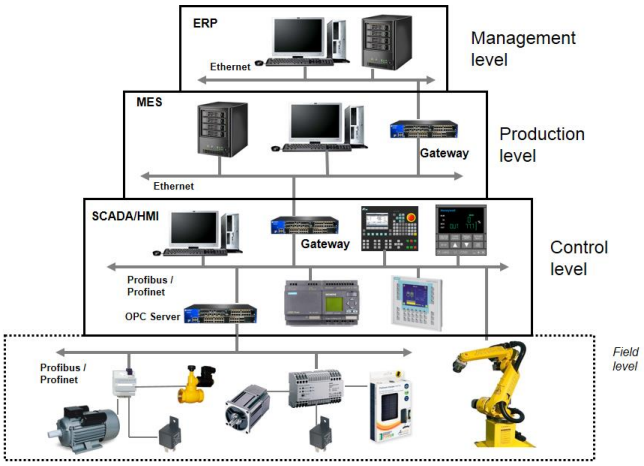


Fig. 1. Hierarchy of the industrial control system

Therefore, also the process control is nowadays being implemented deploying control systems with a hierarchical structure. The model of complex control process, so-called pyramid model, is shown in Fig. 1. [1] [2]

At all levels of the production process control model, large amount of data are produced, collected and stored often resulting in data redundancy. Still, the fact that different levels produce different types of data needs to be respected.

ISBN: 978-1-61804-327-6

A. Control Level

Technology (control) level is the lowest layer of the pyramid model of hierarchical process control and constitutes a basic interface with production. It consists of production lines, machines and equipment, which include integrated sensors and actuators, communicating using technology network with control computers, mainly with PLC (Programmable Logic Controller).[3] At this level, collecting and primary processing of technological parameters is carried out. Data are collected in real-time, with different sampling times, which results in collection of large amounts of data to be saved or archived for further implementation. The cyclic data collection traditionally used to collect data without transmitting the signal differences leads to redundancy. Collected data represents also a part of the noisy industrial environment and contains errors stemming technological information processing. Collected data are often noisy from manufacturing environment and contain errors primary processing of technological information. Removing these adverse conditions, filtering the required data and its subsequent validation are the tasks necessary to be carried out.

B. Production (Supervisory)Level

Supervisory level of pyramid model is a higher (intermediate) level of the hierarchical process control, which is alternatively called as SCADA / HMI (Supervisory Control and Data Acquisition / Human Machine Interface) level. It is used for primary collecting and integrating process data, for monitoring, visualization, evaluation and direct interference in the process. [3] At this level of control, system data are mostly stored in SCADA systems having an apparent purpose: processing alarm, monitoring mixing ratios, batch processes and data history, as well as archiving the operational variables. All other data stored at higher levels of pyramid model of hierarchical process control.

MES systems, typical representatives of information systems at the production level, are responsible for obtaining and collecting data from manufacturing. Obtained data are processed and in real-time stored in an aggregated form into data storage (mostly transactional SQL database). The data saved in the database in a structured form containing the current value of the variable, validity and timestamp (VTQ = Value / Time / Quality).

C. Management Level

Management level of pyramid model covers the previous levels. It consists of database resources for higher levels of control, management information system and tools for internet visualization. It is the level of planning and management. At this level, the data are archived and processed and long-term decisions for production are accepted. [3] At the management level of the pyramid model, data are not directly collected from manufacturing process, but are transferred in the transaction mode from information system of a real-time interface using the ERP Gateway. As the ERP does not operate continuously, continuous data transfer is carried out utilising ERP Gateway. Therefore, huge volumes of predominantly structured data arise in ERP systems.

III. Problems Indentification in Manufacturing

From the application area point of view, manufacturing process does not focus only on production itself, but extends and integrates data from all hierarchical control levels. For effective process control and management, not only production data are required, but also data including customers, resources, and suppliers' information from the upper hierarchy control levels are inevitable.

Manufacturing can therefore generate big amounts of data suitable for application of data mining and / or knowledge discovery process and might provide suitable means to deal with the problems arising in the field of production systems.

Existing approaches in manufacturing utilising data mining techniques can be divided into five main application areas [4]:

- Quality analysis of products to correlate output quality and system parameters, esp. machine settings, in order to identify causes for deteriorating product quality
- Failure analysis of production resources, esp. machines, to analyse causes of errors and prevent break downs in the future
- Maintenance analysis to enhance the availability of production resources, e. g., by optimized maintenance planning
- Production planning and scheduling analysis to improve planning quality, e. g., by a higher capacity utilisation of production resources
- Strategically planning and scheduling analysis to improve customer relationship and increase sales, e.g. by identification of customer behaviour

Each of these application areas covers multiple applied techniques and also different approaches form hierarchical model point of view. Therefore, it is impossible to identify a specific technique and an approach that needs to be applied to optimize or solve issues in selected application areas. It should be noted that all application areas span across multiple hierarchical control levels and therefore, it is impossible to assign them to a one specific level. Due to this fact, we identified the dependencies based on various research applied in the manufacturing area utilising data mining techniques.

Generally, the data mining approaches used in main application areas in manufacturing mostly utilise manual processing of specific data collection to analyse specific manufacturing process aspects in various manufacturing specific cases, e.g. machines, equipment, products, quality, etc. Most of the approaches can be integrated into real time support systems; however they mainly focus only on the approaches and methods themselves. [5][6][7]

IV. Current State of Knowledge Discovery in Manufacturing Area

At present, the application of data mining and knowledge discovery is very broad. However, according to recent studies [8][9], data mining is mostly employed in the fields like marketing, consumer analytics, finance, telecommunication,

insurance, health care etc. The usage of data mining in manufacturing is usually between 9 and 10 percent. Major part of this share is created by large international industries.

The weakness of the current form of manufacturing is often in the subjective perception of global production aims (profitability, production efficiency, plant productivity and product quality), frequent and often unforeseen variations in both manufacturing parameters and variables, the subjective decision making, and also in the vast amount of unstructured data provided by various information systems. [10]

The data mining and knowledge discovery process is usually based on data warehouse integrating all data required in this process. This concept of analytic environment is captures in Fig. 2.

Multiple systems operate at various hierarchical control levels, each using its own databases mostly independent from each other. [11] Very often there is no defined relationship between data in each system, e.g. manufactured product identifier has different numbering schema and order across control level data, SCADA, MES and ERP data. Therefore, it is necessary to integrate these data together to perform analytic reporting and knowledge discovery process.

In most large manufacturing companies, data warehouse is used to store the data from various company systems. Data integrated in the data warehouse serve as the basis for decision support, through the corresponding data mart or decision support tools [12]. Therefore, the ETL process transforming data into data warehouse for further use in business intelligence and analytic tools is extremely important.

The obtained data are accessed through data marts, created through ETL process from the data warehouse, providing organised view on the data from various business perspectives. [12] Data marts for various company specific aspects, like management, manufacturing, quality, etc. provide basis for decision making process.

It should be noted, that the data in data marts are not always integrated in the data warehouse itself. In other words, the data warehouse and data mart data can be separated. In order to obtain the complex view on company data for reporting and knowledge discovery, the data stored in data marts must supplement the company data stored in data warehouse.

The discovery platform must be set over all company data stored in data marts. If data warehouse does not integrate all data, discovery platform must be able to obtain and process them. In companies the discovery platform is mostly used for KPI based reporting and quality assurance [13]. The main advantages of data mining and knowledge discovery have still not been fully exploited.

A variety of knowledge discovery and analytic tools used in discovery platform is available. All major software tools provide connectors for relational databases and data warehouses. However most of the data operations must be handled by tools themselves, whether it is a standalone workstation or a client-server solution.

The main advantage of this approach is robustness and stability, due to the widespread deployment and long term real world experience in various enterprise areas. This factor is important due to the fact that the company KPIs, affecting the company management and business bottomline, are based on data from the data warehouse provided by the data marts.

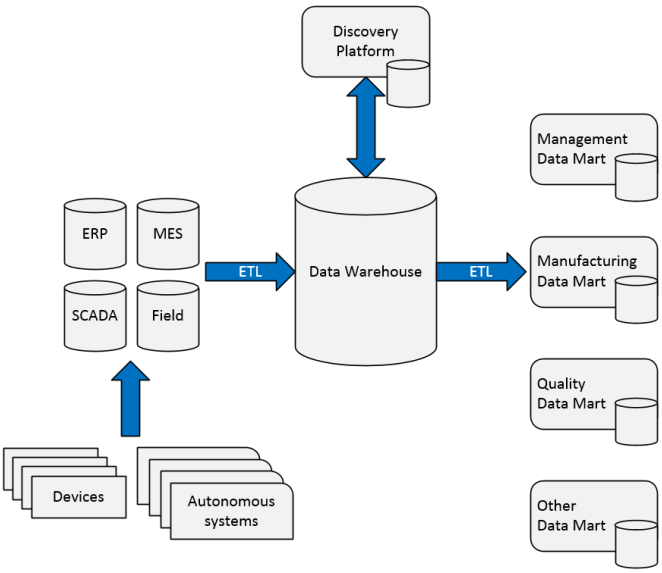


Fig. 2. Current state of knowledge discovery platform in manufacturing

The knowledge discovery process in company analytic environment is usually performed according to company methodology. This methodology can be specific for each individual company. However in recent years, more and more companies are starting to adopt the CRISP-DM methodology. Most companies however, don't adopt this methodology strictly. Due to this fact, the methodology is usually modified to suit the company needs. Since the knowledge discovery and data mining methodology is part of company know-how and not publicly accessible, it is impossible to generalise it as a whole.

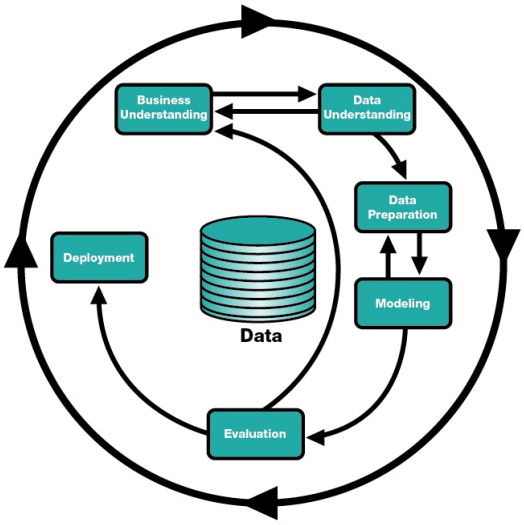


Fig. 3. Relationship between different phases of CRISP-DM [15]

ISBN: 978-1-61804-327-6

However, most of the companies preserve the continuity of the main phases of the CRISP-DM model, as shown in Fig. 3. Hence, the CRISP-DM methodology is, with a certain degree of abstraction, applicable in any manufacturing industry. [14]

Data mining is now used in many different areas in manufacturing engineering to extract knowledge for the use in predictive maintenance, fault detection, design, production, quality assurance, scheduling, and decision support systems. Data can be analysed to identify hidden patterns in the parameters controlling manufacturing processes or to determine and improve the quality of products. It clearly indicates data mining can be used in many different application areas of manufacturing. [16]

However, the manufacturing process brings a huge amount of data stored in databases containing enormous number of records. Every record has attributes needed to be explored to discover useful information and knowledge. All of this factors clearly demonstrates, that the choosing the right methods is crucial to successful discovery of knowledge. [17] Nowadays, there are a lot of types of methods, techniques and algorithms used for data mining process. Kdnuggets in [18] carried out a survey asking companies utilising data mining algorithms in their company. According to this research, the most used algorithms are: decision trees (rules), regression, clustering, statistics, visualization and time series.

Which methods/algorithms did you use for data analysis in 2011? [311 voters]	
Decision Trees/Rules (186)	59.8 %
Regression (180)	57.9 %
Clustering (163)	52.4 %
Statistics (descriptive) (149)	47.9 %
Visualization (119)	38.3 %
Time series/Sequence analysis (92)	29.6 %
Support Vector (SVM) (89)	28.6 %
Association rules (89)	28.6 %
Ensemble methods (88)	28.3 %
Text Mining (86)	27.7 %
Neural Nets (84)	27.0 %
Boosting (73)	23.5 %
Bayesian (68)	21.9 %
Bagging (63)	20.3 %
Factor Analysis (58)	18.7 %
Anomaly/Deviation detection (51)	16.4 %
Social Network Analysis (44)	14.2 %
Survival Analysis (29)	9.32 %
Genetic algorithms (29)	9.32 %
Uplift modeling (15)	4.82 %

Fig. 4. Survey of methods and algorithms usage in data analysis [18]

Many of the methods are exploitable in several areas, but it is very important to perform detailed analysis of the tasks to be solved, because methods are not universally applicable, but depend on the problem to be solved.

V. Knowledge Discovery Approach Proposal in Manufacturing Area

The common analytics environment at most big manufacturing companies includes a data warehouse, or collection of federated data marts, which house and integrate the data for knowledge discovery process. This includes various ranges of analysis function and business intelligence and analytics tools enabling decision support utilising ad hoc queries, dashboards and data mining.

Large manufacturing companies with large investments in their data warehouses have neither the resources, nor the will to replace the existing environment that works well and do what it was designed to. The majority of large companies utilise a coexistence strategy combining the best of data warehouse and analytics environment, with the new trends in big data solutions.

Many companies want to continue to rely on data warehouses for standard BI and analytics reporting, including sales reports, customer dashboards, risk history , etc. The coexistence strategy allows the companies to use data warehouse with its standard workload and storing historical data to establish robust traditional business intelligence and analytics tools. [19]

Despite the robustness of traditional business intelligence and analytics tools, semi-structured and unstructured data from the data collection process do not fit well into traditional data warehouses. Furthermore, data warehouses may not be able to handle the processing of frequently or even continually updated big data sets. As a result, organisations are looking for possibilities to collect, store and analyse big sets of data.

Newer class of technologies including the Hadoop framework and NoSQL systems are often deployed for this task. [20] In some cases, these technologies are being used as staging areas for data before they are transformed into a data warehouse, often in summarised form that is more suitable for relational structures. Big data solution vendors are increasingly pushing the concept of Hadoop Data Lake that is used as central repository for raw data streams present in the company. [21]

This coexistence approach, incorporating Data Lake as the central repository serves as a baseline for our knowledge discovery approach in manufacturing area, captured in Fig. 5.

The proposed knowledge discovery analytic environment is based on common data warehouse approach. The data warehouse integrates various data from heterogeneous systems across various hierarchical control levels. These heterogeneous data are extracted, transformed and integrated into a data warehouse using ETL process. This approach is mostly suitable for discontinuous and non-real-time data from higher hierarchical control levels.

The data marts, created from the data warehouse data, provide organised view on data from business unit perspective (like management, manufacturing, quality, etc.) and provide basis for decision making process in selected area. The data loaded into data marts needs to be extracted and transformed, to create the data structure suitable for further use.

ISBN: 978-1-61804-327-6

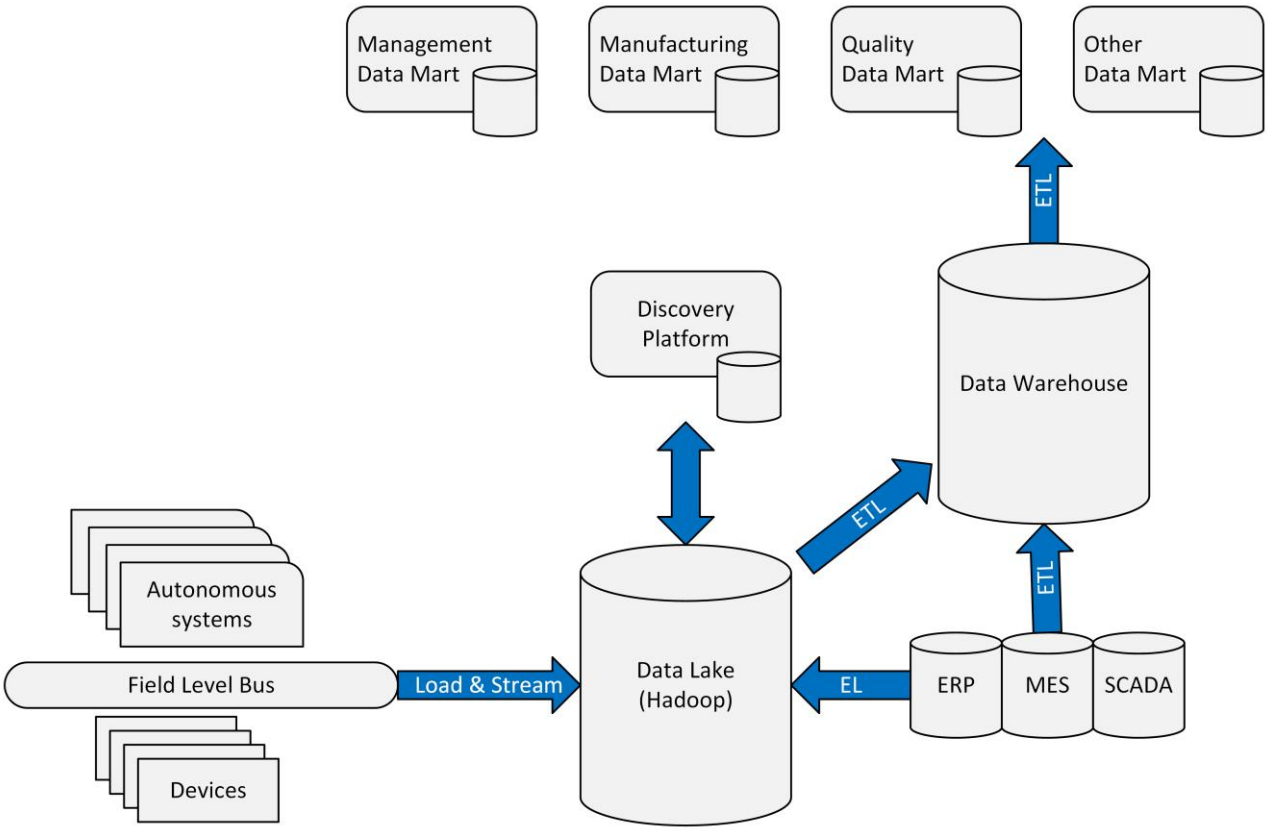


Fig. 5. Knowledge discovery analytic environment proposal

The Data Lake, based on Hadoop framework, provides central data storage for raw manufacturing data. The Data Lake extracts and loads data from heterogeneous database systems and stores them in a raw (original) form. Therefore the data does not have to be transformed to be stored in the Data Lake. In the Hadoop Data Lake cluster, subsets of the data can be analysed using batch query tools, stream processing software and SQL on Hadoop technologies that run interactively or using ad hoc queries in SQL.

The discovery platform in this environment is built on the data integrated in the Data Lake. Due to the use of Hadoop cluster, this environment provides higher performance when working in big data sets than the traditional data warehouse. Big advantage is also the availability of raw data from the manufacturing that cannot be easily stored in the data warehouse.

The offer of tools for discovery platform over Hadoop cluster is not very wide. Most of the standard knowledge discovery tools cannot connect to Hadoop cluster using SQL-on-Hadoop solutions. However, this way most of the data manipulation operations must be performed by the tool itself and not by the Hadoop cluster. In order to enable utilising the full potential of data manipulation performance of Hadoop, software manufacturers offer add-ons or software solutions able to perform selected sets of operations and algorithms directly in a Hadoop cluster. This approach is preferred, since the discovery platform must be able to process the big sets of collected data.

The ETL process transforming data from a Data Lake into a data warehouse is performed only for data not transferred into data warehouse directly. The main use of this particular ETL process is loading the manufacturing data from the field level stored in a Data Lake into a data warehouse. [20]

One of the biggest issues in obtaining manufacturing data is the way of collecting and processing data from the field level of hierarchical control. All these data serves as a basis for decision support at higher hierarchical control levels, the used field level data are usually aggregated into data more suitable for particular decision support task. Therefore the data suitability for business intelligence or analytic tools is very limited.

However with the increasing number of sensors connected to network in production chain, it is easier to collect the production chain data. This feature is provided by Field Level Bus.

The Field Level Bus collects data from various industrial control systems, and loads them into the Data Lake storage. Due to the big amounts of periodic or continuous data collected at this level, Data Lake builds on the Hadoop cluster technology which is the most suitable solution to store the raw field level data.

Main task of Field Level Bus is preparing the data which is a fundamental step for the further use of field level data, as the data can be collected from various, sensors, PLC, devices, systems, etc.

ISBN: 978-1-61804-327-6

Data collected at the field level can also be inconsistent. Therefore, transforming the collected data into cleaned forms storable in Data Lake storage is necessary. This Field Level Bus addresses the need of data analysis aimed at cleaning the raw data. [22]

VI. Conclusion

Knowledge discovery analytic platform proposed in this paper incorporates novel trends and methods used in the knowledge discovery in manufacturing area. The traditional data warehouse approach for knowledge discovery platform is supplemented with Hadoop cluster, to store big data collected at the field level of hierarchical control.

The proposed analytic platform preserves the robustness and well-proven technology for traditional business intelligence and analytic tools, and creates space for knowledge discovery in frequently and continually updated manufacturing data in a raw form. Therefore it represents an ideal compromise between existing traditional tools and the need for strong business intelligence, reporting and analytic platform.

The main disadvantage is the necessity of integrating all data in a Data Lake, which makes it difficult to ensure the integrity and security of company data. In traditional relational databases and data warehouses various approaches, methods and tools for maintaining integrity and security of company data are available. In Data Lake represented by Hadoop cluster, all data needs to be integrated altogether, and the discovery platform must have access to all the data. This is one of the main issues addressed when implementing Data Lake.

The proposed approach focuses on all hierarchical control levels in manufacturing. Therefore, manufacturing area as a whole represents the main application area of this approach. With a certain degree of abstraction, the approach can be applied also in other industrial fields, where lots of data needs to be collected frequently or continuously.

Acknowledgment

This publication is the result of implementation of the project VEGA 1/0673/15: “Knowledge discovery for hierarchical control of technological and production processes” supported by the VEGA.

References

[1] J. Jadlovský, S. Laciňák, M. Čopík and J. Ilkovič, “Technological level of flexible manufacturing system control,” Acta Electrotechnica Informatica, vol.11, No.1, pp. 20-24, 2011.

[2] P. Tanuška, P. Važan, M. Kebísek and D. Jurovatá, “Knowledge discovery from production databases for hierarchical process control,” International Journal of Mechanical, Aerospace, Industrial, Mechatronic and Manufacturing Engineering vol.7, No:11, 2013.

[3] J. Jadlovský, J. Laciňák, J. Chovaňák and J. Ilkovič. “Návrh distribuovaného systému riadenia pružnej výrobnej linky,” In: International Conference – Cybernetics and informatics. Vyšná Boca. 2010.

[4] C. Gröger, F. Niedermann and B. Mitschang, “Data Mining-driven Manufacturing Process Optimization,” In: Proceedings of the World Congress on Engineering 2012 Vol III, WCE 2012. Hong Kong: Newswood 2012, pp. 1475-1481.

[5] K. Wang, S. Tong, B. Eynard, L. Roucoules and N. Matta ,“ Fuzzy systems and knowledge discovery,” FSKD, 2007.

[6] P. Michalik, J. Štofa and I. Zolotová, “Testing the properties of K-means algorithm for data mining applications,” In: LINDI 2013 : 5th IEEE International Symposium on Logistics and Industrial Informatics : Proceedings : September 5-7, 2013, Wildau, Germany. - Piscataway : IEEE, 2013 P. 99-102. - ISBN 978-1-4799-1257-5.

[7] G. Köksal, İ. Batmaz and M. C. Testik, “A review of data mining applications for quality improvement in manufacturing industry,”Expert Systems with Applications, 38 (10) (2011), pp. 13448–13467.

[8] Rexer Analytics, “Data miner survey – 2013 survey summary report,“ 2014, [cit. 20.06.2015]. Available online: http://www.rexeranalytics.com/Data-Miner-Survey-Results-2013.html.

[9] KDnuggets, “Data Mining Community‘s Top Resource,“ 2014, [cit. 23.06.2015]. Available online: http://www.kdnuggets.com.

[10] B. Vorhies, “The big deal about big data: what’s inside – structured, unstructured, and semi-structured data in data magnum blog,“ 2013. [cit. 25.06.2015] Available online: http://data-magnum.com/the-big-deal-about-big-data-whats-inside-structured-unstructured-and-semi-structured-data/.

[11] X.Z Wang, “Data mining and knowledge discovery for process monitoring and control advances in industrial control,“ Springer Science & Business Media, 2012. 251p. ISBN 978-1-44710-421-6.

[12] X. Zhu, “Knowledge discovery and data mining: challenges and realities,“ Challenges and Realities. 2007. Idea Group Inc (IGI). 290p. ISBN 978-1-59904-252-7.

[13] H. Chen, R. H. L. Chiang, and V. C. Storey, “Business intelligence and analytics: from big data to big impact ,“ MIS Q. 36, 4 (December 2012), 1165-1188.

[14] P. Chapman, P. Kerber, J. Clinton, J. Khabaza, T. Reinartz, C. Shearer and R. Wirth, “The CRISP-DM Process Model,”. Discussion Paper. 05-03-99. Marec 1999.

[15] “What is the CRISP-DM Methodology,” [cit. 20.06.2015]. Available online: http://www.sv-europe.com/crisp-dm-methodology/.

[16] J.A. Harding, M. Shahbaz, S. Srinivas and A. Kusiak, “Data mining in manufacturing: a review,” Journal of Manufacturing Science and EngineeringTransactions of ASME, 128(4), 969–976.2006.

[17] A.K. Choudhary, M.K. Tiwari and J.A. Harding, “Data mining in manufacturing: a review based on the kind of knowledge. ” In: Journal of Intelligent Manufacturing. Leicestershire: Loughborough University´s Institutional Repository. 20 (5), s. 501 – 521. 2009.

[18] KDnuggets, “Algorithms for data analysis/data mining. Which methods/ algorithms did you use for data analysis?,” [cit. 19.06.2015]. Available online: http://www.kdnuggets.com/polls/2011/algorithms-analytics-data-mining.html.

[19] T.H. Davenport and J. Dyché, “Big data in big companies,” Thomas H. Davenport and SAS Institute Inc May 2013.

[20] W.Fan and A. Bifet, “Mining big data: current status, and forecast to the future,” SIGKDD Explor. Newsl. 14, 2 (April 2013), 1-5.

[21] M. Rouse, “Big data analytics,” [cit. 20.06.2015]. Available online: http://searchbusinessanalytics.techtarget.com/definition/big-data-analytics.

[22] S. Zhang, C. Zhang and Q. Yang, “Data preparation for data mining, Applied Artificial Intelligence,” An International Journal, Volume 17, Issue 5-6, 2003. Taylor & Francis, 2003. doi: 10.1080/713827180.

Standardization of Electronic Identity Management

Assoc. Prof. Roumen Trifonov, PhD
Department Computer Systems
Technical University - Sofia
Sofia, Bulgaria
e-mail: r_trifonov@tu-sofia.bg

Assoc. Prof. Radoslav Yoshinov, PhD
Telematics Laboratory
Bulgarian Academy of Sciences
Sofia, Bulgaria
e-mail: yoshinov@cc.bas.bg

***Abstract*—** Last year was adopted the Regulation (EU) No 910/2014 on electronic identification and trust services for electronic transactions in the integral market. In accordance with its performance requirements currently in the country a draft law on electronic identity is the process of discussion. These circumstances make it necessary to conduct a serious analysis of international standards in the field of electronic identification and solving a number of issues related to their application

***Keywords*—** electronic identification, international standards, identity management, identity authentication, identity federation

I. INTRODUCTION – THE NEW EUROPEAN REGULATION AND ITS RELATIONS TO THE STANDARDIZATION

On July 23th, 2014, the European Parliament and the Council of Ministers adopted the Regulation (EU) No 910/2014 on electronic identification and trust services for electronic transactions in the integral market and repealing Directive 1999/93/EC [1]. With a view to ensuring the proper functioning of the internal market while aiming at an adequate level of security of electronic identification means and trust services this Regulation:

- lays down the conditions under which Member States recognise electronic identification means of natural and legal persons falling under a notified electronic identification scheme of another Member State;
- lays down rules for trust services, in particular for electronic transactions; and
- establishes a legal framework for electronic signatures, electronic seals, electronic time stamps, electronic documents, electronic registered delivery services and certificate services for website authentication.

The Regulation defines assurance levels, which should characterise the degree of confidence in electronic identification means in establishing the identity of a person, thus providing assurance that the person claiming a particular identity is in fact the person to which that identity was assigned. The assurance level depends on the degree of confidence that electronic identification means provides in claimed or asserted identity of a person taking into account processes (for example, identity proofing and verification, and authentication), management activities (for example, the entity issuing electronic identification means and the procedure to issue such means) and technical controls implemented. Various technical definitions and descriptions of assurance levels exist as the result of Union-funded Large-Scale Pilots, standardisation and international activities.

The regulation addresses the issue of standardization in general, paying attention to their compliance, without focusing on many standards. An example is the following passage: "In adopting delegated acts or implementing acts, the Commission should take due account of the standards and technical specifications developed within the European and international organizations and standards bodies, in particular the European Committee for Standardization (CEN), the European Institute for Telecommunications Standards (ETSI), the International Organization for Standardization (ISO) and the International Telecommunication Union (ITU), with a view to ensuring high levels of security and interoperability of electronic identification and trust services. "

The document mentions specifically only two international standards: ISO 29115 “Entity authentication assurance framework” and ISO 15408 “Evaluation criteria for IT security” (popular under the name “Common Criteria”). The second one is not directly related to electronic identification.

In accordance with the requirements of the Regulation currently in the country a draft law on electronic identity is the process of discussion. These circumstances make it necessary to conduct a serious analysis of international standards in the field of electronic identification and solving a number of issues related to their application.

II. THE MAIN DIRECTION OF THE STANDARDIZATION AND THE MAIN PLAYERS EASE OF USE

Especially in the European Union's privacy legislation technical standards are given significant role, e.g. when the law requires technical mechanisms to be put in place to protect the privacy of users. The specific mechanisms are usually not spelled out in the law itself, as the regulation aims to be “technology neutral”. Effectively, such regulation would allow for a strong harmonization of technical measures in specific domains under the control of the Commission.

Over the past few years have seen increased interest of international standardization bodies to problems of electronic identification. As a result of this intense activity has been adopted many standards that in general beam can be systematized in the following areas:

- Identity Management;
- Identity Authentication;
- Identity Federation.

This standardization work has been realized not only in official intergovernmental standardization organizations (ISO, ITU), but also in a number of technical consortia. It is worth noting that often ISO acts as an integrator of the work of different standards bodies around the world and enjoys a high level of recognition especially with national governments and international institutions.

2A. Standards Related to the Identity Management

The main standards in this area, adopted in recent years, can be marked as follows, broken down by standardization organizations:

a) Standards of the International Standardization Organization (ISO):

- ISO/IEC 24760 A framework for identity management;
- ISO/IEC 29144:2014 The use of biometric technology in commercial Identity management applications and processes;
- ISO/IEC 29003 Identity proofing;

b) Standards (Recommendations) of the International Telecommunication Union (ITU):

- X.1250 Baseline capabilities for enhanced global identity management and interoperability;
- X.1251 A framework for user ISO/IEC 24760 control of digital identity;
- X.1252 Baseline identity management terms and definitions;
- X.1253 Security guidelines for identity management systems;
- X.1255 Framework for discovery of identity management information;

c) Standard of the American National Standardization Institute (ANSI):

- ANSI/NASPO-IDPV-2014 Requirements and Implementation Guidelines for Assertion, Resolution, Evidence, and Verification of Personal Identity.

The basal ITU recommendation X.1252 [2] determined graphically the Identity Management as follows (Fig. 1).

The basal ISO standard ISO/IEC 24760 [3] "specifies a framework for the issuance, administration, and use of data that serves to characterize individuals, organizations or information technology components, which operate on behalf of individuals or organizations. Furthermore, it specifies fundamental concepts and operational structures of identity management."

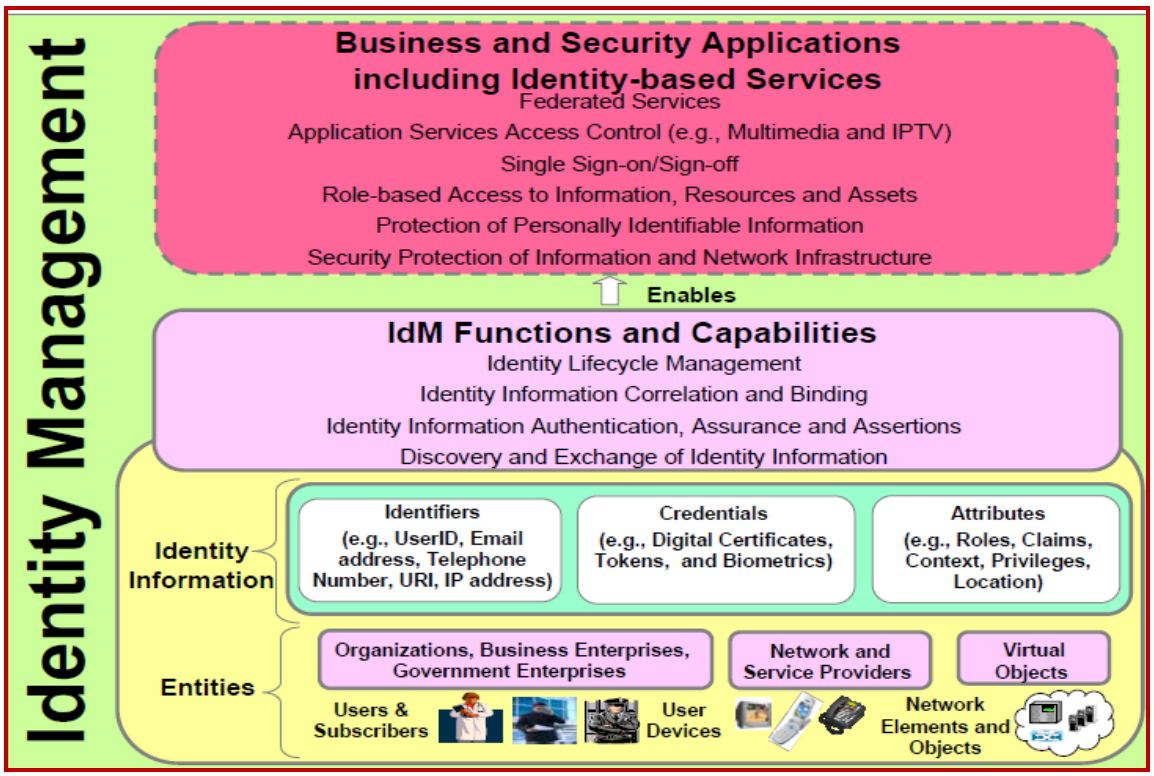


Fig. 1. *Determination of the Identity Management*

ISBN: 978-1-61804-327-6

2B. Standards Related to the Identity Authentication

The main standards in this area, adopted in recent years, can be marked as follows, broken down by standardization organizations (some of the basic standards of ISO coincide with similar recommendations of ITU):

a) Standards of the International Standardization Organization (ISO):

- ISO/IEC 29115 (ITU-T X.1254) Entity authentication assurance framework;
- ISO/IEC 24761:2013 Authentication context for biometrics;
- ISO/IEC 17922 (ITU-T X.1085) Tele-biometric authentication framework using biometric hardware security module;
- ISO/IEC 29146 A framework for access management;
- ISO/IEC 29191:2012 Requirements for partially anonymous, partially unlinkable authentication;
- ISO/IEC 20009:2013 Anonymous Entity Authentication;
- ISO/IEC 20008:2013 Anonymous Digital Signatures;
- ISO/IEC 29100 Privacy Framework;
- ISO/IEC 29101 Privacy Architecture Framework;
- ISO/IEC 29190 Privacy Capability Assessment Model;
- ISO/IEC 24475 Biometric Information Protection;
- ISO/IEC 27018 Code of Practice for Data Protection Control for Public Cloud Computing Services;

b) Standards (Special Publications) of the National Institute for Standardization and Technology (NIST):

- SP 800-63-2 Electronic Authentication Guideline;
- SP 800-118 Guide to Enterprise Password Management;
- SP 800-122 Guide to Protecting the Confidentiality of Personally Identifiable Information (PII);
- FIPS 201-2 Personal Identity Verification of Federal Employees and Contractors.

The authors would like to pay particular attention to the NIST special publication SP 800-63-2 [4], whose specificity is largely coincides with the objectives of e-Governance in Bulgaria. This SP contains the guiding technical principles for federal authorities using electronic authentication. It covers the "online" authentication of users (employees, contractors and individuals) interacting with administrative information systems through open networks. The SP also defines the technical requirements for each of the four levels of protection in the field of proof of identity, the registration, the technical media, the management processes, the protocols for authentication and the related claims. The main part of these processes are illustrated in the Fig. 2.

2C. Standards Related to the Identity Federation

The main standards in this area, adopted in recent years, can be marked as follows, broken down by standardization organizations:

a) Standards of the Organization for the Advancement of Structured Information Standards (OASIS):

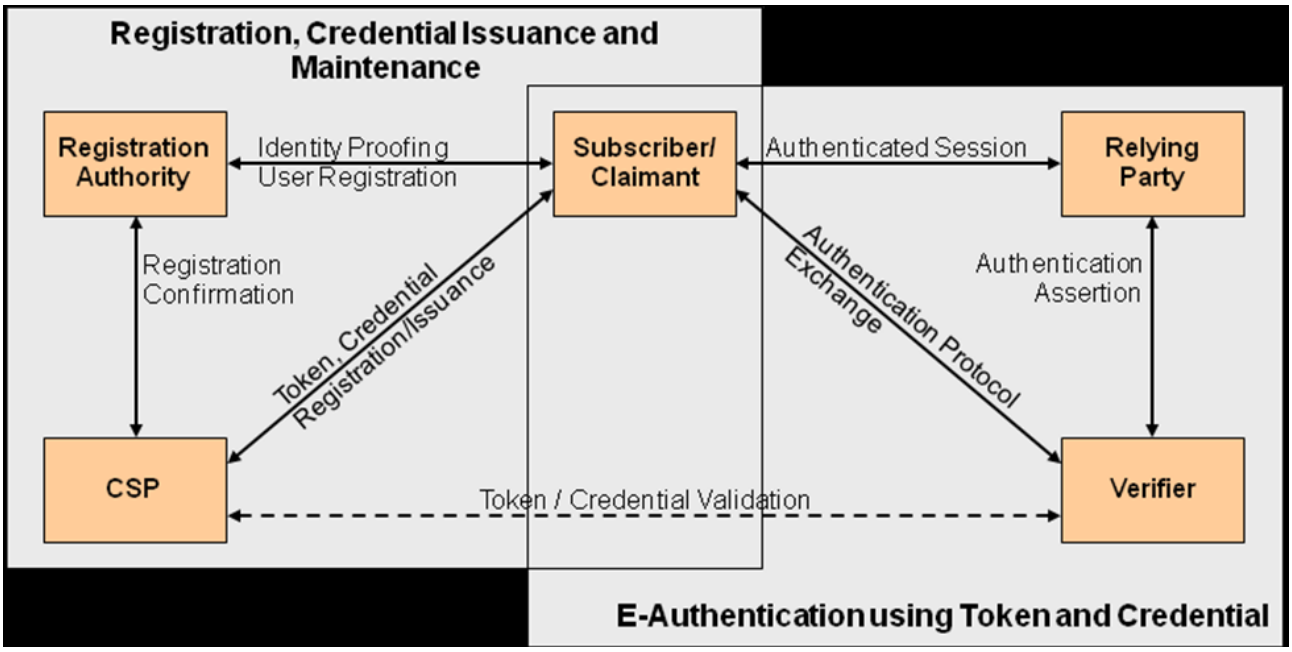


Fig. 2. *E-Authentication of userst*

- Security Assertion Markup Language 2.0 (SAML 2.0);
- Identity Metasystem Interoperability Version 1.0;
- WebServices-Trust Version 1.3

b) Standard (Recommendation) of the International Telecommunication Union (ITU): X.1154 - General framework of combined authentication on multiple identity service provider environments

c) Standard (Specification) of OpenID Foundation: OpenID Connect 1.0;

d) Standard (Specification) of Fast IDentity Online Alliance: FIDO v1.0.

III. . PROBLEMS THAT MUST BE SOLVED IN NATIONAL LEGISLATION

The analysis of the draft law on electronic identification, which was presented for public discussion, in terms of the aforementioned standards shows the presence of a number of "white spots".

Among the identified problems, which should receive unambiguous answer, can be specified the following ones:

- the cited in the Regulation standard ISO / IEC 29115 (ITU X.1154) identifies four levels of authentication assurance (Levels of Authentication - LoA) of the object and describes requirements and guidelines for each of the levels. The choice of the appropriate level should be carried out on the base of the risk assessment for transaction and for respective service. Bringing the levels of exposure to the levels of LoA, the parties of the transaction can determine their necessary LoA. Therefore, it is advisable to legitimize a range of methods for e-ID intended for various applications, in order to meet different LoA;
- moreover, the legislation must answer the question: will it be used in e-Governance applications the principle of so called “Federated Identity Management (FIdM)” and its particular applications, such as so called “Single Sign-On (SSO)”.

IV. CONCLUSIONS

The purpose of this paper is to contribute to the public discussion on the draft of the Electronic Identification Act, which must be adopted in compliance with the relevant EU regulation. It is necessary to take actions aimed targeted to the understanding and the compliance with international standards in this area.

REFERENCES

[1] Regulation (EU) No 910/2014 on electronic identification and trust services for electronic transactions in the integral market and repealing Directive 1999/93/EC

[2] ITU Recommendation X.1252 Baseline identity management terms and definitions

[3] ISO/IEC 24760 A framework for identity management

[4] NIST Special Publication 800-63-2 Electronic Authentication Guideline August 2013

Thermal power analysis of a single family housing

Stan Ivan Felicia Elena, Dinu Radu Cristian

***Abstract*—**this paper presents a study regarding the determination of the necessary heat for a single family house. In the first part is presented an introduction witch describes the main idea of the paper. The second part includes a mathematical model for the necessary heat equations and presents the main equations that are used. The last part of the paper is a calculation example to determine the thermal power of a single family house. The case study includes the constructive details of the house, determining the necessary heat, sizing boilers for hot water and the calculation of the necessary heat for hot water preparation.

***Keywords*—** thermal power, necessary heat, mathematical model, boilers seizing, hot water.

I. Introduction

THIS material presents the calculation of the heat for single-family house, based on current trends existing in this field worldwide.

The analysis presented is carried out in the international context in which the concepts of energy conservation, reduction of emissions and pollutants, energy independence, environmental impact and others, becomes increasingly important significance.

In Romania, most of the residential buildings were completed without significant concerns for the quality of their energy, but in recent years were introduced precise regulations in this area and there is a growing concern for issues regarding the insulation quality and efficient solutions for thermal energy production. [1], [5]

II. The Mathematical Model

Heat housing needs can be determined by adding its three major components:[1]

$$\dot{Q} = \dot{Q}_1 + \dot{Q}_2 + \dot{Q}_3 \qquad [W] \qquad (1)$$

where: $\dot{Q}_1$ - load or thermal power transmitted through the building envelope, [W]

$\dot{Q}_2$ - Load or thermal power due to the charge of ventilation or airflow, [W]

$\dot{Q}_3$ - Load or thermal power for water heating, [W].

Load or thermal power transmitted through the building envelope presents several components, which can be summarized:

$$\dot{Q}_1 = \dot{Q}_{11} + \dot{Q}_{12} + \dot{Q}_{13} + \dot{Q}_{14} + \dot{Q}_{15} \qquad [W] \qquad (2)$$

where: $\dot{Q}_{11}$ - Load or thermal power transmitted through building walls, [W]

$\dot{Q}_{12}$ - Load or thermal power transmitted through ceiling, [W]

$\dot{Q}_{13}$ - Load or thermal power transmitted through the windows, [W]

$\dot{Q}_{14}$ - Load or thermal power transmitted through the floor, [W]

$\dot{Q}_{15}$ - Load or thermal power transmitted through cellar floor, [W]

The thermal loads transmitted through the elements of the building envelope are determined using the calculation relations as:

$$\dot{Q}_{1i} = k_i \cdot S_i \cdot (t_i - t_e) \qquad [W] \qquad (3)$$

where: ki -represents the global heat transfer coefficient through the envelope element, [W/m^{2}K];

S_i -is the surface of the envelope element, m^2;

t_i - is the temperature inside the house, °C;

t_e - is outdoor temperature °C.

The overall coefficient of heat transfer is determined using the relationship:[4]

$$k_i = \frac{1}{\frac{1}{\alpha_i} + \sum \frac{\delta}{\lambda} + \frac{1}{\alpha_e}} \qquad [W/m^2K] \qquad (4)$$

F.E. STAN IVAN is with the Faculty of Electrical Engineering, University of Craiova, Craiova 200440 Romania (email:ely_felicia@yahoo.com; fivan@elth.ucv.ro;e-mail).
R.C. DINU is with the Faculty of Electrical Engineering, University of Craiova, Craiova 200440 Romania (e-mail: rcdinu@ elth.ucv.ro).

where: α_i - is the coefficient of overall convection between the building element and the air inside the building, and his value is considered 8 W/m^{2}K, which corresponds to the natural convection;

α_e - is the global coefficient of convection between the building element and the air from outside the building, and his value is considered 25 W/m^{2}K, which corresponds to the most unfavourable conditions

δ - is the thickness of each layer from the building envelope, m

λ - is the conductivity of each component material from the building envelope, W/mK.

In the tables below are presented the thermal conductivity values for a few usual material as well as values of the global coefficient of heat transfer for some types of windows.

Table 1. Values of thermal conductivity for some common building materials

Material	Conductivity [W/mK]
Concrete	1,45
Brick	0,90
Autoclaved concrete cell	0,40
Plated wood	0,10
Hardwood Oak	0,37
Hardwood pine	0,28
Stone	2,90
Polyurethane	0,018
Extruded polystyrene	0,035
Expanded Polystyrene	0,040
Mineral wool	0,041
Coork	0,054

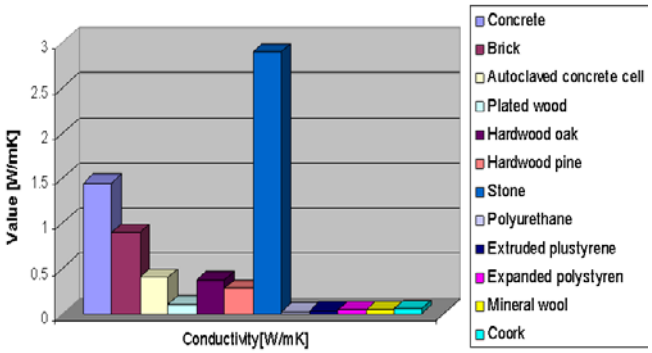


Fig 1. Conductivity variation for different constructions materials

Table 2. Values of the global coefficient of heat transfer for certain types of windows

Window	global coefficient [W/m^{2}K]
Window panes with 3 layers of glass and Kr between starts	0,5
Window panes with 3 layers of glass	0,8
Low E Window panes	1,1
Window panes Float-Float	1,4
Double glazed window	2,0
Window with single glazing	2,5

Because of leaks, for these types of windows in practice global heat transfer coefficient values are higher. Load or thermal power due to ventilation or airflow can be calculated based on heat load transmitted by the tire, with the relationship:

$$\dot{Q}_2 = \varepsilon \cdot \dot{Q}_1 \quad [W] \tag{5}$$

where ε It is a proportionality coefficient, whose value can be regarded as

0,7	0,8	0,9	1
for buildings without thermal insulation	For building with minimal insulation	For buildings with good insulation	for buildings with high performance insulation (houses with low energy consumption or passive houses)

Thermal power for hot water preparation depends on the number of people using home:

$$\dot{Q}_3 = \frac{n \cdot \rho \cdot V \cdot c_p \cdot (t_{we} - t_{wi})}{\tau} \quad [W] \tag{6}$$

where: n – the number of persons

ρ - water density, kg/m^3

V –water daily consume, for one person, m^3

c_p - the specific heat of water, kJ/kgK

t_{we} -the temperature for water to be heated, °C

t_{wi} –the temperature of cold water, °C

τ - the period during which the water is heated, s.

III. CASE STUDY

We consider a building with the following constructive details.

Table 3. Construction details for the analyzed building [2]

Element	Material	S[m^2]
Exterior wall	Exterior plaster	10,25
	Autoclaved concrete	
	Polystyrene	
	Interior plaster	
Plate over basement	Parquet	67,6
	Cement screed	
	Insulation	
	Concrete	
	Polystyrene	
Roof	Gypsum boards	51,34
	Mineral wool	
	Mineral wool	
	Wood	
Windows	Glass	6
	Frame	

Temperatures of calculation are considered t_i=20°C for indoor temperature and t_e=-15°C outdoor temperature (outdoor temperature average during the heating period for Craiova city).

Further will be determinate the thermal parameters for the 3 components of the building examined.

A. Determining the heat demand of housing

According to the relationship (1), it will be determined the thermal power transmitted through the building envelope.

ISBN: 978-1-61804-327-6

Table 4 presents the thermal characteristics parameters for the main components of the building and are determined: the temperature difference, heat resistance and the corrected global heat transfer coefficient, and in the final are calculated all the elements of the heat demand of housing.

Table 4. The thermal characteristics parameters for the analyzed building components [2]

Element	Material	S [m²]	1/α_i	1/α_e	δ[m]	λ[W/mK]	R[m²K/W]	Δt[°C]	R't [m²K/W]	U [W/m²K]
Exterior wall	Exterior plaster	10,25	0,13	0,04	0,02	0,87	0,02	35	9,64	0,10
	Autoclaved concrete				0,18	0,21	0,86			
	Polystyrene				0,3	0,035	8,57			
	Interior plaster				0,02	0,87	0,02			
Plate over basemen	Parquet	67,6	0,13	0,04	0,022	0,13	0,17	35	7,46	0,13
	Cement screed				0,05	1,05	0,05			
	Insulation				0,03	0,04	0,75			
	Concrete				0,16	2,1	0,08			
	Polystyrene				0,25	0,04	6,25			
Roof	Gypsum boards	51,34	0,13	0,04	0,1	0,7	0,14	35	11,08	0,09
	Mineral wool				0,2	0,04	5,00			
	Mineral wool				0,2	0,04	5,00			
	Wood				0,1	0,13	0,77			
Windows	Glass	6	0,13	0,04				35	0,91	1,1
	Frame									

In the table 5 is presented the thermal power transmitted to the building envelope components.

Table 5. Determination of heat transmitted through the building envelope of the analyzed building components

Element	Q_1[W]
Exterior wall	37,20
Plate over basemen	317,03
Roof	162,14
Windows	231

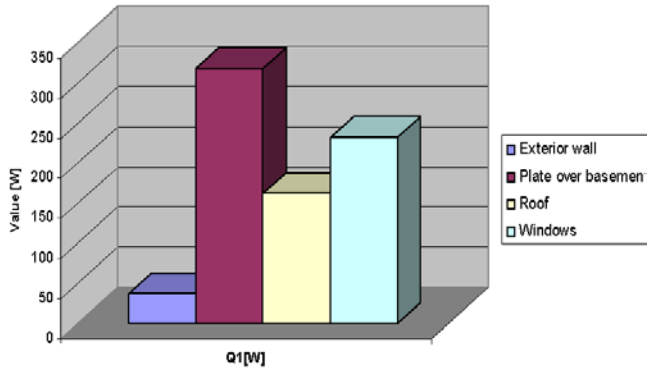


Fig 2. Conductivity variation for different constructions materials

The thermal power transmitted through the building envelope, $\dot{Q}_1$ is:

Q_1=747,37 W $\Rightarrow$ Q_1 = 0,747371 kW

The thermal power due to ventilation or airflow $\dot{Q}_2$ is:

Q_2=672,63W $\Rightarrow$ Q_2 = 0,672634 kW

Table 6. The characteristic parameters of the boiler sizing process for water consumption

n	ρ[kg/m3]	V[m3]	c_p[kJ/kgK]	t_{acm}[°C]	t_{ar}[°C]	τ[s]
4	1	0,085	4,19	65	5	1200

Load and thermal power for hot water preparation is:

Q_3 =490,23W $\Rightarrow$ Q_3 =0,49023 kW

Heat housing needs, $\dot{Q}$ it can be determined by adding its three major components:

$$\dot{Q} = \dot{Q}_1 + \dot{Q}_2 + \dot{Q}_3$$

Q=1910234,31W $\Rightarrow$ Q= 1910,234 kW

B. Sizing boilers for hot water

The calculation for sizing hot water preparation boilers, aims to determine their amount at least equal to the volume of hot water daily needs. In the adjacent tables, according to international standards, it can be observed that for the domestic hot water preparation at 45°C the amount of water should be higher than the DHW to 60°C, to completely cover daily consumption.

Table 7. Consumption of hot water in homes

Temperature	Consume type		
	Low comfort, [l/pers/day]	Normal comfort[l/pers/day]	High comfort [l/pers/day]
60°C	10...20	20...40	40...70
45°C	15...30	30...60	60...100

Table 8. Consumption of hot water in hotel establishments, guesthouses and hostels

Temperature	Room type			
	With bathroom and shower, [l/pers/day]	With bathroom [l/pers/day]	With shower [l/pers/day]	Pensions, hostels
60°C	115...175	90...135	50...90	25...50
45°C	170...260	135...200	75...135	40...75

Indicative sizing, in terms of thermal system for domestic hot water housing, the use of renewable energy sources, it can be considered normal hot water consumption of 50 litters / person/day at 45°C. If the beneficiary is estimated that exceed normal hot water consumption indicated in the table, will take into account this and be sized boiler water consumption indicated by the beneficiary. Minimum volume of the tank V_{bmin} can be calculated with the formula:

$$V_{bmin} = \frac{n \cdot C_{zn} \cdot (t_{acm} - t_{ar})}{t_b - t_{ar}} \quad [l] \qquad (7)$$

where: n – the number of persons;
C_{zn} – normal daily consumption per person considered;
t_{acm} - DHW temperature at the point of consumption;
t_{ar} – cold water inlet temperature in the boiler;
t_b – hot water temperature in the boiler.

ISBN: 978-1-61804-327-6

Taking into account the above, the boiler volume V_b is calculated by the relationship:

$$V_b = f \cdot V_{b\,min} = f \cdot \frac{n \cdot C_{zn} \cdot (t_{acm} - t_{ar})}{t_b - t_{ar}} \text{ [l]} \quad (8)$$

where: f = 1,5…2, for the case of using the solar energy or heat pumps;

f = 1, when using classic fuels, solid biomass, biogas or electricity.

Next will be considered a particular case sizing DHW cylinder, considering a home with four people, a normal consumption of hot water net daily consume(C_{zn}) = 50l/person/day and different energy sources.

Table 9. Determining the minimum amount V_{bmin} and boiler volume V_b

f	n	C_{zn}[l]	t_{acm}[°C]	t_{ar}[°C]	t_b[°C]	V_{bmin}[l]	V_b[l]
1	4	50	45	10	60	140	140

C. Calculation of the necessary heat for hot water preparation

The thermal load required for hot water preparation is determined by the relationship:

$$\dot{Q}_{acm} = \frac{m \cdot c_w \cdot (t_b - t_r)}{\tau \cdot 3600} \ [kW] \quad (9)$$

where: m is the amount of hot water prepared:

$$m = n \cdot C_{zn} \cdot \rho \ [kg]$$

where: ρ is the density of water, which varies depending on temperature, but for calculation it can be assumed indicative ρ = 1000 kg /m^3;

c_w is the specific heat of water - can be considered, $c_w = 4{,}186$ kJ/kgK

t_b is the cylinder temperature, so the temperature to which the water is heated;

t_r is the temperature of the cold water, with seasonal variation and depending on the geographical position

Generally in summer t_r = 12...17°C, in the winter t_r =5...10°C. For calculations can be considered indicative tr = 10°C;

τ [h] is the time the water is heated.

The thermal load required for hot water preparation:

$$\dot{Q}_{acm} = \frac{n \cdot C_{zn} \cdot \rho \cdot c_w \cdot (t_b - t_r)}{\tau \cdot 3600} \text{ [kW]} \quad (10)$$

The thermal load required for hot water preparation needed daily for a person can be calculated with the previous relationship, considering n = 1:

$$\dot{Q}_{acmp} = \frac{C_{zn} \cdot \rho \cdot c_w \cdot (t_b - t_r)}{\tau \cdot 3600} \ [kW] \quad (11)$$

Table 10. Heat determination Q_{acm} required for hot water preparation

m[kg]			c_w[kJ/kg]	t_b[°C]	t_r[°C]	τ[h]	Q_{acmp}[KW]	Q_{acm}[W]
n	C_{zn}[l]	ρ[kg/m^3]						
4	50	1	4,186	45	10	8	1,02	1017,431

The thermal load required for hot water preparation needed daily for a person:

Q_{acmp} = 0,25 KW ⇒ Q_{acmp} = 254,3576 W

Heat required Q_{acm} for hot water preparation:

$$Q_{acm} = Q_{acm} \cdot \tau \ \text{[kWh]}$$

⇒ Q_{acm} = 8,14 [KWh]

IV. DETERMINATION THE LOSSES THROUGH SHEATH HEATER [7]

Boiler located in the basement

The annual amount of heat dissipated by the shell boiler located in the basement of an existing building (cold room) is determined by the relationship:

$$Q_{pbolier} = \frac{0{,}001 \cdot S_{lat}}{0{,}10 + \frac{\delta_m}{\lambda_m} + \frac{\delta_{iz}}{\lambda_{iz}}} \cdot \sum_k n_{h_k} (t_{acb} - t_{sbk}) \ [kWh/year] \quad (12)$$

where: S_{lat} - lateral surface of the cylinder without insulation, [m^2]

δ_m - tank wall thickness (metal), [m]

δ_{iz}- average thickness of the insulation, [m]

λ_m - metal thermal conductivity [W/mK]

λ_{iz} - thermal conductivity depending on the condition of the insulation [W/mK]

t_{acb}- average water temperature in the boiler

$$t_{acb} \approx 0{,}70 \cdot t_{ac0} \left(t_{ac0} = 50^{\circ} C ....60^{\circ} C \right)$$

V. CONCLUSION

Due to a continuous increase of energy prices, it is necessary the thermal rehabilitation of homes, administrative, commercial and industrial buildings. From the annual energy consumption of homes, 55% is for heating. In a house properly isolated, the heat losses are about 20-25% for the exterior walls, basement and roof by 20-30% and 20-25% level for non-performing windows and doors.

ISBN: 978-1-61804-327-6

The analyzed building has a good insulation for the exterior walls and the windows are efficient.

From the values of the heat demand of housing results that the power transmitted through the building envelope and the thermal power due to ventilation or airflow are relatively low, Q_1 = 0,747371 kW and Q_2 = 0,672634 kW. Also the heat housing needs has a value that demonstrate the fact the building has an efficient insulation Q = 1910,234 kW.

As a novelty brought by this article is the approach of the heat requirement for a family, starting from the mathematical model. In the paper it is presented a numerical example in order to determine: the heat requirement, the load and the thermal power for hot water, the sizing boilers for hot water and the calculation of heat for domestic hot water.

This approach is quite complex and a great novelty in the field of energy efficiency in buildings.

Future directions of approach are to create a mathematical model and a computer program possibly using common programming languages in order to develop the application, increase computing speed and determination of several parameters in a relatively short time.

REFERENCES

[1] Analiza termoenergetică a locuinţelor unifamiliale, Capitol realizat în colaborare cu: Ş.l. dr. ing. Lorentz JÄNTSCHI şi ing. Margareta Emilia PODAR

[2] Eficienţa energetică şi economică a clădirilor, Stan Ivan Felicia Elena**,** Mircea Ion**,** Editura SITECH, Craiova 2014, ISBN 978-606-11-3880-7

[3] Metodologie de calcul al performanţei energetice a clădirilor partea a III-a - Auditul şi certificatul de performanţă a clădirii indicativ Mc 001/3 – 2006

[4] Caracteristici termoenergetice ale anvelopei clădirilor, http://instal.utcb.ro/site/proiectecoordonare/serefen/ctec.pdf

[5] Normativ pentru expertizarea termică şi energetică a clădirilor existente şi a instalaţiilor de încălzire şi preparare a apei calde de consum aferente acestora", indicativ NP 048-2000, aprobat de MLPAT cu ordinul nr. 324/N/4.12.2000, http://www.algorithm.ro/resurse/audit_energetic/NP-048-2000.pdf

[6] Reabilitarea termică a locuinţelor şi clădirilor publice, comerciale şi industriale,

[7] INCERC, Buletinul Constructiilot, Vol. 4, 2001

Felicia Elena P. Stan Ivan was born in 1982 in Craiova, Romania. In 2005 she graduated Economic Engineering in power engineering at University of Craiova and in 2009 she received Ph.D. in Energy Engineering from University of Craiova.

At present she is a university assistant at the Faculty of Electrical Engineering, University of Craiova. Her field of interest is energy efficiency of the buildings operation, analysis and value engineering.

Radu Cristian Dinu was born in 1974 in Deva, Romania. In 2000 he graduated Thermoelectric plants at Electrotehnic Faculty University of Craiova. In 2007 he received his Ph.D. in Civil Engineering, Technical University of Civil Engineering Bucharest,

Presently he is a lecturer in Faculty of Electrical Engineering, University of Craiova. His field of interest is Introduction to energy engineering, Thermoenergetics and district heating, industrial energy, modern solutions for energy production.

Modeling the probability of failure-free operation of control systems

MICHALCONOK German [a]
Faculty of Material Sciences Technology in Trnava
Slovak University of Technology,
917 24 Trnava, Slovak Republic
[a]german.michalconok@stuba.sk

KORYTAR Marek [b] and NEMETH Martin[c]
Faculty of Material Sciences Technology in Trnava
Slovak University of Technology,
917 24 Trnava, Slovak Republic
[b]marek.korytar@stuba.sk, [c]xnemethm1@stuba.sk

The aim of this paper is to model the probability of the failure of the safety functions in the control systems. We propose a complete failure probability model for the safety functions of the control system. This model is designed to compute the intensity of the critical failure for the standard channel architectures. These architectures were designed with respect to the standard IEC 61508 and were implemented in Matlab.

Keywords: Control systems, safety functions, channel architectures, standard IEC- 61508

I. INTRODUCTION *(HEADING 1)*

Management of the safety-critical processes requires special approach from the engineering viewpoint. The aim is to eliminate risks that come from the operation of the safety-critical technological devices. The engineering design of such devices leads to the safe operation, predicts possible failures, weaknesses and defines resources, which can lead to reveal such failures. The specification of the requirements on safety-critical system should consist of []:

- Danger identification,
- Danger analysis (effect),
- Risk analysis,
- Risk management,
- Acceptable risk,
- Risk reduction through protection layers,
- Risk reduction through additional protection layers.

There are many techniques to support the identification of danger and risk. However, there is no technique that can solve everything. Risk study uses various methods and techniques.

Methods of danger and risk detection []:

- Checklists,
- What if study,
- Failure mode and effect analysis (FMEA),
- Hazard and operability analysis (HAZOP),
- Dynamic flow graph methodology (DFM).

Techniques for the risk analysis:

- Event tree analysis (ETA),
- Fault tree analysis (FTA),
- Cause consequence analysis.

Risk reduction techniques:

- Event tree analysis (ETA),
- Layer of protection analysis (LOPA).

The aim of this paper is the application of mathematical modeling principles in terms of risk analysis of the complex control system with the use of the FTA method.

II. SAFETY ANALYSIS OF CONTROL SYSTEMS

Safety is closely linked to the risk and therefore safety should be understood relatively. Total system safety is practically unattainable. The safety level of the safety system corresponds to the defined requirements []. Quantitative representation of the safety of the control system is understood as the probability of absence of any dangerous state in the control system with given conditions and time interval. Quantitative safety is mathematically defined as follows:

$$P_T(t) \geq P_R(t) > 0, \tag{1}$$

Where $P_T(t)$is the probability of the acceptable danger of the control system and $P_R(t)$ is the real probability of danger of the control system. Probability of safe operation means that the object can safely process the required function with given condition and time interval(t_1, t_2):

$$R_S(t_1, t_2) = 1 - F_H(t_1, t_2), \tag{2}$$

Where $F_H(t_1, t_2)$is a distribution function, which expresses the probability, that the object cannot process the required function in given time interval (t_1, t_2)

$$F_H(t_1, t_2) = \int_{t_1}^{t_2} f_H(t) \cdot dt, \tag{3}$$

Where $f_H(t)$ is the probability density of the dangerous failure of the object. The intensity of the dangerous failures $\lambda_H(t)$ is the limit of the conditional probability ratio, that the time of emerging of the dangerous failure will be found in the interval $(t, t + \Delta t)$ to the length of the time intervalΔt, where $\Delta t \to 0$.

$$\lambda_H(t) = \frac{f_H(t)}{R_S(t)} \tag{4}$$

The probability evaluation of the failure events has a significant importance in optimizing investments to the safety system. For example modernization of the revealed elements, which are responsible for not meeting the safety level requirements.

The characteristics of safety-critical system are changing in the time of its life cycle. One of them is the probability of

failure emerging $P(t)$ in time interval $(t, t+\Delta t)$ where applies:

$$P(t \le T < t+\Delta t) = F(t+\Delta t) - F(t) \qquad (5)$$

The distribution function $F(t)$ of the stochastic failures of the electrical objects is defined as []

$$F(t) = 1 - e^{-\lambda t}, \qquad (6)$$

Where λ is the parameter of the exponential distribution of failures (failure intensity). Function (2) can be developed to the form of Taylor's order:

$$1 - e^{-\lambda t} = \lambda \cdot t + \sum_{k=2}^{n} (-1)^{k-1} \cdot \frac{(\lambda t)^k}{k!} \qquad (7)$$

If applies, that $\lambda \cdot t \ll 1$ then we can consider:

$$1 - e^{-\lambda t} \cong \lambda t \qquad (8)$$

Then the failure emerging probability in the time interval$(t, t+\Delta t)$, when the failure did not occur, is:

$$P(t \le T < t+\Delta t) = (1 - e^{-\lambda \Delta t}) \cong \lambda \Delta t \qquad (9)$$

Safety function is a function, which can cause a safety hazard. Safety functions are technical actions to reduce the risk on to a tolerable level. Efficiency of the safety function is determined using SIL method (safety integrity level). The standard [29] defines the safety integrity as the probability, that the system will process all required safety functions in given order and in given time. Safety integrity consists of [Rastocny]:

- safety integrity against systematic failures, which is assessed by inspection of actions against the failures;
- safety integrity against the random failures, which is assessed by probability calculations.

The standard [29] defines four levels of safety integrity SIL with operation mode with low request and for systems with high request.

TABLE 1. Safety integrity levels

Safety integrity level (SIL)	Demand Mode of Operation	Continuous / High Demand Mode of Operation
4	$\ge 10^{-5}$ *to* $< 10^{-4}$	$\ge 10^{-9}$ *to* $< 10^{-8}$
3	$\ge 10^{-4}$ *to* $< 10^{-3}$	$\ge 10^{-8}$ *to* $< 10^{-7}$
2	$\ge 10^{-3}$ *to* $< 10^{-2}$	$\ge 10^{-7}$ *to* $< 10^{-6}$
1	$\ge 10^{-2}$ *to* $< 10^{-1}$	$\ge 10^{-6}$ *to* $< 10^{-5}$
1	$\ge 10^{-2}$ *to* $< 10^{-1}$	$\ge 10^{-6}$ *to* $< 10^{-5}$

This standard defines SIL to one safety function. Therefore it is important, to focus on what safety functions SCS (safety-critical control system) are realized and what parts of SCS are taking part on execution of each safety function. Main factors, which have impact on safety of SCS are [1]:

- architecture of SCS;
- failure-free components SCS;
- failure diagnostics SCS (diagnostics coverage, failure detection time);
- failure negation time (Induction of the safe state after the detection of a dangerous state);
- recovering of the SCS after failure (method, and time of restoration);
- technical independence (independence of the SCS elements when dealing with multichannel SCS) and organizational independence (independence of persons involved in the designing process of the SCS).

A person should be involved in the process of SCS architecture design. Human factor is considered as one of the most important aspects to ensure safe operation of potentially dangerous technical objects.

The "patchwork" approach prevails on all levels of control in the process of designing automation systems for technological processes. Decision making is in extreme situations left to personnel. The goal is to increase the level of automation of the technological processes, which are protected from the human failure.

III. PROBABILITY CALCULATIONS FOR TYPED SCS ARCHITECTURES

Safety functions are most often realized on the process level and on the operational level of control. Main parts of the block scheme of the SCS on the process level are:

- power supply;
- input and output interface;
- sensors;
- control logic;
- end elements (actuators, contactors ...).

The input interface can be as a standalone part, or can be a part of sensors or control logic, which depends on the technical realization on the SCS. Simplified block scheme of SCS is in the fig.1.

The average probability of failure on demand of a safety function for the E/E/PE safety-related system is determined by calculating and combining the average probability of failure on demand for all the subsystem, which together implement the safety function.

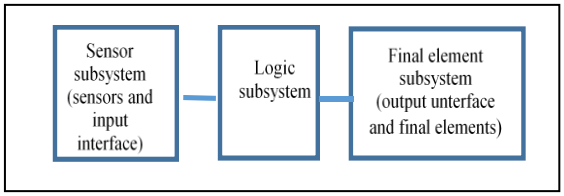


Figure 1. Subsystem structure

Since in this annex the probabilities are small, this can be expressed by the following (see Figure 1):

$$PFD_{SYS} = PFD_S + PFD_L + PFD_{EE} \qquad (10)$$

,where

PFD_{SYS} is the average probability of failure on demand of a safety function for the E/E/PE safety-related system;
PFD_S is the average probability of failure on demand for the sensor subsystem;
PFD_L is the average probability of failure on demand for the logic subsystem; and
PFD_{EE} is the average probability of failure on demand for the final element subsystem.

I. MODELING OF THE PROBABILITY OF SYSTEM FAILURE

The FTA method describes analytical technique, which is used to assess the failure probability, respectively the reliability of the systems. It is possible to use it in many areas, mainly in the field of risk control, quality management or safety management. It is usable as a preventive method for analysis of the existing problem or failure.
The FTA analysis (failure tree analysis) method is the most common used method to assess the reliability of the systems. However, it is not a complex method. This procedure provides a quick, orderly and transparent description of potential failures within the system, which may lead to a pre-defined adverse event. These properties make this method perspective.

In our case, the FTA method is based on the investigation of reliability of complex systems. This method is based on the given key issue – dangerous phenomenon, which is decomposed to partial problems, events and initial phenomena.

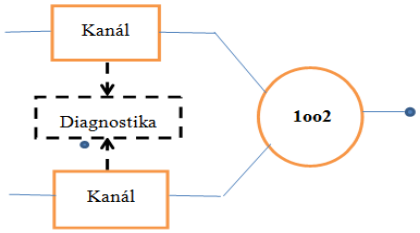
The failure state tree monitors the failure from its symptoms to the cause and provides information about the failure causes on various levels of integration. The analysis of the results of failure state tree modeling method allows investigating the failure causes.
Modeling of the probability of the elementary events occurrence leads to the increase of the reliability of the system, because it specifies the likely occurrence of the analyzed key problem. Based on the results of the analysis it is possible to reduce the probability of the dangerous phenomenon occurrence. The advantage of this method is, that the results can be deducted from the quantitative assessment of the probability of failure events. For example the analysis of the failure tree of the safety-critical system is a function of time.
The realization of the required safety on each control level assumes fulfillment of the whole set of hardware and software requirements. The selection of the safety subsystems architecture determines the reliability of the system in the process of the system design. The standard safety subsystems architectures and computation methods for determining the failure intensity are listed in the standards IEC 61508 a IEC 61511. These standards contain information tables with the results of these computations for selected parameter values. When realizing complex safety systems, the computation of reliability is a problem because of the following reasons:

- each subsystem has its own set of parameters;
- the computation method according to the standard IEC 61508 assumes the determination of minimal architecture reliability at the end of the time interval between tests T_1;
- the method FTA does not take into account, that the reliability of the subsystem is a function of time in the interval between tests. This is why there exists the probability of their "resonance".

The design of the dynamical computation models of standard architectures gives the possibility to model safety architecture of the whole control system. The base of the design is the failure tree according to the FTA method. The result of the modeling process is the time diagram of changes of the whole system. The analysis of the time diagram of the reliability changes will help to determine the values of the maximal probability of failures and design optimizing change of the diagnostic tests plan.

We will use the architecture **1oo2** (fig.2) as a example of the design of the dynamical model SCS. The basic principle of this architecture is the parallel connection of two channels. The safety function can be executed by each channel individually.

Figure. 2. Block scheme of the 1oo2 architecture



In the case of the failure of both channels, the failure of the safety function occurs. [].
The equivalent average time of the channel shutdown t_{CE} is

$$t_{CE} = \frac{\lambda_{DU}}{\lambda_D}\left(\frac{T_1}{2} + MRT\right) + \frac{\lambda_{DD}}{\lambda_D} MTTR \qquad (11)$$

The equivalent average time of the voting group channels shutdown with the equal system shutdownt_{GE}:

$$t_{GE} = \frac{\lambda_{DU}}{\lambda_D}\left(\frac{T_1}{3} + MRT\right) + \frac{\lambda_{DD}}{\lambda_D} MTTR, \qquad (12)$$

Where
λ_{DU} is dangerous Undetected failure rate (per hour) of a channel in a subsystem;
λ_D is dangerous failure rate (per hour) of a channel in a subsystem, equal to 0,5 (assumes 50 % dangerous failures and 50 % safe failures);
λ_{DD} is detected dangerous failure rate (per hour) of a channel in a subsystem (this is the sum of all the detected dangerous failure rates within the channel of the subsystem)

ISBN: 978-1-61804-327-6

T_1 is proof test interval (hour);
MRT is mean repair time (hour);
$MTTR$ is `mean time to restoration (hour).

The average probability of the failure on the request PFD_G for the architecture **1oo2** is

$$PFD_G = 2((1-\beta_D)\lambda_{DD} + (1-\beta)\lambda_{DU})^2 t_{CE} t_{GE} + \beta_D \lambda_{DD} MTTR + \beta \lambda_{DU} \left(\frac{T_1}{2} + MRT\right) \quad (13)$$

where
PFD_G is average probability of failure on demand for the group of voted channels;
β_D is of those failures that are detected by the diagnostic tests, the fraction that have a common cause (expressed as a fraction in the equations and as a percentage elsewhere)
β is the fraction of undetected failures that have a common cause (expressed as a fraction in the equations and as a percentage elsewhere)

Then the mathematical model of the probability of the failure of the architecture **1oo2** SCS in the time is

$$PFD_G(t) = 2\left((1-\beta_D)\lambda_{DD} + (-\beta)\lambda_{DU}\right)^2 \cdot \cdot \left(\frac{\lambda_{DU}}{6\lambda_D}\left(\int_0^{T_1} dt_{CE} + MRT\right) + \frac{\lambda_{DD}}{\lambda_D} MTTR\right)^2 + + \beta \frac{\lambda_{DU}}{2}\left(\int_0^{T_1} dt_{CE} + MRT\right) + \beta_D \lambda_{DD} MTTR \quad (14)$$

Mathematical models of every other typed architectures SCS **1oo1, 1oo2, 2oo2, 1oo2D, 1oo3, 2oo3**, were similarly designed for the mode of low request. The library of the typed SCS architectures was used in "SIMULINK" in the Matlab environment.

One of the library models mask (model of the architecture SCS **1oo2)** is in the fig.3. Scheme of the model is shown in the fig.4. The base of the simulation model is the discrete integrator with the integration time according to the testing period T_1 of the function SCS. One second of the simulation step corresponds to one hour of the testing period. For the testing purposes we are analyzing the failure tree in the process of the combustion of pollutants. (see fig.5). The tree consists of the messages SCS:

1. Failure of the burner combustion pollutants (SCS 1oo1, proof test interval T1 = 2190 h);
2. Failure of the cooling fan exhaust air (SCS 1oo1, proof test interval T1 = 4380 h);
3. The failure of the regeneration unit of xylen (SCS 1oo2, proof test interval T1 = 8760 h);
4. Failure of the temperature sensor (SCS 2oo3, proof test i

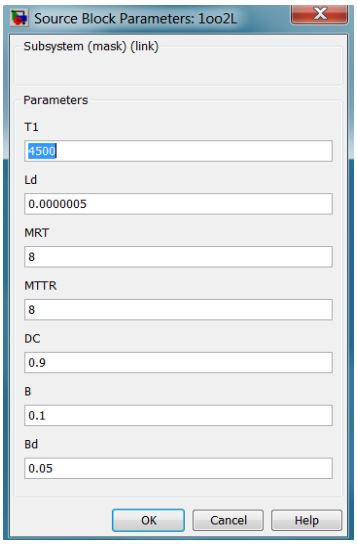


Figure. 3. The mask of the 1oo2 architecture model

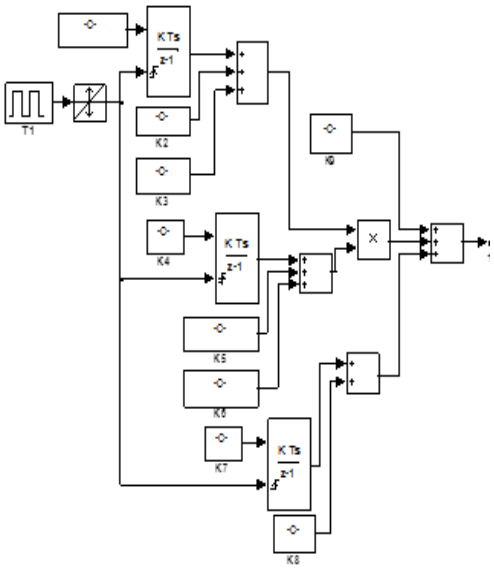


Figure. 4. The realization of the 1oo2 architecture model

In the fig.6 is the failure tree model of the combustion process and in the fig.7 is the time diagram of changes of the average probability of failure on demand of a safety function of the process PFD in 3 years

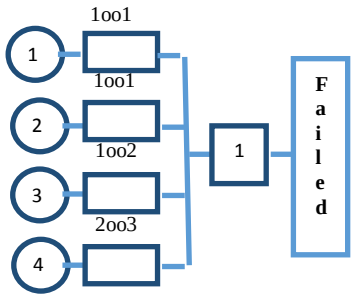


Figure. 5. Failure tree

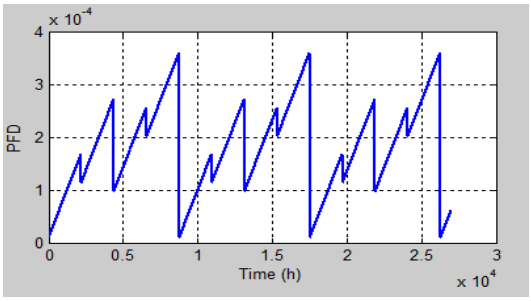


Figure.6. The average probability of failure on demand

TABLE 2. Parameters of the model

Event	T1(h)	λ	MRT	MTTR	DC	β	β_D
1	2190	0.0000005	8	8	0.9	-	-
2	4380	0.0000005	8	8	0.9	-	-
3	8760	0.000004	8	8	0.9	0.1	0.05
3	4380	0.0000005	8	8	0.9	0.1	0.05

The analysis of the time diagram shows, that the maximal probability that the process will be heard has a year interval (8760 hours). It is because the diagnostic times are regularly repeated. Therefore it is needed to exclude similar synchronization.

CONCLUSION

By the failure tree we can create a clear and systematic visual representation, from which it is clear, how the basic elements contribute to the failure rate of the system. This includes the search for the weak points and revelation of the aspects, which are important for the reliability. This method can be used for the qualitative and quantitative analysis and it allows fairly simple procedure, which is useful in the field of projecting and operating the technological processes.

REFERENCES

[1] Dr David J Smith, Reliability, Maintainability and Risk. Practical methods for engineers. Eighth Edition, Published by Elsevier Ltd, 2011, p.435.ISBN 978-0-08-096902-2

[2] K. Rastocny and J.Zdansky."Riadiace systemy so safety PLC". Zilinska univerzita v Ziline, EDIS-vydavatelstvo Zilinskej univerzity, 2013, ISBN 978-80-554-0681-7

[3] Failure in Safety-Critical Systems: A HANDBOOK OF INCIDENT AND ACCIDENT REPORTING. Chris Johndson. Publicity Services, No.2 The Square, University of Glasgow, Scotland. 2003, ISBN 0-85261-781-4..

[4] Joint Software System Safety Committee. SOFTWARE SYSTEM SAFETY HANDBOOK. A Technical & Managerial Team Approach. Joint Services Computer Resources Management Group, U.S. Navy, U.S. Army, and the U.S. Air Force. December 1999

[5] IEC-61508-6:Functional safety of electrical/electronic /programmable electronic safety-related systems. 2010.

[6] STN EN 60204-1: Safety of machinery – Electrical equipment of machines. Part 1: General requirements. 2007.

Network Proximity and Physical Web

Yousef Ibrahim Daradkeh
College of Engineering at Wadi Aldawaser
Prince Sattam bin Abdulaziz University, Saudi Arabia
daradkehy@yahoo.ca

Dmitry Namiot
Faculty of Computational Mathematics and Cybernetics
Lomonosov Moscow State University
Moscow, Russia
dnamiot@gmail.com

***Abstract*—** The Physical Web is a tool (an approach) to connect any physical object to the web. The Physical Web lets "render" physical objects in web, usually, with the help of mobile devices. This approach lets us navigate and control physical objects in the world surrounding mobile devices. There are different ways to enumerate physical objects. In this paper, we will describe the model based on the network proximity. In this model, the circulated information depends on the proximity to the network nodes in the wireless networks. We will discuss the possibility to use network proximity for interactions with physical objects.

***Keywords*—**Physical Web; network proximity; Bluetooth; Wi-Fi

I. INTRODUCTION

The Physical Web is a term that describes the process of presenting everyday objects in Internet. It aims to offer users the way to manage their daily tasks at using everyday objects that are soon to become smart and remotely controllable. This approach lets us navigate and control physical objects in the world surrounding mobile devices. Also, it helps perform everyday tasks depending on the surrounding physical objects. Of course, one of the first questions on this path is the way to enumerate physical objects.

One of the most often used approaches for physical objects markup is the deployment of wireless tags. Wireless tags can support standard protocols like Bluetooth and Wi-Fi. So, for mobile devices (mobile users) the detection of tags is actually the detection of wireless nodes. Note, that in this approach other mobile devices can play a role of tag too. And the network proximity here describes data models based on the detection of surrounding network nodes.

In this paper, we would like to discuss several approaches for building mobile systems based on the detection of physical objects via network proximity. Note, that the classical models for interaction with physical objects are a subject of Internet of Things (Web of Things) [1]. In this paper, we will mostly discuss the services which could be initiated by the presence of surrounding physical object. Such services do not always incur two-way data exchange with the physical objects. In the most cases, it is enough to detect and identify the object.

The proximity is a very conventional way for context-aware programming in mobile world. There are many practical use cases, where the concept of the location can be replaced by that of proximity. Proximity can be used as a main formation for context-aware browsers [2]. The context-aware browser will reveal data chunks depending on the current context.

The usage of network proximity for context-aware systems is very transparent. At his moment, network modules are most widely used "sensors" for mobile phones. All smartphones nowadays have Wi-Fi (Bluetooth) modules. So, Wi-Fi (Bluetooth) related measurements are included into standard interfaces of mobile operating systems. The above-mentioned measurements include the visibility for network nodes and signal strength. By the definition, the distribution for Bluetooth signal, for example, is limited. So, if any Bluetooth node is visible from a mobile device (a mobile phone, for example), then this device is somewhere nearby that node (it is so-called Bluetooth distance). The same is true for Wi-Fi access point. And this proximity information (network proximity) can replace location data. There are two main reasons for this replacement. At the first hand, we can target here all indoor application [3]. Obtaining GPS (Global Positioning System) data indoor is not reliable and sometimes even impossible. In the same time, modern offices usually have plenty of wireless nodes. The second reason is much more interesting. The wireless node could be moveable. So, our context information will follow to the moved object.

For network proximity-based context-aware applications, any existing or even especially created Bluetooth node could be used as a presence-sensor that can play the role of a trigger. This trigger can open access to some content, discover existing content, as well as cluster nearby mobile users [4,5].

The rest of the paper is organized as follows. In Section II, we discuss iBeacons. In Section III, we discuss Google Physical Web. In Section IV, we describe Bluetooth Data Points.

II. IBEACONS

The iBeacon is a wireless tag (beacon), based on Bluetooth Low Energy (BLE) standard [6]. Shortly, any beacon is set to transmit a set of numbers several times per minute, so that any mobile device with BLE support nearby can detect it. The beacon's repetitive transmission is called also as "advertising". The BLE standard specifies a structure for the

ISBN: 978-1-61804-327-6

data that must be transmitted. An application on a mobile device can then detect this parcel of information, unpack it, and use it for providing context-aware services.

The above-mentioned advertising includes a unique ID for a tag and two application-dependent numbers (so-called minor and major).

As per Apple's manual, a proximity universally unique identifier (UUID) is 16 Bytes, and major and minor codes are 2 Bytes each. The common usage for UUID is the identification for a place. For example, it could be a particular shop, café, etc. Major and minor codes could be used to a description of an area within a physical space associated with the above-mentioned UUID. For example, a retailer might use the major and minor code to identify, respectively, a given retail store and a specific shelf, where a beacon will be placed.

On iOS, a given application can scan for up to 20 tags (proximity UUIDs). It is, probably, one of the biggest limitations for iBeacons technology. The mobile application should statically declare UUIDs for the tags in questions [7]. For a mobile application, this declaration lets register to be notified if a Beacon with a given UUID comes within range (or goes out of range) of the device [8]. From the notification, a mobile application can obtain minor and major codes and they can then be used to uniquely identify a given beacon.

The application can then use this data, often (almost always) in tandem with a cloud service, to decide what action to take, if any, when the beacon is detected.

Beacons could be placed anywhere where potential users might wish to either trigger some form of action in a mobile application, or have that application log the fact that it came near to the beacon. For example, commuters in London are to be targeted with branded messages directly to their smartphones, as 500 buses in the capital are equipped with Bluetooth iBeacon technology [9].

There are legal and technical problems behind iBeacons. The legal problems are associated with the company Apple, who owns this technology. The main technological problem is the need for the static description of observer tags. Of course, the underlying system (iOS) can read data from all tags in the proximity, but dispatches only some of them to an application. It means that the only one company (Apple) has the whole picture.

Google comes with the own protocol for BLE [10]. Eddystone is the protocol specification that defines a Bluetooth low energy (BLE) message format for proximity beacon messages. It describes several different frame types that may be used individually or in combinations to create beacons that can be used for a variety of applications. At this moment, we can see the following frames (types of data) in the protocol:

Eddystone-UID: an opaque, unique 16-byte Beacon ID composed of a 10-byte namespace ID and a 6-byte instance ID. The Beacon ID may be useful in mapping a device to a record in external storage. The namespace ID may be used to group a particular set of beacons, while the instance ID identifies individual devices in the group. It is an analog for a minor/major pair in iBeacon from Apple.

Developers typically can use the namespace ID to signify own company or organization, so they know the owner for a beacon.

You can generate a namespace identifier with a UUID generator. But because standard UUIDs are 16-byte identifiers and namespace identifiers are only 10 bytes, we can simply drop the middle six bytes from the UUID. Google also prescribes a second technique a one-way hashing algorithm for generating a UID out of a URL. So you can algorithmically convert a domain name into a unique namespace ID. The instance identifier is meant to uniquely identify a specific beacon. Because the field is 6 bytes long, there are 2^{48} = 281 trillion combinations.

Eddystone-URL: a URL in a compressed encoding format. Once decoded, the URL can be used by any client with access to the Internet. It is a link to the Google Physical web, we will discuss below.

Eddystone-TLM frame broadcasts telemetry information about the beacon itself such as battery voltage, device temperature, and counts of broadcast packets. It contains the packet version (always a one-byte value of 0 for now), the beacon temperature (2 bytes), the beacon battery level (2 bytes), the number of seconds the beacon has been powered (2 bytes) and the number of "PDU" packet transmissions the beacon has sent (2 bytes.)

Actually, the general idea (pattern) is the same as with the "classical" iBeacons. Tags broadcast some ID, an application uses ID for getting data from the cloud. URL here is just a special case for ID. We can simulate URL transmission just by mapping tag's ID to some URL in the cloud-based datastore. Anyway, with obtained URL application should get access to the Internet for obtaining data.

Google provide Proximity Beacon API for setting attachment (data associated with) for BLE tags [11]. This API supports the following actions:

- Registering tags
- Publishing attachments to tags (associate data with tags)
- Retrieving attachments (data from tags)
- Monitoring beacons

Registered tag has got the following attributes:

- Advertised ID (required).
- Status.
- Expected stability.
- Latitude and longitude coordinates.
- Indoor floor level.
- Google Places API Place ID.
- Text description.

ISBN: 978-1-61804-327-6

Attachment is a string up to 1024 bytes long. It could be a plain string, JSON data or even encoded binary data. Attachments are stored in Google's scalable cloud.

There is also a very important remark for the development: on Android platform it is possible to obtain information about all "visible" tags.

Eddystone is a part of Nearby API [12]. Nearby uses a combination of Bluetooth, Wi-Fi, and inaudible sound (using the device's speaker and microphone) to establish proximity (Figure 1). Its implementation has been just announced

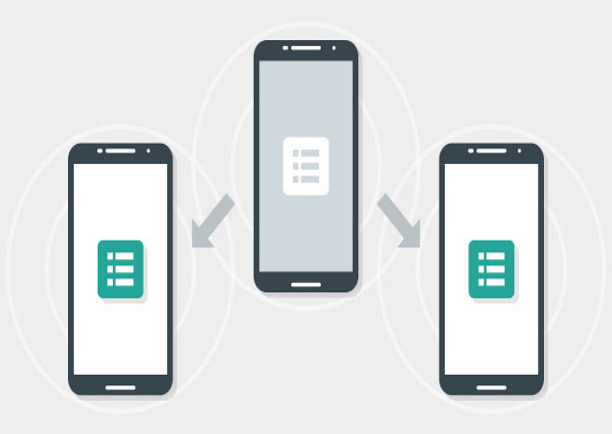


Fig. 1. Nearby API

III. Google Physical Web

Google Physical Web project is an example of integration Web technologies and physical world. Actually, it is part of a more generic problem: how to integrate Internet of Things and web technologies [13]. At its base, the Physical Web is a discovery service: a smart object broadcasts relevant URLs that any nearby device can receive. This simple capability can unlock exciting new ways to interact with the Web [14].

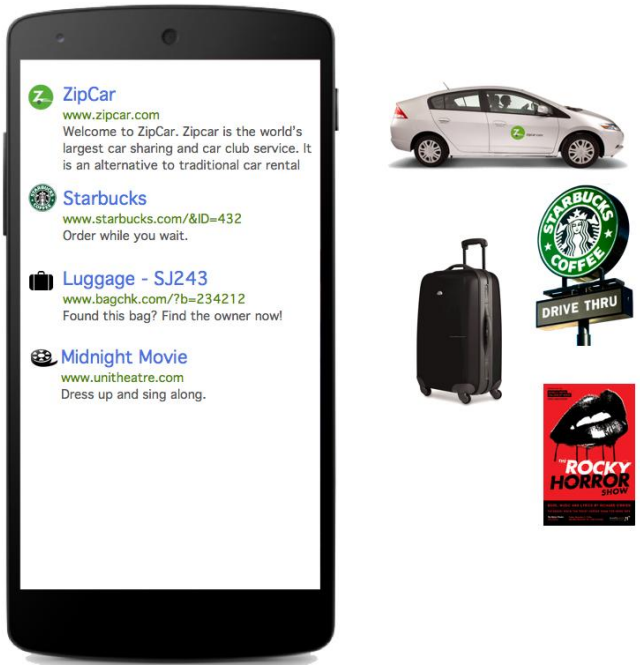


Fig. 2 The physical web

Figure 2 illustrates the basic idea. Actually, the physical objects here are (in the most cases) the same Bluetooth tags. In the current implementation, URL broadcast method involves a Bluetooth broadcast from each tag. The user's phone obtains this URL without connecting to the beacon. As per Google, this ensures the user is invisible to all beacons, meaning a user can't be tracked simply by walking past a broadcasting beacon. It is a very important principle, we would like to discuss separately. As per Google, this was very much by design to keep user's silent passage untrackable. But it is assumes also, that URL detection should be performed automatically, on the background. Again, as per Google's manual, once the user does click on a URL, they are then known to that website. With this solution, Google mostly follows to iBeacon usage (deployment model). Application on the mobile device automatically discovers nearby objects, obtains associated data (URLs in this case) and pushes this information to the user. It is a true push, iBeacons are using push notifications, supported by mobile OS [15]. Notification service is a popular functionality provided by almost all modern mobile OS (iOS, Android, etc). To facilitate customization for developers, mobile platforms support highly customizable notifications. And yes, the third-party push notification customization may allow an installed trojan application to launch phishing attacks or anonymously post spam notifications [16]. So, why do not switch to browsing mode instead of push notifications? Mobile applications may still obtain iBeacons data automatically, but show them only when a user directly requests them. It should like browsing. We must see the direct intention from mobile users to obtain nearby data. In this case, our application should form dynamically a web page (like CGI-script in the web) and show it to the user.

Technically, the Physical Web can use not only BLE as a transport layer. We can mention in this context UPnP

ISBN: 978-1-61804-327-6

technology, for example [17]. The UPnP (Universal Plug and Play) architecture offers pervasive peer-to-peer network connectivity of PCs, intelligent appliances, and wireless devices. The UPnP architecture is a distributed, open networking architecture that leverages TCP/IP and the Web to enable seamless proximity networking in addition to control and data transfer among networked devices [18]. The UPnP architecture defines a base set of standards and conventions for describing devices and the services they provide. It is designed to bring easy-to-use, flexible, standards-based connectivity to ad-hoc or unmanaged networks. In case of the Physical Web, the provided service is just an URL, associated with discovered device.

As the next possible solution, we can mention mDNS [19]. mDNS - multicast DNS service discovery, also known as zero configuration. It is an interface being used to announce and query services on the local network. Using mDNS allows a client to advertise the services of a given host without the direct help of a centralized DNS server. Again, the service here is just an URL.

We can mention here a very simple approach for creating the Physical Web for any Bluetooth/Wi-Fi device. We can define a SSID (name) for Wi-Fi access point (Bluetooth node) as some URL. SSID for Wi-Fi access point (Bluetooth node in the discoverable mode) is broadcasted (being advertised in terms of the Physical Web). Of course, this setup could be done programmatically. And mobile application (programmatically also) can get a list of available Wi-Fi access points (Bluetooth nodes in the so-called discoverable node). This list includes SSIDs (URLs). So, it is a typical Physical Web, even without BLE. This approach will work even without the dedicated tags. Wi-Wi access point (Bluetooth node in the discoverable mode) could be set programmatically (it is true for Android) right on the mobile phone. So, any smart phone could be turned into a Physical Web tag and provide advertising for some URL.

In the more generic form, this approach could be described as Beacon stuffing [20]. It is a low bandwidth communication protocol for IEEE 802.11 networks that enables Wi-fi access points to communicate with clients without association. This enables clients to receive information from nearby access points even when they are disconnected, or when connected to another access point. Originally, this scheme was developed for Wi-Fi as complementary to the 802.11 protocol. It works by overloading 802.11 management frames while not breaking the standard. The beacon-stuffing protocol is based on two key observations. First, clients receive beacons from access points even when they are not associated to them. Second, it is possible to overload fields in the beacon and other management frames to embed data. Access point embeds content in Beacon and Probe Response frames, while clients overload Probe Requests to send data. By the similar manner, this scheme will work for Bluetooth [21]. And of course, for the Physical Web we do not need the two-way communication.

Actually, the beacon-stuffing was the inspiration point for Bluetooth Data Points.

IV. Bluetooth Data Points

Bluetooth Data Points (BDP) [22] let us turn any Bluetooth node into tag. The main idea behind BDP is to associate some user-defined data with existing (or even especially created) wireless networks nodes. Originally, the project targets Bluetooth nodes in the discoverable mode, but the same principles will work for Wi-Fi access points too. This association is similar to the above-mentioned data attachments for beacons. The main difference is the definition (the description) for a tag. BDP is based on the idea of "zero scene preparation". For example, any mobile users should be able to create (open) Bluetooth node right on the own mobile phone, associate some data with this node and so, make them available for other mobile users in the proximity. Figure 3 illustrates this idea. As existing node, we see here Bluetooth node in the car. Many modern cars nowadays are actually Bluetooth nodes. Car's owner can attach data to the own node. Other mobile users in the proximity can "see" Bluetooth node and use its identification (SSID, MAC-address) as key for obtaining associated data from the cloud. It is so-called hyper-local data concept. Data not only present some local information but could be prepared locally also. Instead of the car (Bluetooth node in the car) we can use just another mobile phone. A Bluetooth node (a tag) could be created programmatically. And programmatically we can attach some data to it. So, just one mobile application (in publishing mode) is enough for creating a new data channel. And the same mobile application (in browsing mode) could be used for reading data in the proximity.

Fig. 3. BDP data flow.

The simplest use case is a mobile classified. A mobile users creates an advertising (announce), links it to the wireless node on the own mobile phone and so, it becomes available for reading for other mobile users in the proximity. If the mobile phone (the mobile tag) is moved, all associated data will be "moved" too. Data are not associated with latitude/longitude pair (as in geo-location systems), but with ID of wireless node. Data are visible in the proximity of the node (in the proximity of the author) only.

Google does not describe the above mentioned database for beacons data attachments. BDP uses the classical key-value model for data. Attachments area described individually for the each Bluetooth point. So, there is a key (MAC-address) and a JSON text with linked data. It is a typical key-value data model. This data model is one of the most used models for NoSQL approach. One of the available examples of Open Source key-value stores is Apache Accumulo [23]. It is distributed key-value store. Actually, the whole database for BDP is a distributed hash table. The table rows as key-value pairs to provide a fast way to look up by a key item as attribute given by the value of a column qualifier of a row. In order to support lookups by more than one attribute of an entity, additional indexes can be built.

Data is represented as key-value pairs, where the key is comprised of the following elements: *RowID*, *Column* (*Family*, *Qualifier*, *Visibility*) and *Timestamp*. All elements of the Key and the Value are represented as byte arrays except for *Timestamp*, which is a *Long*. Accumulo sorts keys by element and lexicographically in ascending order. Timestamps are sorted in descending order so that later versions of the same Key appear first in a sequential scan. Tables consist of a set of sorted key-value pairs [24].

In terms of data design, BDP store contains the following information:

(recordID, MAC_address, data_array)

Each record describes a one data chunk (information element) for the given (MAC_address) Bluetooth node. Of course, we could have more than one information element for the same node.

The typical query requests data by MAC_address. So, it is a direct scan via the primary index and it will be fast.

JSON array for data chunks let present the various elements within data attachments. For example, plain text, phone number, email address, hyperlink, link to Twitter/Facebook/Linkedin profile or even an encoded image.

The basic algorithm as it is described above is based on the ideas of browsing data rather than push them to mobile users. BDP's context-aware "browser" obtains a list of the visible Bluetooth node. Than for the each node browser can perform database scan (lookup) and get data associated with this node. This request simply returns nothing in the case of Bluetooth nodes without attachments. All collected data could be packed in JSON array and this array will be returned back to the "browser". And the browser will perform data rendering. Nodes in the array could be sorted by the obtained RSSI (signal strength). Figure 4 illustrates this.

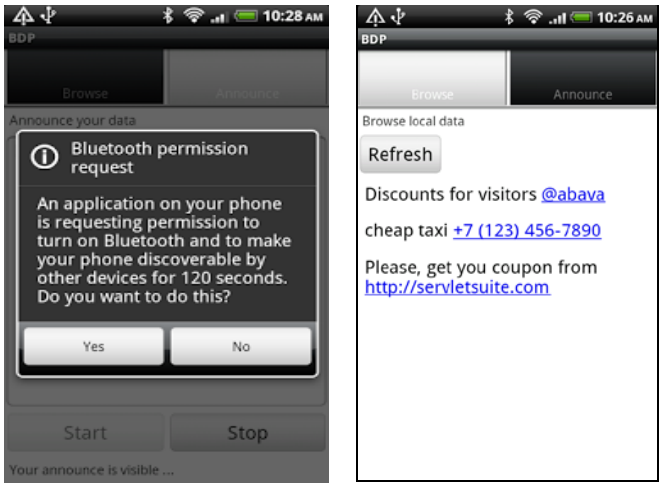


Fig. 4. BDP

As per collected statistics, the system can accumulate "browsing" events. An event here is the fact states that the device with address MAC_1 requests a data chunk provided by the device MAC_2 at the time *t*.

In the normal case, most of the nearby Bluetooth nodes will be "empty" (they will be out of BDP circle). So, we can decrease the number of database lookups with some cache. BDP uses a Bloom filter [25]. The Bloom filter is a method for representing a set of n elements $A = \{a_1, a_2,\ldots,a_n\}$ also called keys to support membership queries. Elements here are MAC-addresses for Bluetooth nodes.

V. Conclusion

In this paper, we discuss existing and upcoming software systems based on the network proximity. As a main result, we can present our list of requirements to the flexible solution, based on the wireless tags.

Any proposed system should support software-based tags. It should be possible to define tags and linked data with existing wireless infrastructure and/or existing mobile devices. We do not reject the idea of using dedicated hardware tags. We just highlight the fact that software-based systems are much more flexible, cheaper and finally allow much more services.

The wireless modules (Bluetooth, Wi-Fi) in the mobile devices make them the most popular and widely distributed sensors. With software based wireless tags, it is a most simple and convenient approach for context-aware programming in Smart Cities environments.

The push-based data delivery in case of wireless tags has got serious usability limitations. By our opinion, the browsing is a more promising approach for getting data in the proximity of tags. Finally, we can predict, that at the end of the day, mobile browsers will incorporate data about proximity. Any mobile browser is a mobile application too. And nothing prevents it, for example, to scan nearby wireless devices.

In the "browsing" mode collected statistics about data scanning is a direct analogue of web log, collected by any web server. This statistics is an important part of data mining for

ISBN: 978-1-61804-327-6

analyzing the behavior of mobile users and should be collected by beacons supporting frameworks.

Acknowledgment

We would like to thank prof. Manfred Sneps-Sneppe for the valuable discussions.

References

[1] Namiot, Dmitry, and Manfred Sneps-Sneppe. "On IoT Programming." International Journal of Open Information Technologies 2.10 (2014): 25-28.

[2] Namiot, Dmitry. "Network Proximity on Practice: Context-aware Applications and Wi-Fi Proximity." International Journal of Open Information Technologies 1.3 (2013): 1-4.

[3] Namiot, Dmitry. "On Indoor Positioning." International Journal of Open Information Technologies 3.3 (2015): 23-26.

[4] Namiot, Dmitry, and Manfred Sneps-Sneppe. "Geofence and network proximity." Internet of Things, Smart Spaces, and Next Generation Networking. Springer Berlin Heidelberg, 2013. 117-127.

[5] Sneps-Sneppe, Manfred, and Dmitry Namiot. "Spotique: A new approach to local messaging." Wired/Wireless Internet Communication. Springer Berlin Heidelberg, 2013. 192-203.

[6] Gomez, Carles, Joaquim Oller, and Josep Paradells. "Overview and evaluation of bluetooth low energy: An emerging low-power wireless technology." Sensors 12.9 (2012): 11734-11753.

[7] Cavallini, Andy. "iBeacons Bible 1.0." Online im Internet: http://meetingof ideas. files. wordpress. com/2013/12/ibeacons-bible-1-0. pdf 22 (2014).

[8] What are iBeacons? http://radar.oreilly.com/2015/04/what-are-ibeacons.html Retrieved: Jun, 2015

[9] Hundreds of London Buses to be Fitted with iBeacons http://performancein.com/news/2015/07/01/hundreds-london-buses-be-fitted-ibeacons Retrieved: Jul, 2015

[10] Eddystone https://github.com/google/eddystone Retrieved: Jul, 2015

[11] Proximity Beacon API https://developers.google.com/beacons/proximity/guides Retrieved: Jul, 2015

[12] Nearby API https://developers.google.com/nearby/ Retrieved: Jul, 2015

[13] Want, Roy. "The Physical Web." Proceedings of the 2015 Workshop on IoT challenges in Mobile and Industrial Systems. ACM, 2015.

[14] The physical web https://github.com/google/physical-web Retrieved: Jul, 2015

[15] Hansen, J., Gronli, T. M., & Ghinea, G. (2012). Towards cloud to device push messaging on Android: technologies, possibilities and challenges.

[16] Xu, Z., & Zhu, S. (2012, August). Abusing Notification Services on Smartphones for Phishing and Spamming. In WOOT (pp. 1-11).

[17] Miller, B., Nixon, T., Tai, C., & Wood, M. D. (2001). Home networking with universal plug and play. Communications Magazine, IEEE, 39(12), 104-109.

[18] UPnP, http://www.upnp.org, Retrieved: Jul, 2015

[19] Nordman, B., & Christensen, K. (2010). Proxying: The next step in reducing it energy use. Computer, 43(1), 91-93.

[20] Chandra, R., Padhye, J., Ravindranath, L., & Wolman, A. (2007, March). Beacon-stuffing: Wi-fi without associations. In Mobile Computing Systems and Applications, 2007. HotMobile 2007. Eighth IEEE Workshop on (pp. 53-57). IEEE.

[21] Banerjee, N., Agarwal, S., Bahl, P., Chandra, R., Wolman, A., & Corner, M. (2010). Virtual compass: relative positioning to sense mobile social interactions. In Pervasive computing (pp. 1-21). Springer Berlin Heidelberg.

[22] Namiot, D., & Sneps-Sneppe, M. (2014, October). CAT??? cars as tags. In Communication Technologies for Vehicles (Nets4Cars-Fall), 2014 7th International Workshop on (pp. 50-53). IEEE.

[23] Sen, R., Farris, A., & Guerra, P. (2013, June). Benchmarking apache accumulo bigdata distributed table store using its continuous test suite. In Big Data (BigData Congress), 2013 IEEE International Congress on (pp. 334-341). IEEE.

[24] Namiot, D., & Sneps-Sneppe, M. (2015). On Mobile Bluetooth Tags. arXiv preprint arXiv:1502.05321.

[25] Broder, A., & Mitzenmacher, M. (2004). Network applications of bloom filters: A survey. Internet mathematics, 1(4), 485-509.

Framework Design for Statistical Fraud Detection

A.A. Ojugo., A.O. Eboka., R.E. Yoro., M.O. Yerokun and F.N. Efozia

***Abstract*—**An unstable economy is rife with fraud. Perpetrated on customers, it ranges from employees' internal abuse to large fraud via high-value contracts cum control breaches that impose serious consequences to biz. Loyal employees may not perpetrate fraud if not for societal pressures and economic recession with its rationalization that they have bills to pay and children to feed. Thus, the need for financial institutions to embark on effective measures via schemes that will aids both fraud prevention and detection. Study proposes genetic algorithm trained neural net model to accurately classify credit card transactions. Compared, the model used a rule-based system to provide it with startup solution and it has a fraud catching rate of 91% with a consequent, false alarm rate of 9%. Its convergence time is found to depend on how close the initial solution space is to the fitness function, and for the recombination and mutation rates applied.

***Keywords*—**Classifiers, fitness function, stochastic, mutation, recombination, products, credit-card, transactions

I. Introduction

FRAUD is illegal acquisition of valuable data or resources via willful misrepresentation. Others view it as a criminal act involving larceny, theft and embezzlement. Legally, it is a state where a criminal makes a false, material statement and an unsuspecting victim rely such a statement; while, the criminal benefits from the entire process (Ojugo et al, 2014). It is perpetrated mainly by internal member of an organization and also those external to it. Fraud benefits an organization, its part or internal/external persons to the organization (Marane, 2011) and it has been classified into types as in fig. 1; We note also that fraud involves all products of financial transactions, and grouped as in fig 2 broadly into: (a) ***cheque/check*** fraud involves use of counterfeit or altered checks via forgery of an account holder signature, and (b) C***ybercrimes*** involves use of ***key logging*** (here, keys struck as a user gains access into his system is tracked in a covert manner so that the user is unaware that his action are monitored), and ***hacking*** (intruder gains access of user's PC without their permission). They gain access via a user's USB or the user is a victim of a virus attack via web-surfing, email attachments and/or online purchase. Example of malware techniques includes (Ojugo et al, 2014):

a. **Trojan-Horse** is a form of malware (virus), hidden in a file, program, free online games downloaded also called shareware, hyper- or related links, emails attachments and screensavers.
b. **Spyware** collects bits of user data at a time without the knowledge of an unsuspecting user. Its presence is hidden and difficult to detect. It is prevented by noting the sites visited, and reading emails sent before following any link.
c. **Phishing** attempts to acquire sensitive/confidential details from account holders, masquerading as trustworthy entity in an electronic communication. Perpetrated via online-payment, social and auction sites, phishing lures users via emails or instant messaging into fake site with deceptive login and hyperlink to real sites. Similarly, voice phishing (***vishing***) are attacks in which account holder is contacted mainly by phone, to check if their account status is been compromised. Rather than refer to a site, the unsuspecting customer is redirected to call a toll-free number.

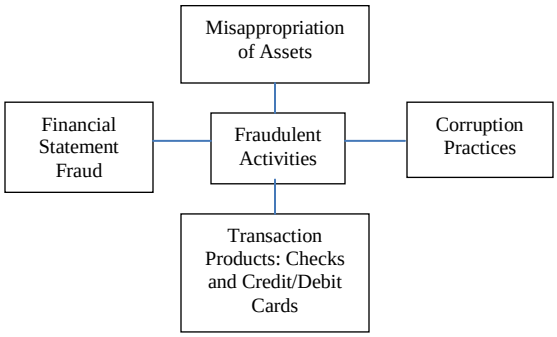


Fig. 1: Types of Fraud and Fraudulent Activities

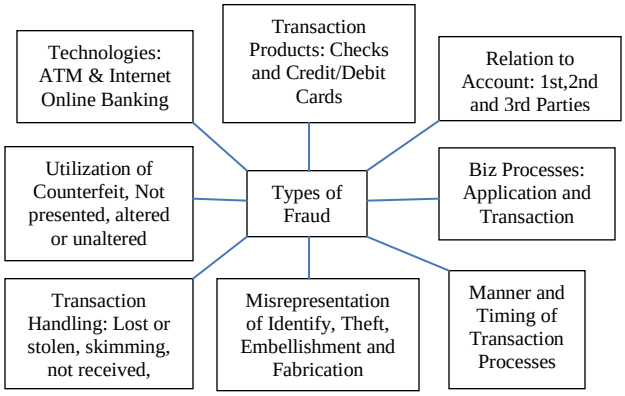


Fig. 2: Forms of Transaction Product Fraud (Delamaire et al, 2009)

A. Fraud Prevention and Detection

When financial institutions lose money to fraud, credit card (account) holders (partially, if not wholly) pay for such lose via reduced benefits, higher interest rate and membership fees. It is to the interest of both the financial institution and her customers to reduce fraud on transaction products. This makes imperative and imminent that financial institutions embark on early detection methods, to savor the situation.

Fraud prevention describes measures to stop the occurrence of fraud in the first place. Measures include elaborate designs, fluorescent fibers, multi-tone drawings, watermarks, laminated metal strips, holographs on banknotes, personal identification numbers for bankcards, security for credit card transactions, password on PCs, Subscriber Identity Module card for mobile phones and telephone bank accounts. With no perfect system, a tradeoff is struck between its inconvenience and expense to a customer and effectiveness of the system to the entire financial process. Thus, fraud detection emerges if its prevention fails, and it involves ***techniques*** for identifying fraud as quick as possible, once it has been perpetrated. In practice, detection is

A.A. Ojugo: Dept. of Math/Computer Sci., Federal University of Petroleum Resources, Effurun, (ojugo.arnold@fupre.edu.ng)

used continuous, as we may be unaware when our prevention technique in place, has failed (Bolton and Hand, 2002).

No matter how carefully we guard/prevent credit card fraud and not beyond mistakes, our confidential details may again be stolen or lost, to end in the wrong hands. Then, we must be able to detect, as soon as possible, that and when fraud is being perpetrated. Fraud detection is a continuously evolving discipline as intruders evolve their strategies to hack into a known detection system. With new criminal minded intruders constantly entering the field, many will however be unaware of previous fraud detection methods that have been successful as they adopt varying strategies. This will however, lead to identifiable frauds and implies that both the earlier and latest detection tools need to be applied always. This also implies that an up-to-date is required to monitor and detect intrusion (Delamaire et al, 2009; Bolton and Hand, 2002).

B. Statistical Fraud Detection

Various statistical methods have been successfully employed in fraud detection as tools. Though diverse in their different applications in both size and type, they have common themes. These tools aim to compare the observed (***historic***) data with its expected (***computed***) data-values. The expected value can be derived cum interpreted in varied ways, depending on the context as they may be single numerical summaries of some aspect of a dataset (feats) behavior with often simple graphical summaries in which an anomaly is readily apparent. They can also be shown as a more complex (multivariate) behavior of data profiles. Such profiles are based on past behavior of the system understudied (e.g. how an account has been previously used), or be extrapolated from other similar systems (Bolton and Hand, 2002). Events are further complicated in that, a given user may behave in a fraudulent manner at some point and not at other times (Delamaire et al, 2009) resulting to true-negatives and false-positive results in fraud detection.

Statistical fraud detection methods may be ***supervised*** or ***unsupervised***. In supervised, samples of both fraudulent and non-fraudulent dataset records are used to construct the model to allow it classify new observations into one of two classes. It requires precise and distinct feats of the original dataset used to build the model, for its classification of the observations into both (***true***) classes as well as examples of both classes. Thus, it can only be used to detect frauds of a type which have previously occurred. Conversely, ***unsupervised*** mode simply seeks those accounts, customers and so forth which are most dissimilar from the norm. These are further examined to detect outliers. Outliers are a basic form of non-standard observation. Such tools will aim at checking data quality as well as separate between false-positives (detection of accidental errors) from true-negatives (detection of deliberately falsified data or data that accurately describe fraudulent pattern). Statistical method alone cannot ascertain if fraud is committed. Such analysis should be viewed as alerting us to anomalous observations that are more likely to be fraudulent than others, so that it can then be investigated in more detail. Thus, the objective of the statistical analysis can be seen as to return a *suspicion score* (higher score is more suspicious than a lower score). Since there are varied ways in which fraud is perpetrated, there also exists many different ways to compute suspicion scores (Bolton and Hands, 2002).

Suspicion scores is computed for each record in database (for each customer with a bank account or credit card, for each owner of a mobile phone, for each desktop computer and so on), and these can be updated as time progresses. These scores are then ranked and investigative attention focused on those with highest scores or on those that exhibit a sudden increase. With cost as an issue here, given that it is too expensive to undertake a detailed investigation of all records, the algorithm investigates only those thought most likely to be fraudulent.

The main ***idea*** here is to employ hybrid genetic algorithm trained neural network model with rule-based preprocessor in computing the suspicion score, as means for fraud detection. This method will seek a solution to this complex and dynamic task, for which conventional method do not yield complete, cost-effective solutions. It analyzes historic data, investigating data features of interest by computing statistically-dependent underlying probabilities between observed and predicted data, to simulate tractability, low-cost, robustness and effective results, with high tolerance to uncertainties, ambiguity, partial truth, imprecision and noise that may have been applied along with its input data.

II. Statement of Problem

The problem statements are as follows:

1. Exchange of idea in fraud detection is often limited since it is unwise to describe in public domain (fraud detection techniques in great detail). It will further equip intruders with adequate information required to evade detection. Thus, we employ statistical fraud detection method and heuristics as in Section III.
2. The complex, dynamic and chaotic nature of fraud, and its range of complications as providing a backdoor to allow for other crime makes imperative and critical, early and accurate detection**.** Supervised detection alone via careful monitoring and management of network is insufficient as intruders often evade such as it often yields inconclusive results for ***unknown*** inputs. This leads to increased rate of false-positives and true-negatives. Our proposed model will effectively classify fraudulent from non-fraudulent activities using soft-computing heuristics as in Section IV.
3. Unavailability of fraud datasets and its censored results – makes fraud detection techniques and studies difficult to assess. Dataset also consist of ambiguities, imprecision, noise and impartial truth that must be resolved via robust search in the bid to classify observations and expected values effectively as in Section III and IV.
4. Classification (as resolved in Section IV) using predictive models is a complex and difficult task due to the chaotic and dynamic nature of ***unsupervised*** evolutionary models. The model also need to resolve effectively and efficiently, statistical dependences and conflict imposed on the model by dataset used in approximating the data feats of interest.
5. Use of hill-climbing methods often has speed constraint imposed on it as the solutions are often trapped at local maxima. This is resolved with hybridization of statistical methods as in Section III/IV. Also, search for optimal via evolutionary heuristics can be quite cumbersome (though no one method yields better optimal than hybrids). Model must also resolve the statistical dependencies imposed on

it by hybridization. The proposed model resolves this in Section III and IV.

6. Search for optimal solution, may also yield overtraining and over-fitting of the model (resolved in Section IV) as it aims to find underlying probability of the data feat(s) of interest. Also, improper selection of parameters may also lead to over-parameterization (resolved in Section V).
7. Some model aim at a single suspicion score to globally classify statistical fraud. Studies show however, that some cases may be a result of true-negatives and false-positives scores as resolved in Section IV.

Proposed genetic algorithm trained neural network will seek to use unsupervised (improved) classification method that will help propagate observed data in model as it seeks data feats of interest to yield an output (Ojugo et al, 2013). Evolutionary models have been successfully applied to enhance accurate prediction in its search for optimal solution, chosen from a set of possible solution space, to yield an output that is guaranteed of high quality and void of ambiguities. These models, further tuned can become robust and perform quantitative processing to ensure qualitative knowledge and experience, as its new language (Heppner and Grenander, 1990).

III. The Proposed Hybrid Model

Proposed model hinges on 3-basic frameworks as in fig 1 – further explained as:

A. Rule-Based Preprocessor

We employ the rule-based system for 3 reasons: (a) it serves as benchmark to measure how well other heuristics perform on comparison, (b) a simplified version yields a sensible solution and generation of rules in the model; Rather, than choosing completely random points swaps and mutations. Thus, greatly improves the proportion of moves accepted, and (c) used as preprocessor to other models, it yield a good starting solution with minimized false-positive cum true-negatives. The rule-based model consists of heuristics and conventional recursion routines to assist in carrying out fraud detection classification, as suited for the problem domain at hand where the following holds (Saleh Elmohamed et al, 1998; Michalewicz, 1998):

a. The Account Data Structure of each account holder in an institution to hold values of each account holder.
b. The Card data structure of cards issued and their types, capable of being used across E-Tranzact and InterSwitch platforms for online banking.
c. Data structure for time periods to keep track of which, when and what transaction took place.
d. Fraud data structure of each time-stamped data containing a knowledgebase of historic data classified into fraudulent and non-fraudulent transactions so as to minimize true-negatives and false positives.

The model's basics function is: given data files of fraudulent and non-fraudulent classes, user account details, time-stamped transactions for account holders, account holders transaction-to-credit card details matrix, the inclusion data that allows for further classification of fraudulent and non-fraudulent dataset. Even when administrator maintains transactions in time, it should reflect classified as it has been noted that fraud results mainly from internal breaches, employees and tips. Using these structures above, the system builds an internal database (knowledgebase) to perform the computation of the suspicion score as well as classification of transaction into classes. It involves a number of essential sub-processes such as checking all time-stamped transaction, entry access point by the account holder as well as conflict resolution that arise in transactions (such as rollback to debit an account holder via online access from credit card) amongst other transaction processes (Ojugo et al, 2014).

The rule-based model is an iterative approach that allows rule generation as its basic procedure for each move as thus. Model first scans through all currently unassigned transactions by choosing the transactions in order of account holder's details, the transaction type and then computes suspicion score for each transaction sampled. It then attempts to assign these transactions between fraudulent and non-fraudulent classes by satisfying the rules governing their consequent classification. Only transactions that satisfy all the rule constraints can be classified. Model then searches through classified (fraudulent and non-fraudulent) knowledgebase of transaction and selects those with higher suspicion score. Selecting threshold values for defining what is considered a high score for each case is a subjective procedure (though, it is also a more straightforward approach to choose a reasonable value). Thus, the rule-based model yields a partial output, as it is unable to assign the given transactions into their varying classes (Ojugo et al, 2014).

B. Genetic Algorithm (GA)

GA is inspired by Darwinian genetic evolution (survival of fittest) consists of population (data) chosen for selection with potential solutions to a specific task. Each potential solution is an individual for which optimal is found using four operators: initialize, select, crossover and mutation (Coello et al, 2004 and Reynolds, 1994). Individuals with genes close to optimal are fit. Its fitness function determines how close an individual is to the optimal solution. Its operators are (Ojugo et al, 2012):

a. Initialize encodes data into forms suitable for selection. Each encodings type used has its merit. Binary encoding is computationally more expensive. Decimal encoding has greater diversity in chromosome and greater variance of pools generated. Float-point encoding or its combination is more efficient. The *fitness* function evaluates how close a solution is to its optimal – after which they are chosen for reproduction. If solution is found, function is *good*; else, *bad* and deselected for crossover. Fitness function is the only part with knowledge of the task, and if more solutions are found, the higher its fitness value.
b. Selection: With best fit data chosen to mate, the larger the number selected, the better the chances of yielding fitter data. This continues until one is chosen, from the last 2 or 3 remaining solutions, to become the selected parents of prospective new offspring. Selection ensures fittest data is chosen to mate. The selection that only mates the fittest is an *elitist* and often leads to converging at local optima.
c. Crossover ensures best fit data (genes) are exchanged to yield a new and fitter pool. There are two crossover types (depends on encoding type used): (a) ***simple*** crossover for binary encoded pool, which allows single- or multi-point cross with all genes from single parent, and (b) ***arithmetic***

crossover allows new pool to be created by adding a data or individual's percentage to another.

d. Mutation alters chromosomes by changing its genes or its sequence, to ensure new pool converges to global minima (instead of local optima). Algorithm stops if optimal is found, or after number of runs if new pools are created (though computationally expensive), or when no better solution is found. Genes may change based on probability of mutation rate. Mutation improves the much needed diversity in reproduction and its algorithm is as thus:

Cultural GA (a variant) has belief space as thus: (a) Normative (has specific value ranges to which data is bound), (b) Domain (has data about task domain), (c) Temporal (has data about events' space is available), and (d) Spatial (has topographical data). In addition, an influence function mediates between its belief spaces and the pool – to ensure altered data conforms to the belief space. Thus, the data pool does not violate its belief space. This helps reduce number of possible individuals GA generates till an optimum is found (Reynolds, 2004).

C. Artificial Neural Network (ANN)

Inspired by human brain, an ANN consists of interconnected neurons that learn by example. Its neurons share data signals by adjusting its weight/bias (as connection strengths between synapses, axons and dendrites), which are summed, depending on the task by an activation function to yield an output that modulates its inputs and nonlinear feats exhibited as in Eq. 1 (Ojugo et al, 2012):

$$\emptyset = f(net) = f\sum_{i=1}^{m} Xi * Wij \qquad (1)$$

Thus, it translates into mathematics, principles of biological processing so as to generate predictive evolution outcomes of a task in the fastest time. Its derives outcomes from experience and recognizes behavior(s) as feats of interest from historic data, in order to suggest an optimal solution of high quality, irrespective of modification introduced to it by other methods in such a multi-agent space, which constantly affects the quality of any solution (Dawson and Wilby 2001). Encoded as a 3-layered unit of input, hidden and output, it is configured into ***feedforward*** and ***recurrent*** net based on task, data feat(s) to be approximated and connection requirements. Its learning allows a trained net to effectively classify patterns based on the employed learning rule into supervised, unsupervised and reinforcement (Ojugo et al, 2013).

The nature of fraud detection requires previous knowledge. Thus, we adopt the recurrent Jordan net to help us incorporate historic dataset feats of interest and previous output to be fed back as input into the model's hidden units to yield the next output. Its correlated weights are interconnected so that $W_{i,j}$ is weight between input and hidden layers, $W_{o}j$ is bias and x_i is diabetes input data. Its output is generated via tangent/sigmoid transfer function, which sums weighted input as in Eq. 2 and Eq. 3 (Minns, 1998). To resolve statistical dependency in the model's structure imposed by the data and heuristics adopted, our network used its ability to store earlier data generated from previous layer(s) as in Kuan and White (1994).

$$Z_{ij} = w_o j + \sum_{i=1}^{m} x_i * w_i j \qquad (2)$$

$$F(Z_{ij}) = \frac{2}{1 + e^{-2*Zij}} - 1 \qquad (3)$$

The Jordan network is more plausible and computationally more powerful than other models. Its back-propagation in time algorithm allows for advanced training/learning so that output at time *t* is used along with a new input to compute its output at time *t+1* in response to dynamism (Mandic and Chambers, 2001). It computes its output y^k, using the Tansig function that sums all inputs, receives target value of input training pattern and computes its error data, which is then sent back from its output to input nodes via error back-propagation to correct and update its weights and biases (c_j^k and c_o^k respectively) using mean square error. It then finds weights that approximate the target output with selected accuracy and modifies its weights by minimizing the error between ***target*** and ***computed*** outputs at end of each forward pass. If error is higher than selected value, process continues with reverse pass; else, stop training. Weights are updated till minimal error is found (Ojugo et al, 2013; Ursem et al, 2002; Guo and Xue, 2011).

Our Jordan net is constructed by modifying the multilayered feedforward with addition a ***context*** layer to help retain data between observations. At each move, new inputs are fed to the net. Previous contents of hidden layer is passed into context layer and later fed back into the hidden layer in the next time step. The context layer contains nothing initially. Output from the hidden layer after the first input will be same as if there is no context layer (Ojugo et al, 2013). Weights are computed same way for new connections to/fro the context layer from its hidden layer. Training aims at best fit data weights computed via Tansig function that assumes approximation influence of data points at its center so that the function decreases from its center (Perez and Marwala, 2011) with an Euclidean length (r_j) which yields distance between $y = (y_1,...,y_m)$ vector and its center ($w_{1j},...,w_{mj}$) given by Eq. 4 to Eq. 6 respectively as:

$$r_j = \left|\left| y - Y^j \right|\right| = \left\{ \sum_{i=1}^{m} (y_i - w_i j)^2 \right\} ½ \qquad (4)$$

The suitable transfer function is applied to r_j:

$$\emptyset(r_j) = \emptyset \left|\left| y - Y^j \right|\right| \qquad (5)$$

Finally, output k receives weighted combination as:

$$y^k = w_o + \sum_{j=1}^{n} \left(c_j^k * \emptyset(r_j) \right) = w_o + \sum_{j=1}^{n} \left(c_j^k * \emptyset ||y - Y^j|| \right) \qquad (6)$$

IV. Experimental Framework

The proposed system will resolve existing problems via:

a. Perform repetitive tasks without emotional defects
b. Embody the knowledge of human experts with the help of special software tools, manipulate data to solve problems and make decisions in that domain.
c. Processes are better formalized and defined on machines.
d. Automatic updating of the knowledgebase.
e. Processes are better formalized and defined on machines.

A. Material and Methods

Dataset contains 33,000 records of credit card transactions. Each record has 23-fields and our nondisclosure agreement prohibits us from revealing the details of the database schema as well as the contents of the data. But, it suffices to know that it is a common schema used by banks in Africa and Nigeria as part of the harmonization scheme. It contains information that

banks deem important for identifying fraudulent transactions. The dataset was already classified into fraudulent or non-fraudulent classes. From records, 38.2% are fraud transactions (emanating from product transaction, asset misappropriation, corruption and financial statement fraud). The sampled data is for a 24-month period. Note that the number of fraud records for each month varies, and the fraud percentages for each month are different from the actual real-world distribution.

B. Data Preprocessing

From original dataset, we prepared the data as suitable for use by the model by removing redundant fields. This helps to reduce the data size as well as speed up the learning heuristics, simplified the learning patterns and made the learned patterns more concise (as adapted from Stolfo et al, 2015). We also compared results of learning between our suitable data versus the original data, and saw no loss in accuracy.

Also, observed data had a skewed distribution of 34% fraud and 66% non-fraud). We adopt 34% fraud class distribution as complete dataset (training data for fraud is always insufficient and we are not expecting an artificially, higher fraud rate to accurately compute suspicion score for fraud patterns). We also must determine ***suspicion*** score for each rule generated by the rule-based model in conjunction with the GA operators to help optimize functions for our training data. And though there are no rules for splitting data, we split it as 50% training, 25% cross-validation and 25% testing for fraud distribution, which also yielded the best classifier for the model. Thus, we demonstrate that even with outliers and noise in dataset and with imprecision and ambiguities applied at its input, model effectively classifies transactions into its proper classes. Thus, GANN effectively scales up learning algorithms void of over-parameterization, over-training and over-fitting of data feats; while maintaining overall performance accuracy.

C. Proposed Model Framework

From Fig 1, the proposed model design employs these:

a. The rule-based system consists of *classifier* to propagate the IF-THEN rule values of selected data, enhanced them as predefined variables classification into intrusion types for fraud detection. The rule-based model is a production system with four (4) components: (i) ***rule set*** containing in each rule a pattern that determines applicability of the rule and corresponding operation to be performed if rule is applied, (b) ***knowledgebase*** (previous transaction set, classified into fraudulent and non-fraudulent using if-then rules as selected data feats), (c) ***control*** strategy specifies the order in which the rules are compared to those in the knowledgebase to find a match and it seeks also a way to resolve conflicts that arise when several rules are matched at the same time, and (d) a rule ***applier*** (Rich, Knight and Nair, 2009).

b. Jordan network provides a self-learning ability, optimized by the CGA's recombines and mutation of the rule-based dataset to train/test the system so that it autonomously classifies transaction into its class types.

c. Decision support – consists of predicted value output with automatic update of the knowledgebase, as transactions are diagnoses on its encounters of new data as in fig 1.

D. Genetic Algorithm Trained Neural Network (GANN)

GANN is initialized with the if-then rules as individuals, whose fitness is computed. 30-individuals are then selected via tournament method as new pool. It then determines mating individuals to yield solutions. We use a multi-point crossover and mutation to help the network to learn all the dynamic and non-linear feats in the dataset (as feats of interest). With mutation, suspicion score for each rule between 1-to-30 is then randomly generated using Gaussian distribution corresponding to crossover points (all genes are from single parent). As new parents contribute the rest to yield new individuals whose genetic makeup is a combination of both parents, mutation is also applied to yield 3-random genes. These further undergo mutation and are then allocated new random values that still conform to the belief space. These random values will range between 0 and 1, which yields the suspicion score for each transaction as generated for each account holder.

The number of mutation applied depends on how far CGA is progressed on the network (how fit is the fittest individual in the pool), which equals fitness of the fittest individual divided by 2. New individuals replace old with low fitness so as to create a new pool. Process continues until individual with a fitness value of 0.8 is found – indicating that the solution has been reached (Ojugo et al, 2013).

Initialization/selection via ANN ensures that first 3-beliefs are met; mutation ensures fourth belief is met. Its influence function influences how many mutations take place, and the knowledge of solution (how close its solution is) has direct impact on how algorithm is processed. Algorithm stops when best individual has fitness of 0 (Dawson and Wilby 2001).

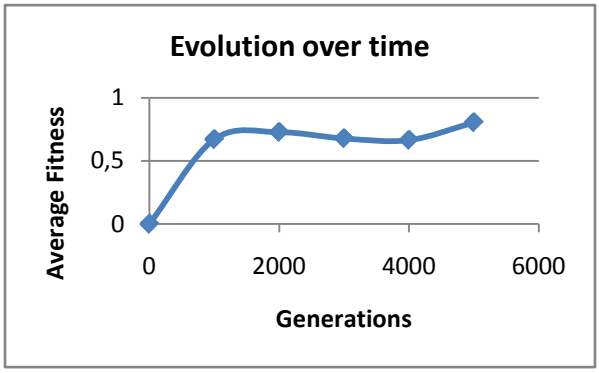


Fig. 1: Fitness Function Computation by Model

V. Result Findings And Discussion

For fraud detection, the performance rating of any detection mechanism is in its fraud catching rate and its false alarm rate. These are critical metrics such that a low fraud catching rate implies that a large number of fraudulent transactions will go through the system; Thus, costing the banks a lot of money (and the cost will eventually be passed to the consumers. Also, a high false alarm rate implies that large number of legitimate transactions will be blocked by the detection system. Thus, supervised intervention, monitoring and management will then be required to authorize transactions. This will frustrate many customers, while also adding operational costs (Stolfo et al, 2015). Note that the fraud catching rate is more important and critical than the false-alarm rate (true-negatives and false-

positives). Ideally, a cost function that takes into account true-negatives and false-positive rates, should be used to compare the classifiers. For lack of cost information from the banks, we rank our classifiers using first the fraud catching rate

Ojugo et al (2013) Performance is evaluated via computed values: mean square error and coefficient efficiency as thus:

Table 1. Model Convergence Performance Evaluation

Model	*MSE*	*COE*	*Fraud Catching Rate*	*False Alarm*
ANN	0.230	0.31	76%	24%
CGA	0.134	0.28	86%	14%
Rule-System			73%	27%
CGANN + Rule-based	0.313	0.219	91%	9%

After training and testing, compared to the models ANN, CGA and rule based system, the results are as follows: (a) ANN was run 24times and it took 223seconds to find solution after 98-iterations (best) and its fraud catching rate ranks at 76%. But, its demerit is that its solution is often trapped at local maxima, (b) GA was run 15-times to eradicate biasness and solution was found each time. It took 98seconds to find the solution after 123-iterations (best) and its fraud catching rate is 76%. Its convergence time depends on how close the initial population is to the solution as well as on the mutation applied to the individuals in the pool. Its demerit is that it seeks a global optima (in this case, a single rule that can be applied to all transactions. This would be appropriate if the transaction platforms are not considered as user are allowed to make transactions from various places – using varying devices that grants them access to their account at any point in time, and (c) CGANN with rule-based preprocessor hybrid was run 152times and its time varied between 29- and 245seconds to find solution after 102-iterations (best) and its fraud catching rate ranks at 91%. Its solution was made even closer using the fuzzy variable dataset (as a preprocessor).

Hybrids have proven to be intelligent modules to transform transaction with adaptive results that provides potential model for fraud detection. Its generated rule set has an accuracy of 92%, sensitivity of 91%, and failure analysis (true-negative and false-positive rate) of 14% respectively. However, the extracted rules are sound and agree with outcome of relevant fraud detection norms and studies.

A. Related Studies

Khasei et al (2012) adopted a feed-forward multi-perceptron network in their study. Its network was extended to represent complex dynamic patterns and cases to treats all data as new – so that previous data do not help to identify data feats chosen, even if such observed datasets exhibits temporal dependence. Consequently, this has practical implementation difficulty as large networks are not easily implemented. However, ***Jordan*** network overcomes such difficulty via its use of its internal feedbacks, which also makes it appropriately suitable for such dynamic, non-linear and complex tasks as its output unit is fed-back as input into its hidden unit with a time delay, so that its outputs at time $t-1$, is also input at time t.

Dheepa and Dhanapal (2009) also reviewed three methods to detect fraud by exploring the different views of same task to see what can be learned. They include: (a) clustering via data clustering of regions of parameter value, (b) Gaussian mix by a measure of the probability density of credit card user's past behavior, to compute probability of current behavior so as to detect anomalies from past behavior, and (c) Bayesian net is used to describe statistics of a particular user and the statistics of different fraud scenarios. He also suggested a combination of all three classifiers.

Stolfo et al (2015) employed meta-learning heuristics (ID3, Bayes, CART and RIPPER) in their study to properly classify transactions. In comparison, the Bayes, RIPPER and CART as base classifiers performed well in its classification of 80% fraud-catching and 19% false alarm result rates respectively.

B. Rationale for Statistical Technique

A major reason for using statistical-data analytics to tackle fraud owes from the fact that a lot of internal control systems have serious weaknesses such that for effective management, institutions must monitor and investigate every transaction that takes place and test it against established parameters, across apps and across systems, from dissimilar applications and data sources. Also, in implementing internal system, some controls are never even turned on. Thus, our rationale for adopting stochastic technique is based on Peter (2014) in which he compared convergence behavior against other machine learning techniques for task classification. In his comparison, decision trees takes 90iterations to converge; neural network approach takes 70iterations, clustering takes 40iterations; while, the hybrid model takes 30 iterations to converge. We also note that while trying to balance model's speed and greater accuracy of classification, more number of rule-set are generated and the knowledgebase consequently updated for optimality and greater functionality. Model trades off speed and accuracy for memory ***resource*** management.

VI. Conclusion and Recommendations

Hybrids are tedious and difficult to implement. Also, dataset accompanying the model must be appropriately encoded so that model can effectively exploit numeric data, which will in turn help the model efficiently, explore the problem space and yield optimal solution. Modelers must seek appropriate data feats and parameter selection alongside proper adjustment of weights/biases so as to avoid ***over-fitting***, ***over-training*** and ***over-parameterization*** of model. If data is properly encoded in a model's structured learning, it will help resolve the conflict imposed on its by the dataset and statistical dependencies of the varying heuristics used. Hybridization using CGA curbs rampant deviates along its output, imposed on model as agents in the space create and enforce their own behavioral rules on the dataset. This is achieved using the CGA belief space.

Models serve as educational tools to compile knowledge about a task, as new language to convey ideas as experts gain better insight to investigate input parameter(s) crucial to a task (Perez and Marwala, 2011), and its sensitivity analysis helps to reflect on theories of systems functioning. Simple model may not yield enough data; while complex model may not be fully understood. Detailed model helps us develop reasonably-applicable models even when not operationally applicable in a larger scale Their implementation should seek its feedback as more critical rather than seeking an accurate agreement with historic data. Since, a balance in the model's complexity will

help its being understood and its manageability, so that the model can be fully explored as seen here (Ojugo et al, 2012).

REFERENCES

[1] Bolton, R.J and Hand, D.J., (2002). *Statistical fraud detection: a review*, Statistical Science, 17(3), pp235-255.

[2] Chakraborty, R., (2010). *Soft computing and fuzzy logic*, Lecture notes, retrieved from http://www.myreaders.info/07_fuzzy_systems.pdf

[3] Coello, C., Pulido, G and Lechuga, M., (2004). *Handling multiple objectives with particle swarm optimization*, Proc. of Evolutionary Computing, 8, pp 256–279.

[4] Dawson, C and Wilby, R., (2001). *Comparison of neural networks in river flow forecasting*, J. of Hydrology and Earth Science, SRef-ID: 1607-7938/hess/2001-3-529.

[5] Delamaire, L and Abdou, H., (2009). *Credit card fraud and detection techniques: a review*, Banks and Bank Systems, 4(2), pp57

[6] Dheepa, V and Dhanapal, R., (2009). *Analysis of Credit Card Fraud Detection Methods*, Int. Journal of Recent Trends in Engineering, 2(3), pp126

[7] Guo, W.W and Xue, H., (2011) *An incorporative statistic and neural approach for crop yield modelling and forecasting*, Neural Computing and Applications, 21(1), pp109-117

[8] Heppner, H and Grenander, U., (1990). S*tochastic non-linear model for coordinated bird flocks*, In Krasner, S (Ed.), *The ubiquity of chaos* pp.233–238. Washington: AAAS.

[9] Kennedy, C and Porter, A., (2013). *Fraud detection and prevention*, [online]: retrieved July 2015 from www.moss-adam.com/fraud_reviews

[10] Khashei, M., Eftekhari, S and Parvizian, J (2012). *Diagnosing diabetes type-II using a soft intelligent binary classifier model*, Review of Bioinformatics and Biometrics, 1(1), pp9-23.

[11] Kuan, C and White, H., (1994). *Artificial neural network: econometric perspective*, Econometric Reviews, Vol.13, Pp.1-91 and Pp.139-143.

[12] Mandic, D and Chambers, J., (2001). *Recurrent Neural Networks for Prediction: Learning Algorithms, Architectures and Stability*, Wiley & Sons: New York, pp56-90.

[13] Michalewicz, Z., (1998). *A survey of constraint handling techniques in Evolutionary computation methods*, www.dhpc.adelaide.edu.au

[14] Minns, A., (1998).*Artificial neural networks as sub-symbolic process descriptors*, published PhD Thesis, Balkema, Rotterdam, Netherlands

[15] Ojugo, A., Eboka, A., Okonta, E., Yoro, R and Aghware, F., (2012). *GA rule-based intrusion detection system*, Journal of Computing and Information Systems, 3(8), pp 1182 - 1194.

[16] Ojugo, A.A., Emudianughe, J., Yoro, R.E., Okonta, E.O and Eboka, A., (2013). *Hybrid neural network gravitational search algorithm for rainfall runoff modeling*, Progress in Intelligence Computing and Application, 2(1), doi: 10.4156/pica.vol2.issue1.2, pp22–33.

[17] Ojugo, A.A., Ben-Iwhiwhu, E., Kekeje, D.O., Yerokun, M.O and Iyawa, I.J.B., (2014). *Malware propagation on time varying network*,Int. J. Modern Edu. Comp. Sci., 8, pp25 – 33.

[18] Perez, M and Marwala, T., (2011). *Stochastic optimization approaches for solving Sudoku*, IEEE Transaction on Evol. Comp., pp.256–279.

[19] Peter, S., (2014). *An Analytical Study on Early Diagnosis and Classification of Diabetes Mellitus*, Bonfring International Journal of Data Mining, 4(2), pp7-13.

[20] Reynolds, R., (1994). *Introduction to cultural algorithms*, Transaction on Evolutionary Programming (IEEE), pp.131-139.

[21] Saleh Elmohamed, M.A, Fox. G and Coddington, P., (1998). *A comparison of annealing techniques for academic course scheduling*, Notes on Intelligence Computing, DHCP-045, pp 1-20. www.dhpc.adelaide.edu.au.

[22] Stolfo, S.J., Fan, D.W., Lee, W and Prodromidis, A.L., (2015). *Credit card fraud detection using meta learning: issues and initial results*, [online]: http://www.researchgate.net/publication/2282588

[23] Ursem, R., Krink, T., Jensen, M.and Michalewicz, Z., (2002). *Analysis and modeling of controls in dynamic systems*. IEEE Transaction on Memetic Systems and Evolutionary Computing, 6(4), pp.378-389

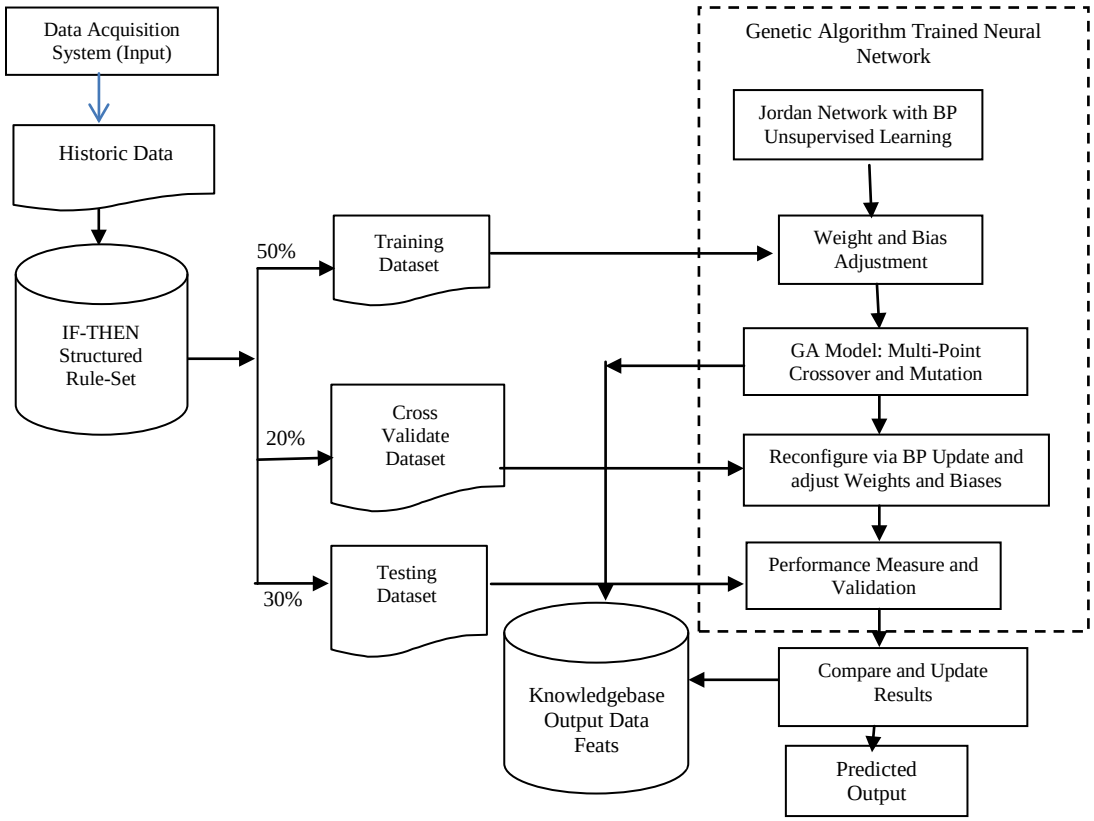


Fig. 2: Dataflow Diagram of the Hybrid Model Fuzzy Genetic Algorithm Trained Neural Network

The diversity of management theories for SME's development

Deniss Sceulovs
Riga Technical University
Faculty of Engineering Economics and Management
Riga, Latvia
deniss.sceulovs@rtu.lv

Elina Gaile-Sarkane
Riga Technical University
Faculty of Engineering Economics and Management
Riga, Latvia
elina.gaile-sarkane@rtu.lv

Abstract— the aim of the proposed scientific research is to evaluate SME managers' readiness to practical application of different management theories to SME (small and micro enterprises) in the Latvia. Authors analyzes and show importance of the management functions in business and on the basis of different field researches and their results evaluates possibility to apply of management functions in small companies. The authors conclude that SME managers are looking for ready-made tools in business development and not ready to demand special business management approaches, which are tailor-made for specific/individual business needs and a company development. The management of SME faces many factors and issues what affects success of a small business. Among them are success factors, modern business models, e-environment, managerial skills, knowledge, individual competences, entrepreneurial mindset etc. For successful management of a small company manager (owner) should choose his own approach in diversify of many modern business theories and practices.

***Keywords*—**SMEs, management functions, management thories and functions, success factors, business models.

I. INTRODUCTION

During last decades balanced development was the aim of any company, ensuring the company's existence in a long-term, for promoting the overall growth of the national and global economy. In Latvia and in other European Union countries, numeric majority of companies are micro, small, and medium enterprises (SME), furthermore, due to the lack of resources, small companies are successful in local market and find it difficult to develop on the globally.

Small and medium enterprises have to choose strategies (tools, methods, approaches etc.) of product development and market promotion that are different from big companies' (corporations), applying only those tools from the available ones what can increase their competitiveness and are most suitable for the particular company's needs and specific.

SMEs are the engine of the European economy. Analysis of the statistical information (see TABLE 1) confirms dominance of small and even micro size of companies in Latvia.

SME are an essential source of jobs, create entrepreneurial spirit and innovation in the EU and are thus crucial for fostering competitiveness and employment [1].

TABLE I. ECONOMICALLY ACTIVE STATISTICAL UNITS IN LATVIA BY SIZE [2]

Unit by size	Year			
	2009	**2010**	**2011**	**2012**
Micro	115939	123924	129394	138628
Small	10254	9732	9970	10436
Medium	2065	1956	2033	2079
Large	351	355	381	400
Total	128609	135967	141778	151543

Latvian business development similarly like in other countries is affected by several factors – starting with the economic, socio-political situation of the country, the global market conditions, and ending with the individual knowledge, skills and competencies of a company's management and staff, their motivation to start and continue a business, business aptitude, business model etc.

Reviewing the managerial functions described in the theory, the authors conclude that they are generally applicable to middle or large companies and corporations which mainly operates in large markets. The authors decided to pursue a more pragmatic approach by exploring the actual functions of micro and small companies in order to build a list of managerial functions needed in the business environment of Latvia. The authors proposed following research question: what SME managers consider as the best drivers for a company development with and aim to shift from small to large enterprise? For answering on research question following research design had been chosen: quick review of management theory and analysis of the results is done. Field research was performed and according to the results main factors influencing SME development were named. Research results are validated in focus group discussion.

ISBN: 978-1-61804-327-6

II. MANAGEMENT THEORIES AND MANAGEMENT FUNCTIONS

There are many possibilities to describe development of management theories. According to one of such approaches historical theories of management can be grouped following [5]:

- Scientific Management Theory (1890-1940) for large industries, describing ongoing tasks, mechanic and routinized activities.
- Bureaucratic Management Theory (1930-1950) focused on dividing organizations into hierarchies, establishing strong lines of authority and control.
- Human Relations Movement (1930-today). More attention was given to individuals and their unique capabilities in the organization.

Also there was mentioned Contemporary Theories of Management [5]:

- Contingency Theory. Contingency theory asserts that when managers make a decision, they must take into account all aspects of the current situation and act on those aspects that are a key to the situation at hand.
- Systems Theory recognizes the various parts of the organization, and, in particular, the interrelations of the parts, e.g., the coordination of central administration with its programs, engineering with manufacturing, supervisors with workers, etc.
- Chaos Theory This trend continues until the system splits, combines with another complex system or falls apart entirely.

Yasin Olum in his article identifies the Scientific Management School comprising the works of Frederick W. Taylor and Lillian Gilbreth's motion study, among others; the Classical Organizational Theory School comprising the works of Henri Fayol's views on administration, and Max Weber's idealized bureaucracy, among others; Behavioral School comprising the work of Elton Mayo and his associates; the Management Science School which I discuss at the end of this section; and Recent Developments in Management Theory comprising works such as Systems Approach, Situational or Contingency theory, Chaos theory, and Team Building approach. For lack of time and space, this discussion will provide a general description of some of the scholars in each of these management theories and the successes that they achieved [6].

Many authors stress that management is science *on the edge* among philosophy, economics, history, etc. Olum in his article writes: "Management is the art, or science, of achieving goals through people. Since managers also supervise, management can be interpreted to mean literally "looking over" – i.e., making sure people do what they are supposed to do. Managers are, therefore, expected to ensure greater productivity or, using the current jargon, 'continuous improvement'" [6]. And logically there is a place for discussing functionality of management or management functions.

At first, one views management functionally, such as measuring quantity, adjusting plans, meeting goals, and foresighting/forecasting. This applies even in situations when planning does not take place. From this perspective, Henri Fayol (1841–1925) considers management to consist of six functions: forecasting, planning, organizing, commanding, coordinating and controlling. He was one of the most influential contributors to modern concepts of management [7].

Mišetić in his research has written: "During the development of management theory, the five management functions have changed over time. In order to show a comparison in the way management theory has developed, Buble compared H. Fayol's (planning, organizing, commanding, coordinating, and controlling) and P. Drucker's (setting goals, organizing, motivating and communicating, measuring and evaluating results, and human resources development) theories of management functions." [8]

By analysis of different management functions described in the scientific articles, text books and internet resources by management professionals (Drucker, Cole, Hill, Praude, Abizāre, Forands etc.) the authors concluded that there are main five management functions: Planning, Organizing, Staffing, Leading, Controlling [9].

According to the Strategic management theory proposed by Charles W.L. Hill and Gareth R. Jones [10] there are three main levels of strategic planning: corporate level, business level and functional level. Mentioned management functions can be applied mainly to corporate level, partly to business level, and minimally to functional level (see Fig. 1.).

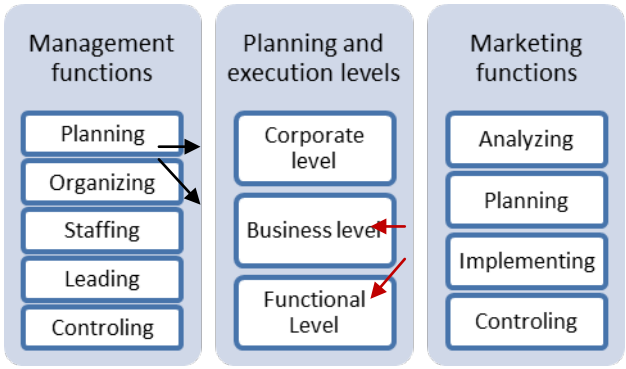


Fig.1. Application of functions in different planning and execution levels in a company [10]

Modern business sustainable development also depends on how effective business is using information, therefore we believe that the information could be taken as a source and mean of the modern business sustainable development.

It is well known that marketing is one of core activities for development and delivering the product or service to the customers. According to Philip Kotler [11] there are four main functions in the marketing: analyzing, planning, implementing and controlling.

From the quick theoretical overview one can conclude that there is different management theories that can be applied to management of SME, at the same time not all management

ISBN: 978-1-61804-327-6

functions can be applied to every small, medium of micro size company. Functions, methods, management style etc. varies from size of the company, field of business, aims of company and other important factors. Therefore it is important to identify is there any other factors influencing management of SME.

III. Small business managerial issues and affecting factors

There are two fundamental characteristics of small businesses which separate them from large companies. One is their smallness, and the other is their rate of turnover and failure rate [12]. In his research Peackok has identified following problems of small business failure [12]:

- The owner is the business,
- Isolation,
- Lack of information,
- Lack of time,
- Pressure,
- Product dedication,
- External changes are critical,
- Small businesses are seldom in equilibrium or even near it,
- High business risk.

In the same research Peackok have mentioned Dunand Bradstreet studies which have consistently found that causes due to poor management predominate in failures [12]:

- US business failures, 92% due to management,
- US 17,000 business failures, 94% due to management, and Canada 2,598 business failures, 96% due to management.

According to this research is obvious that there are some misapply and misinterpretation of main management functions in small companies. And there is a question – should these companies apply the same functions to the business development? Do the managers in small companies need the same competences and capabilities?

Studies conducted in Latvia produce a strong link between a company's success and the entrepreneur him/herself. Thus, in a Swedbank entrepreneurship competence study, the main traits were determined, which, according to the opinion by Latvian inhabitants, must describe a person, who wishes to begin a business venture. The courage to take risks is the most important quality for starting a business (48 % respondents), confidence (45 %), and intuition (42 %) were the most frequently mentioned traits. However, the willingness to make profit ranked second to last on the list of desirable qualities in an entrepreneur important for starting a new business.

Another survey [13] carried out in December 2011 points to the following factors, which could affect business development: an open market (not oversaturated) with relatively low barriers and niche market with development potential; availability of resources that have become scarce in other markets, such as land, timber and low-cost labor; the business culture and mentality congenial to that of the European countries; a multitude of interesting business projects, enterprises that have performed well before investment, well-established production traditions; interesting developments, enterprises with great potential.

For analysis of factors, influencing Latvian SME's the authors has conducted focus group. The focus group method was applied to this research since "focus groups concentrate clearly on a specific topic and involve interactive discussion among its participants" [3]. The focus group was held February, 2012. The total number of respondents was 14, aged 22 to 58, both sexes the size of focus group was determined according to the methodology described by Griffin and Hauser [4]. Within focus group discussion success and failure factors of SME's were discussed. Among focus group discussions results were pointed out that SME managers as best solution for development of company (for turning company from small to large) see employment of e-environment, application of existing e-tools and changing business model towards e-business. Also the focus group members confirmed that main failure of small company is lack of skills, competencies and entrepreneurial mindset of the owner of a business (company).

The authors of the article would like to draw attention of Latvian entrepreneurs to the fact that they must start to think globally, they must be aware of the fact that the consumer purchasing ability is not restricted to the Latvian market. The companies, which will take the step to expand their geographic business boundaries, will definitely gain success, because entrepreneurs have all the competences for further development and expansion of their business. Moreover, company founders and owners must take into account that, as a company develops, its team also must develop, developing and improving their business etc. knowledge.

IV. E-environment as a tool for small business development

Over a period of eighteen years, the number of internet users in the world has increased from 16 mill. people in 1995 up to 2749 mill. people in 2013 (data as of 03.2013) [14].

In Latvia, internet started taking shape around 1992. In 2000, there were around 150 000 people who used internet in Latvia in 2000 [15], whereas in early 2013, the number of internet users of Latvia has grown to 1.277 mill. people [16].

Development of internet and information technologies directly affected entrepreneurship, as a new type of entrepreneurship emerged – electronic entrepreneurship, including e-commerce.

Consumers have an important role in modern-day business models. Firstly, owing to market development and general progress, consumer values have changed (desire to receive individual attitude; the types and rate of buying and payment processes; beliefs, opinions, and expectations as

regards product prices, and other; individual shopping experience and habits, etc.). Secondly, owing thanks to information resources, mainly the internet, and the consumer can quickly obtain a big information amount about the product of interest. Thirdly, in the new e-business models, the consumer and the seller can quickly generate feedback about a product.

Regardless of extensive availability of the e-environment and its elements, there are entrepreneurship sectors (spheres) in Latvia, which actively use the e-environment, as well as those, in which the use of e-environment is not particularly widespread.

The use of e-environment tools in Latvian enterprises differed also depending on the company's size. Thus, in big companies of Latvia, the parameters of using such e-environment tools as a computer, internet, and a website were better than in small or medium enterprises. Companies with 10 to 49 employees use the e-environment tools least.

Among Latvian companies, there are few of those that use internet for selling goods. According to the data of the Central Statistical Bureau of the Republic of Latvia for the year 2012, only 9.1 % of Latvian companies performed e-sales and only 23.6 % of companies have performed e-purchases. Automated data exchange is ensured only in 53 % of Latvian companies.

Regardless of rapid development of the e-environment on a global scale and extensive use thereof in entrepreneurship, the Latvian companies are not employing the e-environment and its tools for entrepreneurship purposes to a sufficient extent.

The authors assume that one of the important reasons that affects and is closely related to the use of the e-environment is the aspects of technology adaptation. The e-environment consists of and interacts with various technology elements (internet, devices, software, etc.), which, for their part, are related to various continuously developing technologies, including information, production etc. technologies.

A. Adaptation of technologies in companies

Time is required for introducing and adapting the new technologies. This time is necessary to prepare a company for introducing new technologies – to streamline and improve the company's processes affected by the new technologies, to train employees, as well as carry out testing and other activities related to new technologies in the company.

It is exactly the differing experience of adaptation of technologies in companies that, according to the author, could be the explanation to why some companies are actively employing the e-environment for entrepreneurial needs and some do not or use it to a lesser extent.

B. Business models in the e-environment

The e-business model is based on mutual integration of key flows and values and implementation thereof between e-market participants, through the use of the e-environment. Three main e-business model elements and integrity can be distinguished: *flows*, *participants*, *value*. The term e-business model describes a broad spectrum of informal and formal models, which may be used in companies to depict various business aspects, such as operational processes, organisational structures, and financial forecasts.

C. Values – a constituent element in modern business models

There are several well-known and popular value theories, such as, the five forces model [17], shareholder value model [18] as well as the "value map" theory, intended for analysing the economic gain for consumers [19] etc.

Various theories were developed many years ago, when the electronic market was not yet developed, and hence are suitable for the conventional market. Due to this reason, the authors of the paper suggests that companies use the Alexander Osterwalder's value proposition concept or the approach that is a constituent element of the author's developed business model canvas) [20].

The Osterwalder's business model was formed based on Freeman's stakeholder theory [21]. The model is adapted to today's market needs and conditions, and the importance of the electronic environment, i.e. of the electronic market, in entrepreneurship is taken into account. The authors of the paper wish to draw attention to Osterwalder's "value life cycle" consisting of five stages: value creation, appropriation, consumption, renewal, and transfer. All life cycle stages are linked to value consumption, using the electronic environment: value creation (based on information and communication technologies (ICT) – adaptation of various products for the needs of an individual consumer, e.g., personal computer, footwear, etc. Value appropriation – "a single click purchase" at an internet shop. Value consumption – listening to music, watching a movie, etc. Value renewal – various software updates, value transfer – disposal of old computers and other machinery, handing over unnecessary books and equipment for further use, etc.

The value approach is broadly used in various business models, including e-business models. The value is at the basis of several modern business models, which are topical also for small business.

D. The use of electronic environment for ensuring development of enterprises in Latvia

The authors of the paper previously have conducted a study on the use of the e-environment by Latvian SMEs in entrepreneurship [22]. The aim of the study is to establish which e-environment tools are used by companies, what is the intensity of their use, and what are the companies' skills in using them.

Time of the survey: May 2012 – October. Sample base – an internet survey created on an online e-survey site www.visidati.lv, as well as by posting an announcement with

a call to participate in a survey in social networks – www.linkedin.com, www.facebook.lv, www.skype.com, www.draugiem.lv. The database of Lursoft Ltd. was also used. Overall, the survey was sent to 2100 Latvian companies, 1600 of which were SMEs. The call to fill in the questionnaire was sent to Latvian company managers and leading employees. The questionnaires were sent out twice, and a part of companies were called by phone and asked to fill in the questionnaires. Filled-in questionnaires with answers to all questions were received from 526 companies. The authors has processed and analyzed the obtained data in SPSS 20 environment. Results were processed with SPSS. 20 and factor analysis and interpretation, ten factors affecting the use of e-environment in the respondent companies were elucidated.

Taking into account the obtained variables and correlation quotients, the author has performed their interpretation and has created factor classification:

1. company managers' and management's understanding of the types of use of e-tools;
2. understanding of communications tools and their use in entrepreneurship;
3. understanding of e-business models (the author's remark: taking into account the variables in this factor, as well as correlation quotients, it can be concluded that companies have no clear understanding of the meaning of e-business models. This factor explains answers given to question 10 of the survey);
4. company's communication with the stakeholders in online mode;
5. company's website, which is a tool of sales and marketing (the author's remark: companies have these tools, but have no understanding of the aims that the tools are intended for and how to use them);
6. the use of e-environment tools in market and consumer studies (the author's remark: companies are informed of the existence of such tools, but are not aware of how they could be used);
7. understanding of state-offered services (the author's remark: the factor is affected not only by weak use of the business tools, but deficiencies of several e-government tools, which encumber their use);
8. acceptance of e-environment tools among employees;
9. versatility of e-environment tools in a company;
10. use of state e-services for business and private needs.

Based on the study results, the authors of the article conclude that overall companies are using e-environment tools. There is an apparent lack of information and poor knowledge of entrepreneurs regarding the use of e-environment tools in entrepreneurship. A positive conclusion – the SMEs understand that with e-environment tools, insufficient or missing resources can be replaced / supplemented. Companies also are willing to gain knowledge and learn how to use e-environment tools for entrepreneurship needs. Several answers point to managers' and employees' partial acceptance of e-environment tools, which is a hindrance for comprehensive use thereof. Factor No. 8 is very important, indicating to acceptance of information technologies, referred to by the authors in Section B.

CONCLUSION

Summarizing results of the study authors came to a conclusion that small business manager (owner) should use various managerial approaches, but there are not ready to apply existing management theories to development but looking for ready-made solutions. Mainly such solutions they find in e-business tools and models.

Research results clearly show that information technology mainly affects company's success and information flow within company. Considering that information resource has today become one of the prevail resources for sustainable development, it must be outlined that this resource economic evaluation becomes of the high importance.

Small and micro companies' success depends on personal characteristics of a company owner (company manager).

The use of e-environment in entrepreneurship proved that a range of e-tools are available simultaneously, which can ensure more effective use of company's resources and development in a long-term, however entrepreneurs do not know all of them or do not know how to or do not wish to use them.

By informing the small and medium company managers about the available e-tools and their use, considerable increase in competitiveness of companies can be achieved, which concurrently ensures creation of additional values for the company's clients and stakeholders, as well as promotes company's development.

Taking into account that SME managers are looking for ready-made tools in business development and not ready to demand special business management approaches, which are tailor-made for specific/individual business needs and a company development the authors suggest to continue research with and aim to develop semi management theory what would cover needs of small and micro companies.

REFERENCES

[1] G. Verheugen, The new SME definition User guide and model declaration. 2011.

[2] Economically active statistical units in Latvia by size group. Central Statistical Bureau of Latvia. [Electronic resource], viewed on January 13, 2014, http://data.csb.gov.lv/Selection.aspx?px_path.

[3] D. Carson, A. Glimore, C. Perry, K. Gronhauigh, Qualitative Marketing Research. Sage, London, 2001.

[4] A. Griffin, J. R. Hauser, The Voice of the Customer, Marketing Science. 12th Edition, Volume No, 1993.

[5] Free Management Library, Authencticity Consulting, LLC, [Electronic resource], viewed on May 1, 2012 http://managementhelp.org/management/theories.htm

[6] Y. Olum, MODERN MANAGEMENT THEORIES AND PRACTICES. Being a paper presented at the 15th East African Central Banking Course, held on 12th July 2004, at Kenya School of Monetary Studies. [Electronic resource], viewed on April 20, 2012 http://unpan1.un.org/intradoc/groups/public/documents/aapam/unpan025765.pdf.

[7] R Barrett."Vocational Business: Training, Developing and Motivating People", Business & Economics - 2003. - Page 51.

[8] D. Mišetić, Management Functions Realization with Small Enterprises in Transition . Enterprise in Transition: International Conference.

[9] E. Gaile-Sarkane, "Semi management theory and business sophistication", WMSCI 2012 - The 16th World Multi-Conference on Systemics, Cybernetics and Informatics, Proceedings, Volume 1, 2012, Pages 25-30.

[10] C.W.L Hill, G. R. Jones, Strategic Management Theory: An Integrated Approach, 9th edition. - South-Western Cengage Learning, 2008 – 454pp.

[11] Ph. Kotler Marketing Management, 11th edition. – Prentice hall, 2002, 815 pp.

[12] R. Peacock, FAILURE AND ASSISTANCE OF SMALL FIRMS [Electronic resource], viewed on February 12, 2012, http://www.sbeducation.info/downloads/sbfail.pdf

[13] Survey: Attieksme pret uzņēmējdarbības uzsākšanu. Uzņēmējdarbības kompetences centrs. [Electronic resource], viewed on January 11, 2014], http://www.manambiznesam.lv/lv/2012/03/19/attieksme-pret-uznemejdarbibas-uzsaksanu/.

[14] Internet World Stats: History and Growth of the Internet from 1995 till Today [Electronic source], viewed on March 14, 2015, http://www.internetworldstats.com/emarketing.htm.

[15] Internet World Stats: Latvia Internet Usage Stats and Telecom Report , [Electronic source], viewed on March 14, 2015 http://www.internetworldstats.com/eu/lv.htm.

[16] Latvijas Interneta asociācija: Statistika, [Electronic source], viewed on march 15, 2015, http://lia.lv/statistika.

[17] M. Porter, "The Five Competitive Forces That Shape Strategy", Harvard business Review, January 2008.

[18] W. Fruhan Jr., "The NPV Model of Strategy - The Shareholder Value Model," in Financial Strategy: Studies in the Creation, Transfer, and Destruction of Shareholder Value. – Homewood: Irwin, 1979. – 124 p.

[19] A. Kambil, A. Ginsberg, M. Bloch, "Rethinking Value Propositions. Working Paper, NYU Centre of Research of Information Systems, 1997.

[20] Business Model Foundry: The Business Model Canvas, [Electronic source], viewed on February 17, 2014 http://www.businessmodelgeneration.com/canvas.

[21] R. Freeman R., Edward Strategic Management: A stakeholder approach. - Boston: Pitman, 1984. – 276 p.

[22] D.Ščeulovs, Use of Electronic Environment in Entrepreneurship Development, Doctoral Thesis. – Riga: RTU Publishing House, 2013. – 164 p.

Design Research for Building an Automated Decision Support System for Intensive Care Units

Pedro Gago
School of Technology and Management
Polytechnic Institute of Leiria
Leiria, Portugal
pedro.gago@ipleiria.pt

Álvaro Silva
Intensive Care Unit
Centro Hospitalar do Porto
Instituto de Ciências Biomédicas Abel Salazar
Porto, Portugal
moreirasilva@me.com

Manuel Filipe Santos
Algoritmi Research Centre
University of Minho
Guimarães, Portugal
mfs@dsi.uminho.pt

Salazar, M., Quintas. C
Serviço de Sistemas de Informação
Centro Hospitalar do Porto
{cesar.quintas, msalazar}@chporto.min-saude.pt

***Abstract*—**We are all aware of the enormous quantities of data that are routinely gathered all around us. Huge databases are being continuously fed sensor data resulting in more data being stored than can ever be analyzed by human experts. Health services in general and Intensive Care Units in particular are one of the domains where data collection is well under way. Even though data are possible to gather, manual data analysis is unfeasible due to both its sheer volume and the lack of available medical experts. Nevertheless, using knowledge discovery from databases techniques it has been possible to extract some knowledge from medical data. However, knowledge discovery from databases still requires expert involvement and is potentially very time consuming. In this paper we propose an architecture for an automated decision support system for intensive care units built using a design research approach. We illustrate the development of the architecture by showing the changes to the initial design that were the result of the design research cycles performed.

***Keywords*—**design research; multi agent systems; decision suppor; knowledge discovery from databases; intensive care

I. INTRODUCTION

As with most human activities nowadays, health care relies heavily on Information and Communication Technologies. More and more patient data is stored in digital formats as institutions strive to achieve a paperless operation allowing them to improve the quality of care [1][2]. Intensive Care Units (ICUs) are one of the most data intensive environments in healthcare as they usually are equipped with a wealth of sensors for real time patients monitoring. However, being able to collect information on dozens or even hundreds of variables does not guarantee better patient outcomes as it is known that humans are unable to simultaneously process more than a handful of variables. Nevertheless, data availability makes it possible the development of Clinical Decision Support Systems (CDSS) through the use of Knowledge Discovery in Databases (KDD).

Even if KDD techniques allow for knowledge discovery from the collected data human expert involvement is a requirement. Moreover, as more data is collected, it is likely that the knowledge gets outdated and that the whole process must be repeated. As such there is a repeated need for human experts to be available in order for this process to be completed. We propose an architecture for a CDSS capable of functioning autonomously after the first cycle of the KDD process.

In order to develop the proposed architecture we followed a design research approach allowing us to use new information as it became available in order to improve the initial idea.

In the next section we explain the background, namely the environment for which this architecture was first developed. In section III design science research is presented. Next, the development process is illustrated and finally we present future work and conclusion.

II. BACKGROUND

A. Clinical Decision Support Systems

CDSS can be found for many objectives, including the following:

- alerts and reminders;
- diagnostic assistance;
- therapy critique and planning;
- prescribing decision support;
- information retrieval;
- image recognition and interpretation.

Despite the potential gains, CDSS are not as disseminated in health care as could be expected [4]. Some hurdles that must be overcome relate do difficulty of having access to adequate databases and adequate sets of rules. Another kind of obstacle is related to the CDSS ease of use, especially in terms of data

This work has been supported by FCT - Fundação para a Ciência e Tecnologia within the Project Scope UID/CEC/00319/2013 and the contract PTDC/EEI-SII/1302/2012 (INTCare II).

entry. It has been shown that CDSS that require significant data entry are not likely to be used on a regular basis [5].

Our proposal aims at avoiding these difficulties by both having automatic rule creation using KDD techniques and automated data collection thus guaranteeing the quality of the available knowledge while avoiding the need for manual data entry.

B. Application Setting

While aiming at building the architecture for a CDSS capable of automating the KDD process we are doing so in the context of a real setting. Our system is being tested in the Intensive Care Unit of Centro Hospitalar do Porto (Oporto, Portugal). This gave us the opportunity to deploy a prototype and to collect some valuable insights.

The available data included the case mix - information that remains unchanged during the patient's stay in the ICU (e.g. age or admission origin) and data from four variables gathered from the bed side monitors and from the nursing records. These four variables were the Heart Rate, the Systolic Blood Pressure, the Oxygen Saturation and the Urine Output. A group of clinical specialists determined the intervals considered normal for each one of these parameters [6].

Prediction objectives of interest to the doctors included: patient outcome at time of discharge (dead or alive), SOFA scores for the next day and length of stay.

III. Design Research

The development of the architecture follows the Design Research methodology, a recognized framework for Information Systems Research that "seeks to extend the boundaries of human and organizational capabilities by creating new and innovative artifacts" [7].

The Design Research methodology encompasses three main components: Environment, IS Research and Knowledge Base. The Environment includes all stakeholders both people and organizations with their diverse views and (often conflicting) interests. From the environment Business Needs may emerge that if relevant are fed into the IS Research part of the methodology. Drawing from the Knowledge Base and with the goal of fulfilling the Business Needs, this phase may include the development and assessment of new theories and/or artifacts. This is an iterative process as the developed artifacts should be refined in light of the assessments made (Fig.1).

Effective design research should follow seven guidelines [7]:

- Design as an Artifact (research must produce a viable artifact in the form of a construct, a model, a method or an instantiation);
- Problem relevance (develop solutions to important business problems);
- Design Evaluation (the utility, quality and efficacy of a design artifact must be rigorously demonstrated);
- Research Contributions (must provide clear and verifiable contributions);
- Research Rigor (both for construction and evaluation of the design artifact);
- Design as a search process (requires utilizing available means to reach desired ends while satisfying laws in the problem environment);
- Communication of research (design science research must be presented effectively both to technology-oriented as well as management-oriented audiences);

One must keep in mind that in this context an artifact is not a fully-grown information system. It may be an idea, a design or a prototype.

In a later paper, Hevner further details this methodology by explicitly including three cycles (Fig. 2): a relevance cycle, a design cycle and a rigor cycle [8]. This new formulation emphasizes the iterative nature of design research in all its constituent parts. The process is designed as to be flexible enough to accommodate changes that may be required after each evaluation step. The relevance cycle initiates the design research process providing the requirements for the research along with acceptance criteria to be used in the evaluation of the search results. Moreover, the rigor cycle provides past knowledge to the research project to ensure its innovation. It also includes the addition of discovered knowledge to the knowledge base. Finally, the design cycle is the internal cycle of the process and relates to the construction and evaluation of an artifact repeated as many times as necessary. As is expected, evaluation results may prompt changes in the artifact that in turn must be evaluated and so on.

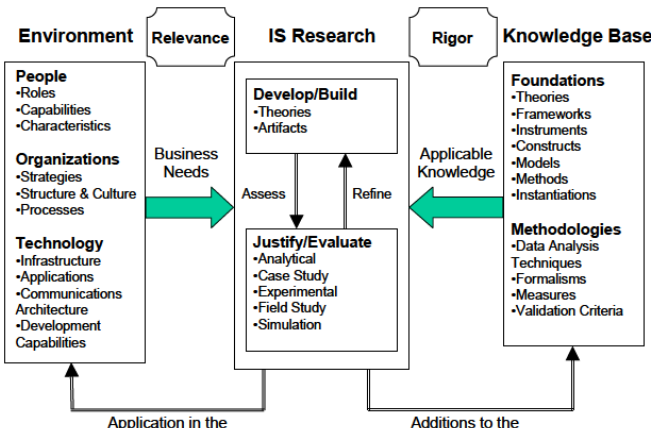


Fig. 1. Information Systems Research Framework [7]

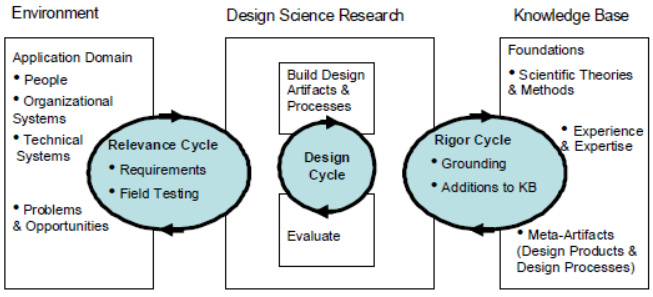


Fig. 2. The Three Cycle view of Design Research [8]

In [9] further considerations regarding Design Research are made and twelve thesis are proposed but overall these authors agree in the cyclical nature of the design research process and that the path to an artifact is not straightforward as evaluation results may prompt new paths to be followed.

IV. Building INTCare

This section describes the ongoing work of developing an autonomous CDSS for use in an ICU by following a design research methodology.

A. Relevance Cycle

In the center of the Design Research methodology, the Design Cycle depends heavily on the work initiated on the Relevance Cycle. This larger cycle includes not only the definition of requirements for the research but also the definition of evaluation criteria for success. While the requirements are fed into the Design Cycle, the criteria for success are stored to be used after all the iterations of the Design Cycle in order to assess the quality of the resulting artifact. Work in the Relevance Cycle involves the understanding of the application domain: people, existing systems and processes. This understanding is what usually allows for the discovery of interesting problems to be solved.

In the context of an ICU, we started by characterizing the decision making process, that can be seen as unfolding in four phases [10] [11] [12] (Fig. 3):

Intelligence - The medical and nursing staff collect physiological data regarding the patient's condition. Then, based on this information, the physicians evaluate the illness state to determine the pertinence of immediate action;

Design - Doctors conceive the possible therapeutic scenarios. Two main categories of scenarios can be considered: immediate action assumption (where there is a pressing need for intervention) and delayed group decision (guidelines for discharge or long term treatments). In the former case, if needed, the physician can consult internal or external colleagues. Delayed group decision choices include: patient's discharge, suspension of life-supporting treatment, writing of do-not resuscitate orders and applying a specific treatment;

Choice – After the identification of the alternatives and entwined with the previous stage, a final decision is taken (either individually or group based); and

Implementation – The corresponding therapeutic procedure is applied.

At that time, doctors relied on their experience and on the available clinical data (e.g. patients chart sheets) to determine if there was a need for action. No use was made of the vast amounts of data, which were automatically collected through bed-side monitors. To the authors' knowledge, although several prognostic scores have been developed (e.g. SAPS [13] or APACHE [14], there were no intelligent DSS systems completely developed and based on the KDD paradigm in use in ICUs for end-of-life decision making. Yet, both patient and physicians could benefit if high-level, reliable and timely information was to be made available. Hence, the INTCare system aimed at filling this gap even with the knowledge of this being a particularly sensitive area with strong physician resistance to be expected [15] .

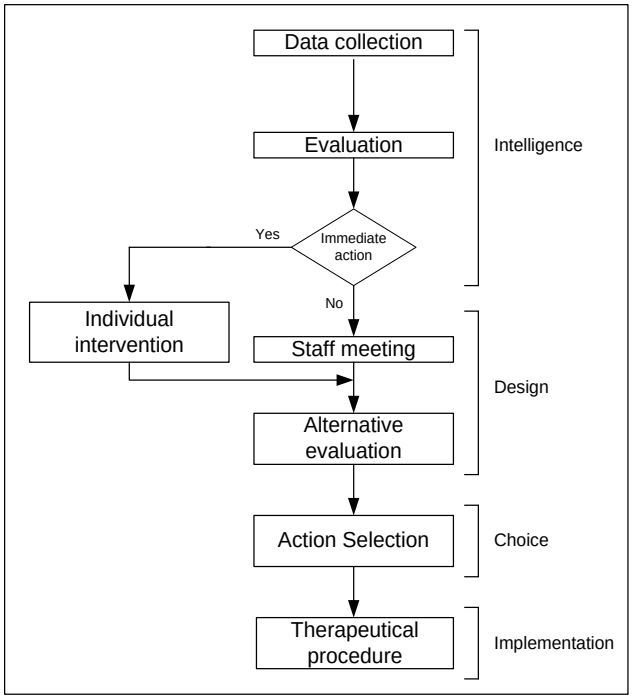


Fig. 3. The CHP ICU decision making process.

Thus, our objective in this study is to investigate "what characteristics should a KDD based CDSS possess in order to be capable of autonomous operation in an ICU setting". Appart from requesting predictions the aim was that human intervention in this CDSS should be limited to an initialization step where a first run of the KDD process is performed so that future runs are automatically executed.

The evaluation criteria chosen is a simple one: to have a prototype of the system in place capable of autonomous operation for at least six months. It is expected that the final evaluation will still take some time as it depends on having enough data and prediction models and there are a number of factors preventing this to be achievable on the short run. With that limitation in place, we opted for using off line data from a previous study to perform sectorial evaluation in order to confirming that the algorithms that allow for the system's adaptation perform as expected. This evaluation was performed in the Design Cycles described in later sections.

As work progressed, the Hospital's Information System grew and new functionalities were made available. Consequently a new Relevance Cycle had to take place where we concluded that given the functionalities made available for data entry by the nursing staff [16] INTCare no longer needed to include a Clinical Data Entry agent and so that agent was removed from consideration. Other functionalities, also described in [16] increased the quality of available data and the possibility of including derived variables in the recorded data. Of particular relevance was the detection of clinical relevant events in real time.

ISBN: 978-1-61804-327-6

B. Rigor Cycle

The interest in Knowledge Discovery from Databases (KDD) and Data Mining (DM) arose due to the rapid emergence of electronic data management methods. In 1997, the Gartner Group suggested that DM is one of the top five key technologies that will have a major impact in the industry within the next years [17]. In effect, these techniques are now widespread and applications can be found in diverse areas, such as: marketing [18], banking [19], manufacturing and production [20], brokerage and securities trading [21] and health care [22]. References for other applications can be found in various publications, including [12] and [23].

Within the Medicine arena, huge databases, with large, complex and multi-source information (e.g. text, images or numerical data), are commonplace. However, human experts are limited and may overlook important details. Furthermore, the classical data analysis (e.g. logistic regression) breaks down when such vast amounts of data are present. Hence, an alternative is to use automated discovery tools to analyze the raw data and extract high level information for the decision-maker [24].

The above goals may involve the application of Machine Learning algorithms such as Decision Trees (DT)[25][26] and Artificial Neural Networks (ANN) [27][28]. Also good results are achieved when using ensembles for supervised learning, where a set of models are combined in some way to produce an answer [29].

C. Design Cycle

Several iterations of the Design Cycle were performed over the years. At first an initial prototype was used to prove to all stakeholders the feasibility of building such a DSS. It was a necessary step in order to establish confidence that the final objective was achievable. After some fine tuning and analysis it was possible to prove that automated data acquisition from the bed side monitors was possible and reliable. This stage uncovered some issues related to user interference with the automatic data acquisition equipment and an initially unplanned meeting with the nursing staff had to be scheduled inform them of the systems objectives and some procedure changes. This evaluation cycle included the definition of prediction models from the data gathered from the bed side monitors and the Hospital's records. Prediction results were encouraging [30]. The architecture proposed is presented in Fig. 4.

Scenario Evaluation was included in this first architecture proposal but it proved less reliable as it made longer term predictions which in turn were based on shorter term predictions. Even if it continues to be interesting to have the possibility of scenarios evaluation in such a system, the scenario evaluation agent was removed from the architecture due to its low accuracy at present.

Evaluation from the first cycle, both in terms of prediction results and insight from medical staff allowed for a new proposal where the data mining agent was complemented by an ensemble agent and all predictions were made using hybrid ensembles of neural networks and decision trees (Fig. 5). Also, the idea of having several of these systems spread through

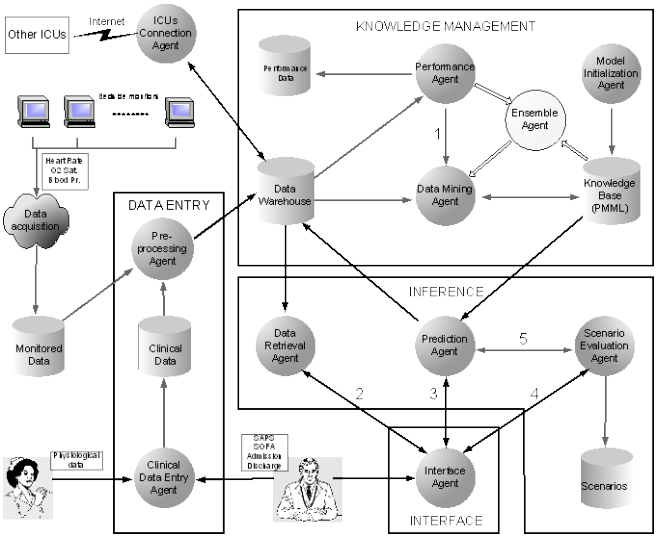


Fig.4. Initial INTCare architecture proposal.

several ICUs was abandoned and the team decided to concentrate its efforts in building a system for one ICU.

The algorithm used for ensemble evolution is based on the Dynamically Weighted Majority algorithm [31]. Using this algorithm, the system has an ensemble of individual predictors with weights associated and the final prediction found by weighing all individual predictions. By adjusting the individual weights to reflect each individual model performance we had the weights vary so that better performing models were given more relevance.

Evaluation of the ensemble agent had to be performed with resort to external data as it would be unfeasible to wait until enough data was collected by our system. In order to be able to evaluate the potential gains of having ensembles as opposed to single prediction models, some experiments were conducted using data gathered in ICUs across Europe in the EURICUS 2

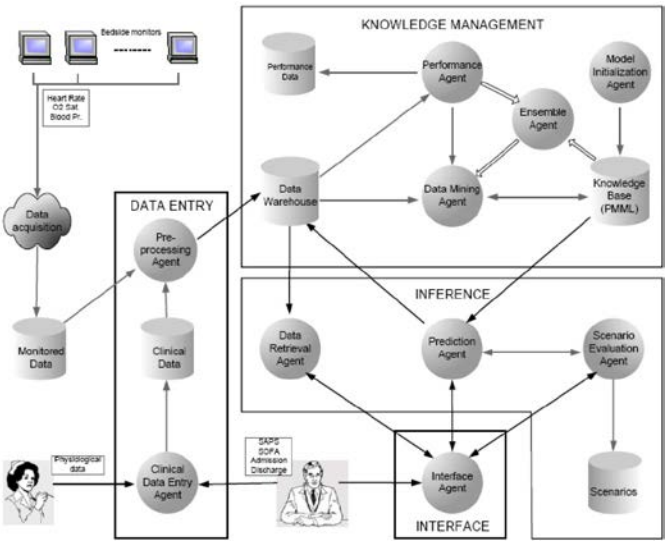


Fig.5. Revised INTCare architecture proposal.

ISBN: 978-1-61804-327-6

study [32] .

The results from the experiments showed that in this context ensemble models are capable of better results than single prediction models and that the chosen ensemble strategy was capable of updating itself while maintaining or even improving its prediction accuracy [33].

Even if we now had an acceptable architecture for an autonomous CDSS based on the KDD process, this experiments gave rise to new questions related to the evolution mechanism. Indeed, the next Design Cycle was devoted to investigating the effects of allowing the evolution of the ensemble include the creation and elimination of individual models.

To evaluate the results we used the average of the values of the area under the Receiver Operating Characteristic curve (AUC ROC) obtained after 30 runs of each experiment. The ROC curves are often used in the medical area to evaluate computational models for decision support, diagnosis and prognosis [34][35].

Results achieved showed a consistent increase in performance thus validating the algorithm. The experiments designed allowed the evaluation of three different strategies, with configuration A being a static ensemble, configuration B an ensemble where the individual model's weights were updated as to reflect each model's performance and configuration C expanded on configuration B by having the addition and deletion of models from the ensemble when their weights crossed predefined thresholds. Results achieved showed configuration C to consistently provide the best results in terms of AUC ROC [36] (Fig. 6). Consequently, this is the algorithm to be used for making predictions in INTCare.

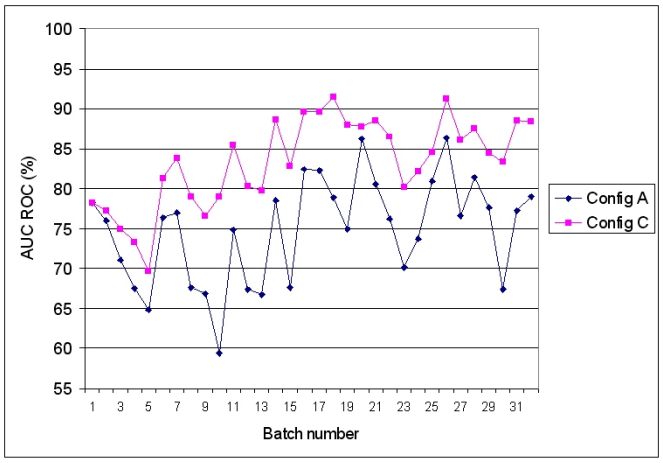


Fig.6. Evaluation of the performance of the ensembles

V. Conclusions And Future Work

Design science research is being successfully used in building a CDSS proposal for use in ICUs. Our objective is the definition of the architecture as building and evaluating such a system in a real setting is a very time consuming process and is outside of the scope of our work. However, the main components have been shown to works as expected and the concept has been shown to be a valid one. It has been shown that using the weights update algorithm, the accuracy of the system is consistently better than the accuracy achieved using a static ensemble.

When in operation, INTCare is expected to provide timely predictions using all data available as it is designed as to be able to connect to every available data source. Such capacity eliminates the need for manual data entry from de medical and nursing staff and thus increases the system's ease of use and likelihood of being used.

Future work will include evaluating the possibility of having several predictions being made and presented to the medical staff (as opposed to the current proposal of presenting only one prediction). Other iterations will include a longer term cycle for a time extended evaluation of a more complete prototype.

Acknowledgment

The authors would like to FCT (Foundation of Science and Technology, Portugal) for the financial support through the contract PTDC/EEI-SII/1302/2012 (INTCare II). This work has been supported by FCT - Fundação para a Ciência e Tecnologia within the Project Scope UID/CEC/00319/2013.

References

[1] G. Manjoney, "Clinical information systems market – An insiders view" Journal of Critical Care, vol. 19, no 4, pp. 215-220, 2004.

[2] E. Zanier, F. Ortolano, L. Ghisoni, A. Colombo, S. Losappio and N. Stocchetti, "Intracranial pressure monitoring in intensive care:clinical advantages of a computerized system over manual recording" Journal of Critical Care, vol. 11, no 1, R7, 2007.

[3] E. Coiera, *Guide to Health Informatics*, 2nd Ed, Hodder Education, 2003.

[4] J. Carter, E. Berner (eds) *Clinical Decision Support Systems – Theory and Practice Design and Implementation Issues*, pp. 64-98, Springer, 2007.

[5] P. D. Clayton and G. Hripcsak, "Decision Support in Healthcare", International Journal of Biomedical Computing, vol. 39, pp. 59-66, 1995.

[6] A. Silva, P. Cortez, M. F. Santos, L. Gomes and J. Neves, "Multiple organ failure diagnosis using adverse events and neural networks". In: Seruca I. et al. (eds), Enterprise Information Systems VI, Springer, 2005.

[7] A. Hevner, S. March, J. Park and S. Ram, "Design Science in Information Systems Research" MIS Quarterly, vol. 28, pp 75-105, 2004.

[8] A. Hevner, "A Three Cycle View od Design Science Research" Scandinavian Journal of Information Systems, vol 19, no 2, pp 87-92, 2007.

[9] J. Iivari, "A Paradigmatic Analysis of Information Systems as a Design Science", Scandinavian Journal of Information Systems, vol 19, no 2, pp 39-64, 2007.

[10] H. Simon, *The New Science of Management Decisions*, Rev. ed. Englewood Cliffs. Prentice Hall, New Jersey, 1977.

[11] M. Mora, J. Forgionne and J. Gupta, *Decision Making Support Systems – Achievements and Challenges for the new decade*, Idea Group Publishing, 2003.

[12] E. Turban, J. Aronson and T. Liang, *Decision Support Systems and Intelligent Systems*, Prentice Hall, New Jersey, 2004.

[13] J. R. Le Gall, S. Lemeshow and F. Saulnier, "A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study", JAMA, vol. 270, no. 24, pp. 2957-2963, 1993.

[14] W. A. Knaus et al., "The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults". Chest, vol. 100, no. 6, pp. 1619-1636, 1991.

[15] A. Barnato and D. Angus, "Value and role of intensive care unit outcome prediction models in end-of-life decision making". Crit Care Clin, vol. 20, no. 3, pp. 345-362, 2004.

[16] F. Portela, P. Gago, M. F. Santos, J. Machado, A Abelha, A. Silva and F. Rua. "Implementing a Pervasive Real-time Intelligent System for Tracking Critical Events with Intensive Care Patients". IJHISI - International Journal of Healthcare Information Systems and Informatics. Vol 4, pp.1-16, 2013.

[17] S. Lee and K. Siau, "A review of data mining techniques". Industrial Management & Data Systems, vol. 101, no. 1, pp. 41-46, 2001.

[18] Y. Kim and W. N. Street, "An intelligent system for customer targeting: a data mining approach". Decision Support Systems, vol. 37, no. 2, pp. 215-228, 2004.

[19] N. C. Hsieh, "An integrated data mining and behavioral scoring model for analyzing bank customers". Expert Systems with Applications, vol. 27, no. 4, pp. 623-633, 2004.

[20] A. Musaev, "Analytic information technologies in oil refinery". Expert Systems with Applications, vol. 26, no.1, pp. 81-85, 2004.

[21] M. Lam, "Neural network techniques for financial performance prediction: integrating fundamental and technical analysis". Decision Support Systems, vol. 37, no. 4, pp. 567-581, 2004.

[22] L. Li, H. Tang et al., "Data mining techniques for cancer detection using serum proteomic profiling". Artificial Intelligence in Medicine, vol. 32, no. 2, pp. 71-83, 2004.

[23] S-h. Liao, "Knowledge management technologies and applications-literature review from 1995 to 2002". Expert Systems with Applications, vol. 25, no. 2, pp. 155-164, 2003.

[24] D. Hand, H. Mannila and P. Smyth, *Principles of Data Mining*. MIT Press, Cambridge, MA, 2001.

[25] J. Quinlan, "Induction of Decision Trees", Machine Learning, vol. 1, no.1, p. 81-106, 1986.

[26] N. Lavrac, E. Keravnou and B. Zupan, *Intelligent Data Analysis in Medicine and Pharmacology*. Kluwer, Boston, 1997.

[27] S. Haykin, *Neural Networks – A Compreensive Foundation (2nd Edition)*, Prentice-Hall, New Jersey, 1999.

[28] R. Dybowski, "Neural Computation in Medicine: Perspectives and Prospects". In: Malmgreen et al. (eds) Proceedings of the ANNIMAB-1 Conference (Artificial Neural Networks in Medicine and Biology), Springer, Berlin Heidelberg New York, pp 26-36, 2000.

[29] T. Dietterich, "Ensemble methods in machine learning". In: Kittler and Roli (eds), Multiple Classifier Systems, LNCS 1857, Springer, pp 1-15, 2001.

[30] P. Gago, M. F. Santos, A. Silva, P. Cortez, J. Neves and L. Gomes. "INTCare: a knowledge discovery based intelligent decision support system for intensive care medicine". Journal of Decision Systems, vol. 14, no. 3, pp. 241–259, 2005.

[31] J. Z. Kolter and M. A. Maloof, `Dynamic weighted majority: A new ensemble method for tracking concept drift', Third IEEE International Conference on Data Mining, 2003.

[32] D. Miranda, R. Nat, A. Nijk, W. Schaufeling and G. Lapichino, "Nursing activities score". Crit Care Med, vol. 31, no. 2, pp. 374-382, 2003.

[33] P. Gago, A. Silva and M. F. Santos, "Adaptive Decision Support for Intensive Care", *Progress in Artificial Intelligence*, pp. 415-425, Springer, 2007.

[34] M.H. Zweig and G. Campbell, `Receiver-operating characteristic (roc) plots: a fundamental evaluation tool in clinical medicine', Clin Chem, vol. 39, no. 4, pp. 561-577, 1993.

[35] T. Lasko, J. Bhagwat, K. Zou, and L. Ohno-Machado, `The use of receiver operating characteristic curves in biomedical informatics', J. of Biomedical Informatics, vol. 38, no. 5, pp.404-415, 2005.

[36] P. Gago and M. F. Santos, "Evaluating Hybrid Ensembles for Intelligent Decision Support for Intensive Care", in O. Okun and G. Valentini (eds) Applications of Supervised and Unsupervised Ensemble Methods Studies in Computational Intelligence, Sprionger, Vol. 245, pp 251-265, 2009.

Feature selection for detecting patients with weaning failures in Intensive Medicine

Sérgio Oliveira, Filipe Portela, Manuel Filipe Santos, José Neves, Álvaro Silva and Fernando Rua

***Abstract*—** In Intensive Care Units most of the admitted patients are mechanically ventilated. The process of ventilator weaning is delicate and it is conducted by following a set of steps. Normally a weaning tentative is executed based in the patient condition (by analyzing ventilation parameters) and physician's knowledge. In some cases this process fails and it causes long term injuries to the patients. The main goal of this work it is to detect patterns to non-successful weaning in order to avoid a wrong tentative and consequently improve patient condition. Clustering data mining was used to select and identify the features and the patterns associated to failures. As result an Index-Davies Bouldin of 0.9819 was achieved. This result represents the better variables symmetric among the clusters created.

***Keywords*—** Ventilation Weaning, Intensive Medicine, Intensive Care Unit, Respiratory Diseases, Data Mining, Clustering, INTCare, Mechanical Ventilation, Extubation, Feature Selection.

I. Introduction

Respiratory failure is one of the most common causes of Intensive Care Unit (ICU) admission. 75% of the patients admitted in an ICU require mechanical ventilation. Despite its benefits, these procedures might have some serious drawbacks contributing to promote lungs injury.

Mechanical Ventilation (MV) is one of the most delicate processes in Intensive Medicine. Although MV had an important role in patient life support, a wrong procedures or configuration can provoke long injuries to the patients.

Mechanical ventilation can have negative effects and its mortality rate ranges is from 41% to 65% [1]. The number of re-intubations vary from 2% to 25% [2].

Automatic control of mechanical ventilation can significantly improve patient care in the ICUs, reduce the mortality and morbidity rates associated with provision of inappropriate ventilator treatments and reduce healthcare costs.

The main goal of this work it is characterize patients whom it is not advised to make a weaning. In this work the objective is not to predict if a patient can be or not extubated but create patterns of patients that should not be put in a weaning process.

This work was supported by the use of Clustering Data Mining techniques (K-Means and K-Medoids) to create patters of weaning failures. Clustering is considered a task of grouping a set of data in such a way that objects in the same group (called a cluster) are more similar.

The work is framed in INTCare project and it used real data collected in real-time from the ICU of Hospital Santo António, Centro Hospitalar do Porto (CHP), Portugal.

As result it was possible to make a feature selection and identify a set of patient characteristics associated to weaning failures. The Index-Davies Bouldin achieved a symmetry inter-clusters of 0.9819. The features with most impact and identified by the better cluster were: CDYN, CSTAT, Flow and Support Pressure.

The paper is divided in seven sections after introduce the work a set of related concepts are presented in the second section. In the third section is presented the work context. Section four present the work developed following CRISP-DM methodology. In section five the results achieved are analyzed having in consideration the main target. Finally some conclusion are written and future work defined.

II. Background

A. Intensive Care Units

The patients who are admitted to Intensive Care Units (ICU) are constantly monitored. They have a set of sensors connected from the body to the bedside monitors. The most common monitoring process is patient vital signs and mechanical ventilation. In fact these patients need intensive care and typically they are in a risk-life condition, being their life condition supported by ventilators.

The main goal of intensive medicine it is use the artefacts available in ICU and the intensivists knowledge to diagnose and treat patients with serious illnesses, restoring them to their previous health condition [3].

Most recently became more difficult to make decision in ICU taking in attention all the data collected from the patient.

This work was FCT - Fundação para a Ciência e Tecnologia within the Project Scope UID/CEC/00319/2013.

Sérgio Oiveira is with Algoritmi Research Centre, University of Minho, Portugal.

Filipe Portela is with Algoritmi Research Centre, University of Minho, Guimarães, Portugal. (Corresponding author to provide phone: +351253510319; fax: +351253510300; e-mail: cfp@dsi.uminho.pt).

Manuel Filipe Santos, is with Algoritmi Research Centre, University of Minho, Guimarães, Portugal. (e-mail: mfs@dsi.uminho.pt).

José Maia Neves is with Algoritmi Research Centre, University of Minho, Braga, Portugal. (e-mail: {jneves@di.uminho.pt).

Álvaro Silva and Fernando Rua are with Intensive Care Unit of Centro Hospitalar do Porto, Portugal (e-mail: {moreirasilva@me.com; fernandorua.sci@hgsa.min-saude.pt).

The existence of vital signs monitors and ventilators allow a continuous data streaming. This situation difficult the data analysis in a short period of time due to a high number of data collected. ICU is a potential source of implicit knowledge. This knowledge can be used to improve the decision making process.

B. Mechanical Ventilation and Weaning

Respiratory failure is a syndrome in which the respiratory system fails in one or both of its gas exchange functions: oxygenation and carbon dioxide elimination [4].

The goal is to reduce lung injury due to over distention. However, the efficacy of this approach has not been established [5]. To overcome this problem the patients are ventilated using artificial ventilation.

Nowadays mechanical ventilator, are only used by the clinicians to consult the patient values. The data observed are not stored in a database. This situation results in a wasting of data that could be transformed in knowledge and it could be very useful to the decision-making process.

In addition, the process of ventilator weaning is based in a medical assumption [6] and in a tentative-error procedure, which sometimes seriously compromises the patient condition.

Mechanical ventilation is commonly used in ICU and it is very important to treat many different illnesses, however is relatively costly [2].

Weaning is a gradual process of liberation from, or discontinuation of, mechanical ventilator support resulting in an extubation. In Intensive Medicine an extubation process is considered successful when a patient can breathe from himself for a period upper than one hour.

The Intelligent Decision Support Systems (IDSS) for mechanical ventilation can be grouped in two types: an expert advisory systems or an automatic control of ventilation or weaning [7]. In the ICUs there is a set of IDSS system to ventilators, however, most of them are rule-based system. They are not adaptive and they do not use the results obtained to improve the models.

After an overview [7] it is also possible to verify that most of the existing systems is not using data mining to predict the results. The most far as they can go it is in the input data that can be based in clinical rules and guidelines. Many of its rules can be adaptive and can be derived on the basis of physiological models.

C. INTCare

In the ICU of CHP was deployed a Pervasive Intelligent Decision Support System (PIDSS). This PIDSS is in a continuous developing and test and it is a result of INTCare project [8].

INTCare system is able to monitoring the patient condition in real-time by collecting, processing and displaying the information collected from the bedside monitors and other hospital sources in an intuitive and easy way.

It also has a module to support the decision process through Data Mining models. This module can induce in real-time and using online learning several models able to predict clinical events, as is for example patient outcome [9], organ failure [9], length of stay [10], readmission [11, 12] and barotrauma [13], among others.

INTCare uses intelligent agents [15, 16] to perform their tasks automatically and without human intervention.

This work is inserted in the second phase of the project where the main concern is the respiratory system.

After make a first research to predict barotrauma [13, 14] now it is time to explore a new field: weaning and extubation.

D. Data Mining

Data Mining is the process of using artificial intelligence techniques and statistical and mathematical functions to extract knowledge from the data stored in the database. The achieved knowledge can be presented in multiple forms: business rules, similarities, patterns or correlations [17]. Clustering is inserted in the group of Data Mining problems.

Clustering has as main goal divides the data collected in datasets with similar values. The groups created by the clusters represents a natural catchment of the data and data aggregations. The groups created should make sense, be helpful or both. Clustering rules are not pre-defined. They are discovered along the clustering process. The clusters are characterized by a great internal homogeneity and external heterogeneity [18].

The use of cluster to identify groups of variables is an important asset in many areas like psychology and social sciences, biology, statistical, pattern recognition, information recovery, machine learning and data mining [19, 20].

Clustering offers a high number of algorithms. The choice of the best algorithm to use it is depending from the data collected and project goal. The majority of the clustering methods are grouped into five categories.

The hierarchical methods execute a hierarchical decomposition of the data. These methods can be divisive or agglomerative.

Divisive methods behave the other way. The density-based methods are useful to filter outliers or discovering data with arbitrary form.

The agglomerative methods start with singular objects to create an isolated group. Then the groups or objects are successively merged until a group is missing.

The Partition Methods build a set of partitions on the data, where each partition represents a cluster.

Grid-based methods restrict the space of objects to a finite number of cells forming a grid structure. The Model-based methods formulate a model hypothesis for each cluster and find the best fit the data to the model [21].

Clustering assessment can be done by laying on two factors: compactness and separability. The compactness expresses how much the cluster elements are near. How lesser the variance value it is, greater it will be the cluster compactness. The calculation of the intra cluster distance is very useful to assess this characteristic. The separability evaluates how diverse the clusters are. This can be evaluated by the inter-cluster distance. How higher the distance is better the clusters are [22].

III. Material And Methods

Cross Industry Standard Process for Data Mining (CRISP-DM) was the methodology chosen to conduct this work. CRIS-DM is a cycle process divided in six steps: Business Understanding, Data Understanding, Data preparation, Modelling, Evaluation and Deployment. These steps provide a structured approach to planning a data mining project. In this work only the last phase was not performed.

To do this work R tool were used. R is presented as an environment of statistical programming language for development [23]. The library "cluster" was used primarily to implement the k-means algorithms and Partitioning Around Medoids (PAM) and then for graphing. For optimum number of cluster number it was used the library "fpc". To evaluate the clusters the library "clusterSim" and the Davies-Bouldin Index were used.

IV. Knowledge Discovering Process

The process of knowledge discovery using DM techniques is very complex. As mentioned the Cross Industry Standard Process for Data Mining (CRISP-DM) methodology was followed to guide the present study.

A. Business Understanding

To identify the patient ventilation variables (feature selection) which interfered in a non-successful weaning / extubation is the main goal of this work. The goal is not to predict if a patient is prepared to be weaned but to identify patterns and clusters of data associated to weaning failures. The clusters were designed using only data monitored by ventilators; the values used were numeric and they were from discrete quantitative type.

Clinically it is expected to create new knowledge to the intensivists helping them to take the better decision in the moment of weaning a patient.

B. Data Understanding

The data used to conduct this study were exclusively collected from ventilators in the ICU of CHP. This data corresponds to 50 patients comprising a period between 2014-09-19 and 2015-02-03 in a total of 15325 records. Each record contains thirteen fields:

- CDYN – (F_1): Dynamic compliance in mL/ cmH_2O;
- CSTAT – (F_2): Static compliance from inspiratory pause measured in mL/ cmH_2O;
- FIO2 – (F_3): Fraction of inspired oxygen (%);
- Flow – (F_4): Peak flow setting in litters per minute;
- RR – (F_5): Respiratory rate setting in berths per minute;
- PEEP – (F_6): Positive End-Expiratory Pressure in cmH_2O;
- PMVA – (F_7): Mean airway pressure in cmH_2O;
- Plateau pressure –(F_8): End inspiratory pressure in cmH_2O;
- Peak pressure – (F_9): Maximum circuit pressure in cmH_2O;
- RSTAT – (F_10): Static resistance from inspiratory pause measured in cmH_2O/L/s;
- Volume EXP – (F_11): Exhaled tidal volume in litters;
- Volume INS – (F_12): Tidal volume settings in litters;
- Support Pressure – (F_13): Exhaled minute volume litters;

The values obtained with the coefficient of variation show that the distributions of most fields are mixed. Only the fields: Plateau Pressure, PMVA, Peak Pressure and FIO2 have a coefficient of variation lower than 20%. This measure of dispersion is the ratio between the standard deviation and the average.

Table 1 provides a statistical analysis of the fields used in this study. For each of the numerical values it was analyzed their minimum (MIN) and maximum (MAX), average (AVG), standard deviation (SD) and coefficient of variation (CV).

Table 1 – Distribution of variables

	MAX	MIN	AVG	SD	CV
Plateau Pressure	36	6.2	19.98	3.83	19.18
CDYN	200	0	46.31	40.45	87.34
CSTAT	71	0	20.88	20.93	100.22
PMVA	18	3.1	10.52	1.84	17.48
Peak Pressure	40	9	20.53	3.85	18.75
RSTAT	29	0	8.89	8.60	96.74
FIO2	100	35	49.68	7.43	14.96
FLOW	60	0	23.83	23.61	99.06
RR	24	0	1.69	5.50	324.73
PEEP	10	3	5.09	1.04	20.45
Volume EXP	2.46	0	0.53	0.17	31.83
Volume INS	0.56	0	0.26	0.25	98.06
Support Pressure	27	4	14.1	3.56	25.24

C. Data Preparation

To carry out this study it was necessary to identify the patients who were not able to be submitted to the extubation process, even though they were very close to being subjected to the process.

In order to identify these occurrences it was necessary to determinate the respective scenario. Patients who were not submitted to extubation process were patients who had mechanical ventilation variations or the support pressure level was continuous for more than one hour but they never were an attempt to extubation under this scenario. In this process five levels were identified:

- -2: Support pressure is constantly changing (less than 3 minutes);
- -1: Support pressure is continuous between 4 minutes and 30 minutes;
- 0: Support pressure is constant from 30 minutes to 60 minutes
- 1: Maintains the same support pressure for more than 60 minutes but the patient is not extubated;

- 2: Maintains the same support pressure for more than 60 minutes and the patient is extubated;

To this work only the patient with level equal to 1 were considered. After this first selection the records which presented at least one null value were deleted. This operation was of utmost importance, because the models cannot be induced if the dataset has null values. At same time clinical ranges were used to validate the values collected. Based in those information it was possible to consider only correct values (after be validated by an intelligent agent). In this process the records without quality and inconsistences (null values and values out of the normal range) were eliminated.

After all the changes are applied to the initial dataset, the number of records decreased to 13135 records, which is associated to 28 patients.

D. Modelling

The algorithms used to create clusters were: k-means and K-medoids. The choice of these two algorithms was fundamentally based on two characteristics: the principle of partition method and the difference in sensitivity to outliers.

K-means is a simple iterative clustering algorithm which partitions the dataset into a priori number. The algorithm is simple to implement and run, being their implementation faster [25]. The K-means is sensitive to outlier, because the objects are far from the majority, which it can significantly influence the average value of the set. This inadvertently affects the allocation of other objects to the clusters. This effect is particularly exacerbated by the use of the squared error function [24].

On the other hand the k-medoids instead of using the value of a cluster object as a reference point, takes on real objects and represents the clusters, dare one object at a cluster. The remaining objects are assigned to the most similar cluster. The partitioning method is then performed based on the principle teaches the sum of the differences between each of p and its corresponding object representative object [24]. K-medoids algorithm is similar to the K-means, except that the centroids must belong to the set of data that are grouped [25].

The implementation of the algorithms required some settings. The identification of the appropriate number of K was achieved by implementing the calculation of Square Error Sum (SSE). The SSE determines the squared distances between each member of the cluster and the cluster centroid.

$$SSE(o_i) = \sum_{i=1}^{n}\sum_{j=1}^{k} w_{ji}^{p}\, dist(o_i, c_j)^2$$

In general, when the number of clusters increases SSE value decreases. When it is identified a dramatic decrease in the SEE value, the number of K can be considered as being great [24].

There were ten plays for each dataset used for each model. The metric used to calculate the dissimilarity being the observations was "Euclidean". The developed models can be represented by the following expression:

$$M_n = \{A_f;\ F_i;\ D_x;\ AG_y\}.$$

The model M_n belongs to an approach (A) and it is composed by a set of fields (F), a type of variable (TV) and an algorithm (AG):

$$A_f = \{Discription\ (Clustering)_1\}$$
$$F_i = \{F\_1_1, F\_2_2, F\_3_3, F\_4_4, F\_5_5, F\_6_6, F\_7_7, F\_8_8, F\_9_9, F\_10_{10}, F\_11_{11}, F\_12_{12}, F\_13_{13}\}$$
$$TV_x = \{Qualitative\ variables\ ordinal_1\}$$
$$AG_y = \{K-means_1, K-medoids(PAM)_2\}$$

E. Evaluation

In this phase, the Davies-Bouldin Index (DBI) was used to evaluate the models generated.

Among the algorithms used the one which presented the most satisfactory results was the K-means algorithm.

Some of the models developed had interesting results but only one appears to present satisfactory results (DBI). Table 2 shows the most relevant models.

Table 2 – Models for clustering

Model	Fields	Number Clusters	Algorithm	DBI
M_1	$F_{\{1,2,3,4,5,6,7,8,9,10,11,12,13\}}$	8	AG_2	1.31
M_2	$F_{\{2,3,4,5,6,13\}}$	5	AG_2	1.10
M_3	$F_{\{1,2,4,13\}}$	7	AG_1	0.98
M_4	$F_{\{1,2,3,4,13\}}$	10	AG_1	1.10
M_5	$F_{\{1,2,3,4,11,13\}}$	10	AG_1	1.20

Analyzing Table 2 it is possible to identify that the model M_3 is best model generated. The Davies Bouldin Index tends to $+\infty$ however M_3 model has an index of 0.98. This is not the ideal case but it is the model closest to 0.

In order to present a representation of the data segments generated by the best model, the figure 1 was designed.

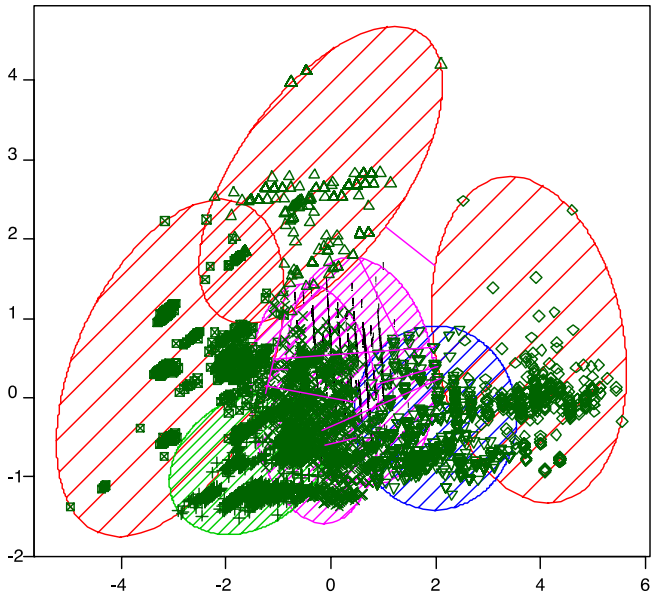


Figure 1: Graphical representation of the clusters presented in table 2

Since the goal is select and identify patient features which were not capable of being subjected to extubation process, it was necessary to identify the cluster that best characterizes these patients.

In order to identify the cluster that best identifies such patients, it was necessary to find the cluster which presented the lowest variability.

Cluster 3 was identified as being the best cluster. It has the lowest sum of squares with a ratio of 1247.02. This cluster contains about 2382 occurrences. Thus cluster 3 has the lowest average of the sum of the squares value: 0.52.

Table 3 shows the patient characteristics / features of Cluster 3. For example in the case of CDYN (F1) the minimum value is 13 and the maximum values is 20.

When a patient presents a value for each one of the four fields between the minimum and the maximum values of the cluster, he is associated to cluster 3.

Table 3 – Distributions of cluster 3

Cluster	Fields	MIN	MAX	MEAN	SD	CFV
Cluster 3	$F_{\{1\}}$	13	20	16.49	1.27	7.73%
	$F_{\{2\}}$	0	48	7.61	14.47	190.30%
	$F_{\{4\}}$	6.9	12	9.14	0.73	7.94%
	$F_{\{13\}}$	8	14	10.72	1.28	11.88%

V. Discussion

The best model M_3 created 7 clusters. Cluster 3 is the only able to properly identify the properties of the patients that should not be weaned / extubated.

The amount of Cluster 3 registers corresponds to 18.13% of the dataset used. The third cluster presents a significant amount of records and it displays relevant characteristics of the distribution of its variables.

There is great heterogeneity in the variables that make up the Cluster 3. $F_{\{2\}}$ is the variable that demonstrates to have more homogeneous values. In general the cluster presents variables with a little dispersion.

Although the clusters generated by the M_3 model have at least one interception with another cluster, cluster 3 is better prepared to target the patients who do not meet the conditions to be extubated. Since it has an average of sum of squares lesser than all the other clusters, presenting a value of 0.52.

VI. Conclusion

After conclude this study it was possible to identify a set of patient features / variables that have great similarity on weaning / extubation failures.

The most satisfactory result is attained by the model M3 obtaining a Davies-Bouldin Index of 0.98.

M3 is the model which presented results closest to the optimum value: zero. It should be noted that most of the variables used in cluster 3 has an acceptable dispersion, since its value is 13.47. This result was obtained with the implementation of the k-means algorithm, thus demonstrating to be the most suitable algorithm to carry out the data segmentation.

The achieved results gives confidence in continues this research work in order to make a better feature selection of the variables associated to a wrong weaning.

The better cluster Davies Bouldin Index is acceptable being it above to the maximum acceptable level (1). It presents a result of 0.98 which means that there is similarities in the clusters created.

Presenting the cluster in an intuitive way and easy to understand and combining it with the clinical experts it is possible to provide best care to the patients.

This work is a starting point to avoid wrong extubation and avoid long injuries associated to respiratory system (e.g. lungs injuries).

VII. Future Work

Having in consideration the achieved results it is expected to improve these results by adding other variables and explore other clustering techniques.

At same time it will try induce data mining models to predict the probability of a patient have a weaning and extubation.

In the future all the results achieved will be included in INTCare system in order to give more options to the intensivist in the moment of deciding.

Acknowledgment

The authors would like to FCT (Foundation of Science and Technology, Portugal) for the financial support through the contract PTDC/EEI-SII/1302/2012 (INTCare II). This work has been supported by FCT - Fundação para a Ciência e Tecnologia within the Project Scope UID/CEC/00319/2013.

References

[1] A. S. Fauci, "Harrison's Principles of Internal Medicine, 17e," ed: Silverchair Science: Minion, 2008.

[2] F. T. Tehrani, "Automatic control of mechanical ventilation. Part 2: the existing techniques and future trends," *Journal of clinical monitoring and computing*, vol. 22, pp. 417-424, 2008.

[3] F. Portela, M. F. Santos, Á. Silva, F. Rua, A. Abelha, and J. Machado, "Adoption of Pervasive Intelligent Information Systems in Intensive Medicine," *Procedia Technology*, vol. 9, pp. 1022-1032, 2013.

[4] A. Kaynar and S. Sharma. (2010, Respiratory Failure. 39. Available: http://emedicine.medscape.com/article/167981-print

[5] T. E. Stewart, M. O. Meade, D. J. Cook, J. T. Granton, R. V. Hodder, S. E. Lapinsky, *et al.*, "Evaluation of a Ventilation Strategy to Prevent Barotrauma in Patients at High Risk for Acute Respiratory Distress Syndrome," *New England Journal of Medicine*, vol. 338, pp. 355-361, 1998.

[6] S. P. Stawicki, "Mechanical ventilation: weaning and extubation," 2007.

[7] F. T. Tehrani and J. H. Roum, "Intelligent decision support systems for mechanical ventilation," *Artificial Intelligence in Medicine*, vol. 44, pp. 171-182, 2008.

[8] F. Portela, M. F. Santos, J. Machado, A. Abelha, Á. Silva, and F. Rua, "Pervasive and intelligent decision support in Intensive Medicine–the complete picture," in *Information Technology in Bio-and Medical Informatics*, ed: Springer, 2014, pp. 87-102.

[9] F. Portela, M. F. Santos, J. Machado, A. Abelha, and Á. Silva, "Pervasive and Intelligent Decision Support in Critical Health Care Using Ensembles," in *Information Technology in Bio-and Medical Informatics*, ed: Springer Berlin Heidelberg, 2013, pp. 1-16.

[10] R. V. Filipe Portela, Sérgio Oliveira, Manuel Filipe Santos, António Abelha, José Machado, Álvaro Silva and Fernando Rua, "Predict hourly patient discharge probability in Intensive Care Units using Data Mining," *ScienceAsia Journal (ICCSCM 2014)*, 2014.

[11] Pedro Braga, F. Portela, and M. F. Santos, "Data Mining Models to Predict Patient's Readmission in Intensive Care Units," in *ICAART - International Conference on Agents and Artificial Intelligence*, Angers, France, 2014.

[12] R. Veloso, F. Portela, M. F. Santos, Á. Silva, F. Rua, A. Abelha, *et al.*, "A clustering approach for predicting readmissions in intensive medicine," *Procedia Technology*, vol. 16, pp. 1307-1316, 2014.

[13] F. P. Sérgio Oliveira, Manuel Filipe Santos, José Machado, António Abelha, Álvaro Silva and Fernando Rua, "Intelligent Decision Support to predict patient Barotrauma risk in Intensive Care Units," in *Procedia Technology - HCIST 2015 - Healthy and Secure People*, Elsevier, Ed., ed, 2015.

[14] S. Oliveira, F. Portela, M. F. Santos, J. Machado, A. Abelha, Á. Silva, *et al.*, "Predicting Plateau Pressure in Intensive Medicine for Ventilated Patients," in *New Contributions in Information Systems and Technologies*, ed: Springer, 2015, pp. 179-188.

[15] L. Cardoso, F. Marins, F. Portela, M. Santos, A. Abelha, and J. Machado, "The Next Generation of Interoperability Agents in Healthcare," *International Journal of Environmental Research and Public Health*, vol. 11, pp. 5349-5371, May 2014.

[16] M. F. Santos, F. Portela, M. Vilas-Boas, J. Machado, A. Abelha, and J. Neves, "INTCARE - Multi-agent approach for real-time Intelligent Decision Support in Intensive Medicine," in *3rd International Conference on Agents and Artificial Intelligence (ICAART)*, Rome, Italy, 2011.

[17] E. Turban, R. Sharda, and D. Delen, *Decision Support and Business Intelligence Systems*, 9th Edition ed.: Prentice Hall, 2010.

[18] S. Tufféry, *Data mining and statistics for decision making*: John Wiley & Sons, 2011.

[19] H. C. Koh and G. Tan, "Data mining applications in healthcare," *Journal of Healthcare Information Management—Vol*, vol. 19, p. 65, 2011.

[20] P.-N. Tan, Steinbach, M., & Kumar, V, *Introduction to Data Mining* 1ed.: Addison-Wesley Longman Publishing Co., Inc., 2005.

[21] J. Han, M. Kamber, and J. Pei, *Data mining: concepts and techniques*: Morgan kaufmann, 2006.

[22] J. C. Krzysztof, W. S. Roman, and A. Lukasz, "Data Mining: A Knowledge Discovery Approach," ed: Springer, 2007.

[23] L. Torgo, *Data mining with R: learning with case studies*: Chapman & Hall/CRC, 2010.

[24] J. Han, M. Kamber, and J. Pei, *Data mining: concepts and techniques: concepts and techniques*: Elsevier, 2011.

[25] X. Wu and V. Kumar, *The top ten algorithms in data mining*: CRC Press, 2009.

Role Of Document Attributes In Information Retrieval

Benjamin Ghansah
Data Link Institue
School Of Computer Science
Tema, Ghana
ben@datalink.edu.gh

Nathaniel Ghansah
Data Link Institue
School Of Computer Science
Tema, Ghana
n.ghansah@datalink.edu.gh

Abstract— This paper was to explore how different fields or attributes of a document that is available to a retrieval model could influence retrieval performance, and also how it would perform when a diversity algorithm is applied: is it better to use only the title text for retrieval than using the abstracts? Or do we get the best performance when we combine all three fields together? We used the Vector Space model with TFIDF weighting retrieval algorithm to rank documents in each collection(version) and later rerank with the Expected Reciprocal Rank on the Cystic Fybrosis (CF) collection. The result shows that, different kinds and magnitude of information about collections(documents) influence retrieval performance and in particular, the more information which is made available to a retrieval model, the more effective it would perform.

Keywords—diversification; information retrieval; ERR;

I. INTRODUCTION

Studies have shown that, the Web has become the major and most visited 'unrestricted' source of satisfying a user informational need[7], with Web search engine serving as the primary method for finding relevant information on the Web. However, according to [3], Search engines usually provide enormous search-result output that contains relevant results which are relatively similar. This phenomenon is usually as a result of web searchers specifying their information needs by providing a single query term. These query terms, in most cases, are ambiguous, multifaceted and have more than one interpretation [2, 9]. A situation properly known as *query abandonment* occurs when a user does not find any relevant document from the list of returned search result, hence current study on search result diversification mainly aims at minimizing query abandonment by providing diversified search results in top positions[1] .

Consider, for example, a single query *apple*. This may refer to the Apple company, the Apple fruit, Apple TV or the Apple store. Without further information to make the user intent more explicit, one has to come up with strategies that can best produce a set of diversified results, capable of covering these diverse interpretations. A more complex situation arises, when for instance, users submitting the query *Apple* company can still have very diverse intents, such as finding a particular apple product, obtaining some reviews by a specific apple product, discovering forums, checking latest releases, etc.[11]. Intuitively, the problem of search result diversification can be stated as follows: Given a set *G* of *m* available items and a constraint *k* on the number of wanted results, the objective is to select a subset *x* of T items out of the *m* available ones, such that, the diversity among the items of *x* is maximized. One of the earliest work on diversification is the work by [3]. They proposed the maximal marginal relevance (MMR) ranking strategy to balance the relevance and the redundancy among the returned documents. In their work, a tradeoff between novelty (a measure of diversity) and the relevance of search results is made explicit through the use of two similarity functions, one measuring the similarity among documents, and the other the similarity between document and query. A parameter controls the degree of tradeoff. In this technique, there is no categorization of either the document or the query, diversification is conducted through the choice of similarity functions. The similarity measure used in MMR is query and intent independent, which renders the control over the match between relative redundancy levels and user intents difficult at best. Note that most diversification algorithms share the shortcoming of MMR; the similarity function is query and intent independent [8, 10, 13]. In summary MMR maximizes the margin relevance of the documents and iteratively select the document that is not only relevant to the query but also dissimilar to the previously selected documents. Though MMR diversification technique is pivotal for most results diversification algorithms and it led to several follow-ups, we use the Expected Reciprocal Rank (ERR) metric proposed by [4] due to the shortcoming of MMR discussed above. We do not claim ERR is the best diversification technique, we adopted this algorithm because of the following reasons: 1) ERR is a generalization of NDCG [5] 2) ERR was one of the metrics used to evaluate the diversity task in the Web track of TREC 2010 [1] .

[1] www.plg.uwaterloo.ca/~trecweb/2010.html

ISBN: 978-1-61804-327-6

This study seeks to explore the most cost effective ways in achieving efficiency and also the kind of information needed from text collections to achieve effectiveness, especially in enhancing search result diversification. In particular, we want to explore attributes of a collection with: i)**Titles only**; a collection containing only the text of the title fields for each document ii)**Abstract only**; a collection containing only the text of the abstract fields for each document iii) **Subject headings only**; a collection containing only the text of the major and minor subject headings for each document and finally iv)**Title + abstract + subject headings**; a collection containing the text from all three fields combined.

These four different collections will enable us to explore how different fields influence retrieval performance in terms of diversification: is it better to use only the title text for retrieval than using the abstracts? Or do we get the best performance(in terms of relevance and diversity) when we combine all three fields together? Based on the questions above, we investigate which field would produce an effective and efficient search result that is efficient and effective in terms of both *relevance* and *diversity*. We use a surrogate of the Cystic Fibrosis Database [12]. Section 2 presents the Experimental Setup and Section 3 presents the empirical results.

II. EXPERIMENTAL SETUP

A. Cascade-based metrics for ERR

We used the cascade user model proposed by [6] and described in algorithm 1. Note that idea of a cascade-based metric is to use the relevance labels as approximation of the probability that the user will be satisfied by the document in position t. Specifically, it has been suggested by [4] to estimate Rt as R(rt).

Algorithm 1 The cascade user model

Require: R1,….R10 the relevance of the 10 documents on the result page.
1: t = 1
2: User examines position t.
3: if random(0,1) ≤ Rt then
4: User is satisfied with the t-th document and stops.
5: else
6: t ← t + 1; go to 2
7: end if

Given a declining utility function φ, the metric is defined as the expectation of '(k), where k is the position where the user finds the document that satisfies him. This quantity turns out to be:

$$\sum_{k} \varphi(k)\mathcal{R}(r_k) \prod_{j=1}^{k-1}(1 - \mathcal{R}(r_j)) \quad (1)$$

Note the ERR [4] algorithm is an instantiation of equation 1 above, with φ (k) = 1/k:

We used the Lemur retrieval toolkit to perform our retrieval experiments on the Cystic Fybrosis (CF) collection [12], the Cystic Fibrosis (CF) database is composed of a collection of 1,239 documents describing clinical and research studies associated with CF. These documents contained in the collection were published between 1974 and 1979 and indexed with the term Cystic Fibrosis in the National Library of Medicine's A4EDLINE database. It also has a set of 100 queries with the respective relevant documents as answers.

Each query includes a query number and text, the record number of each relevant document in the answer, and relevance scores.

The relevance scores are from 4 different sources: REW (one of the authors), faculty colleagues of REW, post-doctorate associate of REW, and JBW (other author and a medical bibliographer).

The relevance scores vary from 0 to 2 with the following meaning:

2 HIGHLY relevant
1 MARGINALLY relevant
0 NOT relevant

As stated above, we used separate versions for our experiments. We have done this for two reasons: (1) the original collection is not in a format that can be indexed by Lemur [2], and (2) not all of the fields in these CF documents are of equal value in retrieving relevant documents. We have created four separate versions of the CF collection: Titles only, Abstract only, Subject headings only and Title + abstract + subject headings versions. We used the simple Vector Space model with TFIDF weighting retrieval algorithm for all our retrievals since it is the default in Lemur. After each retrieval, we rerank each search result with the ERR algorithm described above to obtain diversity.

III. EXPERIMENTAL RESULTS

Table 1: Document Precision on CF database with Top 20 Documents of each Collection

Collection (version)	P@5	P@10	P@20
Titles only	0.04	0.08	0.18
Subject headings only	0.11	0.12	0.22
Abstract only	0.25	0.31	0.35
Title + abstract + subject headings	0.45	0.63	0.69

[2] http://www.lemurproject.org/

ISBN: 978-1-61804-327-6

The result from table 1 shows that, the more information available to the retrieval model, the more accurate the retrieval. The results show that *Title + abstract + subject headings*, which is also referred to as *snippet* of the available document in the collection, based on the results we can conclude that different fields influence retrieval performance, and in particular, the more information available to the retrieval model, the more effective it performs.

Table 2: Document Precision on CF database with Top 20 Documents of each Collection with a diversity measure ERR

	ERR		
Collection (version)	**P@5**	**P@10**	**P@20**
Titles only	0.16	0.19	0.28
Subject headings only	0.31	0.32	0.39
Abstract only	0.35	0.42	0.45
Title + abstract + subject headings	0.49	0.53	0.89

The results from table 2 shows that, reranking the search results with the ERR diversity algorithm further increases the precision in all metrics. This shows that explicitly considering relevance and diversity indeed improves a user search experience more effectively, instead of concentrating only on relevance. Again the results shows that *Title + abstract + subject headings* outperform the other versions. The reason is similar in the explanation above.

IV. CONCLUSION

In conclusion we have shown that different fields influence retrieval performance, and in particular, the more information which is made available to a retrieval model, the more effective it would perform. Future work could explore the possibility of assertion in a more complex environment, such as a distributed information retrieval environment.

REFERENCES

1. Agrawal, R., S. Gollapudi, A. Halverson, and S. Ieong. *Diversifying search results.* in *Proceedings of the Second ACM International Conference on Web Search and Data Mining*. 2009. ACM.
2. Anagnostopoulos, A., A.Z. Broder, and D. Carmel, *Sampling search-engine results.* World Wide Web, 2006. **9**(4): p. 397-429.
3. Carbonell, J. and J. Goldstein. *The use of MMR, diversity-based reranking for reordering documents and producing summaries.* in *Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval*. 1998. ACM.
4. Chapelle, O., S. Ji, C. Liao, E. Velipasaoglu, L. Lai, and S.-L. Wu, *Intent-based diversification of web search results: metrics and algorithms.* Information Retrieval, 2011. **14**(6): p. 572-592.
5. Clarke, C.L., M. Kolla, G.V. Cormack, O. Vechtomova, A. Ashkan, S. Büttcher, and I. MacKinnon. *Novelty and diversity in information retrieval evaluation.* in *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*. 2008. ACM.
6. Craswell, N., O. Zoeter, M. Taylor, and B. Ramsey. *An experimental comparison of click position-bias models.* in *Proceedings of the 2008 International Conference on Web Search and Data Mining*. 2008. ACM.
7. Ghansah, B. and S. Wu. *Distributed Information Retrieval: Developments and Strategies.* in *International Journal of Engineering Research in Africa*. 2015. Trans Tech Publ.
8. Lafferty, J. and C. Zhai. *Document language models, query models, and risk minimization for information retrieval.* in *Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval*. 2001. ACM.
9. Radlinski, F. and S. Dumais. *Improving personalized web search using result diversification.* in *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*. 2006. ACM.
10. Rafiei, D., K. Bharat, and A. Shukla. *Diversifying web search results.* in *Proceedings of the 19th international conference on World wide web*. 2010. ACM.
11. Santos, R.L., C. Macdonald, and I. Ounis. *Exploiting query reformulations for web search result diversification.* in *Proceedings of the 19th international conference on World wide web*. 2010. ACM.
12. Shaw, W.M., J.B. Wood, R.E. Wood, and H.R. Tibbo, *The cystic fibrosis database: Content and research opportunities.* Library & information science research, 1991. **13**(4): p. 347-366.
13. Wang, C., Y. Liu, M. Zhang, S. Ma, M. Zheng, J. Qian, and K. Zhang. *Incorporating vertical results into search click models.* in *Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval*. 2013. ACM.

Proposal Study Desert Forest near Es-Sider Oil Port NW Libya

Fathi Elosta

***Abstract*—-** Located that oil port of siders following waha Oil Company which is located west of sirt on covering the moving sand dunes pot dominated system desertification.

The solar energy will be used to evaporated brine and generate cool air pure water in order to grow food inside the greenhouses where the sea water will be used facilities.

Concentrated idea of the project to exploit an area of about 20km through which pipeline to flow of sea water to the ground which may allow the lows of Nature which produce fresh water from sea water and change climate conditions to humid climate contributes to the rainfall and characterized the technology that will be used in the project to its ability to transform large tracts of desert into green produce commercial quantities of food and energy crops and freshwater used the project is based on the establishment of greenhouse where sea water used provide growth conditions cold and wet of vegetables in addition to the production of fresh water and will be connected to the solar with thermal units for the desalination of sea water supported by technology and will allow pilot also cultivation of algae in the system of bioreactors optical and farming systems in basins exposed produces plant food ,water, energy between greenhouse protected salt water and solar power intensive and cultivation of algae and technological and scientific . the basic idea of this project is to bring sea water into the desert then vaporized process centered on what is known as sea water where the evaporating sea water by solar energy and then intensify freshwater not only that ,but maintains a moist interior cool and are ideal for the cultivation of crops ,it was the first model to sea water greenhouse a unit area of 2000 square meters in port Augusta in Australia and has won its first crop tomatoes in the last months ago ,Sahara forest project combines the technology and solar power uses mirrors to focus the light on the complex to heat which in turn produces steam for the management of turbine generator at a supplying greenhouse powered and operated and the greenhouse extended solar power plant with water marking them more efficient that the project is promising and important in the field of renewable energy and to take advantage of the potential and of solar energy in our integrated project and ambitious aims to cultivate bio-fuels to generate an energy clean aimed at protecting the environmental may work on climate change

The project is based on the new Technologies in carpeted into solar energy and seawater Greenhouse and cultivation of green algae and for the cultivation of arid areas and the provision of food and fresh water as well as bio-fuels to generate clean energy.

The idea of the project is simple and Natural based on the use of solar energy in the evaporation of salty seawater to generate cold air and water and clean energy in deserts and Technology use Mirrors sunlight to form steam to Run turbines which in turn generate electricity is expected to shift the equipment desert to the territory to the a prosperous agricultural out about the need to drill water wells which increase water shortage and possible planting any kind of vegetables in greenhouse sea water are essential in the project which cost to be cooling desert air warm and ? Wet with sea water before being admitted to the greenhouse and this air humidifier will feed crops inside greenhouse then dryer where flowing sea water heated warmly sun? and soon meet air humidifies of channels containing water condenses fresh water and dripped drops on the external aspects of channels that can be assembled the process mimic what happens in Nature evaporates sea water as a result of the sun's heat and work is underway to cooling leads to the formation of clouds and condense these vapors of water to fall water like rain only 10-15 % of moist air condenses into fresh water and the rest out of the surrounding tress thus the greenhouse creating green surroundings around the will be through energy intensive and through the use of Mirrors to focus sunlight on pipes and boilers for water district cause radiation intensive production high heat inside the pipe provides enough power to turn Turbines team which in turn would for electrical power and six cell additional power to the people are a population of benefiting says a specialist and is Neil Crompton one energy experts there is huge potential for these new Technologies unit focused and one covering only 9% of the area of deserts ground and can provide five global need for energy as well as can invest this technique streaming tons of sea water per day using only 20 thousand greenhouse .

***Keywords*— L** I B YA **T** R I P O L I

I. INTRODUCTION

Characterize the oil port of Sidra good location along the Libyan coast to the east so that it can extend the line to fit the tube through which sea water Turn Desert areas a distance of 6 km.

Sidra oil port of a key Libyan oil export ports and the largest, is about 180 km east of Sirte. And is used in the export of Libyan oil through pipelines to transport oil from the south up to its distance is about 1,400 km. Facilities were completed in 1962.

Among its facilities are equipped with four berths for cargo ships, and 19 storage tanks with a capacity of up 6.2 million barrels of crude oil for. *Exposed to Libya since about 400* million years ago to the activities of geology led to the formation of the natural attritions of Libya and the formation of a group basins oil and water in the water of the development the initial system of tectonic processes in the two phases upper Paleozoic and Permian pushed seawater TTS to inundate vast areas of Libya territory to reach the mountains Tibsty then retreat to happen tater in the Pleistocene of the modern era changing climate led to said rain and from as group of units rocky hills and seas of vast sand to create conditions more desert led to the emergence of a group of sand

waha oil company email: Hmbstf@yahoo.co.uk

dunes that make up the vast expanses of Libyan territory which is about

Fig.1 Es.sider Map

1.7000,000 km^2 and believed that the time fourth geological from most of the features of Libyan territory located Libya within the climate semi-dry which is characterized Balk rate of precipitation which contributed to the spread widely to the problem of desertification and the transformation of vast areas of ferule land high production areas to poor flora and fauna causing the poorest to the ecosystem , because the spread of desertification to the climate conditions of long-range of the low rate of precipitation as well as the occurrence of the region in the climate semi-desert to high rate of evaporation and transpiration low soil moisture and high common sulfate and calcium carbonate which led to the spread of marshes salt and therefore to say vegetation in the range western also heavy irrigation has led to the increate of salinity in the soil for the soil to turn saline soil suitable for agriculture but not for the building and also logging and over sponsor of the other problems that have led to increased desertification in the region .

LOCATION GEOGRAPHY

Sidra oil port of a key Libyan oil export ports and the largest, is about 180 km east of Sirte. And is used in the export of Libyan oil through pipelines to transport oil from the south up to its distance is about 1,400 km

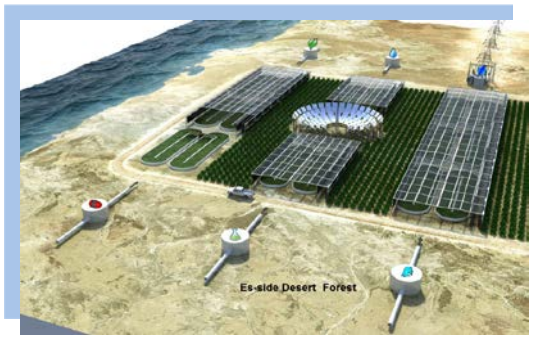


Fig. 2 Sample Es-eide Forest

ENVIRONMENT DEPOSITS

The Libya oil field are located within the arid desert regions lack the necessities of life because of the widespread natural sand dunes I said water summed up the ides of instating the Sahara forest project on an area of about 10 hectares which is a huge for the flow of sea water toward the earth where pipeline it is possible that the laws of nature allow the production of fresh water characterized the technology that will be used in the project in its ability to covert large tracts of desert into green land and the process mimic what in nature evaporates sea water as a result of the sun's heat and work is under way on the cooling leads to the formation of clouds and condense the vapors of water to came k fall water style rainfall only 10 to 15% of the air humidifier condenses into fresh water and the rest outside the green surroundings trees thus the greenhouse creating green surroundings ground and is Neil LEC texts one energy experts there is a huge potential for these new technologies power unit CSP one covered only 1% of space Sahara earth and can provide five global need for energy as well as can invest this technique steaming million tons of sea water per day using 20.000 greenhouse only produce commercial of food crops growth conditions cold and wet vegetable in addition to the production of fresh water. It will be connected to solar with thermal unit for seawater desertification technology backed it will provide experimental station also cultivate algae in the system of bioreactors optical systems and cultivation in ponds exposed exposed produce plant food water and energy university of greenhouse protected cooled salt water and the generation of solar electricity intensive and the cultivation of algae and technological re-greening of the desert what constitutes several precedents that the basic idea of this project is to bring seawater to the desert then vaporized by solar energy and then intensify fresh water not only this but retains environment damp and cold are ideal for the cultivation of crops , it was the first model for the sea water greenhouse a unit with are 2000 square meters in port Augusta in Australia and has won its first crop of tomatoes last month's Sahara forest project combines the technological and solar power plant unlike the PV system which convert sunlight into electricity the CSP uses mirrors to focus the light on the complex to heat which in turn produces steam to run a turbine generator at a facility forest desert the solar power plant supplying gas power and operation and the greenhouse extended solar power plant with water making then more efficient project he is promising and important in the field of renewable energy and to take advantage of the potential of solar power in the Arab region green house and sea water are essential in a project cost will be cheap as cabling the hot desert air and moistened with sea water before entering into greenhouse and his air humidifier feed crops inside the greenhouse and then passes through the dryer where the flow of sea water heated .

Immediately after the heat of the sun meet air humidifier of channels containing sea water cold fresh water condenses and drops dripped on the external aspects of the channels so they

ISBN: 978-1-61804-327-6

can be assembled the process mimic what happens in nature evaporates sea water as a result of heat sun and work is underway on the cooling leads to the formation of clouds and

OBJECTIVE STUDY

This study aims to units of climate change through water vapor to provide by sea water vaporization heats the sunshine so that we have a water vapor could fall all the way large amounts of rain, God willing, and work is underway on cooling, leading to the formation of clouds and condense the vapors water falls to fall water like rain from 10 to 15 per cent of the air humidifier condenses into fresh water and the rest outside the surrounding trees.

Condense the vapors of water to come K fall water like rainfall only to 10 – 15% of the air humidifier condenses into fresh water and the rest outside the surrounding trees, thus the greenhouse creating

Fig. 3 sample forms

process of using energy intensive and through the use of mirrors to focus sunlight on black pipes and boilers for the water district radiation cause intensive production of high heat inside the pipe provides enough energy run Turbine steam which in turn will generate electrical power and six cell additional energy to residents of the area population to have

Fig. 4

benefited says one

specialist and is Neil LEC texts one energy experts there is a huge potential for these new technologies power unit CSP one

Covered only 1% of the are desert in the earth and can provide five global need for energy as well as can invest this technique steaming million tons of sea water per day using 20.000 and only a greenhouse can is know that this project is about the Libya oil field that lack the green land.

CONCLUSION

Environment is a system dynamically complex includes the components and elements of interlocking, multi, where knowledge and information related to this system greatly evolved in recent decades, especially with its association largely being an important dimension of sustainable development dimensions in the community and the fact that the concept of scientific environment he finishes means basic signal to the study of nature... Its elements and organic physical and chemical neighborhood surrounding the object, the ecosystem in nature and accordingly represents the result of the balance between elements with factors and forces that interact with each other to happen balance and end the imbalance in the system Environmental. The idea of the project focused on the possibility of creating a new environmental conditions in the oil field areas through the establishment a new climate regime to increase the proportion of evaporation from solar heating to create an opportunity for rainfall Falls to the primitive desert development phase so that abound trees and farms specialized and natural lakes through climate change dry to wet climate copious rain.

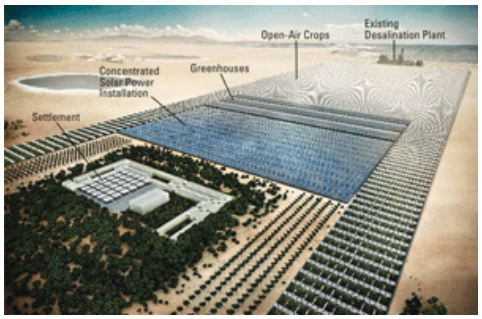


Fig. 5 Lakes

REFERENCES

1-Irrigated agriculture and its impact on the groundwater Jafara basin .Dr. Hassan .m.eljadi

2-ground water between theory and application Dr.Maftih Slaoui

3- Groundwater index earthquake research Dr.Asma Al-keeiabi.

4-Tripoli sheet map, industrial research center

5-Magazine water, General authority for water

Hybrid Model for Early Diabetes Diagnosis

A.A. Ojugo., A.O. Eboka., R.E. Yoro., M.O. Yerokun and F.N. Efozia

***Abstract*—**Diabetes Mellitus (silent killer or sugar disease) is a metabolic disease characterized by high glucose levels, either in a body with insufficient insulin to breakdown glucose, or body that is resistant to effects of insulin. To improve early diagnosis, data-mining tools are used to help physicians effectively classify the disease. Study presents a hybrid fuzzy, genetic algorithm trained neural network model as a decision support system for diabetes classification. Adopted data is split into: training, cross validation and testing to aid model validation with appropriate weights and biases set for each variables. Results indicate that age, obesity and family relations (in first and second degree), environmental conditions are critical factors to be watched; While in gestational diabetes, mothers with or without a previous case of GDM is confirmed if there is: (a) history of babies with weight > 4.5kg at birth, (b) resistant to insulin showing polycystic ovary syndrome, and (c) have abnormal tolerance to insulin.

***Keywords*—Diabetes, Gestational, fuzzy classifiers, linguistics variable, membership function, mutation,**

I. Introduction

DIABETES mellitus is a metabolic disease, characterized by presence of hyperglycemia resulting from defective insulin secretion or its processing in the body. Simply put, it results from bodies that do not have enough insulin to breakdown glucose (byproduct of carbohydrate), or bodies resistant to the effects of insulin. Glucose, as a main source of energy for cells that makes up the muscles and other tissues, is produced from the ***food*** we eat and in our ***liver***. Sugar (or glucose) is absorbed in the bloodstream and enters into a cell by the help of insulin. The liver stores glucose as glycogen so that, if glucose becomes low, the liver reconverts the stored glycogen into glucose to normalize the glucose level (Ojugo et al, 2014). Diabetes is diagnosis from glycemia associated with microvascular disease (Goldenberg and Punthakee, 2013).

Khashei et al (2013) Extreme difficulty in early detection of disease by experienced physicians, has led a continued quest for methods to effectively and precisely classify the disease. Ojugo et al (2014) The models are grouped into ***supervised*** classification models (here, its input variables for diagnosis are ***known***), and ***unsupervised*** classification models (variables for diagnosis are ***unknown***). In both instance, a ***critical*** feat for selecting the appropriate classification model is, its accuracy of classifying the task. Thus, researches continue to improve on the model's efficiency and accuracy.

Diabetes is classified into:

a. **Type-1** (insulin-dependent) is the chronic condition/state, in which the pancreas produces little or no insulin (the hormone required to allow sugar/glucose enter the cells in other to produce energy). With no insulin to release glucose into the cells, sugar builds up in the bloodstream causing life-threatening complications. Despite continued research, type-1 has no cure. Its ***symptoms*** are: increased thirst, extreme hunger, fatigue, frequent urination, blurred vision, bedwetting in children who previously do not bed-wet at night, irritability and mood changes, unintended weight loss, and in females, vaginal yeast infection (American Diabetes Association, 2009).

   Causes of type-1 is ***unknown***; But, in most cases, the body's own immune that normally fights harmful bacteria and viruses, mistakenly destroy insulin-producing (islets) cells in the pancreas. Also, genetics plays a significant role alongside the exposure to environmental factors such as viruses, which may trigger ***type-1*** diabetes (Vaarala et al, 1999). Known risk factors are: genetics, family history, age, geography, exposure to environment (in bacteria and Epstein-Barr virus), early exposure to cow milk, intake of nitrate-contaminated water, low vitamin D, early/late introduction to cereal/gluten into baby's diet, a mother with preeclampsia during pregnancy and if a baby is born with jaundice (American Diabetes Association, 2009).

b. **Type 2** (adult onset or noninsulin-dependent) diabetes is a chronic condition that affects how the body metabolizes sugar (glucose). It often develops slowly since the body either resists the effects of insulin as produced or does not produce enough insulin to maintain a normal glucose level. Though common in adults, this type is increasingly common now to children with obesity issues. While, there is also no cure for type-2, it can however be managed through proper eating habits, exercising, maintaining a healthy weight and sometimes, diabetes medications or insulin therapy. Its symptoms are increased thirst/hunger, weight loss, frequent urination, fatigue, blurred vision, acanthosis nigricians (areas of darkened skin) amongst others (Canadian Diabetes Association, 2014).

Whichever the case, both types of diabetes greatly increases one's risk for a range of complications (for which, monitoring and proper management can help prevent). Diabetes remains the leading cause of many complications such as blindness, kidney failure as well as continues to be a critical risk factor for heart diseases, stroke and amputations (Ojugo et al, 2014).

c. **Gestational** diabetes represents glucose intolerance with the onset or is first recognized during pregnancy. It causes high blood sugar that can affect pregnancy and the baby's health, though the blood sugar usually returns to normalcy soon after delivery. A patient with gestational diabetes is at the risk of type-2 diabetes with each pregnancy and it does not cause any noticeable signs (Canadian Diabetes Association, 2014).

Chinenye and Young (2011) It is fact that type-2 diabetes has asymptomatic preclinical phase which is not benign and thus underscores the need for primary prevention and population screening in order to achieve early diagnosis and consequent treatment. Prevalence of undiagnosed diabetes has been found

A.A. Ojugo: Dept. of Math/Computer Sci., Federal University of Petroleum Resources, Effurun, (ojugo.arnold@fupre.edu.ng)

to range from 4.76% of outpatients attending a family practice clinic, to as high as 18.9% in Nigeria. And such prevalence is higher by 68% in persons of higher socioeconomic status.

The main ***idea*** here is to present a hybrid fuzzy (rule-based), genetic algorithm trained neural network model for diabetes diagnosis. The Soft-computing heuristics aim to provide through its predictions, possible solution to complex-dynamic task for which conventional method do not yield cost-effective, complete results. It analyses data by investigating features of interest through its underlying probabilities to simulate tractability, robustness and low-cost, effective solution with high tolerance to uncertainties, partial truth, ambiguities, imprecision and noise that may have been applied along with its input (data). Soft-computing have been successfully applied to various disciplines, our study adopts a hybrid fuzzy (rule-based), genetic algorithm trained neural network model to aid diabetes diagnosis (as decision support model) for proper treatment.

II. Statement of Problem

The problem statements are as follows:

1. The complex nature of diabetes, its complications and risk factors, makes early and accurate diagnosis ***critical*** as this has deterred many patients from carefully monitoring and management. Supervised diagnosis has become somewhat redundant and time-wasting as it often yields inconclusive results for ***unknown*** inputs – leading to an increased rate of false positive results. The proposed model effectively classifies varying types of diabetes (Section IV).
2. Classification (as resolved in Section IV) using predictive models is quite a complex and difficult task due to the chaotic and dynamic nature of ***unsupervised*** evolutionary models to effectively classify diabetes.
3. Many datasets often consist of ambiguities, imprecision, noise and impartial truth that must be resolved via robust search. Also, speed constraint that often gets such solution trapped at local minima (resolved in Section III).
4. Search for optimal solution, may also yield overtraining and over-fitting of the model (resolved in Section IV) as it aims to find underlying probability of the data feat(s) of interest. Also, improper selection of parameters may also lead to over-parameterization (resolved in Section V).
5. Previous research aim at single heuristic that will globally classify various diabetes-types. Studies show that some case of diabetes result from drug use, related diseases and in some cases, is not even be diabetes – though, with symptoms that are quite similar or mimics any of the various types or diabetes class) as resolved in Section IV.
6. The search for optimal solution via evolutionary heuristic techniques is quite cumbersome. Also, hybrids require resolutions of its many underlying statistical dependencies that exist between the varying searches. The proposed model resolves this in Section III and IV.

As researches continues to advance for early diagnosis of ***type-2*** and gestational diabetes, we propose a hybrid fuzzy (rule-based), genetic algorithm trained neural network model for unsupervised classifications. Prediction is only an improvised search technique by which a model allows the propagation of a set of observed data, as the model seeks data feats of interest (even if such data consists of ambiguities, noise and impartial truth at its input) to yield an output via optimization methods and taxonomy (Ojugo et al, 2013). Such evolutionary models have been successfully applied to enhance accurate prediction that aims at an optimal solution of a task, chosen from a set of possible solution space – to yield an output that is guaranteed of high quality and void of ambiguities. These models, further tuned can become robust and perform quantitative processing to ensure qualitative knowledge and experience, as its new language (Ojugo et al, 2013, Heppner and Grenander, 1990).

III. The Proposed Hybrid Model

Proposed model hinges on 3-basic frameworks: (a) fuzzy (rule-base) logic, (b) genetic algorithm and (c) artificial neural network, as in fig 1 – which is further explained as:

A. Fuzzy Logic (Rule-Based Variables)

A fuzzy system chooses between different control actions and seeks to transform them into a fuzzy set value. The fuzzy system is divided into:

a. The Fuzzy classifier assigns a class label to an object, based on its description, so that it can predict each class label. Object descriptions are vector values with feats and attributes relevant for such classification. It then learns to predict a class label via the training algorithm and its accompanying dataset. If such dataset is unavailable, the classifier is designed to learn *apriori* so that trained, it can effectively classify the objects. The classifier focuses on IF-THEN-rules of actions and possible outcomes that are constructed as a user specifies the class rules and the accompanying linguistic variables (fuzzy set). These help to tune the fuzzy set in line with the class rules (Ludmila, 2008; Chakraborty, 2010). For example:
If Math Error is medium and is small, Then Class 1
If Math Error is medium and is large, Then Class 2etc

b. Fuzzy Cluster algorithm groups the data points (linguistic variables) into homogeneous classes known as clusters so that items in the same class are as similar as possible and vice-versa. Clustering is a compression method, where a large number of samples are converted into small number of representative clusters (Giles and Draeseke, 2001). Depending on data and task, different types of similarity measures help to identify classes and control how clusters are formed. Values, used as similarity measures include distance, connectivity, and intensity (Chakraborty, 2010).

In some cases called non-fuzzy (hard clusters), data is divided into crisp clusters such that each data point belongs to exactly one cluster; While, in fuzzy clusters, it is quite possible for the data points to belong to more than one cluster, and associated with each of the points are membership grades which indicate the degree to which the data points belong to these different clusters (Nascimento, 1991). Thus, we employ a Fuzzy Cluster Means (FCM) algorithm that partitions a finite collection of elements into a collection of fuzzy clusters with respect to some criteria. FCM linguistic description can be implemented by fuzzy logic (Berks et al, 2000) as thus:

a. Select the number of clusters c ($2 \leq c \leq n$), exponential weight μ ($1 < \mu < \infty$), initial partition matrix U^0, and the termination criterion ε. Also, set the iteration index 1 to 0.

b. Calculate fuzzy cluster centers $\{V_i^1 | i=1, 2,..., c\}$ using U^1.
c. Calculate new partition matrix U^{1+1}via $\{V_i^1 | i=1, 2,..., c\}$.
d. Compute new partition matrix = $\| U^{1+1} - U^1 \| = | U_{ij}^{1+1} - U_{ij}^1 |$. If > ε, set l = l + 1. Go to step (b); Else stop if ≤ ε.

Initial cluster centers are computed via (a) arithmetic means of all data points, or (b) running FCM several times, each starting with a different initial cluster centers. For this study, we adopt the first method. To implement the fuzzy logic system, we perform the following:

a. Define control goals/criteria: What do I wish to control? How do I achieve the control? What response is needed? What are the possible (probable) system failure modes?
b. Define input/output relationships and choose minimum number of input variables to the system as well as define the error rate and rate-of-change-of-error.
c. Using the rule-structure, define the control problem into a series of IF-X-AND-Y-THEN-Z rules that define desired model output response for given input case. The number and complexity of rules depends on the number of input parameters to be processed and the number of variables associated with each parameter. If possible, use atleast a variable and its time derivative (though, possible to use an instantaneous error parameter without knowing its rate of change, which can cripples model's ability to minimize overshoot for a step inputs).
d. Create Fuzzy Logic membership functions that define the meaning (values) of Input/output terms used in the rules.
e. Create necessary pre- and post-process Fuzzy set routines if implementing as software; Else, hardwire the rules into the Fuzzy Logic hardware engine.
f. Test/retest the system, evaluate results and tune rules and membership functions, until it yields satisfactory result(s).

Linguistic variables are non-precise variables used to convey a surprising amount of data about our environment or an object under observation. In common usage, linguistic variables often overlap. Linguistic variables need a formal way of describing a linguistic variable in crisp terms that the computer can deal with. To indicate the relationship between measured distance and linguistic term far. Each individual may have slightly differing ideas about the exact distance measurement that *far* actually represents – though, the said distance be consistent (Berks et al 2000). At some point, some individuals agree that *it is not far* and at some point, they also all do agree that *it is far*. The space between *far* and *not far* is the distance measure that is to some degree, a bit of both. Thus, the horizontal axis of such graph shows a *crisp value* of the distance; while, its vertical axis describes degree to which the linguistic variable fits with the crisp measured data (Inan and Elif, 2005).

B. Genetic Algorithm (GA)

GA is inspired by Darwinian genetic evolution (survival of fittest) consists of population (data) chosen for selection with potential solutions to a specific task. Each potential solution is an individual for which optimal is found using four operators: initialize, select, crossover and mutation (Coello et al, 2004 and Reynolds, 1994). Individuals with genes close to optimal are fit. Its fitness function determines how close an individual is to the optimal solution. Its operators are (Ojugo et al, 2012):

a. Initialize encodes data into forms suitable for selection. Each encodings type used has its merit. Binary encoding is computationally more expensive. Decimal encoding has greater diversity in chromosome and greater variance of pools generated. Float-point encoding or its combination is more efficient. The *fitness* function evaluates how close a solution is to its optimal – after which they are chosen for reproduction. If solution is found, function is *good*; else, *bad* and deselected for crossover. Fitness function is the only part with knowledge of the task, and if more solutions are found, the higher its fitness value.
b. Selection: With best fit data chosen to mate, the larger the number selected, the better the chances of yielding fitter data. This continues until one is chosen, from the last 2 or 3 remaining solutions, to become the selected parents of prospective new offspring. Selection ensures fittest data is chosen to mate. The selection that only mates the fittest is an *elitist* and often leads to converging at local optima.
c. Crossover ensures best fit data (genes) are exchanged to yield a new and fitter pool. There are two crossover types (depends on encoding type used): (a) ***simple*** crossover for binary encoded pool, which allows single- or multi-point cross with all genes from single parent, and (b) ***arithmetic*** crossover allows new pool to be created by adding a data or individual's percentage to another.
d. Mutation alters chromosomes by changing its genes or its sequence, to ensure new pool converges to global minima (instead of local optima). Algorithm stops if optimal is found, or after number of runs if new pools are created (though computationally expensive), or when no better solution is found. Genes may change based on probability of mutation rate. Mutation improves the much needed diversity in reproduction and its algorithm is as thus:

Cultural GA (a variant) has belief space as thus: (a) Normative (has specific value ranges to which data is bound), (b) Domain (has data about task domain), (c) Temporal (has data about events' space is available), and (d) Spatial (has topographical data). In addition, an influence function mediates between its belief spaces and the pool – to ensure altered data conforms to the belief space. Thus, the data pool does not violate its belief space. This helps reduce number of possible individuals GA generates till an optimum is found (Reynolds, 2004).

C. Artificial Neural Network (ANN)

ANN as data processing model is inspired by neurons in the human brain, and consists of interconnected neurons (nodes) with capability to learn by example that makes them universal estimators (Ojugo et al, 2012). To processes data that allows for learning to take place, the nodes share signals by adjusting its weights and biases representing its connection strengths between synapses, axons and dendrites respectively. Signals are summed and depending on the task, its activation function limits its output (Ojugo et al, 2014) to modulate associated inputs and nonlinear feats exhibited as in Eq. 1 below:

$$\emptyset = f(net) = f \sum_{i=1}^{m} Xi * Wij \qquad (1)$$

ANN are trials, in an attempt to translate into mathematical models, principles of biological processing so as to generate in

the fastest time, implicit and predictive evolution outcomes of a task. It derives its possible outcomes from experience and can recognize behaviour(s) or feat(s) of interest from historic dataset, in order to suggest an optimal solution of high quality, irrespective of modification made to its by other methods in a multi-agent space, which constantly affects the quality of any solution (Dawson and Wilby 2001). ANN is encoded as a 3-layered input, hidden and output unit; And, is configured as either: (a) ***feedforward*** (data flows from input to output, and extends over multi-layers), and (b) ***recurrent*** (has feedback with dynamic feats to evolve the net as it undergoes relaxation to a stable state where its activation values and output changes no more). In some task, its output change is significant and the dynamic behavior of the net constitutes its output (Ojugo et al, 2013). Its configuration depends on task, dataset feat(s) and connection requirement. It learns ***explicit*** (apriori knowledge) and/or implicit (post-priori knowledge) to allow trained net to effectively classify patterns that change its bias and weight based on a rule (Ojugo et al, 2013). Learning is grouped into:

a. ***Supervised*** in which an input vector has a set of desired responses, one for each node as relayed to the output. A forward pass measures error between ***desired*** and ***actual*** response for each node in the output, which is used to determine weight changes in the net based on the learning algorithm. Desired output is provided by external teacher via back-propagation, delta rule and perceptron rule
b. ***Unsupervised*** or self-organized allows its output trained to respond to a clusters of patterns at its input so that the model discovers statistically, salient feats within the input dataset – such that the model has no prior knowledge how patterns are grouped. Rather, the model develops its own representation of the input dataset.
c. ***Reinforcement*** in which network learns what to do, map states to actions to maximize a numerical reward data. Network then discovers actions that yield most reward by trying them. Also, such actions affect not only immediate data, but the rest states (Caudill, 1987).

The nature of diabetes diagnosis requires previous knowledge. Thus, we adopt the recurrent Jordan net to help us incorporate historic dataset feats of interest and previous output to be fed back as input into the model's hidden units to yield the next output. Its correlated weights are interconnected so that $W_{i,j}$ is weight between input and hidden layers, $W_o j$ is bias and x_i is diabetes input data. Its output is generated via tangent/sigmoid transfer function, which sums weighted input as in Eq. 2 and Eq. 3 (Minns, 1998). To resolve statistical dependencies conflicts in model's structure imposed by dataset and methods adopted, our network used its ability to store earlier data generated from previous layer(s) as in Kuan and White (1994)

$$Z_{ij} = w_o j + \sum_{i=1}^{m} x_i * w_i j \quad (2)$$

$$F(Z_{ij}) = \frac{2}{1 + e^{-2*Zij}} - 1 \quad (3)$$

The Jordan's net is more plausible and computationally more powerful than other adaptive models. Its back-propagation in time algorithm allows for advanced training/learning so that its output at time *t* is used along with a new input to compute the net's output at time *t+1* in response to dynamism (Mandic and Chambers, 2001). Thus, output is computed via Tansig function given as y^k, which sums input, receives target value of input training pattern, and computes error data, weight correction updates in its layers c_j^k and bias weights correction updates c_o^k. This error is sent from its output back to input nodes via error back-propagation to correct weights. Thus, it finds weights that approximate the target output with selected accuracy. Weights are modified by minimizing error between target and computed outputs at the end of each forward pass. If error is higher than selected value, process continues with reverse pass; else, stop training. Weights are updated via MSE till minimal error is found (Ursem et al, 2002; Guo and Xue, 2011).

Our Jordan recurrent net is constructed by modifying the multilayered feedforward with addition a ***context*** layer to help retain data between observations. At each move, new inputs are fed to the net. Previous contents of hidden layer is passed into context layer and later fed back into the hidden layer in the next time step. The context layer contains nothing initially. Output from the hidden layer after the first input will be same as if there is no context layer (Ojugo et al, 2013). Weights are computed same way for new connections fro/to the context layer from the hidden layer. Training aims at the best fit data weights computed via Tansig that assumes approximation influence of data points at the center so that function decreases with distance from its center (Perez and Marwala, 2011). Its Euclidean length (r_j) yields distance between $y = (y_1,...,y_m)$ vector and its center ($w_{1j},...,w_{mj}$) via Eq. 4:

$$r_j = \left|\left| y - Y^j \right|\right| = \left\{ \sum_{i=1}^{m} (y_i - w_i j)^2 \right\} ½ \quad (4)$$

The suitable transfer function is applied to r_j:

$$\emptyset(r_j) = \emptyset || y - Y^j || \quad (5)$$

Finally, output k receives weighted combination as:

$$y^k = w_o + \sum_{j=1}^{n} \left(c_j^k * \emptyset(r_j) \right) = w_o + \sum_{j=1}^{n} \left(c_j^k * \emptyset || y - Y^j || \right) \quad (6)$$

IV. Experimental Framework: Materials/Methods

Table 1: Fuzzy Encoded Universe Discourse Dataset Values

Code	Fuzzy Set (Parameters)	Membership Function Degree for Diabetes		
		Type-1	Type-2	Gest.
P01	Frequent Urination	0.50	0.00	0.00
P02	Increased and Unusual Thirst	0.50	0.00	0.50
P03	Extreme Hunger	0.50	0.00	0.00
P04	Unusual Weight Loss	0.50	0.50	0.00
P05	Extreme Fatigue	0.50	0.00	0.50
P06	Serious Irritation	0.00	0.00	0.50
P07	Frequent Infection	0.00	0.00	0.50
P08	Blurred Vision	0.00	0.50	0.00
P09	Slow Healing of bruises/cuts	0.00	0.50	0.00
P10	Tingle/numbness in hands/feet	0.00	0.50	0.00
P11	Regular skin/bladder infection	0.00	0.50	0.50
P12	Nausea/vomiting	0.00	0.00	0.50
P13	Haemoglobin test > 10	0.20	0.20	0.20
P14	Leg cramp	0.20	0.20	0.20

The retrieved dataset is presented in table 1 below, obtained via research survey, utilizing questionnaires as the tool. A total of hundred questionnaires were distributed to various medical (diabetic) experts spread across fifteen hospitals in Five Geo-

political regions in Nigeria and cross-referenced by the study of Edo et al (2015). The sixth geo-political zone was not left out due to insurgency. To generate a fuzzy Linguistic variable Universe of Discourse, all questionnaires responses were tuned via proposed equation:

$$PFCMUDE = \sum (A, B, C, D, E) * X \qquad (7)$$

A,B,C,D,E = picked questions; X(0.00) = unpicked option and X(0.02) = Assigned questions option fuzzy range value

A. Model Design Objectives

The proposed system will resolve existing problems via:

a. Perform repetitive tasks without emotional defects
b. Embody the knowledge of human experts with the help of special software tools, manipulate data to solve problems and make decisions in that domain.
c. Processes are better formalized and defined on machines.
d. Automatic updating of the knowledgebase.
e. Processes are better formalized and defined on machines.

B. Proposed Model Framework

From Fig 1, the proposed model design employs these:

a. Fuzzy system consists of a *classifier* which propagates if-then rule values of selected data, enhanced as predefined linguistic variables classification into the diabetes classes, and the *fuzzy-cluster-means-universe-discourse-equation*, which enhances the linguistic variables partitioning it into data-point that cumulate into the universal of discourse generated via a survey exercise utilizing questionnaires.
b. Jordan network provides a self-learning ability, optimized by the CGA optimizer that recombines and mutates the rule-based fuzzy dataset to train and test the system so that it autonomously classify diabetes into its class types.
c. Genetic algorithm helps train ANN so as to optimize our collated-answers within the tuned fuzzy dataset linguistic variable (symptoms) universe of discourse values in other to yield a centralized, fuzzy-scaled function boundary in determining high/low degree membership function.

The proposed model consists of these parts:

a. Knowledgebase– consists of historic structured data feats and the database of diabetes symptoms, the fuzzy if-then rules, its global universe discourse linguistic variables and optimized membership functions. Its houses the optimized universe discourse values as represented by fuzzy-if-then, linguistic variables (rule-based) as selected data feats.
b. Inference engine– consists of hybrid fuzzy logic (rule-based) genetic algorithm trained neural network model. Inference engine infers conclusion derived from genetic algorithm trained neural network from the selected data feats encoded as fuzzy-if-then conditions with possible outcomes and consequent action upon criteria being met.
c. Decision support– consists of the predicted output and the output database that is updated automatically in time as patients are diagnoses as long as it encounters and read sin new data. The decision support predicts system output based on the cognitive and the emotional filers as display by the output device. This is seen in fig 1.

Model starts off with Proposed Fuzzy Classifier and Cluster Mean Universe Discourse Equation (PFCMUDE) in which the answers from the collected questionnaire are expressed thus:

C. Genetic Algorithm Trained Neural Network (GANN)

GANN is initialized with the fuzzy (if-then rules) linguistic variables. Individual fitness is computed as 30-individual are selected as new pool via *tournament* method. It determines mating individuals and solutions. Crossover and mutation is applied to help network learn dynamic and non-linear feats in the dataset and feats of interest using a multi-point crossover. With mutation, data between 1 and 30 is randomly generated using Gaussian distribution corresponding to crossover points (all genes are from single parent). As new parents contribute the rest to yield new individuals whose genetic makeup is combination of both parents, mutation is also applied to yield 3-random genes that also undergoes another mutation as they are allocated new random values that still conforms to belief space. Number of mutation applied depends on how far CGA is progressed on the network (how fit is the fittest individual in the pool), which equals fitness of the fittest individual divided by 2. New individuals replace old with low fitness so as to create a new pool. Process continues until individual with a fitness value of 0 is found – indicating that the solution has been reached (Ojugo et al, 2013).

Initialization/selection via ANN ensures that first 3-beliefs are met; mutation ensures fourth belief is met. Its influence function influences how many mutations take place, and the knowledge of solution (how close its solution is) has direct impact on how algorithm is processed. Algorithm stops when best individual has fitness of 0 (Dawson and Wilby 2001). GANN model is as thus:

INPUT:
1. Poolsize (k), crossover (c), mutation (v), influence function (Ifnc) and n;
   // *Initialization and Selection*
2. Randomly generate *K* possible solution
3. Save solution in population *Kok;*
   // Loop till terminal point
4. For m = 1 to n do;
   // *Crossover*
5. Number of crossover nc = (k –Ifnc)/2;
6. For u = 1 to n do;
7. Select two solutions randomly E_A and F_G for *K;*
8. Generate G_V and H_N by 2-point crossover to E_A and $F_{G;}$
9. Save G_V and H_N to *K2;*
10. End For;
    //*Mutation*
11. For u = 1 to n do;
12. Selection a solution Y_h from *K2*;
13. Mutate each bit of Y_h under Ifnc
14. Generate a new solution Y_h^i
15. If Y_h^i is impossible
16. Re-compute Y_h^i with possible solution by modifying Y_h^i
17. End if
18. Re-compute Y_h with Y_h^i in *K2*
19. End for
    //*Re-compute*
20. Re-compute *K = K2;*
21. Return ***Best*** solution in Y

Model stops if stop criterion is met. GANN utilizes number of epochs to determine stop criterion. Initial selection is given as:

1. **R1:** If R01 Then C1 = 0.50
2. **R2:** If R01 AND R02 Then C1 = 0.50
3. **R3:** If R01, R02 AND R03 Then C1 = 0.50
4. **R4:** If R01, R02, R03 AND R04 Then C2 = 0.13

5. **R5:** If R01, R02, R03, R04 AND R05 Then C3 = 0.17
6. **R6:**If R01, R02, R03, R04, R05 AND R06 Then C3 = 0.03
7. **R7:** If R01, R02, R03, R04, R05, R06 AND RO7 Then C3=0.03
8. **R8:** If R01, R02, R03, R04, R05, R06, RO7 AND R08 Then C3 = 0.17
9. **R9:** If R01, R02, R03, R04, R05, R06, RO7, R08 AND R09 Then C3 = 0.17
10. **R10:** If R01, R02, R03, R04, R05, R06, RO7, R08, R09 AND R10 Then C3 = 0.17
11. **R11:** If R01, R02, R03, R04, R05, R06, RO7, R08, R09, R10 AND R11 Then C3 = 0.17
12. **R12:** If R01, R02, R03, R04, R05, R06, RO7, R08, R09, R10, R11 AND R12 Then C3 = 0.17
13. **R13:** If R01, R02, R03, R04, R05, R06, RO7, R08, R09, R10, R11, R12 AND R13 Then C3 = 0.20
14. **R14:** If R01, R02, R03, R04, R05, R06, RO7, R08, R09, R10, R11, R12, R13 AND R14 Then C3 = 0.20

Fitness function (*f*) is resolved with initial pool (Parents) as:

R1=50	R2=50	R3=50	R4=13	R5=17	R6=3	R7=3
R8=17	R9=17	R10=17	R11=17	R12=17	R13=20	R14=20

Table 2: 1st and 2nd Generation of population from Parents

S/N	Selection	Chromosomes (Binary 0 or 1)			Fitness Function
		Parent 1st Gen	Crossover	Parent 2nd Gen	
1	50	110010	1 and 9	1100**00**	48
2	50	110010	2 and 14	1100**11**	51
3	50	110010	3 and 13	1100**00**	48
4	13	001101	4 and 12	0011**01**	13
5	17	010001	5 and 6	0100**11**	19
6	3	000011	5 and 6	0000**01**	1
7	3	000011	7 and 8	0000**01**	1
8	17	010001	7 and 8	0100**11**	19
9	17	010001	mutation	0100**10**	18
10	17	010001	mutation	0100**10**	18
11	17	010001	mutation	0100**10**	18
12	17	010001	4 and 12	0100**01**	17
13	20	010100	3 and 13	0101**10**	22
14	20	010100	1 and 14	0101**10**	22

The Proposed Fuzzy Classifier Diabetes training Algorithm:

INPUT:

1. Diabetes Types (Type1,Type2,Gestational,MODY,LADA)
   No. of Symptoms (P1, P2,…,Pn)
   P = Fuzzy parameters (Symptoms Codes)
   N = 15
   More than five symptoms = Serious
   Exactly four symptoms = Moderate
   Three symptoms and below = Minor
2. *Glucose Level* (125md/dl) = High
3. *Age Range (R)*
   1 – 21yrs of age = teenager; 30 – 40yrs of age = pre-Adult;
   > 41yrs of age = post-Adult; > 50 = pre-menopause
4. *Origin (descent)*
   Caucasians: Americans, Europeans, Asians, North-Africa
   Blacks: African, African-Americans, Blacks Indians etc
   Plus; either Caucasians or blacks
5. // INITIALIZATION
6. Randomly pick a patient *K;*
7. Save identification (diagnosis) Result in *Knot;*
8. // Loop till terminal point
9. For P = 1 to n do;
10. // Type 1 diabetes
11. Diagnose for Type 1 Diabetes;
12. If Type I symptoms is *serious*, age is *teenager*, glucose level is *high*, patient origin is *Plus* and pancreas destruction is *swift* THEN Type 1;
13. Else: May be Type 1;
14. Else: May Not Type 1;
15. // Type 2 diabetes
16. Diagnose for Type 2 Diabetes;
17. If TYPE 2 symptoms is *serious*, patient age is *post-Adult*, glucose level is *high* and patient origin is *black* THEN Type 2;
18. Else: Might be Type 2;
19. Else: May Not Type 2;
20. //Gestational diabetes
21. Diagnose for Gestational Diabetes;
22. If Gestational symptoms is *serious*, patient age is *pre-menopause*, glucose level is *high*, patient origin is *plus* and patient is *pregnant* THEN Gestational diabetes;
23. Else: May be Gestational Diabetes;
24. Else: May Not Gestational Diabetes;
25. // MODY diabetes
26. Diagnose for MODY Diabetes;
27. If MODY symptoms is *serious*, patient age is *teenager*, glucose level is *high*, patient origin is *plus* and parents (siblings) diagnosed with *MODY* THEN MODY diabetes;
28. Else: May be MODY diabetes;
29. Else: May Not MODY diabetes;
30. // LADA diabetes
31. Diagnose for LADA Diabetes;
32. If LADA symptoms is *serious;* patient age is *pre-Adult*, glucose level is *high*, patient origin is *Caucasians* and pancreas destruction is *progressive* THEN LADA diabetes;
33. Else: May be LADA diabetes;
34. Else: May Not LADA diabetes;
35. //Save results in *Knot;*
36. Return diabetes result for patient K

Table 3: 2nd and 3rd Generation of population from Parents

SN	Selection	Chromosomes (Binary 0 or 1)			Fitness Function
		Parent 2nd Gen	Crossover	Parent 3rd Gen	
1	48	110000	1 and 2	110011	51
2	51	110011	1 and 2	110000	48
3	48	110000	3 and 13	110010	50
4	13	001101	4 and 12	001101	13
5	19	010011	5 and 9	010000	16
6	1	000001	6 and 8	000011	3
7	1	000001	mutation	000010	2
8	19	010011	6 and 8	010001	17
9	16	010000	5 and 9	010011	19
10	16	010000	10 and 15	010010	18
11	16	010000	mutation	010010	18
12	17	010001	4 and 12	010001	17
13	22	010110	3 and 13	010100	20
14	22	010110	10 and 14	010100	20

Table 4: 3rd and 4th Generation of population from Parents

	Selection	Chromosomes (Binary 0 or 1)			Fitness Function
		Parent 3rd Gen	Crossover	Parent 4th Gen	
1	51	110011	1 and 2	1100**00**	48
2	48	110000	1 and 2	1100**11**	51
3	50	110010	3 and 13	1100**00**	48
4	13	001101	4 and 12	0011**01**	13
5	16	010000	5 and 9	0100**11**	19
6	3	000011	6 and 8	0000**01**	1
7	2	000010	Mutation	0000**01**	1
8	17	010001	6 and 8	0100**11**	19
9	19	010011	5 and 9	0100**00**	16
10	18	010010	10 and 14	0100**00**	16
11	18	010010	Mutation	0100**01**	17
12	17	010001	4 and 12	0100**01**	17
13	20	010100	3 and 13	0101**10**	22
14	20	010100	10 and 14	0101**10**	22

Tables 2, 3 and 4 are generation of optimized fuzzy linguistic variables (symptoms) by single parent (using first and second bits from left as our crossover and mutation points). Fourth generation is stop criterion with best fitness of 51 (row 2). It implies clusters of various universe of discourse variable set

for the fuzzy set searched, are optimized to 0.51 (with a base value of 0.50). Its parameter combination yields a membership function **< 0.50 = Low Degree Membership Function**; while those **≥ 0.50 = High Degree Membership Function**. Further optimization with 0.50 fuzzy-scaled value yields table 5:

Table 5: Optimized Dataset for Diabetes (Scale 0.00- 1.00)

Code	Fuzzy Set (Parameters)	Membership Function Degree Diabetes				
		T-1	T-2	Gest.	MODY	LADA
P01	Frequent Urine	0.50	0.00	0.00	0.00	0.50
P02	Unusual Thirst	0.50	0.00	0.00	0.00	0.50
P03	Extreme Hunger	0.50	0.00	0.00	0.00	0.50
P04	Weight Loss	0.50	0.00	0.00	0.00	0.50
P05	Extreme Fatigue	0.50	0.00	0.00	0.00	0.50
P06	Irritability	0.00	0.00	0.50	0.50	0.00
P07	Freq. Infection	0.00	0.00	0.50	0.50	0.00
P08	Blurred Vision	0.00	0.50	0.00	0.50	0.00
P09	Slow Healing	0.00	0.50	0.50	0.00	0.00
P10	Tingle/numbnes	0.00	0.50	0.00	0.50	0.00
P11	Bladder infectin	0.00	0.50	0.50	0.00	0.00
P12	Nausea/vomit	0.00	0.50	0.50	0.00	0.00
P13	Haemoglobin test > 10	0.20	0.20	0.20	0.20	0.20
P14	Leg cramp	0.20	0.20	0.20	0.20	0.20

Table 6**:** Fuzzy Classifier Diabetes Training Algorithm Class

Code	Fuzzy Set (Parameters)	Membership Funct. Degree Diabetes				
		T-1	T-2	Gest.	MODY	LADA
P01	Freq. Urination	S	M	M	M	S
P02	Unusual Thirst	S	M	M	M	S
P03	Extreme Hunger	S	M	M	M	S
P04	Weight Loss	S	M	M	M	S
P05	Extreme Fatigue	S	M	M	M	S
P06	Irritability	M	M	S	S	M
P07	Freq. Infection	M	M	S	S	M
P08	Blurred Vision	M	S	M	S	M
P09	Slow Healing Cut	M	S	S	M	M
P10	Tingle/numbness	M	S	M	S	M
P11	Bladder infect	M	S	S	M	M
P12	Nausea/vomit	M	S	S	M	M
P13	Haemglobin > 10	M	M	M	M	M
P14	Leg cramp	M	M	M	S	M

As below, S = serious and M = minor case of diabetes, T-1 = Type 1 diabetes, T-II = type 2 diabetes, Gest. = Gestational.

D. Result Findings and Discussion

Table 6 assigns linguistic labels to class based on the fuzzy classifier diabetes training algorithm; While, table 5 highlights the 5-types of diabetes as thus:

a. **Type I** – is any patient in age range 5-15yrs that exhibits symptoms (P01-P05), glucose ≥ 125mg (7.0mmol/L) with swift and rapid destruction of the pancreas.
b. **Type II –** patients above 40years with symptoms (P08 - P12) and glucose level of more than 125mg (7.0mmol/L).
c. **Gest.:** Mothers having symptoms (P06,P07,P09,P11,P12), glucose level ≥ 125mg during/after gestation.
d. **MODY:** patients in age range (1 - 21yrs) with symptoms (P06-P08,P10-P14) from birth and who siblings/parents have MODY and/or Type-2.
e. **LADA –** patient above 35yrs with symptoms (P01 - P05), a glucose level of more than 125mg (7.0mmol/L) with a slow, gradual and progressive destruction of the pancreas.

Each symptom falls into P_1 _ P_{14} with five clusters: Type-I, Type-II, Gestational, MODY and LADA diabetes – makes up the degree of membership/intensity. Example, P13 in cluster 1, 2, 3, 4 and 5 respectively has 0.20 value, which implies that the degree of symptoms of P13 matches **20% of Type I, 20% of Type II diabetes, 20% of Gestational diabetes, 20% in MODY and 20% in LADA.**

The fuzzy partition for each input feat consists of clinical symptoms of the varied forms of diabetes (frequent urination, unusual thirst, extreme hunger, unusual weight loss, extreme fatigue, irritability, frequent infection, blurred vision, slow to heal cuts/bruises, tingling/numbness in hands/feet, regular skin/bladder/gum infection, nausea/vomiting, hemoglobin A1c test (HbA1c) >10 and leg cramp). However, it can occur that a fuzzy partition for a diabetes type is not correctly set up, or the number of linguistic terms for the input feats is not large enough – to result in some patterns being misclassified.

a. Not exhibiting any class diabetes (C_1)
b. Might be exhibiting a class of diabetes (C_2)
c. Exhibiting a class of diabetes (C_3)

If a patient exhibits at least three or less of symptoms of a class of diabetes THEN (C_1), if patient exhibits exactly four of the symptoms of a class of diabetes THEN (C_2) and if the patients is exhibits five or more of the symptoms of a class of diabetes THEN (C_3). IF-THEN rules generated from the fuzzy partitions of classification of varied for diabetes is thus:

1. **R1**: IF patient exhibit frequent urination and its serious THEN class C_1.
2. **R2**: IF patient exhibit frequent urination and unusual thirst and both symptoms are serious THEN class C_1.
3. **R3**: IF patient exhibit frequent urination, unusual thirst and severe hunger and symptoms serious THEN class C_1.
4. **R4**: IF patient exhibit frequent urination, unusual thirst, extreme hunger and unusual weight loss and symptoms are serious THEN C_2.
5. **R5**: IF patient exhibit frequent urination, unusual thirst, extreme hunger, unusual weight loss and extreme fatigue and the symptoms are serious THEN class C_3.
6. **R6**: IF patient exhibit frequent urination, unusual thirst, extreme hunger, unusual weight loss, extreme fatigue and irritability and the symptoms are serious THEN class C_3.
7. **R7**: IF patient exhibits frequent urination, unusual thirst, extreme hunger, unusual weight loss, extreme fatigue, irritability and frequent infection and these symptoms are serious THEN class C_3.
8. **R8**: IF patient exhibits frequent urination, unusual thirst, extreme hunger, unusual weight loss, extreme fatigue, irritability, frequent infection and blurred vision and these symptoms are serious THEN class C_3.
9. **R9**: IF patient exhibit frequent urination, unusual thirst, extreme hunger, unusual weight loss, extreme fatigue, irritability, frequent infection, blurred vision and slow to heal cuts and these symptoms is serious THEN class C_3.
10. **R10**: IF patient exhibit frequent urination, unusual thirst, extreme hunger, unusual weight loss, extreme fatigue, irritability, frequent infection, blurred vision, slow to heal cuts/bruises and tingling/numbness in hands/feet and these symptoms are serious THEN class C_3.

ISBN: 978-1-61804-327-6

11. **R11**: IF patient exhibit frequent urination, unusual thirst, extreme hunger, unusual weight loss, extreme fatigue, irritability, frequent infection, blurred vision, slow to heal cuts/bruises, tingling/numbness in hands/feet and regular skin/bladder/gum infection and these symptoms are serious THEN class C_3.
12. **R12**: IF patient exhibit frequent urination, unusual thirst, extreme hunger, unusual weight loss, extreme fatigue, irritability, frequent infection, blurred vision, slow to heal cuts/bruises, tingling/numbness in hands/feet, regular skin/bladder/gum infection and nausea/vomiting, and these symptoms are serious THEN class C_3.
13. **R13:** IF patient exhibit frequent urination, unusual thirst, extreme hunger, unusual weight loss, extreme fatigue, irritability, frequent infection, blurred vision, slow to heal cuts/bruises, tingling/numbness in hands/feet, regular skin/bladder/gum infection, nausea/vomiting, and hemoglobin >10 and symptom is serious THEN class C_3.
14. **R14:** IF patient exhibits frequent urination, unusual thirst, extreme hunger, unusual weight loss, fatigue, irritability, infection, blurred vision, slow healing, tingling/numbness, skin/bladder infection, nausea/vomiting, hemoglobin >10 and Leg Cramp symptoms are serious THEN class C3.

E. Model Performance

Ojugo et al (2013) Performance is evaluated via computed values: mean square error, mean absolute error, mean relative error and coefficient efficiency as thus:

Table 1. Model Convergence Performance Evaluation

Model	*MSE*	*MRE*	*MAE*	*COE*	*Classification*
ANN	0.87	0.79	0.75	0.781	0.82
CGANN	0.76	0.81	0.62	0.753	0.90
FLCGANN	0.76	0.77	0.76	0.688	0.96

After training, retraining and testing the model, the results indicate that: (a) GA took 21seconds to find the solution after 98-iterations (at best). CGANN was run 15 times (to eradicate non-biasness), it found optima each time – and the time varied significantly between 0.89econds and 21seconds; And, its convergence time depends on how close the initial population is to the solution as well as on the mutation applied to the individuals in the pool. The solution was made even closer using the fuzzy variable dataset (as a preprocessor).

V. Related Study and Result Findings

Khasei et al (2012) adopted a feed-forward multi-perceptron network in their study. Such networks must be expanded and extended to represent complex dynamic patterns and/or cases such as this, since it treats all data as new – so that previous data signals do not help to identify data feats of interest, even if such observed datasets exhibits temporal dependence. Consequently, this has practical implementation difficulty as large nets are not easily implemented. However, the ***Jordan*** network overcomes such difficulty through the use of its internal feedbacks that also makes it appropriately suitable for such dynamic, non-linear and complex tasks as its output unit is fed-back as input into its hidden unit with a time delay, so that its outputs at time $t-1$, is also input at time t.

Barakat et al (2010) adopts SVM by using an additional intelligent module to transform black-box SVM model to an intelligent SVM's diagnostic model with adaptive results that provides a potential model for diabetes prediction. Its logical rule set generated had prediction accuracy of 94%, sensitivity of 93%, and specificity of 94%. Extracted rules are medically sound and agree with outcome of relevant medical studies.

The rationale for the choice of machine learning technique adopted is based on Peter (2014). He compared convergence behavior and other statistic of machine learning techniques on the task classification in which (for instance), decision trees approach takes 90iterations for convergence; neural network approach takes 70iterations whereas the clustering approaches takes 40iterations for convergence, hybrid approaches takes 30 iterations to convergence. Thus the hybrid approaches are very significant when compared to other optimization approaches.

We ***note***, model's speed is traded-off for greater accuracy of classification, more number of rule set generated to update the knowledge database for optimality and greater functionality.

VI. Conclusion and Recommendations

Hybrids are difficult to implement and its accompanying data must be appropriately encoded so that model can exploit numeric data and efficiently explore the domain space to yield an optimal solution. Modelers must seek proper parameter selection and adjustment of weights/biases so as to avoid ***over-fitting***, ***over-training*** and ***over-parameterization*** of the model. Encoded through the model's structured learning, this will help address issues of statistical dependencies between the various heuristics used, highlight implications of such a multi-agent populated model as well as resolve conflicts in data feats of interest. Thus, as agents create/enforce their own behavioral rules on the dataset, hybridization should be able to curb this (as CGA does for this model in its belief space and operators as applied) to display the underlying probabilities of interest.

Models serve as educational tools to compile knowledge about a task, as new language to convey ideas as experts gain better insight to investigate input parameter(s) crucial to a task (Perez and Marwala, 2011), and its sensitivity analysis helps to reflect on theories of systems functioning. Simple model may not yield enough data; while complex model may not be fully understood. Detailed model helps us develop reasonably-applicable models even when not operationally applicable in a larger scale Their implementation should seek its feedback as more critical rather than seeking an accurate agreement with historic data. Since, a balance in the model's complexity will help its being understood and its manageability, so that the model can be fully explored as seen here (Ojugo et al, 2012).

References

[1] American Diabetes Association. *Standards of Medical Care in Diabetes – 2009*. Diabetes Care, 32: S13-61.

[2] Barakat, N.H., Bradley, A.P., Barakat, M.N.H., (2010). *Intelligible support vector machines for diagnosis of diabetes mellitus*, IEEE Transactions on Information Technology in Biomedicine, 14(4), pp1114-1120

[3] Berks, G., Keyserlingk, D., Jantzeen J., Dotoli, M and Axer H. (2000). *Fuzzy clustering: versatile means to explore medical database*, ESIT, Aachen, Germany

[4] Canadian Diabetes Association. *Standards of Medical Care in Diabetes*

– *2014*. Journal of Diabetes Care, 32: S13-61.

[5] Caudill M., (1987). *Neural Networks Primer*, AI Expert, pp46-52.

[6] Chakraborty, R., (2010). *Soft computing and fuzzy logic*, Lecture notes 1-37, retrieved from http://www.myreaders.info/07_fuzzy_systems.pdf

[7] Chinenye, S and Young, E., (2011). State of diabetes care in Nigeria: a review, The Nigerian Health Journal, 11(4), pp101-106

[8] Coello, C., Pulido, G and Lechuga, M., (2004). *Handling multiple objectives with particle swarm optimization*, Proc. of Evolutionary Computing, 8, pp 256–279.

[9] Dawson, C and Wilby, R., (2001). *Comparison of neural networks in river flow forecasting*, J. of Hydrology and Earth Science, SRef-ID: 1607-7938/hess/2001-3-529.

[10] Edo, A.E., Edo, G.O., Ohehen, O.A., Ekhator, N.P and Ordiah, W.C., (2015). Age and diagnosis of type-2 diabetes seen in Benin City Nigeria, African Journal of Diabetes Medicine, 23(1), pp18-19.

[11] Fausett L., (1994). *Fundamentals of Neural Networks*, New Jersey: Prentice Hall, pp.240.

[12] Giles, D and Draeseke, R., (2001). *Economic Model Based on Pattern recognition via fuzzy C-Means clustering algorithm*, Economics Dept Working Paper, EWP0101, University of Victoria, Canada.

[13] Goldenberg, R and Punthakee,, Z (2013). *Definition, classification and diagnosis, prediabetes and metabolic syndrome*, 37(1), S8-S11.

[14] Guo, W.W and Xue, H., (2011) *An incorporative statistic and neural approach for crop yield modelling and forecasting*, Neural Computing and Applications, 21(1), pp109-117

[15] Heppner, H and Grenander, U., (1990). S*tochastic non-linear model for coordinated bird flocks*, In Krasner, S (Ed.), *The ubiquity of chaos* pp.233–238. Washington: AAAS.

[16] Inan, G and Elif, D., (2005). *Adaptive neuro-fuzzy inference system for classification of EEG using wavelet coefficient*, retrieved from http://www.rorylewis.com/PDF/04

[17] Jang, J.S., (1993). *Adaptive fuzzy inference systems*, IEEE Transactions on Systems, Man and Cybernetics, vol.23, pp.665–685.

[18] Khashei, M., Eftekhari, S and Parvizian, J (2012). *Diagnosing diabetes type-II using a soft intelligent binary classifier model*, Review of Bioinformatics and Biometrics, 1(1), pp9-23.

[19] Kuan, C and White, H., (1994). *Artificial neural network: econometric perspective*", Econometric Reviews, Vol.13, Pp.1-91 and Pp.139-143.

[20] Ludmila I. K. (2008), *Fuzzy classifiers* Scholarpedia, 3(1), pp2925

[21] Mandic, D and Chambers, J., (2001). *Recurrent Neural Networks for Prediction: Learning Algorithms, Architectures and Stability*, Wiley & Sons: New York, pp56-90.

[22] Minns, A., (1998).*Artificial neural networks as sub-symbolic process descriptors*, published PhD Thesis, Balkema, Rotterdam, Netherlands

[23] Nascimento L., (1991). *Comparative Genomus of two Leptospira interrogans Serovirs*, http://jb.asm.org/cgi/content/full/186/7/2164

[24] Ojugo, A.A, (2012). *Artificial neural networks gravitational search algorithm for rainfall runoff modeling in hydrology*, unpublished PhD Thesis, CS Department: Ebonyi State University Abakiliki, Nigeria.

[25] Ojugo, A., Eboka, A., Okonta, E., Yoro, R and Aghware, F., (2012). *GA rule-based intrusion detection system*, Journal of Computing and Information Systems, 3(8), pp 1182 - 1194.

[26] Ojugo, A.A., Emudianughe, J., Yoro, R.E., Okonta, E.O and Eboka, A., (2013). *Hybrid neural network gravitational search algorithm for rainfall runoff modeling*, Progress in Intelligence Computing and Application, 2(1), doi: 10.4156/pica.vol2.issue1.2, pp22–33.

[27] Ojugo, A.A., Ben-Iwhiwhu, E., Kekeje, D.O., Yerokun, M.O and Iyawa, I.J.B., (2014). *Malware propagation on time varying network*,Int. J. Modern Edu. Comp. Sci., 8, pp25 – 33.

[28] Perez, M and Marwala, T., (2011). *Stochastic optimization approaches for solving Sudoku*, IEEE Transaction on Evol. Comp., pp.256–279.

[29] Reynolds, R., (1994). *Introduction to cultural algorithms*, Transaction on Evolutionary Programming (IEEE), pp.131-139.

[30] Ursem, R., Krink, T., Jensen, M.and Michalewicz, Z., (2002). *Analysis and modeling of controls in dynamic systems*. IEEE Transaction on Memetic Systems and Evolutionary Computing, 6(4), pp.378-389

[31] Vaarala, O., Knip, M., Paronen, J., Hamalainen, A.M., Muona, P., Vaatainen, M., Ilonen, J., Simell, O and Akerblom, H.K., (1999). *Cow's milk formula feeding induces primary immunization to insulin in infants at genetic risk for type-1 diabetes*, Diabetes, 8(7), pp1389-1394.

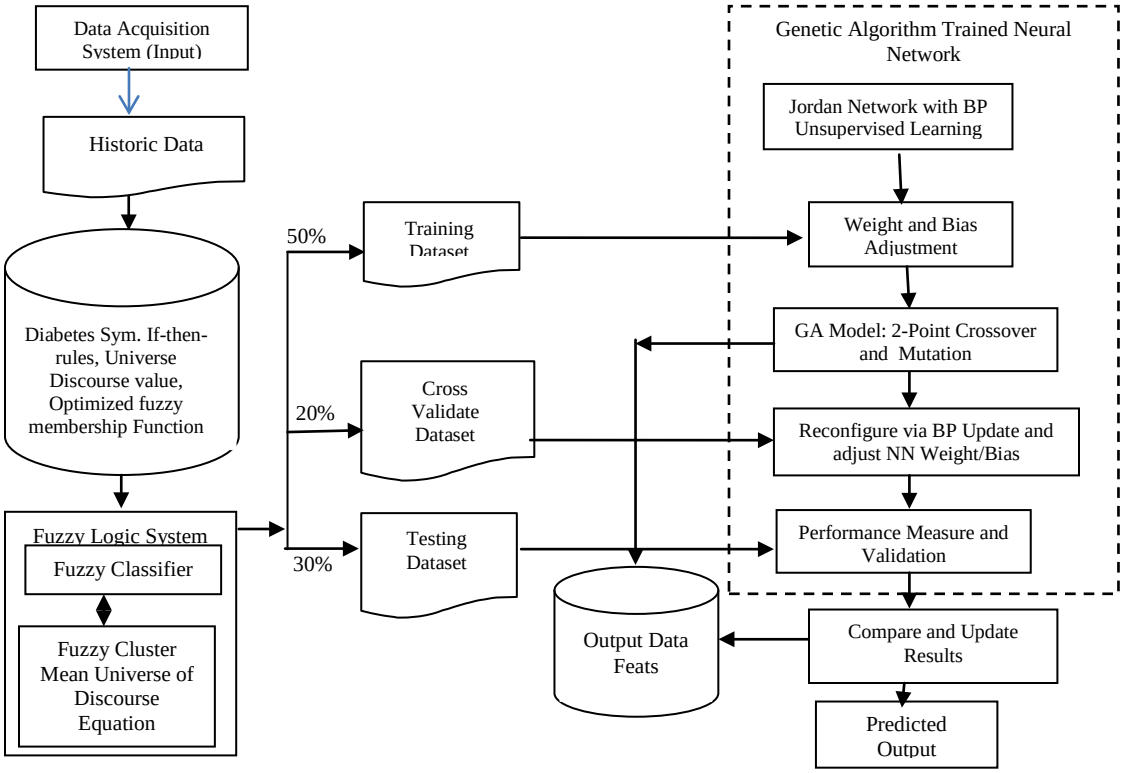


Figure 1: Dataflow Diagram of the Hybrid Model Fuzzy Genetic Algorithm Trained Neural Network

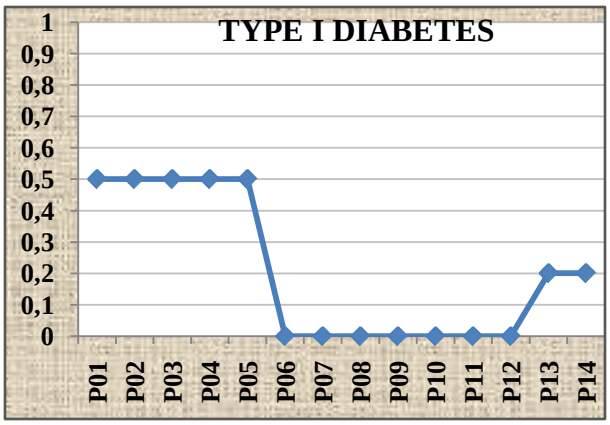


Figure 2a: Cluster I representation Type-I Diabetes.

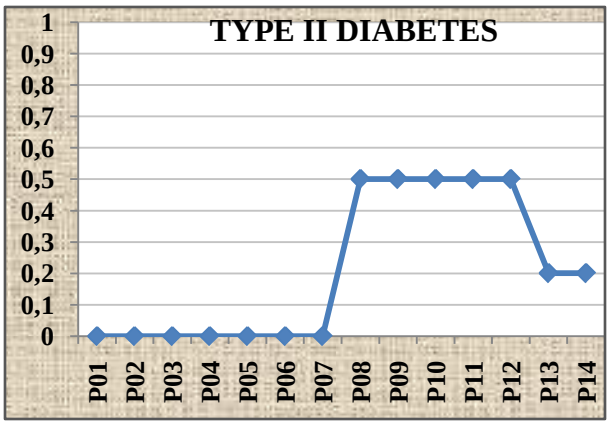


Figure 2b: Cluster II representation Type-II Diabetes.

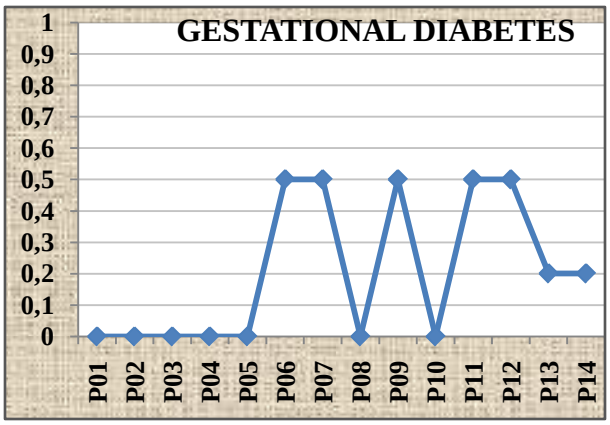


Figure 2c: Cluster III representation Gestational Diabetes.

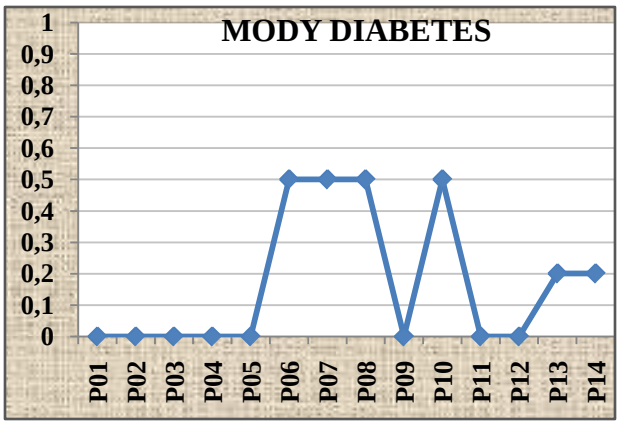


Figure 2d: Cluster IV representation MODY-Diabetes.

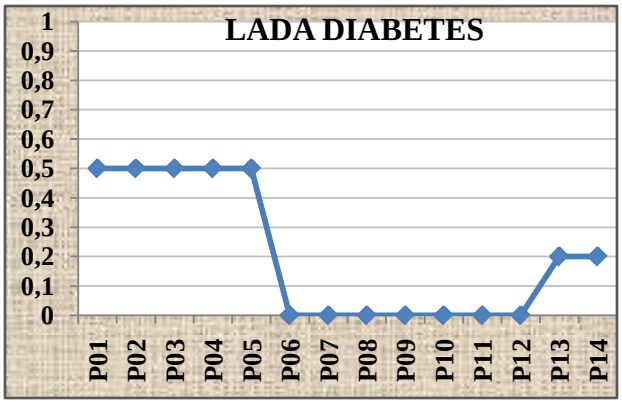
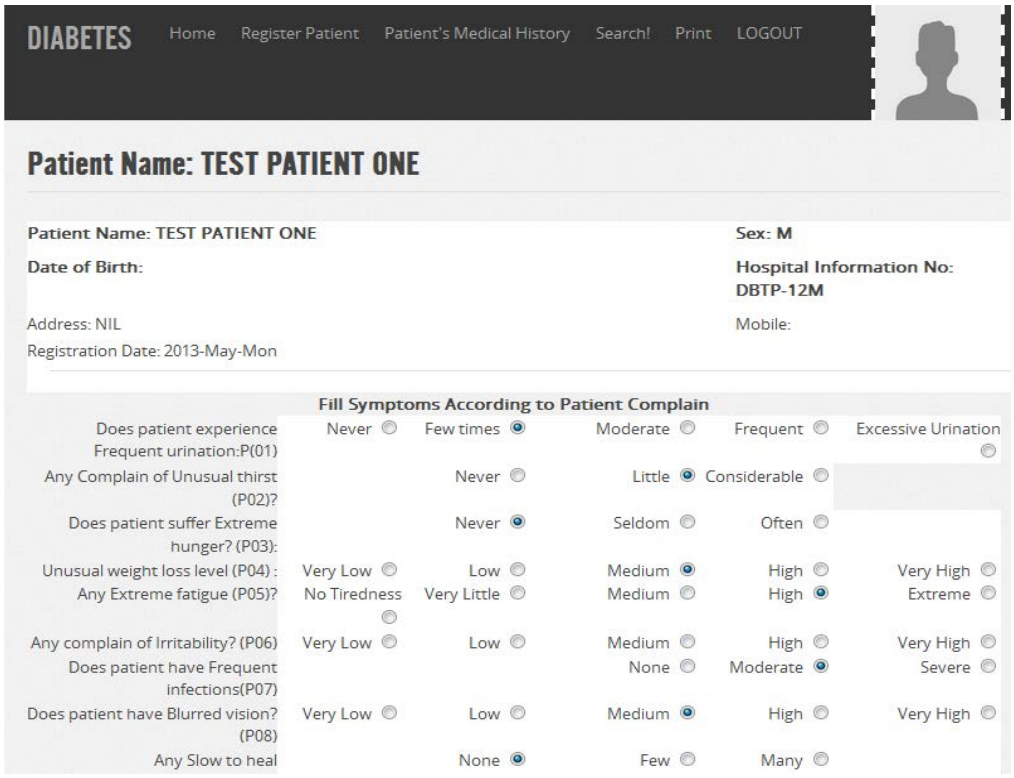
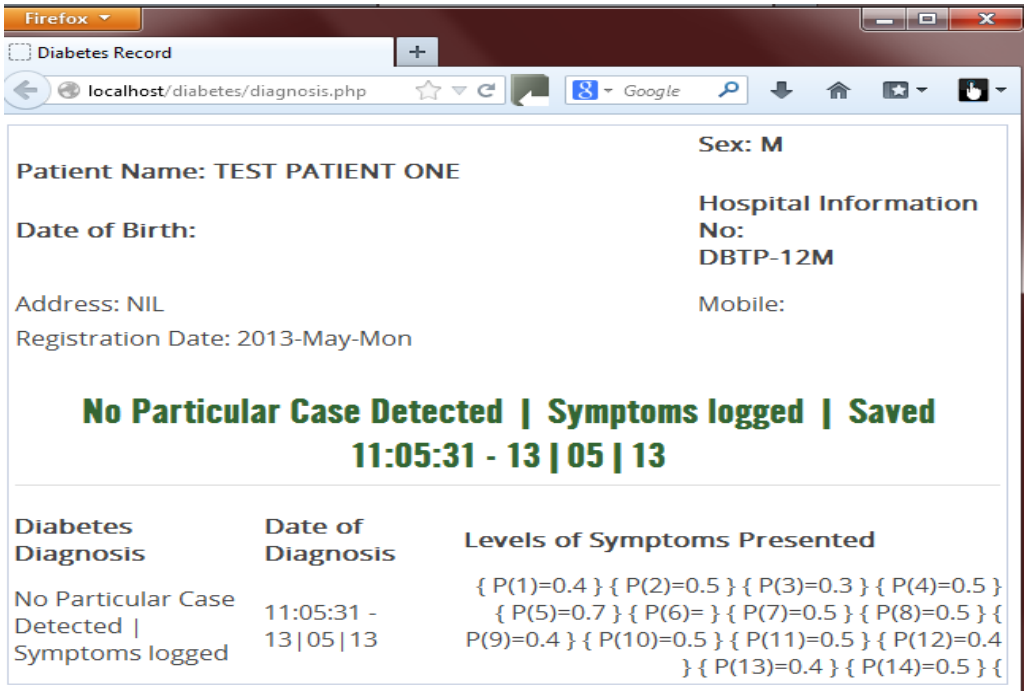


Figure 2e: Cluster V representation LADA-Diabetes.

ISBN: 978-1-61804-327-6

Sample Output of Decision Page and Report Page





Optimization Approach to the Solving of the Problem of N-version Software Systems Design

I. V. Kovalev and P.V. Zelenkov, D.I. Kovalev
Siberian State Aerospace University, P.O. Box 486, 660014-Krasnoyarsk, Russia

Abstract - The problem of developing an optimal structure of N-version software system presents a kind of very complex optimization problem. This causes the use of deterministic optimization methods inappropriate for solving the stated problem. In this view, exploiting heuristic strategies looks more rational. In the field of pseudo-Boolean optimization theory, the so called *method of varied probabilities* (MVP) has been developed to solve problems with a large dimensionality. Some additional modifications of MVP have been made to solve the problem of N-version systems design. Those algorithms take into account the discovered specific features of the objective function. The practical experiments have shown the advantage of using these algorithm modifications because of reducing a search space.

Keywords - N-version software, optimal structure, software system design, pseudo-Boolean optimization.

I. Introduction

Development of high-reliable fault-tolerant systems is an interesting engineering problem having not only technical meaning but also social importance. Systems of this kind determine the stability in social and technical environments, and multiple examples of such systems' crashes prove the strong need for more reliable constructions which can be realized through the use of up-to-date methods and approaches.

The rapid progress of computer technique of late years has made the software an essential part of any complex automated system. The reliability of software component may determine the reliability of whole the hardware-software system. That's why during last years large attention is paid to the development of the methodologies of designing high-reliable software complexes [1-5].

Practically, multi-channel tools increasing the system reliability at the expense of a multiple duplication of certain structure elements are very much in evidence. This approach has given a good account of itself in the designing of hardware parts of complex systems. The use of this methodology leads to a sizable decreasing of appearance probability of random errors having the physical nature. In turn, this approach is not an influence on software reliability, since it doesn't trace so called *dormant* (or sleeping) errors which could arise while writing the program code by a stated specification [6].

The multi-version programming, as a methodology of the fault-tolerant software systems design, allows successful solving of the mentioned tasks. The idea of multi-version programming has been introduced by A. Avizienis in 1977 [7]. The term N-version programming (NVP) used in the literature is of equal meaning and often takes place in papers on the observed methodology. A. Avizienis introduced NVP as *an independent generation* of $N \geq 2$ functionally equivalent software modules from the same *initial specification*. The *concurrent execution tools* are provided for such the modules. In cross-check points (cc-points) software modules generate cross-check vectors (cc-vectors). The components of the cc-vectors and the cc-points are to be determined in the specification set.

The use of N-version programming approach turns out to be effective, since the system is constructed out of several parallel executed versions of some software module. Those versions are written to meet the same specification but by different programmers. Where, the writing process of each version of concrete software module in any way must not intersect with or depend on another version code writing. This is done to avoid the presence of same dormant (or sleeping) errors in separate software designs. This kind of errors is typical for software components.

The problem of developing the optimal structure of an N-version software system (NVS) is the following: to choose a set of software modules, so as to provide the highest reliability

ISBN: 978-1-61804-327-6

for the system subject to the budget constraint. Since a description of any possible system configuration is made through such the positioning of its components, we can say that an observed problem has the binary essence [8]. Moreover, the existing theory of pseudo-Boolean functions and their optimization contains strong tools for solving problems of this kind [9]. And that fact makes the use of binarization algorithms more affordable.

The process of a problem binarization consists in setting relationships between the system states and the binary space elements. In the case of our system model, we need to determine some Boolean vector the elements of which will characterize the system structure. Each element of such the Boolean vector will signify either presence or absence of corresponding system component [10].

In that way, before starting to describe the exact process of binarization, all the necessary terms should be coined and the presented system model should be overviewed in details.

II. Optimization Model for Structuring NVS

The structure of N-version software system is determined consisting of a set of tasks (a set $\mathbf{I}, \operatorname{card}\mathbf{I} = I$). All the tasks are divided into classes, i.e. a set of task classes is introduced as well ($\mathbf{J}, \operatorname{card}\mathbf{J} = J$).

To solve the tasks belonging to a certain class, there is a software module, which can be realized by any of its versions. Thus, $\mathbf{K}, \operatorname{card}\mathbf{K} = J$ – the set of software modules. Let us introduce the vector $\mathbf{S} = \{S_j\}\ (j = \overline{1, J})$, each component of which is equal to a number of module versions (S_j – the number of versions of module solving a task of class j) [11].

To describe the task belonging to particular classes, in [12] the authors define sets of tasks for every task class. That is introduced as two-dimensional array in programming terms. Since the numbers of tasks belonging to different classes are not equal, that may cause some difficulties when translating the analytic expressions into a program code.

Here, it is proposed to use only one set the capacity of which is equal to the number of tasks in a system. And each element of this set is equal to the number of class a task belongs to. So, the set $\mathbf{B}, \operatorname{card}\mathbf{B} = I$ is the set of class membership of tasks, i.e. the element B_i of the set B presents the number of class the i-th task belongs to.

Using the introduced notations, lets us determine a common analytic form of the number of versions solving i-th task. If the element B_i of the set B is the number of class the i-th task belongs to, then an element of the set S, the index number of which is equal to B_i, determines the number of versions in a module solving i-th task. Therefore, this number can be written like this S_{B_i}.

Basing on that, we will introduce the Boolean variables X_s^i to describe the control implication of different module versions:

$$X_s^i = \begin{cases} 1, & \text{the } s-\text{th } (s = \overline{1, S_{B_i}}) \text{ version of module } B_i \text{ is used to solve the } i-\text{th } (i = \overline{1, I}) \text{ task,} \\ 0, & \text{the } s-\text{th } (s = \overline{1, S_{B_i}}) \text{ version of module } B_i \text{ is } \mathbf{not} \text{ used to solve the } i-\text{th } (i = \overline{1, I}) \text{ task.} \end{cases} \quad (1)$$

(1)

Expanding the introduced variables into the implication vector is the head moment in applying pseudo-Boolean optimization methods to the considered systems design.

Since a vector component number is specified by only one index and we deal with two-index variables, it is necessary to establish an algorithm forming an implication vector and an algorithm determining the component indices of this vector. Following section contains the algorithms to convert a problem of optimal structure design for N-version systems to a problem of pseudo-Boolean optimization and vice versa.

A. Conversion Algorithms

The algorithm of an implication vector forming acts in the following way (see the scheme on fig. 2). The first component of an implication vector X describes the first version of a module to be involved in the solving of the first system task. If the software module which solves the first task has more than one version, the next component of vector X characterizes the second version of the first task module. In this way, all the versions of all software modules are overhauled.

ISBN: 978-1-61804-327-6

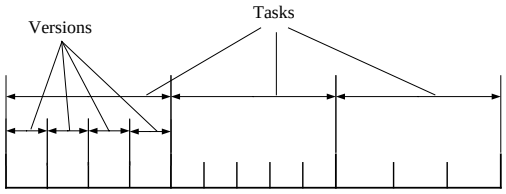


Fig. 1. An example of the implication vector

(2)

Hence, in order to determine the number of a vector X component, being aware of corresponding number of a task i and a number of version s, it is necessary to sum the number of versions in the modules solving the first (i-1) tasks and to add s to obtained value.

Analytically this conversion appears as follows:

$$pos = \begin{cases} s & , if\ i = 1, \\ \sum_{j=1}^{i-1} S_{B_j} + s & , if\ i > 1. \end{cases} \quad (2)$$

(3)

In order not to recalculate the first sum (the number of versions in the first (i-1) tasks) every time when optimizing a system, it would be better to count those sums depending on different i and to memorize them in an index vector:

$$G_i = \sum_{j=1}^{i} S_{B_j}$$, or in recurrent form

$$G_i = \begin{cases} S_{B_1} & , if\ i = 1, \\ G_{i-1} + S_{B_i} & , if\ i > 1. \end{cases} \quad (3)$$

(4)

Therefore, the value of the i-th component of vector G equals the number of versions in modules solving the tasks from the first up to the i-th. It results from this that the value of the last vector G component is equal to n – the implication vector dimensionality, i.e.

$$\sum_{i=1}^{N} S_{B_i} = n$$.

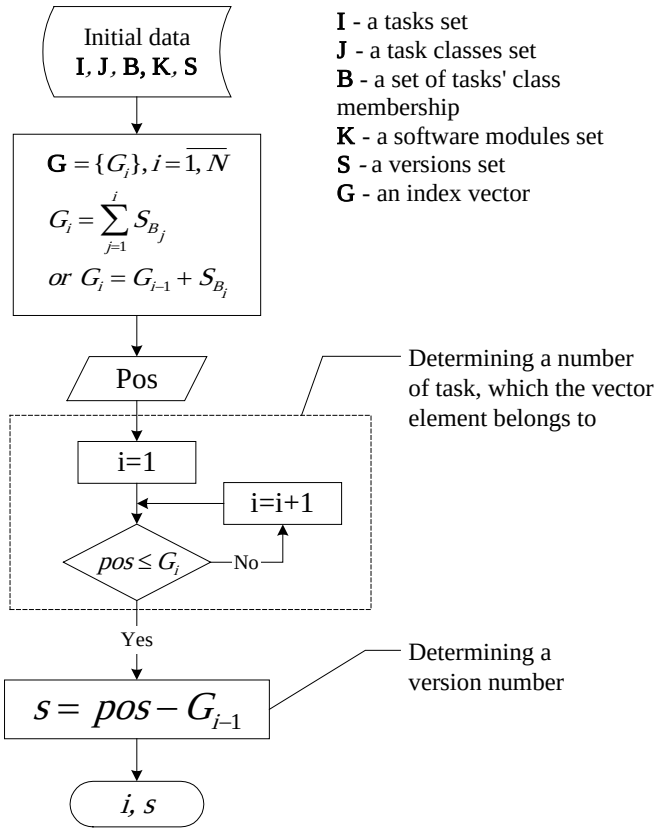
(5)

Once the index vector is introduced, the analytic record of a calculation of the implication vector component number takes the form of the following:

$$pos = \begin{cases} s & , if\ i = 1, \\ G_{i-1} + s & , if\ i > 1. \end{cases}$$

(6)

The reverse conversion task (a conversion of the implication vector component number to the numbers of task and version) consists of the consecutive determining of i and s. The flowchart of this algorithm is show on the fig. 2.



(7)

Fig. 2. A conversion of the implication vector component number into the numbers of task and version

Since the i-th element of the index vector equals the number of versions in modules solving the tasks from the first up to the i-th, the task number is determined by comparing the index of the implication vector component with the elements of the index vector. The comparison is being made from the first element of the index vector till the last one sequentially. And when the value of the parameter pos turns out to be less than or equal to the value of the index vector component, the required task number takes the value of this element number.

Then subtracting the number of versions in all the tasks from the first up to the (i-1)-th (equal to G_{i-1}) from pos we get a version number of software module solving the i-th task corresponding to pos.

The two presented algorithms are the core of applying binary approach to solve the stated

problem. Thus, having received the tools for a problem conversion, it became possible to use the methods developed within the confines of pseudo-Boolean optimization study. Some the features of the considered problem are discussed in the following section. Basing on this the conclusions about the relevant methods are made.

B. The mathematical Statement of the Problem

The converting algorithms considered above allow to describe the NVS structure in a form of a Boolean vector. As it was noticed previously, the optimal design of control system is held subject to different parameters: the reliability (it should be as big as possible), the cost (it should be as small as possible or at least it shouldn't exceed some limit), the allocation & scheduling and so on [14].

In terms of optimization theory, a system reliability function of a system structure is nothing else but an objective function. And conditions imposed on the system structure are the constraints set to limit the objective function domain [13]. Since we are able to associate a system structure with a Boolean vector, an objective function is a pseudo-Boolean one. And an optimization problem becomes a pseudo-Boolean one too.

In the framework of the presented model we will use a system reliability function as the objective function and the system cost will be the constraint imposed on the system [15-19]. In analytic form this problem can be written as follows:

$$\max R = \prod_{i=1}^{I} R_i ,$$

$$R_i = 1 - \prod_{s=1}^{s_{B_i}} (1 - R_{B_i s})^{X_s^i}$$

where

subject to

$$\sum_{i=1}^{I} \sum_{s=1}^{s_{B_i}} X_s^i \cdot C_{B_i s} \le B .$$

(8)

Here, $R_{B_i s}$ and $C_{B_i s}$ are the reliability and the cost of the software version s from module which solves the task of class Bi

III. Optimization Algorithms to form the NVS Structure

To derive an optimal dependability solution by means of an systematic, the exhaustive comparison algorithm would mean that all potential system configurations have to be tentatively generated, checked for the fulfillment of the side conditions and processed to compute the corresponding overall system reliability. This usually would cause a computing complexity that is untractable even for the most modern high speed computers: if, e.g., we consider a system consisting 64 modules, all of which are to be triplicated, thereby selecting each of the module versions from 5 different candidate modules, we would have to consider $[5!/3!\cdot 2!]^{64} = 10^{64}$ different system configurations! Assuming e.g. 1 nsec for processing each system configuration (of course, a value by far too optimistic!), the resulting 1055 sec of needed computation time would exceed the estimated age of the universe of about 1017 sec by many orders of magnitude! Therefore, here only stochastic search methods appear possible to provide, in a heuristic way, an optimal solution.

A. The Method of Varying Probabilities

In the field of pseudo-Boolean optimization theory, the so called method of varied probabilities has been developed to solve complicated problems, especially those ones with a large dimensionality [8]. The method of variable probabilities (MVP) presents a family of heuristic algorithms based on the common scheme: in order to find an extremal solution of a pseudo-Boolean optimization problem, a probability vector of dimensionality of sought solution vector is formed. Each component of the probability vector presents a probability of assigning a one value to the correspondent component of a Boolean vector. In the terms of developing NVS structure, it looks like a probability to include a version-candidate into the system structure.

The initial values of the probability vector components describe a situation when every software version has the equal probability to be included into the system structure. Then, at a computational phase, random decisions are generated according to the probability distribution specified by means of the probability vector. Each time the objective function is calculated in several random points, the values of the probability vector components are updated, so changing a probability distribution form. The way of changing these values defines a separate algorithms of MVP scheme. The common approach for updating a probability vector can be characterized by the rule: the better result received with a one-valued binary vector component the

bigger probability is assigned for it to get the value of one in the final solution.

These scheme can be augmented whether by some special methods for updating the probability vector or through involving the peculiar procedures of generating random solutions at a computational phase of an algorithm. This paper discusses the two methods for updating the probability vector (ARSA and Modified ARSA ver. 1) and the two procedures of generating random solutions (the independent generation of random points and the generation of non-zero solutions) giving thus as a result four different realizations (algorithms) of MVP.

The Adaptive Random Search Algorithm (ARSA) plays a role of the background for the rest of the algorithms of MVP scheme [13]. Initially, ARSA has been developed for the problem of pattern recognition to select an informative subsystem of attributes. The main disadvantage of this algorithm is a potential problem of updating values of probability vector components. Namely, in some cases it is possible to get the values of intermediate solutions which do not let the probability vector components to be changed. To correct the defect, the modification of ARSA has been developed (Modified ARSA ver.1). The statistical data of applying the modified version of ARSA display the better behavior of the algorithm when solving problems of developing a structure of NVS.

Next, applying to the stated optimization problem, ARSA doesn't provide a technique of avoiding zero-solutions when solving the problem of designing NVS structure. To protect an algorithm against spending both computational and time resources for calculating the objective function values in the points of this kind, the particular technique of generating random non-zero solutions has been developed. This technique is utilized in the MVP based algorithm named NVS MVP (mentioning the strict field of using the algorithm).

Making use of both of the mentioned enhancements gave a great raise in the efficiency of applying the MVP based algorithms to the problem of NVS structure development. The statistical results presented in the final part of the paper show it. Different algorithms have been tasted on the same optimization problem with the same quantity of objective function calls.

The objective function of the presented optimization problem has several specific features which can assist to reduce a search domain, thus allowing to decrease the searching time. The objective function as a function of the whole system reliability represents the product of reliabilities of separate software modules. Consequently, when a reliability of any of the modules is equal to zero the overall system reliability turns into zero value also. Physically, it represents a case when there are no versions chosen for (at least) some of the software modules. The implication vector components corresponding to such the software modules will be assigned zeroes as well. Obviously, it is necessary to avoid computing the objective function in such the points.

The number of system structures having at least one software module without versions-candidates assigned can be determined as the difference of the number of all the possible structures and the quantity of the structures which provide every software module with at least one candidate, i.e. $N_0 = N_{all} - N_{R>0}$. The number of all possible structures is determined through the dimensionality of an implication vector n as follows $N_{all} = 2^n$. The second intermediate value is found basing on the multiplication principle from combinatorics as a number of all possible structures with software module combinations each without one of them (that with no versions assigned). Formally, it is described in the following way: $N_{R>0} = \prod_{i=1}^{I} (2^{k_i} - 1)$, where I is the number of software modules, k_i represents a number of versions for the i-th software module.

Then, the final expression determining a sought value looks like this:

$$N_0 = 2^n - \prod_{i=1}^{I} (2^{k_i} - 1) \quad (9)$$

The value of this expression depends on an overall number of candidates (a dimensionality of the optimization problem), a number of software modules I and the numbers of versions for each of the software module (ki, $i = \overline{1, I}$). In general case, this expression takes grand values counting up to $0.9 N_{all}$, i.e. 90% of all the possible solutions. This means in this case that in order to find a solution it is sufficient to search through only 10% of the definitional domain of the objective function.

Unfortunately, there is the other side of the question making this result not so optimistic. Namely, for the problems of large dimensionalities reducing the search domain to

10% means diminishing the dimensionality of a problem by very small value. For instance, for a test problem of dimensionality n=117, avoiding all the null-valued points lowered the problem dimensionality only down to $n_{R>0} = 116$.

Nevertheless, exploiting this feature of the objective function has given satisfactory results when applying the algorithms of the method of varied probabilities (MVP). The modification of the MVP with the ability of avoiding null-valued points called NVS MVP has its own way of generating random points at iterational steps of the algorithms. In NVS MPV, random points are generated so that to provide each software module with at least one version.

At every iterational step, the whole implication vector generated is concerned as consisting of parts each describing the structure of a separate software module. Thus indeed, random vector generating consists of generating of random structures of modules. This approach allows having only non-zero solutions in result.

B. The Random Search of Boundary Points

Another stochastic algorithm to optimize the structure of NVS is the algorithm of random search of boundary points [8]. It is based on the proved fact that a solution of the stated optimization problem is a so called boundary point. Or in terms of binary space topology, a point neighboring to the set of infeasible solutions. Such a point describes a system structure which can not be updated through including a software versions additionally without violating the resource conditions, i.e. no version can not be added to a system structure of this kind paying attention to restrictions. The algorithm of random search of boundary points constitutes a generating of multiple boundary solutions and comparing the objective functions values in them.

The constraint in this optimization problem partitions the whole binary space into two domains – the domain of solution satisfying the constraint function and a set of points not satisfying to the constraint. It is shown that these domains represent the connected sets and that a solution of correspondent optimization problem is a point neighboring to the set of infeasible solutions. This kind of solution can be called a boundary point.

Basing on the results stated above, it is clear that it is sufficient to search among only boundary points in order to find the best value of the objective function. Thus, the problem of finding a best solution becomes a problem of an exhaustive search on the boundary points set.

The following is the algorithm of generating boundary point for the problem of developing the optimal structure of NVS (Fig. 3).

Different boundary points can be reached using this algorithm when different combinations of ways to choose i at the second step of the algorithm will be followed.

Hence, the algorithm of searching boundary points will have the following scheme.

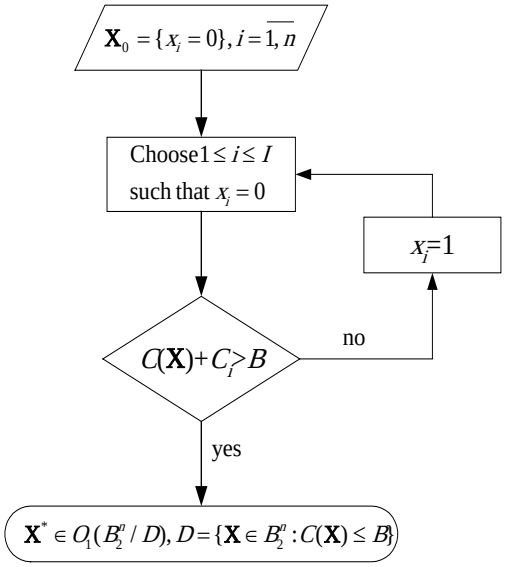
The initializing step: i=0.

Determine a boundary point Xbi (b – as an index means “boundary”).

Calculate the objective function value Fi=F(Xbi).

If the stopping condition is satisfied go to p. 5, otherwise i=i+1 and go to p. 2.

The solution is $F^* = \max_i F_i$.



(10)

Fig. 3. Generating a boundary point

Separate variations of the algorithm of boundary points search may differ from each other in a stopping condition and in ways of reaching boundary when generating boundary points. For the optimization problems of high complexity it is more rational to use stochastic version of the algorithm when boundary points are reached in a random way and this process is executed repeatedly.

IV. The Comparison of the Random Optimization Procedures

Concluding the paper, let us cite the comparative data of the computational results for different random optimization procedures. To gather such the information, the presented algorithms have been used to solve the test NVS structure optimization problems. The efficiency of the random search algorithms has been judged by the values of the objective and constraining functions.

The problem of dimensionality 117 has been chosen as the test problem, i.e. the developed software system included 117 software versions. It's worth mentioning that every of the random search algorithms needed approximately same period of time for calculating under equal conditions. That's why the time has not been set as an efficiency characteristic.

Table 1 contains the computational results of algorithms testing. The best searching capabilities have been revealed with the use of NVS MVP algorithm and the algorithm of

boundary points search. The latter displayed the highest stability of the solutions found, although using NVS MVP it is sometimes possible to find more reliable system configurations.

V. Conclusion

The problem of structuring an N-version software system is specified by the binary character, what made it plausible to apply the methods of pseudo-Boolean optimization. Within the limits of the discrete optimization a set of methods and algorithms has been proposed. The search capabilities of each of the algorithms realized have been investigated by solving the test problems. It was shown that the modification of the method of varying probabilities for NVS MVP together with the algorithm of boundary points search provide the best searching capabilities concerning the time efficiency and the solution quality.

TABLE 1. The results of random search algorithms working.

№	Budget constraint B	Number of iteraions	The random search algorithms			
			NVS MVP		Random search of boundary points	
			$R(X^*)$	$C(X^*)$	$R(X^*)$	$C(X^*)$
1.	800	15000	*0.7872*	789	*0.8074*	796
			0.7916	791	*0.7998*	797
			0.7907	791	*0.8177*	794
		30000	*0.8136*	786	*0.8318*	794
			0.8054	775	*0.8331*	797
			0.8118	784	*0.8377*	798
2.	900	15000	*0.9040*	850	*0.9149*	899
			0.9207	896	*0.9164*	899
			0.9039	887	*0.9148*	898
		30000	*0.9076*	867	*0.9192*	897
			0.9082	875	*0.9167*	897
			0.9155	890	*0.9177*	892
3.	1000	15000	*0.9701*	995	*0.9622*	993
			0.9523	986	*0.9609*	998
			0.9546	989	*0.9635*	998
		30000	*0.9651*	994	*0.9652*	997
			0.9554	988	*0.9661*	995
			0.9712	997	*0.9631*	996

REFERENCES

[1] Laprie J.-C. et al. Hardware- and Software-fault tolerance: definition and analysis of architectural solutions// Proceedings of the IEEE, 1987, Pp. 116-121.

[2] Anderson T., Barrett P.A., Halliwell D.N., Moudling M.L., “An evaluation of software fault tolerance in a practical system”, Proc. Fault Tolerant Computing Symposium 1985, pp. 140-145.

[3] Scheer, S., Maier, T.: Towards Dependable Software Requirement Specifications. In: Daniel, P. (ed.) Proceedings of SAFECOMP 1997, New York (1997)

[4] Kovalev I. *et al* The control of developing a structure of a catastrophe-resistant system of information processing and control // In: 2015 *IOP Conf. Ser.: Mater. Sci. Eng.* 70 012008. doi:10.1088/1757-899X/70/1/012008.

[5] *Kovalev I.V.*, Engel E.A., Tsarev R.Ju. Programmatic support of the analysis of cluster structures of failure-resistant information systems. Automatic Documentation and Mathematical Linguistics. 2007. T. 41. № 3. C. 89-92.

[6] *Keene, S. J.* Comparing Hardware and Software Reliability. Reliability Review, 14(4), December 1994, pp. 5-21.

[7] *Avizienis A.* The methdology of N-version programming// In: Software fault tolerance/ edited by M.R. Lyu, Wiley, 1995, Pp. 23-47.

[8] Antamoshkin A., Schwefel H.P., Torn A., Yin G. and Zilinskas A. System Analysis, Design and Optimization. Ofset Press, Krasnoyarsk, 1993.- 312 p.

[9] Antamoshkin, A.N. Random Search Algorithm for the p-Median Problem / A.N.

Antamoshkin, L.A. Kazakovtsev // Informatica – 2013. - № 3(37). – P. 127–140.

[10] Lev Kazakovtsev, Predrag Stanimirovic, Idowu Osinga, Mikhail Gudima and Alexander Antamoshkin / Algorithms for location problems based on angular distances. Advances in Operations Research. – 2014. – Vol. 2014. Articale ID 701267. 12 pages. - http://www.hindawi.com/journals/aor/raa/701267/

[11] *Kovalev I.* Optimization problems when realizing the spacecrafts control// In: Advances in Modeling and Analysis, C, Vol. 52, No. 1-2, 1998, pp. 62-70.

[12] Kovalev I.V., Dgioeva N.N., Slobodin M.Ju. The mathematical system model for the problem of multi-version software design. Proceedings of Modelling and Simulation, MS'2004 AMSE International Conference on Modelling and Simulation, MS'2004. AMSE, French Research Council, CNRS, Rhone-Alpes Region, Hospitals of Lyon. Lyon-Villeurbanne, 2004.

[13] *Kovalev I., Grosspietsch K.-E.* Deriving the Optimal Structure of N-version Software under Resource Requirements and Cost/Timing Constraints// Proc. Euromicro' 2000, Maastricht, 2000, IEEE CS Press, pp. 200-207.

[14] Kovalev I. et al. The Minimization of Inter-Module Interface for the Achievement of Reliability of Multi-Version Software / Kovalev I., Zelenkov P., Ognerubov S // 2015 IOP Conf. Ser.: Mater. Sci. Eng. 70 012006. doi:10.1088/1757-899X/70/1/012006.

[15] *Ashrafi N. et al.* Optimal Design of Large Software-Systems Using N- Version Programming// IEEE Trans. on Reliability, Vol. 43, No. 2, 1994, pp. 344-350.

[16] *Ashrafi N., Berman O.* Optimization models for selection of programs, considering cost and reliability// IEEE Trans. on Reliability, Vol. 41, No. 2, 1992, pp. 281-287.

[17] *Ashrafi N., Berman O.* Optimization models for reliability of modular software systems// IEEE Trans. on Software Engineering, Vol. 19, No. 11, November 1993, pp. 1119-1123.

[18] Kovalev I. et al. Model of the reliability analysis of the distributed computer systems with architecture "client-server". IOP Conf. Series: Materials Science and Engineering 70 (2015) 012009. doi:10.1088/1757-899X/70/1/012009.

[19] Kovalev, I.V. Fault-tolerant software architecture creation model based on reliability evaluation / I.V. Kovalev, R.V.Younoussov; Advanced in Modeling & Analysis, vol. 48, № 3-4. Journal of AMSE Periodicals, 2002, pp. 31-43.

Solving Motif Finding Problem on Heterogeneous Cluster using CPUs, GPUs, and MIC Architectures

H. M. Faheem
Professor of Computer Systems,
Ain Shams Univ., Egypt
hmfaheem@cis.asu.edu.eg

***Abstract*—** Motif Finding Problem MFP is one of the computationally intensive problems in bioinformatics domain. Solving such problem on heterogeneous cluster consisting of CPUs, CUDA GPUs, and Intel Many Core (MIC) architectures is considered a challenging problem. This paper solves the MFP on a heterogeneous cluster using a scheduling strategy intended to schedule tasks on heterogeneous architectures. The main idea is to solve the problem using suitable parallel computing paradigms such as MPI, OpenMP, and CUDA on individual architectures then to estimate the number of tasks that should be assigned to each one based on its speed. We can find that the total execution time will be significantly improved. The paper then shows how to modify the code to assign the tasks to the architectures.

***Keywords*—**Heterogeneous Architectures, Motif Finding Problem, Task Scheduling.

1. Introduction

Modern high performance computing (HPC) clusters have traditional multicore microprocessors (CPUs), graphics processor units (GPUs), and Intel many integrated core (MIC) architectures. This in turn leads to more heterogeneity among the computational resources within a single cluster. Writing an efficient code that can optimally utilizes these heterogeneous resources depends mainly on the capabilities of the developer and the parallel computing paradigm he uses. Moreover, the scheduling strategy dealing with such heterogeneity is considered one of the most important factors affecting the performance of the heterogeneous systems.

In this paper, we will solve one of the computationally intensive problems in the bioinformatics field. The problem is called "Motif Finding Problem". We will use brute force algorithm to solve this problem three times. The first will be on multicore CPUs, while the second will be on GPGPU, and the third will be on MIC. Consequently, the actual run time for each will be calculated. Eventually, we will use a specific scheduling strategy to assign proper number of workload to the architectures to achieve near optimal hardware resource utilization. We will also see how to modify the code to cope with the deployed scheduling strategy.

The rest of this paper is organized as follows: section 2 describes the motif finding problem. Section 3 shows the implementation on different architectures. Section 4 briefly explains the deployed scheduling strategy. Section 5 presents the changes to the parallel code to fulfill the scheduling strategy requirements. Section 6 contains some concluding remarks and directions for future work.

ISBN: 978-1-61804-327-6

2. Motif Finding Problem

The Motif Finding Problem (MFP) can be simply considered as a string matching problem. Solving the MFP to find a motif of length *L* with permitted mutation *d* can be implemented using a brute-force algorithm. All the possible *L-mers* (4^L) are compared with each possible motif of length *L*. If we have a sequence of size *N* then we can have (*N-L+1*) motifs. Pevzner and Sze [1] presented the challenge problem (15, 4) where the first number is a specific length *L* and the second number a specific mutation *d*. In this paper, we present a problem in which the motif has a length *L*=15, allowed mutations *d*=4, and the number of sequences we are searching in is *T*=20 each of size *N*=600. Solving such computationally intensive problems can be implemented using a set of heterogeneous platforms [2, 3, 4, 5, 6, 7, 8, 9].

3. Implementing MFP on different Architectures

All the experiments on the system have been implemented using Intel Compiler 2015 and Intel MPI V5. GPU experiments implemented using CUDA V6 and GCC compiler and OpenMPI. MIC experiments are implemented using Intel Compiler 2015 and Intel MPI V5 using native mode for MIC. Table 1, 2, and 3 show the typical architectures of regular CPU node, MIC (Xeon Phi) node, and NVIDIA GPGPU (CUDA) node respectively. Infiniband network is used to connect different compute nodes.

Table 4 shows the scalability of the problem when using different number of CPU cores and MPI and OpenMP parallel computing paradigms. We can see that as the number of cores increases the time needed to find a solution considerably decreases. The table also shows the implementation time for solving the problem on Xeon Phi and CUDA.

The problem now is how to efficiently utilize such heterogeneous resources. Of course, the problem can be solved in 56 seconds using 256 regular CPU nodes which mean 6144 CPU cores. Using CUDA will solve the problem in 3234 seconds while solving the problem using MIC will take 22,446 seconds. Searching for an efficient solution in terms of time constitutes one of the major challenges.

4. Scheduling Strategy

In this section we will examine the impact of using scheduling strategy described in [10] on assigning tasks to different architectures. Consequently, we will show how the code will be affected based on such scheduling strategy. The objective of this scheduling strategy is to minimize the time needed to solve the MFP. Since we have 4^{15} tasks, each task will compare one L-mer with all the possible windows extracted from all the given sequences, hence the total number of comparison operations in each task is described in (1) while the total number of comparison operations for all tasks is shown in (2). Table 5 shows the speed differences between architectures for implementing one task. The number of tasks assigned to the architectures and the implementation time are listed in Table 6.

$$C = (N - L + 1) * T \quad (1)$$

Where C is the total number of comparison operations in each task.

$$C_T = 4^{15} * (N - L + 1) * T \quad (2)$$

Where C_T is the total number of comparison operations for all tasks.

5. Modifications to the Code

The pseudo code shown in Fig. 1 represents the OpenMP implementation for finding the motif. The subroutine takes two input arguments. Those arguments are encoded in Quaternary numeral system as we have only 4 letters in the DNA string. i.e. if the Start = $(202033011211)_4$ this can be translated to GAGATTACCGCC. Given the Start and End motifs the MFP_OMP subroutine will search all possible values in this range of Motifs to find the highest score motif. The number of sequences that each Motif will be matched against is 20 sequences each of 600 charcters.

ISBN: 978-1-61804-327-6

Table 1: Regular CPU-based Compute Node

Attribute	Value
Architecture	x86_64
CPU op-mode(s)	32-bit, 64-bit
Byte Order	Little Endian
CPU(s)	24
On-line CPU(s) list	0-23
Thread(s) per core	1
Core(s) per socket	12
Socket(s)	2
NUMA node(s)	2
Vendor ID	GenuineIntel
CPU family	6
Model	62
Stepping	4
CPU MHz	2399.852
BogoMIPS	4799.33
Virtualization	VT-x
L1d cache	32K
L1i cache	32K
L2 cache	256K
L3 cache	30720K
Memory	96 GB
NUMA node0 CPU(s)	0-11
NUMA node1 CPU(s)	12-23

Table 2: Xeon Phi Compute Node

Attribute	Value
Total No of Active Cores	60
Voltage	897000 uV
Frequency	1052631 kHz

The pseudo code shown in Fig. 2 represents the MPI implementation for finding the motif. Similar to MFP_OMP subroutine the MFP_MPI subroutine divided the search space and each MPI rank uses the MFP_OMP routine to search its subspace. The root rank (rank 0) will collect and apply max reduction to find the global Motif. The same code is also used on MIC using Intel compiler with the following switch "-Mmic".

Fig. 3 shows the CUDA implementation for MFP. The CUDA takes very long time to process a single job which causes a timeout so we divided the search space into chunks of size 512*512*512*1 which is a 2D job. The thread index is considered the Motif that must be matched against the 20 sequences. As we have multiple batches for the job an offset is applied to the thread index to differentiate different chunks. The host side starts the CUDA jobs and collects the results in a global array. Finally the host finds the highest score in the global array.

Fig.4 shows the code used to schedule the MFP on the different architectures. The code is designed to get the available MFP implementations from the database. The code then checks if the required architecture(s) is online for each algorithm or not. The list of algorithms that can be implemented on the on-line architectures will have a pre-calculated ratio stored into the database. This ratio is generated by the scheduling strategy. These ratios should be normalized before dividing the chunks. This step is necessary in case one of the algorithms cannot run in the current state of the system due to unavailability of the required architecture.

Once the framework finds the available algorithms and matches them with the currently available hardware it starts the execution of the algorithms and divides the problem size among them according to the scheduling strategy.

ISBN: 978-1-61804-327-6

Table 3: NVIDIA CUDA Compute Node

CUDA Driver Version / Runtime Version	6.0 / 6.0
CUDA Capability Major/Minor version number	3.5
Total amount of global memory	5120 MBytes (5368512512 bytes)
(13) Multiprocessors, (192) CUDA Cores/MP	2496 CUDA Cores
GPU Clock rate	706 MHz (0.71 GHz)
Memory Clock rate	2600 Mhz
Memory Bus Width	320-bit
L2 Cache Size	1310720 bytes
Total amount of constant memory	65536 bytes
Total amount of shared memory per block	49152 bytes
Total number of registers available per block	65536
Warp size	32
Maximum number of threads per multiprocessor	2048
Maximum number of threads per block	1024
Max dimension size of a thread block (x,y,z)	(1024, 1024, 64)
Max dimension size of a grid size (x,y,z)	(2147483647, 65535, 65535)
Maximum memory pitch	2147483647 bytes

Table 4: Implementation Results of solving MFP on CPUs, Xeon Phi, and NVIDIA CUDA

Trial Number	Platform	Result (seconds)
1	1 Regular Node (OpenMP)	13373
2	1 Regular Node (MPI + OpenMP)	13263
3	2 Regular Nodes (MPI + OpenMP)	6590
4	4 Regular Nodes (MPI + OpenMP)	3353
5	8 Regular Nodes (MPI + OpenMP)	1688
6	16 Regular Nodes (MPI + OpenMP)	851
7	32 Regular Nodes (MPI + OpenMP)	430
8	64 Regular Nodes (MPI + OpenMP)	216
9	128 Regular Nodes (MPI + OpenMP)	109
10	256 Regular Nodes (MPI + OpenMP)	56
11	1 XEON Phi Node (Native Mode + OpenMP)	22446
12	1 GPU Node (CUDA)	3234

Table 5: Speed difference of architectures to complete one task (in sec.)

Architecture	CPU Node	GPU Node	Phi Node
Task Execution Time (in Sec.)	1.24E-05	3.01E-06	2.09E-05

ISBN: 978-1-61804-327-6

Table 6: Tasks assigned to architectures based on their speeds and the total execution time

Platform	CPU Tasks	CUDA Tasks	MIC Tasks	Result (Sec.)
1 Regular Node (OpenMP) (CPU) + 1 CUDA + 1 MIC	1.87776959E+08	7.74391252E+08	1.11573613E+08	2333.49
1 Regular Node (MPI + OpenMP) + 1 CUDA + 1 MIC	1.88640585E+08	7.73636387E+08	1.11464852E+08	2330.11
2 Regular Nodes (MPI + OpenMP) + 1 CUDA + 1 MIC	3.22317450E+08	6.56794062E+08	0.946303125E+08	1978.20
4 Regular Nodes (MPI + OpenMP) + 1 CUDA + 1 MIC	4.91150392E+08	5.09223025E+08	0.733684070E+08	1533.72
8 Regular Nodes (MPI + OpenMP) + 1 CUDA + 1 MIC	6.72283226E+08	3.50901078E+08	0.505575197E+08	1056.87
16 Regular Nodes (MPI + OpenMP) + 1 CUDA + 1 MIC	8.25285603E+08	2.17166991E+08	0.312892296E+08	654.08
32 Regular Nodes (MPI + OpenMP) + 1 CUDA + 1 MIC	9.31970990E+08	1.23916984E+08	0.178538504E+08	373.23
64 Regular Nodes (MPI + OpenMP) + 1 CUDA + 1 MIC	9.97518032E+08	0.66624581E+08	0.095992112E+08	200.66
128 Regular Nodes (MPI + penMP) + 1 CUDA + 1 MIC	10.3387509E+08	0.34846130E+08	0.050205998E+08	104.95
256 Regular Nodes (MPI + penMP) + 1 CUDA + 1 MIC	10.5288327E+08	0.18231744E+08	0.026268138E+08	54.91

Fig. 1 Pseudo code for OpenMP implementation for Motif Finding Problem

```
MOTIF  MFP_OMP(Start, End)
Begin
      Motif.value = Start;
      Motif.Score = 0;

      $OMP Directive
      For N = Start to End
      BEGIN
            For SeqIdx = 0 to SequenceCount
            BEGIN
                  Score[SeqIdx] = Score(Sequence[SeqIdx], N)
            END
            if (Score > Motif.Score)
            BEGIN
                  Motif.value = N
                  Motif.Score = Score-
            END
      END
      return Motif;
END
```

Fig. 2 Pseudo code for MPI implementation for Motif Finding Problem

```
MFP_MPI (Start, End)
BEGIN
      MyStart = (End-Start) / Size * Rank
      MyEnd = (End-Start/Size * (Rank+1) – 1
      LocalMotif = MFP_OMP(MyStart, MyEnd)
      MPI_AllReduce(MAX, LocalMotif, Root)
END
```

ISBN: 978-1-61804-327-6

Fig. 3 Pseudo code for CUDA implementation for Motif Finding Problem

```
MFP_CUDA_Device(Offset)
BEGIN
    For SeqIdx = 0 to SequenceCount
    BEGIN
        Score[SeqIdx] = Score(Sequence[SeqIdx], ThreadIdx+Offset)
    END
    Return Score
END
```

```
MFP_CUDA_Host (Start, End)
BEGIN
    //Single CUDA call can process 512*512*512*1.
    While (End-Start > 0)
    BEGIN
        Call MFP_CUDA_Device(Start)
        GetResultsFromDevice()
        Start = Start + 512*512*512*1
    END
    FindMaximumScore()
END
```

Fig. 4 Scheduling routine pseudo code to assign workloads to different architectures.

```
ScheduleJob(AlgorithmName, ProblemSize, ExtraArguments)
BEGIN
    Algorithms = GetListOfAvaliableImplementations(AlgorithmName)
    SpawnCmd[Algorithms.Length]
    Start = 0;
    End = 0;

    For Idx = 0 to Algorithms.Length
    BEGIN
        IF Start eq. End
            End = Algorithms[idx].Ratio * ProblemSize
        ELSE
            Start = End
            End = End + Algorithms[idx].Ratio * ProblemSize
        END
        System.Start(Algorithms[idx].ExecCommand, Start, End, ExtraArguments)
    END

END
```

```
GetListOfAvaliableImplementations(AlgorithmName)
BEGIN
    Algorithms = GetAlgorithmsFromDB(AlgorithmName)
    For Idx = 0 to Algorithms.Length
    BEGIN
        IF Architecture is not available for Algorithms[idx].Architecture
        BEGIN
            RemoveAlgorithm(Idx)
        END
    END
    NormalizeRatios(Algorithms)
    Return Algorithms
END
```

ISBN: 978-1-61804-327-6

6. Conclusion

Solving computationally intensive problems on heterogeneous architectures can significantly improve and speedup the implementation time of the problem solution when proper scheduling strategy and suitable parallel computing paradigms are used. Having a look to the results in tables 4 and 6 can give us an idea about the improvement in the total implementation time. For example; implementing the brute force algorithm using one regular node, one CUDA node, and one Xeon Phi node will reduce the implementation time from 13373 seconds on a single regular node to 2333 seconds with a speedup factor 5.7 while using four regular nodes, one CUDA node, and one Xeon Phi node will reduce the implementation time from 3353 seconds on pure 4 regular nodes to 1533 seconds with a speedup factor of 2.18. Since the number of CUDA nodes and Xeon Phi nodes are fixed in our cluster then we can find that as the number of regular nodes increases, the speedup factor decreases. Future work may include the use of different scheduling strategies and intelligent selection criteria to choose the best scheduling strategy to solve a given computationally intensive problem. We believe that this paper is a step towards a complete system to solve computationally intensive problems on heterogeneous architectures.

Acknowledgement

Computation for the work described in this paper was supported by King Abdulaziz University's High Performance Computing Center (Aziz Supercomputer) (http://hpc.kau.edu.sa).

References

[1] P. Pevzner and S. Sze, " Combinatorial approaches to finding subtle signals in DNA sequences," Proceedings of the 8th International Conference on Intelligent Systems for Molecular Biology, 269–78, 2000.

[2] H. M. Faheem, "Accelerating Motif Finding Problem using Grid Computing with Enhanced Brute Force," The 12th International Conference on Advanced Communication Technology (ICACT), Korea, 2010.

[3] M. Raddad, N. El-Fishawi, and H. M. Faheem, "Implementation of Recursive Brute Force for Solving Motif Finding Problem on Multi-Core," International Journal of Systems Biology and Biomedical Technologies, 2 (3):1-18, 2013.

[4] R. Inta and D. J. Bowman, "An FPGA/GPU/CPU hybrid platform for solving hard computational problems," in Proceedings of the eResearch Australasia, Gold Coast, Australia, 2010.

[5] S. J. Park, D. R. Shires, and B. J. Henz, "Coprocessor computing with FPGA and GPU," in Proceedings of the Department of Defense High Performance ComputingModernization Program:Users Group Conference—Solving the Hard Problems, pp. 366–370, Seattle,Wash, USA, 2008.

[6] M. Showerman, J. Enos, A. Pant et al., "QP: a heterogeneous multi-accelerator cluster," in Proceedings of the 10th LCI International Conference on High-Performance Cluster Computing, vol. 7800, pp. 1–8, Boulder, Colo, USA, 2009.

[7] D. B. Thomas, L. Howes, andW. Luk, "A comparison of CPUs, GPUs, FPGAs, and massively processor arrays for random number generation," in Proceedings of the 7th ACM SIGDA International Symposium on Field-Programmable Gate Arrays (FPGA '09), pp. 63–72, Monterey, Calif, USA, 2009.

[8] K. Underwood, "FPGAs vs. CPUs: trends in peak floatingpoint performance," in Proceedings of the 12th International Symposium on Field-Programmable Gate Arrays (FPGA '04), pp. 171–180, New York, NY, USA, February 2004.

[9] H. Khaled, H. M. Faheem, and R. El-Gohary, "Design and implementation of a hybrid MPI-CUDA model for the Smith-Waterman algorithm," International Journal of Data Mining and Bioinformatics, 12(3): 313-327, Inderscience, 2015.

[10] H. M. Faheem and B. König-Ries, "A New Scheduling Strategy for Solving the Motif Finding Problem on Heterogeneous Architectures," International Journal of Computer Applications, 101(5):27-31, 2014

Car's detection by Gaussian receptive field features, the eigenvalues and MLP

Sarah
BENZIANE HACHEMI
SIMPA,
University of Sciences et de la Technology of Oran,
B.P. 1505, M'Nouer, Oran, Algeria.

Abdelkader BENYETTOU
SIMPA,
University of Sciences et de la Technology of Oran,
B.P. 1505, M'Nouer, Oran, Algeria

ABSTRACT

The car's detection is one of the current research topics; it is characterized by the presence or absence of cars in a special place. Several recent works were made in this context; there's three class of car detections method, one based on the study of the move, another based on a geometrical model and the last one based on the use of the learning process.

In this paper, we implement a method for car detection based on the monocular vision by Gaussian receptive fields features; when getting features extracted, we reduce the size of the vectors samples, which were used as input for our learning step and learning procedure based on the neural network MLP.

Keywords

Monocular vision, cars detection, learning process, Gaussian receptive field's features, eigenvalues technics, neural network MLP.

1. INTRODUCTION

The cars detection is a very delicate task because of the many disturbing facts as, the presence of unfavorable condition's capture (shadow) in a part, and in another part the multi shape of the cars (colors, and size).

Many works has been made in the context of the cars detection, within them we can quote [1], where authors has proposed two steps, the first one is the HG hypothesis generation application and the second one the is HV hypothesis verification of application. In the first step, the hypothesis of generation was made based on the edge detection and this whith the aim to locate the car, while the second part was the feature extraction with Haar parameter's combined to SVM. Other technics tested in [2], where BGF was used for feature extraction for an SVM training. [3], outlined a technic based on a representation of the vehicle in the shape of a geometric model, after an energy function include information about this shape (the symmetry and the shadow of the car) was calculated. The used the genetic algorithm for reduce the space of the features, by selecting the good features. Jun Kong and his colleagues proposed a new approach [4] for extraction of the image from the background and the update of the features based on the quantification of the gray level and the mitigation of the weight in order to reduce the effect of the lighting of the light. At the end, they use a a discrimination function to separate between the two parts the object and the background in order to locate the area in move. The paper [5], talks about a new technic combining approach by appearance, a geometric model and approach in motion. In the first approach, we used the Adaboost training algorithm. A geometric model was used in order to have a good localization of the cars, and the last approach was applied to determine al the area in move. The merger of the various information obtained has been made by adopting the principle of the Bayes.

In our method, the cars detection was based on training; with Gaussian receptive fields features, eigenvalues and the MLP (neural network). Our approach was tested on the "UIUC"database's.

2. THE PROPOSED APPROACH FOR CARS DETECTION

in this method, we get a good cars detection by the training process devised in three steps : feature extraction based on the Gaussian receptive field (Eigenvalues) and the training by the neural network MLP.

A. Features extraction

Features extraction with Gaussian receptive fields was applied using five cores shown in the figure below.

Because of the variability of the size of the cars in the pictures, this phase was applied on an intrinsic space and for the three levels of size, largest to smallest.

The figure 1 presents the results obtained for the feature extraction step by the Gaussian receptive field on one of the three cores.

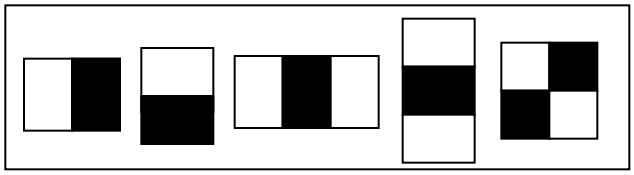


Figure 1 the cores of the Gaussian receptive fields

After applying the different kernels to extract the features of Gaussian receptive field, we organized these features into vectors.

Two types of vector were built, the haves vectors positive examples or the healthy carriers of negative examples. Of course, this operation was used separately for each intrinsic level.

For example, on an intrinsic level of size 40 * 100, one can built a vector of size 5 * 40 * 100, knowing that 5 represents the number of cores used.

Note that the vector size is very large (20,000 items). If one considers that the number of positive examples is 500 examples and the number of negative examples is 500 examples, calculates becomes very heavy. To remedy this problem, it has been applied

in this phase by the clean space called a technical method for reducing the size of these vectors; in effect a covariance matrix Q uses the relationship in (1):

$$Q = S.S^T \quad (1)$$

Knowing that S is a matrix of size 20000 * 1000.

After the SVD, we can find three types of matrices U, V, and UT

Reducing the size of the vectors was made using the first eigenvectors of the matrix U.

Tests and tests with different image sizes, ranging from 5, 10 through 100 and reaching up to 500 led us to limit the size to 100.

Finally, this technique allows us to reduce the size of our sample training vectors in size to a size of 20 000 100.

A. The neural network MLP

The last step was performed by applying an artificial neural network MLP.

A multilayer perceptron (MLP) is a feedforward artificial neural network model that maps sets of input data onto a set of appropriate outputs. A MLP consists of multiple layers of nodes in a directed graph, with each layer fully connected to the next one. Except for the input nodes, each node is a neuron (or processing element) with a nonlinear activation function. MLP utilizes a supervised learning technique called backpropagation for training the network. MLP is a modification of the standard linear perceptron and can distinguish data that are not linearly separable.

Figure 2 shows the general scheme of a multilayer perceptron neural network. If a multilayer perceptron has a linear activation function in all neurons, that is, a linear function that maps the weighted inputs to the output of each neuron, then it is easily proved with linear algebra that any number of layers can be reduced to the standard two-layer input-output model (see perceptron).

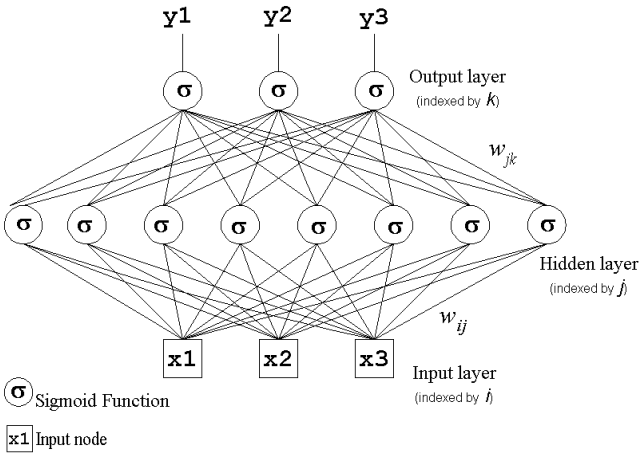


Figure 2 Multilayer Perceptron Scheme

In this phase, we applied an artificial neural network MLP with one hidden layer, further setting the number of neurons in the hidden layer was achieved by adopting (2):

$$N_c = \left| N_e - N_s \right| \quad (2)$$

Nc : Number of neurons in the hidden layer.

Ne : Number of neurons in the input layer.

Ns : Number of neurons in the output layer.

Th best results obtained was with, 98 neurons in the hidden layer. The activation function was sigmoid. Matrices weight and bias vectors were fixed randomly, the number of iterations equal to 100, and the error of the gradient was set at 0.01.

After initialization of the artificial neural network MLP, come the learning step to finally get all the reference models.

To calculate the amount of evaluation of our vehicle detection method, we applied formula (3):

$$TR = \frac{BD}{FD + BD}.100 \quad (3)$$

TR : recognition rate.

BD : good detection.

FD : bad detection.

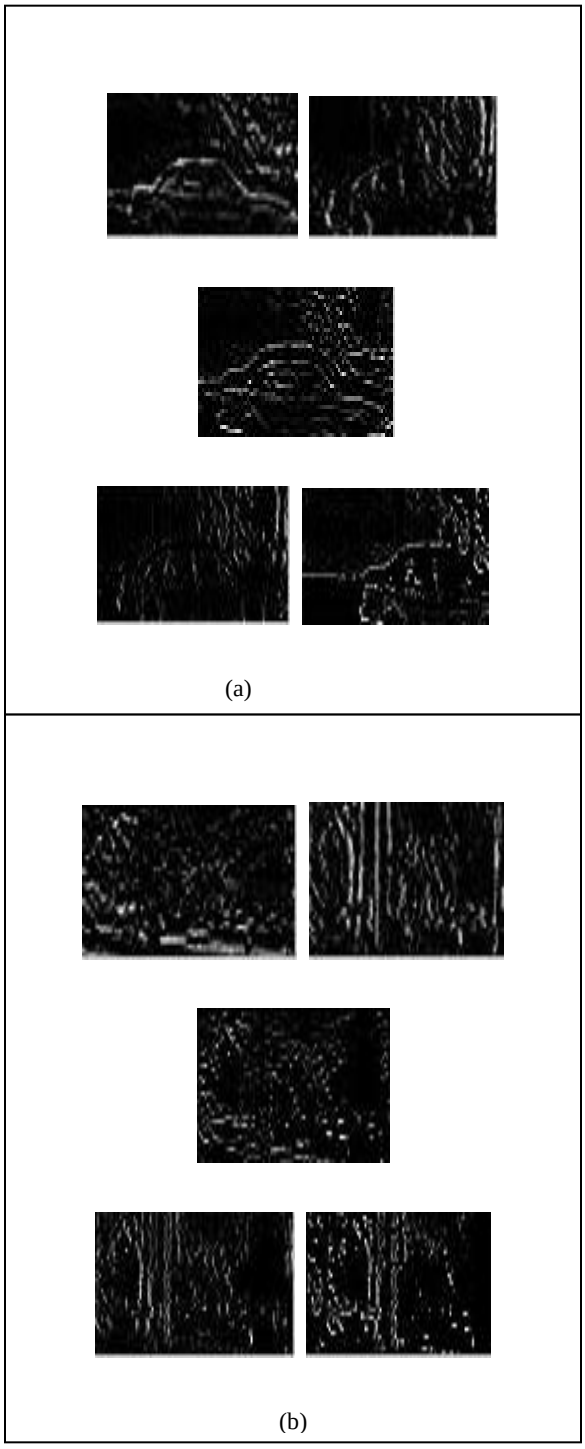
Our tests were on a database of (500 examples of vehicles and 500 examples of non-vehicles), we managed to achieve a recognition rate of 95.8%.

Following is a comparison between our features used and the characteristics of the wavelet Haar as regards the variation of the recognition rate depending on the variation of the length of the vectors.

Table 1. Comparative study between the Gaussian receptive field features type and the wavelet Haar features.

Vectors length \type des features	5	10	50	100	200	500
CRG	43.25	58.47	85.39	95.8	95.54	95.04
OH	40.05	46.21	61.25	83.47	83.28	81.04

Figure 3 displays the results obtained after features extraction by Gaussian receptive fields with the negative and the positive example.



(a)

(b)

Figure 3 The features of the type of Gaussian receptive field: (a) positive example, (b) negative example.

.

ISBN: 978-1-61804-327-6

3. The final results of our vehicle detection method

After having presented our vehicle detection method. Here is the presentation of some final results of the tests.

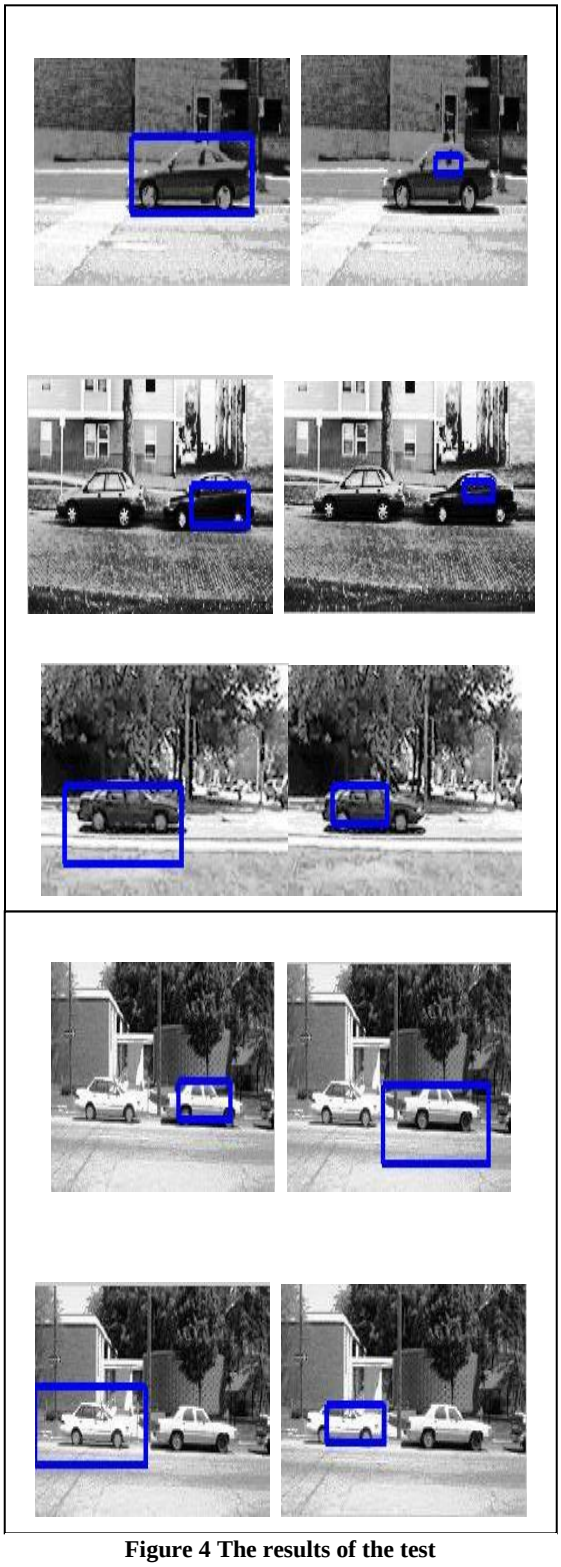


Figure 4 The results of the test

4. Conclusion et perspective

In this study we realized an effective method for the detection of vehicles, and the application of a learning process allowed us to have very good results with a very high detection rate.

As a perspective, we try to use the principle of this method in a work referred to detect vehicles in highways.

5. REFERENCES

[1] Zehang Sun, 2Ronald Miller, 1George Bebis and 2David DiMeo, “A Real-time Precrash Vehicle Detection System”, 1 Computer Vision Lab. Department of Computer Science, University of Nevada, Reno, and 2e-Technology Department, Ford Motor Company, Dearborn, U.S.A, 2002.

[2] Hong Cheng, Nanning Zheng, Chong Sun,” Boosted Gabor Features Applied to Vehicle Detection”, Institute ofArtificial Intelligence and Robotics, Xi’anJiaotong University, China, 2006.

[3] C. Hilario, J.M. Collado, J.M. Armingol and A. de la Escalera, “Pyramidal Image Analysis for Vehicle Detection, Intelligent Systems Laboratory, Department of Systems Engineering and Automation, University Carlos III of Madrid, Span, 2005.

[4] 1,2Jun Kong , 1,2Ying Zheng , 1Yinghua Lu, 2Baoxue Zhang, “A Novel Background Extraction and Updating Algorithm for Vehicle Detection and Tracking”, 1Computer School, Northeast Normal University, Changchun, Jilin Province, China, and 2Key Laboratory for Applied Statistics of MOE, China, 2007.

[5] Ying Zhu, 1Dorin Comaniciu, 1Visvanathan Ramesh, and 2Martin Pellkofer, 2Thorsten Koehler, “An Integrated Framework of Vision-based Vehicle Detection with Knowledge Fusion”, 1Siemens Corporate Research Princeton, NJ, USA, and 2Siemens VDO Automotive AG Regensburg, Germany, 2005.

[6] John Ruiz-Hernandez, James L. Crowley, et Nicolas Gourier, “ Détection de visages embarquée pour la téléphonie mobile”, Équipe PRIMA INRIA Rhône-Alpes, France, 2007.

[7] Yoshida Tatsuya, Kagesawa Masataka, and Ikeuchi Katsushi,“local-features based vehicle recongnition system using parallel vision board “, institute of industrial science, university of Tokio, Japon, 2000.

[8] 1Xin Li, 1XiaoCao Yao, 1Yi L. Murphey, 2Robert Karlsen, 2Grant Gerhart, “A Real-Time Vehicle Detection and Tracking System in Outdoor Traffic Scenes”, 1Department of Electrical and Computer Engineering, University of Michigan-Dearborn, U.S.A, and 2U. S. Army - TARDEC, Warren, U.S.A, 2004.

[9] L. Redhouane, B. Sarah, and B. Abdelkader, “Dorsal hand vein pattern feature extraction with wavelet transforms,” in Networks, Computers and Communications, The 2014 International Symposium on, 2014, pp. 1–5.

[10] S. S. Keerthi and C.-J. Lin, “Asymptotic behaviors of support vector machines with Gaussian kernel,” Neural Comput., vol. 15, no. 7, pp. 1667–1689, 2003

Authors Index

ISBN: 978-1-61804-327-6