Scientific
Conference
of Young
Researchers
2015

15th Scientific Conference of Young Researchers

May 19th, 2015
Herľany, Slovakia

Proceedings from Conference

Faculty of Electrical Engineering and Informatics
Technical University of Košice

Sponsors

**15th Scientific Conference of Young Researchers
of Faculty of Electrical Engineering and Informatics
Technical University of Košice**

Proceedings from Conference

Published: Faculty of Electrical Engineering and Informatics
Technical University of Košice
I. Edition, 334 pages, number of CD Proceedings: 50 pieces

Editors: Prof. Ing. Alena Pietriková, CSc.
Ing. Milan Nosáľ
Ing. Jana Šťastná
Ing. Emília Pietriková, PhD.

ISBN 978-80-553-2130-1

Program Committee of 15th Scientific Conference of Young Researchers of Faculty of Electrical Engineering and Informatics Technical University of Košice

General chair:	Prof. Ing. Liberios Vokorokos, PhD.
Editorial board chairman:	Prof. Ing. Alena Pietriková, CSc.
Proceeding reviewers:	Prof. Ing. Roman Cimbala, PhD. Prof. Ing. Dušan Kocur, CSc. Prof. Ing. Jozef Juhár, CSc. Prof. Ing. Daniela Perduková, PhD. Prof. Ing. Alena Pietriková, CSc. Prof. Ing. Ján Šaliga, CSc. Assoc. Prof. Ing. Ján Gamec, CSc. Assoc. Prof. Ing. Ján Genči, PhD. Assoc. Prof. Ing. Kristína Machová, PhD. Assoc. Prof. Ing. Jaroslav Porubän, PhD.

Organization Committee of 15th Scientific Conference of Young Researchers of Faculty of Electrical Engineering and Informatics Technical University of Košice

Members:	Prof. Ing. Roman Cimbala, PhD. Prof. Ing. Alena Pietriková, CSc. Assoc. Prof. Ing. Jaroslav Porubän, PhD. Ing. Ivana Olšiaková Ing. Jozef Dziak Mgr. Lukáš Hubač Ing. Tomáš Koctúr Ing. Lukáš Lisoň Ing. Peter Lukács Ing. Miroslava Muchová Ing. Milan Nosáľ Ing. Jana Šťastná Ing. Peter Talian Ing. Emília Pietriková, PhD.

Contact address: Faculty of Electrical Engineering and Informatics
Technical University of Košice
Letná 9
042 00 Košice
Slovak Republic

Foreword

Dear Colleagues,

SCYR (Scientific Conference of Young Researchers) is a Scientific Event focused on exchange of information among young scientists from Faculty of Electrical Engineering and Informatics at Technical University of Košice - series of annual events that was founded in 2000. Since 2000 the conference has been hosted by FEI TUKE with rising technical level and unique multicultural atmosphere. The Eleventh Scientific Conference of Young Researchers (SCYR 2011), conference of Graduates and Young researchers, was held on 17$^{\text{th}}$ May 2011. The primary aims of the conference, to provide a forum for dissemination of information and scientific results relating to research and development activities at the Faculty of Electrical Engineering and Informatics has been achieved. 105 participants mostly by doctoral categories were active in the conference.

Faculty of Electrical Engineering and Informatics has a long tradition of students participating in skilled labor where they have to apply their theoretical knowledge. SCYR is opportunities for doctoral and graduating students use this event to train their scientific knowledge exchange. Nevertheless, the original goal to represent a forum for the exchange of information between young scientists from academic communities on topics related to their experimental and theoretical works in the very wide spread field of electronics, telecommunication, electrotechnics, computers and informatics, cybernetics and Artificial intelligence, electric power engineering, remained unchanged.

15$^{\text{th}}$ Scientific Conference of Young Researchers at Faculty of Electrical Engineering and Informatics Technical University of Košice (SCYR 2011) was organized in a beautiful village Herľany. The Conference was opened in the name of dean prof. Ing. Liberios Vokorokos, PhD. by the vicedean of faculty, doc. Ing. Roman Cimbala, PhD. In his introductory address he noted the importance of the Conference as a forum for exchange of information and a medium for broadening the scientific horizons of its participants and stressed the scientific and practical value of investigations being carried out by young researchers.

Traditionally, the program includes two parallel sessions:

- Electrical & Electronics Engineering
- Informatics & Telecommunications

with approximately 105 technical papers dealing with research results obtained mainly in university environment. This day was filled with a lot of interesting scientific discussions among the junior researchers and graduate students, and the representatives of the Faculty of Electrical Engineering and Informatics. This Scientific Network included various research problems and education, communication between young scientists and students, between students and professors. Conference was also a platform for student exchange and a potential starting point for scientific cooperation. The results presented in papers demonstrated that the investigations being conducted by young scientists are making a valuable contribution to the fulfillment of the tasks set for science and technology at Faculty of Electrical Engineering and Informatics at Technical University of Košice.

We want to thank all participants for contributing to these proceedings with their high quality manuscripts. We hope that conference constitutes a platform for a continual dialogue among young scientists.

It is our pleasure and honor to express our gratitude to our sponsors and to all friends, colleagues and committee members who contributed with their ideas, discussions, and sedulous hard work to the success of this event. We also want to thank our session chairs for their cooperation and dedication throughout the whole conference.

Finally, we want to thank all the attendees of the conference for fruitful discussions and a pleasant stay in our event.

Liberios VOKOROKOS
dean of FEI TUKE

May 19th 2015, Herľany

Contents

Electrical & Electronics Engineering

Informatics & Telecommunications

Section:
Electrical & Electronics Engineering

A review of Artificial Intelligence Applications in Power Electronics and Motor Drives

[1]*Martin LEŠO (1st year)*
Supervisor: [2]Jaroslava ŽILKOVÁ

[1,2]Dept. of Electrical Engineering and Mechatronics, FEI TU of Košice, Slovak Republic

[1]martin.leso@tuke.sk, [2]jaroslava.zilkova@tuke.sk

***Abstract*—Artificial Intelligence (AI) techniques, particularly neural networks and fuzzy controllers, are recently having significant impact on power electronics and motor drives. This paper presents an introduction and perspective of neural network and fuzzy controllers in the intelligent control and estimation in power electronics and motors drives area.**

***Keywords*— fuzzy controllers, artificial neural networks, power electronics, motor drives**

I. Introduction

Recently AI based techniques, such as Fuzzy Logic (FL), Artificial Neural Network (ANN), Expert System (ES) and Genetic Algorithm (GA) have been applied widely in different areas and applications such as image processing, industrial process control, diagnostics, predictions, medicine, information management system etc. The main objective of AI is to apply human intelligence in a computer so that the computer could think intelligently, like a human being [1].

This paper primarily deals with the applications of FL and ANN, in the area of power electronics and motor drives. Typical applications of these methods in area of power electronics and motor drives are identification, estimation, control, design, condition monitoring and diagnosis.

The rest of this paper is organized as follows: In section II, the reader can find an overview of applications of ANN in power electronics and motor drives, section III summarises applications of fuzzy controllers in power electronics and motor drives.

II. Applications of ANN

A. Introduction

ANNs can be characterized as ‘computational models’, which are inspired by modeling networks of biological neurons in brain. Among basic features of ANN are the ability to adapt or learn, to generalize or to cluster and organize data. ANNs consist of interconnected processing units called artificial neurons (Fig. 1) that usually operate in parallel and are configured in regular collective behavior of an ANN, like a human brain [2].

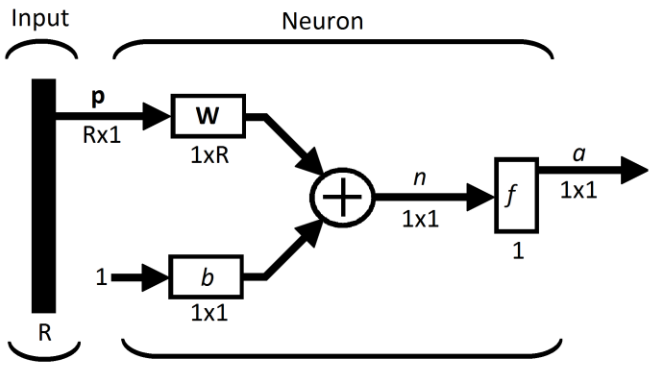


Fig. 1. Single artificial neuron [3]

The interconnection of artificial neurons results in ANN, and its objective is to emulate the function of a human brain in certain domain to solve scientific, engineering and many other real-life problems. However, the structure of biological neural network is not well-understood and therefore, many ANN models have been proposed [1], [3].

The most commonly used ANNs topologies in power electronics and motor drives are feedforward multilayer backpropagation (Fig. 2) and recurrent ANNs.

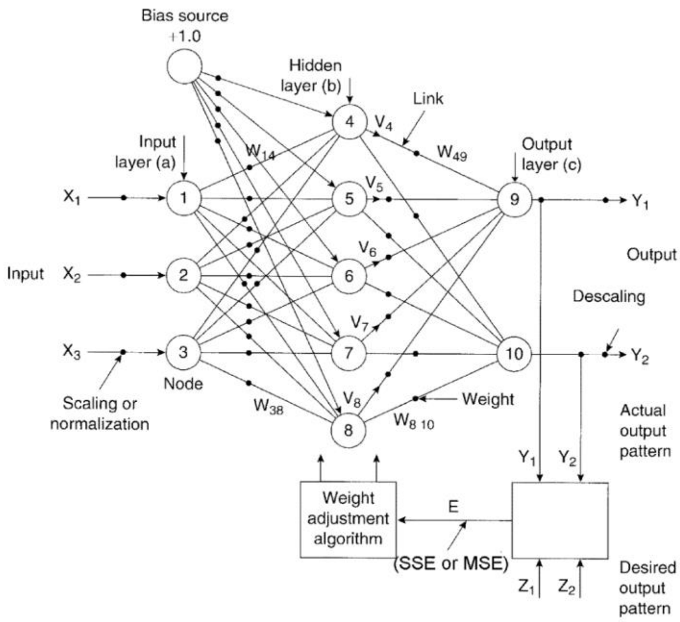


Fig. 2. Feedforward multilayer backpropagation ANN topology [1]

B. *Waveform Processing and Delayless Filtering*

In paper [4] is presented static nonlinear mapping property of feedforward neural networks for various waveform processing and delayless filtering that are applicable to voltage or current waveforms having constant frequency and variable-magnitudes. This case is mainly important for power electronic that operate on a utility system and general-purpose constant-frequency converters power supplies. In other case the ANN was designed for voltage or current waveforms with variable-frequency and variable-magnitude, which is important for the adjustable-speed AC drives area. In both cases the performance of the ANN was found excellent. The ANN has the properties of noise immunity and fault tolerance which particularly important in the distorted waveform processing applications. For different tasks there is number of layers and number of neurons in layers is different. All ANNs were trained offline using MATLAB Neural Network Toolbox with Lavenburg-Marquardt based fast backpropagation learning algorithm [1], [4].

C. *ANN based space vector modulation controllers*

Space Vector Modulation (SVM) has recently grown as a very popular pulse width modulation (PWM) method for voltage – source inverters because of its very good harmonic quality and extended linear range of operation. However, a drawback of SVM is that it requires complex online computation that usually limits its operation only up to several kilohertz of switching frequency [5].

Several researches of ANN implementation of SVM have been worked out. In [5] is presented the development of a complete ANN based space vector modulation and diagnostic controller scheme (ANN-SVM-DIAG controller) for a Voltage Source Inverter (VSI) that operates very well in undermodulation and overmodulation regions. In [6] is described a neural network based space vector pulse width modulation of three-level voltage fed inverter induction motor drives that covers both undermodulation and overmodulation regions with uniformly linear transfer characteristics. Author in the reference [7] propose the ANN for implementing the space vector PWM for a five-level three-phase diode-clamped inverter.

In all cases backpropagation-type ANNs (Fig. 3) were used. The ANN can be conveniently trained offline with the data generated by calculation of the SVM algorithm. Note that the ANN has inherent learning capability that can give improved precision by interpolation unlike the standard lookup table method [6].

D. *ANN based controllers*

The fact that a neural network can approximate any nonlinear function is very significant from the control point of view. The traditional methods for the analysis and synthesis of nonlinear controllers for specifics classes of nonlinear systems exist: linearization techniques, phase plane methods, describing functions etc. However, the ability of an ANN to represent nonlinear mappings and thus to model nonlinear systems, is the main feature which can be readily exploited in the synthesis of non-linear controllers. [14]

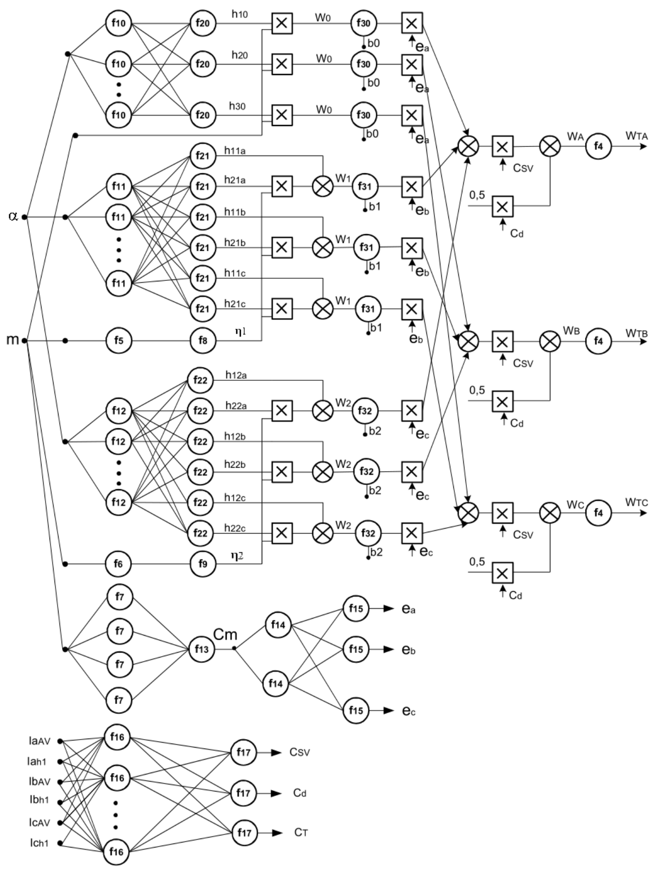


Fig. 3. Complete ANN-DIAG-SVPWM Controller for VSI [5]

In [8] is proposed ANN controller for high-voltage Full Bridge Series-Parallel Resonant (FBSPR) DC-DC converter with high switching frequency and soft switching operation to decrease the losses and optimize the efficiently of converter. The ANN controller was compared with controller based on Sequential State Machine (SSM). The controllers were tested for step changes in both output voltage in steady state and under transient conditions. It is concluded that in general ANN controller provides better characteristics than SSM controller in terms of overshoot, rise-time and setting time. This superiority was more noticeable for great step changes in load current reference voltage [9].

In paper [10] the predictive and adaptive properties of ANN for fast estimation of the compensating current to predictive controller for shunt active power filter (APF) are analysed. An integration of predictive and adaptive ANN-based controller for shunt-type (APF) was presented to improve the convergence and reduce the computational requirement. The predictive algorithm was derived from an ANN based PI controller used to regulate the DC-link voltage in APF. An experimental prototype was produced in the laboratory. The results from experiments match well with the simulation, confirming the usefulness of the proposed technique.

In [11], a practical design problem of the voltage source converter control scheme for industrial micro grid is issued and an ANN based PI controller is proposed with backpropagation weight change algorithm to solve the problem (Fig. 4). The advantage of the ANN based Pi regulator in heavy load with large load change is proved in the paper. Because the PI parameters are adapted with the online tuning mechanism of the ANN algorithm in different phase of the regulation proves, the response of the controller is faster and simultaneously the oscillation are weaken.

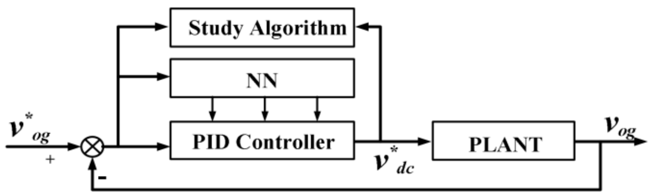


Fig. 4. ANN based control scheme [11]

E. *ANN based parameter identification and estimation*

Mathematical models of the system elements are fundamental to system studies, such as system design, operation and control. Recurrent ANN has also been used for system identification and parameter estimation.

In [12] is presented an adaptive discrete-time grid voltage sensorless interfacing scheme for grid-connected distributed generation inventers, based on ANN identification and deadbeat current regulation. Knowledge of the interfacing parameters R, L, and the grid voltage is required in order to implement a robust voltage-sensorless interfacing scheme. The estimation unit has been designed with low computational demand to estimate in real-time, the interfacing parameters and current grid voltage vector simultaneously. A reliable solution to the nonlinear estimation problem is presented by combining an ANN interfacing parameter identifier and ANN grid-voltage estimator.

Online stator and rotor resistance estimation scheme using ANN for vector controlled speed sensorless induction motor drive has been discussed in the literature [13]. The error between the rotor flux linkages based on an ANN model and a voltage model is back propagated to adjust the weights of ANN model for the rotor resistance estimation. For stator resistance estimation, the error between the measured stator current and the estimated stator current using ANN is back propagated to adjust the weights of ANN. Both the rotor and stator resistance variations (Fig. 5) can be successfully estimated using the adaptation capabilities ANN. The implementation of these neural network techniques required only a small increase of the computation times. The feasibility and validity of the proposed identification has been proved by the experimental results.

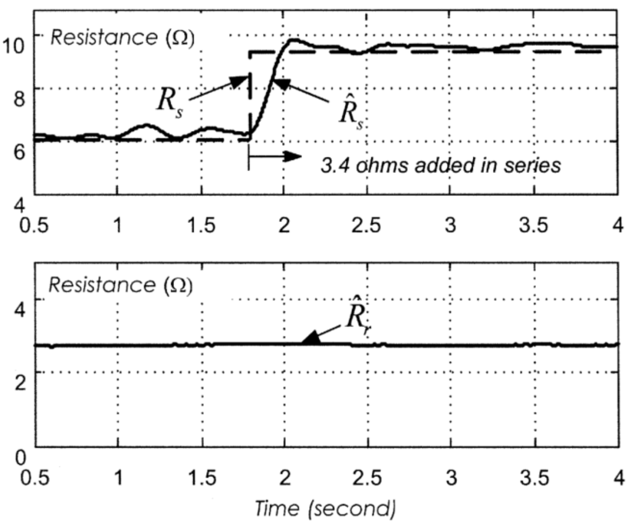


Fig 5. Estimated stator resistance in experiment [13]

III. Applications of fuzzy logic system

A. *Introduction*

Fuzzy Logic Systems (FLSs) are universal function approximators. The heart of a FLS is a linguistic rule-base, which can be interpreted as the rules of a single 'overall' expert, or as rules of 'subexperts' and there is inference mechanism, where rules are considered in appropriate manner to generate outputs. These rules may directly originate from experts, but if the experts are not available, they can also be obtained by the appropriate processing (e.g. clustering) of available input/output data. [14]

B. *Fuzzy logic based controllers*

Fuzzy control systems have demonstrated their enhanced performance in different areas. Although initially fuzzy systems were associated only to the AI that has refrained to the development of theoretical fuzzy systems, in 1985 Japanese researchers Seiji Yasunobu and Soji Miyamoto demonstrated the superiority of fuzzy control systems for the Sendai railway. From that moment on, many applications have taken the advantage of the inherent potential offered by fuzzy controllers. Some notable works on the applications of fuzzy controllers are inverted pendulum balancing by Takeshi Yamakawa, improved vacuum cleaners by Panasonic Corporation, stable CCD development by Canon Inc., energy efficient air conditioners by Mitsubishi Companies and fuel efficient automatic space docking by NASA. [15]

Recently, the fuzzy logic controller has generated a great deal of interest in various applications and has been introduced in the power electronics field [16]-[20]. The advantages of fuzzy logic controllers over the conventional PI controllers consist in fact that they do not need an accurate mathematical model, they can work with imprecise inputs, can handle nonlinearity, and may be more robust than the conventional PI controller [16]. However, several structures of fuzzy logic controllers (FLC) have been proposed.

In papers [16]-[18] are proposed intelligent hybrid PID-FLC (Fig. 6), which automatically varies the coefficients Kp, Ki and Kd. The proposed hybrid controllers are quite suitably applicable to nonlinear objects control, which are difficult and delicate to be properly modeled with constant parameters. In [16] is a PID-FLC used to improve the dynamic response of the output active power and reactive power of a grid connected photovoltaic inverter. The comparison of simulation results, obtained separately from a PID-controller and a PID-Fuzzy controller embedded in a PV inverter shows the advantages of the proposed algorithm in terms of time response, overshooting and stability error. It also shows that the proposed controller maintains its response better than the PID controller in many aspects, regardless of changes of grid parameters during operation.

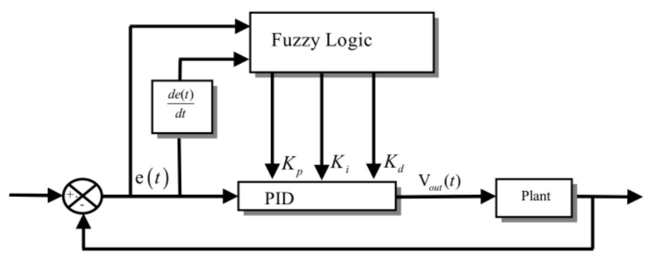


Fig 6. The structure of the PID-FLC [17]

In [19], digital PID-type and fuzzy type controllers are compared for application to the buck and boost dc–dc converters. Experimental results showed that fast transient response and stable steady-state responses could be achieved for both the buck and the boost converters using fuzzy controllers. For the buck converter, comparable results were obtained using PID and PI controllers and fuzzy controller (Fig. 7). In the case of boost converter control, there were some conditions in which the fuzzy control approach yielded superior performance, but there were also some tradeoffs. During startup, the linear and fuzzy control yielded similar responses for most input voltages. In one case, the FLC yielded less overshoot during the startup transient. Under load changes, the FLC resulted in less oscillation, but the peak variation was higher in few cases (decreasing load). In most cases, the fuzzy controller also yielded superior settling time, particularly under load increases.

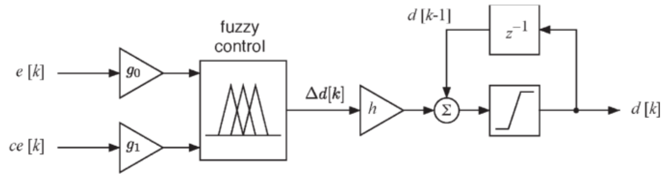


Fig 7. The structure of the Fuzzy logic controller [19]

To improve the performance of the boost converter the Neuro-Fuzzy Controller is proposed in [20], which gives small overshoots and has much superior performance compared to the local PI controllers. The intelligent controller behaves effectively like an adaptive local tuned controller designed for each operating point and gives an improved performance compared to the conventional PI controller.

IV. Conclusion

This paper presents an overview of AI methods, especially FLC and ANN in area of power electronics and motor drives that includes several application examples.

Despite of the fact that the ANN basic technology has advanced tremendously in recent years, currently, industrial applications of neural networks in power electronics and motor drives area appear very seldom. In other hand, FL based controllers can be found in many real-life applications and have very good perspective for the future.

One of the challenges in AI is also their implementation by microcontrollers. At present, the most ANN and FLC implementations are based on slow and sequentially has been executed by digital signal processors (DSPs). Of course, the speed of DSP has been improving dramatically in recent years. Concurrently, there has been tremendous advancement of Very Large System Integration (VLSI) technology, and the ASIC chips including FPGAs have been applied.

The main goal of my PhD. project is to apply methods of AI in area of power electronics and motor drives, where the classical methods are not effective and better results can be achieved using the AI methods.

Acknowledgment

This work was supported by scientific grant agency of the Ministry of Education of the Slovak Republic and of Slovak Academy of Sciences, project VEGA 1/0464/15.

References

[1] Bose, B.K., "Neural Network Applications in Power Electronics and Motor Drives—An Introduction and Perspective," *IEEE Transactions on Industrial Electronics*, vol.54, no.1, pp.14,33, Feb. 2007

[2] Jeon, JongKoo, *Fuzzy and neural network models for analyses of piles* (Book style). Raleigh, NC, 2007, pp. 10

[3] MathWorks Neural Network Toolbox User's Guide, 2014

[4] Jin Zhao; Bose, B.K., "Neural-network-based waveform Processing and Delayless filtering in power electronics and AC drives," *IEEE Transactions on Industrial Electronics*, vol.51, no.5, pp.981,991, Oct. 2004

[5] Phan Quoc Dzung; Le Minh Phuong; Pham Quang Vinh; Nguyen Van Nho; Dao Minh Hien, "The development of artificial neural network space vector PWM and diagnostic controller for voltage source inverter," *Power India Conference, 2006*

[6] Cong Wang; Bose, B.K.; Oleschuk, V.; Mondal, S.; Pinto, J.O.P., "Neural-network-based space-vector PWM of a three-level inverter covering overmodulation region and performance evaluation on induction motor drive," *Industrial Electronics Society, 2003. IECON '03. The 29th Annual Conference of the IEEE* , vol.1, pp.1,6 vol.1, 2-6 Nov. 2003

[7] Saqib, M.A.; Kashif, S.A.R., "Artificial neural network based space vector PWM for a five-level diode-clamped inverter," *20th Australasian Universities Power Engineering Conference (AUPEC)*, pp.1-6, 5-8 Dec. 2010

[8] Imanieh, M.J.M.; Malekjamshidi, Z., "Design, simulation and implementation of a Full Bridge Series-Parallel Resonant DC-DC converter using ANN controller," *International Siberian Conference on Control and Communications (SIBCON)* , pp.97-103, 15-16 Sept. 2011

[9] Malekjamshidi, Z.; Jafari, M.; Imanieh, M., "Implementation of a Full Bridge Series-Parallel Resonant DC-DC converter using ANN and SSM controllers," *International Siberian Conference on Control and Communications (SIBCON)*, pp.120-127, 15-16 Sept. 2011

[10] Bhattacharya, A.; Chakraborty, C., "A Shunt Active Power Filter With Enhanced Performance Using ANN-Based Predictive and Adaptive Controllers," *IEEE Transactions on Industrial Electronics*, vol.58, no.2, pp.421-428, Feb. 2011

[11] Jinbang Xu; Zhizhuo Wu; Xuan Yang; Jie Ye; Anwen Shen, "ANN-based Control Method Implemented in a Voltage Source Converter for Industrial Micro-grid," *Sixth International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA)*, pp.140-145, 2011

[12] Mohamed, Y.A.-R.; El-Saadany, E.F., "Adaptive Discrete-Time Grid-Voltage Sensorless Interfacing Scheme for Grid-Connected DG-Inverters Based on Neural-Network Identification and Deadbeat Current Regulation," *IEEE Transactions on Power Electronics*, , vol.23, no.1, pp.308-321, Jan. 2008

[13] Karanayil, B.; Rahman, M.F.; Grantham, C., "Online Stator and Rotor Resistance Estimation Scheme Using Artificial Neural Networks for Vector Controlled Speed Sensorless Induction Motor Drive," *IEEE Transactions on Industrial Electronics*, vol.54, no.1, pp.167-176, Feb. 2007

[14] Vas, P., *Artificial-intelligence-based electrical machines and drives* (Book style). New York: Oxford University Press, 1999, pp. 39, 155

[15] Sohail Iqbal. Preface in: *Fuzzy Controllers – Recent Advances in Theory and Applications*. Rijeka, Croatia: InTech, 2012

[16] Mortazavi, S.S.; Kianinezhad, R.; Ghasemi, A., "Auto tuned robust active power filter for power quality improvement under fast load variation," *Third International Conference on Electric Utility Deregulation and Restructuring and Power Technologies, 2008. DRPT 2008.*, pp.1646-1651, 6-9 April 2008

[17] Louzazni, M.; Aroudam, E., "Control and stabilization of three-phase grid connected photovoltaics using PID-Fuzzy logic," *IEEE International Conference on Intelligent Energy and Power Systems (IEPS)*, pp.279-284, 2-6 June 2014

[18] Nguyen Gia Minh Thao; Mai Tuan Dat; Tran Cong Binh; Nguyen Huu Phuc, "PID-fuzzy logic hybrid controller for grid-connected photovoltaic inverters," *International Forum on Strategic Technology (IFOST)*, pp.140-144, 13-15 Oct. 2010

[19] Liping Guo; Hung, J.Y.; Nelms, R.M., "Evaluation of DSP-Based PID and Fuzzy Controllers for DC–DC Converters," *IEEE Transactions on Industrial Electronics*, , vol.56, no.6, pp.2237-2248, June 2009

[20] Jawhar S, J.; Marimuthu, N.S.; Singh N, A., "A Neuro-Fuzzy Controller for a Non Linear Power Electronic Boost Converter," *International Conference on Information and Automation*, pp.394-397, 15-17 Dec. 2006

Acoustic Feature Examination for Emotional Speaker Verification in Slovak Language

[1]*Lenka MACKOVÁ (3rd year)*
Supervisor: [2]Dr.h.c. prof. Ing. Anton ČIŽMÁR, CSc.

[1,2]Dept. of Electronics and Multimedia Communications, FEI TU of Košice, Slovak Republic

[1]lenka.mackova@tuke.sk, [2]anton.cizmar@tuke.sk

***Abstract*—This paper is dedicated to the experimental investigation of the best acoustic features in gender-dependent text-independent speaker verification from emotional speech. In an experimental setup different acoustic feature sets (MFCC, LPC, LPCC and PLP) where compared on basis i-vector representation. In process of system evaluation emotional recordings from newly created Slovak emotional database –SUS- were compared according the rules of Mahalanobis distance metric with use of EFR (Eigen Factor Radial) normalization for intersession compensation. The results of the experiment showed the MFCC representation as the best fitted for speaker verification from Slovak emotional speech with recognition rate higher than 80%.**

***Keywords*—emotions, i-vectors, speaker verification, total variability.**

I. Introduction

Emotions are inseparable part of everyday human interaction and can be found not only in speech but gestures, facial expression and "body language" as well. To the human communication emotions fill very important implicit information according which humans are able to identify the meaning of the explicit spoken message [1]. In the human-computer interaction (HCI) the role of emotions became essential as well since it may enhance the interaction between user and the machine.

There is quite a range of HCI applications where emotional perspective is important, for example in systems which can alert a user to signs of emotions that call for attention, in automatic tutoring, in special medical systems (diagnostic of neurological disorders and diseases) or in forensics (speaker verification, etc.). Speaker recognition/verification may be relevant to applications with speaker`s voice access control and since everyday speech is unlikely to be expressed in neutral emotional state only emotion recognition became very important even in this field of science.

II. Initial Status

Recognition of emotion from speech is not a new idea. In the literature there is many experiments dedicated to this field of expertise. Nowadays emotional clues present in the speech may be observed in various features extracted from excitation source, vocal tract system and prosodic components of speech and different classification method emotion recognition systems may be used [2].

After the study of theoretical background about emotional recognition from the speech signal we have specified the following dissertation thesis:

1. The application of new methods for the selection of emotional characteristics from expressive speech, which would include information on:
   a. shape of the excitation impulse.
   b. F0 and its local and global changes.
2. The application of new methods of feature vectors selection and reduction.
3. Analysis of emotional information in speech signal in process of speaker recognition suitable for forensics applications.

III. Solved Tasks in The Previous Year

Tasks which are summarized in the following section were solved in last year of postgraduate study based on dissertation thesis.

A. Creation of Slovak Emotional Database

For purposes of emotional speaker verification several emotional databases in foreign languages were available. To meet our requirements for training and testing the speaker model with satisfactory volume of emotional dataset none of those corpuses were suitable enough. Therefore we decided to create our own native emotional database which is presented in this paper.

Emotional audio sessions from which utterances of different subjects were extracted were captured from free FTA DVB-T transmission using PCI digital capture card.

All of the emotional utterances were manually labeled. The emotional evaluation was provided on the whole sentences or segments of the sentences so that the explicit meaning of the utterance was captured in recording. Evaluated sentences/segments were then labeled with capital letters representing specified emotions (e.g. CU for curiosity, N for neutral, etc.) and cut into separate emotional utterances of individual speakers with duration from 5 to 6 seconds. Emotional range of recorded utterances covers mostly emotions of neutral state and curiosity as well as negative emotions (anger, aggressiveness, etc.).

The SUS database which up to this date consists of approximately 2000 emotional utterances of 7 speakers is classified as an induced database. However the number of speakers of the database is still quite small and dispersion of emotional utterances is not balanced.

B. Extraction of acoustics features from Slovak emotional utterances

Based on our previous research of speaker verification [5] with use of emotional utterances from Emo DB (Berlin Emotional Database) [6] we focused on performing the speaker verification with Slovak emotional speech.

For training and testing speaker models short utterances (5-6 s.) of three male speakers (spk1, spk2, spk3) from SUS database in emotions of neutral and curiosity were chosen. Those emotions were the most common to extract from the SUS sessions.

In process of feature extraction from emotional utterances different number of different acoustic features were used, namely MFCC (Mel-frequency Cepstral Coefficients), LPC (Linear Prediction Coefficients), LPCC (Linear Prediction Cepstral Coefficients) and PLP (Perceptual Linear Prediction) coefficients. Firstly 19 then 22 coefficients of mentioned types per frame were extracted at frame rate of 10 ms with 25 ms Hamming window. To the specified number of basic static parameters the log-energy and additional time derivates were added for enhancement of the system. The recognition performance of the system based on use of individual acoustic features was compared and the result in Table I, Table II and Table III showed the MFCCs as the best acoustic representation for emotional speaker verification which may be related to the fact that MFCCs are the best features to represent perceptual aspect of short-term speech spectrum.

C. Application of i-vectors method on extracted feature vectors

To reduce the dimension of extracted features vectors from emotional utterances the i-vector principle which emerges from JFA [4] was used. Extracted i-vectors were then normalized by EFR (Eigen Factor Radial) normalization technique. To find out the best performance of the recognition system the dimension of the i-vectors were set to 10, 15 and 20 respectively and results were compared. The best results were obtained when the dimension of i-vectors was set to number of 10.

D. Approaches of speaker model training

In process of training the speaker model two different approaches were applied. In the first approach (mixed model) the speaker model per speaker was trained using all of the emotions of the speaker from training set. In the non-mixed approach two speaker models per speaker were created, each trained with only one of the emotions. Created speaker models where then compared to the testing set of utterances of the speakers in testing phase.

The UBM (Universal Background Model) was trained with 32, 64, 128 and 256 Gaussians and the results of verification performance were compared. The best results were obtained mostly using 128 Gaussians except when extracting LPCCs form emotional utterances as showed in Table III.

E. Speaker verifications scoring methods

Based on the experiment carried in [7] to score the recognition system Mahalanobis distance metric was used. The results are showed in Table I, Table II and Table III. In comparison with state-of-the-art systems our results are satisfying, in [8] with use of the MFCCs and GMM classification recognition accuracy was 85%, in [9] 81.5% (MFCC, HMM).

TABLE I SPEAKER RECOGNITION RATE (%) FOR MIXED MODEL WITH 128GMM/UBM TRAINING (I-VECTOR DIMENSION = 10)

	22 MFCC			19 LPC			19 PLP		
	spk1	**spk2**	**spk3**	**spk1**	**spk2**	**spk3**	**spk1**	**spk2**	**spk3**
spk1	**80**	8	12	**73**	15	12	**69**	11	20
spk2	0	**82**	18	2	**73**	25	6	**70**	24
spk3	14	7	**79**	8	22	**70**	25	7	**68**

TABLE II SPEAKER RECOGNITION RATE (%) FOR NON-MIXED MODEL WITH 128GMM/UBM TRAINING (I-VECTOR DIMENSION = 10)

	22 MFCC			19 LPC			19 PLP		
	spk1	**spk2**	**spk3**	**spk1**	**spk2**	**spk3**	**spk1**	**spk2**	**spk3**
spk1	**83**	2	15	**73**	15	12	**69**	11	20
spk2	5	**85**	10	4	**71**	25	6	**70**	24
spk3	12	8	**80**	8	22	**70**	25	7	**68**

TABLE III SPEAKER RECOGNITION RATE (%) FOR 22 LPCC WITH 64 GMM/UBM TRAINING (I-VECTOR DIMENSION = 10)

	Mixed model		
	spk1	**spk2**	**spk3**
spk1	**67**	11	22
spk2	6	**67**	27
spk3	32	3	**65**

	Non-mixed model		
	spk1	**spk2**	**spk3**
spk1	**67**	12	21
spk2	5	**67**	28
spk3	30	5	**65**

IV. FUTURE WORK

Or future work will be dedicated to the further enlargement of the SUS database. We also plan to focus on testing the gender-dependent text-independent speaker verification system on female emotional dataset of the SUS corpus.

ACKNOWLEDGMENT

This publication is the result of the Competence Center for Innovation Knowledge Technology of production systems in industry and services (ITMS project code 26220220155) (50%) and VEGA 1/0075/15 (50%).

REFERENCES

[1] D. Ververidis, C. Kotropoulos, "Emotional speech recognition: Resources, features, and methods," Speech communication 48.9, 2006, pp.1162–1181.

[2] L. Macková, A. Čižmár, "Emotion recognition," *RTT*, 2008. – Bratislava, STU, 2008, pp. 4.

[3] N. Dehak, R. Dehak, P. Kenny, N. Brummer, P. Ouellet, P. Dumouchel, "Support vector machines versus fast scoring in the low-dimensional total variability space for speaker verification," Interspeech, Vol. 9, 2009, pp. 1559-1562.

[4] P. Kenny, G. Bouilianne, P. Ouellet, P. Dumouchel, "Joint factor analysis versus eigenchannels in speaker recognition," *Audio, Speech, and Language Processing*, IEEE Transactions, 2007, pp. 1435-1447.

[5] L. Macková, A. Čižmár, "Speaker recognition from emotional speech using I-vector approach," *Journal of Electrical and Electronics Engineering*. Vol. 7, no. 1(2014), 2014, pp. 93-96.

[6] F. Burkhardt, A. Paesche, M. Rolfes, W. Sedlmeier, B. Weiss, "A database of German emotional speech," *Interspeech*, Vol. 5, 2005, pp. 1517-1520.

[7] L. Macková, A. Čižmár, "Emotional speaker verification based on i-vectors," *Cognitive Infocommunications (CogInfoCom)*, 2014 5th IEEE Conference on. IEEE, 2014. p. 533-536.

[8] S. G. Koolagudi, K. Sharma., K.S. Rao, "Speaker Recognition in Emotional Environment," *Eco-friendly Computing and Communication Systems*, Springer Berlin Heidelberg, 2012, pp. 117-124.

[9] I.M.A. Shahin,"Employing both gender and emotion cues to enhance speaker identification performance in emotional talking environments," *International Journal of Speech Technology*, 16.3, 2013, pp.341-351.

Application of exoskeletons in the process of patients' rehabilitation

[1]Ing. Peter BALOG (1st year)
Supervisor: [2]prof. Ing. Stanislav SLOSARČÍK, CSc.

[1,2]Dept. of Technologies in Electronics, FEI TU of Košice, Slovak Republic

[1]peter.balog.2@tuke.sk, [2]stanislav.slosarcik@tuke.sk

***Abstract*— The paper analysis exoskeletons as auxiliary equipment used in the military and medicine. More attention is devoted to rehabilitation exoskeletons, which are promising devices in the process of patients' rehabilitation. In this paper exoskeletons are divided into three categories according to the extent of injury and areas of use in the rehabilitation process. The purpose of these systems is in the case of patients with residual force the practicing of motoric abilities and recognition of the intended movement. In the case of the para and tetraplegic patients, without residual force exoskeletons provides except practicing the verticalisation and allow the transfer to the different distances, to avoid degradation due to immobilization.**

***Keywords*—exoskeleton, rehabilitation, paraplegia**

I. Introduction

Harms of spine and spinal cord are one of the most serious injuries. In general, the extent of functional disabilities invalidates the patient. Some are dependent on the use of the wheelchair, some need another person to help and some are permanently bedridden. Statistics indicate that spinal cord and spine injuries occurs in 20 to 40 person per million. It is a typical accident of healthy young men, and most of the victims are in working age from 15 to 40 years [1]. The result of spinal cord injury is paralysis of limbs and trunk of varying range - the development of paraplegia (complete paralysis of both lower limbs) or tetraplegia (complete paralysis of all four limbs).

The overall aim of rehabilitation is full or at least partial return of the lost function. Rehabilitation is performed by rehabilitation devices. Lower limb exoskeletons are one of possible rehabilitation devices. These devices include most of the lower limb joints and have anthropomorphic (construction similar to construction of the human body) character. These devices are adapted to the human body and work parallel with them.

In the past, the robotic exoskeletons have been used only to replace manpower. Today, the trend is oriented to development of personally-oriented robotic devices where human operator brings the intelligence and robotic device brings the power. Interaction between the robot and the operator has been solved by the integration of human and robotic devices into one system. This integration has brought an opportunity to create support technologies in the field of biomedical, industrial, military and many others [2]. According to areas of application the exoskeletons and active orthoses can be divided to two large groups. The first large group is the performance enhancing exoskeletons that have been developed for military purposes. After successes with exoskeletons increasing performance engineers began to develop exoskeletons for medicine. These systems help patients with paralyzed limbs. Similarly was created another group called rehabilitation exoskeletons.

II. Performance enhancing exoskeletons

The aim of performance-enhancing exoskeletons is to increase the strength, robustness and resilience of the human body and thus facilitate the operator's hard work. These exoskeletons is not designed for rehabilitation purposes, but can be used by healthy operators. Performance-enhancing exoskeletons are developed and used for military purposes.

The most outstanding performance-enhancing exoskeleton is exoskeleton BLEEX (Berkeley Lower Extremity EXoskeleton). BLEEX (Fig.1) is a robotic system which is worn by operator and provides the ability to perform difficult tasks with minimal effort in any type of terrain. This capability is achieved through four important functions:

- reliable control electronics,
- high-power, compact power supply sources,
- special communication protocols,
- simplicity and low power consumption of all architecture.

Fig. 1 Exoskeleton BLEEX [3]

The system BLEEX has anthropomorphic architecture, meaning that the proposed system will attempt to follow movement of human lower limbs. However, one of the greatest problems is that human joints cannot be easily duplicated, so developers are focused to achieving of maximum degrees of freedom, and to development of exoskeleton which can be easily adapted to the human body.

BLEEX exoskeleton consists of two powered legs, power supply and backpack, which serves as a frame to accommodate the payload. BLEEX do not require active control management, but emulate the movement of the operator as a pair of artificial legs instead.

Another important exoskeleton in the field of military are exoskeletons XOS and XOS2. Exoskeletons XOS and XOS2 were developed by Sarcos Research Corporation. XOS2 (Fig. 2) is an improved version of XOS exoskeleton, which weighs only 68 kg, which is 10% less than the weight of the first generation. XOS2 provides an amplification of about 17 times of operator’s strength with 50 % less energy consumption in compare with first generation, which also provided an amplification of 6 times only [4].

Fig. 2 Exoskeleton XOS2 [5]

III. REHABILITATION EXOSKELETONS

The trend of the current design and development in rehabilitation robotics is a device that would be able to recognize the motoric and intended movement of the patient and give him feedback about it. These devices should assist to the patient only when it is necessary, and should have the same effect as a qualified therapist. In the field of exoskeletons exists two elementary techniques of rehabilitation. The first uses controlled movements with the patient’s limbs and the second uses functional electrical stimulation (FES) of limb muscles. An important element in most of lower limb rehabilitation devices based on the first mentioned technique is a mechanical construction which supports the patient. The support system helps the patient to transfer the weight of his own body, what make patients movement easier. Furthermore, this type of systems also ensures the stability.

Machine support at the rehabilitation process of para and tetraplegic patients is divided into three groups according to the position in which are rehabilitation performed. The first is rehabilitation in a horizontal position, the second is rehabilitation in a vertical position and the last is also rehabilitation in a vertical position, but there are devices which allow the patient to move in various distances.

A. Rehabilitation in a horizontal position

For rehabilitation of musculoskeletal system in the horizontal position are most commonly used motor splints. Motor splints provide passive motion therapy - CPM (Continuous Passive Motion) therapy. These devices are designed for the treatment of knee, hip, shoulder, elbow and ankle joint. Furthermore, there are motor plates for rehabilitation of wrist and fingers, including the thumb.

Medical rehabilitation by motor splints is used mainly to prevent injuries caused by immobilization, for the early resumption of painless movement of joints and accelerate the process of therapy with a good functional outcome.

MOTOmed is a rehabilitation device also designed for paraplegic and tetraplegic patients whose are immobile, or for patients with a certain level of residual strength (Fig. 3).

Fig. 3 MOTOmed rehabilitation device [10]

In the case of immobile patients the rehabilitation consists of passive practicing of musculoskeletal system, while lower respectively upper limbs must be fixed with orthoses. The movement is generated by an electromotor. Patients with a certain level of residual strength rehabilitation are carried out by active exercising of functional muscle groups, while it is possible to regulate the load level.

B. Rehabilitation in a vertical position

To avoid degradation due to immobilization the patients should be verticalize as soon as possible. Verticalization is not based on the sitting position on the bed, but in a vertical standing position. Tetraplegic patient is verticalized on the table or by trestles. Paraplegic patient is verticalized by orthoses, or with the help of functional electrical stimulation intended for standing and walking. To activate different muscle groups are used rehabilitation devices for making active and passive movements of the patient.

Currently, the solution of this problem consist of isolated rehabilitation models, which supports movement of large joints of the lower limbs using motors, or allows the swing gait of the lower limbs using special orthoses. Rehabilitation device in the vertical plane may be divided into cBWS - cable Body Weight Support and sBWS - structural Body Weight Support.

In the system cBWS the patient is hanged in the belts, while the system sBWS provides the patient's stability through arms which supports the patient in the waist and back.

In the rehabilitation process are frequently used the trainers systems with motorized belt. These systems consist of supporting frame, robotic orthosis which perform the desired movement and a running belt. An example of this type of rehabilitation device - Lokomat is commercially available. It is the result of the team work at the Hocoma company in Zurich. It is a cBWS system with controlled motion of the hips and knees in the sagittal plane (plane parallel to the medial plane which passes between the legs). The system includes force sensors, which provide monitoring of movement.

An example of a sBWS rehabilitation system is ALEX (Active Leg EXoskeleton). ALEX is a motorized orthosis and it is based on an active orthosis that allows performing of controlled movements of the hips and knees. ALEX was created by modification of a passive orthosis called GBO (Gravity Balancing leg Orthosis) [12].

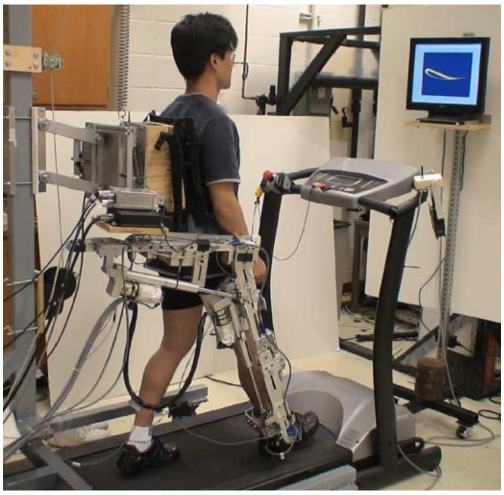


Fig. 4 Exoskeleton ALEX [12]

Rehabilitation device ALEX was developed to assist the patients during rehabilitation when it is necessary. The main advantage of this device is that patients will actively participate in the process of rehabilitation in comparison with other rehabilitation systems.

C. Verticalization devices allowing patients to move to different distances

The aim of this method is to simulate the rehabilitation of lost motoric function and inducing of stepping automatism of the lower limbs in order to transfer the patient to different distances.

The purpose of these verticalization devices is to provide sufficient mechanical and control flexibility for the implementation of everyday movements, which are summarized in ADL (Activities of Daily Living). ADL is a summary of general human activities, which has been defined for the medicine purposes in order to shorten the rehabilitation process and to define the healthy human lifestyle [2]. This includes walking, walking up and down the stairs, sitting down and getting up.

The first presented device from this group is WalkTrainer. It is based on construction with wheels which provide hanging apparatus for supporting a patient. WalkTrainer belongs to the group of cBWS systems, which includes orthoses (pelvic and limb) and closed circuit for functional electrical stimulation. Pelvic orthoses have 6 degrees of freedom and at active walking uses control of ankle joint. Limb orthoses includes sensors for sensing articulated angles (hip, knee and ankle) and interaction forces between the patient and the device [2].

In 2008, the company Argo Medical Technologies presented exoskeleton ReWalk. ReWalk controls hip flexion (reducing the angle between bones), extension (the opposite of flexion) and knee flexion. During walking the patient uses crutches to keep the balance.

Fig. 5 Exoskeleton ReWalk [7]

Using sensors placed on the chest, the device detects the angle formed between the legs and hips and with this information adequately controls the position of the legs. Use of ReWalk is limited to patients who have appropriate height and weight.

A similar solution has been revealed in 2011 by company Ekso Bionics called Ekso GT. Ekso GT is a rehabilitation exoskeleton that allows paralyzed patients to stand up from the wheelchair and walk. Walking is controlled by control electronics, with actuators consisted of battery-powered electric motors, which are replacing the neuromuscular function of the patient. Exoskeleton provides functional rehabilitation under the supervision of a physiotherapist. EKSO GT is equipped with a stepping generator, whereby the patient can activate the selected step.

Exoskeleton provides four control modes:

- "FirstStep" - Physiotherapist controls the patients steps by pressing buttons, the patient can practice standing up and walking with walker or crutches.
- "ActiveStep" - The patient uses crutches or walker, and take control of actuating their steps via buttons on the crutches or walker.
- "ProStep" - The patient moves his hips to one side or forward and thereby initiates the beginning of the next step without button.
- "ProStep Plus" - The next step is initiated by moving the legs forward [14].

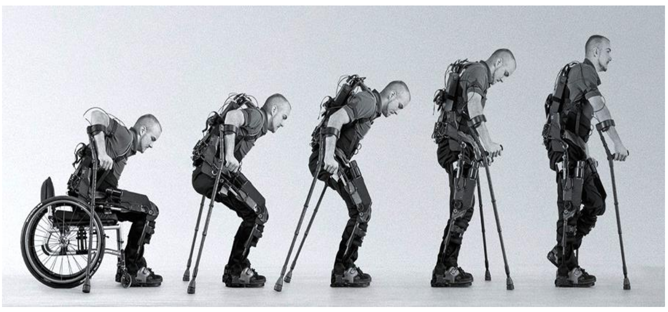


Fig. 6 Exoskeleton Ekso GT [14]

The next and simultaneously most famous rehabilitation exoskeleton is the exoskeleton HAL (Hybrid Assistive Limb). Exoskeleton HAL has been developed by professor Yoshiyuki Sankai for patients with gait disturbances. It was designed to support and enhance the physical abilities of its users, especially for people with motoric function disabilities. There are two basic versions of the system: HAL 3, which provides support only for the lower limbs and HAL 5, which is a whole-body exoskeleton.

Fig. 7 Exoskeleton HAL 3 [6]

HAL is designed for physically disordered persons and seniors to help with their daily activities, but can also be used to support workers who make physically difficult work, such as emergency units. HAL is used especially for injured patients in hospitals and can be modified for long-term rehabilitation. The latest version of the exoskeleton can also be used only for one side of the body in case of one side straining or injury.

To control the exoskeleton HAL are used signals of the patient nervous system. When patient tries to make a move, nerve signals are sent from the brain to the muscles via motoric neurons. These biosignals with low amplitude can be detected on the skin surface of the limbs. Device HAL detect these signals through a sensor attached to the skin of the wearer. Based on the measured signals the actuators amplify desired movement.

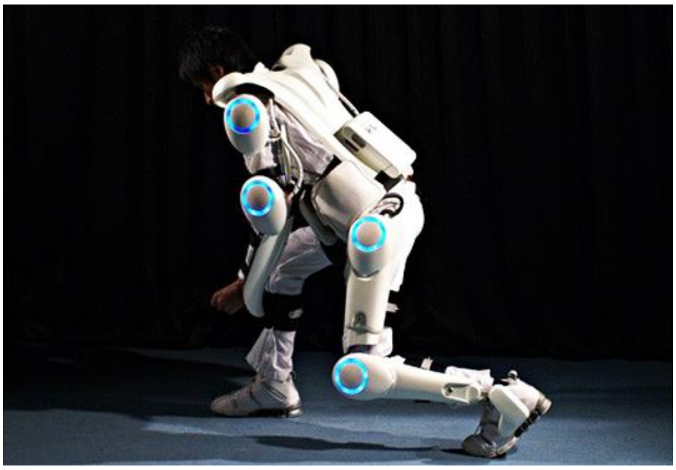


Fig. 8 Exoskeleton HAL 5 [6]

Exoskeleton HAL has two control systems. The first control system so called CVC (Cybernic Voluntary Control) is activated by the user and a second system called CAC (Cybernic Autonomous Control) responsible for supporting automatic movements.

IV. Conclusion

Our goal is to design a simple vertical rehabilitation system which will allow rehabilitation in standing position by controlled movements of the lower limbs. The aim is to design a system that imitates human walking and induce stepping automatism of the walk, what extend the possibilities of rehabilitation and physiological transfer of patient to different distances. The solution includes design of evaluation electronics which will be used to transfer the information from the sensor integrated in the bottom of the skeleton, to the patient. Integrated sensor has to analyze contact of skeleton with the floor during steps automatism. (VEGA 1/0074/15)

References

[1] MUDr. DOLAN, I. – MUDr. ŠRÁMKOVÁ, T.: Úrazy miechy, incidencia na Slovensku a prevencia[online]. Klinika úrazové chirurgie FNBrno.2013.<http://www.solen.sk/pdf/c5c876b0e2ea0744c095d2d506427d02.pdf>

[2] VÍTEČKOVÁ, S. - JIŘINA, M. - KRUPIČKA, R.: Exoskelety a aktivní ortézy dolních končetin[online]. Pohybové ústrojí. 2011.<http://www.pojivo.cz/pu/PU_34_2011.pdf>. ISSN: 1212-4575.

[3] Robots: Humanoids [online].<http://wordinfo.info/unit/3716/ip:4/il:R>.

[4] NOSOWITZ, D.: Real-Life Iron Man Exoskeleton Gets a Slimmer, More Powerful Sequel[online]. Popular science. 2010. <http://www.popsci.com/technology/article/2010-09/real-life-iron-man-suit-gets-sequel>.

[5] MELANSON, D.: Popular Science blows out the Sarcos XOS exoskeleton[online]. <http://www.engadget.com/2008/04/11/popular-science-blows-out-the-sarcos-xos-exoskeleton/>.

[6] Cyberdyne HAL[online].<http://www.dvice.com/archives/2009/04/cyberdyne_hal_e.php>.

[7] Walk Again[online]. <http://www.rewalk.com>.

[8] Rex Bionics[online]. <http://www.rexbionics.com>.

[9] Medicínska Technika. Kolenní motorová dlaha PRIMA ADVANCE - Kinetec[online].<http://www.medicinskatechnika.cz/kolenni-motorova-dlaha-prima-advance-kinetec-p-619.html>

[10] MOTOmed. MOTOmed including FES for active movement[online]. <http://www.motomed.com/en/models/motomed-including-fes.html>

[11] MEIER, Ch.: Rehabilitation robotics[online]. <http://web.ethlife.ethz.ch/e/articles/sciencelife/armin.html>

[12] ROBERTS, Karen, B.: Agrawal discusses role of robotic exoskeletons in stroke rehabilitation[online]. 26.10.2010 <http://www.udel.edu/udaily/2011/oct/agrawal-robotic exoskeletons102610.html>

[13] WalkTrainer.: Mobile equipment for verticalization and locomotion[online]. 19.08.2011 <http://lsro.epfl.ch/page-34300-en.html>

[14] Ekso[online]. Dostupné na internete: <http://intl.eksobionics.com/ekso>

[15] ZOSS, A. - KAZEROONI, H. - CHU, A.: On the Mechanical Design of the Berkeley Lower Extremity Exoskeleton (BLEEX). University of California, Berkeley, CA, 94720, USA.

[16] ZELINSKY, A.: Robot Suit Hybrid Assistive Limb. IEEE Robotics & Automation Magazine, december 2009.

Arrayed Waveguide Grating Based Demultiplexers in WDM Systems

[1]*Ján TÓTH (2nd year)*
Supervisor: [2]Ľuboš OVSENÍK

[1,2]Dept. of Electronics and Multimedia Communications, FEI TU of Košice, Slovak Republic

[1]jan.toth@tuke.sk, [2]lubos.ovsenik@tuke.sk

Abstract—**This paper presents AWG demultiplexer as one of the key components in optical communication systems. WDM networks will most likely substitute traditional network infrastructure in the near future. So far there is a problem with costs and sustainability of existing communication technologies. It is worthy to pay attention of finding some cheaper, more effective, more reliable and perhaps simpler innovation in terms of WDM components.**

Keywords—**AWG demultiplexer, hybrid filters, nonlinear effects, WDM systems.**

I. Introduction

In order to classify the most recent network architectures, the WDM (Wavelength Division Multiplexing) systems take their place. WDM systems are presented in two main standards which are commonly used nowadays. The first standard refers to so called "coarse" WDM which defines 20 nm channel spacing. On the other hand there is a newer standard called "dense" WDM (DWDM). This standard specifies 0.8 nm channel spacing in its basic form. It is obvious that DWDM systems provide a considerable savings in terms of bandwidth reserved per channel. Both standards have their advantages as well as drawbacks. CWDM system is usually cheaper and doesn't require high quality filters to separate particular informational channels [1]. Contrary to that, DWDM systems provide high number of channel available which turns to more efficiently using of a fiber capacity. These systems are mostly pricy and high quality components (filters) are necessary in order to provide a sufficient service. Moreover, DWMD systems meet some difficulties which are negligible in CWDM systems. Nonlinear effect, mainly Self Phase Modulation (SPM) and Four Wave Mixing (FWM) have quite considerable impact on a transmitted optical signal. These negative effects deteriorate signal in frequency domain. WDM links allow transmitting a huge amount of information through a single mode fiber (SMF). Perhaps the most complicated task (for WDM system) is to separate particular channels by using some kind of WDM filters. In fact there are more kinds of filters (demultiplexers) available nowadays. Getting through the list of commonly used multiplexers, there is one which should be considered more precisely [2]. Arrayed Waveguide Gratings (AWG) de/multiplexers turn to be promising solution for future WDM based networks.

II. Arrayed Waveguide Grating Based Demultiplexers

AWG demultiplexer is used to spatially separate a mixed optical signal which usually carries N – channels. Each channel is represented by one wavelength λ_{chi}. A finite number of channels – originally combined together by an AWG multiplexer – which is a complementary device to AWG demultiplexer travels through SMF and finally approaching demultiplexer.

A. *Mach – Zehnder Interferometer as AWG demultiplexer*

It is necessary to understand an operation of Mach – Zehnder interferometer (MZI) in order to understand AWG demultiplexer functionality. MZI can be used as a simple wavelength filter which takes advantage of constructive/destructive interference effect when a total phase shift $\Delta\Phi$ equals 2π. The phase is a function of a propagation constant β (1) as well as a physical distance which an optical signal needs to pass

$$\beta = \frac{2\pi}{\lambda} n_{eff}. \tag{1}$$

Parameter n_{eff} refers to an effective refractive index of Mach – Zehnder interferometer and λ is a wavelength.

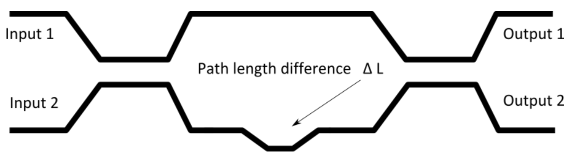


Fig. 1. A simple wavelength filter based on Mach – Zehnder interferometer principle.

Considering the Mach – Zehnder's outputs: *Output 1* equals $\pi/2 + \beta\Delta L + \pi/2$ and *Output 2* is $\pi/2 + \beta\Delta L - \pi/2 = \beta\Delta L$ it is obvious that the most important parameter in these forms is ΔL – a path length difference between two arms. WDM systems usually operate with multiple channels (wavelengths) $\lambda_1,\ldots,\lambda_i,\ldots\lambda_N$. These wavelengths λ_i which satisfy equation (2) for odd m_i values will appear at the *Output 1* of MZI.

$$\frac{n_{eff}\Delta L}{\lambda_i} = \frac{m_i}{2}. \tag{2}$$

The rest of wavelengths λ_i which match with even m_i values in equation (2) will appear at the *Output 2*. It is obvious that this device takes advantage of a path length difference ΔL which is tied with a propagation constant β. MZI device acts as a simple AWG demultiplexer with only two waveguides in a waveguide array [3]. Such a device has many practical implementations and it's essential to understand it.

B. Arrayed Waveguide Grating demultiplexer principle

AWG demultiplexer is made of two Free Propagation Regions (FPRs), input/output waveguides and the arrayed waveguides. This device combines functionality of diffraction gratings (which separates particular wavelengths within optical signal) as well as constructive/destructive interference principle – based on path length difference. Constructive interference is achieved at the spots where phases match 2π. A certain number of light sources is required in order to accomplish satisfactory high light intensities at the image plane of the second FPR as shown in figure (Fig. 2.). Higher number of light sources (a number of waveguides) equals to brighter spots at the image plane. In other words AWG demultiplexer has to spatially distinguish particular wavelengths (channels) at the image plane. FPRs have to be precisely designed and constructed for the actual wavelengths dedicated to a particular xWDM standard [4].

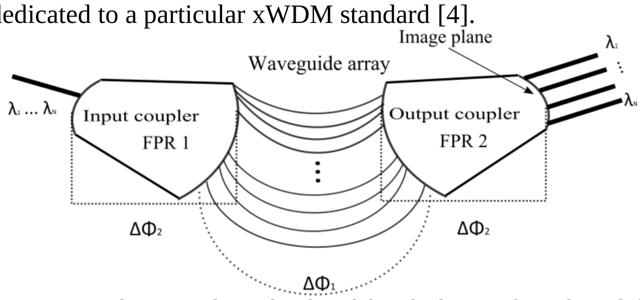


Fig. 2. Arrayed Waveguide Grading based demultiplexer with its phase shifts for particular regions (FPRs, waveguide arrays).

As it has been already mentioned the constructive interference can be achieved when phase shift matches 2π condition shown in equation (3).

$$\Delta\Phi_1 + 2\Delta\Phi_2 = n2\pi \quad (3)$$

AWG demultiplexer has basically two FPRs ($\Delta\Phi_2$ – phase shift) and a waveguide array with phase shift $\Delta\Phi_1$. Based on Rowland circle design (for FPRs physical proportions) the total phase shift can be expressed as show in the following equation (4)

$$2\pi\Delta L\frac{n_{eff}}{\lambda} + \frac{4\pi}{\lambda}R\sqrt{\left(1 + \sin\frac{\theta}{2}^2 - 2\sin\frac{\theta}{2}\cos\gamma\right)} = n2\pi, \quad (4)$$

where n is an integer, R is a radius of a Rowland circle shown in figure (Fig. 3.), n_{eff} refers to an effective refractive index of waveguide array (WA), θ and γ are the angles from Rowland circle, λ is a particular wavelength [4],[5].

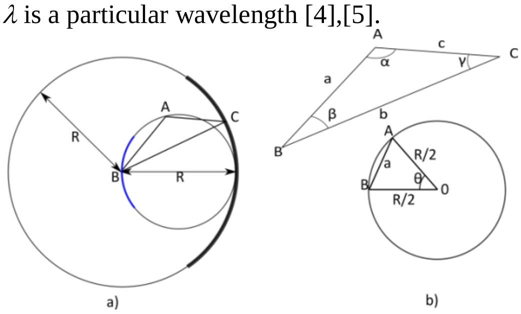


Fig. 3. Rowland circle principle required in construction of FPR for AWG demultiplexer.

III. Future work and Proposal

AWG based approach for demultiplexing of WDM optical signal is the one which has numerous advantages. Lately there have been proposed similar designs which use similar principles as AWGs. Planar Concave Gratings (PCG demultiplexers) and hybrid demultiplexers have a strong potential to compete or improve a traditional approach. There is a great ambition to design the entire optical link in order to investigate possibilities to improve AWG functionality and performance.

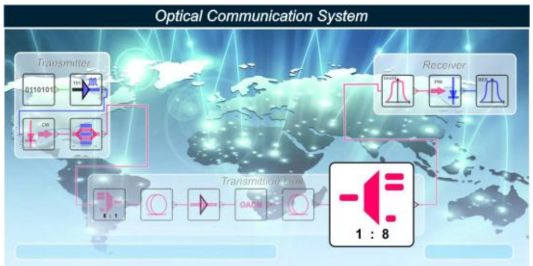


Fig. 4. Software prototype for optical link simulation.

To achieve this task we created a simulation software prototype to puzzle the optical link with its essentials optical elements shown in figure (Fig. 4.). AWG demultiplexer plays a key role in signal separating process. That's why it's quite important to start up with this particular device.

IV. Conclusion

In this paper we presented an idea of designing the all optical communication link. This system should be used in future to simulate a real operation. Such a system will be helpful to investigate AWG demultiplexer itself as well as some possible improvements. On the other hand we would like to take under the scope nonlinear effects which certainly appear in long distance optical systems with higher input power.

Acknowledgment

This work was supported by Cultural and Educational Grant Agency (KEGA) of the Ministry of Education, Science, Research and Sport of the Slovak Republic under the project no. „059TUKE-4/2015 - Implementation of the FSO Long-Term Experimental Measurement Results into the Educational Process in Optoelectronic Sensor and Communication Systems Field", and project no. „006TUKE-4/2014 - The Use of TUKE PON Experimental Model in Teaching". This work is also the result of the contract no. "APVV-0025-12 - Mitigation of Stochastic Effects in High-Bitrate All Optical Networks" supported by the Slovak Research and Development Agency.

References

[1] M. S. Wartak, "Computational photonics: An introduction with MATLAB," Waterloo, Ontario, Canada, 2013.

[2] G.T. Reed, A. P. Knights, "Silicon Photonics an introduction," John Wiley & Sons, Ltd, England, 2004.

[3] T. K. Chan, J. Krap, R. Jiang, N. Alic, S. Radic, CH. F.Marki, J.E. Ford , "1092 Channel 2-D Array Demultiplexer for Ultralarge Data Bandwidth," Journal of Lightwave Technology, Vol. 25, No. 3, March, 2007. Online:<http://psilab.ucsd.edu/publications/42.%20JLT%20wavelength%20demux.pdf>.

[4] M. K. Smit, C. Dam, "PHASAR-Based WDM-Devices: Principles, Design and Applications," IEEE Journal of Selected Topics in Quantum Electronics, Vol. 2, No. 2, June, 1996. Online: <http://repository.tudelft.nl/assets/uuid:37a5ebe1-766e-42cf-9511-573882edc4b8/ieee_smit_1996.pdf>.

[5] J. Brouckaert, W. Bogaerts, P. Dumon, D. Thourhout, R. Baets, "Planar Concave Grating Demultiplexer Fabricated on a Nanophotonic Silicon-on-Insulator Platform," Journal of Lightwave Technology, vol. 25, no. 5, May, 2007.

Biological Effects of Low Frequency Electromagnetic Fields

[1]*Peter CHALUPECKÝ (1st year)*
Supervisor: [2]Alexander MÉSZÁROS

[1,2]Dept. of Electric Power Engineering, FEI TU of Košice, Slovak Republic

[1]peter.chalupecky@student.tuke.sk, [2]alexander.meszaros@tuke.sk

***Abstract*—This article deals with biological effects caused by low frequency electromagnetic fields to humans and animals.**

***Keywords*— emf biological effects, emf exposure**

I. Introduction

The biological effects of low frequency electric and magnetic fields (EMF) have become a topic of considerable scientific scrutiny during the past two decades. The flurry of research in this area has contributed greatly to our understanding of the complex electromagnetic environment to which we are exposed but it has not abated the controversy associated with the harmful effects of electromagnetic fields.

Controversy is the norm when complex environmental issues with substantial economic and health consequences are scientifically scrutinized. Asbestos, lead, acid rain, tobacco smoke, DDT, PCBs (and more recently estrogen mimics) were all contentious issues and were debated for decades in scientific publications and in the popular press before their health effects and the mechanisms responsible were understood. In some cases the debate was scientifically legitimate, while in others interested parties deliberately confuse the issue to delay legislation. The public, uncomfortable with scientific controversy and unable to determine the legitimacy of a scientific debate, wants a clear answer to the question, "Are electric and magnetic fields harmful?" [1]

The question is valid and timely. The answer is likely to have far reaching consequences, considering our growing dependence on electric power, computer technology, and wireless communication, and it is likely to be of interest to a large population using, manufacturing, selling, and regulating this technology [1].

II. Trend in EMF Biological Effect Publications

The Biological Effects Policy Advisory Group (BEPAG), created by The Institution of Engineering and Technology (IET) in England and Wales, published an overview of published papers that provided data regarding biological effects of low-level electromagnetic fields. *Figure 1* shows the number of papers that provided the data for each of the position statements since 2004. The overall numbers have declined by about 35% since the peak in 2008. A comparison of numbers of papers for power and mobile frequencies shows a recent decline in the former with the latter remaining approximately constant. Approximately 40% of the papers fell outside these specific frequency categories, predominantly split between other radio and low frequencies.

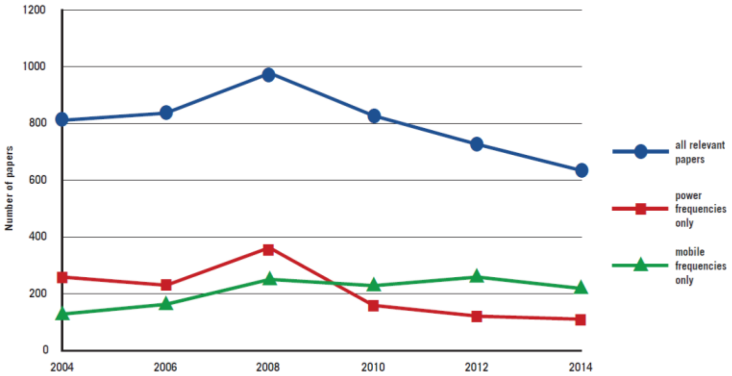


Figure 1: Trend in EMF Biological Effect publications

Figure 2 shows the split in papers based on their model type or methodology. Cellular studies are the most numerous, probably because of the relative ease of carrying out such studies. The largest decline can be seen in the 'physics' category, which includes dosimetry studies and perhaps reflects the absence of any new mechanisms of interaction and the maturity of the topic of environmental dosimetry. Animal studies have continued to increase steadily, which may reflect the increasing number of papers on EMFs originating from outside Europe and the U.S.A [2].

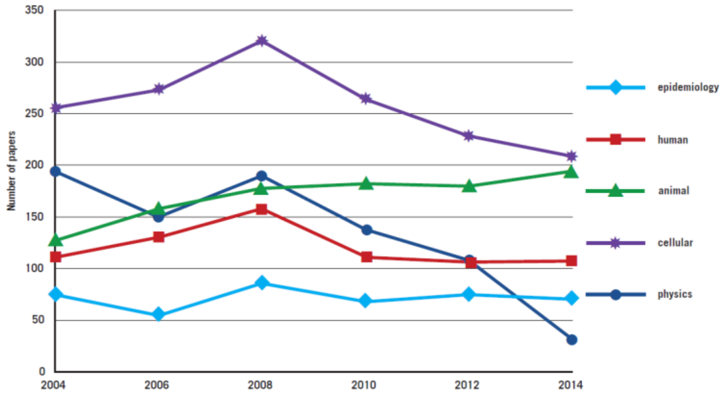


Figure 2: Paper numbers by subject

III. SOURCES, MEASUREMENTS AND EXPOSURES

Electric and magnetic fields exist wherever electricity is generated, transmitted or distributed in power lines or cables, or used in electrical appliances. Since the use of electricity is an integral part of our modern lifestyle, these fields are ubiquitous in our environment. The unit of electric field strength is volts per metre (V m^{-1}) or kilovolts per metre (kV m^{-1}) and for magnetic fields the flux density is measured in tesla (T), or more commonly in millitesla (mT) or microtesla (μT) is used. Residential exposure to power-frequency magnetic fields does not vary dramatically across the world. The geometric-mean magnetic field in homes ranges between 0,025 and 0,07 μT in Europe and 0,055 and 0,11 μT in the USA. The mean values of the electric field in the home are in the range of several tens of volts per metre. In the vicinity of certain appliances, the instantaneous magnetic-field values can be as much as a few hundred microtesla. Near power lines, magnetic fields reach approximately 20 μT and electric fields up to several thousand volts per metre. Few children have time-averaged exposures to residential 50 or 60 Hz magnetic fields in excess of the levels associated with an increased incidence of childhood leukaemia. Approximately 1% to 4% have mean exposures above 0,3 μT and only 1% to 2% have median exposures in excess of 0,4 μT. Occupational exposure, although predominantly to power-frequency fields, may also include contributions from other frequencies. The average magnetic field exposures in the workplace have been found to be higher in “electrical occupations” than in other occupations such as office work, ranging from 0,4–0,6 μT for electricians and electrical engineers to approximately 1,0 μT for power line workers, with the highest exposures for welders, railway engine drivers and sewing machine operators (above 3 μT). The maximum magnetic field exposures in the workplace can reach approximately 10 mT and this is invariably associated with the presence of conductors carrying high currents. In the electrical supply industry, workers may be exposed to electric fields up to 30 kV m^{-1} [3].

IV. MAIN CONCLUSIONS FROM PREVIOUS RESEARCH

A. Neurobehaviour

Exposure to power-frequency electric fields causes well-defined biological responses, ranging from perception to annoyance, through surface electric charge effects. These responses depend on the field strength, the ambient environmental conditions and individual sensitivity. The thresholds for direct perception by 10% of volunteers varied between 2 and 20 kV m^{-1}, while 5% found 15–20 kV m^{-1} annoying. The spark discharge from a person to ground is found to be painful by 7% of volunteers in a field of 5 kV m^{-1}.
Thresholds for the discharge from a charged object through a grounded person depend on the size of the object and therefore require specific assessment. High field strength, rapidly pulsed magnetic fields can stimulate peripheral or central nerve tissue; such effects can arise during magnetic resonance imaging (MRI) procedures, and are used in transcranial magnetic stimulation. Threshold induced electric field strengths for direct nerve stimulation could be as low as a few volts per metre. The threshold is likely to be constant over a frequency range between a few hertz and a few kilohertz. People suffering from or predisposed to epilepsy are likely to be more susceptible to induced extremely low frequency (ELF) electric fields in the central nervous system (CNS). The function of the retina, which is a part of the CNS, can be affected by exposure to much weaker ELF magnetic fields than those that cause direct nerve stimulation. A flickering light sensation, called magnetic phosphenes or magnetophosphenes, results from the interaction of the induced electric field with electrically excitable cells in the retina. Threshold induced electric field strengths in the extracellular fluid of the retina have been estimated to lie between about 10 and 100 mV m^{-1} at 20 Hz. There is, however, considerable uncertainty attached to these values. The evidence for other neurobehavioural effects in volunteer studies, such as the effects on brain electrical activity, cognition, sleep, hypersensitivity and mood, is less clear. There is some evidence suggesting the existence of field-dependent effects on reaction time and on reduced accuracy in the performance of some cognitive tasks, which is supported by the results of studies on the gross electrical activity of the brain.

In animals, the possibility that exposure to ELF fields may affect neurobehavioural functions has been explored from a number of perspectives using a range of exposure conditions. Few robust effects have been established. There is convincing evidence that power-frequency electric fields can be detected by animals, most likely as a result of surface charge effects, and may elicit transient arousal or mild stress. In rats, the detection range is between 3 and 13 kV m^{-1}. Rodents have been shown to be aversive to field strengths greater than 50 kV m^{-1}. Other possible field-dependent changes are less well-defined; laboratory studies have only produced evidence of subtle and transitory effects. There is some evidence that exposure to magnetic fields may modulate the functions of the opioid and cholinergic neurotransmitter systems in the brain, and this is supported by the results of studies investigating the effects on analgesia and on the acquisition and performance of spatial memory tasks [4],[5].

B. Neuroendocrine system

The results of volunteer studies as well as residential and occupational epidemiological studies suggest that the neuroendocrine system is not adversely affected by exposure to power-frequency electric or magnetic fields. This applies particularly to the circulating levels of specific hormones of the neuroendocrine system, including melatonin, released by the pineal gland, and to a number of hormones involved in the control of body metabolism and physiology, released by the pituitary gland. Subtle differences were sometimes observed in the timing of melatonin release associated with certain characteristics of exposure, but these results were not consistent.

From the large number of animal studies investigating the effects of power-frequency electric and magnetic fields on rat pineal and serum melatonin levels, some reported that exposure resulted in night-time suppression of melatonin. The changes in melatonin levels first observed in early studies of electric field exposures up to 100 kV m^{-1} could not be replicated. The findings from a series of more recent studies, which showed that circularly polarised magnetic fields suppressed night-time melatonin levels, were weakened by inappropriate comparisons between exposed animals and historical controls. The data from other experiments in rodents, covering intensity levels from a few microtesla to 5

mT, were equivocal, with some results showing depression of melatonin, but others showing no changes. In seasonally breeding animals, the evidence for an effect of exposure to power-frequency fields on melatonin levels and melatonin-dependent reproductive status is predominantly negative. No convincing effect on melatonin levels has been seen in a study of non-human primates chronically exposed to power-frequency fields, although a preliminary study using two animals reported melatonin suppression in response to an irregular and intermittent exposure.
Overall, these data do not indicate that ELF electric and/or magnetic fields affect the neuroendocrine system in a way that would have an adverse impact on human health and the evidence is thus considered inadequate [6],[7].

C. Neurodegenerative disorders

It has been hypothesized that exposure to ELF fields is associated with several neurodegenerative diseases. For Parkinson disease and multiple sclerosis the number of studies has been small and there is no evidence for an association with these diseases. For Alzheimer disease and amyotrophic lateral sclerosis (ALS) more studies have been published. Some of these reports suggest that people employed in electrical occupations might have an increased risk of ALS. So far, no biological mechanism has been established which can explain this association, although it could have arisen because of confounders related to electrical occupations, such as electric shocks. Overall, the evidence for the association between ELF exposure and ALS is considered to be inadequate.
The few studies investigating the association between ELF exposure and Alzheimer disease are inconsistent. However, the higher quality studies that focused on Alzheimer morbidity rather than mortality do not indicate an association. Altogether, the evidence for an association between ELF exposure and Alzheimer disease is inadequate [8],[9].

D. Cardiovascular disorders

Experimental studies of both short-term and long-term exposure indicate that while electric shock is an obvious health hazard, other hazardous cardiovascular effects associated with ELF fields are unlikely to occur at exposure levels commonly encountered environmentally or occupationally. Although various cardiovascular changes have been reported in the literature, the majority of effects are small and the results have not been consistent within and between studies. Whether a specific association exists between exposure and altered autonomic control of the heart remains speculative. Overall, the evidence does not support an association between ELF exposure and cardiovascular disease [10].

E. Immunology and haematology

Evidence for the effects of ELF electric or magnetic fields on components of the immune system is generally inconsistent. Many of the cell populations and functional markers were unaffected by exposure. However, in some human studies with fields from 10 μT to 2 mT, changes were observed in natural killer cells, which showed both increased and decreased cell numbers, and in total white blood cell counts, which showed no change or decreased numbers. In animal studies, reduced natural killer cell activity was seen in female mice, but not in male mice or in rats of either sex. White blood cell counts also showed inconsistency, with decreases or no change reported in different studies. The animal exposures had an even broader range of 2 μT to 30 mT. The difficulty in interpreting the potential health impact of these data is due to the large variations in exposure and environmental conditions, the relatively small numbers of subjects tested and the broad range of endpoints.
There have been few studies carried out on the effects of ELF magnetic fields on the haematological system. In experiments evaluating differential white blood cell counts, exposures ranged from 2 μT to 2 mT. No consistent effects of acute exposure to ELF magnetic fields or to combined ELF electric and magnetic fields have been found in either human or animal studies.
Overall therefore, the evidence for effects of ELF electric or magnetic fields on the immune and haematological system is considered inadequate [11].

F. Reproduction and development

On the whole, epidemiological studies have not shown an association between adverse human reproductive outcomes and maternal or paternal exposure to ELF fields. There is some evidence for an increased risk of miscarriage associated with maternal magnetic field exposure, but this evidence is inadequate.
Exposures to ELF electric fields of up to 150 kV m^{-1} have been evaluated in several mammalian species, including studies with large group sizes and exposure over several generations. The results consistently show no adverse developmental effects.
The exposure of mammals to ELF magnetic fields of up to 20 mT does not result in gross external, visceral or skeletal malformations. Some studies show an increase in minor skeletal anomalies, in both rats and mice. Skeletal variations are relatively common findings in teratological studies and are often considered biologically insignificant. However, subtle effects of magnetic fields on skeletal development cannot be ruled out. Very few studies have been published which address reproductive effects and no conclusions can be drawn from them. Several studies on non-mammalian experimental models (chick embryos, fish, sea urchins and insects) have reported findings indicating that ELF magnetic fields at microtesla levels may disturb early development. However, the findings of non-mammalian experimental models carry less weight in the overall evaluation of developmental toxicity than those of corresponding mammalian studies.
Overall, the evidence for developmental and reproductive effects is inadequate [12],[13].

G. Cancer

The IARC classification of ELF magnetic fields as “possibly carcinogenic to humans” (IARC, 2002) is based upon all of the available data prior to and including 2001. The review of literature in this EHC monograph focuses mainly on studies published after the IARC review.

H. Epidemiology

The IARC classification was heavily influenced by the associations observed in epidemiological studies on childhood leukaemia. The classification of this evidence as limited does not change with the addition of two childhood leukaemia studies published after 2002. Since the publication of the IARC monograph the evidence for other childhood cancers remains inadequate.

Subsequent to the IARC monograph a number of reports have been published concerning the risk of female breast cancer in adults associated with ELF magnetic field exposure. These studies are larger than the previous ones and less susceptible to bias, and overall are negative. With these studies, the evidence for an association between ELF magnetic field exposure and the risk of female breast cancer is weakened considerably and does not support an association of this kind.

In the case of adult brain cancer and leukaemia, the new studies published after the IARC monograph do not change the conclusion that the overall evidence for an association between ELF magnetic fields and the risk of these diseases remains inadequate.

For other diseases and all other cancers, the evidence remains inadequate.

I. Laboratory animal studies

There is currently no adequate animal model of the most common form of childhood leukaemia, acute lymphoblastic leukaemia. Three independent large-scale studies of rats provided no evidence of an effect of ELF magnetic fields on the incidence of spontaneous mammary tumours. Most studies report no effect of ELF magnetic fields on leukaemia or lymphoma in rodent models. Several large-scale long-term studies in rodents have not shown any consistent increase in any type of cancer, including haematopoietic, mammary, brain and skin tumours.

Two groups have reported increased levels of DNA strand breaks in brain tissue following in vivo exposure to ELF magnetic fields. However, other groups, using a variety of different rodent genotoxicity models, found no evidence of genotoxic effects. The results of studies investigating non genotoxic effects relevant to cancer are inconclusive.

Overall there is no evidence that exposure to ELF magnetic fields alone causes tumours. The evidence that ELF magnetic field exposure can enhance tumour development in combination with carcinogens is inadequate.

J. In vitro studies

Generally, studies of the effects of ELF field exposure of cells have shown no induction of genotoxicity at fields below 50 mT. The notable exception is evidence from recent studies reporting DNA damage at field strengths as low as 35 μT; however, these studies are still being evaluated and our understanding of these findings is incomplete. There is also increasing evidence that ELF magnetic fields may interact with DNA-damaging agents.

There is no clear evidence of the activation by ELF magnetic fields of genes associated with the control of the cell cycle. However, systematic studies analysing the response of the whole genome have yet to be performed.

Overall conclusion

New human, animal and in vitro studies, published since the 2002 IARC monograph, do not change the overall classification of ELF magnetic fields as a possible human carcinogen [14],[15],[16].

V. RECOMMENDATIONS FOR FUTURE RESEARCH

In future research we would like to provide in vitro laboratory studies focusing on cellular changes related to exposure to radiation of varying intensity and frequency. Priority should be given to reported responses for which there is at least some evidence of replication or confirmation, that is strong enough to allow mechanistic analysis and that occurs in mammalian or human systems.

Regarding high density of high frequency electric fields, more tests should be focused on retina defects and relevant CNS disorders in the future.

ACKNOWLEDGMENT

The author wish to acknowledge Scientific Grant Agency of The Ministry of Slovak Republic and Slovak Academy of Science for funding of experimental works in the frame of VEGA No.1/0132/15 grant Research of the penetrating of high frequency electromagnetic waves through ecological building materials.

REFERENCES

[1] Magda Havas, "Biological effects of low frequency electromagnetic fields", in D. Clements-Croome (Ed.). 2004. Electromagnetic Environments and Health in Buildings. Spon Press, London, 535 pp.

[2] Are there harmful Biological Effects of Low/Level Electromagnetic Fields at frequencies up to 300GHz ?, The Institution of Engineering and Technology, Position Statement, 2014

[3] Floderus B, Persson T, Stenlund C (1996). Magnetic field exposures in the workplace: reference distribution and exposures in occupational groups. Int J Occup Environ Helath, 2:226-238

[4] Cook MR et al. (1992). A replication study of human exposure to 60-Hz fields - Effects on neurobehavioral measures. *Bioelectromagnetics*, 13(4):261-285.

[5] ICNIRP - International Commission on Non-ionizing Radiation Protection. Possible health risk to the general public from the use of security and similar devices. Report of a concerted action within the project: "Environment and Health, Health impact of electromagnetic fields" of the 5th Framework Programme of the European Commission. Bernhardt JH, McKinlay A, Matthes R, eds. Munich, International Commission on Non-ionizing Radiation Protection, 2002 (ICNIRP 12/2002).

[6] Warman GR et al. (2003b). Circadian neuroendocrine physiology and electromagnetic field studies: precautions and complexities. *Radiat Prot Dosimetry*, 106(4):369-373.

[7] Quinlan WJ et al. (1985). Neuroendocrine parameters in the rat exposed to 60-Hz electric fields. Bioelectromagnetics, 6:381-389.

[8] AGNIR - Advisory Group on Non-Ionising Radiation. ELF electromagnetic fields and neurodegenerative disease. Chilton, National Radiological Protection Board, 2001a (Documents of the NRPB, Vol. 12, No. 4).

[9] Coyle JT, Puttfarcken P (1993). Oxidative stress, glutamate, and neurodegenerative disorders.*Science*, 262(5134):689-695.

[10] Finkelstein MM (1999). Re: magnetic field exposure and cardiovascular disease mortality among electric utility workers - letter to editor. *Am J Epidemiol*, 150(11):1258.

[11] Environmental Health Criteria 238, Extremely Low Frequency Fields, World Health Organization,2007, p.8.

[12] Mevissen M, Buntenkotter S, Löscher W (1994). Effects of static and time-varying (50-Hz) magnetic-fields on reproduction and fetal development in rats. Teratology, 50(3):229-237.

[13] Negishi T et al. (2002). Studies of 50 Hz circularly polarized magnetic fields of up to 350 microT on reproduction and embryo-fetal development in rats: exposure during organogenesis or during preimplantation. *Bioelectromagnetics*, 23(5):369-389.

[14] Neutra RR, Del Pizzo V (1996). When "Wire Codes" predict cancer better than spot measurements of magnetic fields. Epidemiology, 7(3):217-218.

[15] Olsen JH, Nielsen A, Schulgen G (1993). Residence near high-voltage facilities and risk of cancer in children. *Br Med J*, 307(6909):891-895.

[16] Phillips JL, McChesney L (1991). Effect of 72 Hz pulsed magnetic field exposure on macromolecular synthesis in CCRF-CEM cells. *Cancer Biochem Biophys*, 12(1):1-7.

Comparison of valuation methods in a competitive electricity market

Martin PLOTH (2st year)
Supervisor: Alexander MÉSZÁROS

Dept. of Electric Power Engineering, FEI TU of Košice, Slovak Republic

martin.ploth@student.tuke.sk, alexander.meszaros@tuke.sk

Abstract **- This paper deals with my work in previous year. The paper deals with the issue of pricing and valuation of selected approaches based of network flows. As a reference method is a method postage stamp, which is used to compare with the methods MW- km and methods GGDF factors on standard IEEE 14- bus test system.**

Keywords **- method postage stamp, method MW - km, method GGDF factors.**

I. INTRODUCTION

One of most difficult issues carrying on each restructuring of the electricity market is a fair distribution of the costs of transmission between each performed transactions and transmission facilities. The allocation of the cost of each transactions so that it is fair, i.e. reflect the real use of transmission system is a complex and a demanding task. There are many methods of pricing transmission, which more or less reflects the real use of the transmission network. At present in the territory of the Slovak Republic, i.e. area operated by SEPS a.s. is used pricing method postage stamp. [1] [2] A drawback of the method is no consideration of the length of the transmission and not already real use and physical flows caused by the transaction. Its advantage is the ease of cost allocation. This method will in this comparison serve as a reference to the two methods discussed widely in the world, it is a method based on network namely method MW-km [3][4][5][6] and a method based on the basis of flows in the network, namely method of distribution coefficients use GGDF factors. [3] [5] [7] [8]

II. VALUATION ISSUES

As previously mentioned, these two approaches to pricing problematic are significantly different. Network - based approach, notes into account the real flows in the network, but not their impact on the entire network. Specifically method MW-km is method line-by-line and fixed price has two components: the size of the transmitted power, as well as distance. It is one of the first pricing strategy, is based on DC power flow calculation. On the other hand, is a method, that is based on the power flows in the network. This is a method of distribution coefficients in this case the Generalized Generation Shift Distribution Factors (GGDFs coefficients or D coefficients) determining effect on each generator in system for active power flows. [4] Provides overall use facilities of the transmission system. The coefficients are based on DC current model. These pricing strategies are often modified, especially the method MW-km, so that the offset shortcomings, particularly with a view to the unused portion of the network. Complex view on the issue of pricing, I processed in paper "A review of issue of valuations transmission services in restructured electricity market", as a contribution to the conference SCYR 2014. [5]

III. DESCRIPTION SOLUTIONS PROBLEM

A comparison of two approaches was chosen standard IEEE 14-bus system with two generators. Generators are connected to bus 1 and bus 2. [5] One important factor is that the reference method postage stamps does not consider the length of the line system, and therefore the values are very low compared to both methods. The calculation is carried out for 20 lines k. Where calculation methods postage stamp, line cost C_k we set on 10 €. Calculated costs C_kMW_{1k} are for generator G_1 and lines k 1 to 20 as follows: 368,86; 260,5; 321; 265,39; 98,04; 249,73; 129,89; 74,51; 185,59; 27,09; 33,36; 75,27; 0; 129,89; 28,47; 44,75; 11,57; 6,27 and 21,49. For generator G_2 are costs C_kMW_{2k} same. Results are in Table 1.

TABLE I

CALCULATION METHOD POSTAGE STAMP			
Line k	Line cost	G_1	G_2
	C_k (€)	$C_kMW_{1,k}$	$C_kMW_{2,k}$
Total	200	2531,3	2531,3
Total cost of each generator (€)		2556,09	2556,09
Total cost (€)		5062,6	
Each transaction cost (€)		100	100
Cost (€/MW)		0,61	0,76

We assume, that the length of the lines are proportional to the line reactance. As a method MW-km, same as well as the method GGDF used DC power flows. The values of line costs C_kL_k (€) for lines k 1 to 20 are for method MW-km following: 52,64; 198,43; 176,13; 156,87; 154,7; 152,16; 37,46; 8,9; 8,9; 8,9; 265,93; 227,59; 115,9; 8,9; 8,9; 75,18; 240,55; 170,88; 177,83 and 220,66. Calculated costs

$C_kL_kMW_{1k}$ are for generator G_1 and lines k 1 to 20 as follows: 2446,29; 6002,63; 6326,58; 4625,26; 3555,72; 1701,17; 1069,61; 129,76; 74,42; 185,85; 541,49; 854,59; 981,65; 0; 129,76; 239,59; 1206,36; 224,82; 125,36 and 752,38. For generator G_2 are costs $C_kL_kMW_{2k}$ same.

TABLE II

CALCULATION METHOD MW-KM			
Line k	Line cost	G_1	G_2
	C_kL_k (€)	$C_kL_kMW_{1,k}$	$C_kL_kMW_{2,k}$
Total	2466,87	31173,29	31173,29
Total cost (€)		62346,58	
Each transaction cost (€)		1233,69	1233,69
Cost (€/MW)		8,03	11,68

The values of line costs C_kL_k (€) for lines k 1 to 20 are for method GGDF same as in method MW-km. Calculated costs $C_kL_kMW_{1k}$ are for generator G_1 and lines k 1 to 20 as follows: 5657,65; 9122,69; 7195,44; 4919,29; 3464,48; 1755,6; 1079,95; 152,14; 87,26; 222,75; 671,93; 1018,36; 1173,4; 0; 152,14; 270,85; 1402,26; 237; 153,02 and 936,56. For generator G_2 are costs $C_kL_kMW_{2k}$ follows: 764,91; 2882,56; 5457,73; 4331,24; 3646,97; 1646,74; 1059,27; 107,37; 61,57; 148,95; 411,04; 690,82; 789,91; 0; 107,37; 207,47; 1010,47; 212,4; 97,15 and 578,22.

TABLE III

CALCULATION METHOD GGDF			
Line k	Line cost	G_1	G_2
	C_kL_k (€)	$C_kL_kMW_{1,k}$	$C_kL_kMW_{2,k}$
Total	2466,87	39662,77	24212,16
Total cost (€)		63874,93	
Each transaction cost (€)		1532,56	935,31
Cost (€/MW)		9,96	8,54

Results are in Table 2 and Table 3. The issue of valuation using the method of distribution coefficients has already been dealt with in paper " Metódy oceňovania prenosových služieb". [4]

IV. CONCLUSION

Currently, the method used postage stamp in area of Slovak republic. This method does not reflect the real use of transmission network and is unfair to every transaction. The method postage stamp have undergone comparison with the method MW- mile and GGDF factors. Based on the results, it can be stated, that the method postage stamp does not reflect transmission costs sufficiently, and the results compared methods can be stated that the method GGDF distributes costs most fairer among the compared methods.

These methods have been tested on 14- bus system, but for verification of these conclusions, it is necessary to verify the results on the larger transmission test systems. Other methods are valuation perspective, such as Kirscher and Bialek tracing method, which will focused on the future.

ACKNOWLEDGMENT

This work was supported by the Scientific Grant Agency of the Ministry of Education of Slovak Republic and the Slovak Academy of Sciences under the contract No. Vega 1/0388/13.

REFERENCES

[1] Chemišinec, I.; Marvan, M.; Nečesaný, J.; Sýkora, T.; Tůma, J.,: " Obchod s elektřinou,": Conte Praha , 2010

[2] Hassan, M.Y.; Majid, M.S.; Hussin, F.; Rahman, H.A.; Lo, K.L;: "Certain Considerations In Pricing Unbundled Transmission Services", In: First International power and Energy Concerence PECon 2006: November 28-29th, 2006, Purtajaya, Malaysia, p272-275

[3] Shahidehpour, M. ; Yamin, H.; Li, Z., : " Market Operations in Electric Power Systems": Library of congress Cataloging in Publication Data, 2002, ISBN 0-47L-44337-9

[4] Lo, K. L.; Hassan, M. Y.; Jovanovic, S.,:"Assessment of MW-mile method for pricing transmission services: a negative flow-sharing approach," Generation, Transmission & Distribution, IET , vol.1, no.6, pp.904,911, Nov. 2007

[5] Orfanos, G.A.; Tziasiou, P.S.;Georgilakis, P.S.; Hatziargyriou, N.D.: "Evaluations of transmission Pricing Methodologies for Pool Based Electricity Martkets", 2011 IEEE Trondheim PowerTech, 2011, Trondheim, Norway, p1-8

[6] Jiuping, P.; Yonael, T.; Saifur, R.,: "Review of Usage-Based Transmission Cost Allocation Methods under Open Access", IEEE Transactions on power systems, Vol 15,No,4, Nov. 2000, p.1218-1224

[7] Qian Zhang; Jihui Yu, "Research on transmission services pricing based on allocation of fixed transmission cost," *Services Systems and Services Management, 2005. Proceedings of ICSSSM '05. 2005 International Conference on* , vol.2, no., pp.1369,1374 Vol. 2, 13-15 June 2005

[8] Kilyeni, S.; Pop, O. ; Slavici, T.; Craciun, C.; Andea, P.; Mnerie, D."Transmission cost allocation using the distribution factors method", MELECON 2010 -2010 15th IEEE Mediterranean Electrotechnical Conference, pp 1093,1098, 26-28 April 2010, ISBN 978-1-4244-5793-9

[9] Ploth, M.; Mészáros, A.; Zbojovský,J.,: " Metódy oceňovania prenosových služieb " , in Electrical Engineering and Informatics V Proceeding of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice, Košice 2014 S.869 -874 .ISBN 978-80-553-1704-5

[10] Ploth, M.; Mészáros A., :" A review of issue of valuations transmission services in restructured electricity maerket ", In: SCYR 2014 : proceedings from conference : 14th Scientific Conference of Young Researchers : May 20th, 2014, Herľany, Slovakia. - Košice : TU, 2014 S. 15-18. ISBN 978-80-553-1714-4

[11] Murali ,M. ; Sailaja Kumari, M. ; Sydulu, M.,: " A review of transmission pricing methods in restructured electricity market and case studies ", Electricity Engineering department , National Institute of Technology; Warangal, International Electrical Engineering Journal (IEEJ) Vol. 5 (2014) No. 1, p. 1186-1197

Context-perceptive Human Computer Interface for Smart Environment

[1]*Ladislav JACHO (1st year)*
Supervisor: [2]Branislav SOBOTA

[1,2]Dept. of Computers and Informatics, FEI TU of Košice, Slovak Republic

[1]ladislav.jacho@tuke.sk, [2]branislav.sobota@tuke.sk

***Abstract*— Body language together with spoken words describes the way we communicate with other people. Our natural ways of communication could hardly take a role in relationship of human and computational system for decades. We got used to work with artificial devices such as keyboard and mouse. In these days we interact with ubiquitous computing presented in our closest living environment. Computational power of various types of devices is used to provide the user interface as comfortable as possible. This paper aims at definition and evaluation of new and progressive user interfaces of smart environments with perception of human behavior.**

***Keywords*—human-computer interaction, smart environment, ambient intelligence, natural user interface**

I. INTRODUCTION

Miniaturization of computers without direct proportion on the loss of the performance together with increase of stability of wireless communication standards provides an advantage to almost hide the presence of various types of devices. Some of them could be purchased at affordable price. Their coordination with the use of proper software can lead to building up of complex unit able to sense the events and context of specific environment. Acquired data about inhabitants and environment is crucial to perform prediction and automated task.

Connection of all mentioned facts covers the term of smart environment (SmE) or ubiquitous computing (UbC) and ambient intelligence (AmI).

Smart environments could bring extended solutions in:

- Safety, security and privacy
- Health, wellness and ergonomics
- Resource saving
- Entertainment and convenience.

The field of a smart environment is strongly related to other disciplines and achievements in areas of computer science shown in Fig. 1 [1].

Despite desirable automation of various tasks at workplaces and living environments there will still remain tasks with the need of human attention. It is connected with natural will of controlling the space, where we exist. In the connection with SmE there is a challenge to define suitable user interfaces (UI). UI is the research subject of human-computer interaction (HCI).

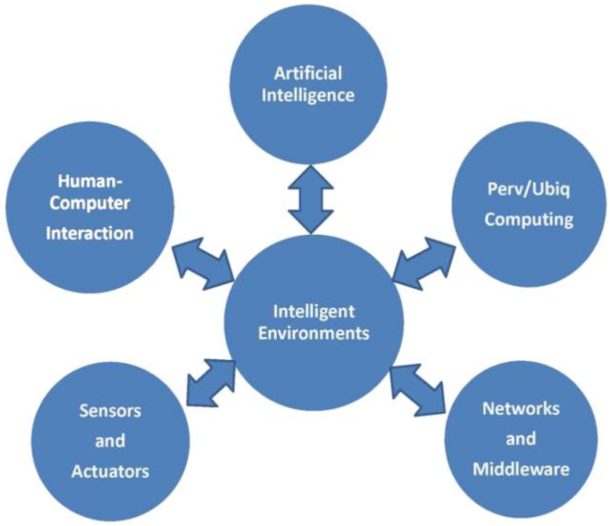


Fig. 1 Interaction between intelligent environments and other disciplines [1]

II. HUMAN-COMPUTER INTERACTION AND SMART ENVIRONMENTS

The way we interact with computers is the result of either technological but also non-technological changes of our society. Some of changes are caused by increased level of stability and reliability level of particular technology with commercial success. Most significant non-technological triggers include increase of expectation from computational system and the loss of fear from using computers. People are becoming more liberal. The phenomenon is the result of coming of new technology consumer's generations.

SmE is divided into three groups of components according their main functionality:

- Sensors
- Controllers
- Actuators.

Sensors are core devices which are used to record data about the space and its inhabitants. It includes physical properties like pressure, temperature and humidity or biometric properties like capturing face, voice or gait. Extended list of properties contains TABLE I [2].

Controllers dispose with computational performance to classify, store, retrieve but also analyze and disseminate data from sensors. Analysis of different information in context is demanding great performance that causes delegation of computation per external network to cloud technologies.

Actuators are used to provide changes of environment based on commands from the user or analysis of context data. Mediator of both cases is usually a controller. Actuators consist from electric and mechanical parts like electro motors, electric valves, and thermostats.

Functionality of each component type could be embedded into one device. Smartphone or other wearable devices could accomplish functions of all three types in parallel. Fig. 3 presents smart environment in layers according to the interaction level with the user [3]. In the lowest layer are appliances, distributed smart devices together with robot agents which are getting more capable to replace the user in many activities and also capable to satisfy social needs in the rising scope [4]. Cooperation of SmE and robots could dramatically simplify robotic tasks connected with communication and information exchange with the user [5].

TABLE I
MEASURABLE PROPERTIES OF SMART ENVIRONMENTS* [2]

Category	Measurand
Physical properties	pressure, temperature, humidity, flow
Motion properties	position, velocity, angular velocity, acceleration
Contact properties	strain, force, torque, vibration
Presence	tactile/contact, proximity, distance/range, motion
Biochemical props.	biochemical agents
Identification	RFID/person ID, biometrics

* Table includes measurable properties of the surrounding space and the user as a part of the environment

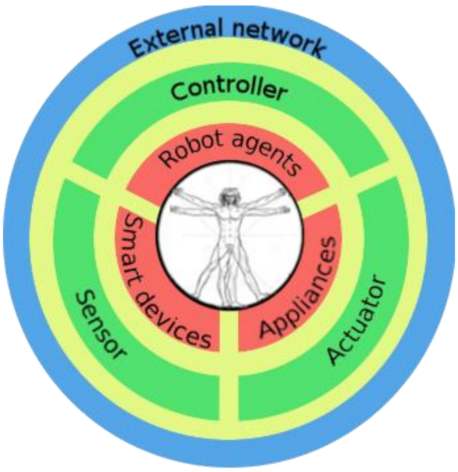


Fig. 3 Schematic view of layers of smart environment presenting components in interaction with the user [3]

According to Mark Weiser's origins of UbC presented in [6], the user interface of smart environment is highly based on human-centered design without necessary addition of abstraction level like desktop metaphor and window system known from graphical user interface (GUI). With increase of requirements of easy understanding and learning to work with the user interface, modern user interfaces are inspired by natural way of communication using posture, gestures, spoken words and handwriting. This approach is characteristic for natural user interface (NUI).

The user interaction and SmE consists of both, *conscious* or *direct interaction* and *subconscious* or *indirect interaction*.

A. *Conscious (direct) interaction*

Conscious interaction is connected with performing tasks of control and SmE setup or obtaining information which are related or unrelated with SmE. Remote control is not considered to be a part of this category.

The user should be allowed to use the way of communication which is the most comfortable for him/her. Basic element of NUI is a touch on the surface. UI based on touch is widely used in mobile devices with embedded capacitive display. There are known and frequently used gestures which were outlined [7] or arise through evolution of touch interface. Mostly used set of gestures contains tap, swipe, pinch, zoom and rotate. Entire system controlled by touch interface could avoid the loss on scalability option per providing possibility of creation the user defined gestures. Similar requirement are set for voice commands and gesture-based commands in 3D space. *VRPITA* [9] is the system with custom gesture creation support presented in Fig. 2.

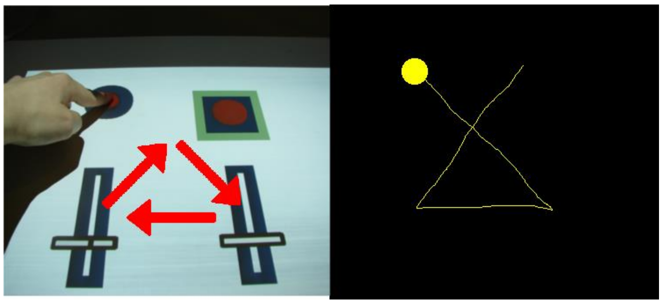


Fig. 2 Left – touch user interface of VRPITA system (red arrows describe direction of fingertip movements), Right – acceptable gesture pattern

B. *Subconscious (indirect) interaction*

Interaction with physical world can play a role in communication with other people. For example in particular context if someone enters the room all other people are asked to stand up. This kind of signal could be used to turn on the lights in the room or in the opposite turning off heating in case of leaving. Interaction with physical objects could create inputs for computational system without further interventions and needed attention of the user. Physical objects are becoming building elements of spatial [10] or grasp user interfaces [11].

III. INPUTS AND CONTEXT DATA

Recognition of the user behavior represents the input data for SmE. In the case of high support and assistance in daily living it is necessary to perceive context of the user's outer activity. Contextual data basics include a location where the activity occurs. Inputs in the form of contextual data considering location are widely used in navigation systems of mobile devices. In case of SmE it is necessary to deal with rich content consists from heterogeneous data including time and other measurable values describing status of the entire environment [8]. Human behavior is connected with emotions in many cases. This fact introduces new requirements on capturing and data analysis from different perspective. Captured data need to be compared to the rich information base. Credible outputs highly depend on quality of the model which needs to be precisely chosen mostly on decisions of human beings [12].

There are many challenges in building appropriate UI considering desirable properties. Many of them could not be implemented using known methods of building traditional user interfaces. Entire system of AmI should be adaptable and highly customizable according to variety of living spaces.

All human activities of interaction with physical space and other people are significant through properties:

- activities are continuous and nested (one activity could be a part of another)
- activities are abruptly interrupted because of unexpected events or lose of attention (they could be resumed)
- activities rarely have exact starting and ending point so determination requires flexible approach
- exchange of particular steps could describe the same activity

For appropriate activity determination there is a need to create associative information model [3].

IV. Distribution and form of outputs

Outputs of computational process are in case of SmE in the form of status change or delivering additional information. This information should be reached in many forms (image, sound, video, text) using different ways. Content should be available online via web browsers, mobile devices and embedded in environment. Broad range of UI modalities on each device requires flexible approach to layout generation according to the context. Generation and the placement of UI elements should consider limitation of display devices, distance, orientation and angle of view. Our physical environment is exhibited to less or more significant changes which could be hardly predictable. Displaying surface could also be the subject of change for example in the case of projection. GUI needs to be continuously adjustable in this case [13]

Another challenge is to keep UI consistent on all devices in real time. UI need to be changed according to internal or external triggers. In [14] is presented multi-access service platform which keeps the statement of UI actual on different devices per usage of services, templates, and internal UI model.

It is important to notice that human attention and awareness is limited. The design of UI should not provide overwhelming amount of information.

Implementation of SmE depends on parallel processing of recording, segmentation and analysis of acquired data, computational process and representation of outputs. Processes and connections are described in Fig. 4.

V. Usability evaluation

The user interfaces of SmE should be effective in variety of different activities in wide variety of contexts. Building efficient UI for such a complex system as smart environment is challenging, time consuming and costly process. Most of effort is aimed at lowering resources without impact on quality of UI. There are several known and used approaches and methods for UI usability evaluation. [15]

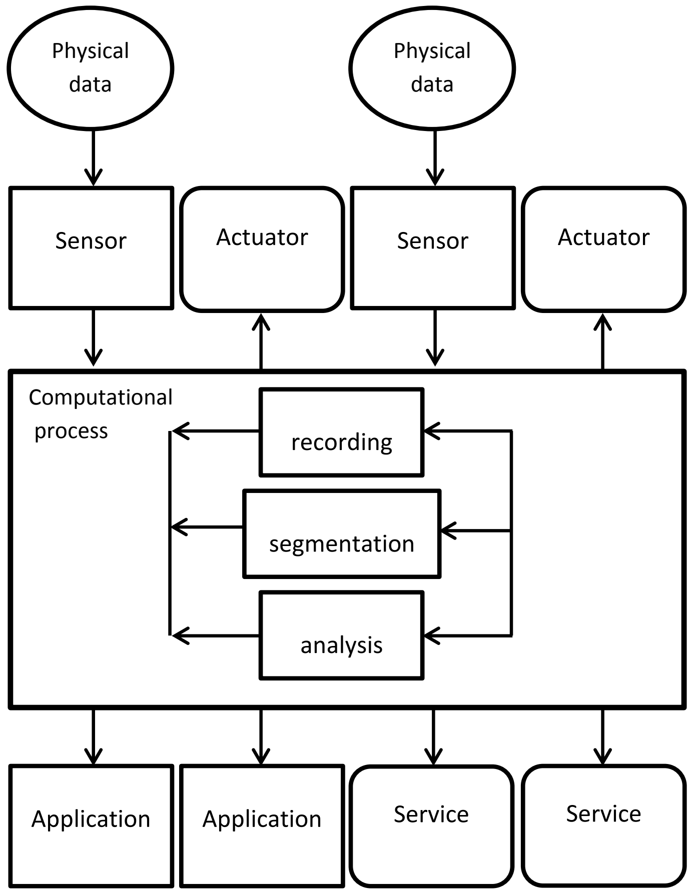


Fig. 4 Processes and connections of smart environments

A. *Model-based approach*

Model-based evaluation is using models to describe human behavior in interaction with the system. This type of evaluation could be used to replace or supplement empirical testing. Model-based evaluation could bring appropriate results in early stages of UI development in case that implementation of components is incomplete. Usability tests are in the form of calculation and simulations. This approach represents engineering or analytic models to building UI. There are three main types of models according to the level of analysis. [16]

Task network model is used to describe the human behavior in the use cases of developed system. All operations should be decomposed into component tasks which can be also further decomposed. The model is usually presented in the graphical form of PERT charts. Processes are a set of nodes and connections between them represents the control flow. Tasks in the same level of hierarchy could be performed in the series or in parallel. There is not fixed level of model abstraction but achieved results are more useful in engineering work considering human factors than to studies of cognition. Despite known cases when some aspects of human behavior are considered like workload measurement, memory requirements and Fitts law, cognition is not usually represented. Determination of the task order is also based on randomness, probabilistic or scenarios and developed conditional rules. [17]

Cognitive models integrate theories of psychological mechanisms of complex learning, cognition and human performance. These models assume that human activity is purposeful, adaptive and computed by information processing mechanisms realized by physical and biological processes. Cognitive architecture implemented in connection of usability tests is in the form of computer programs which represent agents or virtual users operating in simulated task environment. The effort is aimed at creation of complex behavioral models which are used in a domain of basic research projects and at current state of art there is very limited experience of their use in actual design settings. [18]

GOMS models represent reduction methods of user's interaction into physical, cognitive or perceptual actions. Through determined elementary actions (operations) is UI further studied according to goals which are achieved with the use of methods and selection rules [16]. GOMS models produce valuable evaluation data based on computation in short period of time with low cost and effort. UI is described by the knowledge of procedures that the user must have to operate the system. Despite several known techniques of building up these models there is not sufficient method to include user behavior affected by social surrounding, fatigue and environment. Evaluation also works in conditions typical for expert users without considering of novices.[18]

B. Inspection based approach

Inspection based evaluation approach stand in the opposite of empirical testing with target users. There are several methods including heuristic evaluation, cognitive walkthrough, pluralistic walkthrough and feature, consistency and standard inspection which could be applied on all systems without further specification. Most of them rely on experts which evaluate UI in terms of effectivity, efficiency and user's satisfaction. Advantage of usability evaluation based on inspection is the usage in the stage of design. Mentioned method of evaluation doesn't produce statistical outputs and quality strictly depends on experience of evaluators. [19]

C. Empirical testing

Empirical evaluation is usually performed by real users with different levels of expertise. This method includes choosing tasks which will be tested in the work with the system and UI properties which will be evaluated. During a testing are user's actions or behavior in general including facial movements recorded and analyzed. Final part of testing is spontaneous thinking aloud which could be source of major improvements. Evaluation based on empiricism is usually in final stages of UI development because of high demands on time and resources. [20]

Diversity of approaches to UI development for SmE requires combination of known evaluation methods.

VI. Conclusion

This paper introduces various challenges and current approaches connected with design, implementation and evaluation of user interfaces for smart environments. The user interface strictly depends on other properties of the SmE and their improvements. Because of this fact authors rely on prototyping using virtual reality. Desirable outcomes could highlight benefits of SmE in the meaning of communication with handicapped people. Identified properties of complex user interfaces and insufficiencies of known solutions become the subject of further research in this area.

Acknowledgment

This work was supported by the KEGA grant no. 083 TUKE-4/2015 “Virtual-reality technologies in the process of handicapped persons education”.

References

[1] J. C. Augusto et. al., “Intelligent Environments: a manifesto”, Human-centric computing and Information Sciences, 2013, 3:12.

[2] D. J. Cook – S. K. Das, “How smart are our environments? An updated look at the state of the art,” Pervasive and Mobile Computing, Vol. 3, pp. 53-73.

[3] D. J. Cook – S. K. Das, *Smart Environments: Technologies, Protocols and Applications*, John Wiley & Sons, Inc., New Jersey: Hoboken, 2005, ISBN 0-471-54448-5 .

[4] M. Broxvall et. al., “PEIS Ecology: Integrating Robots into Smart Environments,” in *Robotics and Automation,* Orlando, 2006, pp. 212-218

[5] S. Coradeschi – A. Saffiotti, “Symbiotic Robotic Systems: Humans, Robots, and Smart Environments,” in *IEEE Intelligent systems,* 2006, vol. 21, pp. 82-84.

[6] M. Weiser, “The origins of ubiquitous computing research at PARC in the late 1980s”, in *IBM Systems Journal,* 1999, pp. 693-696.

[7] Mac Basic: Multi-Touch gestures [online]. [Accessed 10.3.2015]. Available on: https://support.apple.com/en-us/ht4721

[8] S. Bartolini, Smart Sensors For Interoperable Smart Environment, PhD. Thesis, The University of Bologna, 2009, 248 p.

[9] L. Jacho, “Virtual control desk based on virtual reality systems – visualization part”, Master thesis, Technical university of Košice, 2014, 75 p.

[10] J. Ch. Lee, “Spatial User Interfaces: Augmenting Human Sensibilities in a Domestic Kitchen”, Master thesis, National Cheng Kung University, Taiwan, 2005, 116 p.

[11] B. Taylor, V. M. Bove, “Graspables: Grasp-Recognition as a User Interface,” in *CHI 2009 Techniques for Mobile Interaction,* Boston, 2009, pp.917-925.

[12] B. Jou – S. Bhattacharya – S. Chang, ‘’Predicting Viewer Perceived Emotions in Animated GIFs”, in *ACM MM Grand Challenge MM,* 2014, Orlando, US

[13] V. Schwartze – S. Feuerstack – S. Albayrak, “Behavior-sensitive User Interfaces for Smart Environments” in *Digital Human Modeling,* Vol. 5620, 2009, pp 305-314

[14] M. Blumendorf – S. Feuerstack – S. Albayrak, “Multimodal User Interaction in Smart Environments Delivering Distributed User Interfaces” in *Constructing Ambient Intelligence,* 2008, vol. 11., pp. 113-120.

[15] C. Santoro, *A Task Model-Based Approach for the Design and Evaluation of Innovative User Interfaces,* Presses universitaires de Louvain, Louvain la Neuve, Belgium, 2005, ISBN: 2-930344-86-5

[16] D. Kieras, “Model-based Evaluation” in J. Jacko & A. Sears (Eds.), *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications,* 2nd Ed., Mahwah, New Jersey: Lawrence Erlbaum Associates, 2007, ISBN-13: 978-0805858709

[17] W. Warwick et. al., “Integrating Architectures: Dovetailing Task Network and Cognitive Models” in *Conference on Behavior Representation in Modeling and Simulation,* 2008, ISBN: 978-1-60560-124-3

[18] P. Pirolli, “Cognitive Architectures and Cognitive Engineering Models in Human-Computer Interaction” in F. T. Durso (Ed.) *Handbook of Applied Cognition*, John Wiley and Sons., 2007, ISBN: 978-0-470-01534-6

[19] Jacob Nielsen: Usability Inspection Methods [online]. [Accessed 12.3.2015]. Available on: http://www.sigchi.org/chi95/proceedings/tutors/jn_bdy.htm

[20] M.Y. Ivory – M.A. Hearst, “The State of the Art in Automating Usability Evaluation of User Interfaces” in *ACM Computing Surveys,*2001, vol. 33, no. 4, pp. 471-516

Contribution to Analysis of High Frequency Elements from the View of Different Materials and Technologies Applications'

[1]*Tibor ROVENSKÝ (2^{th} year)*
Supervisor: [2]Alena PIETRIKOVÁ

[1,2]Dept. of Technologies in Electronics, FEI TU of Košice, Slovak Republic

[1]tibor.rovensky@tuke.sk, [2]alena.pietrikova@tuke.sk

***Abstract*— This paper is summarization of last year of post gradual study. Solved research tasks and published articles are presented here. This work involves design and measurements of microstrip narrow-band notch filter on LTCC substrate Murata LFC; stability of LTCC substrates in high frequency area after accelerated aging testes; impact analysis of LTCC materials in microstrip filters' behavior up to 13 GHz; correlation between simulations and real measurements of microstrip filters based on LTCC in high frequency area and usability of various LTCC in in microstrip filters construction.**

***Keywords*—dielectric properties, filters, LTCC, microstrip structure, scattering parameters**

I. Introduction

With evolution of information technology, the frequency applied to electronic devices has become higher to handle the growing volume dat. These devices are used in various types of environments (temperature, humidity, etc.) and they are loaded by high frequencies.

Because of high frequencies it is not possible to use typical electronic elements (resistors, capacitors, inductors), for that reason all devices for high frequency applications are designed and constructed using planar transmission lines, mostly by microstrip lines. Microstrip structures' design is strictly depending on dielectric properties, dielectric constant and loss tangent, of used substrate.

The most significant problems caused by higher frequencies are losses as well as stability and reliability of mentioned dielectric properties. For that reason it is necessary to find out behavior, stability and reliability of dielectric substrates in different environments loaded by high frequency [1].

II. Initial Status

Nowadays, there are a lot of PCB and LTCC dielectric substrates available on the market but not all of them are suitable for high frequency applications. Only few of these substrates were designed and have sufficient properties for high frequency applications, which means they have high dielectric constant and low loss tangent. However, in general each application has unique requirements which have to be fulfilled. Except suitable dielectric properties their stability in high frequency area and different environment is very important. For these reasons it is necessary to find suitable dielectric substrates and characterized their behavior, because no research in this area was done so far.

After studying theoretical background about complex permitivity, dielectric polarizations and losses [2] we want to solve these PhD theses:

1. Design methodology for measurements of dielectric properties in high frequency area.
2. Dielectric properties analyses' of PCB and LTCC materials using planar structures in high frequency area.
3. Microscopic analyses of crystallization process in various processing conditions' and their impact on dielectric properties.
4. Analyses of accelerated ageing on dielectric properties of PCB and LTCC materials.
5. Applique analyses results' in fabrication of selected high frequency element based on PCB and LTCC.

III. Solved Tasks in The Previous Year

A. Design and Measurements of Microstrip Narrow-band Notch Filter Based on LTCC Substrate Murata LFC

Microstrip narrow-band notch filter from 2.4 GHz to 2.5 GHz were designed, simulated, fabricated and measured. Return losses and forward transmission coefficient were analyzed and compared with simulation and requirements. Fabricated filter has attenuation only -37.9 dB (required -40 dB) and stop band from 2.32 GHz to 2.52 GHz (required from 2.4 GHz to 2.5 GHz) but in spite these facts it can be used in UWB applications. On the other hand it is necessary to investigate influence of filter with quite different than required performance on UWB radar systems' accuracy.

B. Stability of LTCC Substrates in High Frequency Area after Accelerated Aging Tests

Microstrip structures using two different LTCC substrates were fabricated and loaded by accelerated aging test. Before starting tests and after each testing sessions measurements were done. All results were plotted in graphs and analyzed.

Small changes of dielectric properties could lead to stopband's enlargement and degradation of usable signal. HTOL tests have no significant influence on LTCC substrates in comparison with temperature cycling test which changed properties of dielectric LTCC substrates and behavior of microstrip structures (enlargement of primary stopband). Comparison between DuPont GreenTape 951 and 9K7 showed good suitability and stability for HF application loaded by thermal stress in spite of fact that DuPont GreenTape 9K7 was designed especially for HF applications. Importance of pastes conductivity was proved in this article. This work will be basic for future microstructure analyses.

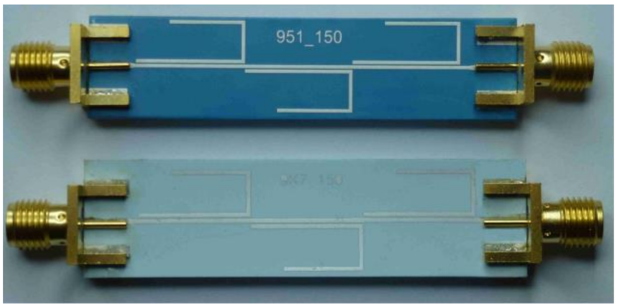


Fig. 1. Fabricated microstrip bandstop filter with L-resonators based on GreenTape DuPont 951 (up) and GreenTape DuPont 9K7 (down)

C. *Impact Analysis of LTCC Materials on Microstrip Filters' Behaviour up to 13 GHz*

We have presented 3 types of microstrip filters (LP, BP, BS) based on various LTCC dielectrics due to meet the aim to investigate the impact of materials on filters' behavior in UWB frequency application. Relatively high dielectric constant of LTCC (GT 951PX, GT 9K7 PX, Murata LFC) enables a significant reduction in the module dimensions. Except size reduction, the proposed filters fabricated using different LTCC substrates are characterised in low dielectric losses. It is not possible to unambiguously determine the most suitable LTCC dielectric for these filter design but temperature-stable dielectric properties of Murata LFC make them a promising ceramic for HF application (repeatability of realised experiments). The novelty of this work lies in unconventional using of LTCC as material with defined dielectric properties proper for HF applications. Processing demands and low dielectric losses are the most essential properties in material selection. This paper confirms suitability and possibility for high frequency applications, especially for filter design.

D. *Correlation between Simulations and Real Measurements of Microstrip Filters Based on LTCC in High Frequency Area*

Three types (LP, BP, BS) of filters were designed, simulated, fabricated and measured. Simulated and measured forward transmission coefficients were plotted in joint graphs to analyse their mutual correlation.

Simulations provided by Ansoft Designer are very ideal and do not correspond with real measurements, this distortion could influence performance of devices in very important manner, on the other hand, simulations provided by HyperLynx 3D EM Designer are in correlation with measured results, in general for all three types of filters. For these reasons HyperLynx 3D EM Designer is considered to be more suitable for design of microstrip structures in HF area.

E. *Usability of Various LTCC in Microstrip Filters Construction*

Introduction of LTCC type 9K7 by DuPont and Murata LFC as a substrate into the HF application brings a new quality due to good microwave dielectric properties. In our research, we tried to find out a method of shrinkage instability elimination as well as a correlation among shrinkage and mechanical properties. It is well known that LTCC shrinks in the three directions of length, width, and thickness during firing. Shrinkage of about 15 % occurs in each direction. Murata demonstrated an exception to this rule. Murata LFC unlike DuPont LTCC (DuPont 951, DuPont 9K7) retains its physical dimensions in the x and y directions and restricts the ceramic shrinkage to only the z-direction (thickness). Murata offers non-shrinkage substrates using its proprietary pressured non-shrinkage firing method that provides superb dimensional accuracy and flatness. Minimal shrinkage of Murata provides excellent electrical characteristics because of its use of low dielectric ceramic material. In our experiments we omit the pressure assisted firing method that is typical for Murata and instead of this we implemented various combinations with DuPont LTCC. We found out that combination of Murata with DuPont 9K7 LTCC in a three layer system brings new quality with near elimination of shrinkage. Murata's LTCC substrates treated by conventional firing process could be accepted in HF applications in combination with DuPont 9K7.

IV. FUTURE (WORK)

In future work we will focus on design, simulation and fabrication of microstrip resonators based on different LTCC substrates (Green Tape 9K7X, Green Tape 951PX, Murat LFC) and PCB (Rogers RO30003C, Rogers RO4003C, Rogers RT/duroid 5880 and Rogers RT/duroid 6010LM).

Fabricated microstrip resonators will be measured. From measured scattering parameters dielectric properties (dielectric constant, loss tangent) of substrates will be calculated in wide frequency range from DC up to 13 GHz.

REFERENCES

[1] Jia-Sheng Hong, *"Microstrip Filters for RF/Microwave Applications"*, John Wiley & Sons, Inc., Hoboken, 2nd edition, New Jersey, 2011, 655.

[2] Agilent: „Basic of Measuring the Dielectric Properties of Materials“, 2013.

[3] T. Rovenský, A. Pietriková, K. Ruman, "Design and measurements of microstrip narrow-band notch filter based on LTCC substrate Murata LFC", In: Electrical Engineering and Informatics 5 : Proceedings of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice. - Košice : TU, 2014 S. 451-454. - ISBN 978-80-553-1704-5.

[4] T. Rovensky et al., "*Stability of LTCC substrates in high frequency area after accelerated aging tests*", In: ISSE 2014: 37th International Spring Seminar on Electronics Technology, Dresden: Piscataway : IEEE, 2014. P. 221-224. ISBN 978-1-4799-4455-2.

[5] A. Pietrikova, K. Ruman, T. Rovensky, I. Vehec, "*Impact Analysis of LTCC Materials on Microstrip Filters' Behaviour up to 13 GHz*", In: Microelectronics International, ISSN: 1356-5362, will be published.

[6] K. Ruman, T. Rovensky, A. Pietrikova, "Corelation between simulations and real measurements of microstrip filters based on LTCC in high frequency area", In: Acta Electrotechnica et Informatica, 2015, will be published.

[7] A. Pietriková, K. Ruman, T. Rovenský, P. Lukács," Usability of Various LTCC in Microstrip Filters Construction" 2014, Key Engineering Materials, 635, 61, will be published.

Contribution to Analysis of Layers Based on Silver Nanoparticles Realized by Inkjet Printing Technology

[1]*Peter LUKÁCS (2nd year)*
Supervisor: [2]Alena PIETRIKOVÁ

[1,2]Dept. of Technologies in Electronics, FEI TU of Košice, Slovak Republic

[1]peter.lukacs@tuke.sk, [2]alena.pietrikova@tuke.sk

***Abstract*—This paper is a summarization of last year of post gradual study. Solved research tasks and published articles are presented here. This work involves the theoretical background of inkjet printing technology, design and comparison between simulations and measurements of UWB antenna based on polyimide substrates and also the measurements of wettability, roughness and micro hardness of most used substrates in inkjet printing technology, such as polyimide, polyester and aramid paper.**

***Keywords*—Inkjet printing, polyimide substrate, surface treatment, UWB antenna.**

I. Introduction

The unstoppable development of nanotechnology brings new possibilities in the area of electronics technology. One of these technologies is inkjet printing technology, which offers a lot of advantages, such as creation conductive, semi conductive, isolation or other function layers onto various flexible substrates. Described non-contact printing method presents the digital printing, which works with small ink quantities with very low viscosity based on nanoparticles of special materials. This printing technology is suitable especially for applications, where the precise printing with high accuracy is desirable [1].

Inks, used for inkjet printing technology may be divided into 3 groups. The first, most used inks are based on nanoparticles of conductive organic (PEDOT:PSS) or inorganic (silver, gold, copper) materials. The second group consists of semi conductive inks, which are based on carbon nanotubes, as well as on organic materials, such as PQT-12 or P3HT. The last group of inks serves to creating of isolation layers. For this purpose, organic (PVP, PMMA) and inorganic (Zr, TiO_2, SiO_2) materials are used [2], [3].

II. Initial Status

Nowadays, printing with high accuracy is desirable mainly in high frequency area, especially in the case of antennas. The requirements to antennas are very strict, what includes a strict defined frequency band, high gain of antenna and last but not least, the size of antenna is also the very important parameter. Flexible polymeric substrates, such as polyimide film Kapton®, are usually used in high frequency area because of their suitable dielectric properties and also good temperature resistance [1], [5].

Printing the structures and lines with width less than 50 µm is extremely difficult. There are many technological factors that affect the quality of printed structures. For this purpose, the quality of surfaces is needed to analyze. The roughness, wettability and surface tension of substrates play a significant role to drop spreading. Surface treatment represents an important step to optimize the technological process of printing [4].

After studying theoretical background about inkjet printing technology, substrates' surfaces and their modifications as well as drop spreading we want to solve these PhD theses:

1. Optimization of technological factors affecting the structure of silver layers realized by inkjet printing.
2. Analysis of adhesion mechanism between layers based on silver nanoparticles and substrates and proposal for methodology of their measurement.
3. Analysis of properties of silver layers realized by inkjet printing technology and analysis the impact of various parameters to the quality.
4. Application of analysis' results for the design of 2 selected electronic components on a flexible substrate by inkjet printing technology.

III. Solved Tasks in The Previous Year

Tasks which are summarized in the following section were solved in the last year of postgraduate study based on dissertation thesis.

A. Analysis of Substrate Surfaces

For the purpose of optimization process of inkjet printing technology, we realized the surfaces' analysis of polymeric substrates, as well as Al_2O_3.

Fig. 1 shows the roughness of the silver layer deposited by inkjet printing technology onto polyimide substrate.

The roughness measurement of substrates and silver layers was realized by 3D optical profiler Senofar Senso Scan Neox. For analysis of the adhesion mechanism, scratch test was also applied. For analysis the substrate's wettability, the contact angle of several types of polymeric substrates, papers and Al_2O_3 was also measured.

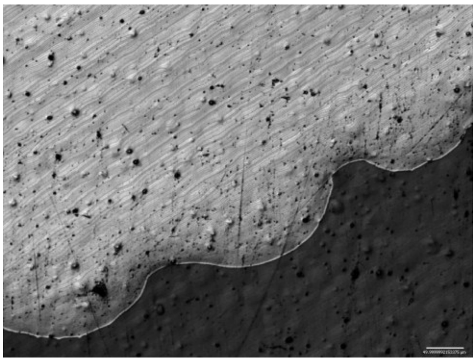
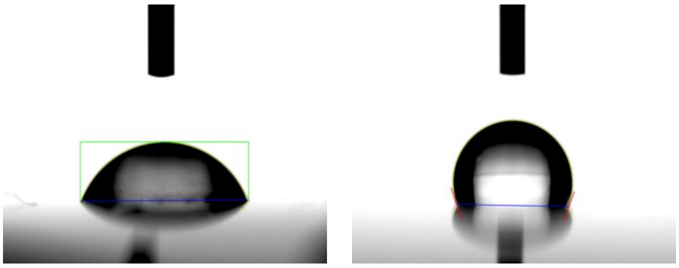


Fig. 1. Surface of silver layer after sintering process printed on polyimide substrate Kapton® HN.

The results of contact angle measurements show, that the contact angle of polyimide foils Kapton® HN without surface treatment is about 60°. After applying the chemical agent EGC-1720 to the substrate, the contact angle grows up to 110°.



a) b)

Fig. 2. Contact angle measurement a) Kapton® HN without surface treatment, b) Kapton® HN treated by EGC-1720.

B. *Design and Realization of UWB Antenna on Flexible Substrate by Inkjet Printing Technology*

The description of design (Fig. 3) and realization of UWB antenna on polyimide substrate Kapton® HN by inkjet printing technology is presented here. The mentioned antenna works in frequency bandwidth from 5.5 to 11.7 GHz with gain of 2.4 dB. The printing of antenna was realized by PixDro PL50 inkjet printer with print head Spectra® SE-128 AA by using conductive ink based on nanoparticles of silver AX JP-6n by Amepox Microelectronics LTD. The impedance matching to 50 Ω was realized by length adjustment of feeding element to λ/4 [5].

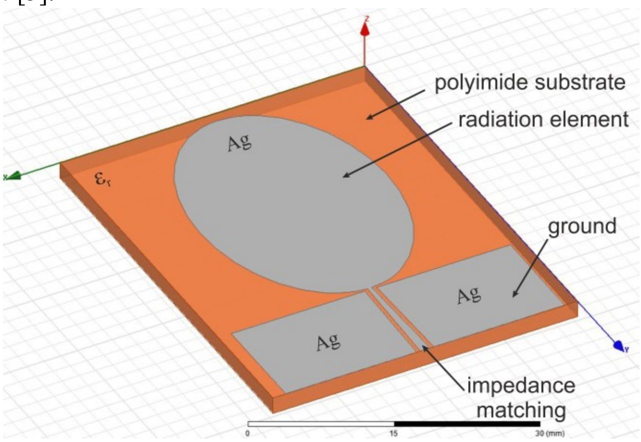


Fig. 3. Antenna design.

The reflection coefficient S_{11} of antenna was measured by vector network analyzer Agilent PNA in range of 0 – 13.5 GHz and it is illustrated in Fig. 4, which shows, that the bandwidth of simulated and realized antenna is very similar and it is more than 6 GHz, from 5.5 to 11.7 GHz. The dielectric constant of DuPont™ Kapton® HN is 3.4 with thickness 50.4 μm [5].

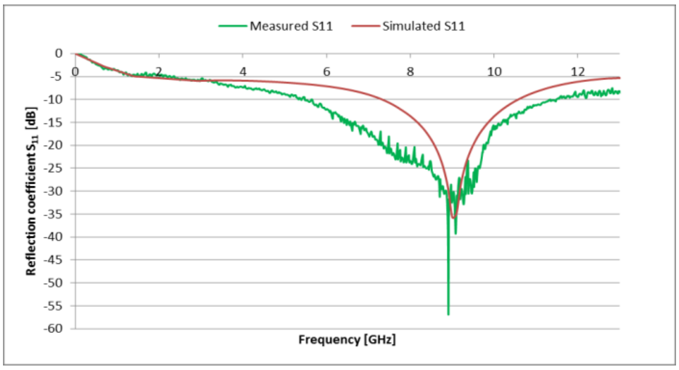


Fig. 4. Measured and simulated reflection coefficient S_{11} of realized UWB antenna.

IV. Future Work

In future work we will focus on surface treatment of several kinds of polymeric substrates by chemical agents and on their analysis. After optimization of substrates' surfaces, the obtained results will be implemented into the inkjet printing technology with respect to PhD thesis.

Awards

The Best Poster Award for Young Scientists for the poster "Analysis of Mechanical Properties of LTCC Substrates", 37th International Spring Seminar on Electronics Technology, 07-11th May, 2014, Dresden, Germany.

The Dean's Award for The Best Presentation in section Electrical & Electronics Engineering, 1st year, 14th Scientific Conference of Young Researchers, 20th May, 2014, Herľany, Slovakia.

Acknowledgment

This work was supported by project VEGA: 1/0218/13.

References

[1] P. Lukacs, "Active Circuit Components of UWB Radar Based on Organic Electronics", In: *SCYR 2014: Proceedings from conference: 14th Scientific Conference of Young Researchers*: May 20th, 2014, Herľany, Slovakia. – Košice: TU, 2014 p. 19-22. ISBN 978-80-553-1714-4.

[2] J. Daniel, "Printed Electronics: Technologies, Challenges and Applications," *International Workshop of Flexible & Printed Electronics*, Korea, 2010.

[3] J. Felba, H. Schaefer, "Materials and technology for conductive microstructures W: Nanopackaging: nanotechnologies and electronics packaging." Ed. James E. Morris. New York, NY: Springer, 2008, pp. 239-263.

[4] G. Tomaszewski, J. Potencki, T. Wałach, Mariusz Pilecki, "Investigation of ink spreding on various substrates in inkjet technology", *38th International Conference of IMAPS-CPMT*, Rzeszów-Czarna, 4 p., 2014.

[5] P. Lukacs, A. Pietrikova, J. Potencki, G. Tomaszewski, "UWB Antenna Based on Nanoparticles of Silver on Polyimide Substrate", *38th International Spring Seminar on Electronics Technology*, Eger, Hungary, May 6-10th, 2015. Will be published.

Contribution to Analysis of Microstrip Filters for UWB Sensor Systems Based on LTCC

[1]*Kornel RUMAN (4th year)*
Supervisor: [2]Alena PIETRIKOVÁ

[1,2]Dept. of Technologies in Electronics, FEI TU of Košice, Slovak Republic

[1]kornel.ruman@tuke.sk, [2]alena.pietrikova@tuke.sk

***Abstract*—This paper describes the tasks and obtained results solved in the previous year of post gradual study. The work is focused on the possibilities of using Low Temperature Co-fired Ceramics (LTCC) technology in the high frequency (HF) area. This paper deals with design, simulation and integration of microstrip Band-Pass (BP) and Low-Pass (LP) filters into one multilayer package based on LTCC. Article investigate the impact of interlayer vias on microstrip LP filters' performances in the HF area. This paper assesses the suitability of LTCC substrate Murata LFC for the production of the microstip stepped-impedance LP filter for HF area. Article demonstrates the compatibility of LTCC substrate Murata LFC with conductor system LL612 Ag designed for GreenTape 9K7. LP, BP and band stop (BS) filters for Ultra Wide-Band (UWB) radar systems were designed, simulated, fabricated and measured using three various dielectric substrates: DuPont GreenTape 951, DuPont GreenTape 9K7 and Murata LFC. We introduce a design of multilayer structure which includes all parts of I – Q demodulator based on LTCC substrate Green Tape 951PX.**

***Keywords*—Low-Pass, Band-Pass, LTCC, High Frequency.**

I. Introduction

The rapidly growing field of UWB applications in various areas pushes the requests for new enhanced UWB radar systems. One of the very promising solutions for the UWB device realization is based on so called M-sequence approach, where the operation is based on a special type of the M-Sequence. This UWB sensor system is simply expandable with I – Q demodulator on side of receiver and mixer on the transmitting side [1]. I – Q demodulation is very useful if we need to obtain both the magnitude and phase of the received signal. Our sensor system is designed in order to match the frequency bandwidth set by ECC organization [2]. Therefore, it is necessary to select received signals at I – Q demodulator inputs by BP filters. Mentioned UWB sensor system operates in frequency band from DC to 2 GHz, for that reason it is necessary to filter signals at I – Q demodulator outputs by the LP filters.

II. Initial Status

Nowadays, filtering in I – Q demodulator is provided by commercial fabricated BP and LP LTCC filters (BFCN-7900+ and LFCN-3400+) from Mini-Circuits Company [3]. However, these filters do not achieve the required properties (cut off frequency and attenuation in stop band). Because the pass band of commercial fabricated BP filter is only 0.3 GHz (from 7.8 GHz to 8.1 GHz) [4] the operating bandwidth (2.5 GHz) set by ECC is not fully utilized. Commercially fabricated LP filter have pass band from DC to 4.1 GHz [5] which is 2.1 GHz more than is required. This causes aliasing and leads to unobserved Nyquist sampling theorem. To solve the problem with filtering in I – Q demodulator, we have specified the following dissertation thesis:

1. Design, measurement and properties evaluations of thick film antialiasing LP prototype filter constructed on LTCC substrate.
2. Design, measurement and properties evaluations of thick film BP prototype filter constructed on LTCC substrate with bandwidth set by ECC organization.
3. Integration of LP and BP prototype filters mentioned in point 1 and 2 to multilayer structure based on LTCC substrate.
4. Study and analysis of 3D technological possibilities on LTCC substrate and assess its suitability for the construction of I - Q demodulator prototype for UWB sensor system based on LTCC substrate.

III. Solved Tasks in The Previous Year

Tasks which are summarized in the following section were solved in last year of postgraduate study based on dissertation thesis.

A. *Integration of Microstrip LP and BP Filters to Multilayer Structure Based on Various LTCC*

According to the simulations multilayer structures with integrated filters (BP and LP) using various LTCC materials were fabricated and measured. We've demonstrated impact of interlayer vias on LP filters' performances and suitability of different LTCC substrates for UWB sensor system (based on M – sequence). Measurement of LP filters where signal is transmitted through vias and where signal is transmitted straight to LP filters shows that vias with fine quality improve properties of transmitted signal. We point out that quality of interlayer vias have major impact on LP filters' performances. We have shown that the more suitable LTCC material for microstrip LP and BP filters for UWB sensor system (frequency range from DC to 13 GHz) is LTCC substrate GreenTape 951PX (despite to bigger value of loss tangent). On the other hand the transmission characteristics of microstrip filters based on both LTCC types of substrates meet requirements and are acceptable for use in I – Q demodulator for UWB sensor system. This project was preliminary for manufacturing multilayer package of I – Q demodulator with all his parts based on LTCC substrate [6].

B. Design of Low Pass Filter Based on Substrate Murata LFC for Ultra Wide-Band Sensor System

The microstrip stepped-impedance LP filters with cut off frequency 2.5 GHz and a minimum attenuation -40 dB in stop band were designed, simulated and experimental tested. The layouts of LP filters were created by the HyperLynx 3D EM Designer (from Mentor Graphics) and simulated with full-wave electromagnetic simulation. Comparisons between simulated and measured S parameters of designed microstrip stepped-impedance LP filters were presented. We point out that conductor system designed for GreenTape 9K7 (DuPont LL612) is compatible with LTCC substrate Murata LFC. We verified the suitability of the LTCC material Murata LFC for the production of the microstrip stepped-impedance LP filters for high frequency area. Transmission characteristics of microstrip LP filters are acceptable for use in the UWB sensor system as an antialiasing LP filters for I – Q demodulator [7].

C. Impact Analysis of LTCC Materials on Microstrip Filters' Behaviour up to 13 GHz

We have presented 3 types of microstrip filters (LP, BP, BS) based on various LTCC dielectrics due to meet the aim to investigate the impact of materials on filters' behavior in UWB frequency application. Relatively high dielectric constant of LTCC (GT 951PX, GT 9K7 PX, Murata LFC) enables a significant reduction in the module dimensions. Except size reduction, the proposed filters fabricated using different LTCC substrates are characterised in low dielectric losses. It is not possible to unambiguously determine the most suitable LTCC dielectric for these filter design but temperature-stable dielectric properties of Murata LFC make them a promising ceramic for HF application (repeatability of realised experiments). The novelty of this work lies in using various LTCC as active part of electronic devices which gives LTCC another option and advantage in HF applications. Processing demands and low dielectric losses are the most essential properties in material selection. This paper confirms suitability and possibility for high frequency applications, especially for filter design [8].

D. Correlation between Simulations and Real Measurements of Microstrip Filters based on LTCC in High Frequency Area

Three types (LP, BP, BS) of filters were designed, simulated, fabricated and measured. Simulated and measured forward transmission coefficients were plotted in joint graphs to analyse their mutual correlation. Simulations provided by Ansoft Designer are very ideal and do not correspond with real measurements, this distortion could influence performance of devices in very important manner, on the other hand, simulations provided by HyperLynx 3D EM Designer are in correlation with measured results, in general for all three types of filters. For these reasons HyperLynx 3D EM Designer is considered to be more suitable for design of microstrip structures in HF area [9].

E. Modified I – Q Demodulator for M-Sequence UWB Sensor System Based on LTCC

We have developed new multilayer I – Q demodulator with embedded microstrip LP and PB filters based on LTCC dielectric substrate Green Tape 951PX (Fig.1). Multilayer I – Q demodulator was designed, simulated and experimental tested. We point out that developed microstrip LP and BP filters achieved better transmission characteristic in comparison with commercial available LTCC SMD LP [5] and BP [4] filters. We verified the suitability of the LTCC material GreenTape 951 for the production of I – Q demodulator. We've demonstrated impact of interlayer vias on microstrip LP filters' performances. Measurement of microstrip LP and BP filters shows that filters based on microstrip structures are suitable for HF area (up to 10 GHz). We presented advantages of this new "LTCC" approach for the construction of multilayer I – Q demodulator with embedded microstrip filters [10].

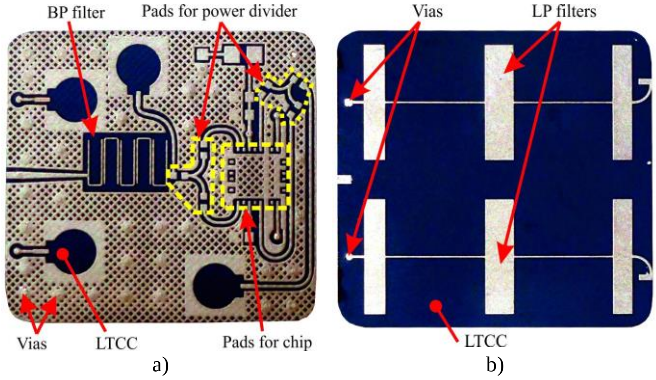


Fig. 1. Multilayer I – Q demodulator: a) top side b) bottom side

References

[1] J. Sachs, M. Kmec, R. Zetik, P. Peyerl, and P. Rauschenbach, "Ultra wideband radar assembly kit," Geoscience and Remote Sensing Symposium, 2005, iGARSS '05. Proceedings. 2005 IEEE International.

[2] ECC Decision (06)04, "The harmonised conditions for devices using Ultra-Wideband (UWB) technology in bands below 10.6 GHz," 2006. [Online]. Available: http://www.erodocdb.dk/docs/doc98/official/pdf/ECCDEC0406.pdf.

[3] LIPTAJ, Martin: Návrh ASIC obvodov pre širokopásmové aplikácie: Dizertačná práca. Košice: TUKE KEMT, 2012. 122 s.

[4] LTCC Bandpass Filter, "LTCC Bandpass Filter BFCN-7900+, 50Ω, from 7800 to8100 MHz," 2015. [Online]. Available: http://194.75.38.69/pages/npa/BFCN-7900+_NPA.pdf.

[5] Ceramic Low Pass Filter, "Ceramic Low Pass Filter LFCN-3400 +, 50Ω, from DC to 3400 MHz," 2015. [Online]. Available: http://194.75.38.69/pdfs/LFCN-3400+.pdf.

[6] K. Ruman, A. Pietrikova, I. Vehec, T. Rovensky and P. Galajda, "Integration of microstrip LP and BP filters to multilayer structure based on various LTCC", In: ISSE 2014: 37th International Spring Seminar on Electronics Technology, Dresden: Piscataway : IEEE, 2014. P. 114-119. ISBN 978-1-4799-4455-2.

[7] K. Ruman, A. Pietrikova, I. Vehec, T. Rovensky, "*Design of low pass filter based on substrate Murata LFC for ultra wide-band sensor system*", In: Electrical Engineering and Informatics 5: proceedings of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice., Košice: FEI TU, 2014. P. 693-697. ISBN 978-80-553-1704-5.

[8] A. Pietrikova, K. Ruman, T. Rovensky, I. Vehec, "*Impact Analysis of LTCC Materials on Microstrip Filters' Behaviour up to 13 GHz*", In: Microelectronics International, ISSN: 1356-5362, will be published.

[9] K. Ruman, T. Rovensky, A. Pietrikova, "Corelation between simulations and real measurements of microstrip filters based on LTCC in high frequency area", In: Acta Electrotechnica et Informatica, 2015, will be published.

[10] K. Ruman, A. Pietrikova, I. Vehec, T. Rovensky and P. Galajda, "*Modified I – Q Demodulator for M-Sequence UWB Sensor System Based on LTCC*", In: ISSE 2015: 38th International Spring Seminar on Electronics Technology, 2015, will be published.

Cooperation of micro grid with distribution power system

[1]*Jozef DUDIAK (2nd year, external form)*
Supervisor: [2]Michal KOLCUN

[1,2]Dept. of Electric Power Engineering, FEI TU of Košice, Slovak Republic

[1]jozef.dudiak@student.tuke.sk, [2]michal.kolcun@tuke.sk

***Abstract*—this paper deals with facts about the regulation and management of micro grid as the part of smart grid. This paper also contains the overview on all my tasks solved in the last year period of time. There are some facts about the regulation of micro grid such as primary, secondary and tertiary control. The grid with implemented voltage level of 400 V in simulating program MODES is also mentioned there. The principle of regulation that is divided into 4 parts is discussed in chapter three. Finally there are aims for implementation and analyze for the next year.**

***Keywords*—Smart grid, micro grid, renewable energy sources, regulation of micro grid**

I. Introduction

Due to the increasing demand of electricity and the decreasing trend in fossil sources are micro-sources becoming an integral part of the energy mix of electricity sources. Micro grids are an effective tool in wide range of applications in cooperation with various renewable sources, which produce electricity in normal operation but also in island operation. Micro grids have great importance considering the enormous potential for application in remote areas. The new ideology of using renewables offers us the possibility to supply the load without financial compensation to the consumer. Renewables in micro grid usually consist of number of different sources, such as wind turbines, photovoltaic cells, together with the storage (eg.: battery, supercapacitor, etc.).[1]

As shown in [2], the direct benefits of micro grids fall into two categories: local benefits are resulting from the internal operation of the network and the wider benefits are from the ways in which micro grid cooperate with its associated distribution systems. Additional benefits include higher reliability of supplying the energy to customers in micro grid and beyond. Operation of micro grid indirectly provides many other benefits that are very difficult to define and quantify. From the environmental way it brings the reduction of emissions of greenhouse gases and other pollutants, the integration of renewable energy sources into the network, reducing the dependence on external sources, fuel prices as well as creation of new jobs.

The main challenge of these intelligent systems is control and ensuring the operation stability. The aim of my thesis is to research and analyze the principles of stable operation in micro grid with integrated wind, photovoltaic and other renewable energy sources, along with the battery systems, as the storage of energy. There are two control tasks that need to be analyzed and defined, that are interaction between micro grid and distribution system, as well as ensuring the operation inside the micro grid.

II. The initial status in the solving of the research task

The main task was to initialize the 400 voltage level to the program MODES and the implementation of battery system to the simulation program. The creator of the program made the patch for MODES which content energy storage systems and made some changes that accept 0,4 kV voltage level. Now we started to choose the area that is best for modeling the smart grid cases. The result of modeling should be cases of operational regulation of the micro grid. There are defined some cases for testing that are: micro grid without renewable, micro grid with renewable, micro grid with energy storage systems, island operation mode of micro grid. It is very important to implement the regulation of micro grid to the simulating program. The study and the analysis of micro grid regulation concept were done in the period of last year.

III. Description of the tasks solved in the previous year

The main task was finishing the written thesis to the dissertation exams. My thesis talks about various projects of smart grid in the world and in the Europe. The result is that many countries and governments invest a lot of money to research and development projects of smart grids. The projects are mainly about integration of renewable energy sources to the power system and their influence on the operation of the grid. There are some projects that talk about storage of energy produced during the generation peak of the renewables. The biggest group of realized projects is about smart metering activities in local distribution network all over the world. The aim of Europe is to implement more 80 % of smart meters to the grid before the year 2020. Asset of smart metering is in decreasing the losses of the grid and in the offering data from the meter to the customer who can adjust its consumption according the prize of the electricity. Slovak republic starts with demonstration project of smart metering. We have time to the end of April to analyze which technology is the best for our infrastructure. In general energy sector is now very interesting for new smart technologies and many companies offer products for testing the benefits of communication infrastructure.

The technical requirements of smart grids are also the part of my thesis. There are processed the functionality as the reliability and monitoring of electricity transmission, management of electricity consumers.

Building the micro grid with implemented renewables and energy storage systems should be the final result of this thesis. There are many captures that must be solved such as planning the grid operation, energy system management, smart substations – automation and protection, monitoring of equipment in substations, and the solution for smart metering.

During the last year we analyze the regulation of the micro grid. There are five levels of control.

A. Level 0: Internal control loops

The control of each device is included in the zero level, which belongs to the level of primary regulation. Its aim is to create linear and non-linear control loop sensing currents, voltages, feedback and prediction of events. These control loops allow regulation of output voltage and current to maintain the stability of the system.

B. Level 1: Primary regulation control

This level of control is necessary to ensure the voltage and frequency deviations within allowed limits of regulation. Deviations are formed in the inner current and voltage loop. The primary control ensures the fastest response to any changes in the production and consumption and increases the stability of the system operation.

C. Level 2: Secondary regulation control

It ensures that the electrical measurement parameters are maintained within the required limits. This level also provides a synchronization loop for easy connection the micro grid to the distribution power system and its disconnection in case of critical situation.

D. Level 3: Tertiary regulation control

This level of regulation ensures the regulation and management of power flow between micro grid and the distribution power system.

Micro grid with AC voltage can be operated in the mode with connection to the distribution power system and without that connection as the island operation. It is necessary to take in to account the possibility of control in grid operation mode and suggest the logical control of micro grid. It is important that micro grid should be able to disconnect from the rest of the system, but also the re-connection, respectively the restoration after complete power failure.

Operating mode of micro grid that is connected to the distribution system is based on rules and operating orders of the distribution system. The transition to island mode is restricted by intentional or unintentional event such as short circuit in the network. We need to define own algorithm in details to detect the transition to island mode. In fact it means that we set operational parameters in simulating program that specify the limits for safe operation of micro grid. In island mode micro grid must provide the desired active and reactive power as well as frequency stability and the voltage stability must me also ensured. Reconnection to the distribution power system requires compliance with phasing conditions. Active synchronization ensures monitoring voltage within allowable limits, frequency and phase angle in micro grid.

IV. Proposal for the next steps in the next period of time

There are some tasks to be solved in the next year time. First of all we need to initialize the primary regulation control on the power sources that are operating in the micro grid. After this step the grid will be prepared for initialization of secondary control. This regulation control is very important for maintaining the balance between production and consumption. The work on mentioned scenarios will be also the aim for the next year. We need to implement logic for cooperation the production from various energy sources together with energy storage systems. The whole regulations of micro grid are based on inverters that are placed on the output of the micro power plant. The regulation is so relative to every power plant and storage system. It means that every power plant connected to micro grid must be able to provide ancillary services. The finalization of my work should also contain some diagrams that are needed for optimization the operation of sources according the peaks that are during the day.

References

[1] BORLASE, S.: Smart Grid, Infrastructure, Technology and Solutions. Boca Raton: CRC Press, 2013. 577s. ISBN 978-1-4398-2905-9.

[2] SHAFIEE Q. et all, Distributed Secondary Control for Islanded Microgrids—A Novel Approach, IEEE transactions on power electronics, vol. 29, no. 2, February 2014.

[3] EKANAYAKE, J.B.: Smart grid: Technology and applications. Chichester: Wiley, 2012. 283s. ISBN 978-0-4709-7409-4.

[4] SAILATITHA, P.: SMART GRID. In: International Journal of Engineering Trends and Technology. roč. 4, 2013, č. 4. ISSN: 2231-5381.

[5] VIKASH, Kumar et al.: Frequency Control of smart Grid – A MATLAB/SIMULINK Approach. In: International Journal of Engineering Research & Technology. roč. 2, 2013, č. 10. ISSN 2278-0181.

[6] ALLARD, S.: Electric vehicles charging in a smart microgrid supplied with wind energy. In: PowerTech (POWERTECH), 2013 IEEE Grenoble, Grenoble: IEEE, 2013. ISBN 978-1-4673-5667-1.

[7] BAYLISS, Colin – HARDY, Brian.: Transmission and Distribution Electrical Engineering – Third Edition. Great Britain: Charon Tec Ltd, 2007. 1010s. ISBN 978-0-7506-6673-2.

[8] MORRIS, G. Y.: A Framework for the Evaluation of the Cost and Benefits of Microgrids, 2012.

[9] Abdelrahman O. I. Adam et all., Distributed Generation for Distribution Network Voltage Enhancement and Energy Loss Reduction, IEEE Innovative Smart Grid Technologies, Kuala Lumpur - Malaysia, May 2014.

[10] Palizban O. and Kauhauniemi K., Microgrid Control Principles in Island Mode Operation, PowerTech IEEE Grenoble, Grenoble, June 2013.

[11] Palizban O. et all., Microgrids in active network management—Part I: Hierarchical control, energy storage, virtual power plants, and market participation, Elsevier journal, January 2014, p. 428 - 439.

Design of High Frequency Converter for Induction Heating Based on Numerical Model

[1]*Matúš Ocilka (4th year)*
Supervisor: [2]Dobroslav KOVÁČ

[1,2]Dept. of Theoretical and Industrial Electrical Engineering, FEI TU of Košice, Slovak Republic

[1]matus.ocilka@tuke.sk, [2]dobroslav.kovac@tuke.sk

***Abstract*—The aim of this research is to design high frequency resonant converter for induction heating. In order to design the converter the parameters of the induction coil and workpiece needs to be calculated. For this purpose the simulation model of induction heating is presented using finite element method of solution. The model provides essential data for designing converter such as equivalent resistance and inductance of induction coil – workpiece coupling during heating process, amplitude and frequency of AC current to achieve desired temperature of workpiece. In order to minimize the switching losses of the tranzistors, an LCL resonant circuit with ZVS switching is designed. The control of the converter is made via Phase Locked Loop.**

***Keywords*—COMSOL Multiphysics, induction heating, LCL resonant circuit, PLL, ZVS**

I. Introduction

The induction heating (IH) is non-contact technique of heating electrically conductive materials. The induction heating system consists of source of high frequency AC current, induction coil and worpiece itself (Fig.1).

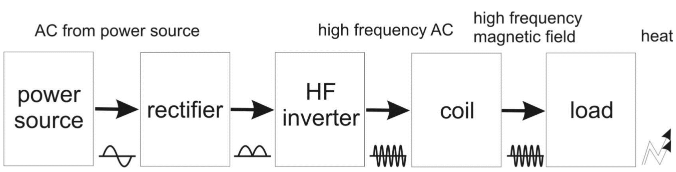


Fig. 1 Induction heating system

As its name implies, this technique relies on currents that are internally induced in workpiece, called eddy currents. The flow of eddy currents in electrically conductive material generates Joule's heat and thus the workpiece is heated. This is primary mechanism of heat generation in every electrically conductive material even non-magnetic. The second mechanism of heat generation is present only in magnetic materials and it's related with the hysteresis losses [1].

The design of induction heating includes knowing the processes ongoing in workpiece such us heat generation, heat transfer in material and heat transfer between the material and external environment. Electromagnetic and heat transfer phenomenon are highly interrelated because of the physical properties of heated metals. These properties are nonlinear functions of magnetic field intensity, temperature, chemical composition and many others [2].

Considering abovementioned facts, the IH phenomenon is a complex problem, thus numerical simulations seem to be a useful tool for the design and the investigation of IH systems [4].

The program used in design is called COMSOL Multiphysic [3]. This model considers the nonlinear properties of metal depended on temperature and allows to design and to observe the relations that are present in workpiece.

II. Initial State Of Research

The induction heating model was created in program COMSOL Multiphysics. In this case the workpiece is bolt M12x60 made from medium carbon steel. Parameters that have the main contribution in induction heating, thus used in simulation model are: relative permeability μ_r, electric resistivity ρ, thermal conductivity k and heat capacity C_p. The simulation result of this model is depicted in Fig. 2. The coil is supplied with current I_{max} = 50 A with frequency f = 100 kHz.

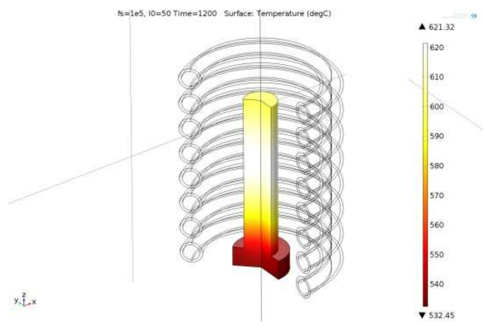


Fig. 2 Model of induction heating in COMSOL Multiphysics

In order to design the converter utilizing LCL resonant circuit, the equivalent resistance (R_{eq}) and equivalent inductance (L_{eq}) of induction coil - workpiece coupling needs to be calculated [5].

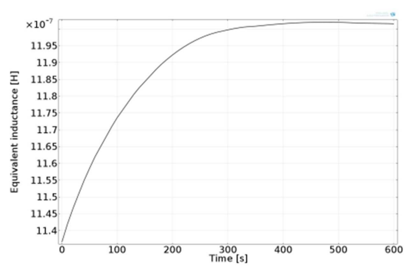


Fig. 3 Equivalent inductance L_{eq} of coil – workpiece coupling during 10 min heating

The results of R_{eq} and L_{eq} during heating process are depicted in Fig. 3 and Fig.4, respectively.

Next step is to design the inverter that feeds the coil with high frequency AC current.

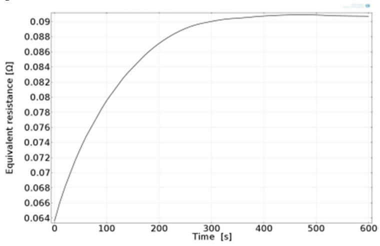


Fig. 4 Equivalent resistance R_{eq} of coil – workpiece coupling during 10 min heating

For our purpose the half-bridge converter with LCL resonant circuit is used. The topology of converter is depicted in Fig.5 [5].

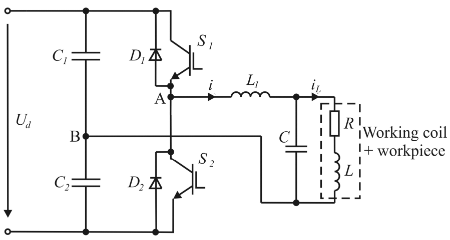


Fig. 5 Half-bridge resonant converter

III. Current State of Research

Induction coil is connected in parallel with capacitor, creating the parallel resonant tank. Adding inductor L_1 into the circuit ensures current lagging in the converter output and ZVS is achieved. Considering the LCL topology and constant values of L_1 and C, resonant frequency during heating does not remain constant and it is changing due to change of R_{eq} and L_{eq}. Calculated resonant frequency of this circuit lies within range f_r = 103.3 – 105.2 kHz. In order to achieve ZVS in whole range the control circuit with phase locked loop control is used [6], [7]. In resonance, phase shift φ between the converter output voltage and voltage of resonant capacitor is φ = 90°. The PLL control ensures that this phase shift is approximately 90° and sets the switching frequency of converter in that manner.

A. Experimental results

In the Fig. 6, the simulation and experimental results of surface temperature for I_{max} = 50 A are shown.

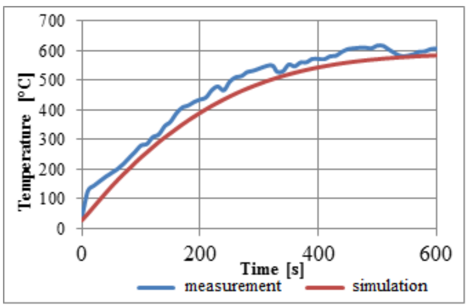


Fig. 6 Simulated and measured temperature of workpiece surface for I_{max} = 50 A

The measured temperature settled down after 450 s to value of 600 °C. On the other hand, the simulated temperature did not reach its maximum even after 600 s. Its value was 584 °C. The switching frequency of converter during heating is depicted in Fig. 7.

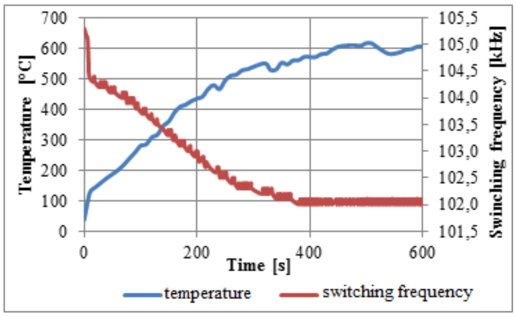


Fig. 7 Experimental measured switching frequency of converter during heating process for I_{max} = 50 A

The difference between the simulation model and measurement results of temperature in dynamic states is closing to 61%. On the other hand, in steady – state states is approximately 5%. This is caused by representation of numerical model and thermal properties of workpiece.

As can be seen in Fig. 7 the switching frequency of converter changes within range f_s = 102.0 – 105.3 kHz which approximately equals calculated f_r. The resonant frequency is interrelated with R_{eq} and L_{eq}, thus can be stated that presented numerical model brings accurate results and it is good approach in modelling similar problems.

IV. Future Research

Future work will be focused on implementing control scheme via PLL in digital signal processor, increasing the switching frequency and amplitude of current flowing through the coil and in the numerical model, all non-linear properties needs to be included.

Acknowledgment

The paper has been prepared under support of Slovak grant project KEGA No. 015TUKE-4/2015.

References

[1] S. L. Semiatin, S. Zin, “Elements of Induction Heating: Design, Control, and Applications,” ASM International, Materials Park, Ohio, USA, 1988

[2] V. Rudnev, D. Loveless, R. Cook, M. Black “Handbook of induction heating,” Marcel Dekker, Inc., New York, USA, 2003, ISBN 0-8247-0848-2

[3] COMSOL Multiphysics User’s Guide, Version 4.3a, Comsol AB, October, 2012

[4] H. Liu, J. Rao, “Coupled Modelling of Electromagnetic-Thermal Problem in Induction Heating Process Considering Material Properties, Information Engineering and Computer Science,” 2009. ICIECS 2009. International Conference on, vol., no., pp.1, 4, 19-20 Dec. 2009

[5] G. M. Segura, “Induction Heating Converter’s Design, Control and Modelling Applied to Continuous Wire Heating,” doctoral thesis, Barcelona, 2012

[6] M. Ocilka, “High frequency converter for induction heating” In: SCYR 2014 : 14th Scientific Conference of Young Researchers proceedings from conference : May 20th, 2014, Herľany, Slovakia. - Košice : TU, 2014 S. 54-55. - ISBN 978-80-553-1714-4

[7] M. Ocilka, D. Kováč, “Simulation and design of induction heating system,” In: SSIEE 2014 : Proceeding of scientific and student's works in the field of Industrial Electrical Engineering : Volume 3. - Košice : TU, 2014 S. 280-283. - ISBN 978-80-553-1711-4

Design of System for Analyzing Transmission Environment for Free Space Communication

[1]*Matúš TATARKO (4th year)*
Supervisor: [2]Ľuboš OVSENÍK

[1,2]Dept. of Electronics and Multimedia Communications, FEI TU of Košice, Slovak Republic

[1]matus.tatarko@tuke.sk, [2]lubos.ovsenik@tuke.sk

***Abstract*—This paper deals with two types of free space communication methods and about system which allows to model the characteristics of the transmission channel. In this paper I will focus on Free Space Optics (FSO) communication and on Radio Frequency (RF) transmission. The fog, dust and small particles are critical parameter for FSO link, because they decrease visibility and availability of FSO systems. The dense rain causes attenuation of radio waves. Combination of these two communication methods creates hybrid link which is more reliable than each link separately. Therefore, it is important to know when and how to do the switching process to maintain high speed data transmission process. There are exists several types of switching methods, but I have proposed new switching methods based on fog sensor.**

***Keywords*—availability, communication, fog sensor, free space optics, radio frequency.**

I. Introduction

As was mentioned in abstract, the main losses of FSO are due to fog, dust, smog. Free space optical beams are absorbed and scattered by the air molecules as well as by the solid and liquid particles diffused in atmosphere. Absorption of the signal causes a decrease in signal strength. Scattering does not cause a decrease in signal strength, but it send off the signal in different directions [1]. These phenomena occur in foggy days, when relative humidity is high and visibility is low. Technology of FSO systems has amount of advantages: it provides broadband full duplex communication; it not requires telecommunications license; easy and fast installation; affordably solution of last mile.

Radio frequency communication transmits data via air by electromagnetic waves. Entire frequency spectrum is cut to smaller bands which have different properties. The electromagnetic waves can be propagated directly or by reflection according to wavelength is used. For this purpose we consider direct visibility between antennas because RF link is the backup link for FSO link which needs direct visibility. We use microwave 60 GHz antennas which can provide data transmission rate similar to FSO.

Known switching methods are hard and soft switching [2]. More information about my previous research in this area and about hard and soft switching; configuration commands and devices for switching are describes in our other articles [3,4].

II. Previous Solutions

The most enterprise data applications are satisfied with greater availabilities than 99 %. However, if mission critical data or voice is transmitted via FSO link, the availability requirement increases to 99,999 % [1]. The solution is based on hybrid FSO/RF link. But there is missing a program package which allows simulation of hybrid link. We have program package, called FSO System Simulator, but it does not include RF link [5]. In this program is able to simulate only FSO link under the different weather conditions, change visibility or distance of given link. It is a good choice, but only for simulation of FSO link.

III. Obtaining Information about Weather

As was mentioned above, critical fades are caused by foggy weather. I have proposed a sensor system for monitoring weather in Technical University of Košice campus. Measured data is stored in database which allows easy future evaluation of data. I created application for collecting data for this system in Java. In Fig. 1 you can see window of mentioned application.

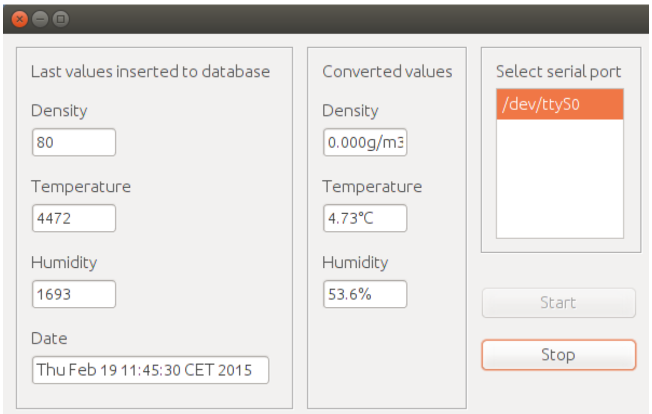


Fig. 1. Application´s window.

Application´s window is divided into three parts (Fig. 1). *Last values inserted to database* obtain raw information from sensor unit and actual date. *Converted values* mean real values about weather with their units. Last part is called *Select serial port*. In this part you can select device´s serial port.
As you can see from Fig. 1 raw value of *Density* is 80 and after conversion to real values it corresponds with 0 g/m^3. It means, that atmosphere was clear, so it was a sunny day

without any fog. Interesting values for us begin with number of *Density* on level 125. This value for continental type of fog means 3,5 km visibility (Fig.2).

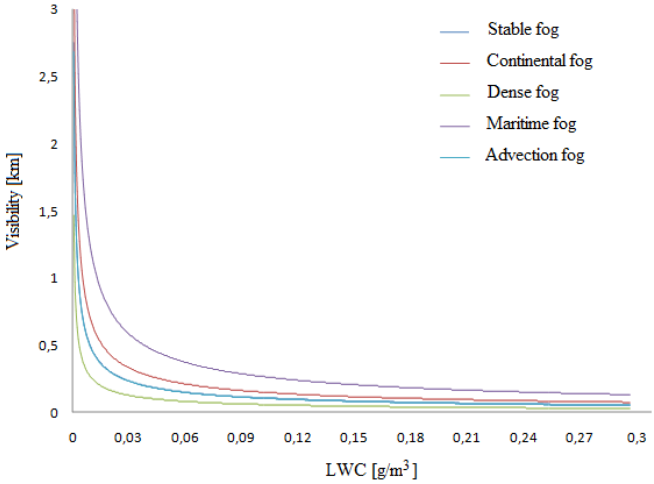


Fig. 2. Dependence of visibility on LWC.

IV. System for Analyzing Transmission Environment

In general, for design and implementation of free space communication in to the environment is necessary to properly set up the parameters of future link. It is important to obtain information about climate in given area, amount of precipitation during the year and character of environment in different seasons. My proposal of system for analyzing transmission environment can be divided to three main parts and it will be programmed in C#. Proposal includes possibility to enter information about RF link too. From obtained information about weather is able to make switching method based on Liquid Water Content (LWC) [6,7].

A. Steady model of Free Space Optics communications

It contains amount of different parameters based on mathematical equations which can be set. All settings we can divide to two areas. First area allows enter device parameters manually or choose them from database. Second area describes channel properties. Due to visibility and weather condition you can simulate different decreases of signal caused by rain or snow. You can also simulate turbulence by the selection of Kim or Kruse model. There is able to choose computing mode via international visibility code [8]. After setting all required properties of the channel, it is necessary to start calculation process. There exist two way, how to show the results. Representation of results is graphically displayed in created window. Detail values of results will be written under the window. They contain information about different types of attenuations: due to atmosphere, turbulences, geometrical, rain or snow and the maximum length of FSO link with given settings under the selected channel properties.

B. Steady model of Free Space Optics/ Radio Frequency communications

This part has the same settings for FSO link, but it allows input device parameter for RF system too. Similar, we can set them manually or choose from database. Setup channel properties are the same for both links. During the calculation process it has been taken in to the account all parameters from the user settings. Representation of result is similar as in *Steady model of Free Space Optics.*

C. Statistical model

This part of program is focus on availability and reliability of free space communication. There is a possibility to create graphs of measured, stored data. Database contains additional functions which help us to filter the stored data. Every day 8640 values of each parameter are stored. If we want to show availability of given FSO or RF systems for a long periods only for visual check, we do not need to count every single measurement. Program allows set the threshold for maximum density of fog. After that it is able to calculate availability of FSO or RF link [%]. User can also see graphs of fog, temperature and relative humidity. Data which exceeds the threshold is counted [sec]. Each pattern is one second long.

V. Conclusion

My next research will be focus on my dissertation work and to make the final version of System for Analyzing Transmission Environment for Free Space Communication. My dissertation work is divided to three main parts. Two parts I have done, but third part including program package is in the final process. I need to complete the graphic interface and make experiment by this program. I have amount of weather measurement, so I need to evaluate them. After that I want to write a manuscript which will be the end of my Phd. study.

Acknowledgment

This work was supported by project no. „006TUKE-4/2014 The Use of TUKE PON Experimental Model in Teaching". This work is also the result of the contract no. "APVV-0025-12 - Mitigation of Stochastic Effects in High-Bitrate All Optical Networks" supported by the Slovak Research and Development Agency.

References

[1] I.I.Kim, E. Korevaar, "Availability of Free Space Optics (FSO) and Hybrid FSO/RF Systems," Proc. of SPIE, vol. 4530, pp. 84-95, 2001
[2] Z. Wenzhe, S. Hranilovic, C. Shi, "Soft-Switching Hybrid FSO/RF Links using Short-length Raptor codes: Design and Implementation," IEEE journal on selected areas in communications, vol. 27, no.9, December 2009
[3] M. Tatarko, Ľ. Ovseník, J. Turán, "Prepínanie hybridnej FSORF linky pomocou prepínača," Electrical Engineering and Informatics 4: proceedings of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice, Košice, 2013, pp. 177-181. ISBN 978-80-553-1440-2
[4] M. Tatarko, Ľ. Ovseník, J. Turán, "Prepínanie hybridnej FSORF linky pomocou smerovača," Electrical Engineering and Informatics 4 : proceedings of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice, Košice, 2013, pp. 182-186. ISBN 978-80-553-1440-2
[5] M. Tatarko, Ľ. Ovseník, J. Turán, "Software package for analyze FSO link," Infocommunication Journal, No. 1, 2013.
[6] M. S. Khan, S. S. Muhammad, M. S. Awan, etc., "Validating Relationship between Aerosol´s Liquid Water Content and Optical Attenuation for Terrestrial FSO Links," 11th International Conference on Telecommunications, Gratz, Austria, June, 2011.
[7] A. Prokes, V. Skorpil, "Estimation of Free Space Optics Systems Availability Based on Meteorological Visibility," IEEE Latin American Conference, Medelin, 2009.
[8] I. I. Kim et. al., "Wireless optical transmission of fast Ethernet, FDDI, ATM, and ESCON protocol data using the TerraLink laser communication system," Optical Engineering, Vol. 37 No. 12, December 1998.

Determining the position of traffic signs in automatic inventory system based on optical correlator

[1]*František RAKOCI (4th year)*
Supervisor: [2]Ľuboš OVSENÍK

[1,2]Dept. of Electronics and Multimedia Communications, FEI TU of Košice, Slovak Republic

[1]ferorak@gmail.com, [2]lubos.ovsenik@tuke.sk

***Abstract*—This article describes the proposal of traffic sign position determining algorithm, for purpose of automatic inventory system, based on optical correlator. Whole inventory process should be processed using the inventory system installed in a vehicle. It should work without any need of human intervention. Because positions of car and recorded traffic sign are different, it is needed to determine the traffic sign position by calculation based on available input parameters. The main input parameter is the exact car GPS position and distance of recorded traffic sign from this vehicle. An additional parameter is a vehicle orientation and input image areas representing the recorded traffic sign.**

***Keywords*—Automatic inventory, GPS position, optical correlator, traffic sign**

I. Introduction

The main purpose of our research is to propose and realize the automatic traffic sign inventory system with use of optical correlator. Automatic inventory makes an evidence, controlling and maintenance activities significantly easier. In addition to data collection, the inventory system can signalize some traffic sign defects, or create a statistic reports for maintenance purposes. In present we have prepared a structure of system software core. Our goal is to implement each of the individual parts described in the block scheme (Fig. 1). The actual important part of inventory system functionality we are implementing is to determine the position of recorded traffic sign using GPS coordinates. Inventory system hardware is usually installed in the vehicle dedicated for this purpose. The position of vehicle and recorded traffic sign is different during the whole inventory process. It is needed to determine the exact traffic sign position using the specific correction, based on distance between the vehicle and traffic sign and on their relative geometric orientation.

II. System Block Scheme

The Fig. 1. is representing the blok scheme of whole inventory system. It consists from software core and I/O peripherals. Software core communicates with the optical correlator, stereoscopic camera system and GPS receiver. Information from peripherals is processed and stored in the inventory database.

A. Software core

Each function block of software core controls the functionality of whole system. The segmentation block obtains pictures from stereoscopic camera system and creates binary images for optical correlator processing. Optical correlator compares these images with binary samples from reference database. Its output image is processed in Match evaluation block and the result is stored in the inventory database. The traffic sign position is calculated in position processing block. The most important input parameters for this block are GPS coordinates and orientation of vehicle. The additional information for position calculation are the segmented binary images from stereoscopic camera system. It is possible to get the distance between the vehicle and the recorded traffic sign based on horizontal shift of image between these images.

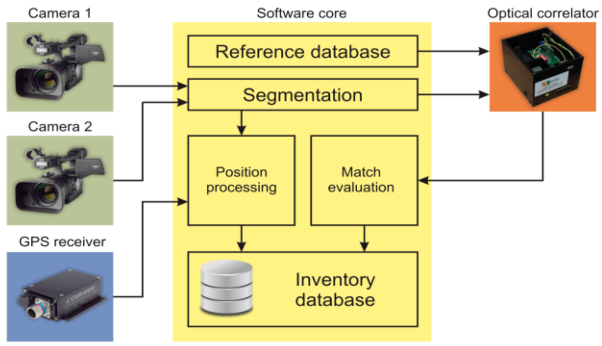


Fig. 1. Block scheme of automatic traffic signs inventory system [2]

B. Optical correlator

Optical correlator is the main computing part of whole inventory system. It provides the optical correlation using a difraction lenses. Because the optical correlator works in very high speed, it is possible to compare binary image samples without significant load of electronic computing system.

C. Stereoscopic camera system

Stereoscopic camera system is used to record images and for measurement of recorded traffic sign distance. Calculation of this distance is based on relative horizontal shift of binary images from both cameras of stereoscopic system.

D. GPS receiver

GPS receiver is used for obtaining the actual GPS coordinates of vehicle. This information is the main input parameter for calculation the traffic sign position. The second important parameter obtained from GPS receiver is the horizontal orientation of vehicle. It is possible to get the direction of motion vector by comparing GPS coordinates of vehicle in specific time intervals.

III. Basic input parameters

A. GPS position

GPS position of inventory vehicle is the key parameter which the additional calculation of exact traffic sign position is based on. The inventory system obtains the GPS coordinates in NMEA sentences form. It is needed to export the necessary information from NMEA structure. [3]

B. Horizontal orientation

The recorded traffic sign has always non-null distance from vehicle, so its GPS coordinates are different from GPS coordinates of the inventory vehicle. For correct calculation of the exact traffic sign position it is important to know the correct relative geometric configuration of vehicle and traffic sign. It is possible to determine the proportion for correction of individual coordinates based on this geometric configuration. The vehicle orientation is calculated as a direction of vehicle motion vector. If the GPS coordinates are obtained in at sufficiently short intervals, the motion vector direction can be calculated using simple trigonometry formula. If the first vehicle coordinates are lat_{v1}, lon_{v1}, and second coordinates are lat_{v2}, lon_{v2}, the angle for correction of individual coordinates, based on vehicle direction, should be calculated as

$$\varphi_v = \tan^{-1}\left(\frac{lat_2 - lat_1}{lon_2 - lon_1}\right) \qquad (1)$$

IV. Additional parameters

A. Traffic sign distance

Distance between the traffic sign and vehicle is determined using the output of stereoscopic camera system. According to relative shift of the corresponding binary images of certain traffic sign between the both pictures, it is possible to determine the distance of this traffic sign from the vehicle. The Fig. 2. describes the basic principle of stereoscopic measurement. Distance can be calculated according to this figure and formula (2).

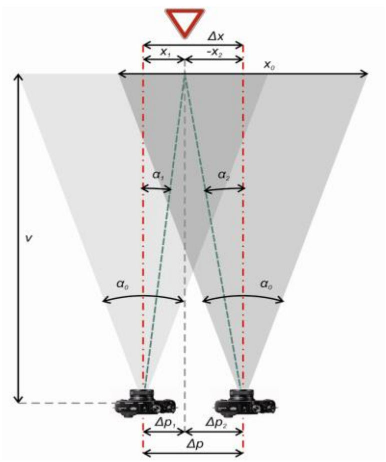


Fig. 2. Principal of stereoscopic measurement [2]

$$v = \frac{\Delta x}{2 \cdot \tan\left(\frac{\alpha_0}{2}\right)(x_2 - x_1)} \qquad (2)$$

B. Position on input scene

During the inventory process, the road signs are recorded using the camera system with specific viewing angle. It is needed to know the right place in the image scene where the recorded traffic sign is placed. As it is described in Fig. 2., we can calculate the corresponding horizontal angle where the recorded traffic sign is placed. Then we can simply add this angle to the vehicle angle for exact GPS coordinates correction (3). The center of images is zero point, so traffic signs placed on the left side have negative values of angle. [1]

$$\varphi = \varphi_v + \varphi_i \qquad (3)$$

V. The final position determining algorithm

- Regular determining vehicle position using GPS receiver
- Calculation of the vehicle direction
- Computing distances between vehicle and traffic sign
- Direction offset, based on input scene sign position
- Calculation of the exact traffic sign coordinates [lat, lon], based on (4) [4]

$$\begin{gathered} dist = 1{,}1515 \cdot 60 \cdot 1{,}609344\ km/° \\ lat = lat_v + \frac{v \cdot \sin(\varphi)}{dist} \\ lon = lon_v + \frac{v \cdot \cos(\varphi)}{dist \cdot \cos(lat)} \end{gathered} \qquad (4)$$

VI. Conclusion

The traffic sign position determining is the only one of many parts of automatic inventory system. It contributes to provide a complex inventory process without need of any human intersection. Compared with another solutions, this approach of position determining allows us to identify the exact position of traffic sign independently of inventory car position. It is needed to solve more accurate segmentation, based on color and corner filters. In our research we are also trying to implement the evaluation of different types of traffic sign damages, based on geometric proportion, or comparing of inner traffic sign textures. [3]

Acknowledgment

This work was supported by Cultural and Educational Grant Agency (KEGA) of the Ministry of Education, Science, Research and Sport of the Slovak Republic under the project no. „059TUKE-4/2015 - Implementation of the FSO Long-Term Experimental Measurement Results into the Educational Process in Optoelectronic Sensor and Communication Systems Field“, and project no. „006TUKE-4/2014 - The Use of TUKE PON Experimental Model in Teaching”. This work is also the result of the contract no. “APVV-0025-12 - Mitigation of Stochastic Effects in High-Bitrate All Optical Networks” supported by the Slovak Research and Development Agency.

References

[1] J. Mrovlje, D. Vrančić, *Distance measuring based on stereoscopic pictures*, 9th International PhD Workshop on Systems and Control: Young Generation Viewpoint, 1.-3. Október 2008 , Izola, Slovenia.

[2] F. Rakoci, *Stereoskopické meranie vzdialenosti v automatickom inventarizačnom systéme dopravných značiek,* Electrical Engineering and Informatics V, Proceeding of the Faculty of Elektrical Engineering and Informatics of the Technical University of Košice, Slovakia, 2014, ISBN 978-80-553-1704-5

[3] J. Ružbarský, Ľ. Ovseník, J. Turán, *Traffic Signs Inventory System*, Carpathian Journal of Electronic and Computer Engineering 6/1, 2013, pages 44-49, ISSN 1844 – 9689

[4] Ed Williams, *Aviation Formulary V1.46.*, published on: http://williams.best.vwh.net/avform.htm.

Effect of series FACTS devices on distance protection

[1]*Miroslav KMEC (3rd year)*
Supervisor: [2]Ľubomír BEŇA

[1, 2] Dept. of Electric Power Engineering, FEI TU of Kosice, Slovak Republic

[1]miroslav.kmec@tuke.sk, [2]lubomir.bena@tuke.sk

Abstract— **The increase in the line complexity, configuration and the power transmitted has introduced a number of problems in relation to distance protection. Distance relays are widely used as the main protection for transmission lines around the world. They operate based on impedance measurement at the relaying point, which is affected by several factors including the fault location, pre-fault condition (loading condition), short circuit levels, arc fault resistance and presence of FACTS devices on a transmission line. In this paper, discussion is made on the basic needs of protective systems, the various schemes of protection and the importance of distance protection in power system.**

***Keywords* — FACTS devices, distance relays, faults, transmission lines.**

I. Introduction

In recent years the application of FACTS devices in power systems has increased to enhance the operational capacity of transmission lines. This has been achieved by pushing transmission lines to operate closer to their thermal and stability limits [1].

Protection is one of the most important areas of the power system operation, where the impact of FACTS devices is considerable and therefore it is essential to be studied. In this context, distance relay is very much prone to mal-operation as its impedance measurement is largely affected by the presence of FACTS devices and their operational strategies. Effects of these devices compounded with the impacts of previously mentioned factors could severely affect the apparent impedance measured at the relaying point and cause mal-operation of distance relays, in the form of over-reaching or under-reaching the fault point. Therefore, in the presence of FACTS devices, the conventional relay characteristics, such as Mho and quadrilateral, may not be suitable under some fault conditions. Impacts of FACTS devices on the measured apparent impedance at the relaying point are more complicated than the previously mentioned factors, since apart from their parameters, their installation points on transmission lines can also affect the measured apparent impedance [1]-[3].

II. General overview of issues

One of the most important FACTS devices which is connected to system in series is Thyristor Controlled Series Capacitor (TCSC) that is used to increase the overall capability of power transmission in a power long transmission line. Disadvantage of it is that although its usage will improve the power transmission capability and Stability on system, it may result i non-coordinated trip operation of circuit-breakers connected to main power system [4].

TCSC mounted on Fig. 1. (Left) is a type of series FACTS compensators. It consists of a capacitance (C) connected in parallel with an inductance (L) controlled by a valve mounted in anti-parallel thyristors conventional (T1 and T2) and controlled by an extinction angle (α). Using TCSC is possible to increase the transfer capability of existing power transmission systems together with the use of other benefits [5].

If we want solve only steady state (published in Elektroenergetika journal, 2015) we can simplify our model of TCSC, considering that ω is a constant representing the parallel connection unchanging capacitance XC and variable inductance XTCR (α), or susceptance BC and BTCR (α). After further simplification, we could replace the entire controller only element with XTCSC (α) respectively BTCSC (α), Fig. 1 (Right).

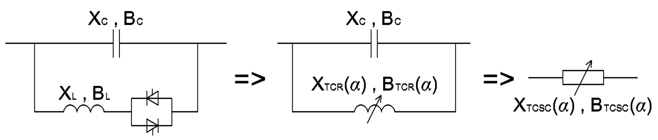


Fig. 1. Structures of simplified TCSC model

If we want to examine what will happen in power system when short-circuit or other dynamic state will occur we need to consider management of extinction angle (α) using controller, presented at Fig.2.

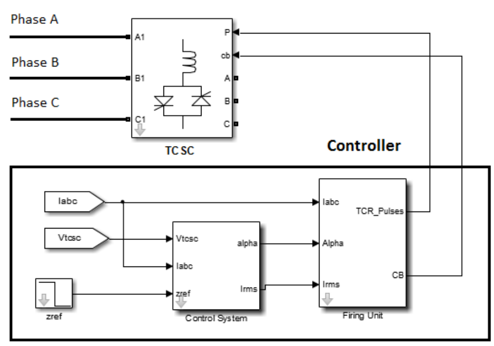


Fig. 2. Structures of dynamic TCSC model

It is well documented in the literature that the introduction of FACTS devices into power systems has a great influence on their dynamics. As power system dynamics changes, many sub-systems are affected, including the protective systems. Therefore, it is essential to study effects of FACTS devices on the protective systems, especially distance protection [6].

III. Current Research Activities

At the beginning of my research I have dealt with theoretical study of operations and settings of digital distance protection (Distance Relay - DR). This study consisted from understanding of several different parts but mainly from operation and relation the internal logic of these devices.

Subsequently I continued my research by creating a functional model of DR in program Matlab / Simulink. In this program was designed the whole internal logic of protection, where individual parts operate as one unit. This part was published in Elektroenergetika journal (Vol.7, No1 (2014)). There is always space for next upgrades and further ideas to improve the model, which I gradually add to the model.

Next steps in extending the DR model lead to wider implementation of fault loop impedance calculation for different types, configurations and operation modes of the network which was published in a conference (Current Problems of Maintenance of Electrical Equipment and Management, 2015) and Acta Electrotechnica and Informatica journal, 2015.

In the past MHO tripping characteristics were most frequently used but these characteristic did not provide sufficient comfort settings.

In my research I worked with modern quadrilateral characteristic (composed of lines), which have variability settings, thereby providing opportunity for future adaptation and changes, which can occur by the addition of new device (in my case it will be FACTS devices) to the existing power system.

After completing basic model DR I came to study the operation of series TCSC device described in the sections above.

I studied all the effects and possible variants simulation of this device in professional publications and books. Now, I'm trying to simulate this device and implement to the network (later more complex network) together with DR.

IV. Direction of future research activities

My next steps would lead to the creation common model of DR. This means that I want to create a block "in Matlab called subsystem", where the user only connects necessary measured values for calculate and sets the basic parameters of the network in user interface. This unit could be set according to the configuration of the power system in which will be implemented.

At this time the block is fully functional and includes all the necessary parts for its unmanned operation. Given that I'm not always satisfied with this block, therefore it is continuously improved and its final version will be ready after the completion of the all joint tests DR with TCSC device.

However, the main objective is the implementation of designed TCSC model to the network and study changes in the measured impedance parameters with subsequent evaluation.

Nowadays, I examine closer dynamic TCSC model implemented in Simulink program and test its behavior in various situations which may arise in power system. Only then can I investigate the impact of TCSC on the DR parameters.

According to the publication, I know that TCSC device have a significant impact on DR and therefore should my final steps have led to description of the impact of TCSC device on quadrilateral characteristic. This description would allow quicker and easier setting of distance relay parameters, if TCSC device will be fitted for the first time to the existing network.

V. Conclusion

In the presence TCSC device in power transmission systems, the conventional distance characteristic are greatly subjected to mal-operation in the form of over-reaching or under-reaching the fault point therefore these characteristic might not be utilized satisfactorily.

Since the distance relay are the most common used protection HV and EHV networks, therefore they should be given to more attention.

References

[1] S. Jamali and H. Shateri "E*ffect of TCSC on distance relay tripping characteristic*" Iran University of Science and Technology, Universities Power Engineering Conference, 2004. UPEC 2004. 39th International, Volume 2, Publication Year: 2004, Page(s): 704 - 708 vol., *IEEE conference publication*.

[2] M. Khederzadeh "*The Impact of FACTS Devices on Digital Multifimctional Protective*", Transmission and Distribution Conference and Exhibition 2002: Asia Pacific. IEEE/PES, Volume: 3 Digital Object Identifier: 10.1109/TDC.2002.1177775 Publication Year: 2002, Page(s): 2043 - 2048 vol.3 Cited by: Papers (30), *IEEE conference publication*.

[3] A. Apostolov "*Modeling of Multifunctional Distance Protection IEDs*" Developments in Power System protection (DPSP 2010), Managing the Change, 10th IET International Conference on Digital Object Identifier: 10.1049/cp.2010.0354 Publication Year: 2010 , Page(s): 1 – 5, *IEEE conference publication*.

[4] S. Jamali and H.Shateri "*Effect Of Tcsc On Distance Relay Tripping Characteristic*", Universities Power Engineering Conference, 2004. UPEC 2004. 39th International (Volume:2), 2004

[5] M. Zellagui and A. Chaghi "*Comparing Effects of TCR and TSC on MHO Distance Protection Setting in 400 kV Algerian Transmission Line*" Leonardo Electronic Journal of Practices and Technologies, ISSN 1583-1078

[6] M. Zellagui and A. Chaghi "*Impact of TCSC on Measured Impedance by MHO Distance Relay on 400 kV Algerian Transmission Line in Presence of Phase to Earth Fault*" J. Electrical Systems 8-3 (2012): 273-291.

Effects of non-ionizing electromagnetic fields exposing biological systems

[1]Michal KOSTEREC (1st year)
Supervisor: [2]Juraj KURIMSKÝ

[1,2]Dept. of Electric Power Engineering, FEI TU of Košice, Slovak Republic

[1]michal.kosterec@tuke.sk, [2]juraj.kurimsky@tuke.sk

***Abstract*—With rapid development of wireless transfer of information the exposition of our bodies to radiofrequency electromagnetic radiation has increased. Every day the radiofrequency electromagnetic field of mobile phones, Wi-Fi internet, TV broadcast and even more effect on us while we cannot see it or even feel it. First part of this paper is focused on literature review of electromagnetic field properties and effects of non-ionizing radiation on biological systems. In the end it evaluates future work and research of effects of non-ionizing radiation on biological systems.**

***Keywords*—electromagnetic field, biological effect of electromagnetic filed, specific absorption rate, temperature.**

I. INTRODUCTION

Whether this electromagnetic exposure is or is not dangerous for our health scientist are trying to find answer for. Electromagnetic pollution is responsible for interference between electrical devices, but it also affects the human body. The organism reacts to the electrical, magnetic and electromagnetic fields around it. The biological effects depend on the type, frequency and strength of these fields. The boom in the use of the cellular phone has led to a very evident increase of electromagnetic pollution [1].

The number of mobile phone users has increased exponentially recently and it has become an important device in human daily life. The world's largest individual mobile operator by subscribers is China Mobile with over 500 million mobile phone subscribers. Over 50 mobile operators have over 10 million subscribers each, and over 150 mobile operators had at least one million subscribers by the end of 2009. In 2014, there were more than seven billion mobile phone subscribers worldwide, a number that is expected to keep growing [2]. Figure 1 shows the increase of mobile users until 2014.

Radio frequency used in mobile communication has the ability to penetrate through the semi-solid substances like meat and living tissue [3]. Researches conducted show that the electromagnetic wave produced by the mobile phone might cause adverse effect to human especially at place near the ear region [4].

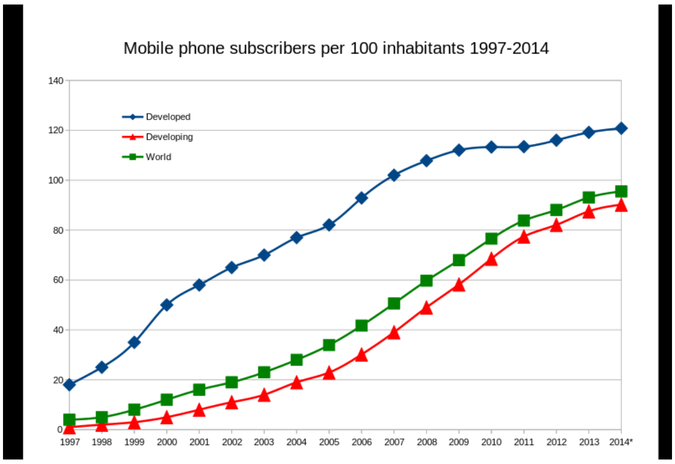


Fig. 1 Mobile phones in use per 100 inhabitants 1997-2014 [2]

II. LITERATURE REVIEW

A. *Electromagnetic field properties*

Constant fields - the direction does not change. Examples include the Earth's natural electric and magnetic fields, the magnetic fields used to run subways and streetcars, and the fields used in magnetic resonance imaging (MRI).

Alternating fields - the direction does change. The number of periods per second is equal to the frequency.

Low-frequency fields (LF): Alternating fields up to 30 kHz. This fields typically Occur in power systems, electric railways and industrial processes, droughts and melting, smelting and welding.

High-frequency, radio-frequency (HF, RF) fields: Alternating fields from 30 kHz to 300 GHz. These fields are found in mobile phones, radio and TV broadcasting, satellite communication, radar systems, etc.

Electromagnetic fields propagate as waves and travel at the speed of light *c*. The wavelength is proportional to the frequency. If the distance to the field source is less than one wavelength, then we are usually in the near field (almost always the case in the low frequency range up to 30 kHz). If the distance is more than one wavelength, then far field conditions usually hold [1].

This distinction between near and far fields is important when it comes to measurements. In the near field, the ratio of electric field strength *E* (V/m) and magnetic field strength *H* (A/m) is not constant, so we have to measure each separately. In the far field, it is enough to just measure one field quantity since we can compute the other [1] [5]. Their ratio being a known constant:

$$Z_\omega = \frac{E}{H} = 120\pi = 377\Omega. \quad (1)$$

For point sources, the far field is at the distances greater than $r > \lambda / 2\pi$, where r is the distance from the source [5].

In the near field (Figure 2), the wave impedance may become any value ranging between:

$$120\pi\left(\frac{2\pi r}{\lambda}\right) < Z_\omega < 120\pi\left(\frac{2\pi\lambda}{r}\right)\Omega. \quad (2)$$

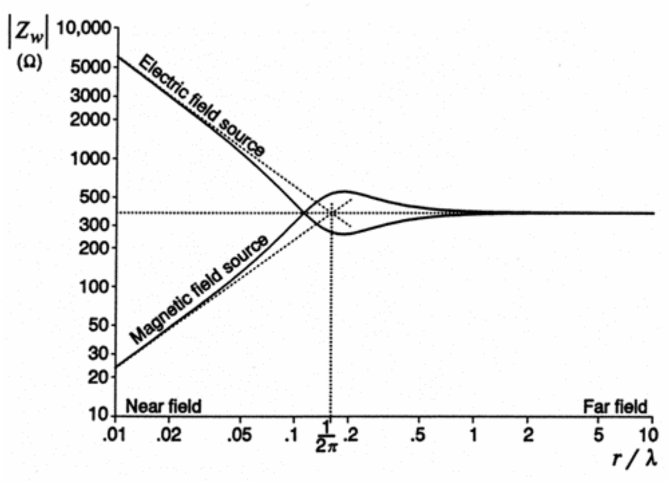


Fig. 2 Wave impedance depended on distance [5]

For antennas with dimensions D, the following requirements must be met for far-field [6]:

$$r > \frac{2D^2}{\lambda};\ r > 5D;\ r > 1{,}6\lambda. \quad (3)$$

B. Global System for Mobile communication

Global System for Mobile communication (GSM) is a digital wireless communication protocol for mobile phones and was developed in the early 1980's. GSM is a cellular network used by mobile phone where mobile phone connect to the cellular network by searching of cells in the immediate surrounding area [7]. GSM technology is invented to eliminate certain problems with the predecessor cellular networks. The existing problems of cellular network are that analog networks could not handle the growing capacity of cellular networks and the existing digital networks were not compatible with each other [4]. GSM technologies had made the data communication easier to build into the system. It is a low-cost standard supported voice calls and short message service (SMS). GSM also provided useful features like security, authentication and the Invention of the SIM card. GSM networks operate in four different frequency ranges which are 900 MHz band, 1800 MHz band, 850 MHz band and 1900 MHz band. GSM 900 and GSM 1800 standard is the most commonly used standard. 850 MHz band and 1900 MHz band is introduced because the 900 MHz and 1800 MHz frequency bands were already allocated [8]. GSM 1800 standard provide more bandwidth and less power requirements than GSM 900 MHz The transmission power in the mobile phone is limited to a maximum of 2 watts in GSM 850 and GSM 900 while maximum power of 1 watt in GSM 1800 and GSM 1900 networks [4] [7].

C. Biological effects of electromagnetic radiation

The measurable effect of the RF electromagnetic radiation is the thermal one [9]. The degree of absorption of electromagnetic waves is a function of the frequency and intensity of the field and the type of tissue. The depth of penetration deceases at higher frequencies. Fields above 10 GHz are absorbed at the skin surface. Only a small portion of the energy penetrates into the underlying tissue. Very high field strength is needed to produce cataracts or skin burns. The organs with the least blood flow are most endangered, e.g. the eyes. In contrast, the heart and brain are better at handling heat due to their better blood flow. Besides obvious damage such as burns, long-term effects are also under discussion, e.g. increased risk of cancer, influences on hormone balance, cell growth and the immune system [10].

RF radiation can affect human body because human body is made up of 65-70% water, electrolytes and ions. The water molecule is a polar molecule which has positive and negative charges separated by a dipole length. Since human body contains water, electrolytes and ions, human body has its own weak electromagnetic field and each of the cells has its own electromagnetic field too [11]. Therefore, weak electromagnetic fields such as RF radiation emitted from mobile phone will interact with human body and affect the human body's own weak electromagnetic fields. As the human body's own weak electromagnetic fields were interfered, the body's natural healing processes also interfered [4] [11].

Human body's own electromagnetic frequencies are essential to human's health, repair cells, reproduction and replication of Deoxyribonucleic Acid (DNA). It will be altered by outside sources of energy such as electromagnetic and electrical radiation then cause biological effect to human. Electromagnetic and electrical radiation can cause the displacement of electrolytes and ions within the body which interfere the body's neurological system and maintaining homeostasis system. This will weaken the defense mechanism of the body as the body expends energy to redress this imbalance [10].

RF radiation emitted from mobile phone has effects on human health which can be categorized as thermal, non-thermal, genotoxical, increased probability of getting brain tumor and non-specific symptoms [12]. Non-specific complaints had been made by mobile phone users such as symptoms like headaches, earaches, blurring of vision, short term memory loss, numbing, itchy, burning sensations, bad sleep, electromagnetic hypersensitivity exhaustion and anxiety when using mobile phone. Researchers had found that symptoms such as headache, fatigue, and difficulty in concentration were more common in people with higher exposures to RF radiation [13].

Experiments performed at Radiation and Nuclear Safety Authority (STUK) also have produced indicators that mobile radiation could cause temporary changes in the functions of cells. These functions include the functions of genes, activation of proteins, and the internal chemical communication within cells. Some studies also show that mobile phone signals affect sleep patterns and possibly delay sleep onset cause of RF radiation exposure. In another clinical study carried out by Sweden's Karolinska Institute and Wayne State University in the US (United States), the authors suggested adverse effects on sleep quality within certain sleep stages is caused by radiation. Electromagnetic hypersensitivity happens in some users of mobile phone who reported feel several unspecific symptoms such as burning and itchy in the skin of the head, exhaustion, sleep disturbances, dizziness, loss of mental attention, reaction times and short term memory loss, headaches, depression, heart palpitations, and disturbances of the digestive system [14].

D. Specific absorption rate

Different concepts and quantities have been proposed and used to characterize propagation and absorption of electromagnetic energy in biological systems. One of the identities is *SAR*. *SAR* is the time rate of energy absorption per gram of tissue from non-ionizing electromagnetic radiation. The term SAR distribution is used to indicate the pattern of *SAR* inside the body [14].In the Figures 1 and 2 there are shown *SAR* distribution towards human head when using mobile phone.

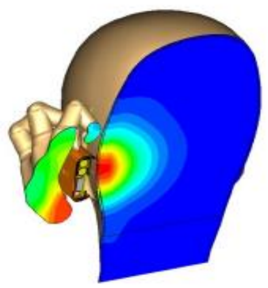


Fig. 3 SAR distributions towards human head when using mobile phone [14]

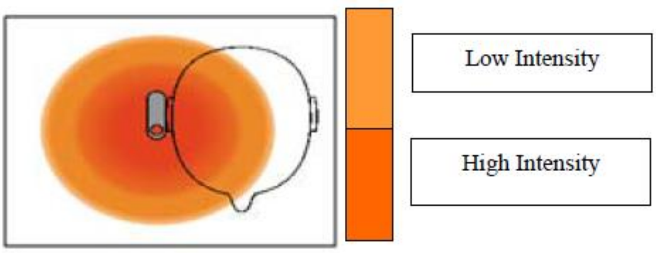


Fig. 4 High intensity radiation of radio frequency toward the skin near the ear skull region [15]

From [8] it is shown that the computed total power absorbed by head tissues is P_{abs} = 0.795W with only mean *SAR* values over tissue balls are used in mobile phone norms. *SAR* values are dependent on distance of the body and mobile phones. As the distance of the body and mobile phones is closed, the *SAR* values will be higher and vice versa. *SAR* value is normally specified at the maximum transmission power. Transmission power will be higher when the mobile phone is used at the area with very low field strength of received signals [16] [17].

Using mobile phones with low *SAR* values and making a call at high reception field strength which permitting low transmission power can reduce the exposure to the high intensity radiation [14]. *SAR* is time rate of energy absorbed in an incremental mass, divided by that mass. Average *SAR* in a body is the time rate of the total energy absorbed divided by the total mass of the body. The units are watts per kilogram (W/kg). *SAR* can be calculated by [18]:

$$SAR = \sigma \frac{|E|^2}{\rho} = c \frac{\Delta T}{\Delta t} \ (W/kg) \qquad (4)$$

where, σ is the conductivity of tissue simulant (s/cm), E is the electric field strength (V^2/cm^2), ρ is the density of tissue simulant (g/cm^3), c is the specific heat capacity of the tissue simulant (J/g/°C) and T is the change in temperature when exposed for time change of Δt.

TABLE I
ELECTRIC PARAMETERS FOR VARIOUS TISSUES [19]

Tissue	ρ (g/cm^3)	Frequency			
		900 MHz		1800 MHz	
		ε_r	σ (S/m)	ε_r	σ (S/m)
Bone	1850	8	0.11	8	0.16
Skin/fat	1100	34.50	0.6	32	0.52
Muscle	1040	58.50	1.21	55	1.7
Brain	1030	55	1.23	53	1.7
Lens	1050	44.50	0.8	41	1.29
Cornea	1040	52	1.85	50	2.32

III. PRESENT AND FUTURE RESEARCH

Our preliminary works are conducted to previous research efforts of the team of Professor Karol Marton [20]. We applied non-ionizing electromagnetic field to expose several biological tissues for different time periods. We have chosen the frequency and power similar as it is with telecommunication devices. The very first experiences revealed that research methodology, laboratory procedures, analytical methods and sample logistics are suitable for designing of future research [21].

For the measurement of non-ionizing radiation in various locations of the Slovak Republic, subsequent analysis and consultation will conclude an agreement with an expert on electromagnetic fields in practice. In the future it is planned to find researchers in biology with required equipment and laboratory for providing biological tissue and for analysis of these tissues before and after electromagnetic exposure. Researching the impact of non-ionizing radiation will be assembled in the workplace with EMC chamber where biological tissue will be exposed to the non-ionizing radiations, which have been previously measured in various locations. For a better idea of penetration, distribution and impact of non-ionizing radiation on biological tissue will be creating a mathematical model and simulation in ANSYS software.

IV. CONCLUSION

According to statistics the trend in various types of wireless communication is rising. It rises localized electromagnetic fields pollution in populated areas. On one side this calls for better bands management from technical, and on the other for deeper studies of possible negative effects from ecological and health risk standpoints.

REFERENCES

[1] M. Salovarda, K. Malaric, “*Measurements of Electromagnetic Smog*,” in Electrotechnical Conference, 2006. MELECON 2006. IEEE Mediterranean, 2006, pp. 470–473.

[2] Wikipedia, Mobile Phone [Online]. Avaible at: <http://en.wikipedia.org/wiki/Mobile_phone> Accesed on 9. March 9, 20152015

[3] M. A. Bhat, Dr. V. Kumar, “*Calculation of SAR and Measurement of Temperature Change of Human Head Due To The Mobile Phone Waves At Frequencies 900 MHz and 1800 MHz*,” Advances in Physics Theories and Applicat, Vol.16, 2013

[4] D. Mat, F. K. W. Tat, K. Kipli, A. Joseph, K. Lias, A. S. W. Marzuki, "*Visualization and Analytical Measurement of Electromagnetic Radiation from Handheld Mobile Phones*," in 2010 Second International Conference on Computer Engineering and Applications (ICCEA), 2010, vol. 2, pp. 246–250.

[5] I. Strauss: "*Near and Far Field – From Statics to Radiation*," Conformity, Feb. 2001.

[6] R. Bansal: "*The far field: how far is far enough?*", Applied Microwave and Wireless, Nov. 1999, pp. 58- 60.

[7] Wikipedia. GSM. [Online]
Available at:
<http://en.wikipedia.org/wiki/GSM>
Access on March 9, 2015.

[8] G. Scarella, O, Clatz, S. Lanteri, G. Beaume, S. Oudot, JP. Pons, S. Piperno, P. Joly, J. Wiart. "*Realistic numerical modeling of human head tissue exposure to electromagnetic waves from cellular phones*." 2006.

[9] L. Ungureanu, M. Baritz, L. Cristea, D. Cazangiu, "*The heating level analysis of facial tissue during the exposure of electromagnetic radiations provided from a mobile phone*," in E-Health and Bioengineering Conference (EHB), 2013, 2013, pp. 1–4.

[10] M. Salovarda, K. Malaric, "*Measurements of Electromagnetic Smog*," Electrotechnical Conference, 2006. MELECON 2006. IEEE Mediterranean, 2006, pp. 470–473.

[11] J.D. Kraus. "*Electromagnetic*," International Edition. 4th ed. New York: McGraw-Hill, Inc., 1992.

[12] D. Coggon. Health Risk from Mobile Phone Base Stations. [Online]
Available at:
<http://oem.bmj.com/cgi/content/full/63/5/298>
Access on March 9, 2015.

[13] Wikipedia. Mobile Phone Radiation and Health. [Online]
Available at:
<http://en.wikipedia.org/wiki/Mobile_phone_radiation_and_health>
Access on March 9, 2015.

[14] D. Mat, F. Kho, A. Joseph, K. Kipli, S. Sahrani, K. Lias, A. S. W. Marzuki, "*The effect of headset and earphone on reducing electromagnetic radiation from mobile phone toward human head*," in 2010 8th Asia-Pacific Symposium on Information and Telecommunication Technologies (APSITT), 2010, pp. 1–6.

[15] G.A. Bhalerao, Prof. R.G. Zope, G.B. Khaire, B.P. Tambe, "*Human Brain with Emr of Cell Phone*," International Journal of Modern Engineering Research, Vol.3, Issue.1, Jan-Feb. 2013 pp-307-310

[16] C. Karger, "*Infrared Thermal Imaging to Measure Local Temperature Rises Caused by Handheld Mobile Phones*," IEEE Transactions on Instrumentation and Measurement, 54 (4); 1513-1519, 2005th

[17] H. Arai, "*Measurement of mobile antenna system*," Artech House Inc. 2,001th

[18] Specific Absorption Rate (SAR) requirements and regulations different regions, the 3rd Generation Partnership Project, 1999

[19] F. Akleman, L. Sevgi. "*FDTD Analysis of human headmobile phone interaction in terms of specific absorption rate calculations and antenna design*," IEEE-APS Conference, Antennas and propagation for wireless communication, vol. 1. pp. 85-88. 1998.

[20] K. Marton, R. Cimbala, I. Kolcunová, "*Biologické systémy a elektromagnetická kompatibilita*", The 6th International Scientific Symposium ELEKTROENERGETIKA 2011, Technical University of Košice, 2011, pp. 146-148.

[21] M. Foľta, "*RF elektromagnetický expozičný system*," Bakalárska práca. Košice: Technická univerzita v Košiciach, Fakulta elektrotechniky a informatiky, 2014

Finite Control Set Model Predictive Control Applied on Permanent Magnet DC Motor

[1] *Viktor Šlapák (3rd year)*
Supervisor: [2] František Ďurovský

[1,2] Dept. of Electrical Engineering and Mechatronics, FEI TU of Košice, Slovak Republic

[1] viktor.slapak@tuke.sk, [2] frantisek.durovsky@tuke.sk

Abstract—**This paper is dedicated to one of predictive control methods – the finite control set model predictive control with application to a permanent magnet DC motor (PMDC). Proposed method uses prediction of future states of PMDC for all possible switching states of power converter as a voltage source. Both speed and current of PMDC motor are controlled by a single controller. Simulation results are included.**

Keywords—**model predictive control, permanent magnet DC motor, cost function**

I. Introduction

Generally, predictive control includes various control methods, but their common characteristic is that they use system model to predict future behavior of the controlled system.

The advantages of the predictive control are:

- easy to understand
- easy to implement
- applicable to various systems
- more variables can be controlled simultaneously by single controller

On the other hand, predictive control is sensitive on the accuracy of used system model and often needs high computation time, thus it was not applied on fast systems, such as power converters and electrical drives in the past. However, with development of fast microprocessors, computation power is fast enough to overcome this disadvantage. A survey of predictive control applied on power electronics and electrical drives can be found in [1].

Finite control set model predictive control (FCS-MPC) uses the fact that the operation principle of power converters offers only limited count of switching states, i.e. three states for H-bridge and eight states for three phase inverter. Therefore, only limited counts of voltage values can be applied on the load. These values are used for prediction of system behavior for each switching state. The FCS-MPC controller then chooses the best action to achieve given demands (e.g. speed control [2], [3], torque ripple minimization [4] or even position control [5]).

In the conventional control methods, output voltage of the converters is commonly created by using pulse-width modulation or by space-vector modulation (in case of three-phase drives) [6]. Pulse-width modulation causes regular switching of all switching elements, e.g. at switching frequency of 10 kHz with variable switch-on/off time. FCS-MPC on the other hand has variable switching frequency, with maximum frequency equal to the system sampling frequency, but in transient states it is also possible to hold only one selected switching state for several sampling cycles. Therefore it is possible to achieve higher dynamics than by conventional control methods.

In this contribution, FCS MPC, applied on a permanent magnet DC (PMDC) motor, is presented in order to describe principle and possibilities of this type of predictive control.

In the second section, a PMDC motor model is described. The third section is dedicated to predictive control design and in the fourth section, the proposed control is verified by simulation.

II. PMDC Motor Model

State space model of PMDC motor used for proposed type of control is as follows:

$$\frac{d}{dt}\begin{bmatrix} i_a \\ \omega \end{bmatrix} = \begin{bmatrix} -\frac{R_a}{L_a} & -\frac{k_t}{L_a} \\ \frac{k_t}{J} & -\frac{B_\omega}{J} \end{bmatrix}\begin{bmatrix} i_a \\ \omega \end{bmatrix} + \begin{bmatrix} \frac{1}{L_a} \\ 0 \end{bmatrix} u_a + \begin{bmatrix} 0 \\ -\frac{1}{J} \end{bmatrix} T_L \tag{1}$$

where: i_a is the armature current, ω is the rotor angular speed, R_a is the armature resistance, L_a is the armature inductance, k_t is the torque constant, J is the motor inertia, B_ω is the friction coefficient and T_L is the load.

FCS-MPC is a discrete time control, therefore the model of PMDC motor was discretized by Taylor series expansion.

III. Predictive Control Design

PMDC motor is driven by conventional H-bridge with four switching elements, i.e. four possible switching states. In fact, PMDC motor can be connected to positive, negative or zero voltage (there are two possible switching states for zero voltage).

According to actual motor states, future states are predicted one step forward for each switching state by using discrete motor model. Predicted values are then compared by cost function; the one with minimal value of the cost function is selected.

For this purpose, following cost function to control both angular speed and motor current was designed:

$$J_p = \lambda_1\left(\omega_{set}(k+1) - \omega(k+1)\right)^2 + \lambda_2\left(\hat{T}_L(k)/k_t - i_a(k+1)\right)^2 + f_{i\lim} \quad (2)$$

where: λ_i is the weighting factor, ω_{set} is the requested angular speed, $\hat{T}_L$ is the estimated motor load and $f_{i\lim}$ is a nonlinear function for current limitation.

The first term in (2) penalizes speed reference tracking error one step forward, whereas the second term controls motor current to minimalize current oscillations. Controlling motor current in such a way also increases motor dynamics in case of load change. Finally, the last term strongly penalize the switching state, which would cause motor overcurrent.

Current limitation is defined as:

$$f_{i\lim} = \begin{cases} 0 \ldots \text{for } i_a < \mathrm{I_{MAX}} \\ Inf \ldots \text{for } i_a > \mathrm{I_{MAX}} \end{cases} \quad (3)$$

where I_{MAX} stands for maximal allowed value of the motor current. In the implementation, a very large constant is used instead of infinite.

Weighting factors in cost function are determined by importance of each term. Incorrectly chosen values could lead to static error in speed following or in large overcomes in waveforms.

In order to accurately predict motor states, the disturbances applied on the motor have to be known or online identified. For that purpose, Kalman filter with torque estimation was designed.

Although the Kalman filter is difficult to tune and needs a lot of computational time, its usage has few important advantages. The first advantage is very good noise filtering capability and the second is the possibility to observe some system variables. Both capabilities can be done by using only single Kalman filter.

The control structure is depicted in Fig. 1, where LKF stands for linear Kalman filter in the form of predictor - corrector.

IV. Simulation Results

Proposed predictive control was verified by simulation, where step changes of the velocity demand and load were applied (Fig. 2). Firstly, in the time of 0.2 s velocity step change to 50 rad.s^{-1} was requested. It can be seen, that the controller chose switching state that applies to the positive voltage for several sampling instants, until motor current reached a maximum value. After that, the maximum value is held by switching between zero and positive voltage, until velocity approaches requested value. In the steady state, without the load, all three switching states can be observed.

At the time of 0.5 s, a load step change to rated torque was applied. As can be seen, this disturbance is quickly compensated.

V. Conclusion and Future Work

Proposed control method with speed and current control by one controller ensures very high dynamics of the controlled system. FCS MPC has also an ability to control more system variables by a single controller in opposite to conventional cascade control structure, where one controller for each controlled variable is needed. Presented principle of FCS MPC can be easily applied also on different types of load and power converters as well.

Disadvantages of such control are computational power requirements and sensitivity to predictive model accuracy.

Future work is focused on applying proposed control method on permanent magnet synchronous motor fed by a three phase inverter with experimental verification as well.

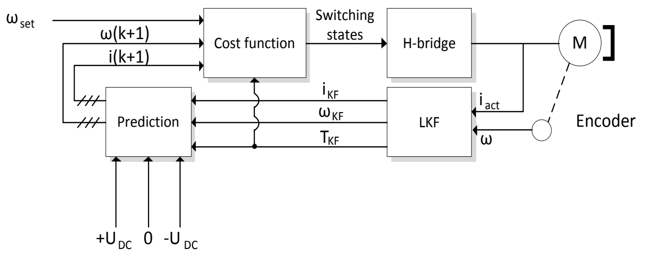


Fig. 1 Finite control set model predictive control of PMDC motor

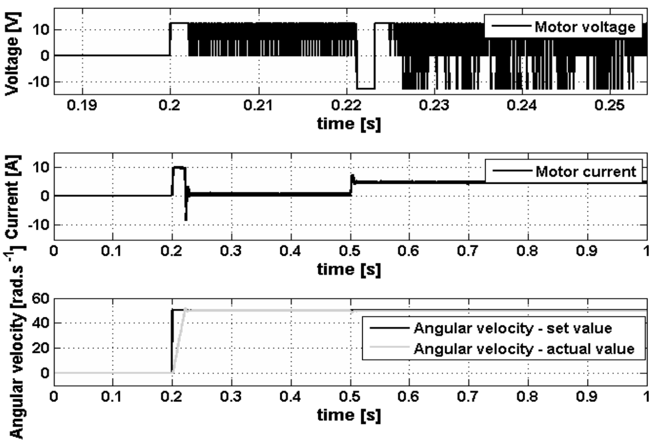


Fig. 2 Simulation results

Acknowledgment

This work was supported by project No.: VEGA 1/0121/15.

References

[1] Cortes, P.; Kazmierkowski, M.P.; Kennel, R.M.; Quevedo, D.E.; Rodriguez, J., "Predictive Control in Power Electronics and Drives," Industrial Electronics, IEEE Transactions on , vol.55, no.12, pp.4312-4324,Dec. 2008

[2] Fuentes, E.J.; Silva, C.A.; Yuz, J.I., "Predictive Speed Control of a Two-Mass System Driven by a Permanent Magnet Synchronous Motor," Industrial Electronics, IEEE Transactions on , vol.59, no.7, pp.2840-2848, July 2012 doi: 10.1109/TIE.2011.2158767

[3] Preindl, M.; Bolognani, S., "Model Predictive Direct Speed Control with Finite Control Set of PMSM Drive Systems," IEEE Transactions on Power Electronics, vol.28, no.2, pp.1007-1015, Feb. 2013 doi: 10.1109/TPEL.2012.2204277

[4] Changliang Xia; Yingfa Wang; Tingna Shi, "Implementation of Finite-State Model Predictive Control for Commutation Torque Ripple Minimization of Permanent-Magnet Brushless DC Motor," IEEE Transactions on Industrial Electronics, vol.60, no.3, pp.896-905, March 2013 doi: 10.1109/TIE.2012.2189536

[5] Fuentes, Esteban; Kennel, Ralph M., "A finite-set model predictive position controller for the permanent magnet synchronous motor," IEEE International Symposium on Sensorless Control for Electrical Drives and Predictive Control of Electrical Drives and Power Electronics (SLED/PRECEDE), 2013, pp.1,7, 17-19 Oct. 2013

[6] Mohan, N.; Nema, R.K., "Performance analysis and comparison of pulse width modulation techniques for two level inverter based induction motor drives," Annual International Conference on Emerging Research Areas: Magnetics, Machines and Drives (AICERA/iCMMD), 2014, vol., no., pp.1,5, 24-26 July 2014 doi: 10.1109/AICERA.2014.6908189

Hardware-in-the-Loop method for electric drives

[1]*Peter TALIAN (1st year)*
Supervisor: [2]Daniela PERDUKOVÁ

[1,2]Dept. of Electrical Engineering and Mechatronics, FEI TU of Košice, Slovak Republic

[1]peter.talian@tuke.sk, [2]daniela.perdukova@tuke.sk

***Abstract*—Hardware-in-the-loop simulation is an increasingly popular engineering tool for its effectiveness in maintaining a balance between the two competing demands of: a) well designed and thoroughly tested systems, and b) reduced development time and costs to remain competitive. This paper gives an introduction to Hardware-in-the-Loop (HIL) simulation and shows how HIL simulation can be used to successfully test control systems.**

***Keywords*—hardware-in-the-loop (HIL), real-time simulation (RTS), programmable logic controller (PLC)**

I. Introduction

MOTOR-DRIVE systems are an important component in industrial applications. A lot of these applications require adjustable speed drives. The demand considering the reliability, safety, flexibility, and capability of those drives is increasing steadily. [1]

Hardware-in-the-Loop (HIL) simulation is a technique that is used for the development and testing of control systems which are used for the operation of complex machines and systems. With HIL simulation the physical part of a machine or system is replaced by a simulation. As the name implies, in HIL simulation, a part of the system is modeled and simulated in real time, while the remainder is the actual hardware, connected in closed loop through various I/O interfaces such as analog-to-digital (A/D). Hardware-in-the-loop (HIL) simulations have been successfully used as essential prototyping means for system development and testing in a wide range of industrial automation and engineering fields, such as aerospace [3], automotive [4], controls [5], manufacturing [6], and naval and defense. [2]

Nowadays more and more developed using this methodology. HIL simulation enables thus to check availability and reliability of drives (machines, power electronics and control) before their insertion on a whole system. Indeed, implementation constraints are taken into account such as sensor accuracy, sampling period, modulation frequency, active limitations and so on.

II. Brief historical overview

One of the first uses of hardware-in-the-loop (HIL) simulation was for flight simulation. Hardware-in-the-loop simulation has found widespread use in testing missile guidance systems in the past 40 years. The Sidewinder program was one of the first to use HIL simulation back in 1972. Soon after, in parallel with missile simulations, NASA was working on the development of highly maneuverable aircraft technology (HiMAT) [7]. The purpose of the HiMAT program was to investigate the use of then advanced concepts such a reduced static stability and fly-by-wire. NASA developed a range of high-fidelity hardware-in-the-loop simulations to enable this research.

More recently, within the past 20 years, hardware-in-the-loop simulation has gained popularity in automotive industry. There hardware-in-the-loop simulation is used for design of anti-lock braking systems (ABS), traction control systems (TCS), suspension systems and others. Other applications of hardware-in-the-loop simulation are mainly in the fields of power engineering and robotics. [8]

III. HIL Simulation For Electric Drives

Hardware in the loop simulation system has advantages of saving development cost, shorting development time, so it has been applied in the field of automotive system, power electronics, electric drives, etc. [11]

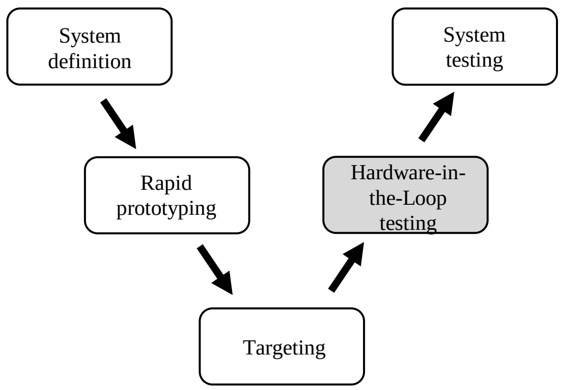


Fig. 1. Design cycle of a control system

In hardware-in-the-loop simulation systems, part of the simulation loop is composed of computer software, while the rest is the actual hardware systems. In practical control systems, the hardware in the loop can either be the controllers or the plant. For example in aerospace and military systems, the hardware in the system is usually the controllers.

However in process control systems, the hardware is usually the plant, and the controllers can be the Simulink models. One advantage of hardware-in-the-loop simulation is that the validation of control results is very straight forward. In industrial control, hardware-in-the-loop simulation techniques can significantly reduce the time required to design controllers and can increase the reliability of the systems. [12]

The conception of hardware-in-the-loop station is illustrated in Fig.2.

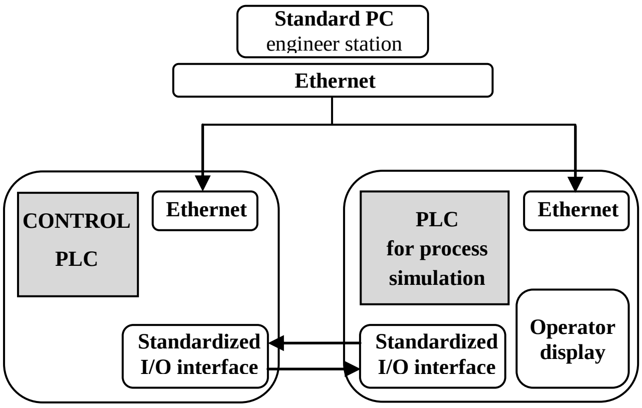


Fig. 2. Conception of hardware-in-the-loop station

The conception of station uses Ethernet connection (100 Mb/s) to communicate between the engineer station and PLCs. The Ethernet connection between target nodes allows parallel computation of models with small time step size. [12]

Fast prototype design is a new approach in controller design and implementation. The controllers can be constructed with Simulink and Stateflow, and the executable code for the controllers can be generated easily, and the parameters of the controllers can be tuned on-line, according to the actual control behavior. The designed controller can be regarded as a prototype, and once the control results are satisfactory, the code can be downloaded to the real controllers and the control actions can be carried out without the use of MATLAB or Simulink. [12]

IV. HIL SIMULATIONS CONCEPTS FOR ELECTRIC DRIVES

An electrical drive can be divided into several subsystems (Fig. 3): the process control, the power electronics set, the electrical machine and the mechanical load to move (the mechanical power train of a vehicle for example). Power devices are connected according to the action and reaction principle. A controller board contains the process control and yields the switching orders of the power electronics converter. [16]

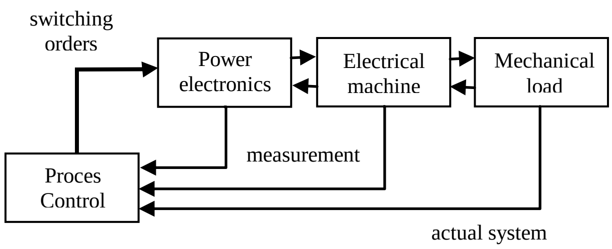


Fig. 3. Subsystems of an electric drive

Measurements of all power parts are inputs for this controller board. In some cases, several controller boards are used. In other case analog devices as FPGA are used to control the faster dynamics and to achieve high-frequency modulations of power electronics.

A. Signal level HIL simulation

Only the controller board (which contains the process control) is tested (Fig. 4). The other parts (power electronics, machine and mechanical load) are simulated in real-time. The simulation system must manage inputs and outputs of the controller board under test. A second controller board is thus used to simulate in real-time the power parts of the system. A specific signal conditioning is required to impose the same inputs and outputs as imposed by the power parts. This method can be called “signal level HIL simulation” because only signals are used at the interface between the system under test and the simulation environment. This kind of HIL has been very often employed in aerospace and automotive applications for assessment of controller boards. [16]

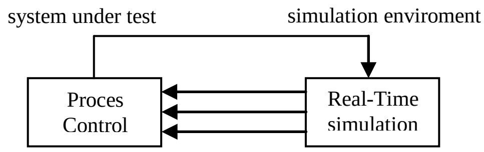


Fig. 4. Signal level HIL simulation

B. Power level HIL simulation

In the second case, the actual controller board and the power electronics converter are evaluated. The other parts (electrical machine and mechanical load) are simulated. The simulation system must impose inputs and outputs for the power electronics and the controller board under test. The simulation environment is generally composed of a second power electronics set (electric load) and a second controller board (real-time simulation) (Fig. 5). This method can be called “power level HIL simulation”. Indeed the interface between the system under test and the simulation environment require signal and power variables. [16]

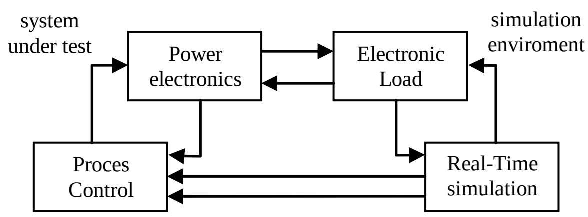


Fig. 5. Power level HIL simulation

C. Mechanical level HIL simulation

In the last case, the whole drive (control, power electronics and electric machine) is tested and the mechanical part is simulated. The simulation system must impose mechanical inputs and outputs to the electrical machine under test. Moreover, measurements on the mechanical part have to be sent to the controller board under test. Another electrical machine (load machine) is often used as controlled mechanical load. It is supplied by a second power electronics set (load supply). A second controller board (real-time simulation) is required to control the load machine and to send fictitious mechanical "measurements" to the controller board under test (Fig. 6).

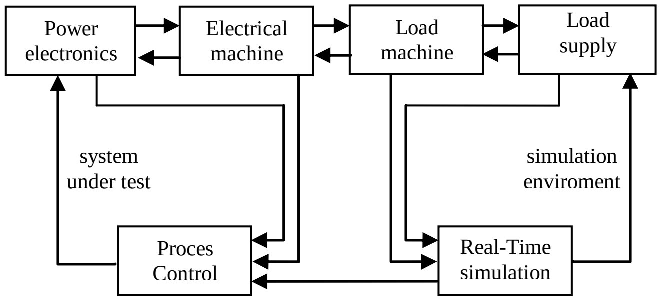


Fig. 6. Mechanical level HIL simulation

This method can be called "mechanical level HIL simulation". Indeed the interface between the system under test and the simulation environment correspond to mechanical variables. [16]

D. *Control System*

Most control systems in the offshore and marine sector are implemented on PLC's. A Programmable Logic Controller or PLC is a digital computer dedicated for the control of machines and systems. PLC's have a widespread use in industry because of their rugged design which make them suitable for operating in severe conditions.

Most PLC brands support these programming languages: function block diagram (FBD), ladder diagram (LD), structured text (ST), instruction list (IL), sequential function chart (SFC) or programming in C-code (#C).

PLC's can use a wide variety of fieldbus systems. Some PLC brands only support their own "closed" fieldbus system while others support a wide variety of "open" fieldbus system.

V. APPLICATIONS

A. *Produce of HIL simulation*

The main steps in HIL Simulation are as follows:

1. Develop a mathematical model. Create a mathematical model of the real environment where the hardware device is meant to be used.
2. HIL Simulation (Software + Hardware). Test your device on a simulated process (mathematical model).
3. Implement your hardware on the Real Process (Hardware only). If everything is OK, you may want to implement your hardware device in the real environment where it meant to be used.

These tasks follow the main idea with a HIL simulation. First step is to simulate your system in software. Next is to test your hardware on the simulated process. Finally you implement your hardware on the real system. [15]

HIL simulation is widely used in developing Embedded Control Systems, such as:

- Medical Devices
- Industrial machines
- Power Generation Systems
- White Goods
- Aerospace
- Automotive
- Process Control

B. *HIL simulation in automotive*

HIL simulation for electric drives has been done for many years now. In automotive HIL simulation, it became more and more important from about 2004 onwards, due to the increasing development efforts for hybrid-electric vehicles. Today, controlled electric drives are used for a large variety of different important and safety-critical systems in modern vehicles. They can be found in hybrid-electric or electric powertrains as well as in electric steering systems or gear box actuation. Most of these applications have in common that they incorporate a complex distributed control system (hardware and software) comprising several ECUs and having significant requirements with respect to reliability and safety. HIL testing is therefore an obvious choice, and many different solutions have been presented. [20]

C. *HIL in power electronics*

Hardware-in-the-Loop Simulation for Power Electronics systems is the next quantum leap in the evolution of HIL technologies. The ability to design and automatically test power electronics systems with HIL simulations will reduce development cycle, increase efficiency, and improve reliability and safety of these systems for large number of applications. Indeed, power electronics is an enabling technology for hybrid electric vehicles, electric vehicles, variable speed wind turbines, solar photovoltaics, electric trains etc. Real-time simulations of switching transitions require digital processor speeds and latencies that can actually be met with off-the-shelf computer systems and with FPGA/CPU platform technologies making it 100 times faster than traditional computational methods to achieve high-resolution HIL for power electronics.

D. *DC Motor Drive*

The main Simulink block diagram which was used for real time closed-loop speed control of a permanent magnet DC motor with an inner current loop is on Fig 7.

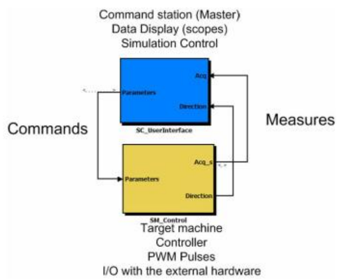


Fig. 7. The Simulink block diagram for PMDC motor control

The model consists of two subsystems, the first subsystem model of DC motor runs on the PLC PP550 and the second subsystem speed and current controller run at the PLC X20CP1484-1.

The simulation results are then transferred to the operator display on PLC PP550for display. The current loop has a PI controller with gains calculated to cancel the pole due to the electrical time constant of the DC motor, and to give the required bandwidth. The speed controller also has a PI controller designed to achieve a certain phase margin and bandwidth. The motor is coupled with another PMDC motor which functions as a generator.

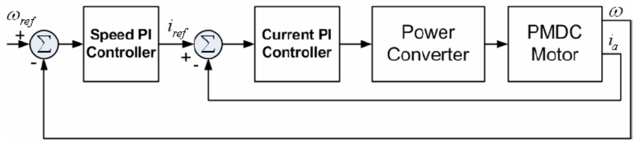


Fig. 8 Block diagram of speed-current controller for an PMDC machine

The generator is supplying a variable resistive load. The controller parameters are controlled by the command station GUI. A fixed-step sampling time was chosen for this setup to avoid overrun. Fig. 8 shows the block diagram of the controller. Fig. 9 shows the Simulink implementation of the controller block diagram.

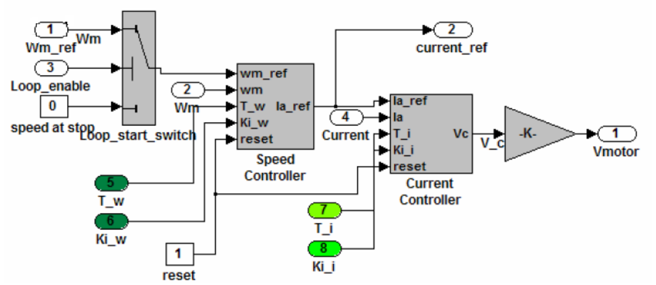


Fig. 9 The controller implementation in Simulink

The PWM signals are generated by RT-EVENTS and then send to a control board that runs the DC motor. The speed is sensed using an optical encoder, and the current is sensed and fed to one of the ADC channels of the PLC. This block diagram is compiled and transferred to the PC. The collected data are then sent to the command station for display. A square pattern change in the speed reference was given using the graphical real time interface. [20]

VI. Conclusion

This paper offers an overview about hardware-in the loop methods for electric drives. HIL simulation reduces development time and enables various tests that cannot be achieved on the actual system for cost or security reasons (fault operation for instance). “Signal level HIL simulation” is very often used in industry to check controller boards and process controls. Its “power level” and “mechanical level” extensions are growing because they are promising intermediary step before integration of electric drives in actual systems. Because of the increasing complexity of the systems under test, the organization of the HIL simulator is of prime importance to assess the best performance. This knowledge I would to use for my research. The main goal of PhD. study is intelligent methods of complex drives control using method HIL.

References

[1] S. Grubic, B. Amlang, W. Schumacher, A. Wenzel, “A High-Performance electronic Hardware-in-the-Loop Drive-Load Simulation Using a Linear Inverter ” In: *IEEE Trans. Industrial Electronic*, vol.57, No. 4, pp. 1208–1216, April 2010.

[2] A. Martin, M. Reza Emani , “An Architecture for Robotic Hardware-in-the-Loop Simulation” *IEEE international conference on Mechatronics and Automation* June 25 - 28, 2006, Luoyang, China.

[3] J. Leitner, “Space Technology Transition Using Hardware in the Loop Simulation,” inProceedings of the 1996 Aerospace Applications Conference, Vol. 2, February 1996, pp. 303-311.

[4] H. Hanselman, “Hardware-in-the-loop Simulation Testing and its Integration into a CACSD Toolset,” *IEEE International Symposium on Computer-Aided Control System Design*, September 1996, pp. 15-18.

[5] M. Linjama, T. Virvalo, J.Gustafsson, J. Lintula, V. Aaltonen, and M. Kivikoski, “Hardware-in-the-loop Environment for Servo System Controller Design, Tuning, and Testing,” In: *Microprocessors and Microsystems*, Vol. 24, No. 1, 2000, pp. 13-21.

[6] G. Stoeppler, T. Menzel, and S. Douglas, “Hardware-in-the-loop Simulation of Machine Tools and Manufacturing Systems,” In: *Computing and Control Engineering Journal*, Vol. 16, No. 1, 2005, pp. 10-15.

[7] B. L. Ballard, R. E. Elwell Jr., R. C. Gettier, F. P. Horan, A. F. Krummenoehl, and D. B. Schepleng, “Simulation Approaches for Supporting Tactical System Development,” *John Hopkins APL Technical Digest* (Applied Physics Laboratory), Vol. 23, No. 2-3, 2002, pp. 311-324.

[8] M. Bacic, “On hardware-in-the-loop simulation”, In: *44th IEEE Conference on Decision and Control, and the European Control Conference 2005*, Seville, Spain, December 12-15, 2005.

[9] R. Isermann, J. Schaffnit, and S. Sinsel, “Hardware-in-the-loop simulation for the design and testing of engine-control systems,” In: *Control Engineering Practice*, vol. 7, pp. 643-653, 1999.

[10] J. S. Cole and A. C. Jolly, "Hardware-in-the-loop simulation at the US Army Missile Command," In: *Technologies For Synthetic Environments: Hardware-In-The-Loop Testing*, vol. 2741, Proceedings Of The Society Of Photo-Optical Instrumentation Engineers (Spie), 1996, pp. 14-19.

[11] Y. Liao, X. Shi, Ch. Fu, J. Meng, “Hardware In-the-Loop Simulation System Based on NI-PXI for Operation and Control of Microgrid”, In: *Renewable Energy Research and Applications* (ICRERA), Nagasaki, Japan, pp. 1-6, Nov. 2014.

[12] D. Xue, C. YangQuan, “Hardware-in-the-loop Simulation and Real-time Control,” *System Simulation Techniques with MATLAB and Simulink*, First Edition. 2014 John Wiley & Sons, Ltd. Published 2014 by John Wiley & Sons, Ltd.

[13] A. M. Osama, N. Y. Abed, "Real-time simulation of electric machine drives with hardware-in-the-loop", *COMPEL - The international journal for computation and mathematics in electrical and electronic engineering*, Vol. 27 Iss: 4, pp.929 – 938, 2008.

[14] T. Schulte, J. Bracker, „Real-Time Simulation of BLDC Motors for hardware-in-the-loop Applications Incorporating Sensorless Control“, *Proc. of IEEE Int. Symp. on Industrial Electronics (ISIE)*, Cambridge, UK, June, 2008.

[15] H. P. Halvorsen, „Hardware-in-the-Loop Simulation“, Telemark University College Department of Electrical Engineering, Information Technology and Cybernetics

[16] A. Bouscayrol, ” Different types of Hardware-In-the-Loop simulation for electric drives,” In: *IEEE International Symposium on Industrial Electronics ISIE 2008*, Cambridge. pp.2146-2151, July 2008.

[17] K. Kyslan, ” Load Torque Emulator based on Industrial Convertors,” In: *SCYR - 10th Scientific Conference of Young Researchers*, Košice, 2010

[18] W. Lhomme, R. Trigui, P. Delarue, A. Bouscayrol, B. Jeanneret, F. Badin, “Validation of clutch modeling for hybrid electric vehicle using Hardware-In-the-Loop simulation”, In: *IEEE-VPPC’07*, Arlington (USA), September 2007.

[19] C. Dufour, T. Ishikawa, S. Abourida, J. Bélanger, "Modern Hardware-In-the-Loop simulation technology for fuel cell Hybrid Electric Vehicles", In: *IEEE-VPPC’07*, Arlington (USA), September 2007.

[20] T. Schulte, A. Kiffe, F. Puschmann, , "HIL Simulation of Power Electronics and Electric Drives for Automotive Applications", In: *Electronics*, Vol. 16, No. 2, December 2012

[21] T. Schulte, J. Bracker, „Real-Time Simulation of BLDC Motors for hardware-in-the-loop Applications Incorporating Sensorless Control“, *Proc. of IEEE Int. Symp. on Industrial Electronics (ISIE)*, Cambridge, UK, June, 2008.

Hybrid RF/FSO Links with Soft – Switching Configurations

[1]Peter STRNISKO (2nd year)
Supervisor: [2]Ľuboš OVSENÍK

[1,2]Dept. of Electronics and Multimedia Communications, FEI TU of Košice, Slovak Republic

[1]ing.peter.strnisko@gmail.com, [2]lubos.ovsenik@tuke.sk

***Abstract*—Hybrid RF/FSO (Radio frequency/Free space optics) links combine the advantages of RF and FSO links. FSO links offer the high (gigabit per second) data rate, but suffer from atmospheric loss due to fog and scintillation. RF links have lower data rate and they are sensitive to rain or wet snow. In this paper the Soft – switching scheme is presented. Soft – switching is a method of switching between the different parts of hybrid link.**

***Keywords*—Channel model, Hybrid RF/FSO, Raptor codes, Soft – switching.**

I. INTRODUCTION

In free-space optical communication an optical carrier is employed to convey information wirelessly. FSO systems have the high data rates with the advantages of quick deployment times, high security, and no frequency regulations. Unfortunately such links are highly susceptible to atmospheric effects. Scintillation induced by atmospheric turbulence causes random fluctuations in the received signal. One method to improve the reliability of these systems is to introduce a radio frequency link to create a hybrid RF/FSO communication system. When the FSO link is blocked by cloud or fog, the RF link maintains reliable communications, but at a reduced data rate. Typically a millimeter wavelength carrier is selected for the RF link to achieve data rates comparable to that of the FSO link. At these wavelengths, the RF link is also subject to atmospheric effects, including rain and scintillation, but less affected by fog. The two channels are therefore complementary: the FSO signal is severely attenuated by fog, whereas the RF signal is not; and the RF signal is severely attenuated by rain, whereas the FSO is not. Both, however, are affected by scintillation [1].

There are two possible configurations for a hybrid RF/FSO link. The first of them is hard - switching configuration. The transmitter and receiver jointly select either the FSO or RF channel for data transmission. The high data rate FSO link is selected only if channel conditions permit reliable communications, otherwise all data are sent over the RF channel. The transmitter and receiver must be coordinated via feedback to select the correct channel for transmission. A key disadvantage of hard-switching is that at any time one link is sitting idle. In practice, the sensitive FSO channel will cause it to be selected rarely [2]. The second configuration is the soft – switching. This configuration will be described below.

The paper is organized as follows: section 2 presents the hybrid RF/FSO communications, primary FSO link and backup RF link, section 3 describes the soft – switching and Raptor codes and finally conclusion is given.

II. HYBRID RF/FSO COMMUNICATIONS

For the point-to-point hybrid RF/FSO system the channel model is a combination of two individual line-of-sight RF and FSO channels [2].

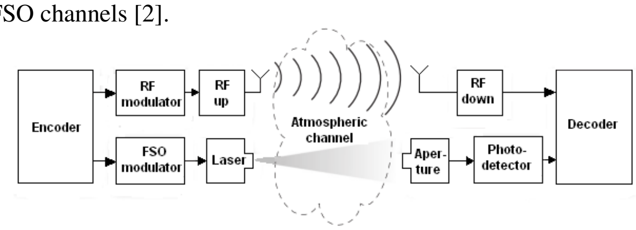


Fig. 1 Hybrid RF/FSO communication system [1]

First, a binary message sequence is jointly encoded into parallel FSO and RF bit streams. Each bit stream is modulated according to a discrete modulation scheme, up-converted to its respective carrier frequency, and transmitted simultaneously through an atmospheric channel to the hybrid receiver. The hybrid receiver simultaneously down-converts each carrier signal, demodulates, and jointly decodes the received sequences to recover the original message [1]. Fig. 1 presents the block diagram of hybrid communication system.

A. *FSO Channel model*

In FSO communication systems, data are transmitted by modulating the instantaneous intensity of a laser source and detecting the received intensity with a photodiode. A conventional channel model is:

$$\gamma = hR\chi + n \qquad (1)$$

where γ a χ is the optical intensity of transmitted and received signal, h is the channel gain, R is receiver responsivity and n is noise. Noise is modeled as Gaussian noise [2].

B. *RF Channel model*

RF link is considered as a complementary link to the FSO channel [3]. Point-to-point RF links achieve several km transmission distance at rates of tens to hundreds of Mbps [4]. In most of the situations the RF channel is modeled as a fading free, additive white Gaussian noise channel.

III. Soft – Switching Using Raptor Codes

As mentioned above, there are two possible configurations for a hybrid RF/FSO link. Hard – switching and soft – switching, as presented in Fig. 2. To overcome the hard – switching disadvantage is to coordinate data transmission in both link (RF and FSO) using channel coding [2], [3].

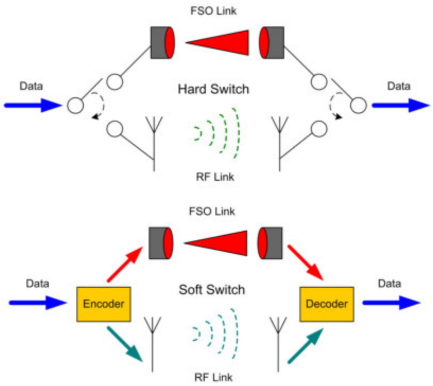


Fig. 2 Configurations for hybrid RF/FSO links [2]

In this case data are encoded by a single LDPC (Low Density Parity Check) code with a portion of the codeword split to RF and FSO links and the rate is adjusted via puncturing according to instantaneous channel conditions. Although this technique improves over hard – switching, channel conditions must be known at transmitter and receiver and complex soft decoding is required which is difficult at FSO data rates [2]. Raptor codes can improve this problem. These codes do not require channel knowledge at the transmitter and are able to adapt their rate to channel conditions with only a single bit feedback per message [4], [5]. Use Raptor codes for hybrid RF/FSO communication links is detailed described in [6], [7]. In [6] a bit – wise Raptor coding is considered in which bits transmitted on FSO and RF links are random linear combinations of message bits [2]. Additionally, soft iterative detection is required at the receiver and the impact of varying atmospheric conditions is not considered.

A. *Design of Short - length Raptor codes*

A packet – level coding technique using Raptor codes is described here. A sequence of Raptor – encoded outputs are generated from the message packets and sent over RF and FSO links simultaneously. There is no coordination between RF and FSO links and packets are sent either on RF or FSO channels at their respective transmission rates. The receiver collects packets that were not discarded by the equivalent RF/FSO erasure channels and the entire message is decoded when sufficient packets are received. It is assumed that the transmitter continues sending encoded packets for a given message until a 1 – bit feedback is received to proceed to the next message. We consider a set of short – length $k = 16 – 1024$ Raptor codes [2]. Raptor codes are random bipartite graph codes in which every encoded packet is a linear combination of a random number of message packets [7]. A message consists of k packets each of length l bits. The degree of an encoded packet is the number of message packets combined, and is chosen randomly and independently for each encoded packet. Raptor codes consist of an LT (Luby Transform) code and a pre-coding LDPC code. We can write:

$$\langle m \rangle = (1+\varepsilon_m)k \qquad (2)$$

where m is the number of received packets necessary for successful decoding of a k packet message, ε_m is the statistical overhead and $\langle \bullet \rangle$ denotes expectation [2].

IV. Conclusion

In this paper a soft – switching configuration for hybrid RF/FSO links with packet – level Raptor codes is presented. The length of Raptor code is $k = 16 – 1024$. This short length achieves high data rate and low decoding cost. We know, from the available literature, that the hardware encoder and decoder for $k = 16$ can be implemented in an FPGA (Field programmable gate array) to demonstrate the practicality of soft – switching in hybrid RF/FSO links. Soft – switching hybrid RF/FSO links based on short – length Raptor codes are practical and efficient.

This paper is the review of the state of the art of this problem and in the future my work will be devoted to hybrid RF/FSO communication systems with soft – switching configuration and MIMO (Multiple Input Multiple Output) communication system. I will test different types of channels, modulation, coding and their combination. The results of these different types of transmission systems I will compare and evaluate. Also, I will deal the application of MIMO hybrid RF/FSO communication systems. Appropriate transmission system, I would like verified experimentally.

Acknowledgment

This work was supported by Cultural and Educational Grant Agency (KEGA) of the Ministry of Education, Science, Research and Sport of the Slovak Republic under the project no. „059TUKE-4/2015 - Implementation of the FSO Long-Term Experimental Measurement Results into the Educational Process in Optoelectronic Sensor and Communication Systems Field“, and project no. „006TUKE-4/2014 - The Use of TUKE PON Experimental Model in Teaching”. This work is also the result of the contract no. “APVV-0025-12-Mitigation of Stochastic Effects in High-Bitrate All Optical Networks” supported by the Slovak Research and Development Agency.

References

[1] S. Arnon, J. R. Barry, G. K. Karagiannidis, R. Schober, M. Uysal, “*Advanced Optical Wireless Communication Systems*,” Cambridge University Press, 2012.

[2] W. Zhang, S. Hranilovic, C. Shi, “*Soft-Switching Hybrid FSO/RF Links using Short-length Raptor codes: Design and Implementation,*” IEEE Journal on Selected Areas in Communications, Vol. 27, No. 9, December 2009

[3] Tsunami MP.11 Model 5054 Technical Specifications, ” http://www.proxim.com/products/bwa/multipoint/mp11“, 2009.

[4] “GigaBeam 5.8 Series PtP datasheet,” http://www.gigabeam.com/PDFs/GigaBeam5.8PtPdatasheet.pdf. [online], last accessed: Mar. 30, 2009.

[5] S. Vangala , H. Pishro-Nik, “*Optimal Hybrid RF-Wireless Optical Communication for Maximum Efficiency and Reliability*,” Proc. IEEE Annu. Conf. Inform. Sci. Syst., pp. 684–689, 2007.

[6] A. AbdulHussein, A. Oka, and L. Lampe, “*Rateless Coding for Hybrid Free-Space Optical and Radio-Frequency Communication*,” submitted toIEEE Trans. Wireless, under review.

[7] A. Shokrollahi, “*Raptor codes*,” IEEE Trans. Inf. Theory, vol. 52, no. 6, pp. 2551–2567, Jun. 2006.

Impact of Degradation Mechanisms for all Optical Fiber Communication Systems

[1]*Ján RUŽBARSKÝ (2nd year),*
Supervisor: [2]*Ján TURÁN*

[1,2]Dept. of Electronics and Multimedia Communications, FEI TU of Košice, Slovak Republic

[1]jan.ruzbarsky@tuke.sk, [2]jan.turan@tuke.sk

***Abstract*—In order reduction means for the transmission of information technology has been developed to allow the transmission of large amounts of data when using existing transmission path. A concrete example is the compression channels of the WDM technology. Currently, the practice of using optical systems with a transmission capacity of 16 Gbps, at very short distances on the order of meters can be achieved by 40 Gbps. With increasing transmission speeds of light waveguides exhibit undesirable nonlinear effects. The aim of postgraduate studies is the investigation of nonlinear phenomena in optical networks.**

***Keywords*—Nonlinear effects, xWDM systems, Stimulated scattering, OADM.**

I. Introduction

Nonlinear phenomena are a hot topic in particular the fact that their emergence me closely connected with increasing transmission speed. Among the nonlinear effects are nonlinear phase modulation and stimulated inelastic scattering. For further investigated stimulated scattering, thus stimulated Raman scattering (SRS) and stimulated Brillouin scattering (SBS) [1].

In a simple quantum-mechanical picture applicable to both SRS and SBS, a photon of the incident field (called the pump) is annihilated to create a photon at a lower frequency (Belonging to the Stokes wave) and a phonon with the right energy and momentum to conserve the energy and the momentum. Of course, a higher-energy photon at the so-called anti-Stokes frequency can also be created if a right of phonon energy and momentum is available. Even though SRS and SBS are very similar in their origin, different dispersion relations for acoustic and optical phonons lead to some basic differences between the two. A fundamental difference is that SBS in single-modefiers occurs only in the backward direction Whereas SRS can occur in both directions [2], [3], [4].

Area of study encompasses inter nonlinear stimulated scatterings also technology of optical add/drop multiplexers (OADM). The investigation is not focused only on stimulated scatterings but as well as their impact on the function of add/drop multiplexers. The aim in the near future is to design its own implementation for add/drop multiplexers, which will seek to suppress non-linear effects [5], [6], [7], [8].

II. Tasks were already solved

In the previous time, all efforts was focused to obtain a sufficient amount of study materials in contemporary issues and understanding the basic operation of the technology. Developments in optical networks is continuous and retain the current overview is therefore very important. The basic objectives were to study the theory of nonlinear phenomena incipient in optical fiber systems. The next item was understanding and studying of basic types of optical add/drop multiplexers (OADM), which are already an integral part of the optical transmission systems. Last but not least, standards of Coarse WDM (CWDM) and Dense WDM (DWDM) systems were studied [9], [10], [11].

The sources for the study were not only books in the field, but also articles published in international conferences and proceedings. Unsubstitutable role played the documents of optical components' manufacturers and vendors.

III. Knowledges and results obtained during a previous year

Last year was devoted to the study of nonlinear effects and finding their practical measurements. The area of optical communications is now highly developed and it is possible to see in household. Availability for experimental purposes, however, a major problem especially in terms of finances. In academic conditions of the Technical University we will work with simulation program OptSim. A part of experiments will be interactive environment Matlab.

OptSim is an intuitive modeling and simulation environment supporting the performance evaluation and design of the transmission level of optical communication systems. A library of simulation algorithms, including more than 400 of the most widely used optical and opto-electronic components, is distributed with OptSim. OptSim end users are typically telecom and datacom R&D staff focused on developing and deploying key infrastructure for access, metro, and long-haul networks, including equipment for DWDM, optical amplification, all-optical network configurations, and link transmission optimization.

The OptSim interface to Matlab enables OptSim users to run Matlab routines that fully interact with the OptSim data structure, either for pre- and post-processing or for co-simulation purposes. The interface enables users to write new Matlab routines for use with OptSim co-simulation or to reuse their own Matlab proprietary code.

A. *Simulations in OptSim program*

Environment of OptSim program provides large amount of different configurations. After installation it is possible to run and to simulate base programs which are available immediately after installation. OptSim allows to create own versions of simulation programs. Scalability is a big advantage of this program.

Recently, simulations focused on optical add/drop multiplexers' presentation of basic features have been executed. OADM can drop (separate) any selected wavelength or wavelengths from transmitted signal. Subsequently, OADM can add new signals on the wavelengths which were droped in the previous step. There is a possible to add also new signals on new wavelengths that have not been used up to now.

In the Fig. 1. is showed basic model of transmission system with OADM. In the transmission path is located optical add/drop multiplexer which one signal drops and at the same time adds another signal [12].

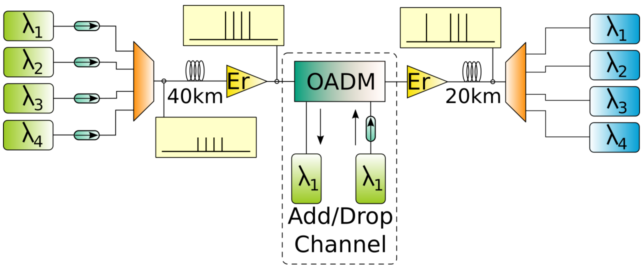


Fig. 1. Basic block diagram of transmitted signal.

The Fig. 2. shows spectrum of transmitted signal. There are four signals which have follow carrier frequencies tx_1 = 193.35THz, tx_2 = 193.45THz, tx_3 = 193.55THz, tx_4 = 193.65THz. The frequencies in this model meet standard DWDM with frequency grid 100GHz.

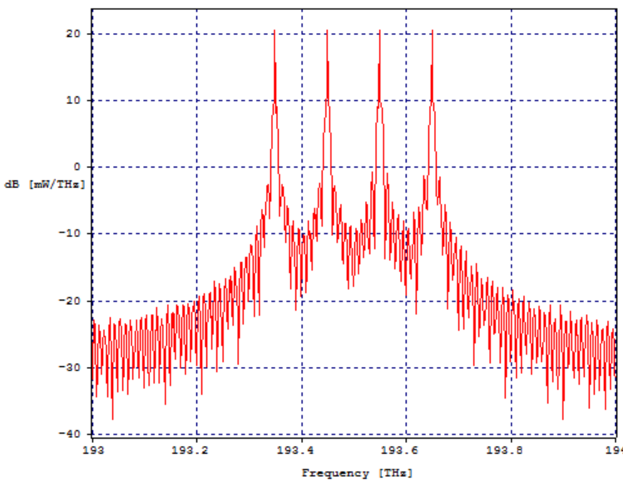


Fig. 2. Spectrum of transmitted signal.

Fig. 3. shows spectrum of signals processed by OADM. New frequency at the value 193.15THz is added to the transmission path.

IV. Conclusion

Our experiments were the first step towards meeting dissertation theses. Simulations performed so far modeled the recommendations and standards for selected technology. In the coming period will continue the simulation with standard CWDM and DWDM. The aim will monitor the impact of nonlinear phenomena on transmitted signals and optical add/drop

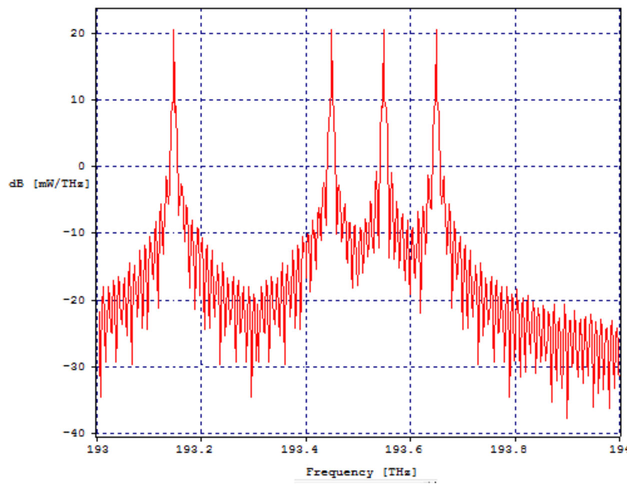


Fig. 3. Spectrum of transmitted signal after one signal is dropped and another is added.

multiplexers. Acquired results will be continuously published in journals, anthologies of domestic and international conferences.

Acknowledgment

This work was supported by Cultural and Educational Grant Agency (KEGA) of the Ministry of Education, Science, Research and Sport of the Slovak Republic under the project no. 006TUKE-4/2014. This work is also the result of the contract no. APVV-0025-12 - Mitigation of Stochastic Effects in High-Bitrate All Optical Networks supported by the Slovak Research and Development Agency.

References

[1] M. F. S. Ferreira, *Nonlinear Effects in Optical Fibers*. New Jersey : John Wiley & Sons, Inc., 2011, 1st. ed.

[2] G. P. Agrawal, *Nonlinear Fiber Optics*. Rochester, NY : Academic Press, 2013, 5th. ed.

[3] J. Ruzbarsky and J. Turan, *Influence of Degradation Mechanisms in a Fully Optical Fiber Communication Systems*, ser. SCYR 2014. 14th Scientific Conference of Young Researchers, Herlany, 2014.

[4] J. Ruzbarsky, T. Ivaniga, J. Turan, and L. Ovsenik, *Optical Networks FTTx and Reduced Attenuation Balance with Passive Optical Splitter*, ser. Thomson Delmar Learning. Carpathian Journal of Electronic and Computer Engineering, 2014, vol. 7, no. 1.

[5] W. I. WAY, *Optimum Architecture for MxN Multicast Switch-Based Colorless, Directionless, Contentionless, and Flexible-Grid ROADM*. the National Fiber Optic Engineers Conference, 2011.

[6] B. Clouet, R. Shimpe, and A. Schex, *On the Design Trade-Offs of ROADM Architectures*. ITG-Fachbericht 248: Photonische Netze, Leipzig, 2014.

[7] T. Watanabe, K. Suzuki, and T. Takahashi, *Multicast Switch Technology that Enhances ROADM Operability*. NTT Technical Review, 2014, vol. 12 No. 1.

[8] C. Pulikkaseril, L. Stewart, M. A. F. Roelens, G. W. Baxter, S. Poole, and S. FRISKEN, *Spectral modeling of channel band shapes in wavelength selective switches*. Optical Society of America, 2011, vol. 19, No. 9.

[9] ITU, *Transmission characteristics of optical components and subsystems, Recommendation ITU-T G.671*, ser. Series G. International Telecommunication Union, 2012, 6th ed.

[10] ——, *Spectral grids for WDM applications: DWDM frequency grid, Recommendation ITU-T G.694.1*, ser. Series G. International Telecommunication Union, 2012, 2nd ed.

[11] ——, *Multichannel DWDM applications with single-channel optical interfaces, Recommendation ITU-T G.698.1*. International Telecommunication Union, 2006, series G.

[12] J. Ruzbarsky and J. Turan, *Optical Add/Drop Multiplexers for All-optical Fiber Communication*, ser. MIPRO 2015. 38th International Convention on Information and Communication Technology, Electronics and Microelectronics, 2015, will be published after conference.

Influence of state of charge level on frequency control reserve provision by energy storage systems

[1]*Ing. Tomáš Košický (5st year)*
Supervisor: [2]Doc. Ing. Ľubomír Beňa

[1]Department of Electric Power Engineering , FEI TU of Košice, Slovak Republic

[1]tomas.kosicky@tuke.sk, [2]lubomir.bena@tuke.sk

***Abstract*— Batteries in transmission grids can provide auxiliary network services, such as primary frequency response, voltage regulation in network nodes or back up power. Battery Energy Storage Systems (BESS) is often the most limiting factor within their more extensive usage for the frequency regulation. Because of this it is necessary to adjust the system charging SoC and the performance of battery systems so that they fulfil the same requirements as those ones which are applied in case of conventional providers of primary frequency regulation. This article says more about the possibilities of using and stability of battery systems and their ability to be a part of the auxiliary services providers.**

***Keywords*— auxiliary services, frequency regulation, battery energy storage systems (BESS), battery charging level (SoC)**

I. Introduction

The share of primary regulation performance in Slovak Republic within the common transmission grid ENTSO-E is ±29MW [1]. According to the technical requirements of the transmission grid operator (TSO) SEPS the maximum share of the primary regulation performance purchase (PRV) from abroad is 30% [2]. The share of PRV ±8 MW represents an import from neighboring transmission grids. The primary regulation performance in Slovakia is provided by two blocks of nuclear power plant Mochovce. Block 3 and block 4 of the nuclear power plant Mochovce are being certified at the moment. Block 5 and block 6 of the thermal plant Vojany, steam-gas power plant Energochem Svit and some generators of hydroelectric plant Gabčíkovo. Considering the fast of a stabile delivery of the primary regulation performance it becomes more and more problematic to be able to replace some blocks of power plants, for example nuclear power plant Jaslovske Bohunice or steam-gas power plant Malženice. Furthermore, due to a performance reservation and providing of PRV, the lifetime of generators decreases. Within the European transmission system ENTSO-E, the pressure to create a common market with primary regulation energy is increasing which also would have an impact on the financial effectiveness if traditional rotary units for PRV are used [3,4].

II. Simulation

The BESS model for Slovakian transmission grid is based on weekly analysis of system frequency to analyse the influence of battery on the primary frequency regulation. On picture 1, there is a flowchart of proposed algorithm of positive or negative performance balance.

To keep the SoC system BESS we decided for a regulating network. After the first 15 minutes of providing of PRV we suppose an activation of secondary frequency regulation (SRV) and a capacity deviation by 50% or more *SoC* will be balanced within next 15 minutes.

$$A = FoC = \frac{\int_{t=1}^{t_n k} P(t)dt}{AF_{avg}}$$

$$B = FoD = \frac{\int_{t=1}^{t_n k} P(t)dt}{AF_{avg}}$$

$$C = \sum_{\mathrm{t=1}}^{t_n} \mathrm{P}_{\mathrm{REAL}\ (ODOD)}(t) = \sum_{\mathrm{t=1}}^{t_n} \mathrm{P}_{\mathrm{TARGET}\ (ODOD)}(t)$$

$$\mathrm{D} = \mathrm{P}_{\mathrm{AB}(DOD)}(t)dt = \sum_{\mathrm{t=1}}^{t_n} \mathrm{P}(t)dt - \sum_{\mathrm{t=1}}^{t_n} \mathrm{P}_{\mathrm{REAL}(DOD)}(t)$$

$$E = \sum_{\mathrm{t=1}}^{t_n} \mathrm{P}_{\mathrm{REAL}\ (DOD)}(t) = \sum_{\mathrm{t=1}}^{t_n} \mathrm{P}_{\mathrm{TARGET}\ (DOD)}(t)$$

$$\mathrm{F} = \mathrm{P}_{\mathrm{AB}(ODOD)}(t)dt = \sum_{\mathrm{t=1}}^{t_n} \mathrm{P}(t)dt - \sum_{\mathrm{t=1}}^{t_n} \mathrm{P}_{\mathrm{REAL}(ODOD)}(t)$$

$$G = \Delta P(t) = \left[\sum_{\mathrm{t=t_n}}^{t_n k} \mathrm{P}_{\mathrm{TARGET}(DOD)}(t) - \sum_{\mathrm{t=t_n}}^{t_n k} \mathrm{P}_{\mathrm{TARGET}(ODOB)}(t)\right] + \left[\sum_{\mathrm{t=t_n}}^{t_n k} \mathrm{P}_{\mathrm{AB}(ODOB)}(t) + \sum_{\mathrm{t=t_n}}^{t_n k} P_{AB(DOD)}(t)\right]$$

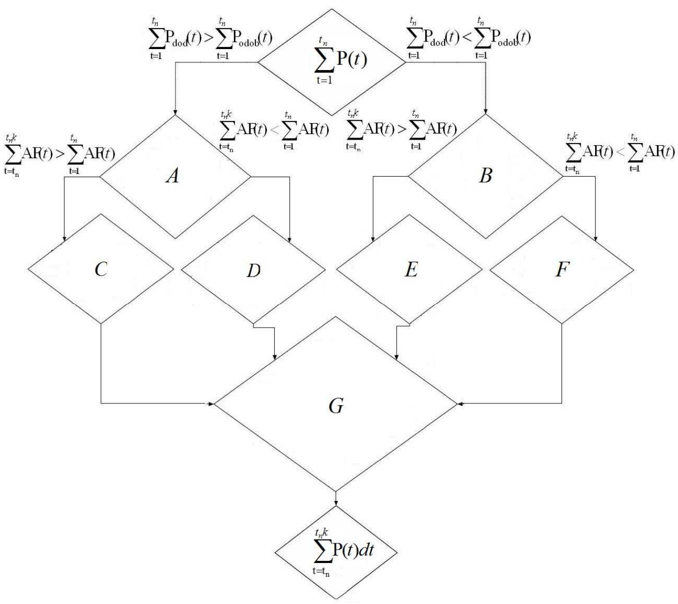


Fig.1. Flowchart of SoC balancing on 50%

As we see on Figure 2, within the analysed day the CoC kept the limits when our model was used. During some 15 minutes intervals at the end of the day, the positive deviation of system frequency caused that the batteries were being charged, the system did not make it within active points AF to reduce the SoC so the batteries were charged on a level of 91,99%. Within the intervals t_n 91 a 92, no AF were outside the dead band. During following intervals t_n 93 a 94 and with a high amount of active points, the system provided a regulation performance and the charging of SoC came back on 50%.

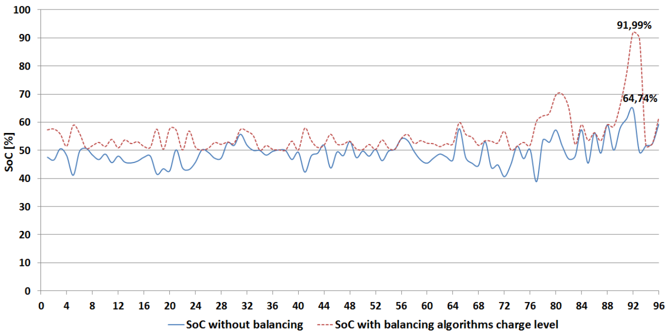


Fig. 2. Development of SoC on analysis day 13.09.2012

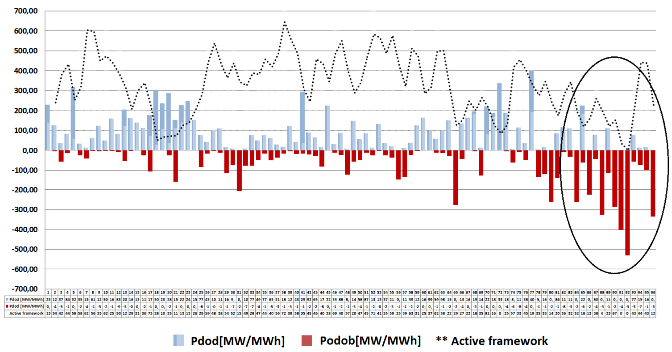


Fig. 3. 15-minutes values of delivered/ consumed performance and AF

tn [s]	Pdod [MW/MWh]	Podob [MW/MWh]	Average Framework	Active Framework [AF]	Frame of Charge [FoC]	Frame of Discharge [FoD]	SoC [%]
90	0	284,803	302	67	1,607	0	57,943
91	0	402,997	178	0	7,741	0	69,138
92	0	530,939	89	0	4,178	0	83,886
93	77,600	56,314	292	435	0	0,053	50,591
94	15,493	73,423	403	448	0,145	0	51,609
95	16,805	100,025	400	430	0,302	0	52,312
96	0	335,179	276	12	0	0	61,522

Chart 1. Balancing of SoC at the analysis day

tn [s]	Pdod [MW/MWh]	Podob [MW/MWh]	Average Framework	Active Framework [AF]	Frame of Charge [FoC]	Frame of Discharge [FoD]	SoC [%]
90	0	284,803	302	67	0,200	0	66,053
91	0	402,997	178	0	0,200	0	77,247
92	0	530,939	89	0	0,200	0	91,995
93	77,600	56,314	292	435	0	0	89,782
94	15,493	73,423	403	448	0,145	0	51,609
95	16,805	100,025	400	430	0,200	0	52,312
96	0	335,179	276	12	0	0	61,556

Chart 2. Balancing of SoC at restricted FoC, FoD

Figures 4. and 5. present performance units FoC/FoD for a 30 minutes interval, where the balancing performance was limited in a way that the real performance matches the tolerance defined by the operator:

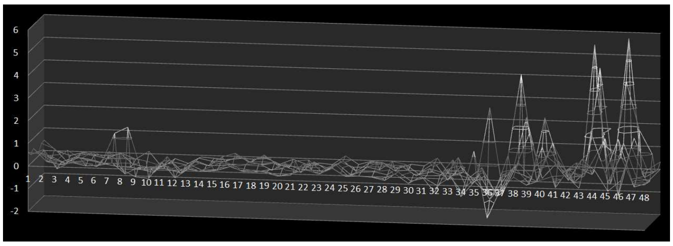


Fig. 4. Values FoC, FoD within two following 15 minutes intervals for analysed week

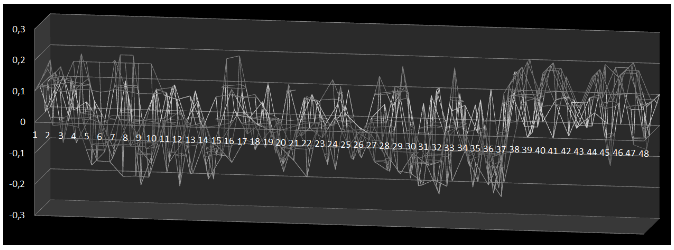


Fig. 5. Values FoC, FoD within two 15-minutes intervals in analysed week if algorithm for PRV quality is kept

III. Conclusion

In this article we analysed the strategies of usage of battery systems for primary frequency regulation. It was proved that keeping the limits of SoC can be explicit executed and there is no need to use external sources within the tolerance defined by the operator. The algorithm we applied did not cause any early capacity loss of BESS. During the simulation the limits of SoC_{min}/SoC_{max} were not exceeded. This analysis also showed the advantages of battery systems and their ability to ensure a frequency regulation under same conditions as guilty for conventional PRV providers.

IV. References

[1] ENTSO-E, Operation Handbook, 2009, ch. Policy 1: Load-Frequency Control and Performance.

[2] Technické Podmienky Prístupu a Pripojenia, Pravidlá Prevádzkovania Prenosovej Sústavy, SEPS 2012

[3] Decentralized demand-side contribution to primary frequency control.Angel Molina-Garcia, Francois Bouard, and Daniel S Kirschen. IEEE Transactions on Power Systems, 26(1):411{419, 2011. doi: 10.1109/ TPWRS.2010.2048223

[4] Optimizing a Battery Energy Storage System for Primary Frequency Control Alexandre Oudalov, Daniel Chartouni, and Christian Ohler. IEEE Transactions on Power Systems, 22(3):1259{1266, August 2007. ISSN 0885-8950. doi: 10.1109/TPWRS.2007.901459.

Inner Product Masking of Hardware AES Implementation in VHDL

Martin PETRVALSKY (3^{rd} year),
Supervisor: Milos DRUTAROVSKY

Dept. of Electronics and Multimedia Communications, FEI TU of Kosice, Slovak Republic

{martin.petrvalsky, milos.drutarovsky}@tuke.sk

***Abstract*—In this paper we developed a countermeasure against the first order differential power analysis attack. The countermeasure is based on an Inner Product masking scheme in which one 8-bit value stored in hardware is represented as four 8-bit shares. We implemented basic building blocks of AES cryptosystems. Firstly, we had to implement Galois field operations in VHDL code - addition, multiplication, squaring and inversion. After, we developed all the operations needed for AES using the Inner Product masking scheme. Moreover, we attacked the developed design using the first order differential power analysis attack. We collected the measurements of the instantaneous power consumption during the operations that are tested. Finally, we suggested improvements for the design to improve resistance against the attack.**

***Keywords*—Advanced Encryption Standard, Inner Product Masking, Side Channel Attacks, VHDL implementation**

I. Introduction

There is a gap between security of proposed countermeasure schemes and implementations of these schemes in current cryptography. We agreed on realization of hardware implementation of promising scheme called inner product (IP) masking [1], which can defend the implementation against the first order DPA attack.

We chose AES algorithm and Altera Cyclone III DISIPA FPGA board [2] as hardware platform. The software implementation for the 8-bit Atmel was already developed for AVR ATMega128 in assembly language [3]. Hardware implementation was the next logical step.

II. Building the IP Masking Scheme Library

The next step was hardware implementation of the IP masking in VHDL for the chosen platform. As the first approach, we implemented VHDL codes with purely combinatorial logic. Before realization of IP masking operations we needed to implement operations of the underlaying field - Galois field GF(2^8):

- Addition realized with XOR gates (addition in GF(2^8) is bitwise XOR of inputs)
- Multiplication realized using Mastrovito's multiplier
- Squaring realized using net of the XOR gates (the net was computed using properties of squaring and AES irreducible polynomial)
- Inversion realized using subfield GF(2^4)

We lately added powering to 4 and 16 in GF(2^8) for better effectiveness. The realization of the powering are the same as squaring. For every operation we created highly commented and readable VHDL code and testbench, which is used for testing implementation against precomputed results (using Matlab).

Then we implemented all basic operation using IP masking scheme needed for AES algorithm:

- masking, unmasking, mask refreshing
- addition, addition of a constant
- multiplication, multiplication of a constant
- squaring, powering to 4 and 16

These basic operations using IP masking scheme are closely described in [3]. The most interesting operation in terms of side-channel attacks is IP masked version of S-box operation because of its nonlinearity. We needed to implement more complex IP masked operations in VHDL code based on the basic operation:

- Inversion
- Affine transformation
- S-box

Similarly as for GF(2^8), we created highly commented and readable VHDL code and testbenches (which were used for testing implementation) for every IP masked operation.

III. Main results

The first main result was developed and tested hardware implementation of IP masked operations. They were written in VHDL language and for Altera Cyclone III FPGA target. We developed 6 entities for underlying field GF(2^8), 12 entities for IP masking, testbenches for every entity and Matlab scripts for testing. All entities were developed using only combinatorial logic with the most complex developed operation (IP masked S-box) that used 13,202 logic elements. Therefore maximal frequency was below 10 MHz, which could not be deployed in real application. In future it will be necessary to simplify developed complex operations using registers, to adhere the temporal division of the shares of masked values and to take into the account the results of DPA measurements.

Next step was DPA attack on basic IP masked operations. Measured data was evaluated using a t-test trying to find statistically significant difference between fixed input and random input measured trace sets [4] (or fixed versus another fixed input). Before attacking on IP masked operation, we always performed measurement with one masked share fixed to zero. These measurements were done for testing if we could detect first order leakage. Example of the test measurement can be found in Fig. 1 (IP squaring with fixed share to zero).

The first measured and evaluated IP masked operation was squaring due to its simplicity. We ran various measurements of IP squaring with slightly different configurations. Introductory measurement analysis showed first order flaw during the cycle with LED toggling. It was caused by toggling input (fixed 0x00 and random) in a way that LED was toggling, which was evaluated as a leakage. LED switching was disabled and input data were chosen randomly between fixed and random. This upgrade fixed the problem with false detection.

After fixing the false detection, we run the attack on IP squaring, where all the masked shares were evaluated at the same time. First order leakage was not detected (as expected). Then we evaluated the data for the second order attack (using second statistical moment using one leakage point in time). The results showed second order flaw exactly at the point in time when the masked data was handled (Fig. 2).

Next measurement was done on IP squaring implementation with separately evaluated shares. Similarly as for the measurement before the first order DPA attack was not showing any leakages. Then we applied second order attack using first statistical moment in two separate leakage points in time (bivariate 1st order attack in). This attack was not successful with the same amount of the measured traces as for univariate second order attack. This can mean that computing each share separately is more secure than computing them at the same time. In my opinion, much more traces are needed (better processing and evaluating) for the successful bivariate attack.

Last measurements were done for IP refreshing of the mask. In [5] it was theoretically proven that IP refresh scheme contains first order flaw. My task was to prove that this leakage can be measured. First, we tried to specify one share (e.g. fixed to 0x00), which produced measurable first order leakage. Then we tried purely combinatorial implementation with random shares. With 45,000 measurements we were not able to detect any first order leakage (second order univariate was, indeed, detectable because all shares were handled in the same cycle). Then we tried to artificially write the vulnerable intermediate value (according to [5]) to register in order to increase leakage (our experience was that the value leaks information if it is written to the register). Even with registered intermediate value we were not able to detect the theoretical first order flaw. We suggested that much more traces were needed for a successful attack on the vulnerable intermediate value.

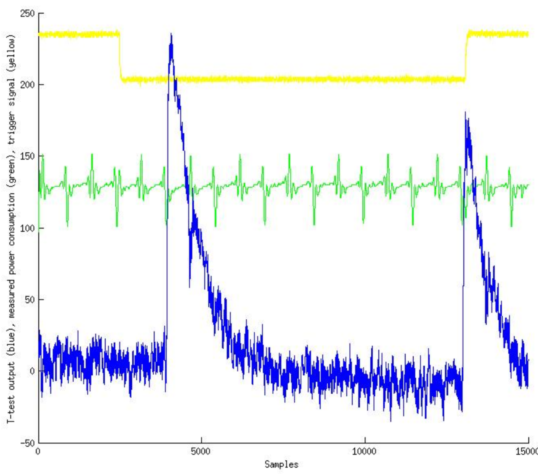


Fig. 1. T-test of IP squaring (fixed 0x00 vs. random) with one input masking share fixed to value 0x00. Squaring of all shares was done simultaneously. Altera Cyclone III FPGA board was used for this measurement. Parameters of the measurement: sample rate 25 GS/s, 500 MHz band, 10,000 measurements. The current flow from decoupling capacitor to FPGA was measured (sensing point 4 on our evaluation board). Left peak is first order flaw from the IP squaring operation with fixed one share, peak on the right is leakage from writing the result to UART.

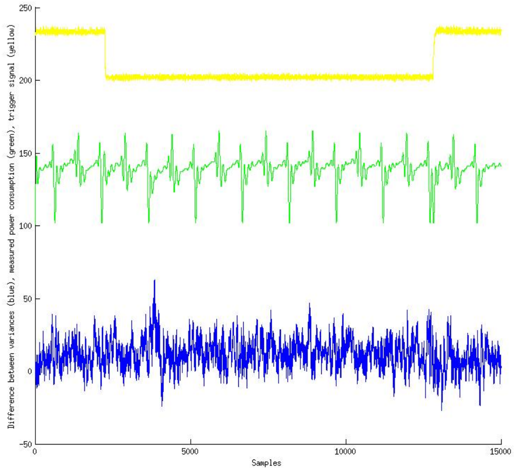


Fig. 2. Analysis of difference between variances (blue graph) of two group of traces during IP squaring (all shares at the same time). First group with input 0, second group was with random input. The figure shows second order leakage univariate (in one point at the time) around sample 4,000 at the exact point where the IP squaring was taking place one clock cycle after trigger (yellow graph) goes to 0. Trigger was down for 7 clock cycles, which can be seen on the green graph of power consumption. Altera Cyclone III FPGA board was used. Parameters of the measurement: sample rate 25 GS/s, 500 MHz band, 45,000 measurements. The current flow from decoupling capacitor to FPGA was measured (sensing point 4 on our evaluation board).

IV. Conclusion

We developed AES library based on IP masking scheme in VHDL. In the future we would like to improve our implementation in terms of hardware area, speed and security in order to make this kind of countermeasure effective.

Acknowledgment

This research was supported by APVV-0586-11 grant and it was performed in the framework of the COST Action IC1204 (Trustworthy Manufacturing and Utilization of Secure Devices).

References

[1] S. Dziembowski and S. Faust, "Leakage-resilient circuits without computational assumptions," in *Theory of Cryptography*, ser. Lecture Notes in Computer Science, R. Cramer, Ed. Springer Berlin Heidelberg, 2012, vol. 7194, pp. 230–247.

[2] M. Varchola et. al., "DISIPA FPGA board: A novel hardware platform for power analysis attacks," Demo Night of International Conference on Reconfigurable Computing and FPGAs (ReConFig), 2013.

[3] J. Balasch, S. Faust, B. Gierlichs, and I. Verbauwhede, "Theory and practice of a leakage resilient masking scheme," in *Advances in Cryptology ASIACRYPT 2012*, ser. Lecture Notes in Computer Science, X. Wang and K. Sako, Eds. Springer Berlin Heidelberg, 2012, vol. 7658, pp. 758–775.

[4] G. Goodwill, B. Jun, J. Jaffe, and P. Rohatgi, "A testing methodology for side channel resistance validation." [Online]. Available: http://www.cryptography.com/public/pdf/a-testing-methodology-for-side-channel-resistance-validation.pdf

[5] E. Prouff, M. Rivain, and T. Roche, "On the practical security of a leakage resilient masking scheme," in *Topics in Cryptology CT-RSA 2014*, ser. Lecture Notes in Computer Science, J. Benaloh, Ed. Springer International Publishing, 2014, vol. 8366, pp. 169–182.

Localisation of Moving and Static Persons in 3D Space

[1]*Peter KAŽIMÍR (3rd year),*
Supervisor: [2]*Dušan KOCUR*

[1,2]Dept. of Electronics and Multimedia Communications, FEI TU of Košice, Slovak Republic

[1]peter.kazimir@tuke.sk, [2]dusan.kocur@tuke.sk

***Abstract*—This paper describes recent advance in research of through-wall person localisation methods based on short-range ultra wideband (UWB) radar measurements. Short range UWB radars are considered to become a valuable tool of living persons localisation in emergency situations such as building fires, natural disasters or acts of terrorism. Through-wall (or through-obstacle) passive person localisation applications developed in the last years, aimed towards use in emergency situations, usually provide a person location estimate in 2D space. However, according to security experts, the person position estimate in 3-dimensional space (3D) is required.**

***Keywords*—UWB radar, localisation, through-wall**

I. Introduction

Presently, we recognize two typical scenario types which represent the need for detection and localisation of persons hidden behind walls, obstacles or otherwise obscured. First group of scenarios are military and security operations in which it is required to detect unauthorised entry to secured areas, or to detect, localise and determine the number of perpetrators of the crime or their victims. Second group of scenarios is aimed towards life rescue in emergency situations such as earthquakes or building fires in which the goal is to estimate the location of motionless survivors to increase the effectiveness of search and rescue operations.

Among existing techniques used in described situations are thermovision cameras, cameras with long optical fibers that are injected into the holes or fissures in the collapsed buildings, acustic systems detecting sounds produced by survivors, search dogs that mark a place where they smell human beings. The use of short-range UWB radar with ability to detect motion was proposed as a complementary or alternative solution to these existing techniques.

The proposed system for human beings detection and localization consists of UWB radar with one transmitting and two or more receiving antennas. Transmitting antenna emits electromagnetic waves to monitored area and receiving antennas capture reflections of electromagnetic waves from objects in this area. Information about presence and position of persons is determined from the measured data by means of signal processing methods. Such system has a potential of increasing the effectiveness of life saving operations, which motivates me to work on the topics related to the improvement of its reliability.

II. UWB radar systems

The acronym UWB refers to a technology which make use of signals occupying a very large bandwidth of frequency spectrum. The Federal Communications Commission (FCC) defines a UWB device as any device where the fractional bandwidth B_f is greater than 0.2 or which occupies the absolute bandwidth B greater than 0.5 GHz. The fractional and absolute bandwidth is defined as follows:

$$B_f = \frac{2(f_u - f_l)}{f_u + f_l}, \quad (1)$$

$$B = f_u - f_l, \quad (2)$$

where f_u and f_l are the upper and lower cut-off frequencies of the emission point $L_{cut-off} = -10$ dBm.

There are several types of UWB radars such as Impulse radar, Impulse Doppler radar, FM-CW (Frequency Modulated Continuous Wave) radar, Stepped-FM radar, Noise and Pseudonoise radar.

UWB radar systems emitting signals occupying a frequency band $DC - 5GHz$ can be used for through-obstacle person localisation, with advantage, because electromagnetic waves with such frequencies propagate through most common building materials with acceptable attenuation, whereas higher frequencies are significantly attenuated. Unfortunately, we have identified a possible unfavourable influence of narrowband interference on the M-sequence sensor performance at the construction of wireless UWB sensor network using FSK modulated signals with the carrier about 868 MHz for sensor network communication infrastructure. This effect has motivated us to study the mentioned problem.

III. Analysis of the interference by narrowband signals

In order to operate with a high efficiency, rescue or law enforcement teams are used to employ at the emergency situations solution usually narrowband wireless personal communication systems operating in the same frequency band as the UWB sensors. In addition, a standard TV and radio broadcasting, and WiFi and mobile communications (e.g. GSM, LTE, etc.) can be active as well. And finally, the particular nodes of the UBW sensor network could communicate by using a narrowband transmission systems with carriers located in the frequency band exploited by the UWB sensors. Hence, the UWB sensor performance under the condition of narrowband interference has to be considered.

In the past, only a few papers and works have been devoted to the analysis of the UWB sensor operation under the narrowband interference. The most comprehensive analysis of that problem have been done in [1] for the impulse UWB radar. In the case of the M-sequence UWB radars, it has been reported that the performance of this kind of UWB devices is quite-well resistant a narrowband interference. This statement has originated in the basic principle of M-sequence UWB radar consisting in the combination of the pulse compression by using M-sequence and sub-sampling (stroboscopic sampling). We further analyzed the problem of narrowband signal interference on the M-sequence sensors in [2]. The partial analysis of the impact of the interference signal power level on the received signal degradation are shown in Fig.1.

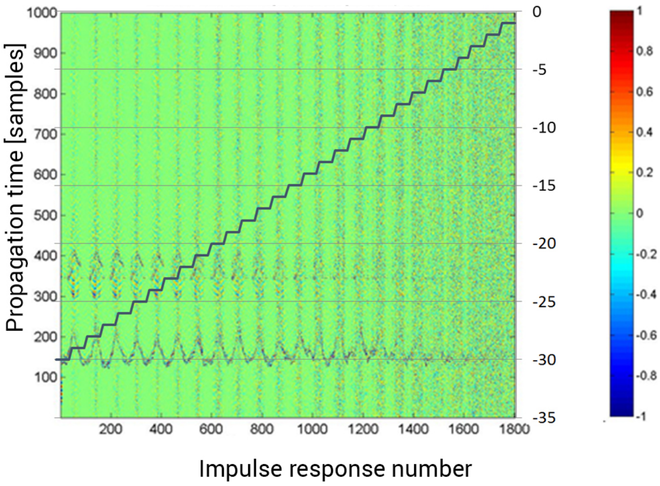


Fig. 1. M-sequence sensor signal degradation by narrowband interference signal scaling from -30dBm to -1dBm

IV. Advances in 3D uncooperative target localisation

In order to obtain data from scenarios similar to real-life situations, we organised a measurement campaign. The campaign took place at Ilmenau University of Technology, Germany, and was mainly focused on localisation of single moving or motionless targets in 3D space. The sensor used in this series set of measurements was an M-sequence sensor with one transmitting and four receiving channels. The campaign was very specific due to the fact, that the data were not measured through the wall only, but through the ceiling and through the floor as well. Among many other scenarios, we simulated the search for the motionless target hidden under the ground level, as well as localisation of moving targets behind the wall or on the upper floor. The data are being processed by our original and novel signal processing methods in an ongoing research on 3D uncooperative target localisation. Partial results are to be published in [3].

V. Conclusion and future work

The UWB sensors have been recently successfully used to locate moving and motionless uncooperative targets. Such sensor application can be improved in terms of performance by using proper antenna array layouts and signal processing methods. Recently, we have responded to the problems of the TOA based multilateration techniques and proved the viability of the novel signal processing methods. Our research in the field of 3D uncooperative target localisation was very positivelly accepted by the peers in our field. We would like to continue the research on the extended antenna array layouts, localisation and tracking signal processing methods and analysis and suppression of the narrowband signal interference.

VI. Extracurricular activities

Alongside with the research I took part in several extracurricular activities such as:

- Commitee member for oral defense of master and bachelor thesis.
- Teaching four courses of *Introduction to programming and networks weekly*.
- Consultant for various student master and bachelor thesis.
- Representation of the university at the ”Noc vyskumnika 2014” event.
- Representation of the university at the ”Den otvorenych dveri” event and other similar events.
- Representation of the university by providing an interview to a local newspaper - *Kosicky Korzar*.

VII. Achievements

Based on the quality of the research in the field of localisation of uncooperative targets in 3D space, I was awarded the following prizes:

- **Young Scientists Contest, Polish Academy of Sciences Prize**
  Awarded at: 15th International Radar Symposium (IRS 2014), Gdansk, Poland
  Awarded for: scientific paper titled ”Simple 3D Localization of Tag-Free Moving Targets by UWB Radar”
- **Best Student Paper Award (2nd place)**
  Awarded at: Progress In Electromagnetics Research Symposium, PIERS 2014), Guangzhou, China
  Awarded for: scientific paper titled ”Localisation of Motionless Persons in 3D Space by UWB Radar”
- **EuRAD Young Engineer Prize**
  Awarded at: 17th European Microwave Week (EuMW 2014), Rome, Italy
  Awarded for: scientific paper titled ”A Simple Approach to Through Wall Localization of Persons Moving in 3-Dimensional Space”
- **Best doctoral work at Technical University of Kosice**
  Awarded at: Best doctoral work contest 2014

Acknowledgment

This work was supported by the Scientific Grant Agency under the contract No. 1/0563/13 and by the Slovak Research and Development Agency under the contract No. APVV-0404-12.

References

[1] A. Nezirovic, I. Damen, and A. Yarovoy, “Narrowband Interference Suppression in UWB Impulse Radar for Human Being Detection,” in *Ultra-Wideband, 2007. ICUWB 2007. IEEE International Conference on*, Sept 2007, pp. 56–61.

[2] D. Kocur, P. Kazimir, and J. Hoffmann, “M-sequence UWB Sensor Signal Degradation By Narrowband Signal,” in *Microwave and Radio Electronics Week 2015*, accepted for publication.

[3] P. Kazimir and D. Kocur, “Localisation of Uncooperative Targets in 3D Space,” in *Acta Electrotechnica et Informatica*, accepted for publication, pp. 203–219.

Magnetic and Electric Properties of Amorphous Ferromagnetic Microwires

[1]*Viktória ŠUHAJOVÁ (4^{th} year)*
Supervisor: [2]Ján ZIMAN

[1,2]Deptartment of Physics, FEI TU of Košice, Slovak Republic

[1]viktoria.suhajova@tuke.sk, [2]jan.ziman@tuke.sk

***Abstract* — The summary of main results of the study of magnetic and electrical properties of two kinds amorphous ferromagnetic microwires during my postgraduate study is presented in this paper. Experimental results concerning single magnetic domain wall dynamics in an inhomogeneous magnetic field in so called bistable microwire are reported. Domain structure contribution to the impedance of amorphous ferromagnetic wire in a zero external field was measured in as cast $Co_{68.2}Fe_{4.3}Si_{12.5}B_{15}$ wire, prepared using the in-rotating-water-quenching technique and in the wire with small uniform helical anisotropy. It was found that the frequency dependence of "domain structure impedance" of these wire have a characteristic shape with a single maximum.**

***Keywords* — ferromagnetic wires, magneto-impedance, domain wall mass.**

I. Introduction

Amorphous ferromagnetic cylindrical wires are novel materials which have been extensively studied in the last two decades. They can be divided into two groups. The first one includes wires prepared by the in-rotating-water quenching technique. Typical radius of wires prepared by this technique is about 100 μm [1]. In the second group there are amorphous glass-coated microwires prepared by rapid quenching and drawing of molten alloy (Taylor-Ulitovsky technique). These microwires consist of metallic nucleus of diameter of 100 nm – 20 μm which is covered by the glass-coating of thickness of 2 – 20 μm [2].

Small size combined with soft magnetic properties (low coercivity) and magneto-transport properties (particularly GMI effect) make these materials very promising for technical applications [3].

II. GMI Effect

Giant magneto-impedance (GMI) effect in thin ferromagnetic wires has been intensively studied for about two decades. The magneto-impedance effect is defined as a relative change in the real and imaginary components of impedance due to present of external magnetic field. In some materials relative changes in impedance can be more than 100%. The main parameter for interpretation of this effect is so-called skin depth δ. This parameter specifies the thickness of the region beneath the surface of the conductor through which the ac current flows [4].

When magnetized in axial direction the wires with high magnetostriction are characterized by the so-called bistable behavior, i.e. rectangular axial hysteresis loop. Assuming that microwire consists of one large axial domain, magnetic reversal takes place by a large Barkhausen jump [2]. Magnetic bistability is a property very interesting for many sensor applications based on a high rate of magnetic reversal [3], [5].

III. Summary of Results

A. *Magnetic domain wall dynamics in an inhomogeneous magnetic field* [6]

An amorphous ferromagnetic $Fe_{77,5}B_{15}Si_{7.5}$ wire, prepared using Taylor-Ulitovsky method, was used for the measurements. Theoretical models for the wall with negligible or fixed length were created. These models were compared with experimental data. Description of experimental set-up and measurement procedure can be found in [6].

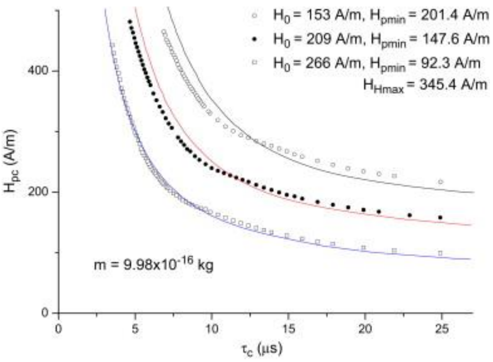


Fig. 1 Comparison of modeled curve (full lines) and experimental ones (point). Curves are modeled for constant mass of the wall

We have found that model curves are in qualitative agreement with experimental results. The best agreement between model and experimental curves was obtained for the case when the wall propagates in a less disturbed magnetic field, i.e. for the lowest value of H_{pcmin} = 92.3 A/m. For higher values of H_{pcmin} the field in which the wall propagates is more disturbed and the difference between model and experimental curve becomes gradually more distinct. It seems that this simple model cannot explain all aspects of the experimental curves. Different parts of the domain walls are influenced by different forces due to the field inhomogeneity. There are two consequences that should probably be taken into account. The first one is the change in the average force acting on the wall for different lengths of the wall. The second one can be the change in wall length due to force inhomogeneity. Therefore the model was modified for a quasi-planar domain walls with a fixed length. We looked for the optimal values of the fitting parameters, the wall length and mass of its unit area.

Agreement between experimental and modeled curves was not much better than that shown in Fig. 1. Moreover big differences in the values of the wall mass per its unit area calculated from the fitting parameters for three experimental curves were obtained.

B. *Effect of domain structure on the impedance of ferromagnetic wire with circumferential anisotropy* [7]

An amorphous ferromagnetic $Co_{68.2}Fe_{4.3}Si_{12.5}B_{15}$ wire, prepared using the in-rotating-water-quenching technique, was used for the measurements. The measurements were carried out on as-cast wire and also on a treated sample. Description of experimental set-ups and treatment procedures can be found in [7].

We have found that the frequency dependence of contribution to the impedance of domain structure for both types of wire have a characteristic shape with a single maximum. This shape probably originates from domain walls oscillations. Additional contribution can originates from the changes in volume magnetization distribution after removal of the domain structure.

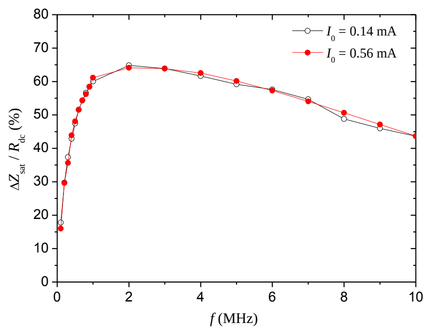


Fig. 2 Frequency dependence of magneto-impedance of domain structure for two different amplitudes of ac current I_0 for as-cast wire

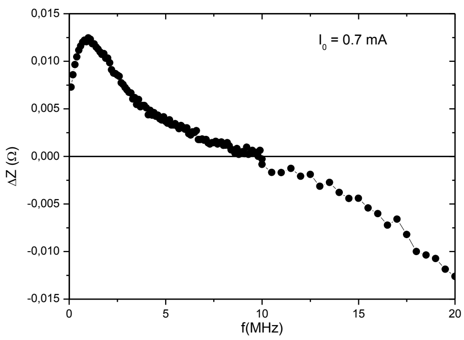


Fig. 3 Frequency dependence of impedance of domain structure for annealed wire with small uniform helical anisotropy

For an as-cast wire after turning off the saturating axial field the domain structure is determined by anisotropy distribution which is of helical type with a small angle by which the easy direction deviates from the circular one. It seems reasonable to assume that the deviation angle has the same sign inside the circular domains, and that the domain walls are located at places where the deviation angle changes its sign. After removal of the domain structure, these places become sources of stray field which results in a reduction of the axial component of magnetization, decrease of transverse permeability and finally decrease in impedance. This mechanism causes positive contribution to magneto-impedance of domain structure. If this mechanism is applied to a treated wire with uniform helical anisotropy, the contribution to the impedance of domain structure is negative. In this case, places where the domain walls are fixed becomes the sources of the stray field which results in the decrease of the axial component of magnetization, decrease of transverse permeability and finally decrease in impedance with respect to the state without domain structure.

C. *Relevant publications and citations*

1. **Impedance and domain wall mass determination in cylindrical wire with circular anisotropy:** J. Ziman, M. Kladivová, V. Šuhajová: sent for publication.
2. **Effect of domain structure on the impedance of ferromagnetic wire with circumferential anisotropy:** J. Ziman, V. Šuhajová, M. Kladivová 2015, In: Sensors and Actuators A: Physical, Vol. 223 (2015), p. 134 – 140.
3. **Single domain wall contribution to the impedance of amorphous ferromagnetic wire:** J. Ziman, V. Šuhajová, M. Kladivová, 2014, In: Acta Polonica A, Vol. 126, no. 1 (2014) p. 82 – 83. ISSN 1898-794X.
4. **Magnetic domain wall dynamics in an inhomogeneous magnetic field:** J. Ziman, V. Šuhajová, M. Kladivová, 2012, In: Physics B: Condensed Matter, Vol. 407, no. 18 (2012), p. 3905 – 3909, ISSN 0921-4526.
   - 2013 [1] TIAN, W., HU, J., PAN, J. et al. In: Review of Scientific Instruments Vol. 84, no. 3 (2013), art. no. 035004 ISSN: 0034-6748
   - 2012 [1] VARGA, R. In: Acta Physica Slovaca Vol. 62, no. 5 (2012), p. 411-518 ISSN: 0323-0465
   - 2014 [1] KOSTYK, J., VARGA, R., VAZQUEZ, M. In: Acta Physica Polonica A Vol. 126, no. 1 (2014), p. 30-31 ISSN: 1898-794X

Acknowledgment

This work was supported by Slovak Research by VEGA grant No. 1/0778/12.

References

[1] M, VÁZQUEZ, "*Soft magnetic wire,*" in Physics B: Condensed Matter, Vol. 299 (2001), pp. 302 – 303.

[2] V.S. LARIN, A.L. TORCUNOV, A. ZHUKOV, J. GONZALES, M. VAZQUEZ,*"Preparation and properties of glass-coated microwires,"* in J. Magn. Magn. Mater, 249 (2002), pp. 39 – 45.

[3] M.VAZQUEZ, A. HERNANDO, "*A soft magnetic wire for sensor applications,*" in J. Phys. D: Appli. Phys. 29 (1996), pp. 939 – 949.

[4] M, VAZQUEZ, "*Giant Magneto-Impedance and Applications in soft magnetic „Wires",* in J. Magn. Magn. Mater., 22ž – 230 (2001), pp. 693 – 699.

[5] A.P. ZHUKOV, M. VAZQUEZ, J. VELÁZQUEZ, H. CHIRIAC, V. LARIN, "*The remagnetization process in thin and ultra – thin Fe – rich amorphous,"* in J. Magn. Magn. Mater., 151 (1995), pp. 39 - 45

[6] J. ZIMAN, V. ŠUHAJOVÁ, M. KLADIVOVÁ, "*Magnetic domain wall dynamics in an inhomogeneous magnetic field":*, In: Physics B: Condensed Matter, Vol. 407, no. 18 (2012), p. 3905 – 3909, ISSN 0921-4526.

[7] J. ZIMAN, V. ŠUHAJOVÁ, M. KADIVOVÁ "*Effect of domain structure on the impedance of ferromagnetic wire with circumferential anisotropy":*, In: Sensors and Actuators A: Physical, Vol. 223 (2015), p. 134 – 140.

Magnetic nanofluids as possible replacement for conventional insulating/cooling fluids

[1]Samuel BUCKO (2nd year)
Supervisor: [2]Roman CIMBALA

[1,2]Dept. of Electric Power Engineering, FEI TU of Košice, Slovak Republic

[1]samuel.bucko@tuke.sk, [2]roman.cimbala@tuke.sk

***Abstract*— This paper deals with my knowledge, experience and results, which I collected during last year on PhD. study. This paper deals with nanofluids, especially magnetic nanofluids which could be replacement for conventional insulating/cooling liquids as mineral oils. Monitoring of condition during thermal ageing of nanofluids might be by measuring as with conventional liquids, for this case the dielectric relaxation spectroscopy was used. The next development in my work for next period is described in last part.**

***Keywords*—nanofluid, magnetic fluid, dielectric spectroscopy, transformer**

I. Introduction and Motivation

Rising of consumption electric power and its quality has strong influence on distribution, transmission and power production devices. Especially, increasing requirements on thermal conduction and insulation have been important for safe power system functioning. Power transformers are part of this power system, where the cooling and insulation have important role. Conventional cooling and insulation system of power transformer consisting of oil-paper system has only one solution for improvement its properties with rising trend of consumption - extended surface of cooling system. The inherently poor thermal conductivity of conventional fluids, as mineral oil puts a fundamental limit on heat transfer. Since 1990's, researchers and engineers developed possible substitution for conventional thermal/insulation liquids composed of nanoparticles dispersed in these conventional liquids - nanofluids. Motivation for works of many researchers is in development of stable suspension without settle and with excellent thermal and the insulation properties. Mineral oil is limited by petroleum reserves and it has non-environmentally effects at leakage. Solution for improvement cooling/insulating properties in transformer is substitution of the used mineral oil by nanofluid based on previous oil, or using new liquids which are environmentally friendly and have electrical, chemical properties similar as mineral oil. Furthermore, these new liquids can be also used at production of nanofluids. [1-3]

Number of power transformers on area of eastern Slovakia is about 60 on distribution level and 13 on transmission level, which all are filled by mineral oil. This aspect is sufficient for local motivation.

II. Present Status of Nanofluids in World and on Our Department

Nanofluid is composed by three parts:

1. **Liquid carrier** - There can be used conventional liquids as water, oils, glycol ethylene. Especially, mineral oil is used for needs cooling and insulation. In some new articles [5, 7] it has been described several investigations with natural and synthetic esters, which are more environmental friendly as mineral oil. On our Dept. of Power Engineering, we used as carrier fluid the mineral transformer oil.
2. **Nanoparticles** - Nanoparticles used in nanofluids have been made of various materials, such as oxide ceramics (Al_2O_3, CuO), nitride ceramics (AlN, SiN), carbide ceramics (SiC, TiC), metals (Cu, Ag, Au), semiconductors (TiO_2, SiC), carbon nanotubes, and composite materials such as alloyed nanoparticles $Al_{70}Cu_{30}$ or nanoparticle core–polymer shell composites.[1] For our cases were used nanoparticles of magnetite with different concentrations.
3. **Surfactant** - Surface active compound, which has been added in order to disperse nanoparticles. Surfactant consists of long chain molecule with polar and non-polar part, which polar part is attracted by nanoparticle and non-polar by fluid. It warrants mutual repulsion of nanoparticles. For magnetite nanoparticles can be used as surfactant acid oleic as well as it was in our case.[8]

Nanofluid consisting of magnetite particles is called the magnetic fluid (MF) with its specific properties described in [8]. In [6] was described prototype of single phase, low power, and medium-voltage electrical transformer. This work is headed for replacement of conventional liquids used in transformer to new MFs with specific properties. One of important tasks in area of liquid dielectrics is monitoring of electrical, chemical a. o. properties and their changes with degradation degree. Spectroscopy of conventional liquids is today on considerably high level, but important questions in case MF are: Can we monitoring degradation of MF with the conventional spectroscopy techniques?, or Shall we develop new spectroscopic techniques for MF and then the replacement of conventional liquids will be more expensive?

III. Solved/Unsolved tasks of last year

Few years ago started some of my colleagues on department with several works, which were based on monitoring dielectric properties and their changes with increasing thermal degradation for MFs.[8] Understanding of polarization processes in dielectrics, measurement methods and knowledge from new area of nanofluids were my task at the beginning of my PhD study. Well-known dielectric spectroscopy in insulating liquids was a good stepping stone for identification changes in NFs. [3, 4]

But question was: Can we measure dielectric properties with equipment for conventional liquids as mineral oil or synthetic ester? Important aspect for spectroscopy of the MFs is measuring of parameter, which could be good reflection for changes inflicted thermal, chemical, electrical ageing. In my case it has been observed by frequency domain relaxation spectroscopy, where were measured and calculated the parts of complex permittivity, that before thermal degradation and after. For better interpretation the linearized Cole-Cole plot and linear regression was used. For better comparison the mineral oil, which was used as carrier liquid in case MF, there was also thermal stressed. From results on Fig. 1 we could observe growth of slope the regression line with increase of time of thermal stress. Similar change in slope was found in both of liquids: MF and mineral oil.

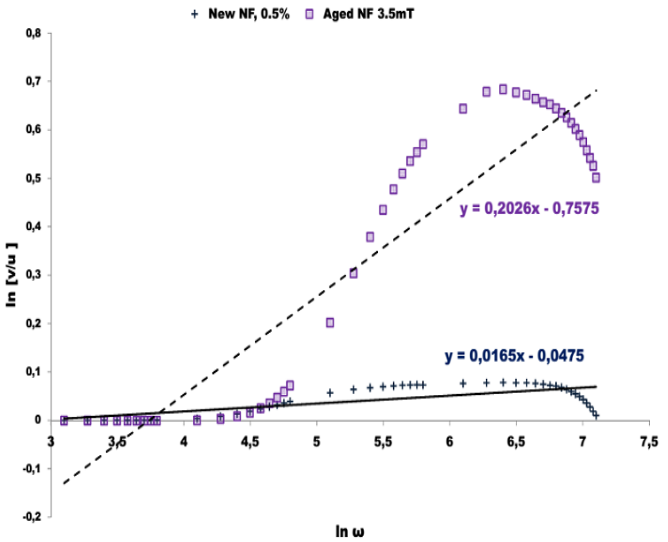


Fig. 1. Linearized Cole-Cole diagram with regression lines for magnetic fluid before and after thermal degradation

Based on these results we can consider that frequency domain relaxation spectroscopy as one of tools for monitoring properties of MF.

The only simple mathematic-physical method for this reflection not exists. Creation of set of many data about properties of MF then can be helpful for monitoring with increasing time of thermal, chemical, electrical degradation. Based on these changes could be predicted the next progress of degradation MFs as well as cellulose insulation.

Furthermore, I cooperated on current contents paper about MFs, on part about the effects of dielectrophoretic and magnetophoretic forces acting on MF in electric and magnetic field. Currently, the paper is waiting on release.[2]

Another part of my work is coordinate graduate students with my supervisor on their diploma works. We work on experiment relating to alternative liquid dielectrics based on natural esters. There were examined different types of liquids like rapeseed oil, flax oil, and thistle oil, and monitored parameter was breakdown voltage. Samples of these oils were thermal stressed and consequently measured by IEC 60156. Results from this work could be used for production MF based on these oils.

Other diploma work deals with time dependence relaxation spectroscopy in MF, where are measured the charge and discharge currents. This experiment is measured in special anechoic shielding chamber with 50Hz network filter for fine measuring without harmonics and interferences.

IV. Next steps

In [2] was thermal stressed only alone MF, but in real transformers is insulation system composed of liquid, paper and pressboard. In my next work, under thermal degradation will be inclusion also cellulose part, for better and realistic imitation conditions in transformers.

I want pick up on idea of my colleague about abrasive qualities of nanoparticles on paper insulation and change in its properties, like breakdown voltage or polarization processes.

Unfortunately, on our department it can't be possible production of magnetic fluids, because of equipments limitation. However, production of nanofluids with different carrier liquids based on synthetic and natural esters or some kinds of nanoparticles could be interesting. This problem could be solved with cooperation with some chemical institute or with Slovak Science Academy. Furthermore, quantity of MF on our department is limited and then we cannot be possible execute breakdown voltage tests, which they are destructive.

Next aim of my dissertation work will improvement of spectroscopy methods and expansion of data set of MF and search further advantages or disadvantages these new composite materials.

V. Acknowledgment

The authors also wish to acknowledge Scientific Grant Agency of The Ministry of Slovak Republic and Slovak Academy of Science for funding of experimental works in the frame of VEGA No. 1/0311/15 grant.

References

[1] Sarit K. Das et al., *Nanofluids: Science and Technology*. New Jersey: John Wiley & Sons, 2008, ch. 1., pp. 1-39

[2] R. Cimbala, J. Király et al., "*Polarizačné procesy v magnetických kvapalinách* ," Chem. listy, accepted for publication.

[3] C. D. Paraskevas et al., "Temperature dependent dielectric spectroscopy in frequency domain of high-voltage transformer oils compared to physicochemical results," IEEE Trans. on Dielectrics and El. Insulation, vol.13, No.3, June 2006, pp 539-546.

[4] Viet-Hung Dang et al., "Comparative study of streamer Phenomena in mineral, synthetic and natural ester oils under lightning impulse voltage," *High Voltage Engineering and Application (ICHVE), 2010 International Conference on* , vol., no., pp.560-563, 11-14 Oct. 2010

[5] M. J. Given, et al., "The influence of magnetite nano particles on the behaviour of insulating oils for pulse power applications," *El. Insulation and Dielectric Ph. (CEIDP),2011 Annual Report Conference on,* pp. 40-43, 16-19 Oct. 2011

[6] L. Pislaru-Danescu, et al., "Prototyping a ferrofluid-cooled transformer, *Trans. on Industry App.,* IEEE, vol.49, no 3, May-June 2013

[7] J. Li, et al., "Investigation of natural ester based liquid dielectrics and nanofluids," *High Voltage Eng. an App.(ICHVE), 2012 Int. Conf.* on pp16-21, 17-20 Sept. 2012

[8] R. Cimbala, J. Kiraly. "Dielectric spectroscopy of transformer oil based ferrofluid from view of i(t) characteristics at initial stage of ageing,"*El. Power Enginering (EPE),* Proccedings of the 2014 15th International Scientific Conference on, pp. 467-472, 12-14 May 2014

Measurement of particle filter accuraccy

[1]*Ján VALISKA (4th year),*
Supervisor: [2]*Stanislav MARCHEVSKÝ*

[1,2]Dept. of Electronics and Multimedia Communications, FEI TU of Košice, Slovak Republic

[1]jan.valiska@tuke.sk, [2]stanislav.marchevsky@tuke.sk

***Abstract*—This paper describes my scientific efforts in particle filters area performed this year. As you can see in almost every scientific paper regarding to object tracking using particle filters, quality or accuracy of this implementations are described by words like: this is good trajectory or this is bad trajectory. I've implemented three types of filters. Every of these filters have quiet good results. Since I like the numbers and also prefer assessment of quality in reasonable metric. This paper describes how to measure this quality by numbers not words.**

***Keywords*—particle filter, histogram, color, accuraccy**

I. Introduction

Start of school year after summer holidays was the same as in early year. But in winter semester I have only one entry in my teaching time agenda. This has forced me to use the time for science in particle filter theory.

One of unsolved problems was that there is no reasonable metric for quality assessment of particle filter tracked trajectories. During the study, I proposed and implemented some types of particle filters. Thees filters are doing good job in object tracking in video sequences. But this quality are defined only by my words. This subjective quality assessment is impaired because of my brain and my eyes are different from other persons. [1]

Problem of objective quality assessment is in that fact, there is need to know the ground-truth object trajectory. This is not reachable in real world, because of relativity. In computer vision are that problem solvable. You can measure requested quality using synthetic video-sequence. This process is described in next chapters.

II. Synthetic video-sequences

Synthetic video is concept of generating video-sequences. Using right software tool is quiet simple to create video-scene that consist of object that are moving from known position to other positions. [2]

Because I have respect to school budget, I decided to use **free software** to solve problem of synthetic video. Free software are tools and programs created by community of peoples that share their work(programming code) to others peoples. Knowing of this source code helps to software debugging or patching. Every software engineer can contribute to this software. [3]

The choice of animating software fell on ***Blender***. Using this big piece of software I have created four types of video sequences. The main role played four-sided object called square. The trajectory of square object described by words have this properties:

- In first passage object moved from start position to next position using three degrees of speed(slow, medium, fast).
- In second, object started to move with constant speed but with rotating 90° clock-wise.
- Next section the object started scaling to 1,5 times dimensions.
- In last section object changes position quickly and rotates in 180° clock-wise.

Trajectory of moving object can be exported directly from Blender software. But it must be performed using *python* script, because lack of this feature in the software. *Python* is popular scripting or interpreted language, and is also frequently used as console language of big software products. Presence of this console language saves me from manually determining of exact position of square object in every frame in video sequence. For a better vision at figure 1 is showed exact characteristics of movement using shapes for x and y directions. [4]

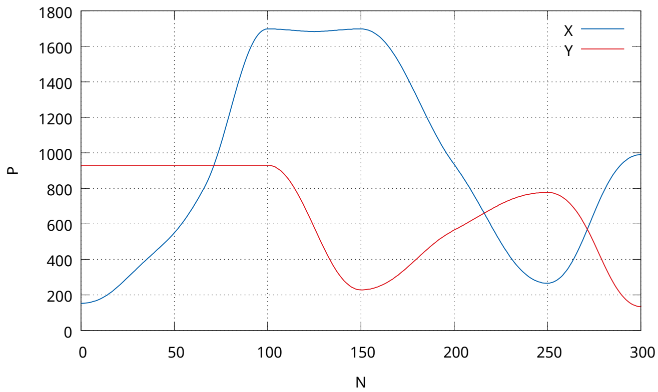


Fig. 1. Exact object movement described by particular axes

As was described above, I created four video sequences for quality assessment of particle filters. In first video sequence moving object have one solid color and background is white. In second, object is divided into four pieces and every part has a different color. Background is also white. Next two video sequences are quiet same as previous but background have multiple color stripes. Small example of object tracking using particle filters on synthetic sequence is at figure 2.

Only to be exact and precise, video-sequence have spatial resolution 1920×1080 pixels. Frame rate is 30 frames per second. Overall number of frames in video sequence is 300, thus duration is 10 seconds. [5]

Fig. 2. Example of synthetic video sequence with multiple color object and multiple color background

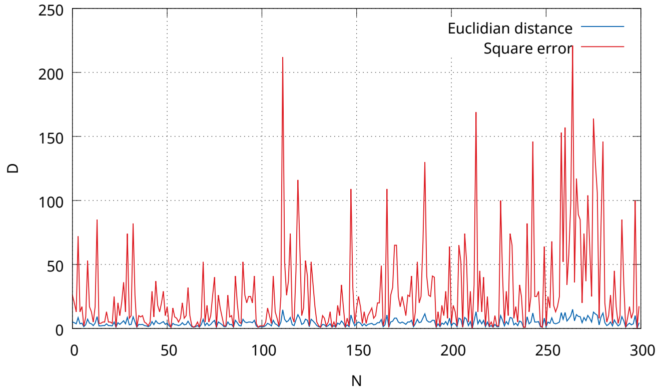


Fig. 3. Visualisation of tracking error using both methods

III. Error metrics

As described above, for measuring of tracking error we need to know original trajectory of moving object. This trajectory can be obtained from synthetic video-sequences using specialised software as described in previous chapter. Next is possible to compare every point in original trajectory with point from tracked trajectory.

I propose two well know methods that are described in next two sub-sections.

A. Euclidean distance

Every point in trajectory can be measured as difference between original and tracked movement as:

$$D_{ED} = \sqrt{(p_x - q_x)^2 + (p_y - q_y)^2} \tag{1}$$

where p is point from original trajectory, q is tracked point. This measure (D) is well known metric named as *Euclidean distance*. Since position point in video sequence is defined using two values for both axes x and y, error must be calculated for every component of this point. [6]

This distance must be expressed for every point in trajectory. For example generated video sequences have 300 frames, thus original trajectory have also 300 position points. [7]

Mean tracking error of entire video sequence can be expressed as mean of euclidean distances, as:

$$MED = \frac{1}{N}\sum_{i=1}^{N} D_{ED_i} \tag{2}$$

where D is function of distance between $i - th$ original and tracked point and N is number of frames. [8]

B. Mean Square Error

On the other side there is also different metric that can be used for trajectory assessment. This method is known as *square error* [9]. Can be calculated as:

$$D_{SE} = (p_x - q_x)^2 + (p_y - q_y)^2 \tag{3}$$

Mean square error for whole trajectory is then computed as:

$$MSE = \frac{1}{N}\sum_{i=1}^{N} D_{SE_i} \tag{4}$$

Results of comparing original and measured trajectories are in plot 3. This results are expressed by both proposed methods.

IV. Conclusion

As described in previous two sections measurement of particle filter accuracy is not straightforward task. In real conditions it is very hard task, because original trajectory must be precisely defined and this must be done using some others methods, like GPS tracking, or others precise position tracking methods. Thus, as you can see, experimental verification of quality of filter, can be done relatively easily but it needs prepared resources as synthetic video sequences and synthetic object trajectory.

Because this is probably last year of my study at Technical university of Košice, I am currently summarizing my whole scientific research work in my dissertation thesis. I hope that this work will be a good basis for people who are investigating the "jungle" of particle filter theory.

Acknowledgment

This work was supported by the Scientific Grant Agency of the Ministry of Education, Science, Research and Sport of the Slovak Republic under the contract No. 1/0563/13 and by the Slovak Research and Development Agency under the contract No. APVV-0404-12.

References

[1] D. Duckworth, "Monte carlo methods for multiple target tracking and parameter estimation," Master's thesis, EECS Department, University of California, Berkeley, May 2012. [Online]. Available: http://www.eecs.berkeley.edu/Pubs/TechRpts/2012/EECS-2012-68.html

[2] A. Doucet and A. M. Johansen, "A tutorial on particle filtering and smoothing: fifteen years later," 2011.

[3] G. Sullivan, J. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (hevc) standard," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 22, no. 12, pp. 1649–1668, Dec 2012.

[4] F. Rezac, M. Voznak, P. Partila, and K. Tomala, "Interactive video audio system and its performance evaluation," in *36th International Conference on Telecommunication and Signal Processing*, Roma, Italy, 2013, pp. 43–46.

[5] S. Marchevský, D. Levický, Ľ. Maceková, S. Pillár, B. Hrušovský, and Ľ. Holováč, *Progresívne technológie v DVB-T*. TUKE, 2011, ch. IPTV: ZÁKLADNÉ TECHNOLOGICKÉ RIEŠENIE, pp. 150–199.

[6] B. Hrušovský, J. Mochnáč, and S. Marchevský, "Error concealment method based on motion vector prediction using particle filters," *Radioengineering*, vol. 20, no. 3, pp. 629–702, 2011.

[7] T. Kodama, T. Yamaguchi, and H. Harada, "A method of object tracking based on particle filter and optical flow to avoid degeneration problem," in *SICE Annual Conference 2010, Proceedings of*, 2010, pp. 1529–1533.

[8] J. Valiska, S. Marchevský, and R. Kokoška, "Object tracking by color-based particle filter techniques in videosequences," in *Radioelektronika (RADIOELEKTRONIKA), 2014 24th International Conference*, April 2014, pp. 1–4.

[9] L. Maceková and S. Marchevský, "A new image and video quality criterion," *Acta Electronica et Informatica*, vol. 4, no. 2, pp. 15–19, 2004.

Measuring of shielding effectiveness of electromagnetic field

[1]Peter KURIMSKÝ (3rd year)
Supervisor: [2]Alexander MÉSZÁROS

[1,2] Dept. of Electric Power Engineering, FEI TU of Košice, Slovak Republic

[1]peter.kurimsky.2@student.tuke.sk, [2]alexander.meszaros@tuke.sk

***Abstract*— This paper deals with my work in previous year. The measurements were focused on measuring of shielding effectiveness in the range of high frequencies from 1 GHz to 9 GHz in the step 0,2 GHz. The measured object was a polystyrene and brick wall with a thickness 25 cm. Size of brick wall were 2x2 m. Measurement was preformed according IEEE Standard, Method for Measuring the Effectiveness of Electromagnetic Shielding Enclosures.**

***Keywords*—shielding of electromagnetic field, electromagnetic field, antenna.**

I. INTRODUCTION

Wide professional community is more and more focused on resources of electromagnetic fields, their impact and the associated concept of electromagnetic compatibility. EMC could be defined as the ability of devices to coexist in the same electromagnetic environment. Electromagnetic interference can cause severe problems, and it should be taken into account in the design of new power plants and other devices. The term Electromagnetic Compatibility is closely related to the notion of shielding electromagnetic field.

Everyone is exposed to a complex mix of electric and magnetic fields, both at home and at work, from the generation and transmission of electricity, domestic appliances and industrial equipment, to telecommunications and broadcasting.

Research activities on my PhD study are focus on the measuring of shielding effectiveness and reflection of electromagnetic fields. Measured objects were building materials – polystyrene and brick.

II. THE INITIAL STATUS IN THE SOLVING OF THE RESEARCH TASK

Quality of shielding materials is determined by three coefficients, shielding effectiveness *SE*, absorption coefficient *A* and a reflection coefficient *R*. The main factors which determine the shielding effectiveness are the capability of shielding materials (the electric and magnetic conductivity and the permeability), the thickness of shielding material and the frequency of the incident wave. If we know all these factors, the shielding effectiveness of materials can be calculated by (1):[1][2]

$$SE = A + R \tag{1}$$

Shielding effectiveness can be calculated according to the relations (2–5) if the value of the transmitted signal is set in logarithmic unit.

$$SE = \left|E_{1\log}\right| - \left|E_{2\log}\right| \text{[dB]} \tag{2}$$

$$SE = \left|H_{1\log}\right| - \left|H_{2\log}\right| \text{[dB]} \tag{3}$$

$$SE = \left|V_{1\log}\right| - \left|V_{2\log}\right| \text{[dB]} \tag{4}$$

$$SE = P_{1\log} - P_{2\log} \text{[dB]} \tag{5}$$

where E_1 and H_1 are the intensity of electric field and magnetic field in the absence of the shielding, respectively E_2 and H_2 are the intensity of electric field and magnetic field within the shielding, V_2 is voltage reading within the shielding, V_1 is voltage reading in the absence of the shielding and P_2 is power detected within the shielding, P_1 is the power detected in absence of the shielding.

Formulas (8), (9), (10) and (11) are used according to the available measuring equipment.[2][3][4]

III. TASK SOLVED IN THE PREVIOUS YEAR

My research was focused on the building materials, polystyrene, brick wall. Block diagram for the purpose of measuring of shielding effectiveness *SE* of the electromagnetic field is shown in Fig.1. This workplace consists of a analog signal generator Agilent N5181A, spectrum analyzer Agilent N9038A MXE EMI, the receiving antenna and transmitting antenna horn type.

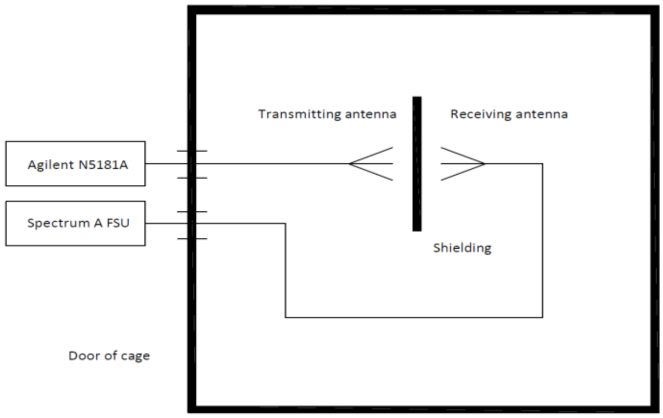


Fig. 1. Block diagram for purpose of measuring of shielding effectiveness of electromagnetic field

View of the measuring in the chamber is shown in Fig. 2. The measuring results are shown in Fig.3 and Fig.4. Dependence of shielding effectiveness *SE* of electromagnetic field in frequency range 1 GHz to 5 GHz for bricks wall shown Fig.2. Dependence of shielding effectiveness of electromagnetic field in frequency range 1 GHz to 9 GHz for bricks wall shown Fig.3. Dependence of shielding effectiveness of electromagnetic field in the frequency range from 1 GHz to 9 GHz for polystyrene with thickness 50 mm show Fig.4

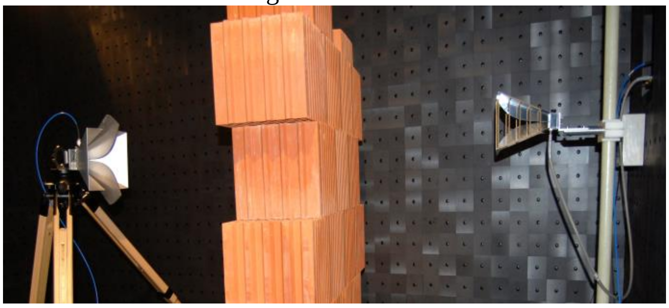


Fig. 2. View of the measuring in the chamber

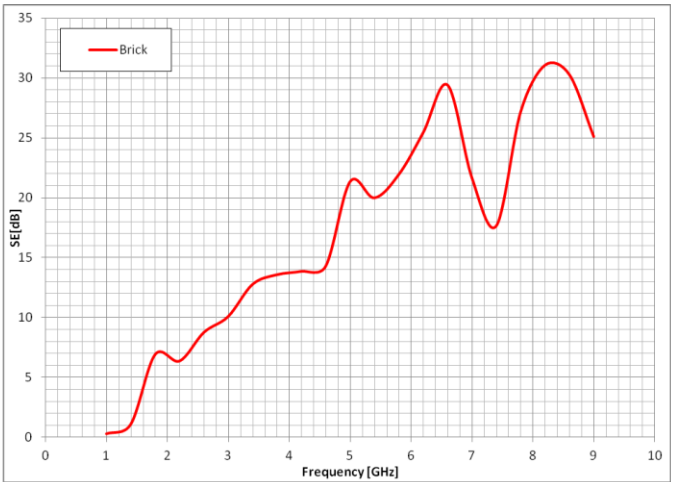


Fig. 3. Dependence of shielding effectiveness SE of electromagnetic field in frequency range 1 GHz to 9 GHz for bricks wall

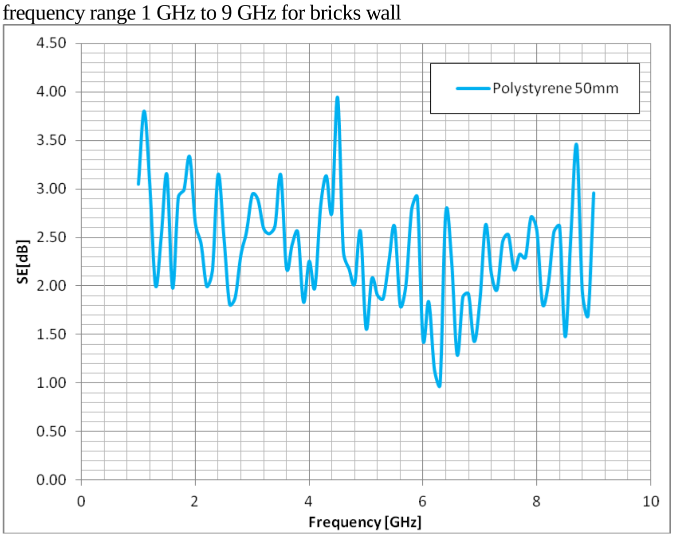


Fig. 4. Dependence of shielding effectiveness SE of electromagnetic field in frequency range 1 GHz to 9 GHz for polystyrene with thickness 50mm

More results can be found in [6][7][7][8] and other.

IV. Proposal for the next steps in the next period of time

In next period I would like to continue with measuring of reflection electromagnetic field of building materials not measured in the quoted publications. From the measured data, it is possible to calculate absorption of the electromagnetic field.

Acknowledgment

This publication is the result of the Project implementation: Protection of the population of the Slovak Republic from the effects of electromagnetic fields, ITMS: 26220220145, Activity 2.1 supported by the Research & Development Operational Programme funded by the ERDF.

Európska únia
Európsky fond regionálneho rozvoja

Operačný program VÝSKUM a VÝVOJ

Agentúra
Ministerstva školstva, vedy, výskumu a športu SR
pre štrukturálne fondy EÚ

We support research activities in Slovakia / Project is cofinanced from EU funds.

References

[1] S. Bhattacharjee, "Protective Measures to Minimize the Electromagnetic Radiation" Advance in Electronic and Electric Engineering, vol. 4, Number 4, pp. 375-380, 2010, ISSN 2231-1297.

[2] T.W. Wieckowski, J.M. Janukiewicz, "Methods for Evaluating the Shielding Effectiveness of Textiles", Fibres & Textiles in Eastern Europe, Vol. 14, No. 5 (59), January / December 2006.

[3] L. Zhang, Y. Zhang, " Simulation Analysis for the Materials Shielding Effectiveness of EMP", Cross Strait Quad-Regional Radio Science and Wireless Technology Conference, July 26-30, pp. 32-35, 2011.

[4] IEEE Standart, Method for Measuring the Effectiveness of Electromagnetic Shielding Enclosures, EMC Society, New York 2006, p.39.

[5] A. Mészáros [et al.], *Meranie intenzity elektromagnetického poľa a určenie účinnosti tienenia vybraných materiálov pri frekvenčnom rozsahu od 500 MHz – 5 GHz*, In: Elektroenergetika. Roč. 7, č. 2 (2014), s. 19-22. - ISSN 1337-6756.

[6] M. Pavlík [et al.], *Meranie účinnosti tienenia a odrazu elektromagnetického poľa pre stavebný objekt polystyrén*, In: Electrical Engineering and Informatics 5 : Proceedings of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice. - Košice : TU, 2014 S. 496-501. - ISBN 978-80-553-1704-5.

[7] M. Pavlík [et al.], *Measuring of shielding effectiveness of electromagnetic field of building material* In: Current Problems of Maintenance of Electrical Equipment and Management. - Košice : TU, 2014 S. 186-190. - ISBN 978-80-553-1818-9.

[8] M. Pavlík, *Meracia metóda pre určenie účinnosti tienenia elektromagnetického poľa*. In: Teória a aplikácia metód technickej diagnostiky: XVII. Ročník medzinárodnej vedeckej konferencie: 28.-29. Október, 2014, Košice, Slovakia. - Košice : TU, 2014 S. 3-6. - ISBN 978-80-553-1858-5.

Overview of Degradation Mechanism in All-Optical WDM Systems

[1]*Tomáš IVANIGA (1st year)*
Supervisor: [2]Ján TURÁN

[1,2]Dept. of Electronics and Multimedia Communications, FEI TU of Košice, Slovak Republic

[1]tomas.ivaniga@tuke.sk, [2]jan.turan@tuke.sk

***Abstract*— This publication deals with Dense Wavelength Division Multiplexing (DWDM) systems and the physical phenomena that affects a high-speed optical signal transmission. The non-linear optical effects which have been intensively researched in recent years when designing a fully optical communication system must be taken into account. DWDM systems have high demands on transmission capacity and speed. Limiting factors are FWM (Four-Wave Mixing) and CPM (Cross-Phase Modulation) which may result in poor draft crosstalk between channels. For this reason, it is important to denote each non-linear phenomena individually in this optical communication systems.**

***Keywords*— CPM, FWM, SPM, standard DWDM,**

I. Introduction

Wavelength division multiplexing (WDM) networks are optical networks which use a principle of wavelength division multiplexing allowing a single fibre to transmit several optical carrier waves, each being at a different wavelength [1]. It is a frequency multiplex in the optical transmission band of the transmission medium as the wavelength corresponds to the frequency of the optical signal carrier. This transfer technology allows deploying optical carrier waves carrying the data to the whole band of the medium permeability, taking into account the available transmission technologies, meaning one fibre is binding more optical wavelength ranges transmitting the data. The transmission at each wavelength can be achieved with a different transmission (bit) rate, another type of modulation and other forms of signal. The idea of WDM networks allows better use of built optical networks by transporting a single fibre with a number of carrier wavelengths, yet the structure of optical networks has not changed [2].

The disadvantage of WDM systems is the increased damping of the network due to the multiplexer and demultiplexer inserted into the transmission chain. In the mid-nineties the term Dense WDM started to be used after managing to transfer more and more channels with significantly smaller spacing in one fibre.

II. Dense Wavelength Division Multiplexing

Dense Wavelength Division Multiplexing (DWDM) is one of the most perfect and most widely used systems in today's optical communications. Spacing between channels are 0.8 nm, theoretically up to 0.1 nm in UDWDM (Ultra Dense Wavelength Division Multiplexing), means the ability to transmit a single fibre to several dozen channels [3]. These channels are transmitted by optical fibres in parallel and without being dependent on each other, the transmission capacity of optical connection increases many times. Modern DWDM systems can on a single physical link operate 96 channels, in each such channel transmit enable a signal rate of 2.5 to 10 Gbit.s^{-1}. In Fig. 1 is displayed a 4-channel DWDM system.

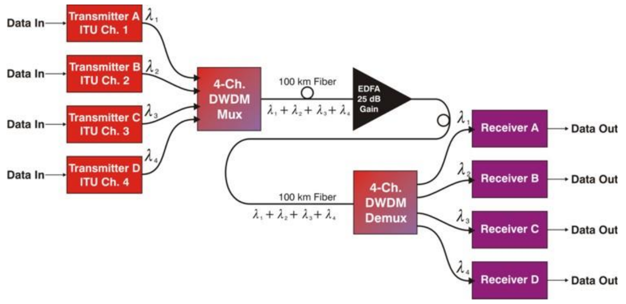


Fig. 1 4-channel DWDM [3]

For proper implementation of this technology is used laser DFB (Distributed FeedBack laser) with an extremely narrow spectral line, amplifier EDFA (Erbium Doped Fibre Amplifier) and spectral filters with high selectivity. These devices are very sensitive to temperature and frequency stability. This is one of the main reasons of its high price. Regulation ITU-T G.694.1 "Spectral grids for WDM applications: DWDM frequency grid" defines the various transmission channels for DWDM wavelengths in the range of 1,490 nm (200.95 THz) to 1620 nm (186.00 THz), S, C and L band [3]. Channel spacing DWDM is based on the normalized initial frequency of 193.1 THz. The raster with the spacing of the individual channels in the range of 100 GHz (0.8 nm), 50 GHz (0.4 nm), 25 GHz (0.2 nm) (UDWDM) and 12.5 GHz (0.1 nm) depends on the frequency. In Fig. 2 is shown channel spacing for DWDM system. The distance between wavelengths and frequencies is given by the equation

$$\Delta\lambda = \frac{\lambda.\Delta f}{f}, \tag{1}$$

where $\Delta\lambda$ represents the wavelength spacing, Δf is the frequency spacing value, λ represents each wavelength and f represents frequency. For a high quality of transmission and the correct operation it is necessary that the wavelength deviation from the designed wavelength will not exceed 0.2, representing tolerance ± 0.16 nm for 100 GHz channel spacing.

For use DWDM technology in access and metropolitan networks should be appropriate to narrow wavelength range to the band "C" only, which represents the area from 1530 nm to 1565 nm [4-6]. For this band, the components (EDFA amplifiers, etc.) are suitable and affordable.

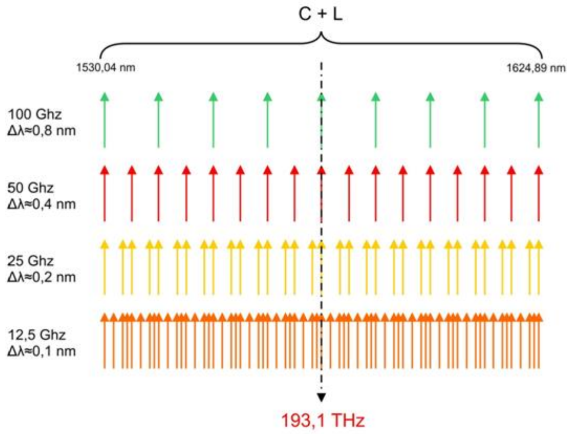


Fig. 2 Channel spacing DWDM system

III. Further Research Field

Non-linear effects in optical fibres arise from the change in the refractive index of the medium with the optical intensity and inelastic scattering [6]. The dependence of the performance and the index of refraction is responsible for Kerr effect [7]. As shown in Fig.3, is displayed linear and non-linear dependence.

Depending on the type of input signal, Kerr effect occurs in three different phenomena, self-phase modulation (SPM), cross-phase modulation (CPM) and four-wave mixing (FWM). At high power levels, the inelastic scattering phenomena causes stimulated effects such as stimulated Brillouin scattering (SBS) and stimulated Raman scattering (SRS). Scattered light intensity increases exponentially when the incident power exceeds a certain threshold. The difference between the Brillouin and Raman scattering is that SBS generates phonons (acoustic) that are coherent and cause macroscopic acoustic waves in the fibre, while SRS generates phonons (optical) that are inconsistent and do not generate macroscopic wave [7].

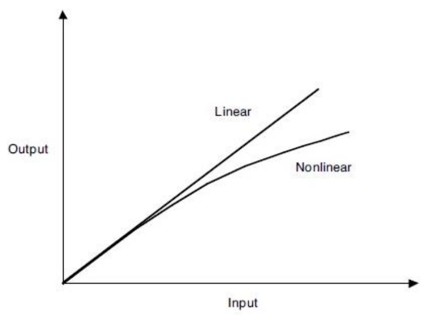


Fig. 3 Linear and non-linear dependence

A. Self-Phase Modulation

Self phase modulation is a nonlinear phenomenon caused by the interaction of light and material. If the light pulse propagates in a nonlinear optical medium due to the optical Kerr phenomenon induced a change in refractive index. This causes a dependence of the pulse phase on the intensity, and this leads to a change in the pulse spectrum [8]. The Fig.4 shows the effect of SPM on a pulse signal in which the frequency of the pulse at the onset decreases and increases during deceleration. The frequency is in the middle of the pulse approximately linear [6-8].

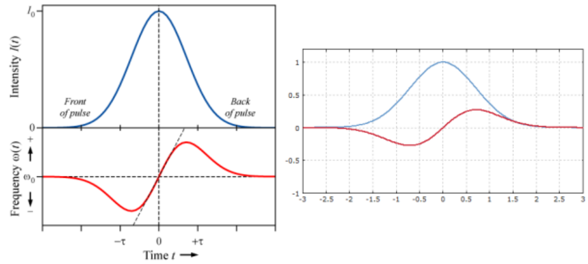


Fig. 4 The impact of self-phase modulation on the signal pulse [9]

Phase changes with time in the same way as the optical signal. Different parts of the pulse go through various phase shifts due to fluctuations in phase. This results in the frequency chirping. The primary phenomenon of self-phase modulation is broadening the pulse. SPM effect is more pronounced in systems with high transmission power as chirping effect is proportional to the signal transmission power. Phase is expressed as follows

$$\phi = \frac{2\pi}{\lambda} nL, \tag{2}$$

where λ is the wavelength of the optical fibre of the transmitted pulse having a refractive index n, and nL is called optical length [9].

The fibre used for high transmission power can be replaced with n_{eff} and L_{eff} respectively i.e.

$$\phi = \frac{2\pi}{\lambda} n_{eff} L_{eff}, \tag{3}$$

or

$$\phi = \frac{2\pi}{\lambda} (n_l + n_{nl} I) L_{eff}, \tag{4}$$

where the right part of the first relationship is a linear portion of a constant phase (ϕ_l) and the second relationship is a nonlinear portion of the constant phase (ϕ_{nl}).

If the intensity is time dependent and the wave is modulated in time, then the phase φ is also time dependent. Substitution phase in time causes a change in the frequency spectrum which is given by

$$\omega = \frac{d\phi}{dt}. \tag{5}$$

In dispersion, the medium in the spectrum changes temporarily and the pulse changes the nature of variations. It is necessary to take into account the Gaussian pulse, which modulates the carrier frequency ω and the new instantaneous frequency becomes

$$\omega' = \omega_0 + \frac{d\phi}{dt}. \tag{6}$$

The sign of the phase shift is SPM negative due to a minus sign in the expression for the phase $(\omega t - kz)$

$$\phi = -\frac{2\pi}{\lambda} L_{eff} (n_l + n_{nl} I), \tag{7}$$

and, therefore ω is

$$\omega' = \omega_0 - \frac{2\pi}{\lambda} L_{eff} n_{nl} \frac{dI}{dt}. \tag{8}$$

Thus the leading edge pulse $\frac{dI}{dt} > 0$ is

$$\omega' = \omega_0 - \omega(t), \tag{9}$$

and the rear edge $\frac{dI}{dt} < 0$ so,

$$\omega' = \omega_0 + \omega(t), \quad (10)$$

where

$$\omega(t) = \frac{2\pi}{\lambda} L_{eff} n_{nl} \frac{dI}{dt}. \quad (11)$$

This shows that the pulse is chirp and the frequency is varied across the pulse. This phenomenon called chirping is caused by SPM leading to an extent pulse spectrum [10].

B. *Four-Wave Mixing*

When tied to a light beam of high intensity, the response of the optical environment ceases to be linear and together with Kerr phenomenon is further define as four-wave mixing [11]. General FWM occurs in binding the three light beams of different wavelength. Optical environment for the fourth wave different from the three other at the optical fibre (referred to as "idler").

The FWM is shown in Fig.5, where the right part of the image shows the effect of FWM [11]. FWM is a type of optical parametric oscillation. Condition of occurring FWM is called "Phase-mismatching" factor

$$\Delta k = k_1 + k_3 - k_2 - k_4, \quad (12)$$

where k_1, k_2, k_3 a k_4 are linearly dependent vectors and the wave frequency is equal to

$$f_{ijk} = f_i + f_j - f_k. \quad (13)$$

This nonlinear phenomenon occurs only in WDM systems where its presence causes complications of transmission. Evolving frequency signals/channels degrade the transmission quality when they interference with the useful signals while increasing the absorption of useful channels whereas transmit power in favor of the new channels [9]. Option filtration newly created channels is possible only for those who are not in the original wavelength. Prevent FWM can be with an appropriate distribution channels, so that the new created wavelengths do not occur on the useful channels and thus do not reduce the transmission quality [12].

Effect of FWM is inversely dependent on chromatic dispersion coefficient, FWM is most apparent at wavelengths zero dispersion.

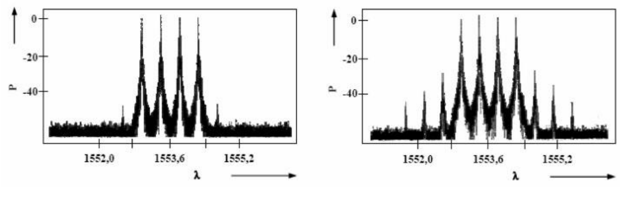


Fig. 5 The emergence of new channels in four channel DWM system. In Figure on the right we can see spectral lines of new channels. [12]

C. *Cross-Phase Modulation*

Cross-phase modulation is a major limitation for non-linear single-channel systems. The intensity dependence of the refractive index leads to a further non-linear process [13]. If two or more optical pulse propagates at the same time, there is a cross-phase modulation accompanied by the self phase modulation. It arises due to the nonlinear refractive index of the optical pulse which depends not only on the intensity of the beam but also on the intensity of other propagating pulse [15-17].

Cross-phase modulation actually performs fluctuation power at a certain wavelength to phase fluctuations in promoting other channels. The result of CPM may be asymmetrical enlargement of the spectral lines and pulse shape distortion [14]. The effective refractive index of the nonlinear medium can be expressed by the input power P and the effective core area A_{eff} as,

$$n_{eff} = n_l + n_{nl} \frac{P}{A_{eff}}. \quad (14)$$

The nonlinear effects are dependent on the light output ratio of the cross-sectional area of the fibre

$$k_{eff} = k_l + k_{nl} P, \quad (15)$$

where *kl* is the linear part of the propagation constant and knl the nonlinear propagation constant. The phase shift caused by nonlinear constant after covering a distance *L* inside the fibres is as follows

$$\phi_{nl} = \int_0^L \left(k_{eff} - k_l\right) dz, \quad (16)$$

consequently, the nonlinear phase shift is expressed by relationship

$$\phi_{nl} = k_{nl} P_{in} L_{eff}. \quad (17)$$

If the optical pulses spread at the same time, the nonlinear phase shift of the first channel depends not only on the strength of the channel but also the signal strength of the other channels. For two channels is a nonlinear phase shift as follows

$$\phi_{nl}^1 = k_{eff} L_{eff} (P_1 + 2P_2). \quad (18)$$

For N-channel transmission system is a shift for the I-th channel expressed by relationship

$$\phi_{nl}^i = k_{nl} L_{eff} \left(P_i + 2\sum_{n \neq i}^N P_n\right). \quad (19)$$

The element 2 in the above equation is the original form of the non-linear susceptibility, indicating that the CPM is twice more effective than the SPM at the same energy. The first part of the equation represents the contribution of SPM and second part the contribution of CPM. CPM is effective only if the influencing signals overlap in time.

CPM limits the performance of the system in the same way as SPM, the frequency "chirping" and chromatic dispersion, although CPM may interfere with the performance of more than SPM. CPM greatly affects the particular system with a large number of channels.

D. *Stimulated Raman Scattering*

Stimulated Raman scattering is among the most important effects, whether in terms of nonlinear phenomena in optical fibres or in adversely affecting the transmission communication systems, although its base is relatively simple. When this phenomenon occurs during the propagation of photons to the molecular vibrations, leads to loss of energy and generation of optical phonons. The generated waves with small wavelengths are called Stokes waves [16].

In stimulated Raman scattering occurs pumping energy from lower wavelengths at higher wavelengths, causes at higher wavelengths suppress signals from lower wavelengths. This phenomenon occurs especially at high bit rates and at high transmission performance. The only way to reduce the impact of Raman scattering is to reduce the input power. In the Fig.6 is shown stimulated Raman scattering [17].

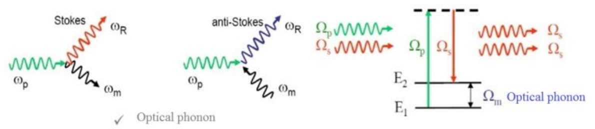


Fig. 6 Stimulated Raman scattering [17]

E. *Stimulated Brillouin Scattering*

The detrimental effect on communication systems called stimulated Brillouin scattering observed in 1964. Its base is similar to the SRS, except that for the SBS, in the dissemination of photons are generated in molecular vibration acoustic phonons [18]. The stimulated Brillouin scattering occurs only in the backward direction of propagation along the fibre, unlike the SRS, which occurs in a straight line. Like the SRS, the SBS pronounced at high transmission speeds and at high transmission performance [19]. The Fig.7 shows stimulated Brillouin scattering.

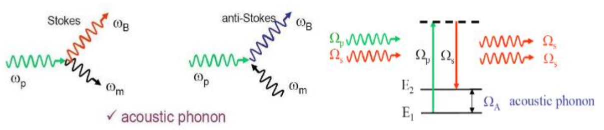


Fig. 7 Stimulated Brillouin scattering [19]

IV. CONCLUSION

This publication consists of basic information about DWDM system and its composition. Research on the non-linear effects applied in WDM systems is done on a very broad scale. My task is to address the non-linear phenomena in DWDM systems and design a system for eliminating the effect of non-linear phenomena in optical amplifier (EDFA). This is an amplifier design in Matlab and OptSim and the subsequent realization of fully optical communication system.

Fig.8 represents the scheme of a fully optical communication system created in Technical University of Košice optoelectronic system laboratories, incorporating the amplifier part which will be my task to research in the coming years.

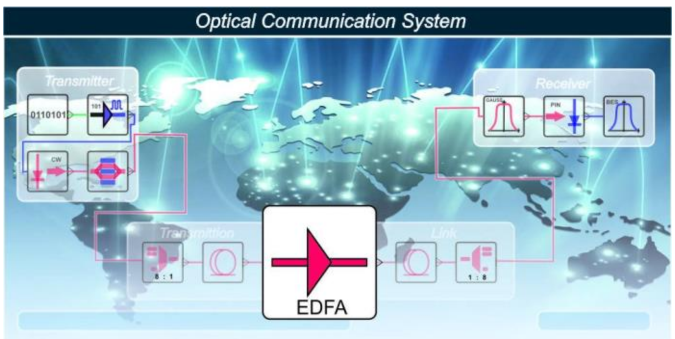


Fig. 8 Optical communication system

ACKNOWLEDGMENT

This work was supported by Cultural and Educational Grant Agency (KEGA) of the Ministry of Education, Science, Research and Sport of the Slovak Republic under the project no. 059TUKE-4/2015and project no. 006TUKE-4/2014. This work is also the result of the contract no. “APVV-0025-12 - Mitigation of Stochastic Effects in High-Bitrate All Optical Networks” supported by the Slovak Research and Development Agency.

REFERENCES

[1] A. Udalcovs, V. Bobrovs, J. Porins , "*Evaluation of SPM-induced optical signal distortions in ultra-dense mixed-WDM system*," Future Generation Communication Technology (FGCT), 2012 International Conference on , vol. 180,no. 184, Dec. 2012, pp. 12-14.

[2] A. Supe, J. Porinš , “*Methods for Estimation of Optical Fiber Non-Linearity Using Self-Phase Modulation Effect*” , Latvian Journal of Physics and Technical Sciences, 2011, Vol.48, No.6, ISSN 0868-8257, pp. 29-40.

[3] P. Lafata, J. Vodrážka, “*Optické Přistupové sítě a přípojky FTTx*”, CTU in Praque – publishing CVUT, Praque 2014, ISBN 978-80-01-05463-5, 294 pp.

[4] Srivastava, M.; Kapoor, V., "*Analysis and compensation of self phase modulation in wavelength division multiplexing system*," Engineering and Systems (SCES), 2014 Students Conference on , vol.1, no.4, pp. 28-30, May 2014.

[5] N. Shibata, R. Braun, R. Warts, "*Phase-mismatch dependence of efficiency of wave generation through four-wave mixing in a single-mode optical fiber*," Quantum Electronics, IEEE Journal of , vol.23, no.7, Jul 1987, pp.1205-1210.

[6] K. Kashyap, H.Singh, P.Singh, Ch.Gupta , ”*Effect of Self phase Modulation on Optical Fiber*” , American International Journal of Research in Science, Technology, Engineering & Mathematics, ISSN (Print): 2328-3491, ISSN (Online): 2328-3580, ISSN,(CD-ROM): 2328-3629, March - May, 2013, pp. 160-164.

[7] G.P. Agrawal, R.-J. Essiambre, ”*Nonlinear limits of SDM transmission*”, (2014) Proceedings - 2014 Summer Topicals Meeting Series, SUM 2014, art. no. 6903044, pp. 174-175.

[8] M. Bhutani, A. Gagneja, ”*Optical Transmission System Simulation for Analysis of Self Phase Modulation Non Linearity*” , ISSN 2229-5518, [online],[cit. 09.03.2015], International Journal of Scientific & Engineering Research, vol. 4, Issue 6, June-2013 1707, Available on the internet: http://www.ijser.org/paper/Optical-Transmission-System-Simulation-for-Analysis-of-Self-Phase-Modulation-Non-Linearity.html.

[9] G.P.Agrawal, ”*Nonlinear Fiber Optics (Fifth Edition)*”, The Institute of Optics University of Rochester, New York, 2013, ISBN: 978-0-12397-023-7, 621 pp.

[10] G.Deshmukh, S. Jagtap , ”*Four Wave Mixing in DWDM Optical System*”, International Journal of Computational Engineering Research, Vol, 03, Issue 6, June 2013, pp.07-11.

[11] T. K. Chiang et al. , “*Cross-phase modulation in fiber links with multiple optical amplifiers and dispersion compensators*, ” Journal of Lightwave Technology, vol.14, no.3, March 1996, pp.249-259.

[12] N. Sharma , “*Fiber Non-linearities : Threat to the Performance Analysis of Optical Fiber System*“ , ISBN: 978-981-08-7932-7, Proc. of the International Conference on Advanced Computing and Communication Technologies (ACCT 2011), 2011, pp. 382-386.

[13] J. Tang et al., “*The Multispan Effects of Kerr Nonlinearity and Amplifier Noises on Shannon Channel Capacity of a Dispersion Free Nonlinear Optical Fiber*,”Journal of Lightwave Technology, vol. 19, No.8 ,August 2001. pp. 1110-1115.

[14] I. Rasheed, M. Abdullah, S. Mehmood, M. Chaudhary , "*Analyzing the non-linear effects at various power levels and channel counts on the performance of DWDM based optical fiber communication system*," Emerging Technologies (ICET), 2012 International Conference on , vol.1, no.5, Oct. 2012, pp. 8-9.

[15] O. Ozolins, G. Ivanovs,” *Estimation of DWDM Transmissionfor Broadband Access with FBG Technology*”, Scientific Institute of Telecommunications, Riga Technical University, ISSN 1392 – 1215, 2011. No.5, pp.11-14.

[16] T. Sabathi, K. Bavithra,” *Combating Cross Phase Modulation and its Polarisation Effects in Dense Wavelength Division Multiplexed Systems*”A. P,Department Of Electronics and Communication Engineering, Mepco Schlenk Engineering College,Sivakasi, Tamilnadu, India. International Journal of Scientific Engineering and Technology, ISSN : 2277-1581, June 2013, Volume No.2, Issue No.6, pp. 528-531.

[17] R. Ramaswami, K.N. Sivarajan, G.H.Sasaki, *”Optical Network”*, A practical perspective third edition,ISBN: 978-0-12374092-2, 2010, 845.pp.

[18] O.Farzao, c.Correia,etc.,”Stimulated *Raman Scattering and its Applications in Optical Communications and Optical Sensors*” The Open Optics Journal, 2009, Volume 3, pp. 1-11.

[19] T, Schneider, D Hannover, M. Junker,“ *Investigation of Brillouin Scattering in Optical Fibers for the Generation of Millimeter Waves*“, Journal of Lightwave Technology, vol. 24, no. 1, 2006, pp. 295-304.

Modelling the distribution of electromagnetic field and calculation of shielding effectiveness

[1]*Ján Zbojovský*(*4rd year*)
Supervisor: [2]*Alexander Mészáros*

[1,2] Dept. of Electric Power Engineering, FEI TU of Košice, Slovak Republic

[1]jan.zbojovsky@tuke.sk, [2]alexander.meszaros@tuke.sk

***Abstract*— This paper deals with the simulation of electromagnetic field and calculation of shielding effectiveness. The impact of shielding materials to its distribution is observed. The simulation is created in ANSYS Workbench, and HFSS based on finite element method.**

***Keywords*—electromagnetic field, simulation, ANSYS, shielding**

I. INTRODUCTION

Exposure to electromagnetic fields is not a new phenomenon. However, during the 20th century, environmental exposure to man-made electromagnetic fields has been steadily increasing as growing electricity demand, ever-advancing technologies and changes in social behavior have created more and more artificial sources. Everyone is exposed to a complex mix of weak electric and magnetic fields, both at home and at work, from the generation and transmission of electricity, domestic appliances and industrial equipment, to telecommunications and broadcasting. At present it is not possible to confirm or deny possible harmful effects of electromagnetic fields on human health. The significant number of completed and published research says that this issue is very actual. Electrical and biological systems must be resistant to other systems and its action may adversely affect the normal operation of other systems and devices. The motivation to solve this research task results from activity 2.1- "Research the impact of building materials to the distribution of electromagnetic field", project: Protection the population of Slovak Republic against the effects of electromagnetic fields founded by the requirement of European Union.

II. PROBLEMATIC OF ELECTROMAGNETIC FIELD

The electromagnetic field is a physical field in which electric and magnetic forces operate in space. Nowadays, the effects of electromagnetic fields on living organisms are discussed by the International Organization for Non-Ionizing Radiation Protection (INCIRP) and the World Health Organization (WHO). Biological effects are measurable changes induced by some stimulus or change in the environment. These changes do not necessarily lead to damaging health. The human body has many compensatory mechanisms and the ability to adapt to various changes. These reactions are normal signs of life. Biological effect, however, could have an adverse effect in the case where the body is exposed to long periods of harmful factors and compensatory mechanisms are insufficient.

We can solve the problems about electromagnetic field by measurement or by modelling and computing specific situation [1], [2], [3], [4], [5].

III. SOLUTION OF ELECTROMAGNETIC FIELD

The aim of this work is the simulation of electromagnetic field and impact of materials with various properties to its distribution. The model and simulations are created in computational software Ansys Workbench 14.5 for static calculations and in HFSS for frequency depending simulations. Calculations are based on the finite element method.

Following pictures shows the distribution of static magnetic field.

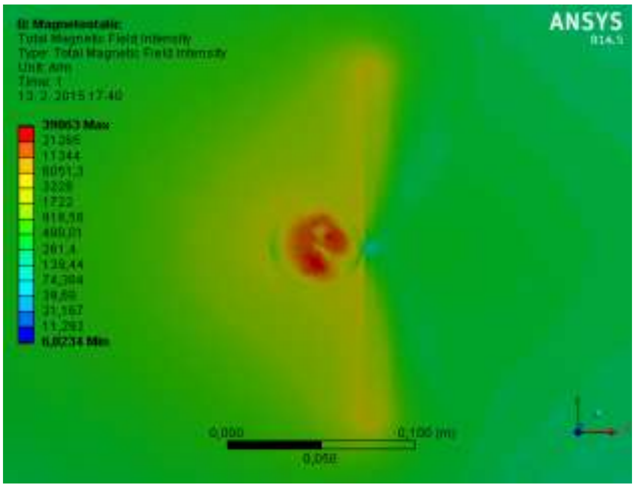


Fig.1. Distribution of the magnetic field

In my simulation the electrical properties of shielding material and frequency where changed. Shielding effectiveness was observed by the change the thickness of shielding materials and their distance from the source. Then relative permeability for magnetic field and isotropic resistivity for electric field was changed. Following picture shows the propagation of electric field in waveguide, in the middle is material from polystyrene, frequency range is from 1,5 to 5GHz.

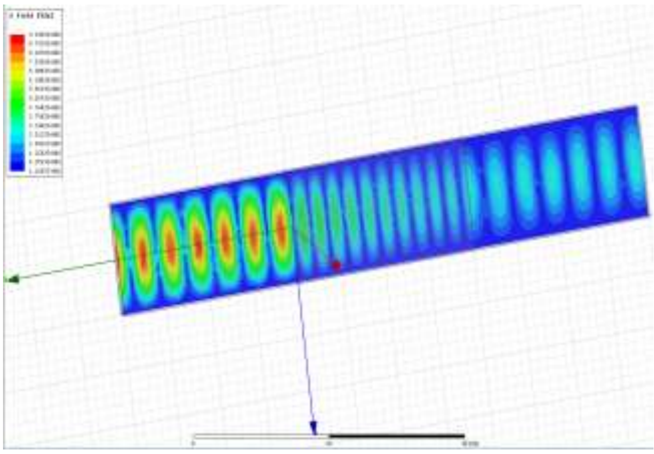


Fig.2. Propagation of electric field through the polystyrene

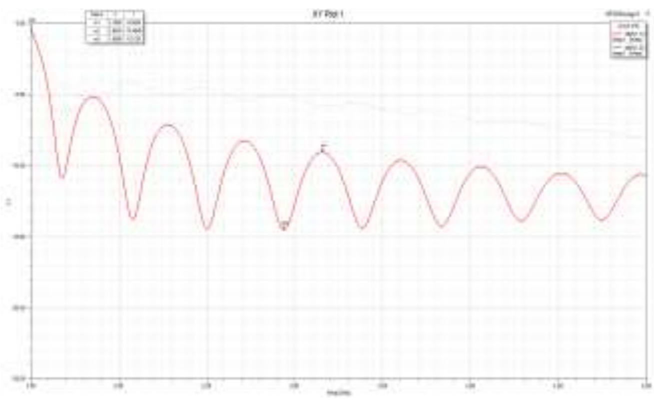


Fig.3. Electric field behind the shielding material (polystyrene)

Some results from simulations are available on [6], [7], [8], [9], and some results from measurement in [10], [11]. Shielding effectiveness was calculated according to [5], [12].

IV. Conclusion

My work is focused on simulations of the electromagnetic field distribution and the ways how to shield it. This was done for static calculations in Ansys Workbench using Emag and in HFSS for frequency depending simulations. All results from simulations are published in my dissertation work. Further will be appropriate to continue the research of electromagnetic field propagation in the next diploma and dissertation thesis. It would be good to focus on the influence of electromagnetic fields to the biological systems. The question is how to solve this problematic by simulation.

Acknowledgment

This publication is the result of the Project implementation: Protection of the population of the Slovak Republic from the effects of electromagnetic fields, ITMS: 26220220145, Activity 2.1 supported by the Research & Development Operational Programme funded by the ERDF .

The author wish to acknowledge the Project: Convergence, Regional Competitiveness and Employment ITMS: 26210120002, 26230120002 supported by the Research & Development Operational Programme funded by the ERDF.

The authors wish to acknowledge Scientific Grant Agency of The Ministry of Slovak Republic and Slovak Academy of Science for funding of experimental works in the frame of VEGA No.1/0132/15 grant Research of the penetrating of high frequency electromagnetic waves through ecological building materials.

We support research activities in Slovakia / Project is cofinanced from EU funds.

References

[1] Vecchia P, et al: "Exposure to high frequency electromagnetic fields, biological effects and health consequences (100kHz to 300GHz)", INCIRP 16/2009

[2] Cook, David M.: The electromagnetic Field, Prentice-Hall, New Jersey 1975, 510 s. 0-13-913293-7

[3] Rusiecki A: "Calculations and measurements of shielding effectiveness of slotted enclosure with built-in conductive stirrer" In: Przegland Elektrotechniczny, ISSN 0033-2097, R. 88 NR 10b/2012, p. 328-329

[4] Mayer, D., Ulrych B.: "Základy numerického řešení elektrických a magnetických polí", SNTL/ALFA Praha 1988, 208str., 04-528-88

[5] S. Celozzi, R. Araneo, G. Lovat, "Electromagnetic Shielding", Electrical Engineering Department La Sapienza University Rome, Italy, IEEE Press.

[6] Zbojovský, J., Mészáros, A. Medveď, D.: "Simulation of electromagnetic fields propagation through the various materials" In EPE 2014: Electric Power Engineering 2014 : proceedings of the 15th International Scientific Conference : May 12-14, 2014, Brno, Czech Republic. ISBN 978-1-4799-3807-0.

[7] Zbojovský J., Mészáros, A., Medveď, D., Kolcun, M., Kolcunová, I., Dolník B., Kurimský, J., Pavlík, M.: "Modelovanie rozloženia elektromagnetického poľa pri použití tieniacej bariéry" In: EPE 2013 : Electric Power Engineering 2013 : proceedings of the 14th International Scientific Conference : May 28-30, 2013, Kouty nad Desnou, Czech Republic. - Ostrava : VSB – Technical University, 2013 P. 379 - 382. - ISBN 978-80-248-2988-3

[8] Zbojovský J., Mészáros, A., Medveď, D., Kolcun, M., Kolcunová, I., Dolník B., Kurimský, J., Pavlík, M.: "Vplyv tieniacich materiálov na rozloženie elektromagnetického poľa" In: Elektroenergetika 2013 : proceedings of the 7th International Scientific Symposium on Electrical Power Engineering : September 18-20, 2013, Stará Lesná, Slovak Republic. - Košice : TU, FEI, 2013 S. 439-442. - ISBN 978-80-553-1441-9

[9] Zbojovský, J., Mészáros, A. Medveď, D., Kolcun, M., Kolcunová, I., Dolník, B., Kurimský, J., Pavlík, M., Kurimský, P,: "Calculation of shielding effectiveness based on the results of simulation the electromagnetic field" In: Current Problems of Maintenance of Electrical Equipment and Management. - Košice : TU, 2014 S. 204-208. - ISBN 978-80-553-1818-9

[10] Pavlík, M., Kolcunová, I., Medveď, D.: "Measuring of the shielding effectiveness of electromagnetic field of brick wall in the frequency range from 1 GHz to 9 GHz" In EPE 2014: Electric Power Engineering 2014 : proceedings of the 15th International Scientific Conference : May 12-14, 2014, Brno, Czech Republic. ISBN 978-1-4799-3807-0.

[11] Pavlík, M, Kolcunová, I., Dolník B., Kurimský, J., Mészáros, A., Medveď, D., Kolcun, M., Zbojovský J. : "Meranie účinnosti tienenia v oblasti vyšších frekvencií" In: EPE 2013 : Electric Power Engineering 2013 : proceedings of the 14th International Scientific Conference : May 28-30, 2013, Kouty nad Desnou, Czech Republic. - Ostrava : VSB – Technical University, 2013 P. 269-273. - ISBN 978-80-248-2988-3

[12] E. Űnal, A. Gőkcen, Y. Kutlu, "Effective Electromagnetic Shielding", IEEE Microwave magazine, 1527-3342, 2006, p. 48 – 54.

Models of nonlinear elements and nonlinear circuit simulation using MATLAB

[1]*Jozef DZIAK (2st year)*
Supervisor: [2]Iveta TOMČÍKOVÁ

[1,2]Dept. of Theoretical and Industrial Electrical Engineering, FEI TU of Košice, Slovak Republic

[1]jozef.dziak@tuke.sk, [2]iveta.tomcikova@tuke.sk

***Abstract*—Paper deal with study the possibility, approach and the correct procedure for nonlinear circuit simulation using mathematical tool MATLAB. Especially, paper describes nonlinear circuit simulation, modeling of nonlinear elements and using graphical user interface in MATLAB.**

***Keywords*—nonlinear circuit simulation, circuit analysis, Matlab, graphical user interface.**

I. Introduction

Circuit simulation is a technique for checking and verifying the design of electrical and electronic circuits and systems prior to manufacturing and deployment. It is used across a wide spectrum of applications, ranging from integrated circuits and microelectronics to electrical power distribution networks and power electronics. Circuit simulation combines mathematical modelling of the circuit elements, or devices, formulation of the circuit equations and techniques for solution of these equations. [1]

Many different kinds of network element are encountered in network analysis. For circuit analysis it is necessary to formulate equations for circuits containing as many different types of network elements as possible. There are various methods for formulation of circuit equation.

Research in this topic consists of these parts:

1. Description of real circuits and their elements.
2. Design of real electrical and electronic components models.
3. Selection of appropriate methods to solve mathematical models of real electrical circuits.
4. Create application to solve nonlinear electrical and electronic circuits using MATLAB.
5. Choosing a suitable method for applications testing.

II. Initial State

Nonlinear circuits consist of linear and nonlinear elements. Before nonlinear circuits simulation simulation and modeling of nonlinear elements we muss describe models of linear elements. All models of active and passive linear elements for simulation of AC and DC steady-state circuits using Sparse Tableau Analysis are described in [1].

Before the start of circuit simulation chooses the method of preparation circuit equations is needed. The process of selecting a suitable method is described in [2]. Sparse Tableau Analysis (STA) with system of circuit equations (1) was selected. Subsequently, linear elements models were adjusted for STA.

$$f(x)=\begin{bmatrix}\mathbf{A} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{I} & -\mathbf{A}^{\mathrm{T}}\\ \mathbf{Z} & \mathbf{Y} & \mathbf{0}\end{bmatrix}.\begin{bmatrix}\mathbf{i}\\ \mathbf{u}\\ \mathbf{v}\end{bmatrix}-\begin{bmatrix}\mathbf{0}\\ \mathbf{0}\\ \mathbf{s}\end{bmatrix} \tag{1}$$

In addition to the above activities options of MATLAB were examined, in particular MATLAB graphical user interface [2].

III. Description of Tasks and Results

Next research followed up the results described in initial status. Simulation of nonlinear circuits using MATLAB consist of:

A. Description and creating models of nonlinear elements.
B. Selection of method for nonlinear equations systems solving.
C. Creating application in graphical user interface of MATLAB.
D. Creating and testing application of nonlinear circuit simulation.

A. Models of nonlinear elements

First step of nonlinear circuit simulation is creation models of nonlinear elements. Models of real circuit elements consist of linear and nonlinear elements. Linear elements are described in [1]. We know three varieties of ideal nonlinear elements:

Nonlinear resistors divided into voltage-controlled with element equation (2) or current-controlled with element equation (3).

$$i_R = f(u_R) \tag{2}$$

$$u_R = f(i_R) \tag{3}$$

Nonlinear cumulative elements are nonlinear voltage-dependent capacitor $C(u)$ and current-dependent inductor $L(i)$.

Nonlinear controlled sources divided into four groups:

- voltage control voltage source (VCVS)
- voltage control current source (VCCS)
- current control voltage source (CCVS)
- current control current source (CCCS)

Process of creating a model for the STA method will be demonstrated on the basic nonlinear element semiconductor diode Fig. 1 case a). STA models of real elements consist of ideal elements and models of ideal elements will be used. Elements equation of semiconductor diode is (4). Therefore, we can use only one ideal element (voltage-control nonlinear resistor) to create model of diode Fig. 1 case b). Companion model of diode Fig. 1 case c) consist of resistor and voltage source.

$$i_D = I_{sat}\left(e^{\frac{u_D}{\eta . U_T}} - 1\right) \tag{4}$$

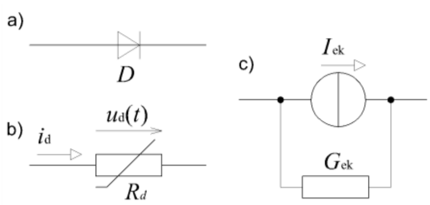


Fig. 1 a) semiconductor diode, b) replacement with nonlinear resistor, c) companion model of diode.

Second way to create a model of nonlinear element is linearization of this element. In this case, only linear ideal elements will be used to create the STA model.

B. Simulation of nonlinear circuits in MATLAB

Linear STA equation system will be extended of nonlinear elements equation written in the vector of nonlinear element equation **G(x)**. G(x) consists of nonlinear functions $g_i(x)$. Every entry $g_i(x)$ is the nonlinear function corresponding to a single nonlinear element in circuit. This extension system called nonlinear STA system (5) is used to simulate nonlinear circuits.

$$f(x) = \begin{bmatrix} \mathbf{A} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & -\mathbf{A}^{\mathrm{T}} \\ \mathbf{Z} & \mathbf{Y} & \mathbf{0} \end{bmatrix} . \begin{bmatrix} \mathbf{i} \\ \mathbf{u} \\ \mathbf{v} \end{bmatrix} + [\mathbf{H}].[\mathbf{G}(\mathbf{x})] - \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{s} \end{bmatrix} \tag{5}$$

C. Solving of nonlinear equations system

We must find solving of nonlinear equation system (5). There are numerous approaches to solving nonlinear systems, most based on using some type of approximation involving linear functions. Nonlinear problems are often treated numerically by reducing them to a sequence of linear problems.

Approximate the graph of $y = f(x)$ by the tangent line at a point $x^{(0)}$ near the desired root, and use the root of the tangent line to approximate the root of the original nonlinear function $f(x)$. This leads to Newton's iterative method (6) for finding successively better approximations to the desired root [4].

$$x^{(k+1)} = x^{(k)} - \frac{f\left(x^{(k)}\right)}{f^{\prime}\left(x^{(k)}\right)} \tag{6}$$

This generalizes to handling systems of nonlinear equations. Let $f(x) = 0$ denote a system of $p = n+2m$ linear and nonlinear equations in $p = n+2m$ unknowns (7). Newton's method for solving this system is given by (8) when k = 0, 1, 2, … . In this, $f'(x)$ is a generalization of the derivative known as the Jacobian matrix of system $f(x)$, and the second equation is a linear system of order p [4].

$$x = (i_1, \ldots, i_m; u_1, \ldots, u_m; v_1, \ldots, v_n) \tag{7}$$

$$\begin{gathered} \mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \delta^{(k)} \\ \mathbf{f}\left(\mathbf{x}^{k}\right)\delta^{(k)} = -\mathbf{f}\left(\mathbf{x}^{k}\right) \end{gathered} \tag{8}$$

An important related class of problems occurs under the heading of optimization. Given a real-valued function $f(x)$ with x a vector of unknowns, a value of x that minimizes $f(x)$ is sought. In some cases x is allowed to vary freely, and in other cases there are constraints on x.

D. Graphical user interface

User-friendly interface should be used to enter information about circuit and print output information from circuit simulation. MATLAB provides a simple graphical interface called Graphical user interface (GUI). GUIs provide point-and-click control of software applications, eliminating the need to learn a language or type commands in order to run the application. Therefore, GUI was used to create applications for nonlinear circuit simulation .

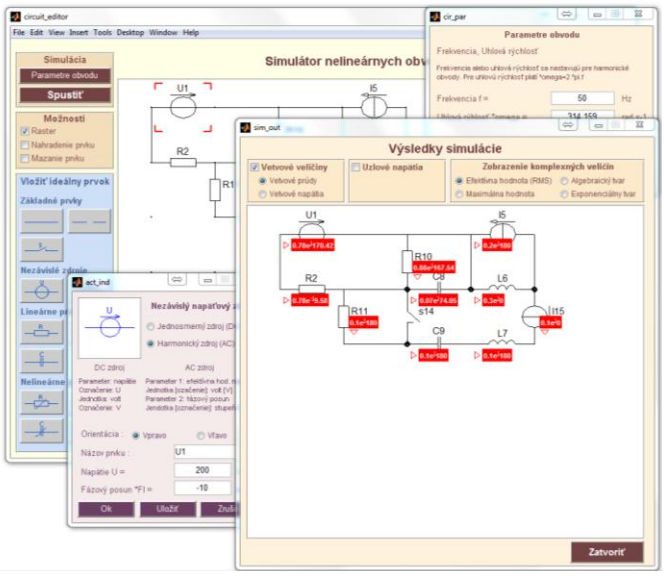


Fig. 2 Circuit simulator in MATLAB Graphical user interface.

IV. Conclusion

Tasks for next research are:

1. Description of real elements and creating models of real circuit elements.
2. Improve graphical application of nonlinear circuit simulation.
3. Testing application of nonlinear circuit simulation.

Acknowledgments

The paper has been prepared under support of Slovak grant project KEGA No. 015TUKE-4/2015.

References

[1] Najm F. N. – Circuit Simulation, John Wiley & Sons, Inc., Hoboken, United States of America, 2010, ISBN 978-0-470-53871-5.

[2] Dziak J. – Selection of method for computer-based circuits simulation. 13th Scientific Conference of Young Researchers. Herľany, 2013. ISBN 978-80-553-1422-8.

[3] Dziak J. – Linear circuit simulation using MATLAB and modeling of nonlinear elements. 14th Scientific Conference of Young Researchers. Herľany, 2014. ISBN 978-80-553-1714-4.

[4] Riečanová Z. and col. – Numerical methods and mathematical statistics. ALFA/SNTL, Bratislava, ISBN 063-559-87.

Modern Trends in Unmanned Aerial Vehicles Development

[1]*Ján BAČÍK (3st year),*
Supervisor: [2]Pavol FEDOR

[1,2] Department of Electrical Engineering and Mechatronics, Technical University of Košice, Slovak Republic

[1]jan.bacik.2@tuke.sk, [2]pavol.fedor@tuke.sk

***Abstract*—The paper deals with short overview of process of the unmanned aerial vehicles (UAV) development. Some the newest trends are presented from the field of sensorial system and navigation and basic principles of 3D simulator development are described.**

***Keywords*— UAV, sensorial system, machine vision, mapping, controlling**

I. Introduction

Complex development of small unmanned helicopter involves many small, but very important subtasks like attitude and heading reference system, inertial navigation system, simulation and control systems design. These all tasks could be labeled as low-level unmanned aerial vehicle design [1, 2, 3].

On other hand, industry misses solutions that enable small UAVs deployment in real life scenarios [4]. Deployment in real life requires solving problems associated with tasks of odometry in real environment, mapping and path planning [5]. All these tasks are associated with concept of point clouds. Point cloud is special data structure, specifically a set of data points in some coordinate system. In a three-dimensional coordinate system, these points are usually defined by X, Y, and Z coordinates, and often are intended to represent the external surface of an object.

The objective of the paper is survey of our work techniques that allow deployment of small unmanned aerial vehicles. These techniques are based on point clouds, stereo vision and intelligent techniques of control systems. We describe step by step each task necessary for design small UAV capable autonomy fly in real environment.

II. Visual Odometry

For our purpose we chose combination techniques of machine vision and classical inertial navigation. Our sensory system consisted of a stereo camera and triple axis inertial measurement unit. The output of stereo camera is pair of images, which are mutually displaced. We can determine true dimensions and distance of each point in the space based on images displacement. Necessary condition, for proper determination of all dimensions in space, is proper calibrated camera. Camera calibration is realized by calibration chessboard with known dimension of squares and their number.

Visual odometry was tested on real data. Data were acquired of flying hexacopter in real environment. For data accuracy we used scenario of industrial hall. Graphic user interface with original stereo camera output, artificial horizon and displacement of stereo images is shown in the Fig. 1.

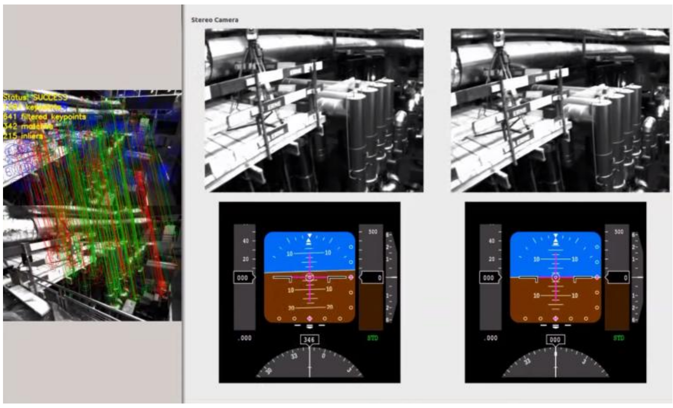


Fig. 1 Graphic user interface for visual odometry visualization

III. 3D Mapping

When we had finished visual odometry, our next step was creating three dimensional reconstruction of real environment and visualizing it as three dimensional maps. For this purpose we used again stereo camera connected to hexacopter. During flight we acquired single point clouds and known position of hexacopter in the hall (from visual odometry). Subsequently, the individual point clouds were registered and connected together. Process of map rendering was visualized on operator’s computer (Fig. 2).

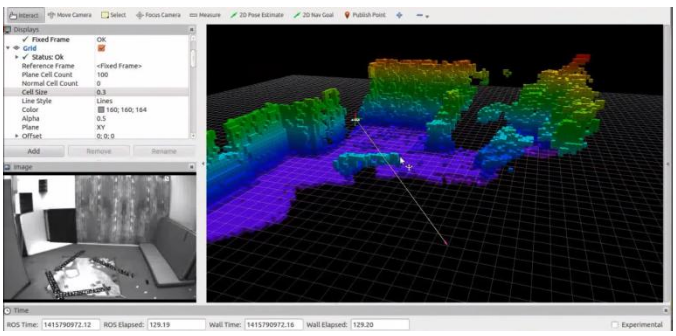


Fig. 2 Creating of 3D map based on stereo camera images

IV. Fuzzy Control

When we had information about hexacopter position in the space and three-dimensional map, out next step was designing controller, which allow hexacopter control position in certain space.

The hexacopter 3D model visualization was developed in Gazebo simulation environment (Fig. 3) which is provided with a robust physics engine, with quality graphics and conventional programming and graphic interface. The model runs in real time which enables its direct control.

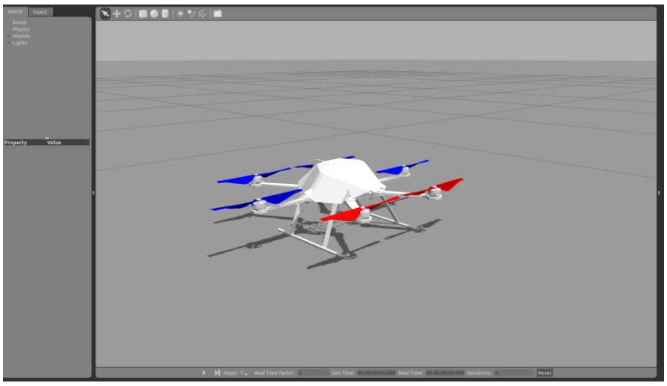


Fig. 3 Hexacoter simulation model in Gazebo enviroment

The objective of the fuzzy controller design was the control of hexacopter position in space, i.e. position control in all three axes X, Y and Z. A separate discrete fuzzy controller with standard PI structure was designed for position control in each axis [5]. The controller inputs are control deviations between the desired and the real position in the individual axes and their difference, and the output is the gain of the corresponding control action. The resulting fuzzy controller diagram is shown in Fig. 4.

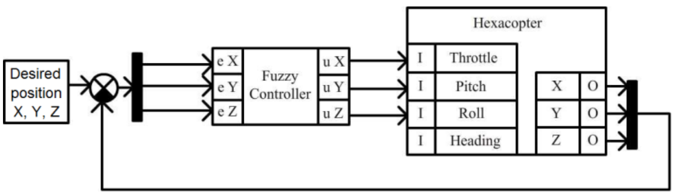


Fig. 4 Block diagram of hexacopter with control structure

The main problem in hexacopter altitude control (motor thrust control) is the fact that its propellers are not capable of generating negative thrust. If we want to stop the hexacopter at a certain altitude, we have to apply the brake before reaching this altitude, while the braking is possible only through reducing motor thrust, or stopping the motors. Altitude control as such therefore means balancing between motor thrust force and gravitational force.

In the design of hexacopter altitude fuzzy controller (position in Z-axis) we used a standard Mamdani type controller, and the fuzzification of variables and proposal of rules were based on the analysis of experimentally measured data and experience of a pilot – expert.

The experimentally measured responses of fuzzy control of hexacopter position in space to step changes of the set-point for the individual axes are shown in Fig. 5.

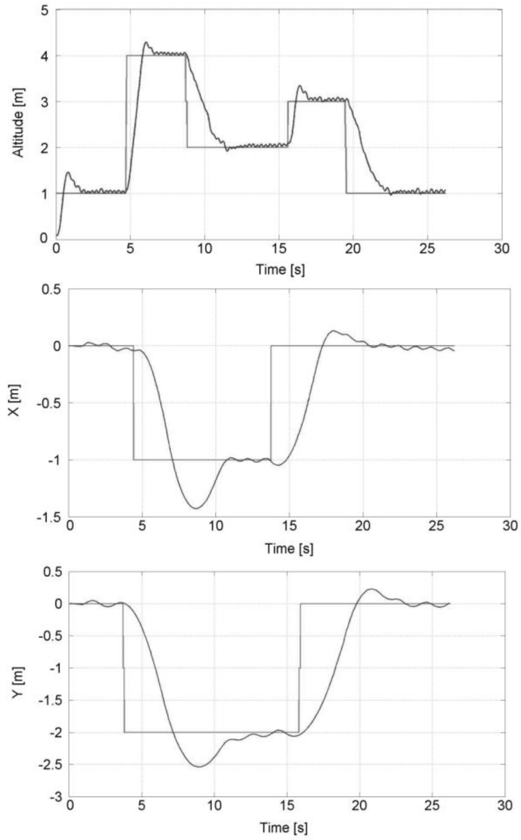


Fig. 5. Experimental data in Z, X and Y-axis

V. Future Work

Next goal of development is applying all mentioned techniques to helicopter T-Rex 600. Our final goal is creating visual real-time gazebo simulation model. Three-dimensional model for simulator, created in CAD program, is in Fig. 6.

Fig. 6 T-Rex 600 - 3D model for application in Gazebo environment

Acknowledgment

The work has been supported by project KEGA 011TUKE-4/2013.

References

[1] R.P.G. Collinson, *Introduction to Avionics Systems, third edition*, Springer, New York, 2011.
[2] D.H. Titterton, J.L. Weston, *Strapdown Inertial Navigation Technology, second edition*, Institution of Engineering & Technology, Reston, 2005.
[3] T. Hsu, *MEMS & Microsystems: Design, Manufacture, and Nanoscale Engineering*, John Wiley & Sons Ltd, New Jersey, 2008.
[4] European Robotics Challenges, http://www.euroc-project.eu/
[5] G. Cook, *Mobile Robots: Navigation, Control and Remote Sensing*, Wiley & Sons Ltd, New Jersey, 2011.

MPPT methods for Converters of Photovoltaic Systems

[1]*Matej BEREŠ (1st year)*
Supervisor: [2]Dobroslav KOVÁČ

[1,2]Dept. of Theoretical and Industrial Electrical Engineering, FEI TU of Košice, Slovak Republic

[1]matej.beres@tuke.sk, [2]dobroslav.kovac@tuke.sk

***Abstract*-The paper presents a brief overview of MPPT (Maximum Power Point Tracking) techniques for Photovoltaic (PV) applications. These techniques include Perturb and Observe (P&O), Incremental Conductance (IC) and Constant Voltage (CV). The MPPT algorithms of P&O, IC, CV methods are described with help of flowchart. The Fuzzy control and Neural Network (NN) are also briefly described.**

***Keywords*- Photovoltaic, Maximum Power Point, Perturb and Observe, Incremental Conductance, Constant Voltage, Fuzzy Control, Neural Network.**

I. Introduction

The growth of emerging economies and their energy consumption, coupled with strong concerns about the environmental conservation and the excessive production of CO have generated an urgent need to find economically viable, clean and sustainable solutions [1]. Photovoltaic energy stands out as a free power source and essentially distributed throughout the planet, however this technology is still considered expensive considering factors such as high initial investment, the limited life span and low energy efficiency rates [2].

The characteristic of PV panel is that the available maximum power is provided only in a single operating point given by a localized voltage and current known [3], called Maximum Power Point (MPP).

In order to gain maximum power, MPPT (Maximum Power Point Tracking) is an essential part of a PV generation system. Because of the nonlinear voltage-current characteristics of PV cells, the power versus voltage (P-V) curve in solar cells has more complicated nonlinear relationship when solar illumination and ambient temperature change, so the MPP is difficult to solve analytically, and therefore numerous techniques have been proposed to realize MPPT [4]. The literature proposes many MPPT strategies that guarantee to operate at the Maximum Power Point (MPP) of the I-V curve of the PV array.

This paper is organized as follows: Section II presents characteristic of photovoltaic cell. Section III presents the some methods for achieve MPP.

In the first instance it's necessary to know what the PV cell is and what characteristics have.

II. Characteristic of Photovoltaic Cell

Photovoltaic cells consist of a silicon P-N junction that when exposed to light releases electrons around a closed electrical circuit. Electrons from the cell are excited to higher energy levels when a collision with a photon occurs. These electrons are free to move across the junction and create a current [5].

Characteristic of photovoltaic cell is illustrated in Fig. 1. The figure illustrates solar illumination which is represented as symbol *S*. The *S* variation is from 200W/m^2 to 1000W/m^2, and temperature *T* is constant 40 ºC. Besides the solar illumination, another important factor influencing the characteristics of a photovoltaic module is ambient temperature, as shown in Fig. 2. In this figure is solar illumination constant 1000W/m^2, and temperature *T* vary from 20 ºC to 100 ºC [6].

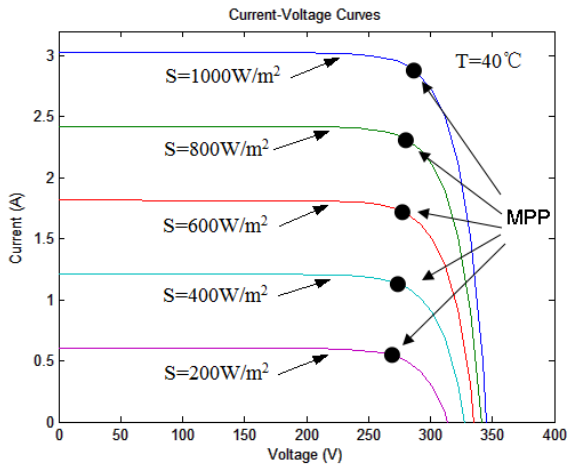


Fig. 1 Current versus voltage curves of PV array influenced by solar illumination [6].

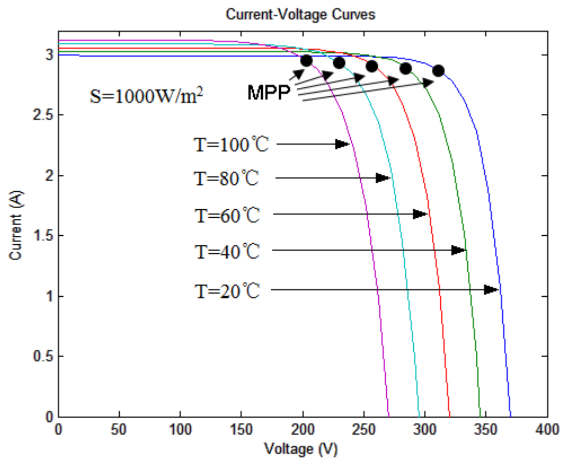


Fig. 2 Current versus voltage curves of PV array influenced by temperature [6].

Next factor that have characteristic influence of solar cell is partial shading, see Fig. 3.

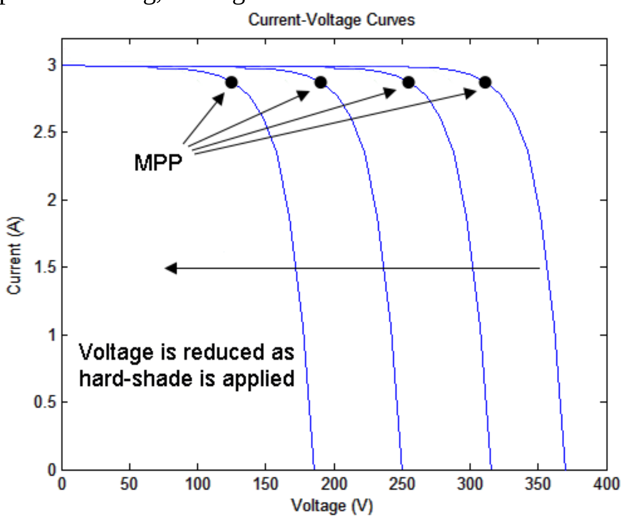


Fig. 3 Current versus voltage curves of PV array influenced by partial shading [7].

It follows that, in the design of algorithm for solar cell converters is necessary to take account of these influences.

III. Control algorithms for maximum power point

The configuration diagram of the photovoltaic generating system for the MPPT control is as shown in Figure 4. It is composed of the PV module to convert solar energy into the electrical energy and the boost converter for step-up voltage the PV voltage. In the PV module, a voltage and current are measured and the power is calculated and the MPPT control is performed about the solar illumination change. By using V_{ref} outputted from the MPPT control, the boost converter is controlled through PWM [17].

In recent years, a large number of techniques have been proposed for tracking the MPP in the literature [8-13] and the most commonly used are the Perturb and Observe (P&O) method, Incremental Conductance (IC) method and Constant Voltage (CV) method. There are also some other methods, which are more or less deviations from the above algorithms. Also the MPP is possible achieve by using fuzzy control or neural network.

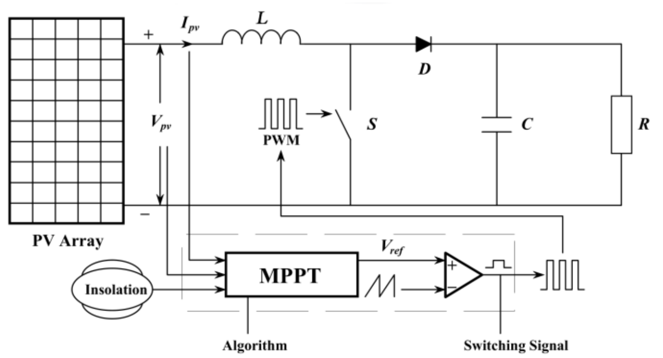


Fig. 4 System configuration diagram for the PV MPPT control [17].

A. *Perturb and Observe method*

The P&O algorithm is very popular and the most commonly used in practice because of simplicity in algorithm and ease of implementation.

The P&O algorithm is also called “hill-climbing”, but both names refer to the same algorithm depending on how it is implemented. Hill-climbing involves a perturbation on the duty cycle of the power converter and P&O a perturbation in the operating voltage of the DC link between the PV array and the power converter [8].

The basic flowchart implementing the P&O algorithm is shown in Fig. 5. In this algorithm the operating voltage of the PV module is perturbed by a small increment, and the resulting change of power, ΔP is observed. If the ΔP is positive, then it is supposed that it has moved the operating point closer to the MPP. Thus, further voltage perturbations in the same direction should move the operating point toward the MPP. If the ΔP is negative, the operating point has moved away from the MPP, and the direction of perturbation should be reversed to move back toward the MPP, [14].

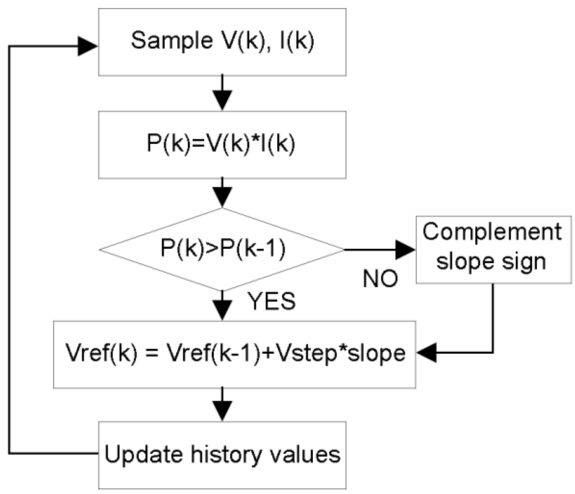


Fig. 5 Flowchart of P&O MPPT algorithm

The shortcomings of the P&O method can be improved by using the incremental conductance method.

B. *Incremental Conductance method*

The incremental conductance algorithm [10] makes use of the fact that the slope of the solar module P-V curve is zero at the MPP, positive on the left of the MPP, and negative on the right of the MPP. The IC method has the advantage over the P&O of not oscillating around the MPP under rapidly changing environmental conditions, but has a more complex circuitry. The accuracy of the method depends on the iteration size, which is usually fixed for the conventional IC method. There are many improvements to the conventional IC [15]. The basic flowchart implementing the (IC) algorithm is shown in Fig. 6.

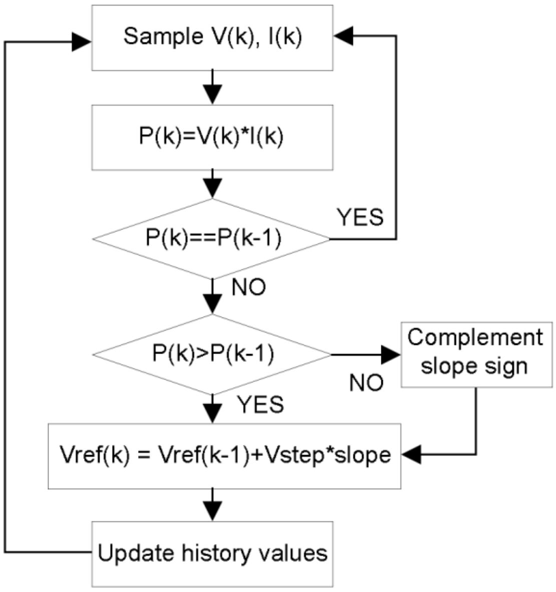


Fig. 6 Flowchart of IC MPPT algorithm

C. Constant Voltage method

The output voltage of the solar cell array has the constant voltage characteristic having the little bit of vibration amplitude about the solar illumination change. Therefore, it can be said to the CV control method in which it sets as V_{ref} and it controls by the constant voltage. This method is not needed the calculated power value for an output. The duty of the Boost converter is determined by the control circuit and the DC voltage of the output terminal is consistently maintained by V_{ref} value. The CV control method has the disadvantage that is unable to track the MPP in solar illumination rapidly change and the power efficiency is reduced [17]. However, there is the advantage of reducing the sensor of an array and DC part and Fig. 7 shows the flowchart of the CV MPPT control method.

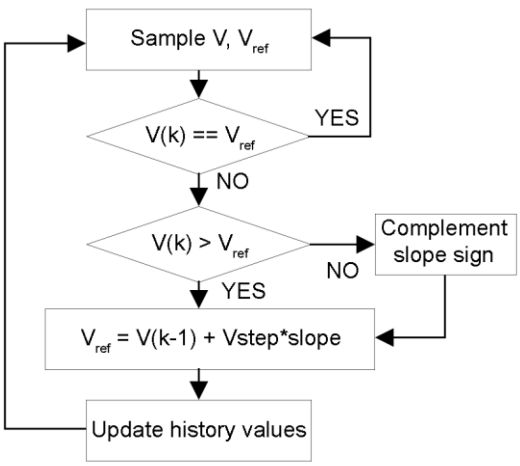


Fig. 7 Flowchart of CV MPPT algorithm

D. Constant Voltage and Incremental Conductance methods together

CV method is a simplified MPPT method and is actually a voltage regulator control. The method keep the output voltage of the PV array at a constant voltage stage, since the MPP will vary with illumination intensity or temperature change, The method could not track the MPP rapidly and precisely especially in fluky weather [16].

From the preceding analysis, the constant voltage method cannot guarantee that the output power of PV array is MPP, and only guarantee the work point is near the MPP. In order to maximize the performance of PV array, the IC method is used in the next step after the control target of constant voltage method is achieved. The method is different from the traditional IC method, and the CV method is adopted when the external environment change. Therefore, the stability of the MPP is the main control objective in the IC method and the step size can be selected much smaller than the traditional IC method, which can effectively eliminate the power oscillation around MPP [16].

The basic flowchart implementing the CV with implementing IC algorithm is shown in Fig. 8. The output voltage and output current of PV array at k time is defined respectively as V_k and I_k, and the working status of the system is determined by the output voltage of PV array as following. In the first instance, the CV method is implemented when the output voltage of PV array is outside of the range of $V_s \pm \Delta V$, where V_s is the set voltage, and ΔV is the range. In the next instance, the IC method is implemented when the output voltage of PV array is in the range of $V_s \pm \Delta V$. The working point of PV array is repeatedly adjusted until $dI / dV = -I / V$, and the system working point is the MPP at this moment.

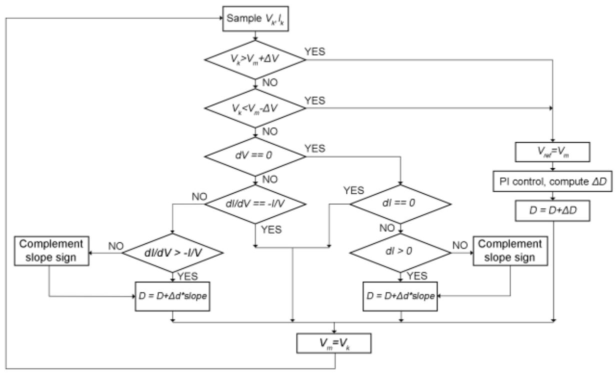


Fig. 8 Flowchart of CV with IC MPPT algorithm [16]

E. Fuzzy control method

The fuzzy control has adaptive characteristic in nature, and can achieve robust response of a system with uncertainty, parameter variation and load disturbance. It has been broadly used to control ill-defined, nonlinear or imprecise system. Fuzzy logic controllers have the advantages of working with imprecise inputs, not needing an accurate mathematical model, and handling nonlinearity. Therefore has been successfully applied in many fields, such as industry controls. Fuzzy control does not require accurate models of control object [6].

A MPP search based on fuzzy heuristic rules, which does not need any parameter information, consists of a stepwise adaptive search, leads to fast convergence and is sensorless with respect to sunlight and temperature measurements. The control objective is to track and extract maximum power from the PV arrays for a given solar insolation level. The maximum power corresponding to the optimum operating point is determined for a different solar insolation level and temperature [6].

The fuzzy logic consists of three stages: fuzzification, inference system and defuzzification. Fuzzification comprises the process of transforming numerical crisp inputs into linguistic variables based on the degree of membership to certain sets. Membership functions are used to associate a grade to each linguistic term. The number of membership functions used depends on the accuracy of the controller, but it usually varies between 5 and 7 [6], [18-19]. The inference system corresponds to a set of fuzzy that have learning capability to approximate nonlinear functions.

The last stage of the fuzzy logic control is the defuzzification. In this stage the output is converted from a linguistic variable to a numerical crisp one again using membership functions. There are different methods to transform the linguistic variables into crisp values. It can be said that the most popular is the center of gravity method. However the analysis of these methods is beyond the scope of this paper, therefore the more information about this control method can be seen in literature [6].

F. *Neural networks*

Another MPPT method well adapted to microcontrollers is Neural Networks (NN) [20]. The simplest example of a NN has three layers called the input layer, hidden layer and output layer, as shown in Fig. 9. The input variables can be parameters of the PV array such as *V* and *I*, atmospheric data as illumination, temperature and partial shading or the combination of these. The output is usually one or more reference signals like the duty cycle or the DC-link reference voltage.

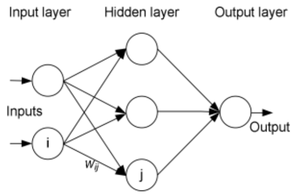


Fig. 9 Neural network layers [20]

The main disadvantage of this MPPT technique is the fact that the data needed for the training process has to be specifically acquired for every PV array and location, as the characteristics of the PV array vary depending on the model and the atmospheric conditions depend on the location. These characteristics also change with time, so the neural network has to be periodically trained.

IV. Conclusion

This paper commonly describes basic MPPT methods which can be used for control of converters. Some methods can be mixed for achieve better results. The paper also describes PV characteristics in the various conditions. In the next case can be these methods used for multiphase converters.

Acknowledgment

The paper has been prepared under support of Slovak grant project KEGA No. 015TUKE-4/2015.

References

[1] European Renewable Energy Council, EREC-2005. [Online]. Available: http://www.erec.org/renewable-energy/photovoltaics.html
[2] S. Liu and R. A. Dougal, “Dynamic multiphysics model for solar array,” IEEE Trans. Energy Convers., vol. 17, no. 2, pp. 285–294, Jun. 2002.
[3] F. Chekireb, C. Larbes, D. Rekioua, F. Haddad. Implementation of a MPPT fuzzy controller for photovoltaic systems on FPGA circuit. Energy Procedia pp. 541-594, 2011.
[4] Cuauhtemoc Rodriguez, and Gehan A. J Amaratunga, "Analytic Solution to the Photovoltaic Maximum Power Point Problem," IEEE Transactions on Circuit and Systems, Vol. 54, No. 9, pp. 2054–2060, September 2007.
[5] Nicola Femia, Giovanni Petrone, Giovanni Spagnuolo, and Massimo Vitelli, " Optimization of Perturb and Observe Maximum Power Point Tracking Method," IEEE Trans. on Power Electronic, Vol. 20, No. 4, pp.963- 972, July 2005.
[6] Jiyong Li; Honghua Wang, "Maximum power point tracking of photovoltaic generation based on the fuzzy control method," Sustainable Power Generation and Supply, 2009. SUPERGEN '09. International Conference on , vol., no., pp.1,6, 6-7 April 2009
[7] D. Sera, R. Teodorescu, J. Hantschel, and M. Knoll, “Optimized Maximum Power Point Tracker for Fast-Changing Environmental Conditions,” IEEE Transactions on Industrial Electronics, Vol. 55, No. 7, July 2008.
[8] T. Esram, and P. L. Chapman, “Comparison of photovoltaic array maximum power point tracking techniques,” IEEE Trans. Energy Convers., vol. 22, no. 2, pp. 439-449, Jun. 2007.
[9] I. Kim, M. kim, and M. Youn, “New maximum power point tracker using sliding-mode observer for estimation of solar array current in the grid-connected photovoltaic system, ” IEEE Trans. Ind. Electron., vol. 53, no. 4, pp. 1027-1035, Aug. 2006.
[10] A Pandey, N Dasgupta, A K Mukerjee. “A Simple Single-Sensor MPPT Solution”. IEEE Transactions on Power Electronics. America, vol. 22, pp. 698-700, March 2007.
[11] L Fangrui, D Shanxu, L Fei, et, al. “A variable step size INC MPPT method for PV systems”. IEEE Transactions on Industrial Electronics. America, vol. 55, pp.2622-2628, July 2008.
[12] T. L. Kottas, Y. S. Boutalis, and A. D. Karlis, “New maximum power point tracker for PV arrays using fuzzy controller in close cooperation with fuzzy cognitive networks”, IEEE Transactions on Energy Conversion, Vol. 21, pp.793-803, 2006.
[13] L. Wang, “An improved multiple fuzzy NNC system based on mutual information and fuzzy integral”, International Journal of Machine Learning and Cybernetics, Vol. 2, No. 1, pp. 25-36, 2011.
[14] Durgadevi, A.; Arulselvi, S.; Natarajan, S.P., "Study and implementation of Maximum Power Point Tracking (MPPT) algorithm for Photovoltaic systems," Electrical Energy Systems (ICEES), 2011 1st International Conference on , vol., no., pp.240,245, 3-5 Jan. 2011.
[15] Q. Mei, M. Shan, L. Liu, and J.M. Guerrero, "A Novel Improved Variable Step-Size Incremental-Resistance MPPT Method for PV Systems," IEEE Transactions on Industrial Electronics, vol. 58, no. 6 (June 2011): 2427- 2434.
[16] Yuansheng Xiong; Suxiang Qian; Jianming Xu, "Research on Constant Voltage with Incremental Conductance MPPT Method," Power and Energy Engineering Conference (APPEEC), 2012 Asia-Pacific , vol., no., pp.1,4, 27-29 March 2012.
[17] Jung-Woo Baek; Jae-Sub Ko; Jung-Sik Choi; Sung-Jun Kang; Dong-Hwa Chung, "Development of novel MPPT algorithm of PV system considering radiation variation," Control Automation and Systems (ICCAS), 2010 International Conference on , vol., no., pp.1235,1241, 27-30 Oct. 2010.
[18] N. Patcharaprakiti, S. Premrudeepreechacharn, "Maximum power point tracking using adaptive fuzzy logic control for grid-connected photovoltaic system," in Proc. IEEE Power Engineering Society Winter Meeting, 2002, pp. 372-377.
[19] Chao Zhang, Dean Zhao, "MPPT with asymmetric fuzzy control for photovoltaic system," in Proc. 4 th IEEE Conference on Industrial Electronics an Applications, 2009, pp. 2180-2183.
[20] Zhao Yong; Li Hong; Liu Liqun; Gao XiaoFeng, "The MPPT Control Method by Using BP Neural Networks in PV Generating System," Industrial Control and Electronics Engineering (ICICEE), 2012 International Conference on , vol., no., pp.1639,1642, 23-25 Aug. 2012.

Navigation of autonomous vehicles based on Intelligent Sensor Systems

[1] *Radovan SIVÝ (3rd year)*
Supervisor: [2]Daniela PERDUKOVÁ

[1,2]Dept. of Electrical Engineering and Mechatronics, FEI TU of Košice, Slovak Republic

[1]radovan.sivy@tuke.sk, [2]daniela.perdukova@tuke.sk

***Abstract*— Paper presents description of experimental vehicle concept and user interface to operate with Kinect device. Part of this paper is devoted to design of the experimental vehicle for the implementation and verification of mapping and navigation algorithms. Last part describes graphical user interface created in RVIZ package in ROS environment.**

***Keywords*—Experimental vehicle, DSP, Kinect, SLAM**

I. Introduction

The current trends of modern robotics tend to the development of small unmanned autonomous robots capable of self and long-term performance of the tasks set in unfamiliar surroundings without full external operator intervention. Dissertation thesis solves highly topical issue in the field of modern robotic systems with a focus on the development of an autonomous mobile robot using operating systems based on the latest technologies for a more efficient, faster and more reliable to fulfill its goals in unknown environment. The concept of the robotic system in the context of the autonomous mobile robots includes the knowledge of methods for locating the position of robotic systems, the methods for mapping the unknown space, and algorithms for their intelligent navigation, decision-making and management. [1][2][3] Important is the area of knowledge of sensor systems to create maps of unknown environment based of camera systems (machine vision). Machine vision is currently the world's most modern trend towards substitution of expensive sensors operating on the principle of laser or ultrasonic distance measurements. Therefore, the paper describes concept and structure of experimental vehicle for processing Kinects data and verification of SLAM methods and algorithms in real environment. The use of this sensor is currently a worldwide trend in development of robotic systems. [4][5]

II. Experimental Vehicle

An important part of the work will be an experimental verification of the mapping and localization results. It is very important to verify obtained theoretical results on real devices, because many of them cannot be verified only by mathematical simulation. Otherwise, they might not be sufficient for use in a real environment. Therefore, experimental model of the robotic ground vehicle was designed for the purpose of implementation and verification of mapping and navigation algorithms. The prototype of experimental model is based on the RC car chassis Traxxas Slash 1:10 TQ RTR (Fig.1). This RC car is equipped with a powerful DC motor Titan 12T 550, transmission Magnum 272 with hardened steel gears and gear ratio of 2.72 to 1. The RC car model was chosen due to its robust design and appropriate size of the chassis. During the design of the experimental vehicle we considered the fact that the chassis will be equipped with a platform. The platform will include the Kinect sensor and mini-PC for data processing. These devices represent an additional weight about 1.5 kg. Therefore, the oil shocks will provide stiffness of the chassis. Also, they are used to adjust the required angle and stiffness of the vehicle. The disadvantage of this RC car is an electronic controller XL-5 LVD, which is not able to actively brake during driving. Turning of the front axle is provided by servo Traxxas 2075.

Fig. 1.RC car chassis

III. Concept of Chassis

Concept of experimental vehicle is based on a platform that divides experimental vehicle into two parts, upper and lower. The platform is composed of shaped Plexiglas.

The lower (hardware) part of the vehicle is designed to store batteries and control devices for the operation of an experimental vehicle. Control devices consists of two circuit boards.

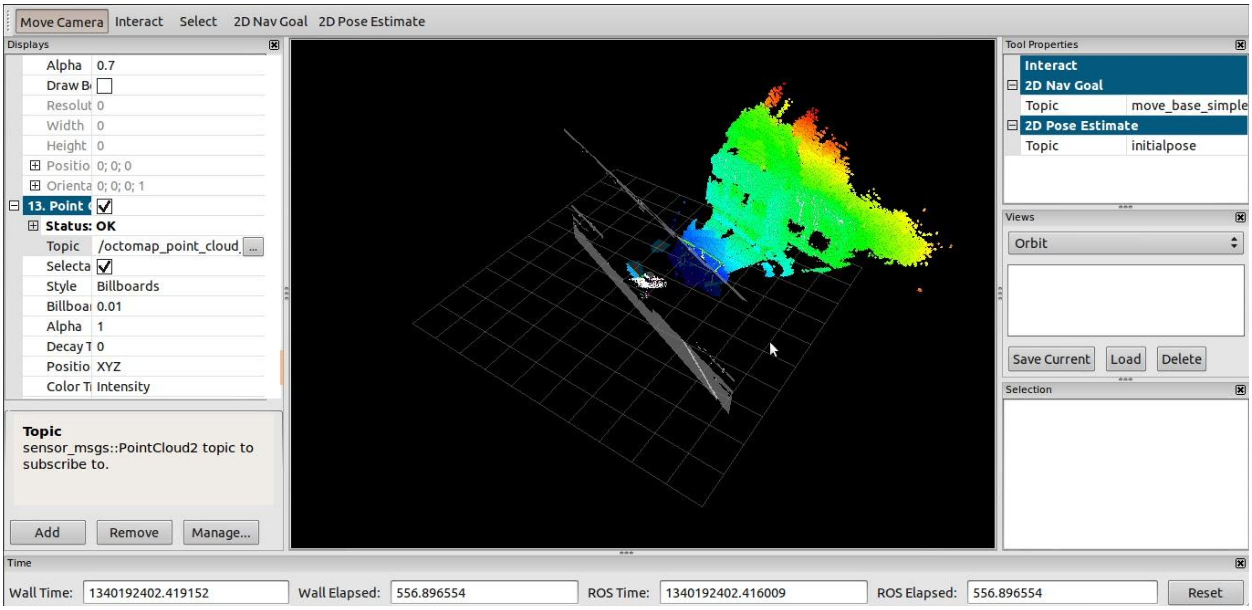


Fig. 2.Visualised data from Kinect in RVIZ

First circuit board ensures the regulation of engine speed, evaluation of the data from additional sensors and creates a communication bridge with a mini-PC through RS232. For these purposes digital signal processor C2000 TMS320F2808 has been used. During the design of the experimental vehicle we considered that the vehicle should be extended with other types of sensors and equipment. Therefore, we extended the PCB by analog and digital outputs from the DSP, PWM for servo/motor, and external power supply for additional sensors and features. For example, the distance sensors which create the front shield. The front shield is feature, formed by a pair of ultrasonic proximity sensors, used to detect obstacles that the main Kinect sensor cannot detect. Limitation of Kinect detection distance is from 0,4m to 5m.

The built-in controller for DC motor used in RC car is in terms of braking process for our solution unsatisfactory. Therefore, it was necessary to design and implement our own controller. It’s based on two H-bridges.

The upper (software) part is still under implementation. It will consist of sensor Kinect and mini-PC for processing the Kinect’s data.

IV. Graphical User Interface

We are using two important technologies in SLAM implementation, 3D depth sensor i.e. Kinect and Robot Operating System (ROS). ROS is a collection of open-source libraries that intend to solve various problems related to robotics like Object Identification, Gesture Recognition, Mobile Robotics, Path planning etc.[6] The graphical user interface (Fig.2) based on sensor Kinect data has been created for interaction between the user and mapping as well as localization process. The main task was to create a graphical interface that will display current workspace data and to render and analyze RGBD 3D map of the environment. Graphic User Interface was created in RVIZ. RVIZ is a 3D visualization ROS package for displaying sensor data and state information from ROS. Using RVIZ, you can visualize a virtual model of the robot. You can also display live representations of sensor values coming over ROS Topics including camera data, infrared distance measurements, sonar data, and more. Our interface shows depth data in distance color spectrum. We are using registered data structure for matching Color and Depth image from sensor Kinect. Next step is to implement Point Cloud Library and use PCL to generate 3D maps of environment.

V. Further Work

The main goal is to connect hardware (lower) part and software (upper) part together through RS232 link. This is the highest priority goal, because without communication we cannot test our SLAM algorithm program in real environment. Other important tasks will be to implement and improve SLAM algorithms. The goal is to improve accuracy of map and vehicle position on the map. At last, we want to add some optional features into experimental vehicle like charging station, headlights and detailed telemetry.

Acknowledgment

The work has been supported by project KEGA 011TUKE-4/2013.

References

[1] J. Stowers, M. Hayes, A. Bainbridge-Smith: Altitude Control of a Quadrotor Helicopter Using Depth Map from Microsoft Kinect Sensor, IEEE Mechatronics, Turkey, 2011

[2] J.Tran, A. Ufkes, M. Fiala, Low-Cost 3D Scene Reconstruction for Response Robots in Real-time, Proceedings of the 2011 IEEE International Symposium on Safety, Security and Rescue Robotics Kyoto, Japan, November 1-5 2011

[3] D. S. O. Correa, D. F. Sciotti, M. G. Prado, D. O. Sales, D. F. Wolf, F. S. Osório: Mobile Robots Navigation in Indoor Environments Using Kinect Sensor, Critical Embedded Systems, Brasil, 2012

[4] J. Z. Sasiadek, A. Monjazeb, D. Necsulescu: Navigation of an autonomous mobile robot using EKF-SLAM and FastSLAM, conference Control and Automation, France, 2008

[5] V. Piratla, S.B.Malode, S. K. Saini, A. Jakhotia, A.K. Sao, B. S. Rajpurohit, G. Haobijam: Autonomous Navigation in GPS Denied Indoor Environment Using RGBD Sensor, Kinect, IIT Manli, India, 2012..

[6] A Gentle Introduction to ROS JasonM. O’Kane, 2014. ISBN 978-14-92143-23-9

Optical Correlator in Image and Video Processing Systems

[1]*Dávid SOLUS (1st year)*
Supervisor: [2]Ľuboš OVSENÍK

[1,2]Dept. of Electronics and Multimedia Communications, FEI TU of Košice, Slovak Republic

[1]david.solus@tuke.sk, [2]lubos.ovsenik@tuke.sk

***Abstract*—Control systems are very important in these days. There are used in many industrial fields such as automotive, medicine and manufacture. Video assistant and video surveillance systems belong to these systems. The systems process large quantities of data in real time, so the processing speed is very important. The use of optical processing technology will increase the speed of processing amounts of data.**

***Keywords*—Colour models, edge detection, Optical correlator, optical Fourier transform.**

I. Introduction

Video assistant systems (VAS) are used to increase safety and comfort of the driver and passengers. The role of VAS is to evaluate the surrounding and warm the driver against potential danger. VAS supports the driver in various driving situation. Many of these systems use camera to acquire information about movement, speed, distance, shape or form. The motor car companies design their own video assistant systems such as Traffic Sign Detection, Audi Active lane assist, Parkpilot or BMW Night Vision.

Video surveillance systems are used to increase quality of product manufacture and to ensure security at airports, banks. More recently, government agencies, businesses and even school are turning toward video surveillance as a means to increase public security. Application of video surveillance include car and pedestrian traffic monitoring, human activity surveillance for unusual activity detection, people counting, controlling manufacturing or quality of product manufacture.

The optical systems that can handle the amount of data in a very short period of time are used more often. Example of such systems is an optical correlator. The optical correlator uses the optical properties of lens in process of correlation.

The paper is organized as follows. In Chapter II, optical correlator, types of optical correlator and optic Fourier transform are described. Chapter III is devoted to basic colour models and edge detection. Chapter IV contains a description of the application of optical correlator. Conclusion is summarized in Chapter VI.

II. Optical correlator

An optical correlator automatically recognizes or identifies the contents of an image by combining an incoming image with a reference image, and the degree of correlation after combining the images determining the Intensity of an output Light beam. Optical correlators can perform complex pattern recognition more rapidly than known Digital techniques. Optical correlators are capable of processing large amounts of Data in a data stream that can be useful in the detection, extraction and classification of desired information included in the data. Optical correlators are employed for many optical signal processing applications, including pattern and character recognition, implementation of optical interconnections in hybrid optoelectronic parallel computers, and in artificial neural Network technologies. Optical correlators are typically found in optical communication systems for signal detection applications involving data that are conveyed by light carriers whose Frequency or operating Wavelength is increasing as the technology continues to advance. The optical correlators are frequently implemented using spatial light Modulator (SLM) technology that involves spatial data serving as reference data and representative of two or three dimensional quantities commonly stored on holograms. Spatial light modulators (SLM) are transducers that modulate Incident Light in a spatial pattern corresponding to an electrical or optical input [1, 2].

A. *Types of Optical Correlator*

There are two main types of optical correlator, the Matched Filter (MF) and the Joint Transform Correlator (JTC). Both types achieve largely the same results, but process the information in different ways.

Matched Filter

The process involves taking the individual Fourier Transforms of the “Input” and “reference” images, taking their product and then Fourier Transforming the result. Usually, the reference transform is done off-line in electronics to provide the “filter”, which is usually displayed on a transmitting Spatial Light Modulator (SLM). Typically, the input transform is produced optically, using an input SLM. The input Fourier Transform is then optically multiplied with the reference filter and the result optically Fourier transformed to produce the correlation. The advantages of this method are a high space-bandwidth product and extremely fast process time.

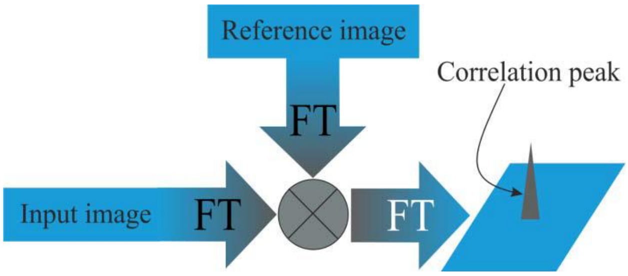


Fig. 1 Matched Filter Correlator [17]

Joint Transform Correlator

The Joint Transform Correlator was invented by Weaver and Goodman in 1966 and is traditionally the less commonly used of the two designs. The process differs significantly from the Matched Filter process, since no reference filter is required.

JTC use "Input" and "Reference" images aligned and displayed alongside each other on SLM as is shown in Figure 2. These before single images are now Fourier Transformed together. A non – linearity (mostly a camera or photodiode array) then captures the intensity distribution of transform to produce Joint Power Spectrum (JPS).

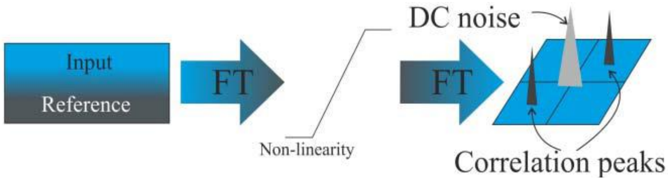


Fig. 2 Joint Transform Correlator [17]

JPS is then binary or threshold processed and this processed image enters to transform process as input image of second Fourier Transform. Output of this transform is correlation plane include correlations peaks per match. This found match is shown as highly localized intensities and the intensity of peaks provides a measure of similarity of the images being compared whilst position of spots denote how the images are relative aligned in the input scene.

The most advantages of this method of correlation are the simplified optical train (and therefore much cheaper to produce) and no strict alignment criteria [3-7, 17, 18].

B. Optical Fourier Transform

A better understanding the principles of the optical Fourier transform is that a simple lens can perform a Fourier transform in real-time as follows (see Fig. 3). Place an image, for example a slide transparency, at the focal length of the lens, and illuminate that slide with coherent light, like a collimated laser beam. At the other focus of the lens place a screen. The lens will automatically perform a Fourier transform on the input image, and project it onto the screen. For example if the input image is a sinusoidal grating, as shown below, the resultant Fourier image will have a bright spot at the centre, the direct current term, with two flanking peaks on either side, whose distance from the centre will vary with the spatial frequency of the sinusoid.

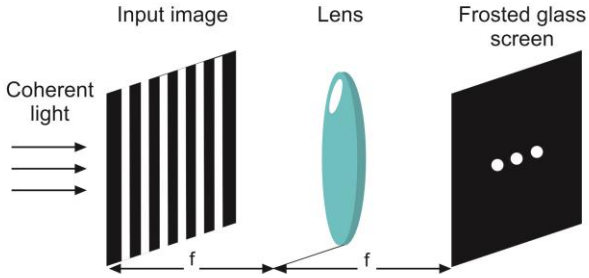


Fig. 3 Optic Fourier transform

The holistic principle behind the optical Fourier transform is that every point on the input image radiates an expanding cone of rays towards the lens, but since the image is at the focus of the lens, those rays will be refracted into a parallel beam that illuminates the entire image at the ground-glass screen. In other words, every point of the input image is spread uniformly over the Fourier image, where constructive and destructive interference will automatically produce the proper Fourier representation (Fig. 4 (a)).

Conversely, parallel rays from the entire input image are focused onto the single central point of the Fourier image, where it defines the central direct current term by the average brightness of the input image (Fig. 4 (b)).

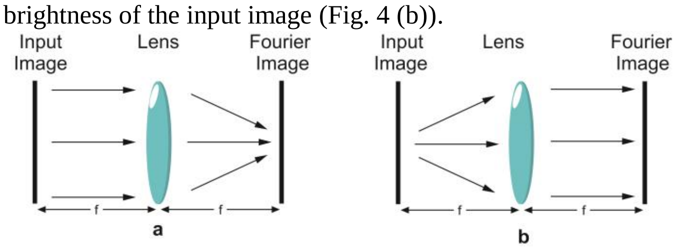


Fig. 4 Direct (a) and inverse (b) Fourier transform

The optical Fourier transformer automatically operates in the reverse direction also, where it performs an inverse Fourier transform, converting the Fourier representation back into a spatial brightness image. Mathematically the forward and inverse transforms are identical except for a minus sign that reverses the direction of the computation [8-12, 17, 18].

III. PRE-PROCESSING OF IMAGE

A. Basic Colour Models

Colour images can be modelled as three-band monochrome image data where each band of data corresponds to a different colour. The actual information stored in the digital image is the brightness information in each spectral band. Ehen the image is displayed on the screen by picture elements that emit light energy corresponding to that particular colour.

Typical colour images are represented as red, green and blue, or RGB images. Using 8-bit monochrome standard as a model, corresponding colour image would have 24 bits/pixel – 8 for each of the three colour bands. RGB colour model is shown in Fig 5.

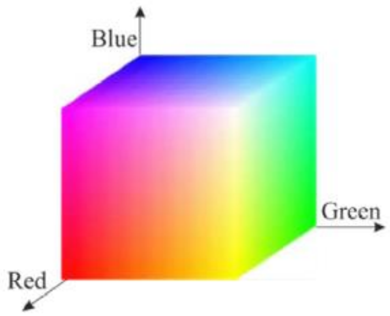


Fig. 5 RGB colour model [16]

The final colour depends on combination of these colour. As it is known, every byte can encode values in range from 0 to 255, so we can create 16777216 various colours by combination of three basic colour values.

In many applications, RGB colour information is transforms into mathematical space that decouples the brightness, or luminance space and a two-dimensional colour space. Now the two-dimensional colour space does not contain any brightness information, but it typically contains information regarding the relative amounts of the different colours. An additional benefit of modelling the colour information is this manner is that it creates a more people-oriented way of describing the colours.

For example, the Hue/Saturation/Value (HSV) colour transform allows the description of colours in terms that we can more readily understand.

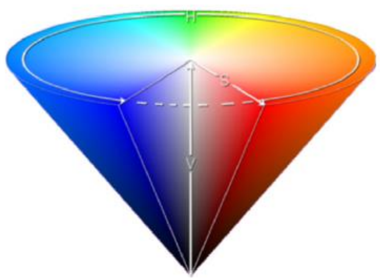


Fig. 6 HSV colour model [16]

The Value represents quantity of white light or specifies the rate of light reflectivity. The Hue is what we normally thong of as colour. It can be in range from 0 to 360 degrees. The value of Saturation labels quantity of grey colour in proportion to basic hue. It can be in range from 0 to 100% [13-16].

B. *Edge Detection*

Many of edge line detection operators are implemented with convolution masks, and most are based on discrete approximations to differential operators. A large change in image brightness over a short spatial distance indicates the presence of an edge. Some edge detection operators return orientation information, whereas others only return information about the existence of an edge at each point. Edge detection methods are used as a first step in the line detection process. Edge detection is used to find complex object boundaries by marking potential edge points corresponding to place in an image where rapid changes in brightness occur. After these edge points have been marked, they can be merged to form lines and object outlines. Edge detection operators are based on the idea that edge information in an image is found by looking at the relationship a pixel has with its neighbours. If a pixel´s grey-level value is similar to those around it, there is probably no an edge at that point. However, if a pixel has neighbours with widely varying grey levels, it may represent an edge point.

The Roberts operator marks edge points only; it does not return any information about the edge orientation. It is the simplest of the edge detection operators and will work best with binary images. The Sobel edge detection masks look for edges in both the horizontal and vertical directions and then combine this information into a single matric. The Prewitt is similar to the Sobel, but with different mask coefficients. The Kirsch edge detection masks are called compass masks because they are defined by taking a single mask and rotating it to the eight major compass orientations. The Hough transform is used to find lines. Line is a collection of edge points that are adjacent and have the same direction. The Hough transform is an algorithm that will take a collection of edge points and find all the lines on which these edge points lie. The primary advantage of the Hough transform is that it reduces the search time for finding lines [13].

IV. FURTHER RESEARCH FIELD

The correlator technology and concepts can be used in a wide range of application areas, which fall into either or both of two categories, namely recognition/validation/ inspection and motion tracking. It is of particular use in applications involving extremely large amounts of data processing. Below is a summary of application areas optical correlator is either currently involved with or has interest in. Other fields include Database Searching, Component Monitoring, CCTV Real Time Analysis, Mass Spectrometry, Automated Vision Systems [3, 17-20].

A. *Optical High Performance Computing Architecture*

By applying the principles of Fourier optics to solving Partial Differential Equations (PDEs), optical correlator is developing a technology which addresses the fundamental limitations of serial electronic processing. The technology is of particular significance to High Performance Computing (HPC) / supercomputer architectures used to solve PDEs in the field of Computational Fluid Dynamics (CFD), which is considered of particular relevance as its demands for increases in data resolution and processing speed cannot be met with the current FFT-based approach.

B. *Traffic Signs Recognition*

Object recognition is currently more and more popular. There are many applications, where the video-based object recognition helps to keep life safe, such as image retrieval, surveillance systems, driver assistance system etc. Each of this system has its own specific requirements for input data (image or video). To extract relevant features from input data, input images or videos, have to be input pre-processed. Correct extracted features are very important for next processing.

Recognition process is based on the cross-correlation matching between ROI extracted form traffic scene and reference traffic sign from reference database. Systems use the optical correlator for this cross-correlation process. Optical correlation enables the rapid identification of targeted sign within an input traffic scene.

C. Medical Analysis

With the NHS plans for screening all over 50's for diseases coming into force over the next few years, the need for high data processing systems is set to become ever more apparent, making the optical correlator a highly attractive technology.

It is still commonplace for teams of workers to examine high-resolution test data to pick out cancerous or abnormal cells (for example Pap smears). The optical correlator is looking into how the correlator may be put to use to diminish the errors and missed warning signs in a variety of medical applications.

D. Biometric Recognition

The ability of the correlator to examine and pick out matches from cluttered or partial data creates an exciting opportunity in the rapidly expanding field of biometric recognition. Full graphical comparisons of captured data are possible in a field where the norm is to look for specific features.

Fingerprint recognition relies upon identifying enough minutiae points from the sample to guarantee a match. However, many people do not have any such minutiae points and cannot therefore be included in normal search process.

In Facial recognition, most processes rely on calculating the distance between the eyes and nose, or on low-resolution faces. However, these rely on capturing a large proportion of the face and suffer from changes in appearance or expression.

V. Conclusion

This paper contains basic information about optical correlators, various types of an optical correlator and principle of operation. Technical University of Košice, department of Electronics and Multimedia Communications, optoelectronic system laboratories has Cambridge optical correlator. My research includes working on applications based on Cambridge optical correlator in image and video processing systems.

Fig. 6 represents detailed block diagram of a proposed system – Vertical Traffic Signs Inventory System. This system contains five blocks. Blocks Database, Pre-processing and Traffic Sign Identification will by designed in C# language. Database contains all necessary information about traffic signs. Pre-processing block includes colour filtering, edge detection, shape detection, extracting a region of interest and as well as other partial functionalities.

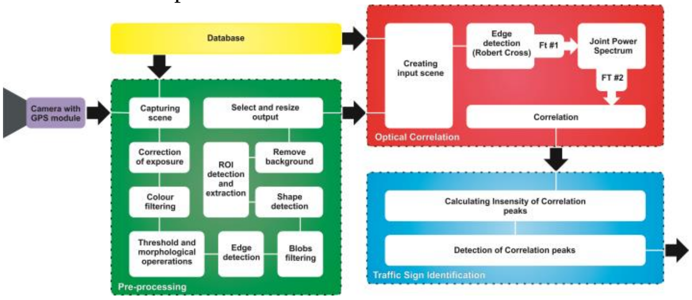


Fig. 7 Block diagram proposed system

In my feature work I will consider possibilities of using an optical correlator in biochemical industry specifically recognition DNA images and optical Fourier transform to solve differential equations.

Acknowledgment

This work was supported by Cultural and Educational Grant Agency (KEGA) of the Ministry of Education, Science, Research and Sport of the Slovak Republic under the project no. „059TUKE-4/2015“ and project no. „006TUKE-4/2014”. This work is also the result of the contract no. "APVV-0025-12" supported by the Slovak Research and Development Agency.

References

[1] Optical Correlator, http://www.timbercon.com/Optical-Correlator.html, February 2015.
[2] Cambridge Correlators, "Fourier Optic Experimenter," User Guide.
[3] S. Lehar, "An Intuitive Explanation of Fourier Theory," http://cns-alumni.bu.edu/~slehar/fourier/fourier.html, February 2015.
[4] J. W. Goodman– GRAY, R. M.: Fourier Transform: An Introduction for Engineers, Kluwer Academic Publishers, Norwell, MA, 1995.
[5] P.Ambs, "Optical Computing, A 60-Year Adventure," in Advances in Optical Technologies Journal, 2010.
[6] T. Naughton, M. Klíma, J. Rott, "Improved Joint Transform Correlation Performance Through Spectral Domain Thresholding," Maynooth,Ireland, 1998, <http://www.cs.nuim.ie/~tnaughton/pubs/varasto/tjn199809oesijtc.pdf>
[7] B. E. A. Saleh and M.C. Teich, Fundamentals of photonics, Wiley - Interscience, 2007.
[8] Základy Fourierovské optiky, http://www.vutbr.cz/www_base/zav_prace_soubor_verejne.php?file_id =7161, February 2015.
[9] Fyzikální praktikum III, Studijní materiály. Základy fourierovské optiky, http://physics.mff.cuni.cz/vyuka/zfp/, February 2015.
[10] Ľ. Ovseník, A. Kažimírová Kolesárová, J. Turán, "Video Surveillance Systems", In: Acta Electrotechnica et Informatica, Košice : FEI TU, Vol. 10, No. 4, 2010, pp. 46-53.
[11] T. Harasthy, Ľ. Ovseník and J. Turán, "Current summary of the practical using of optical correlators," Acta Electrotechnica et Informatica, Vol. 12, No. 4, 30-38, DOI: 10.2478/v10198-012-0042-2, 2012.
[12] S. E. Umbauch, "Computer Vision and Image Processing and A practical Approach Using CVIPtools," Prentice Hall PTR, Upper Saddle River, NJ 07458, June 1999.
[13] Q. Ye, W. Gao and W. Zeng, "Color Image Segmentation Using Density-Based Clustering," ICASSP 2013, China 2003.
[14] L. Luo, X. Li, "A method to Search for Segmentation Threshold in Traffic Sign Detection," Fifth International Conference on Image and Graphics, 2009.
[15] Ch. Poyton, "A Guide Tour of Color Space," John Wiley & Sons, New York, 1996.
[16] B. V. K. V. Kumar, A. Mahalanobis, R. D. Juday, "Correlation Pattern Recognition," Cambridge University Press, ISBN 9780521571036, 2005.
[17] T. Harasthy, Ľ. Ovseník and J. Turán,"Traffic Sign Recognition System based on Cambridge Correlator Image Comparator," Carpathian Journal of Electronic and Computer Engineering. Vol. 5, no. 1 (2012), p. 127-132. – ISSN 1844 – 9689.
[18] T. Harasthy, Ľ. Ovseník, J. Turán,"Cambridge correlator in driver assistance systems," Infocommunications Journal. Vol. 4, no. 3 (2012), s. 15-21. – ISSN 2061-2079.
[19] T. Harasthy, Ľ. Ovseník, J. Turán,"Traffic Signs Recognition with Using Optical Correlator," IWSSIP 201, 18th International Conference on Systems, Signals and Image Processing, Sarajevo, Bosnia and Herzegovina. - Sarajevo: University Sarajevo, Faculty of Electrical Engineering, 2011 P. 239-242. - ISBN 987-9958-9966-1-0
[20] F.P. Paulo and L.P. Correia, "Automatic detection and classification of traffic signs," Image Analysis for Multimedia Interactive Services, WIAMIS, 6-8 June 2007.

Optical Safety System of a Robotic Manipulator Workspace

[1]Marek VACEK (4th year)
Supervisor: [2]Jaroslava ŽILKOVÁ

[1,2]Dept. of Electrical Engineering and Mechatronics, FEI TU of Košice, Slovak Republic

[1]marek.vacek@tuke.sk, [2]jaroslava.zilkova@tuke.sk

***Abstract*—This summative article describes the process of modeling, control and realization of an optical safety system which will be used in a robotic workspace in which the robotic arm will be cooperating with a human.**

***Keywords*—Seriallink Robot Manipulator, Robotics Toolbox, Safety System, MATLAB**

I. Introduction

This article describes the design and realization of a safety system of a robotic workspace capable of working in human proximity or in cooperation with a human. The most suitable solution is to use a camera as a permanent workspace safety monitoring sensor system.

The best solution, using only one camera, is to place the camera directly above the robotic workspace facing downwards and monitor the movement of both the human and the robot at the same time. This system could be supplemented by a warning system, which lets the user know where the robot's working area lies. A projector is used for adaptive placement of virtual barriers. This projector can project the robot's working area, which changes according to the robot's current position, onto the workspace.

II. Prerequisites

The first important decision when designing the robotic workspace pertains to the size of the robot. If the robot should work in the immediate proximity of a person, it is suitable to use a smaller model of an experimental robotic arm - the bigger and heavier the robot, the greater the probability of injury upon contact.

The possibility of using a safety system also depends on the robot's size. If the camera is to be mounted above the robotic workspace and to be able to monitor all of it, bigger robots would make it difficult to meet such a requirement.

Furthermore, as the robotic workplace is to be used primarily in teaching, the following criteria were significant:

- low costs
- small size
- easy manipulation
- openness of the whole system

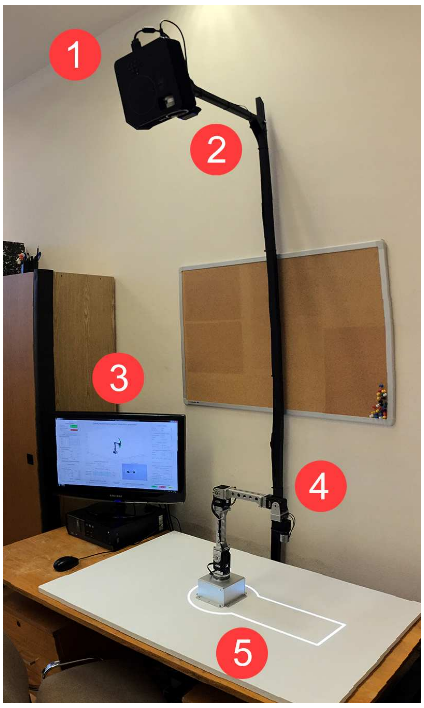


Fig. 1. Experimental robotic manipulator workspace with an optical safety system

1. Projector
2. Camera
3. Graphical user interface control panel
4. Experimental robotic arm
5. Generated optical safety barrier of the robotic arm

III. Workspace Operation

This whole safety system (Fig. 1.) works as follows:

The camera recording the movement of the human and robot is, with the projector, placed above the robotic workspace and monitors the safety during the entire operation.

The projector sends an image signal about the boundaries of the robot's working area. The boundaries of the danger zone are generated according to the current position of the robotic arm.

The camera records any breach of the safety zones. When they are breached, the control system stops the robot to prevent any human injuries.

The control system is composed of a computer with a monitor, to and from which all the workspace's control signals are sent.

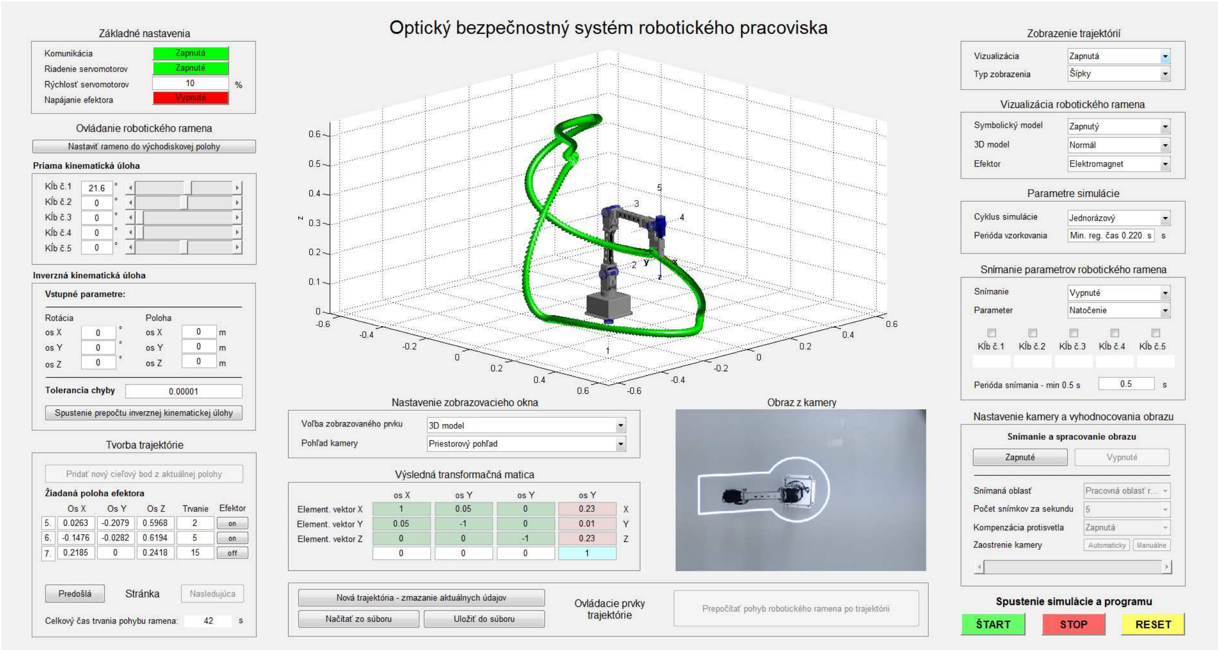


Fig. 2. Control panels of the graphical user interface, which enable the control of the experimental robotic arm's movement and of the robotic workplace's optical safety system.

The whole control system was created in Matlab (Fig. 2.). The graphical and control interface is always visible on the computer screen and it has all the functions needed to control the robotic arm's movement and the workspace's optical safety system.

The user interface can be divided into several subgroups:

- control of the communication between the computer and the robotic arm - it is possible to switch the communication port and control packet transfer on and off. The user can also verify the effector's functionality by requesting that control packets be sent.

- calculation of forward and inverse kinematics with the possibility of entering the position value of individual joints, using numerical values or sliders. The user defined position of individual joints is automatically calculated into a transformation matrix of position, and the 3D robot model in the defined position is displayed in one of the graphical windows.

- creation, saving, loading, and calculation of the robotic arm's trajectories with the possibility of setting the duration of each movement, activating the effector, choosing the packet sending period, and, of course, controlling its operation in one cycle or a repeated cycle.

- possibility to monitor the robotic arm's parameters either just when the arm is moving, or during the whole simulation. The user can choose, which parameter's value they want to read (position, velocity, load etc.) and on which actuator (either individually, or all actuators at once). After the movement or the simulation ends, it is possible to display the results in graphs in the viewing window (the required value of the actuator, the final value of the actuator, or their difference).

- change of what the viewing window shows - the robotic arm (3D model, front view, top view, and side view), its model (full, symbolic, transparent, wire frame etc.), or the parameter graphs.

- change of the camera's parameters with the possibility to change the frame rate, the resolution, the recorded area, and backlight illumination compensation.

IV. Use and Conclusion

The created safety system can be used to test newly-developed robot programs, to verify the principles of control of the robotic arm's actuators, to calculate the robotic arm's trajectories, to process the camera feed, to generate image on the projector, to generate the robotic arm's safety zones, to evaluate the breach of the robotic arm's safety zones, and to monitor the overall safety of the workspace, which is intended for mutual robot-human cooperation.

This system can be further developed to optimalize data processing speeds to improve the reaction times when stopping the robotic manipulator and to speed up the reception, processing, and sending of the input and output data of all used devices.

Acknowledgment

The work has been supported by project KEGA 011TUKE-4/2013.

References

[1] Reza N. Jazar: Theory of Applied Robotics. Springer New York Dordrecht Heidelberg London 2007. 893 s. ISBN 978-1-4419-1749-2.

[2] Corke, Peter: Robotics, Vision and Control. Springer-Verlag Berlin Heidelberg 2011. 570 s. ISBN 978-3-642-20143-1.

[3] Corke, Peter: Robotics Toolbox for MATLAB, Release 9. Available on the Internet: <http://www.petercorke.com/RTB/robot.pdf>.

[4] Mathworks MATLAB: R2012a Documentation [online]. Available on the Internet: <http://www.mathworks.com/help/techdoc/>.

[5] KULIĆ,D.: Safety for Human-Robot Interaction, The University of British Columbia, Vancouver, 2005.

[6] VACEK,M.: "Optický bezpečnostný system robotického pracoviska", Technical university of Košice, Faculty of Electrical Engineering and Informatics, Postgraduate Thesis, 2013.

Optimization the operation of renewable energy sources in electric power system

[1]Martin VOJTEK (1st year)
Supervisor: [2]Michal KOLCUN

[1,2]Dept. of Electric Power Engineering, FEI TU of Košice, Slovak Republic

[1]martin.vojtek@tuke.sk, [2]michal.kolcun@tuke.sk

***Abstract*—Proportion of electricity produced from renewable energy sources records tremendous growth nowadays. This is happening in accordance with strategies, which aim on decrease the production of greenhouse gases and carbon emissions. But, on the other hand it has resulted in many changes and related problems in electric power system especially with its optimal operation control. It can be solved by accumulation of electricity, which can provide better integration of renewable energy sources due to its internment power supply and it also helps to prepare conventional distribution grids to new smart grid concept. In the light of these facts, this is a big challenge and it is necessary to create new approaches to the management and control of renewable sources, energy storages, and also to optimal operation of whole grids.**

***Keywords*—renewable energy sources, energy storage, micro grid, smart grid, optimal operation control.**

I. Introduction

For a number of years, motivated by economics, decentralization policies and sustainability concerns, the trend in the power industry worldwide is to reduce investment and operation costs and to become environmentally friendly. This has resulted in an increased share of renewable energy (such as wind and solar) in the total generation capacity, and an increased number of small and medium sized geographically distributed power plants (such as wind turbines, photovoltaic arrays and district combined heat and power plants) [1].

These types of sources are characteristic with unpredictable and intermittent production of electricity and this is a major problem. In order to provide a power quality and reliability, this unstableness and balancing output power must be firmed in certain way. The variable, intermittent power output from a renewable power plant, such as wind or solar, can be maintained at a committed (firm) level for a period of time using some type of energy storage technology. The energy storage systems can smooth the output power, control the ramp rate (MW/min) to eliminate rapid voltage and power swings on the electrical grid what improve power quality. In dependence on technology it can also provide frequency regulation [2].

All of this causes a rapidly changing network topology and in order to this some new approaches are needed. This topological change in the power system landscape is opening up possibilities to form a new concept called microgrid. Microgrids, also characterized as the “building blocks of smart grids”, are perhaps the most promising, novel network structure.

In summary, distribution grids are being transformed from passive to active networks, in the sense that decision-making and control are distributed, and power flows bidirectional. This type of network eases the integration of DG, RES, demand side integration (DSI) and energy storage technologies, and creates opportunities for novel types of equipment and services, all of which would need to conform to common protocols and standards [3].

II. Microgrids concept

Microgrids are defined as localized groups, which integrate distributed generators, storages systems such as battery and also local loads. These devices are coupled together and act as autonomous power systems with a single point of common coupling to the main electricity network [1][4][5].

Microgrids meet the power quality and reliability requirements of the local customers [6], and also shield itself from issues such as voltage distortion, voltage sag, flicker, and lightning transients [7]. Apart from the technical benefits, microgrids have also economic advantages [8] if appropriate optimal control and management systems are implemented.

To economically operate the microgrid, the central controller will dispatch controllable distributed energy resources including batteries [4]. A microgrid can either operate at the grid connected or autonomous modes [6], [7] when the grid is suffering blackouts. At autonomous modes, voltage and frequency should be supported by a microgrid itself, usually through synchronous generators.

For a microgrid without synchronous generators, the system voltage and frequency would be difficult to maintain without the support of the ac grid. One solution is to use a voltage source converter (VSC) interfaced energy sources to provide voltage and frequency control [9]. In [9]–[11], battery systems are employed to restore system voltage and frequency quickly (several cycles). In practice, applications of battery storage system for grid frequency regulation have been deployed [12] with the maximum capacity of 20 MW.

As was mention before, microgrid consists of some basic parts, which must be operated by intelligent control. Distributed energy sources based on renewable energy, energy storages and loads are described below.

A. *Distributed energy sources based on renewable energy*

During the past decades, the deployment of distributed generation has been growing steadily. Distributed generators are connected typically at distribution networks, mainly at medium voltage and high voltage level, and these have been designed under the paradigm that consumer loads are passive and power flows only from the substations to the consumers and not in the opposite direction [3]. For this reason, many studies on the interconnection of distributed generators within distribution networks have been carried out, ranging from control and protection to voltage stability and power quality [13].

Different microgeneration technologies, such as micro-turbines, photovoltaics, fuel cells and wind turbines with a rated power ranging up to 100 kW can be directly connected to the LV networks. These units, typically located at users' sites, have emerged as a promising option to meet growing customer needs for electric power with an emphasis on reliability and power quality, providing different economic, environmental and technical benefits. Clearly, a change of interconnection philosophy is needed to achieve optimal integration of such units [3].

Most importantly, it has to be recognized that with increased levels of microgeneration penetration, the LV distribution network can no longer be considered as a passive appendage to the transmission network. On the contrary, the impact of micro sources on power balance and grid frequency may become much more significant over the years.

Controllability of intermittent renewable sources units is limited by the physical nature of the primary energy source [6]. Moreover, limiting renewable energy sources production is clearly undesirable due to the high investment and low operating costs of these units and their environmental benefits over carbon emission. Consequently, it is generally not advisable to curtail intermittent renewable energy sources units, unless they cause line overloads or overvoltage problems [3][6].

The operation strategy for intermittent renewable energy sources units can therefore be described as "priority dispatch". It means that intermittent renewable energy sources units are generally excluded from the unit commitment schedule, as long as they do not violate system constraints. Units with independent reactive power interfaces (decoupled from the active power output) can be included in reactive power dispatch to improve the technical performance of the total microgrid [3][14].

B. *Energy storage systems*

Accumulation or energy storage systems are irreplaceable in terms of production and consumption of electricity. Electric energy is very problematic commodity in principle and any efforts to its distribution face the problem of immediate supply and demand. Energy storages are still in development nowadays. However, there are some options how to store electricity with relatively good efficiency [15].

Energy storage systems can be divided into primary and secondary. Primary storages may supply immediately, secondary after polarization of electrodes (initially must be charged). Energy storages can be also divided into many categories, for example in dependence by the output power, accumulation capacity, number of cycles and others. Technological division is described by the block diagram (Fig. 1.).

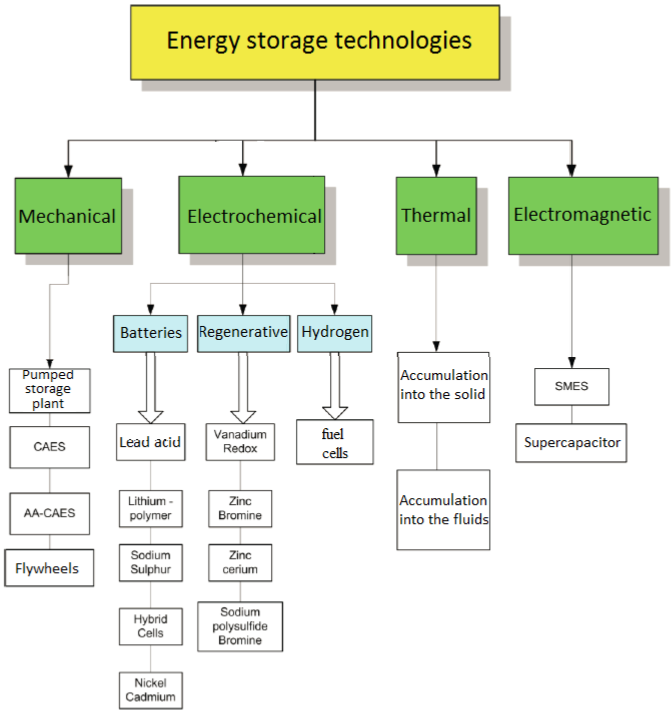


Fig. 1 Block diagram of energy storage technologies

The energy storage device can be installed for each micro-source, or as a shared resource for more microsources, or in some cases one storage devices for the entire microgrid. The criteria to decide the ratio of energy storage device to microsource and also the size of energy storage device is mainly based on the microgrid characteristics, especially the dynamic response of the microsources and the power quality required by the loads [13] [15].

The dynamic response of microsources cannot be seen independently to design energy storage devices, but it should be considered relative to the rate of change of the loads. For instances, a slow-response microsource will just need a small size energy storage device or no energy storage device at all if the microgrid loads have slow rate of change. However, it will require a big size energy storage if the microgrid loads have fast rate of change. Similarly, a fast-response microsource depending on the rate of change of the microgrid loads just requires small size energy storage or no energy storage at all. The idea here is to balance the dynamic response of the microsource and the rate of change of loads in order to have smooth load following or load-generation balancing in transient [16].

Technically, a storage unit could behave either under a load-following paradigm (i.e. balancing applications) or under a price-following paradigm (i.e. arbitrager applications) depending on the purpose of its operation. At the same time, storage units can provide balancing reserves ranging from short-term (milliseconds to minute-level) to long-term (hourly to daily scale) applications. Specifically, for DC-based storage technologies (battery, super-capacitor etc.), a properly designed power electronic interface could contribute to the reactive power balance of the system without incurring significant operational costs [3][15][16].

C. *Loads*

Demand side integration is also referred to as demand side management (DSM) or demand side response (DSR). It is based on the concept [17] that customers are able to choose from a range of products that suit their preferences. The innovative products packaged by energy suppliers will deliver – provided that end-user price regulation is removed powerful messages to consumers about the value of shifting their electricity consumption. Examples of such offers include [16]:

1. time-of-use - higher "on-peak" prices during daytime hours and lower "off-peak" prices during the night and at weekends (already offered in some EU member states)
2. dynamic pricing (including real-time pricing) - prices fluctuate to reflect changes in the wholesale prices
3. critical peak prices - same rate structure as for Time of use, but with much higher prices when wholesale electricity prices are high or system reliability is compromised.

The control of customers load can either be [3][16][17][18]:

1. manual - customers are informed about prices, for example on a display, and decide on their own to shift their consumption, perhaps remotely through a mobile phone
2. automated - customers' consumption is shifted automatically through automated appliances, which can be pre-programmed and can be activated by either technical or price signals (as agreed for instance in the supply contract).

Demand side integration measures in a microgrid are based on forecasts of load and source outputs and will very probably vary from day to day. A requirement for the successful application of microgrid DSI measures is the adoption of smart metering and smart control of household, commercial and agricultural loads within the microgrid. Depending on the criticality of the target load, DSI measures can generally be divided into shiftable loads and interruptible loads. The integration of DSI measures is expected to maximize their benefits in potential "smart homes", "smart offices" and "smart farms" within microgrids [3][18].

III. Microgrids Control Functions

Main control functionalities of microgrid are presented in this section. These functionalities can be distinguished in three groups, as shown in Fig. 2. The lower level is closely related to the individual components and local control (micro sources, storage, loads and electronic interfaces), the medium level to the overall microgrid control and the upper level to the interface to the upstream network.

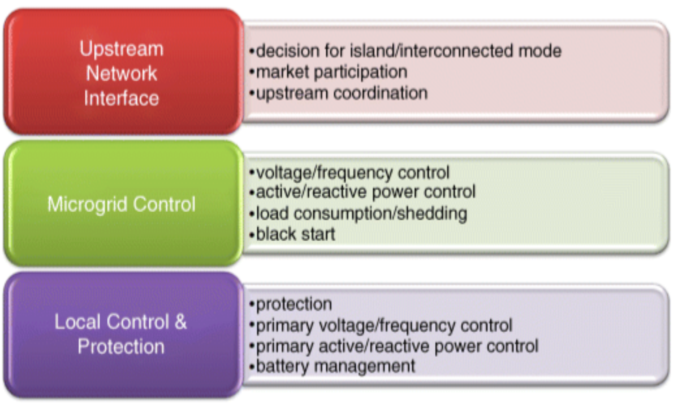


Fig. 2 Microgrid control functionalities [3]

A. Upstream network interface

The core interaction with the upstream network is related to market participation, more specifically the microgrid actions to import or export energy following the decisions of the energy service provider/company. Owing to the relatively small size of a microgrid, the service provider can manage a larger number of microgrids, in order to maximize its profit and provide ancillary services to the upstream network.

B. Internal microgrid control

This level includes all the functionalities within the microgrid that require the collaboration of more than two actors. Functions within this level are:

- load and renewable energy sources forecast,
- load shedding/management,
- unit commitment/dispatch,
- secondary voltage/frequency control,
- secondary active/reactive power control,
- security monitoring,
- black start.

C. Local control

This level includes all the functionalities that are local and performed by a single DG, storage or controllable load, that is:

- protection functions,
- primary voltage/frequency control,
- primary active/reactive power control,
- battery management.

It should be noted that these functionalities are relevant to the normal state of operation. They might need to change in critical or emergency states. The normal state covers both islanded and interconnected mode and does not deal with transition to island mode. The role of information and communication technology is critical for the relevant control functions [3].

IV. Current Status and Future Direction of Microgrid Research in the World

Current research in planning and designing for microgrid mainly focuses on the following aspects: 1) Models and algorithms. 2) Optimal configuration for energy storage system. 3) Distribution network planning containing microgrid. 4) Software development [14]. For example, Basu, A. K. et al. [19] proposed an optimization method using differential evolution technique under real power demand equality constraint, heat balance inequality constraint, and distributed generators capacity limits constraint at the planning of combined heat and power based microgrid, Chen, S. X. et al. [20] presented a new method for optimal sizing of an energy storage system in a microgrid for storing renewable energy at the time of surplus and for rescheduling, Kirthiga, M. V. et al. [21] proposed an approach transforming an existing radial distribution network into an autonomous microgrid using novel sizing and siting strategies for distributed generators and structural modifications for autonomous microgrids, the Distributed Energy Resources Customer Adoption Model (DER-CAM) developed at Berkeley Laboratory is a fully technology-neutral optimizing model of economic DER adoption. Its objective is to minimize the operating cost of on-site generation and CHP systems, for either an individual customer site or a microgrid, and Hybrid Optimization Model for Electric Renewable (HOMER) developed at the National Renewable Energy Laboratory (NREL), USA, designs the most optimized and cost effective hybrid generation system component configuration after a significant amount of simulations for offgrid or grid-connected power systems.

Numbers of research work have been reported related to energy management and control of microgrid with renewable energy and energy storage [22]. Hierarchal control was addressed in [22] and planning based operation discussed in [23]. Intelligent approaches have been presented in [24],[25].Two-layered control architecture for stabilizing microgrid in islanded mode has been discussed in [26] and generation capacity design of microgrid to maintain power quality and energy surety in both grid-connected and islanded modes has been discussed in [27].

As was mentioned above, microgrid has three levels of control functionalities. My future research aims to the lowest control level named Local control and protection, especially on battery or other technology independent energy storage management. This includes optimal operational control in order to decrease environmental impact and operational costs, and on the other hand to increase reliability and power quality. I estimate that to achieve these goals, it will be needed to create a computer model of microgrid including renewable energy sources, energy storages, and loads. After creating a model some optimization methods or algorithm will be proposed and result will be presented.

Acknowledgment

This work was supported by Slovak Research Agency No. VEGA 1/0388/13 project.

References

[1] Mahat, P.; Escribano Jimenez, J.; Rodriguez Moldes, E.; Haug, S.I.; Szczesny, I.G.; Pollestad, K.E.; Totu, L.C., "A micro-grid battery storage management" *Power and Energy Society General Meeting , 2013 IEEE* , pp.1,5, 21-25 July 2013

[2] ABB, "Keeping smart grid in balance" brochure, http://www.abb-energystoragesolutions.com/

[3] Nikos Hatziargyriou, "Microgrids control and architecture," National Technical University of Athens, Greece,

[4] Zhixin Miao; Ling Xu; Disfani, V.R.; Lingling Fan, "An SOC-Based Battery Management System for Microgrids," *Smart Grid, IEEE Transactions* , vol.5, no.2, pp.966,973, March 2014

[5] F.Peng, Y.W.Li, L. Tolbert, "Control and protection of power electronics interfaced distributed generation systems in a customer-driven microgrid," *IEEE Power Energy Soc. Gen. Meet.* , pp. 1–8, Jul 2009

[6] C. Marnay, F. Rubio, and A. Siddiqui, "Shape of the microgrid," *IEEE PES Winter Meeting*, Columbus, Ohio, USA, Jan. 2001, pp. 150–153.

[7] IEEE guide for design, operation, and integration of distributed resource island systems with electric power systems, IEEE Standard 1547TM, July 2011.

[8] Z. Zhang, G. Li, and M. Zhou, "Application of microgrid in distributed generation together with the benefit research," *IEEE PES General Meeting*, Minnesota, USA, Jul. 2010, pp. 1–5.

[9] H. Karimi, H. Nikkhajoei, and R. Iravani, "Control of an electronically-coupled distributed resource unit subsequent to an islanding event," *IEEE Trans. Power Del.*, vol. 23, no. 1, pp. 493–501, Jan. 2008

[10] G. Li, M. Yin, M. Zhou, and C. Zhao, "Modeling of VSC-HVDC and control strategies for supplying both active and passive systems," *Proc. IEEE Power Eng. Soc. Gen. Meet.*, 2006.

[11] H. Chen, "Research on the control strategy of VSC based HVDC system supplying passive network," in Proc. *IEEE Power Energy Soc Gen. Meet.* , Jul. 2009, pp. 1–4

[12] G. L. Soloveichik, "Battery technologies for large-scale stationary energy storage," *Annu. Rev. Chem. Biomol. Eng.* , vol. 2, no. 1, pp. 503–527, 2011.

[13] Zavoda, F., "Advanced distribution automation (ADA) applications and power quality in Smart Grids," *Electricity Distribution (CICED), 2010 China International Conference*, pp.1,7, 13-16 Sept. 2010

[14] Zheng, Q.P.; Wang, J.; Liu, A.L., "Stochastic Optimization for Unit Commitment—A Review," Power Systems, IEEE Transactions, no.99, pp.1,12, 2014

[15] Mastný, P., "Obnovitelné zdroje elektrickej energie," ČVUT v Prahe, 2011. ISBN 978-80-01-04937-2

[16] Zamora, R.; Srivastava, A.K., "Energy management and control algorithms for integration of energy storage within microgrid," *Industrial Electronics (ISIE), 2014 IEEE 23rd International Symposium*, pp.1805,1810, 1-4 June 2014

[17] CIGRE WG C6.09. "Demand Side Integration", Technical Brochure, August 2010.

[18] Llanos, J.; Saez, D.; Palma-Behnke, R.; Nunez, A.; Jimenez-Estevez, G., "Load profile generator and load forecasting for a renewable based microgrid using Self Organizing Maps and neural networks," *Neural Networks (IJCNN), The 2012 International Joint Conference*, pp.1,8, 10-15 June 2012.

[19] Basu, A.K.; Bhattacharya, A.; Chowdhury, S.P.; Chowdhury, S.; Crossley, P.A., "Reliability study of a micro grid system with optimal sizing and placement of DER," *SmartGrids for Distribution*, 2008. IET-CIRED. CIRED Seminar , pp.1,4, 23-24 June 2008.

[20] Chen, S.X.; Gooi, H.B., "Sizing of energy storage system for microgrids," *Probabilistic Methods Applied to Power Systems (PMAPS), 2010 IEEE 11th International Conference*, pp.6,11, 14-17 June 2010

[21] Kirthiga, M.V.; Daniel, S.A.; Gurunathan, S., "A Methodology for Transforming an Existing Distribution Network Into a Sustainable Autonomous Micro-Grid," *Sustainable Energy, IEEE Transactions*, vol.4, no.1, pp.31,41, Jan. 2013

[22] J.M. Guerrero, J.C. Vasquez, J. Matas, L.G. de Vicuna, M.Castilla,"Hierarchical Control of Droop-Controlled AC and DC Microgrids — A General Approach Toward Standardization," *IEEE Trans. on Industrial Electronics*, vol. 58, no. 1, pp. 158 - 172 , Jan 2011.

[23] H. Kanchev, Di Lu, F. Colas, V. Lazarov, B. Francois, "Energy Management and Operational Planning of a Microgrid With a PV - Based Active Generator for Smart Grid Applications," *IEEE Trans. on Industrial Electronics*, vol. 58, no. 10, pp. 4583 - 4592 , Oct 2011.

[24] S.Chakraborty, M.D. Weiss, M. G.Simoes,"Distributed Intelligent Energy Management System for a Single - Phase High - Frequency ACMicrogrid," *IEEE Trans. on Industrial Electronics*, vol. 54, no. 1, pp. 97-109, Feb 2007.

[25] H.S.V.S. K.Nunna, S. Doolla, "Multiagent - Based Distributed Energy Resource Management for Intelligent Microgrids," *IEEE Trans. on Industrial Electronics*, vol. 60, no. 4, pp. 1678 - 1687, Apr 2013.

[26] Jong Yul Kim et al., "Cooperative Control Strategy of Energy Storage System and Microsources for Stabilizing the Microgrid during Islanded Operation," *IEEE Trans. Power Electronic*s, vol. 25, no. 12, pp. 3037 - 3048, Dec. 2010.

[27] Qiang Fu et al., "Microgrid Generation Capacity Design With Renewables and Energy Storage Addressing Power Quality and Surety," *IEEE Transition on Smart Grid*, vol. 3, no. 4, pp. 2019 2027, Dec. 2012

Price Regulation in Electric Power Engineering

[1]*Ing. Ružena Lovasová (5th year)*
Supervisor: [2] prof. Ing. Iraida Kolcunová, PhD.

[1,2]Dept. of Electric Power Engineering, FEI TU of Košice, Slovak Republic

[1]ruzena.lovasova@tuke.sk, [2]iraida.kolcunova@tuke.sk

***Abstract*— This paper refers to one of the network sector of the national economy in which the regulation is seen as essential. It draws attention to the fundamental laws essential for the pricing for electricity. The legislation implies a relatively extensive set of regulated prices. They are processed in the table and reflected in the price composition. The basic price structure is presented in short. The aim of the paper is to summarize a number of factors that affect the amount of expenses of the electric energy end-users.**

***Keywords*— electricity, regulatory policy, price regulation, price of electric energy, end-user.**

I. Introduction

Electric power engineering is a network industry with electric energy as a final product. Providing electric energy as goods is carried out by regulated activities – from the electric energy production trough transmission, operator duty of short-term market with electric energy, distribution to supply of electric energy. Conditions for implementation of those regulated activities are prescribed in the Law of 2012, No 250 of the Regulation in network industries.

The purpose of the regulation is to ensure that the electric energy as goods would be accessible to all market participants under the same conditions. The core subject-matter of regulation protecting the electric energy market against non-transparent and discriminatory approach of the participants is the price. The conditions determining the price, implementation and quality of electric energy supply are provided and elaborated in the regulatory policy [4][5][8][9][10].

II. The Regulatory Policy

The regulatory policy lays down the basis, principles and methods of regulation of the network industries. It is a strategic document adopted by the Council for a regulation. The Council for a regulation is the Regulatory authority for network industries. It is an independent collective state authority of the regulation strategy and management in the network industries. The Regulatory policy establishes the procedures for the Regulatory authority for network industries at the regulation performance in the network industries in relation to the general binding rules for the regulation of Slovak republic. The Council for a regulation has set the length of the regulatory period from the 1st January 2012 to the 31st December 2016. The length of the regulatory period is determined by the evaluation of the previous regulatory period to ensure the stability of the market environment for the regulated subjects [4] [5].

III. Price Regulation

„The price setters must know the laws regarding prices and be sure that their pricing policy is legal. Price regulation within the meaning of the prices Law is an instrument of the state macroeconomic policy to determine or direct guide price levels by control authorities." [2].

The Decree 221/2013 of the Regulatory authority for network industries is implementing rules for price regulation in the electric power engineering.

A calendar year is a regulatory year under the listed Decree. The price regulation, excluding precise determination of the sub-activities of security of electricity supply, divides the entity cost to economically justified and unjustified. It is a unique approach to the cost structure of the business entity. The electric power engineering price regulation thereby is protecting consumer of electricity against unjustified or disproportionate cost of calculation of the regulated entities. It provides a balance of prices in electric energy sales by the precise calculation of amount of reasonable profit [1].

The Decree determines directly or by calculation the following: [1][5][8]

- fixed price, that is not possible to modify,
- maximum price, that is not possible to exceed
- tariff – understood as the fixed price linked to the technical unit.

The method of price regulation implementation is summarized in the following table [8]:

Activities	Fixed price	Maximum price	tariff
Production of electricity from renewable energy sources	x		
Combined electric energy and heat production	x		
Electric energy production from indigenous coal	x		
Connection to the system		x	
Access to the transmission system and electric energy transmission		x	
Access to the distribution system and electric energy distribution		x	x
Providing system services		x	x
Providing support services		x	
Operator duty of short-term market with electric energy		x	x
Electric energy supply by last instance suplier		x	
Electric energy supply to vulnerable consumers		x	

Prices for activities stated in table are transferred into basic electric energy end price structure [1][8][10][11][12][13]:

a) partial prices for electric energy supply,

b) partial prices for access to distribution system and for electric energy distribution,

c) partial prices for losses, for system operation, for system services, for contribution to National Nuclear Fund,

d) tax charge by value added tax eventually by consumer tax.

If end user considers an electric energy supplier change, it is enough to compare only the price for supply because this item is defined by regulatory office as maximal price. This price can be different at each supplier but lower than maximal price. The price for distribution and other price components are stated by regulatory office as fixed prices, therefore they are not competitive among mutual electric energy suppliers [10].

Based on previous studies we can say that the electric energy end price consists of:

- Costs of all process participants beginning from electric energy generation until its supply,
- State expenses for electric energy generation from renewable sources, from domestic coal, for combined electric energy production and for termination of electric energy production in nuclear plants,
- Other state costs which charge end user's price taxes.

IV. CONCLUSION AND PROPOSAL FOR NEXT RESEARCH

The objective of this paper was to refer to a big quantity of influences and complexity respectively structure of electric energy price creation. It is only a brief description that is a summary of a whole range of factors having impact on usually monthly expenses of end users.

Therefore further research will focus on two points of view on this topic:

a) First the electric energy market participator's point of view:

Specification of reasons, influence factors, cost and calculation items of subjects entering the price creation process.

b) On the other side the end user's point of view:

Design a more understandable invoicing model.

ACKNOWLEDGMENT

The authors also wish to acknowledge Scientific Grant Agency of The Ministry of Slovak Republic and Slovak Academy of Science for funding of experimental works in the frame of VEGA No. 1/0132/15 grant.

REFERENCES

[1] Janíček F., Krondiak E., Kultan J., Šedivý J., Vechter J.,: Model trhu s elektrinou. Ekonomické aspekty výroby, prenosu a distribúcie elektriny v Slovenskej republike. Prvé vydanie, Renesans, 2009, str. 58 – 65, ISBN 978-80-89402-03-9

[2] Kádárová J., Janeková J.,: Ekonomika podniku. Prvé vydanie, Technická univerzita v Košiciach, Strojnícka fakulta, 2010, str. 161 – 169, ISBN 978-80-553-0542-4

[3] Lovasová R.: Plánovanie cenotvorby na teplo v tepelnej energetike. Technická univerzita v Košiciach, Fakulta elektrotechniky a informatiky, Katedra elektroenergetiky, písomná práca k dizertačnej skúške, 2010

[4] Regulačná politika na nastávajúce regulačné obdobie 2012 – 2016 v znení dodatkov č. 1 a 2., zverejnená na stránke Úradu pre reguláciu sieťových odvetví

[5] Zákon č.250/2012 Z.z. o regulácii v sieťových odvetviach v znení neskorších predpisov

[6] Zákon č. 251/2012 Z.z. o energetike v znení neskorších predpisov

[7] Zákon č. 309/2009 Z.z. o podpore obnoviteľných zdrojov energie a vysoko účinnej kombinovanej výroby v znení neskorších predpisov

[8] Vyhláška Úradu pre reguláciu sieťových odvetví 221/2013, ktorou sa ustanovuje cenová regulácia v elektroenergetike v znení neskorších predpisov

[9] Vyhláška Úradu pre reguláciu sieťových odvetví 24/2013, ktorou sa ustanovujú pravidlá pre fungovanie vnútorného trhu s elektrinou v znení neskorších predpisov

[10] www.urso.gov.sk, Usmernenie pre odberateľov elektriny v domácnosti, Sieťové poplatky a TPS

[11] Nariadenie vlády SR č. 426/2010 o výške odvodu z dodanej elektriny koncovým odberateľom a spôsobe jeho výberu pre Národný jadrový fond v znení neskorších predpisov

[12] Zákon č.609/2007 Z.z. o spotrebnej dani z elektriny v znení neskorších predpisov

[13] Zákon č. 222/2004Z.z. o dani z pridanej hodnoty v znení neskorších predpisov

Pulsed heating as a method of modifying the magnetic properties of amorphous material FINEMET - type

[1]*Lukáš HUBAČ (3th year)*
Supervisor: [2]Ladislav NOVÁK

[1,2]Deptartment of Physics, FEI TU of Košice, Slovak Republic

[1]lukas.hubac@tuke.sk, [2]ladislav.novak@tuke.sk

***Abstract*— In this paper is described another method for modification of the magnetic properties of amorphous material FINEMET-type. Specifically, it is a pulse annealing of this material in the shape of amorphous ribbons. Furthermore, the apparatus for measurement of the pulse of other physical quantities such as voltage, current, resistance, power, temperature is described. In the next step, we measured the magnetostriction of our sample.**

***Keywords* — amorphous ribbon - FINEMET, pulse heating, measurement of magnetic properties, magnetostriction.**

I. Introduction

Nano-crystallization of amorphous FINEMET creates materials, which have a wide range of applications due to their high magnetic softness. The disadvantage of these alloys is their brittleness. This makes them not suitable for application where shocks or vibrations are present, for example, in automotive or aviation industry. Therefore, our research is focused on improvement of their magnetic properties while keeping the amorphous state of the material. Exceptional magnetic softness (very low H_c), reduced hysteresis losses, modified shape of the magnetization curve can be achieved in these alloys using pulse annealing. Using this kind of treatment it is possible to preserve their mechanical properties, in contrast to the traditional heat treatment, which is used on a large scale in manufacturing nano-crystalline alloys [1, 2].

II. Pulse Source and Measurements of Sample Temperature During Pulse Annealing

Overview last year:

Determination of the temperature achieved during the single impulse was one of the principal tasks. The first idea was to place IR (Infra Red) sensing diodes directly on the sample holder in order to measure the temperature directly during the pulse. Since each IR semiconductor diode that we had at our disposal was made of parts made of ferromagnetic material, we decided to calibrate individual pulses. This was realized contactless with miniature IR pyrometer (CTF-SF15-C3 MICRO-EPSILON). The reason for contactless temperature measurement is a small heat capacity of the sample and a large heat capacity of scanning sensors with short reaction times. Heating rate during thermal treatment was about 3000 °C/s. Cooling was carried out exponentially with average cooling rate of 100 °C/s. Selected pulses the different lengths are depicted in Fig. 1a.

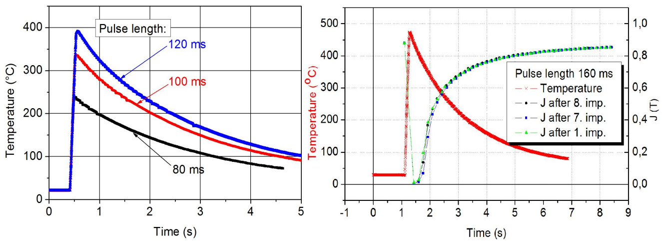


Fig. 1a, 1b. Graphic temperature dependence of time and progress of the magnetic polarization during the pulse. During the pulse we observed the transition from the ferromagnetic to the paramagnetic state and back.

Calibration accuracy was verified by measurement of Curie temperature (T_c). Fig. 1b illustrates magnetic polarization during the pulse. During the pulse we observed the transition from the ferromagnetic to the paramagnetic state and vice versa [3 -5].

In Fig. 2a, the dependence of cubed magnetic polarization on the temperature from which T_c can be accurately obtained, is shown.

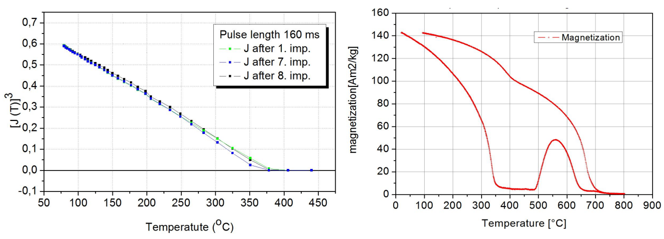


Fig. 2a, 2b. Dependence of cubed magnetic polarization on temperature and dependence of the magnetic polarization on temperature. From this graphs, the temperature T_c is determined.

Subsequently we verified the measured value of T_c using vibration magnetometer in the laboratory of ÚEF SAV in Košice. Output of this measurement is shown in Fig. 2b. From this graph, the temperature T_c is determined as the temperature of inflection point.

During pulse annealing a series of magnetic measurements were carried on FINEMET samples for several times. The results of these measurements were published in articles [3 - 5].

RESULTS ACHIEVED IN THIS YEAR:

Another task was to improve the apparatus for measuring of other physical quantities during the pulse. These were the voltage on the sample, the current flowing through of sample and the relevant temperature for the given number of periods. From these of quantities were subsequently calculated additional physical quantities such as resistance, power and total supplied energy. The reason for this measurement was to better define the impulse. The following experimental set-up (see Fig. 3) depicts the apparatus for impulse heating and its connection with the existing the apparatus for measuring the magnetic properties [6, 7].

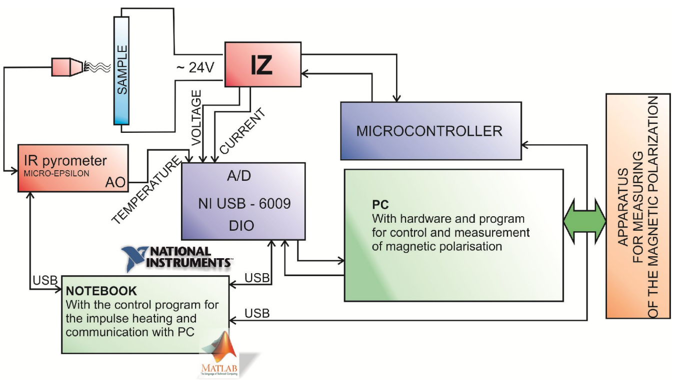


Fig. 3. Experimental set-up for impulse heating.

In the Fig. 4 is plotted progress of single parameters during the impulse.

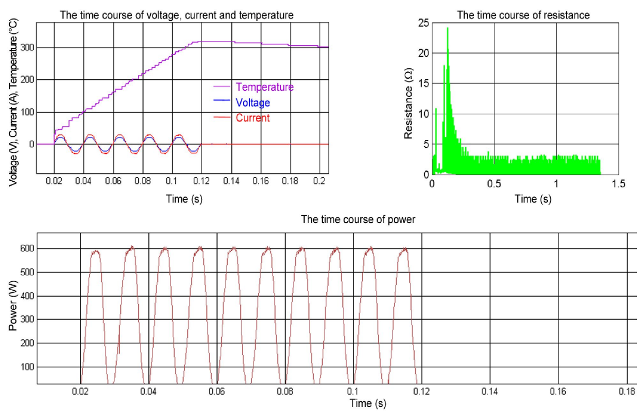


Fig. 4. Progress of selected parameters during the impulse.

In the next step, magnetostriction, our sample was measured as another parameter of magnetic properties. It was measured in the laboratory of PF-UPJŠ using Narita method [8]. Magnetostriction is given by the following equation:

$$a_S = \frac{1}{3} \cdot J_S \cdot \frac{\Delta H}{\Delta \pm} \qquad (1)$$

Where J_S is saturation magnetization, $\Delta\sigma$ is the change of mechanical stress and ΔH is a corresponding change in the magnetic field. Ratio $\frac{\Delta H}{\Delta \pm}$ is a slope fitting line of the curve.

The calculated value of magnetostriction is 21.13×10^{-6} for our sample.

III. RESULTS

Results presentations and publications previous year:

- Active participation in the conference SCYR 2014: Experimental Measurements and Results Obtained during 2^{nd} Year of my Postgraduate Study /*L. Hubač*

- Published paper in proceedings: IMPULZNÝ OHREV, AKO MOŽNOSŤ ÚPRAVY MAGNETICKÝCH VLASTNOSTÍ FEROMAGNETICKÝCH AMORFNÝCH ZLIATIN / *Lukáš Hubač, Ladislav Novák*, In: ELECTRICAL ENGINEERING AND INFORMATICS 5 : PROCEEDINGS OF THE FACULTY OF ELECTRICAL ENGINEERING AND INFORMATICS OF THE TECHNICAL UNIVERSITY OF KOŠICE.

- Paper was published in Current Contents journal ACTA PHYSICA POLONICA - A: PULSE HEAT TREATMENT OF FINEMET ALLOYS UNDER TENSION / *Antal Lovas, Lukáš Hubač, Ladislav Novák*

IV. THE NEXT STEPS

In the future we would like to extend the measuring apparatus to be able carry out uniaxial mechanical stress that is to be controlled by PC. In the next step magnetization on ribbons type FINEMET in the outer magnetic field with uniaxial mechanical stress will be measured. We want explore influence of the pulse annealing during uniaxial mechanical stress, on changes in magnetic properties and the relation of these changes to the structure of these materials.

REFERENCES

[1] Lovas A., Kiss L. F., Varga B., Kamasa P., Balogh I., Bakonyi I., J. Phys. IV (France) 8. Pr2/ 298 (1998)

[2] Varga L.K., Gercsi Zs, Kovács Gy, Kákay A, Mazaleyrat F., J. Magn. Magn. Mater. 254-255 (2003) 477-479

[3] L. Hubač - 2013 . In: SCYR 2013 : *Device for pulse annealing of amorphous magnetic ribbon* : Proceedings from conference : 13th Scientific Conference of Young Researchers : May 14th, 2013, Herlany, Slovakia. - Košice : TU, 2013, S. 250-252. - ISBN 978-80-553-1422-8

[4] Lukáš Hubač, Ladislav Novák – 2013. In: Senzorika a magnetometria 2013: *Magnetické vlastnosti amorfnej zliatiny typu FINEMET v procese spracovania impulzným ohrevom*: zborník prezentácií a príspevkov z vedeckej konferencie s medzinárodnou účasťou : 5. decembra 2013, Košice, Slovenská republika. - Košice : TU, 2013 S. 26-30. - ISBN 978-80-553-1540-9

[5] Antal Lovas, Lukáš Hubač, Ladislav Novák - 2013. In: CSMAG'13 : *Pulse heat treatment of FINEMET alloys under tension:* 15th Czech and Slovak Conference on Magnetism, In: Acta Physica Polonica A. Vol. 126, no. 1 (2014), p. 100-101. - ISSN 0587-4246, impact factor: 0,53.

[6] MathWorks, http://www.mathworks.com/support/, (Visited on 11-12/2014).

[7] National Instruments, http://sine.ni.com/psp/app/doc/p/id/psp-115/lang/cs, (Visited on 11-12/2014).

[8] KENJI NARITA, JIRO YAMASAKI, and HIROTOSHI FUKUNAGA, Measurement of Saturation Magnetostriction of a Thin Amorphous Ribbon by Means of Small-Angle Magnetization Rotation, In: IEEE TRANSACTIONS ON MAGNETICS, VOL. MAG-16, NO. 2, MARCH 1980

Research and future proposal of the UWB low profile antenna based on LTCC

[1]*Ján SCHNEIDER (2nd year)*
Supervisor: [2]Ján GAMEC

[1,2]Dept. of Electronics and Multimedia Communications, FEI TU of Košice, Slovak Republic

[1]ján.schneider@tuke.sk, [2]jan.gamec@tuke.sk

***Abstract*—This paper is summarization of the last year of author's dissertation topic. Firstly, the required theoretical knowledge about future research in the UWB low profile antenna design will be discussed. Next, we briefly introduce the author study activities in the last year. We also discuss about pedagogical part of postgradual study.**

***Keywords*—Ultra wide band, UWB antenna, low temperature cofired ceramic.**

I. INTRODUCTION

The topic of dissertation thesis is low-profile antennas of Ultra wide band (UWB) radars based on Low Temperature Co-fired Ceramic (LTCC). UWB is technology in which its transmission occupies a bandwidth of more than 20% of its center frequency (>500 MHz). Antenna is one of the most important components in UWB systems [9]. Ideally, the UWB antenna should be compact, planar and low cost. Our general research direction is design, simulation and practical realization of several functional samples of UWB antennas on LTCC, which will be compatible with UWB radar system/sensor network. The LTCC technology is a suitable technology to provide low cost, high performance, and high level of integration. This is mainly because the LTCC technology is able to integrate complex 3D structures such as vias or air cavities. Many studies have hence been devoted to integrated LTCC antennas [8]. This paper is divided in two sections. Section one discuss about research tasks and in section two pedagogical tasks solved in previous year are stated.

II. RESEARCH TASKS SOLVED IN PREVIOUS YEAR

First and huge part with high priority of my solved tasks in previous year was to write and to vindicate thesis to the dissertation examination. Dissertation work describes theoretical overview which is useful to requirements of low profile antenna. Work is divided into the three main chapters. First chapter describes basic features of dielectric materials which are permitivity, losses and loss tangent and also work describes two kinds of material used in microwave area: LTCC and PCB (Printed circuit board) consisting of ceramic-filled PTFE composites, hydrocarbon ceramic laminates or filled PTFE (random glass or ceramic). LTCC was chosen for low losses in the UWB frequencies (for example band width by FCC (Federal Communications Commission) is: 3.1-10.6 GHz) and for the ability to create 3D multilayer structures. Second chapter describes evolution progress in development wideband antennas as in the [3] and also a few wideband and multi-band antennas currently used. Next, parameters of the antenna are described. Band widths of antenna, gain, frequency dependence of reflection coefficient are most important. These parameters with physical dimensions will be evaluated in our low profile antenna based on LTCC. Last chapter of thesis to the dissertation examination describes methods of modification of electromagnetic field of antenna. These methods are named EBG (electromagnetic band gap) structures. EBG structures are artificial periodic (or sometimes nonperiodic) objects that prevent/assist the propagation of electromagnetic waves in a specified band of frequency for all incident angles and all polarization states [6]. For our case, in lowprofile antenna design, EBG structures are useful to suppress the surface waves. These structures are also found to be effective for bandwidth improvement, compact microstrip antenna designs, harmonic control, and the radiation pattern control. EBG structure is also an attractive solution to reduce the power loss through the substrate of a conventional antenna [7].

During the summer vacation (from June 29th to July 18th, 2014), I participated with my colleagues Jozef Lipták, Daniel Novák and Marek Godla on a very intensive program of MARTIN Summer School Marine Technology Instrumentation, organized by the SARTI centre of Technical university of Catalonia in Vilanova i la Geltr`u (Spain). Lectures, exercises and projects of this summer school were mainly focused on the technology used in the Navy (inertial navigation, measurement of seawater parameters, autonomous underwater vehicles, underwater communication systems …) and base introduction to LabView (myRio).

Fig. 1 Participation diploma

All projects were presented in front of the committee and awarded by participation diplomas (Fig. 1). The projects were also published in the Instrumentation Viewpoint journal [1][4][5]. Main purpose of our project was draft application in LabView, which monitors rotating engine of Glider, AUV (autonomous underwater vehicles) or ASV (autonomous surface vehicles) and indicates the fault of engine, if detects it by acoustic waves.

I am member of the research team of Prof. Dušan Kocur created for realization APVV project oriented to location of persons using UWB radar system . Last year of study, I have been involved in few events for the general public in which our research team promoted functionality and usage of UWB radar systems for rescue and other security forces.

Also, I participate at different various measurements localization and detection of persons with impulse UWB radar, see in Fig. 2. My task is to set the appropriate parameters of UWB radar antenna system. These parameters include, for example, height and relative spacing of the antennas, type of used antenna. The aforementioned factors affect the adoption of radar signal parameters such as signal-to-noise ratio, attenuation etc.

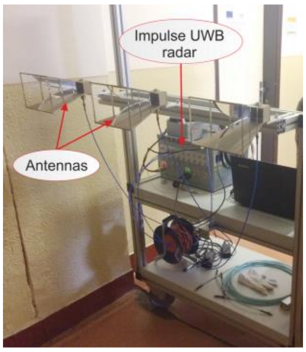


Fig. 2 UWB impulse radar system with antennas

III. Pedagogical tasks solved in previous year

Most important part of doctoral study is also teaching students of 1st and 2nd degree of the study on university. In previous year, I was teaching practice exercises from subject of automotive electronics. For this subject, I prepared brand new set of practical exercises on real KIA Ceed vehicle. Students can test and verify the functionality of the selected sensor (lambda probe, sensor camshaft) in engine compartment with multifunctional diagnostic station, see Fig. 3. Further, students could in addition to testing the functionality of the sensors go through the whole process of emission control with diagnostic station. They create a protocol that can be compared with the emission standards set by the manufacturer or by State Slovak Republic.

IV. Conclusion and proposal for next research

For further research it is planned a proposal of a low profile UWB antenna with a desired frequency bandwidth (~ from 200MHz to 6GHz), with as small geometric dimensions as possible, the standing wave ratio of at least 2.5 in (SWR≤2,5) and maximize gain by focusing on the implementation of the antennas on LTCC. Next, focus and developing optimization process predictability shrinkage of the material is in the process of firing low-temperature ceramic firing, with the component after firing to achieve the desired dimensions with respect to wavelength. And finally, realize a series of functional samples antennas. Verified by measuring, the characteristics of antennas on the basis of electrical parameters. Evaluate the low profile antenna statistical methods and compared with other commercial antennas.

Also, a multi-view e-learning system for remote education [2] will be created from practical exercises and measurements with diagnostic station ACTIGAS on Kia Ceed vehicle created for automotive electronics.

Fig. 3 Diagnostic station ACTIGAS

Acknowledgment

APVV This work was supported by the Scientific Grant Agency of the Ministry of Education, science, research and sport of the Slovak Republic under the contract No.1/0563/13 and by the Slovak Research and Development Agency under the contract No. APVV-0404-12.

References

[1] Ján Schneider, Petr Mazůrek, Guénolé Le Bougeant, “Automatic Detection System Measuring Acoustically the Rotation Rate and Faults in a Rotating Engine”, in Instrumentation Viewpoint. No. 16 (2014), p. 34-35. - ISSN 1886-4864, 2014

[2] Ján Valiska ... [et al.], “A multi-view e-learning system for remote education” In: ICETA 2014 : 12th IEEE International Conference on Emerging eLearning Technologies and Applications : proceedings : December 4-5, 2014, Starý Smokovec. - Danvers : IEEE, 2014 S. 489-494. - ISBN 978-1-4799-7738-3, 2014

[3] Ján Schneider, Ján Gamec, “Overview of UWB Low-Profile Planar Antennas”, in Acta Electrotechnica et Informatica. Roč. 14, č. 2 (2014), s. 55-59. - ISSN 1335-8243, 2014

[4] Daniel Novák, Viktor Fúra, Alexandre Paillot, “Acoustic Channel Frequency Response” in Instrumentation Viewpoint. No. 16 (2014), p. 28-30. - ISSN 1886-4864, 2014

[5] Valentin Bouliou--Hello, Jozef Lipták, Mila Oriol Pascual, “Automatic Morse Transceiver”, in Instrumentation Viewpoint. No. 16 (2014), p. 36-37. - ISSN 1886-4864, 2014

[6] F. Yang and Y. Rahmat-Samii, Electromagnetic Band Gap Structures in Antenna Engineering, Cambridge University Press, Cambridge, UK, 2009

[7] Md. Shahidul Alam, Norbahiah Misran, Baharudin Yatim, and Mohammad Tariqul Islam, “Development of Electromagnetic Band Gap Structures in the Perspective of Microstrip Antenna Design,” International Journal of Antennas and Propagation, vol. 2013, Article ID 507158, 22 pages, 2013

[8] Yang, B.; Yarovoy, A.; Amaldoss, S.E., "Performance analysis of a novel LTCC UWB 60 GHz semi-shielded aperture stacked patch antenna with differential feeding," Antennas and Propagation (EUCAP), Proceedings of the 5th European Conference on , vol., no., pp.1882,1885, 11-15 April 2011

[9] JANGID, S. - KUMAR, M.: "Compact Planar UWB Patch Antenna with Integrated Bandpass Filter & Band Notched Characteristics," Communication Systems and Network Technologies (CSNT), 2012 International Conference, pp.15, 19, 11-13 May 2012

Research of special equipment to improve transient stability of power systems.

[1]*Zsolt ČONKA 3rd year*
Supervisor: [2]Michal KOLCUN

[1,2]Dept. of Electric Power Engineering, FEI TU of Košice, Slovak Republic

[1]zsolt.conka@tuke.sk, [2]michal.kolcun@tuke.sk

***Abstract* — This paper offers information's about research of special equipment's which are utilized to improve transient stability of power systems. These devices have a wide range of applications. For example UPFC, the main function is to regulate the power flow, but it has also positive influences on transient stability as well as at voltage stability. These days the problem of stability of power systems is more recent.**

***Keywords*— FACTS, power system, transient stability, black out.**

I. Introduction

One of the biggest problem in ENTSO-E interconnected power system is the transient stability. Proof of this problem is blackout in the U.S. in 2003 and a few months later in Italy. Over 50 Million American as Canadians stayed in August 2003 without electricity up to 42 hours. Blackout affected large American and Canadian cities New York, Detroit, or Toronto. Just a month later occurs blackout almost the whole of Italy. Without electric power found itself 57 million people. Nowadays, when the number of renewable energy sources rises, the risk of blackout also increases. One option to avoid such a huge power outage is to use advanced special equipment to improve system stability.

II. Recent Research

Research is focused on finding new opportunities to increase stability of power systems. Many international companies invest huge amounts of money in research and development of new devices that can improve the stability of the system. Many of these devices is the object of our research. Research of impact of specialized devices for improve transient stability of power system was investigated in some articles. Reference [1] showed the impact of Thyristor controlled series compensator (TCSC) on transient stability of little power system. The results showed that the TCSC can improve transient stability of power system. The Impact of Unified power flow controller (UPFC) on transient stability of PS was investigated in [2]. The influence of another special Flexible AC transmission system (FACTS) was publicized in [3]. Reference [4] shows the comparison of two special devices on transient stability, that are Static Var compensator (SVC) and TCSC. The results of this research are that the Thyristor controlled series compensator has better influence on transient stability of power system than Static Var compensator. The results of further research have been published in various conferences and journals: [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19].

III. My current research

Due to continued increases of electricity consumption, we are forced to increase the production of electricity. Building of new solar and wind power plants causes a number of new challenges. Building of huge number of wind farms in the seas, especially in northern Germany and the requirement to transfer this huge amount energy over long distances today presents considerable problems for transmission system operators. Another problem is the liberalization of the electricity market in which there is a large transit flow of electricity between power systems. These facts largely cause overloading of lines, deterioration of stability of the system and the risk of total collapse of the power system "Black Out".

Modern devices based on power electronics called FACTS devices can help us in reducing risks from these problems.

These devices are specifically designed to increase the transfer capability of existing lines and power flow controlling over certain transmission lines in the system [2]. Over the past few decades two generations of FACTS devices emerge .First generation of FACTS include the conception of thyristor controlled phase shift capacitors and transformers .Second generation of FACTS Utilize GTO (gate turn off) thyristor switched inverters Which are used as source voltage of inverters.

Static Var Compensator (SVC) is one of the most popular FACTS devices. It is a static source of reactive power. It is known that the SVC with suitable injection of signal may significantly improve the transient stability of the power system. SVC was used successfully to improve the transient stability of the synchronous machines.

Another important facility is Thyristor-controlled series compensator (TCSC). Main use of this device is to control the power flow in the transmission system. But there are other benefits that can be exploited to improve the transient stability of the transmission system. [3]

The emergence and development of new types of FACTS devices were mainly possible with utilization of GTO thyristors. Thanks to the new type of thyristors were developed STATCOM, which is seriously competitor of SVC. In terms of transient stability of the power system STATCOM provides better performance than SVC. It is able to respond quickly to fluctuations in the power system.

Very considerable device is UPFC (Unified Power Flow Controller). It was created as a combination of static synchronous series compensator (SSSC) and static synchronous compensator (STATCOM). The article [19] shows the UPFC influence on transient stability of the power system.

My current research is focused on measuring of the impacts of various FACTS devices on the transient stability of the Slovak transmission system. Further investigate the effects of regulators in Slovakia system on the operation of FACTS devices and vice versa.

IV. FUTURE RESEARCH

My future research will focus on modeling of an accurate model of the transmission system of the Slovak Republic with regulation of generators excitation and power system stabilizers.

Further research will be dedicated to the development of such a controller of FACTS devices that fully cooperate with regulators that are already installed in the power system.

Further research will present other properties of these devices. From previous studies it is known that these devices have other usable features in the power system. For example, the TCSC is proposed mainly due to the ability to regulate power flow, but also from my previous research, it is clear that it has a significant impact on the stability of electricity transmission. Another useful feature is the reduction of losses.

IV. CONCLUSION

This paper presents the importance and possibilities for improving the dynamic stability of the power system. Installation of new equipment in the transmission system in the near future may be necessary to maintain the reliability and quality of electricity supplied. Installation of such equipment is very costly, so it is important that we have chosen the right kind of equipment and choosing the right destination location of that device. Today, the application of these devices is hampered mainly because of the price, but we should consider what economic damage can cause several hours long Black Out, that we could avoid by installing FACTS devices.

ACKNOWLEDGMENT

This work was supported by VEGA 1/0388/13 project.

REFERENCES

[1] Impact of TCSC on the Transient Stability / Zsolt Čonka, Michal Kolcun - 2013. In: Acta Electrotechnica et Informatica. Roč. 13, č. 2 (2013), s. 50-54. - ISSN 1335-8243 Spôsob prístupu: www.versita.com/aei

[2] Vplyv UPFC na dynamickú stabilitu sústavy / Zsolt Čonka ... [et al.] - 2013. In: Elektroenergetika. Roč. 6, č. 2 (2013), s. 13-15. - ISSN 1337-6756 Spôsob prístupu: http://jeen.fei.tuke.sk/index.php/jeen/article/view/300/270

[3] Utilizing SVC for transient stability enhancement / Zsolt Čonka, Michal Kolcun - 2013. In: Elektroenergetika. Roč. 6, č. 2 (2013), s. 16-18. - ISSN 1337-6756 Spôsob prístupu: http://jeen.fei.tuke.sk/index.php/jeen/article/view/301/271

[4] Power system stability improvement using facts devices / Zsolt Čonka, Michal Kolcun - 2013. In: AEI 2013 : international conference on Applied Electrical Engineering and Informatics 2013 : September 02-10, Genoa, Italy. - Košice : FEI TU, 2013 S. 32-36. - ISBN 978-80-553-1459-4

[5] Impact of the regulators on the size of the short-circuit currents / Zsolt Čonka ... [et al.] - 2013. In: Elektroenergetika 2013 : proceedings of the 7th International Scientific Symposium on Electrical Power Engineering : September 18-20, 2013, Stará Lesná, Slovak Republic. - Košice : TU, FEI, 2013 S. 86-89. - ISBN 978-80-553-1441-9

[6] Modelling of excitation system influence to transient stability of power system using PSLF / Matúš Novák ... [et al.] - 2013. In: Elektroenergetika 2013 : proceedings of the 7th International Scientific Symposium on Electrical Power Engineering : September 18-20, 2013, Stará Lesná, Slovak Republic. - Košice : TU, FEI, 2013 S. 531-534. - ISBN 978-80-553-1441-9

[7] Simulation of faults in the power system and their elimination in the EMTP-ATP / Dušan Medveď ... [et al.] - 2013. In: Elektroenergetika 2013 : proceedings of the 7th International Scientific Symposium on Electrical Power Engineering : September 18-20, 2013, Stará Lesná, Slovak Republic. - Košice : TU, FEI, 2013 S. 484-487. - ISBN 978-80-553-1441-9

[8] Zlepšovanie dynamickej stability elektrizačnej sústavy optimalizáciou parametrov PSS / Matúš Novák, Zsolt Čonka, Michal Kolcun - 2013. In: Electrical Engineering and Informatics 4 : proceedings of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice. - Košice : FEI TU, 2013 S. 152-157. - ISBN 978-80-553-1440-2

[9] Possibilities to increase the efficiency of photovoltaic cells / Zsolt Čonka ... [et al.] - 2011. - 1 elektronický optický disk (CD-ROM). In: Renewable Energy Sources 2011 : Proceedings IP 2011 : 20. - 29.5.2011, Špičák - Železná Ruda, Czech Republic. - Plzeň : University of West Bohemia, 2011 P. 21-24. - ISBN 978-80-261-0010-2

[10] Improving transient stability with utilizing FACTS device / Zsolt Čonka, Stanislav Kušnír - 2012. - 1 elektronický optický disk (CD-ROM). In: SCYR 2012 : proceedings from conference : 12th Scientific Conference of Young Researchers : May 15th, 2012, Herľany, Slovakia. - Košice: TU, 2012 S. 355-358. - ISBN 978-80-553-0943-9

[11] The Influence of PSS Parameters to Transient Stability of Power System /Matúš Novák, Zsolt Čonka - 2013. In: SCYR 2013 : Proceedings from conference : 13th Scientific Conference of Young Researchers : May 14th, 2013, Herľany, Slovakia. - Košice : TU, 2013 S. 422-425. - ISBN 978-80-553-1422-8

[12] Modelling of excitation system influence to transient stability of power system using PSLF / Matúš Novák ... [et al.] - 2013. In: Elektroenergetika 2013 : proceedings of the 7th International Scientific Symposium on Electrical Power Engineering : September 18-20, 2013, Stará Lesná, Slovak Republic. - Košice : TU, FEI, 2013 S. 531-534. - ISBN 978-80-553-1441-9

[13] Impact of the regulators on the size of the short-circuit currents / Zsolt Čonka ... [et al.] - 2013. In: Elektroenergetika 2013 : proceedings of the 7th International Scientific Symposium on Electrical Power Engineering : September 18-20, 2013, Stará Lesná, Slovak Republic. - Košice : TU, FEI, 2013 S. 86-89. - ISBN 978-80-553-1441-9

[14] Analysis of influence of PVs on island operation in power system/ Pavol Hocko ... [et al.] - 2013. In: Elektroenergetika 2013 : proceedings of the 7th International Scientific Symposium on Electrical Power Engineering:September 18-20, 2013, Stará Lesná, Slovak Republic. - Košice : TU, FEI, 2013 S. 94-97. - ISBN 978-80-553-1441-9

[15] Transient stabililty studies using different software tools / Matúš Novák ... [et al.] - 2014. In: Current Problems of Maintenance of Electrical Equipment and Management. - Košice : TU, 2014 S. 170-175. - ISBN 978-80-553-1818-9

[16] Relative permitivitty and dissipation factor of oil paper insulation / Lukáš Lisoň ... [et al.] - 2014. In: Current Problems of Maintenance of Electrical Equipment and Management. - Košice : TU, 2014 S. 197-203. - ISBN 978-80-553-1818-9

[17] Influence of photovoltaic power plants on the power system to the prediction for year 2020 / Pavol Hocko ... [et al.] - 2014. In: 14th International Conference on Environment and Electrical Engineering : conference proceedings : Krakow, Poland, 10-12 May, 2014. - Danvers

[18] Impact of connection point of alternative energy sources to the size of short circuit current in the distribution network / Miroslav Kmec ... [et al.] - 2014. In: Proceedings of Intensive Programme 2014 : Perspectives for the development of low-power systems using biomass in the context of the EU energy policy for the Central European region : July 6th-17th 2014, Pardubice, Czech Republic. - Plzeň : Západočeská univerzita v Plzni , 2014 S. 274-279. - ISBN 978-80-261-0356-1

[19] UPFC influence to transient stability of power system / Matúš Novák ... [et al.] - 2014. In: Elektro 2014 : 10th international conference : proceedings : Rajecké Teplice, May 19-20, 2014. - Žilina : ŽU, 2014 S. 343-346. - ISBN 978-1-4799-3720-2

Simulation and Analysis of Wire Rod Rolling

[1]Tomáš BOROVSKÝ (4st year)
Supervisor: [2]František ĎUROVSKÝ

[1,2]Dept. of Electrical Engineering and Mechatronics, FEI TU of Košice, Slovak Republic

[1]tomas.borovsky@student.tuke.sk, [2]frantisek.durovsky@tuke.sk

***Abstract*—A tension control plays an important role in assuring product quality and process stability in steel rolling industry. The paper presents the mathematical model of a part of wire rod roughing mill and it is used to simulation and analysis of the tension control algorithm.**

***Keywords*—wire rod rolling, tension control, rolling torque**

I. Introduction

The roughing mill consists of six stands, individually driven by motors. The workpiece pass through the stands with box-oval and round-oval grooves in horizontal-vertical rolling sequence. There is no rolling force measurement, no interstand tension measurement and no loop is formed between the stands. Hence, motor current or motor torque is used to interstand tension control [1]. In general, rolling with low tension on roughing mill is a fundamental condition to ensure the stable rolling process. However, deviation in rolling parameters as a groove wear, or new set up, can increase or decrease interstand tension [2]. Tensions between roll stands have a great influence on delivery size. When end of a bar passes through a roll, the tension of the bar suddenly turns to zero and the size of bar changes [3]. All these deviations have a big impact on product quality and on the time that is required to set up the mill. Moreover, these changes in cross sectional area along the workpiece can lead to the cobbles on intermediate and finishing mill. Massive tension can cause total unload of given stand that consequently leads to cobbles due to the failure of interstand tension control. Against, no tension causes loopering between the stands, that leads to cobbles on roughing mill.

II. Rolling process and tension control

Rolling process is a complex and difficult phenomenon and its simulation and analysis requires a deep knowledge of rolling, plant and control system parameters. Calculations of the main rolling parameters, such as rolling speed and rolling torque, are described in [1]. Drive, motor and mechanical parameters used for simulation are derived from original ones, used in rolling mill Slovakia Steel Mills, a.s. Strážske. Control structure includes *Cascade control, Impact Droop Function* and *Tension control*.

On a continuous rolling mill fixed speed ratio must be achieved to get stable mass flow. Interaction between upstream and downstream stands are not allowed except the tension control. The feed-forward, so-called *Cascade control function*, is used for these purposes [4].

The *Impact Droop function* is used to get smoother workpiece entry. It is suggested to run subsequent stand faster, with an speed offset. Once workpiece passes the given stand, the speed offset is set to zero, and excess in inertial energy of given stand, due to the higher speed, helps in reducing speed drop [5].

As it was above mentioned, there is no possibility to exactly measure the tension between the stands. Hence, the motor torque, which is proportional to the rolling torque, serves as the tension indicator [6]. In stable gauge conditions, interstand tension is a result of relative rolling velocity variation in successive stands of continuous rolling sequence [7]. Tension torque can be caused by back or front tension, and it has affect on position of neutral point. With the variation of neutral point position along the roll bite length comes motor torque deviation due to the rolling force deviation [7]. This deviation in motor torque can be observed once the workpiece reaches successive stand. Hence the motor torque of given stand before and after the loading of successive stand is used for the *Tension control* [6].

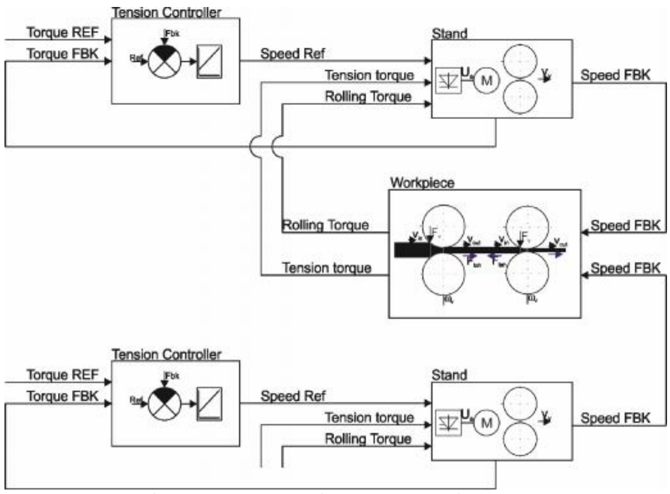


Fig. 1. Scheme of simulation model for two stands of rouging mill.

III. Simulation model

Simulation model is created in *MATLAB Simulink*, and it involves first three stands of roughing mill; STD_{01}, STD_{02}, STD_{03}. Simulation model consists of three main parts:

- model of drive and stand
- model of workpiece
- control structure

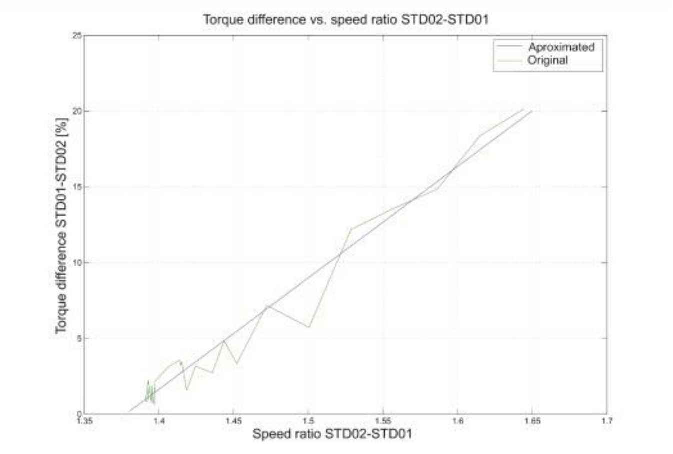


Fig. 2. Relation between speed ratio and front tension M_{vtf} on STD_{01}.

General simulation model includes three models of rolling stands and three models of workpieces. In Fig. 1 is shown a principle scheme of the simulation model for two stands.

A. Model of the drive and stand

Model of drive and stand is based on real machines and it includes a model of drive, gearboxes and rolling stand.

B. Model of the workpiece

As it was above mentioned, load torque is represented by tension free rolling torque Mv_0 and torque deviation from front (M_{vtf}) and back (M_{vtz}) tension. In order to simplify the model, back tension M_{vtz} is neglected. M_{v0} is a static part of the workpiece model and its value was determined experimentally, during tension free conditions on the mill. M_{vtf} represents a dynamic part of the workpiece model and its value is changing with the speed ratio change. The relation between speed ratio and M_{vtf} was identified experimentally, during tension conditions on the mill, and it is shown in Fig. 2. Hence, linear function has been determined as an dynamic part of the workpiece model.

C. Control structure

Control structure has to control the rolling speed of the stands and maintain desired tension conditions. Principled scheme of the control structure is shown in Fig. 3. The main points are *Cascade speed control, Impact droop function* and *Tension control*.

Cascade speed control uses a feed-forward function to achieve stable speed rations between the stands. Hence the tension controller output of given stand is added to the tension controller output of the upstream stand.

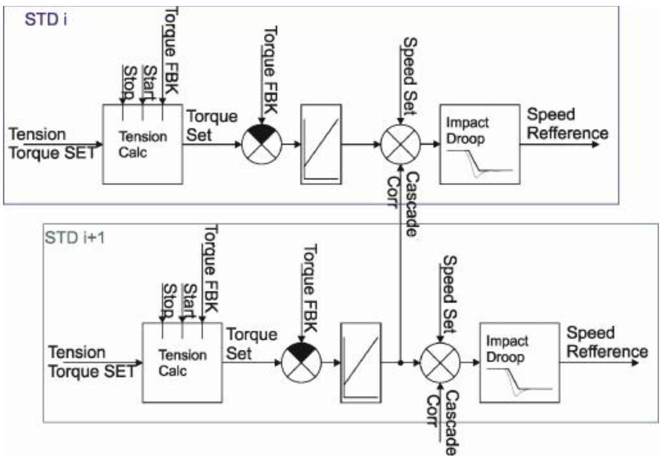


Fig. 3. Scheme of the Control structure for two stands

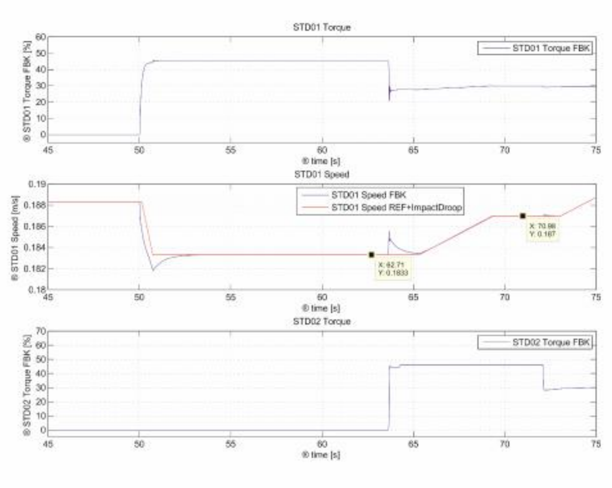


Fig. 4. Simulation of rolling on STD_{01} and STD_{02}. Tension controller increases the speed of STD_{01} to reduce the tension.

Impact droop function starts to decelerate given stand when the load is detected. Results of *Impact droop function* for STD_{01} can be seen in Fig. 4 as a reference speed droop during stand loading.

Tension control is used to maintain low interstand tension. Since there is no rolling force and no tension measurement on roughing mill, motor torque is used to the tension identification. An interstand tension between stands STD_i and STD_{i+1} is controlled by speed of STD_i. Tension free rolling conditions STD_i is represented by time interval, when workpiece is rolled by STD_i but does not reach consequent stand. Once workpiece enters STD_{i+1}, motor torque of STD_i can be affected by tension between STD_i and STD_{i+1} and *Tension control* turns on. Working sequence of *Tension control* during simulation is shown in Fig. 4.

IV. Further work

The next step is to create a model of the intermediate train and propose the new tension control algorithm for roughing and intermediate mill. More robustness control and decrease of the time required to the tension elimination are expected from the new design.

V. Acknowledgment

The authors wish to thank the project VEGA 1/0464/15 for its support.

References

[1] Tomáš Borovský, *Rolling Speed and Torque Prediction in Wire Rod Mill*, SCYR, Herľany, Slovakia, 2013

[2] Cony Eriksson, *Roll Pass Design for Improved Flexibility and Quality in Wire Rod Rolling,* Doctoral Thesis, Royal Institute of Technology, Stockholm, Sweden 2004.

[3] Harutoshi Ogai, Akira Fujii, Kanji Baba, Yukio Nouchi, *Free Tension Set-up of Rod Bar Rolling by Fuzzy Inference Method,* IEEE, Hokkaido, Japan, 1991

[4] Jingrong Liu, *Design andAnalysis of Intelligent Fuzzy Tenion Controllers for Rolling Mills,* University of Waterloo, Waterloo, Canada, 2002.

[5] Scaglia Antonella, Melandri Fraziano, *Impact Speed Drop Compensation Procedure for a new layout Wire Rod Mill,* IEEE, Milano, Italy, 2002

[6] Korhonen A.S.,Nikkilä K.J., *A Study on the Rod Exit Speed in Wire Rod Rolling,* Scandinavian Journal of Metallurgy, Helsinky, Finland, 1974

[7] Youngseog Lee, *Rod and Bar Rolling*. Pohang, Korea, POSCO Technical Research Laboratories, Pohang, Korea, 2004

Simulation of hybrid system

[1]*Miroslav MIKITA (1st year)*
Supervisor: [2]Michal KOLCUN

[1,2]Dept. of Electric Power Engineering , FEI TU of Košice, Slovak Republic,

[1]miroslav.mikita@tuke.sk, [2]michal.kolcun@tuke.sk

***Abstract*— Hybrid renewable energy systems (HRES) are becoming popular for remote area power generation applications due to advances in renewable energy technologies and subsequent rise in prices of petroleum products. Economic aspects of these technologies are sufficiently promising to include them in developing power generation capacity for developing countries. Research and development efforts in solar, wind, and other renewable energy technologies are required to continue for, improving their performance, establishing techniques for accurately predicting their output and reliably integrating them with other conventional generating sources. The paper describes methodologies to model HRES components, HRES designs and their evaluation.. This paper gives simulation results of PV-Wind-Diesel hybrid system. HOMER (Hybrid Optimization Models for Energy Resources) power optimization software by NREL (National Renewable Energy Laboratory) is used to simulate and analyze the PV-Wind-Diesel hybrid system. HOMER is a micro power optimization software used in evaluating designs of both off-grid and grid-connected power systems for a variety of applications.**

***Keywords*— PV; Wind; Battery; Homer; Simulation; Hybrid system; Renewable energy.**

I. Introduction

Hybrid Renewable Energy Systems (HRES) is composed of one renewable and one conventional energy source or more than one renewable with or without conventional energy sources, that works in stand alone or grid connected mode. HRES is becoming popular for stand-alone power generation in isolated sites due to the advances in renewable energy technologies and power electronic converters which are used to convert the unregulated power generated from renewable sources into useful power at the load end. The important feature of HRES is to combine two or more renewable power generation technologies to make best use of their operating characteristics and to obtain efficiencies higher than that could be obtained from a single power source. Hybrid systems can address limitations in terms of fuel flexibility, efficiency, reliability, emissions and economics [1].

A. Wind energy

The energy available in the wind depends on the density and air velocity. The density, as any other gas, changes with the temperature and pressure which varies with the high level of the sea. The energy of a mass of air which is displaced is determined by the Kinetic Energy (K.E) flux.

$$P_0 = \frac{1}{2} \cdot \rho \cdot A \cdot V^3 \tag{1}$$

Where,
P_0 - Power in Watts
ρ- Air Density in kg/m3
A-Turbine Area in m2
V- Wind Speed in m/s.

When wind move across the wind turbine, the static pressure drops to a lower pressure than the atmospheric pressure. As the air follows its trajectory, it takes its atmospheric value again, inducing an extra wind deceleration. By this way, in a distance between upstream of the turbine and downstream, behind the turbine, there is no change in static pressure, but there is a reduction in kinetics energy. This phenomenon is represented by the Betz law in eqn (2).

$$P_{MAX} = \frac{8}{27} \cdot \rho \cdot A \cdot V^3 \tag{2}$$

Where,
P_{max} - Power in Watts
ρ- Air Density in kg/m3
A-Turbine Area in m2
V- Wind Speed in m/s.

Wind is a natural phenomenon related to the movement of air masses caused primarily by the differential solar heating of the earth's surface. Seasonal variations in the energy received from the sun affect the strength and direction of the wind. The ease with which wind turbines transform energy in moving air to rotary mechanical energy suggests the use of electrical devices to convert wind energy to electricity. Wind energy has also been utilized, for decades, for water pumping as well as for the milling of grains

B. Solar Energy

Solar energy is the most promising of the renewable energy sources in view of its apparent unlimited potential. The sun radiates its energy at the rate of about 3.8 x 10^{23} kW per second. Most of this energy is transmitted radially as electromagnetic radiation which comes to about 1.5kW/m^2 at the boundary of the atmosphere. After traversing the atmosphere, a square meter of the earth's surface can receive as much as 1kW of solar power, averaging to about 0.5 over all hours of daylight. Studies relevant to the availability of the solar energy resource in Slovakia have indicated its viability for practical use.

In solar photovoltaic applications, the solar radiation is converted directly into electricity. The most common method of doing this is by the use of silicon solar cells. The power generating unit is the solar module which consists of several solar cells electrically linked together on a base plate. On the whole the major components of a photovoltaic system include the arrays which consist of the photovoltaic conversion devices, their interconnections and support, power conditioning equipment that convert the dc to ac and provides regulated outputs of voltage and current; controller, which automatically manages the operation of the total system; as well as the optional storage for stand alone (non-grid) systems [2].

C. HOMER

HOMER (Hybrid Optimisation Model for Electric Renewables) – another time-step simulation program developed by NREL – adds optimization to basic simulation capability. It simulates the annual performance of many different system configurations for a specified set of energy sources to find a configuration that satisfies technical constraints at the lowest life-cycle cost. It is also possible for the user to define sensitivities (e.g. different mean values for solar irradiation, wind or power consumption) to narrow the range of results. The outcome of the simulation is a list of the possible systems in order of life-cycle costs. A graph depicts the various ranges of the most cost-effective systems over the given operating period, based on the selected criteria. Detailed results can be output for each of the individual simulated systems (graphs, tables, scatter plot, print-out)[3],[4],[5].

II. Hybrid System Parameters

A. Load profile

The load profile is hypothetical higher load. Fig. 1. illustrates this profile. A higher base load of 2 kW occurs throughout the day and night. Small day peak of 4 kW occur in the morning and at noon, while the majority of the load occurs in the evening. This evening load, with a peak load of 6,7 kW. The total daily load averages 85 kilowatt-hours per day.

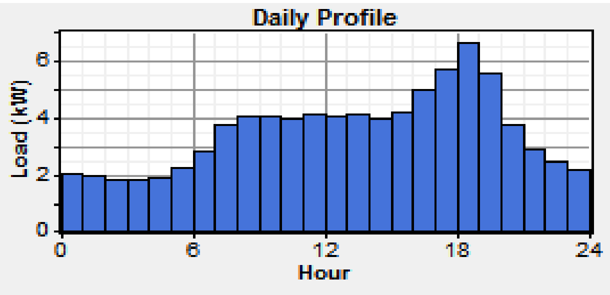


Fig. 1. Hourly load profile

B. Solar resource data

The solar resource was used for a site in Košice at a location of 48° 43' N latitude and 21° 15' E longitude. Solar radiation data for this region was obtained from the NASA Surface Meteorology and Solar Energy web site. The annual average solar radiation for this area is 3,09836 kWh/m^2/d. Fig. 2. shows the solar resource profile over a one-year period.

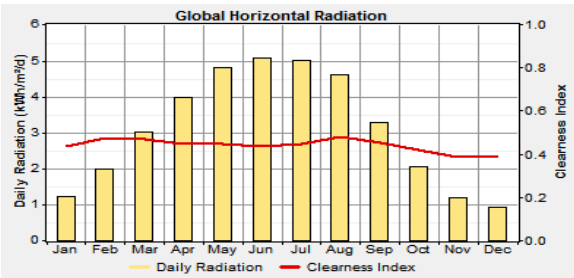


Fig. 2. Solar radiation a clearness index profile for Košice

C. Wind resource data

Wind resource data was also used for a site in Košice. The annual scaled average speed of wind in this site is 4,5 m/s. Figure 3 shows the wind resource profile over a one-year period.

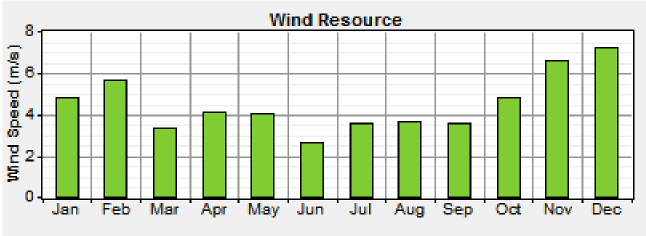


Fig. 3. Wind speed profile for Košice

III. Hybrid Simulation System

After the system components and the equations, Modeling and simulations of the micro power system is carried out. HOMER, optimization model is used to simulate the system. Large number of options are available for different sizes of the components used, components to be added to the system which make sense, cost functions of components used in the system. HOMER's optimization and sensitivity analysis algorithms evaluated the possibility of system configuration.

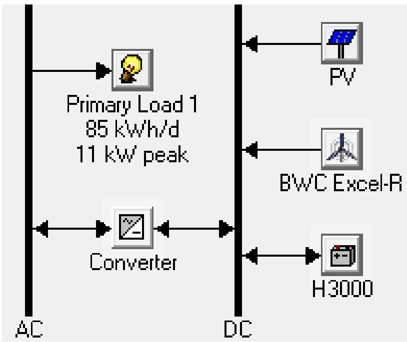


Fig. 4. PV-Wind-Battery hybrid generation system

A. Simulation result

Fig. 5. shows two possibilities of hybrid system sizes which are based at simulation. First and best choice is PV-Wind-Battery system and the second is PV-Batery system. Both of

				PV (kW)	XLR	H3000	Conv. (kW)	Initial Capital	Operating Cost ($/yr)	Total NPC	COE ($/kWh)	Ren. Frac.	Capacity Shortage	Batt. Lf. (yr)
				15	1	15	10	$ 32,938	631	$ 41,001	0.111	1.00	0.09	15.0
				30		50	10	$ 39,125	314	$ 43,137	0.119	1.00	0.10	20.0

Fig. 5. PV-Wind-Battery hybrid generation system

this systems can provide electricity for the default load.

Fig. 6. shows the monthly production from sources PV sources production per year is 20809 kWh and Wind turbine production per year is 21298 kWh. Year production of whole hybrid system is 42107 kWh.

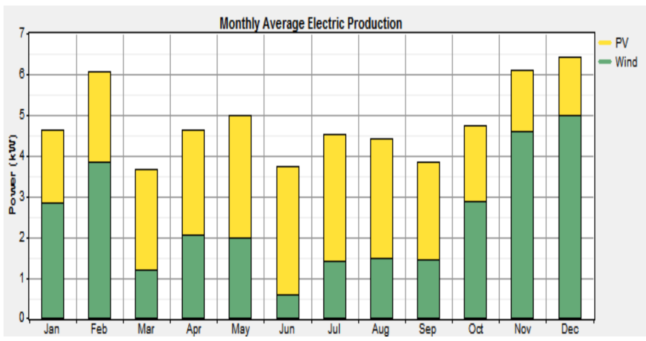


Fig. 6. Monthly average production in PV-Wind-Battery system

B. Sumary of problems and next direction

Problems in sizing optimization of hybrid systems are with chosing right mix of renewable sources to best cover load of system for lowest price of whole hybrid system.

Next direction of my work is designing hybrid systemes conected to grid and analyze benefits of this connectin and behaviour of hybrid system connected to the grid.

ACKNOWLEDGMENT

This work was supported by Slovak Research Agency No. VEGA 1/0388/13 project.

REFERENCES

[1] NEMA, Pragya; NEMA, R. K.; RANGNEKAR, Saroj. A current and future state of art development of hybrid energy system using wind and PV-solar: A review.Renewable and Sustainable Energy Reviews, 2009, 13.8: 2096-2103.Kotler P., Keller K.L., Marketing, Rebis, Poznań 2012.

[2] AJAO, K. R.; OLADOSU, O. A.; POPOOLA, O. T. Using HOMER power optimization software for cost benefit analysis of hybrid-solar power generation relative to utility cost in Nigeria. International Journal of Research and Reviews in Applied Sciences, 2011, 7: 96-102.

[3] KATSIGIANNIS, Y. A.; GEORGILAKIS, P. S. Optimal sizing of small isolated hybrid power systems using tabu search. Journal of Optoelectronics and Advanced Materials, 2008, 10.5: 1241.

[4] GIVLER, T.; LILIENTHAL, P. Using HOMER® software, NREL's micropower optimization model, to explore the role of gen-sets in small solar power systems. Case Study: Sri Lanka, National Renewable Energy Laboratory, Golden, Colorado, 2005.

[5] AL-KARAGHOULI, Ali; KAZMERSKI, L. L. Optimization and life-cycle cost of health clinic PV system for a rural area in southern Iraq using HOMER software.Solar Energy, 2010, 84.4: 710-714.J. Wang, "Fundamentals of erbium-doped fiber amplifiers arrays (Periodical style—Submitted for publication)," *IEEE J. Quantum Electron.*, submitted for publication.

Soft Switching High Frequency DC/DC Converters with Secondary Side Active Rectifier and Snubber

[1]*Róbert ŽATKOVIČ (1st year)*
Supervisor: [2]Jaroslav DUDRIK

[1,2]Dept. of Electrical Engineering and Mechatronics, FEI TU of Košice, Slovak Republic

[1]robert.zatkovic@tuke.sk, [2]jaroslav.dudrik@tuke.sk

***Abstract*—A paper is focused on unconventional isolated DC/DC converters with secondary side active rectifier. By this special type of converter topology soft switching of switches on the primary side of isolated DC/DC converter could be achieved. In the paper a principled idea of this type of topology is described, next overview from published works to this topic is done and also the function of basic version of isolated DC/DC converter with active rectifier is discussed and simply analyzed. Finally, there are offered solutions and possible ways to develop the desired fully soft switching DC/DC converter topologies.**

***Keywords*—active rectifier, DC/DC converter, snubber, soft switching**

I. Introduction

Soft switching isolated DC/DC converters are frequently mentioned type of power converters. They are high frequency operated converters and switching frequency is very important working condition for them. It is known, that increase of their switching frequency causes reduction of their weight and volume. Therefore, tendency is to use very high switching frequencies. Unfortunately, switching losses in transistor switches are proportional to switching frequency. Soft switching technique in general reduces switching losses, ideally to zero, independently on working frequency of power converter. Hence application of some kind of soft switching technique to the high frequency DC/DC converter is very useful and desired to achieve its high efficiency, small weight and size. The DC/DC converter with secondary side active rectifier offers good conditions, which could result in achieving this goal. This type of converter has one or more transistor switches on the secondary side. By these secondary switches soft switching on the primary side in full load range is very simply achieved. With combination of some kind of auxiliary circuit like snubber on secondary side, it could be designed quite simply soft switching high frequency DC/DC converter. Principled block diagram of this type of isolated DC/DC converter is shown in Fig. 1.

II. Overview of Published Solutions

Soft switching DC/DC converter with active rectifier controlled by one switch with passive snubber is presented in [1]–[4]. This kind of topology consists of full bridge inverter, high frequency center tapped transformer, center tapped full

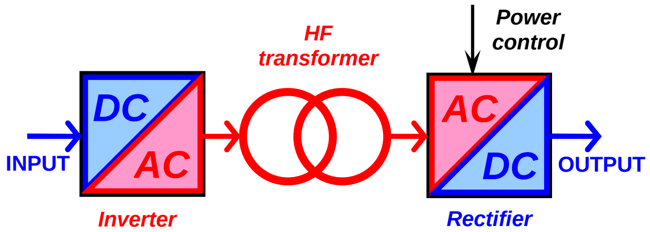


Fig. 1. Block diagram of DC/DC converter with secondary side active rectifier.

wave controlled rectifier with one switch and snubber, and output filter. The mentioned converter is controlled by classical pulse width modulation. This kind of solution is very simple, low cost and easy applicable. Due to few devices in series with load current, conduction losses are small. The problem can occur only at soft switching of rectifier transistor switch, when big change of his duty cycle occurs. Further, in some cases it can be problem that switching frequency of rectifier switch is so high. Soft switching DC/DC converter with active rectifier controlled by one switch with active snubber is presented in [5]. This kind of topology consists of full bridge inverter, high frequency center tapped transformer, center tapped full wave controlled rectifier with one switch and active snubber, and output filter. This converter and also active snubber are controlled by classical pulse width modulation. The main advantage is that this kind of topology ensures soft switching of rectifier transistor switch for big change of his duty cycle. Also due to only few devices in series with load current, the conduction losses can be quite small. The disadvantage is that active snubber consists from a lot of devices and moreover, one additional control signal is required. Soft switching DC/DC converter with active rectifier controlled by two switches is presented in [6]. This topology consists of full bridge inverter with lossless snubber, high frequency center tapped transformer, center tapped full wave controlled rectifier with two switches, and output filter. The mentioned converter is controlled by phase shifted pulse width modulation. This type of topology is a simple and basic solution of DC/DC converter with active rectifier controlled by two switches. It is easy applicable in final product. Significant disadvantage is that inverter transistor switches turn off only at zero voltage, so when IGBTs are used as switches, tail current causes switching losses. The soft switching DC/DC converter with active rectifier controlled by two switches with symmetric passive snubber is presented in

[7]–[12]. This kind of topology consists of full bridge inverter, high frequency center tapped transformer, center tapped full wave controlled rectifier with two switches and symmetric snubber, and output filter. This converter is controlled by modified pulse width modulation. This topology can be used in very high power applications. It can be very useful especially at high output voltage of the converter. The problem can occur only at soft switching of rectifier transistor switches, when big change of their duty cycle occurs. The disadvantage of this kind of topology is also that the symmetric passive snubber is quite complicated and consists of many devices. Soft switching DC/DC converter with active rectifier controlled by two switches is presented in [13]–[18]. This topology consists of full bridge inverter with lossless capacitor snubber, high frequency transformer, full bridge controlled rectifier with two active switches connected in one leg of rectifier, and output filter. The mentioned converter is controlled by phase shifted pulse width modulation. The main advantage is that at this kind of topology simple transformer with only one primary and one secondary winding is used. But conduction losses of whole converter can be high, because there are a lot of devices in series with load current. The main disadvantage is that inverter transistor switches turn off only at zero voltage, so when IGBTs are used as switches, tail current causes some turning off losses.

III. Analysis of the Simplest Version

The basic and simplest version of isolated DC/DC converter with secondary side active rectifier has one transistor switch on the secondary side. The circuit diagram of topology is shown in Fig. 2.

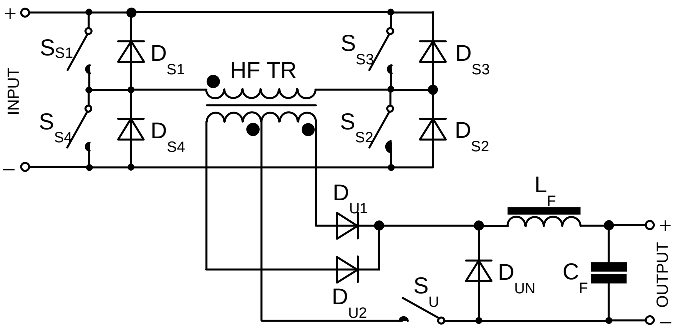


Fig. 2. Circuit diagram of DC/DC converter with active rectifier with one switch on secondary side.

This topology is very simple, easy understandable and analyzable. Moreover, it is very easy applicable in the real converter design. The basic idea of the topology is to turn off load by rectifier transistor switch before the inverter transistor switches turn off. It causes that load current in secondary winding of transformer falls down, and hence falls down load current in primary winding. It results in zero current value in corresponding inverter transistor switches. Of course, magnetizing current of high frequency transformer will still occur, but it is insignificant because of its low value. So we can speak about zero current turning off of inverter transistors. To achieve this goal the rectifier switch has to be turned off before corresponding inverter switches. Otherwise inverter switches turn off full load current and therefore we speak about hard turning off, respectively full current turning off which of course causes high turning off losses. In Fig. 3 simplified analytical time waveforms which can easily describe the operation principle of this kind of topology are shown.

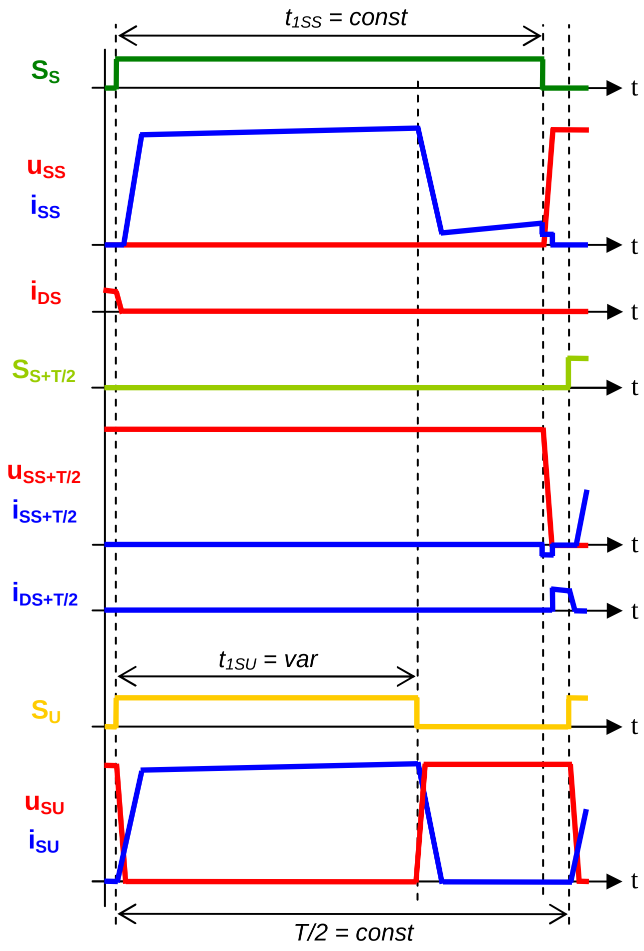


Fig. 3. Simplified time waveforms of DC/DC converter with active rectifier controlled by one switch.

It is obvious that switching frequency of rectifier transistor switch is two times higher than switching frequency of inverter transistor switches. Corresponding inverter transistor switches S_{S1} and S_{S2} (respectively S_{S3} and S_{S4}) cooperate in same constant duty cycle. Between turning off one pair of inverter switches and turning on the other pair of inverter switches, must be ensured switching dead time. It is to protect transistors from their destruction, but also it must be provided to reach desired conditions for soft turning on of corresponding inverter switches. During the dead time, magnetizing current of high frequency transformer discharges the output capacities of inverter transistor switches. Therefore, duration of the dead time must be minimally so long to ensure that these capacities are discharged to zero. Then we can speak about zero voltage turning on of inverter transistors. From analytical time waveforms is obvious that inverter transistor switches also turn on at zero current, so we also speak about zero current turning on of inverter transistors. It is caused by leakage inductance of high frequency transformer which is in series with inverter switches. This inductance slows down the load current rise. The output power of converter is controlled only by rectifier transistor switch. So duty cycle of this transistor switch is variable. But it is important that it must be less then the duty cycle of inverter transistor switches, to achieve zero current turning off of these switches. Rectifier transistor switch turns on at zero current, so we can speak about zero current turning on of rectifier switch. It is like inverter switches, caused by leakage inductance of high frequency transformer which is in series with rectifier switch.

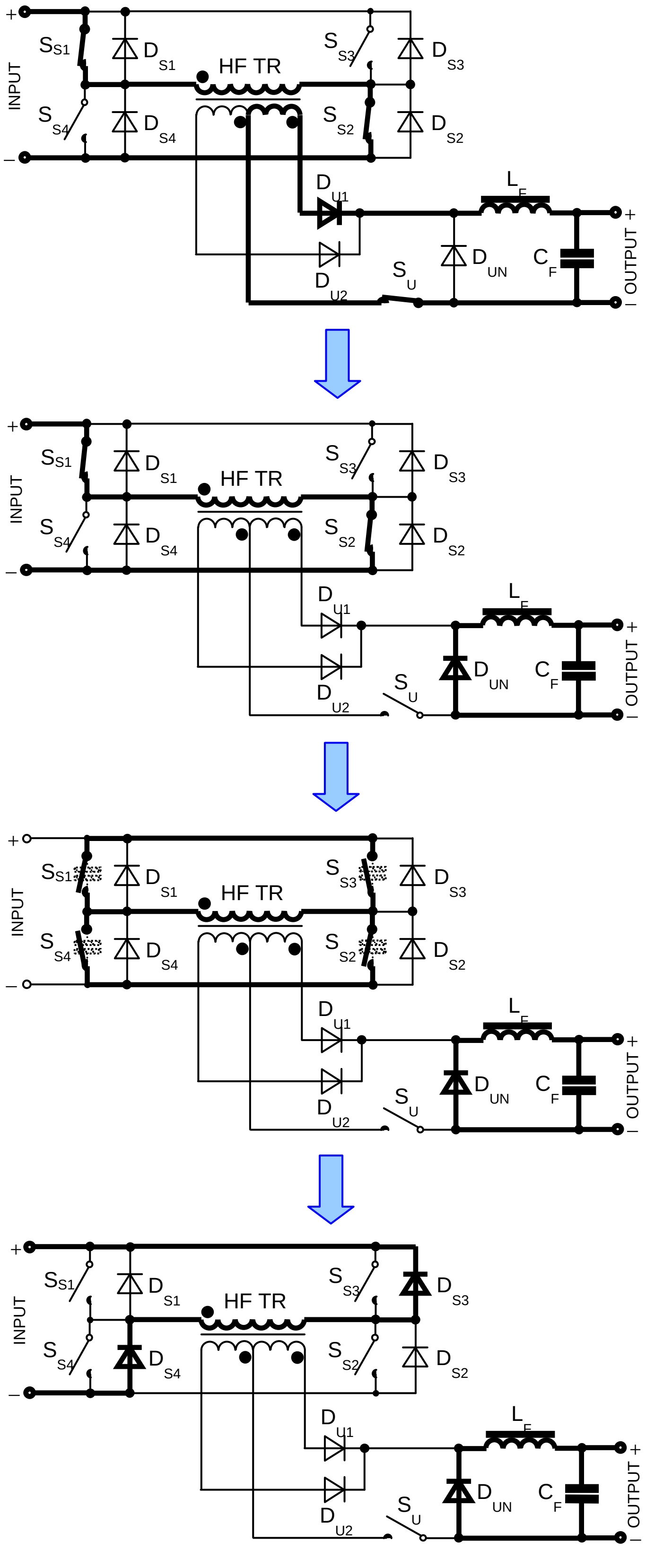


Fig. 4. Simplified sequence of equivalent circuit diagrams of DC/DC converter with active rectifier controlled by one switch.

But it is obvious that turning off of rectifier switch is at full load current, so we can speak about hard turning off. It of course causes high switching losses which are undesirable. Fig. 4 shows simplified sequence from equivalent circuit diagrams of significant period time intervals, by which can be better understood the operation principle of this topology.

Soft turning off of rectifier transistor switch S_U has to be ensured by auxiliary circuit. This auxiliary circuit could be some kind of lossless snubber. Principled circuit diagram of this kind of topology comprising snubber is shown in Fig. 5. In this way, rectifier switch ensures soft turning off of inverter transistor switches and snubber ensures soft turning off of rectifier transistor switch. Thus we can speak about fully soft switching high frequency DC/DC converter with secondary side active rectifier and snubber.

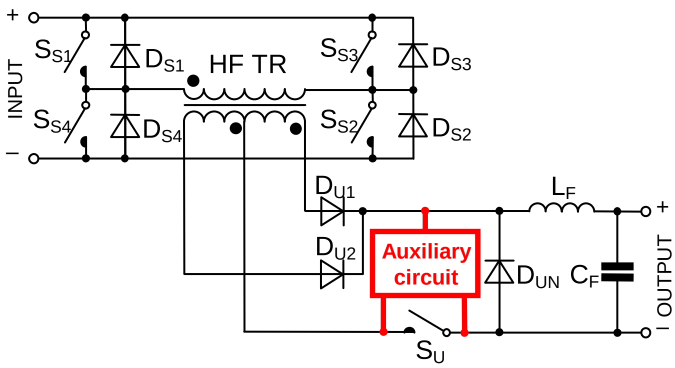


Fig. 5. Principled circuit diagram of DC/DC converter with active rectifier with one switch on secondary side and snubber.

IV. Conclusion

Isolated DC/DC converters with secondary side active rectifier are principally suitable to achieve soft switching of all transistor switches used in converter circuit. In general, active rectifier offers desired conditions to achieve soft switching of switches on the primary side in full load range. This solution of soft switching of inverter switches is very simple and always effective. So in real inverter circuit connection IGBTs can be easily used, even at very high switching frequency of converter and very big change of load. Therefore, high power IGBTs could be used and thus high power DC/DC converters could be designed. The inverter switches will always turn on at zero current and also at zero voltage, so we can speak about zero voltage-zero current turning on (ZVZCS). Turning off of the inverter switches will always be at zero current, so we can speak about zero current turning off (ZCS). Hence, the problem of the inverter soft switching is fully solved. It does not need to develop any other solution because this is a very good way to achieve the final goal. Therefore, a focus of research to this part of problem is not required. The soft switching of transistor switches on the secondary side must be ensured by the other solution. Soft turning on of the secondary switches is ensured. The turning on of the rectifier switches will always be at zero current, so we can speak about zero current turning on (ZCS). But soft turning off of the rectifier switches is not ensured, so it must be resolved. The possible way to achieve this goal is to use some kind of snubber and also select suitable control algorithm for whole converter. For real design of DC/DC converter is ideal to develop the control algorithm which completely solves the described problem. It could be a very simple and convenient solution, but it is usually unreal or limited. So use some kind of snubber is the most realistic way to reach the desired condition. Therefore, design and function of the snubber is very important and the focus of research to this problem is required and important. The developed snubber can consists of passive parts like capacitor, inductor, diode, and also of active parts like transistor. Hence, final developed snubber can be passive or active. In general, passive circuit is simpler than active, but active circuit can be more effective and in many situations more convenient. It must be taken into consideration that control algorithm of the active snubber can significantly influence final function and effect. Therefore, correct combination of snubber and control algorithm is significant. So research and development of various combinations of control algorithms with various snubbers is required. By this way the various fully soft switching DC/DC converter topologies can be developed. Finally, the advantages and disadvantages of every solution should be evaluated. Into consideration range of load, topology of rectifier, conduction losses, complexity of snubber and number of additional control signals should be taken.

Acknowledgment

This work was supported by Slovak Research and Development Agency under project APVV-0185-10 and by Scientific Grant Agency of the Ministry of Education of Slovak Republic under the contract VEGA No. 1/0464/15.

References

[1] J. Dudrik, V. Ruscin, "Voltage fed zero-voltage zero-current switching PWM DC-DC converter," in *Proc. 13th Power Electronics and Motion Control Conference*, Poznan, 2008, pp. 295–300.

[2] J. Perduľak, M. Bodor, "Novel Zero-Voltage and Zero-Current Switching Full-Bridge PWM Converter Using Simple Secondary Active Clamp Circuit," in *Proc. 12th International PhD Workshop*, Owd, 2010, pp. 147–152.

[3] J. Dudrik, "DC-DC menič s riadeným výstupným usmerňovačom s jedným aktívnym spínačom," *EE časopis*, Vol. 20, No. 3, 2014, pp. 12–14.

[4] J. Dudrik, V. Ruscin, "ZVZCS PWM DC-DC converter with controlled output rectifier," *Acta Electrotechnica et Informatica*, Vol. 10, No. 1, 2010, pp. 12–17.

[5] R. Žatkovič, J. Dudrik, "Vysokofrekvenčný DC/DC menič s mäkkým spínaním pri širokom rozsahu zaťaženia," *EE časopis*, Vol. 20, No. 6, 2014, pp. 36–38.

[6] J. Dudrik, "Soft-switching PS-PWM DC-DC converter for arc welding," *Acta Electrotechnica et Informatica*, Vol. 6, No. 3, 2006, pp. 1–4.

[7] J. Dudrik, V. Ruščin, M. Bodor, D. N. Trip, P. Špánik, "Soft switching full-bridge PWM DC/DC converter using secondary snubber," *Journal of Electrical and Electronics Engineering*, Vol. 2, No. 1, 2009, pp. 147–150.

[8] J. Dudrik, M. Bodor, M. Pastor, "Soft-Switching Full-Bridge PWM DC–DC Converter With Controlled Output Rectifier and Secondary Energy Recovery Turn-Off Snubber," *IEEE Transactions on Power Electronics*, Vol. 29, No. 8, 2014, pp. 4116–4125.

[9] V. Ruščin, J. Dudrik, "Soft switching DC/DC converter using auxiliary circuits," *Journal of Electrical and Electronics Engineering*, Vol. 1, No. 1, 2008, pp. 242–245.

[10] J. Dudrik, M. Bodor, D. N. Trip, "Operation analysis of soft switching PWM DC-DC converter with secondary snubber," *Communications*, Vol. 15, No. 3, 2013, pp. 6–12.

[11] J. Dudrik, M. Bodor, "Novel soft switching DC/DC converter with controlled output rectifier," in *Proc. 15th International Power Electronics and Motion Control Conference*, Novi Sad, 2012, pp. DS1b.13-1–DS1b.13-5.

[12] J. Dudrik, M. Bodor, V. Ruscin, "Zero-voltage and zero-current switching DC-DC converter with controlled output rectifier," in *Proc. 14th International Power Electronics and Motion Control Conference*, Ohrid, 2010, pp. T1-39–T1-44.

[13] T. Mishima, K. Akamatsu, M. Nakaoka, "A High Frequency-Link Secondary-Side Phase-Shifted Full-Range Soft-Switching PWM DC–DC Converter With ZCS Active Rectifier for EV Battery Chargers," *IEEE Transactions on Power Electronics*, Vol. 28, No. 12, 2013, pp. 5758–5773.

[14] K. Suzuoka, S. Moisseev, L. Gamage, K. Soshin, K. Nishida, M. Nakaoka, "Boost transformer linked full bridge soft-commutation DC-DC power converter with secondary-side phase-shifted PWM rectifier switches," in *Proc. 29th Annual Conference of the IEEE Industrial Electronics Society*, 2003, pp. 49–54 vol. 1.

[15] T. Mishima, K. Akamatsu, M. Nakaoka, "A zero voltage and zero current soft switching PWM dc-dc converter with secondary-side phase-shifting active rectifier," in *Proc. Energy Conversion Congress and Exposition*, Raleigh, NC, 2012, pp. 2544–2551.

[16] M. A. Al, M. Nakamura, S. Sato, M. Nakaoka, "High frequency link transformer secondary side control DC-DC power converter," in *Proc. Power Conversion Conference*, Osaka, 2002, pp. 593–596 vol. 2.

[17] N. A. Ahmed, M. Miyatake, K. Nishimura, L. Hyun-Woo, M. Nakaoka, "Novel High Frequency Planner Transformer Linked Soft Switching DC-DC Power Converter with Secondary Side-Phase Shifted PWM Active Rectifying Switches," in *Proc. International Conference on Power Electronics and Drives Systems*, Kuala Lumpur, 2005, pp. 129–135.

[18] T. Mishima, K. Akamatsu, M. Nakaoka, "Analysis, design and evaluations of a soft switching PWM DC-DC converter with phase shift-controlled ZCS active hybrid rectifier," in *Proc. 38th Annual Conference on IEEE Industrial Electronics Society*, Montreal, QC, 2012, pp. 244–249.

Substrates and joints in power electronics

[1]Tomáš GIRAŠEK (1st year)
Supervisor: [2]Alena PIETRIKOVÁ

[1,] Dept. of Technologies in Electronics, FEI TU of Košice, Slovak Republic

[1]tomas.girasek@tuke.sk, [2]alena.pietrikova@tuke.sk

***Abstract*—This paper presents possibility of ceramic substrates use in power electronic as well as joints used in this area. Work offers an actual survey of power electronic and brings theoretical introduction as well as it offers description of ceramic, their important properties and basic requirements in this area. This paper detailed describes DBC substrates, which are the most often used in IGBT power electronic modules and it describes their benefits in comparison with organic substrates or substrates with metal core.**

***Keywords*— Power electronic, joints, ceramic substrates, thermal expansion**

I. Introduction

The thesis of my work is in study of material behavior in power electronics and to study and analyze of termomechanical thermal cycling in various power electronics modulus and measuring of their behavior and reliability. Power electronics are now driven by challenging requirements, such as volume, power density, operation temperature and cost. The general objectives of power electronics packaging are to distribute signal and power, dissipate the heat generated from the power chips, and protect he chips and circuits from mechanical damage. Power electronic packaging is critical to performance and reliability of power electronics system. For a high current, high voltage power module, it requires the paralleling of many insulated gate bipolar transistors (IGBTs) and diodes that are mounted on the substrates via large-area die attachment. Lead free solder, e. g. Sn3Ag0.5Cu (SAC305), is a conventional choice for lead-free power assembly. However, SAC305 could not be used at high temperature because of its low homologous temperature (low melting temperature). Therefore in power electronics have to be use other types both solders and joints that can operate at higher temperatures (more than 230°C) [1].

II. The substrates used in power electronics moduls

In power electronic modules are used several types of substrates. On these substrates are placed several requirement. The basic requirements on power electronic substrates are excellent thermal conductivity, suitable thermal expansion, high electrical resistivity, resistance to power cycling and high temperature resistance. All of these requirements are very important when choosing substrate in the end application.

These substrates could be divided to three categories:

1. Organic substrates – glass fiber reinforced epoxy substrates FR4 with thin cooper layers usually used in soft electronic [12].
2. Metal substrates – substrates with metal core and dielectric layers on surface (aluminum PCB, cooper PCB) [10], [11].
3. Ceramic substrates – frequently used DBC (Direct Bond Copper) [9] and DBA (Direct Bond Aluminum) substrates [26].

Circuit boards are usually fabricated from glass fiber reinforced epoxy printed boards (FR4) with cooper layers. Their disadvantage are low thermal conductivity (0.25 – 0.5 W/m.K) and low used temperature (just 150 °C) [23] there are organic substrates with higher working temperature (on based polyimide's or with use Teflon), but they are too expensive. However, for high power applications, large currents and high voltages demanded by the applications require circuit boards that can carry high current loads and efficiently dissipate large amounts of heat [2]. The substrates with metal core have these properties. Cooper that is used as core in these substrates has excellent thermal conductivity (400 W/m.K) [5]. These cooper substrates have on the surface dielectric layer on which are placed circuit layers. Whereas employing large amount of copper and so many steps in production, production these substrates is too expensive, therefore manufacturer's substitutes metal core substrates with ceramic materials. In my future work we will be use ceramic substrates; therefore they are detailed described in next lines.

DBC (Direct Bond Copper) denotes a process in which copper and ceramic material are directly bonded. Normally, DBC has two layers of copper that are directly bonded onto an aluminum-oxide (Al_2O_3) or aluminum-nitride (AlN) [15]. In specific cases can also be used ceramic silicon nitride (Si3N4) [18], [21], beryllium oxide (BeO) [13], silicon carbide (SiC) [17] and boron nitride (BN) [17]. Similarly than DBC are used in power electronic also DBA substrates. They have ceramic base (Al2O3, AlN, BN, BeO, SiC) but on the surface is thin layer of aluminum. They used there where are used a lot of wire bonding, because soldering of aluminum is very difficult.

While any oxide can be used for DBC, most applications use electrically insulating substrates that are thermally conductive [19]. Beryllium oxide has the highest thermal conductivity (260 W/mK) of the ceramics but it is not commonly used because of beryllium toxicity [13]. Aluminum oxide is commonly used despite a moderate thermal conductivity (~ 250 W/mK) because of its low cost, high substrate quality, and high bond strength. Aluminum nitride is also used because of its high thermal conductivity (170 W/mK) and a coefficient of thermal expansion (~4.7 µm/m°C) that is closely matched to silicon [3], [14].

The DBC process yields a super thin base and eliminates the need for the thick, heavy copper bases that were used prior to this process. Because power modules with DBC bases have fewer layers, they have much lower thermal resistance values and because the expansion coefficients matches silicon, they have much better power cycling capabilities (up to 50 000 cycles) [2].

Properties of DBC ceramic substrates (Tab. 1):
1. Good mechanical strength, mechanically stable shape, good adhesion and corrosion resistant.
2. Excellent electrical insulation.
3. Very good thermal conductivity.
4. Environmentally clean.
5. The thermal expansion coefficient is close to that of silicon, so no interface layers are required.
6. Good heat spreading.
7. May be structured just printed circuit boards.
8. The 0.3 mm thick layer permits higher current loading for the same conductor width. Assuming the same cooper cross-section the conductors needs to be only 12% of that of a normal printed circuit board.
9. The excellent thermal conductivity provides the possibility of very close packaging of the chips. This translates into more power per unit of volume and improved reliability of system and equipment.
10. DBC ceramic is the basis for the "chip-on-board" technology which represents the packaging trend for the future.

TABLE I:
BASIC PROPERTIES OF DBC SUBSTRATES [2], [15], [18], [19].

Ceramic	Aluminum Oxide (Al_2O_3)	Aluminum Nitride (AlN)
Purity	≥96%	≥97%
Dielectric strength	10kV/mm	~14kV/mm
Electrical resistivity	$>10^{14}$ Ω/cm	$>10^{14}$ Ω/cm
Thermal conductivity	24-28 W/mK	≥150 W/mK
Thermal expansion coefficient	$7.4 .10^{-6}$ K^{-1} (50 – 200 °C)	$5. 10^{-6}$ K^{-1} (25 – 500 °C)
Thickness	0.63; 0.38; 0.25 mm	0.63 mm
Surface finish	Cu or (electroless) nickel plated Cu	Cu or (electroless) nickel plated Cu
Cu thickness	Standard 0.3 mm	Standard 0.3 mm
Ni thickness	Standard 7 µm	Standard 7 µm
Application temperature (inert atmosphere)	-55 to 850°C	-55 to 850°C

One of the basic parameters of power electronic substrates is the coefficient of thermal expansion (CTE). CTE is also critical to power modules reliability. Thermal expansion is the tendency of materials volume a change in response to temperature change. The CTE for all materials described above is measured in [10^{-6} K^{-1}]. Copper, for example, has a higher CTE than silicon. Given temperature increase for both materials, cooper expands about six times more than silicon. Table 1 shows key materials used in DBC substrates [4].

The CTE of Al_2O_3 is closer to silicon than copper. The CTE of AlN is also closer to that of Si, which reduces stresses in die attach materials (Fig. 1). However, the stresses are actually higher in the joint between the cooper baseplate and DBC because of the greater difference between the net CTE of the Al_2O_3 DBC and of the Cu baseplate. This causes the power module to bend further. This difference causes thermal stresses in the devices, solder interconnections, and substrates because the mismatches are frozen during the assembly process of the module at high temperatures, especially during the soldering process. These stresses can cause mechanical and fatigue failure or changes in operating behavior [4].

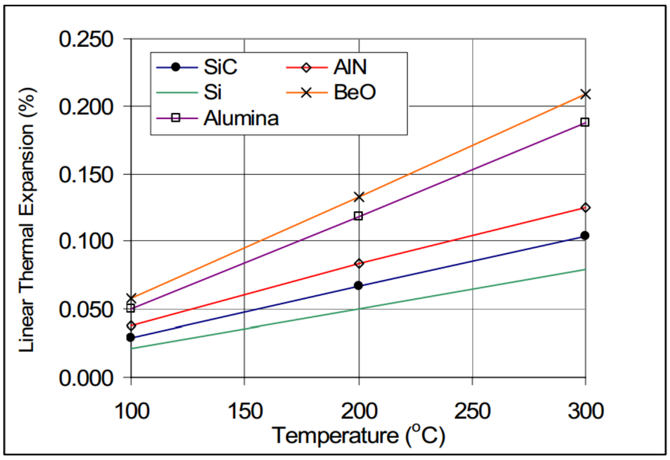


Figure 1: Thermal expansion of common ceramics and device materials vs. temperature [5].

DBC ceramic substrates are the base materials of future for both the construction and the interconnection techniques of electronic circuits. They will be employed as base material for electronic components with high values of power dissipation and demanding requirements concerning their thermal shock behavior as well as their failure rate, whenever normal printed circuit boards are no longer adequate. Typical concept of modules used in power electronic is shown in the Fig. 2.

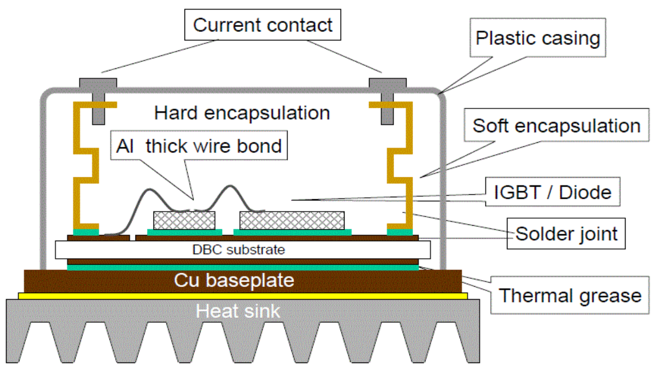


Figure 2: Principal design of power module with DBC substrate [8].

Example use DBC substrates:
1. Power hybrids and power control circuits,
2. power semiconductors,
3. smart power building blocks,
4. electronics heating devices.

III. JOINTS AND BONDS IN POWER ELECTRONICS

Similar to the packaging of other electronic devices, power electronic packaging is also the first level of assembly toward building the power electronic system or an end product. The power electronic packaging mainly involves die attachment, interconnection between die pads and the leadframe or substrate and encapsulation. The particular operating mechanisms and their intended purposes present some challenging requirements [6], [8].

First of all, the joints needs to ensure desirable circuit performance and efficiency by minimizing parasitic effects (resistance, inductance, and capacitance), especially at high switching frequencies. Secondly, the large current and high voltage of power joints requires the packaging materials and architecture to possess not only high current carrying capability but also excellent electrical insulating properties. Thirdly, the increasingly dense integration of power devices forces its packaging to have efficient heat dissipation. Finally, the reliability of power electronic packaging becomes relatively important due to high-power conversion and the accompanying loss. Typically, the following interconnection methods are employed in power electronics packaging [6], [7].

A. *Wire bonding*

With its mature technology, wire bonding was initially widely used and still dominates in electronic power packaging because it can effortlessly accommodate design changes in packaging and make use of existing infrastructure. Moreover, the reliability of wire bonding has been proved. As with the multi-chip wire bonding module, the multiple different switching dies are soldered to a direct-bond copper (DBC) or direct-bond aluminum (DBA) substrate. To increase the current-carrying capability, multiple aluminum wires ranging from 125μm to 625μm are ultrasonically bonded from one pad to a surrounding conductive trace, as shown in the Figure 3. In practice, some disadvantages of wire bonding interconnection in power electronic packaging become obvious, especially for high-power and high-frequency applications. For instance, the thin, long wire leads to large electrical resistance of the entire packaging. Correspondingly, parasitic inductance and thermal resistance also increase [6], [20].

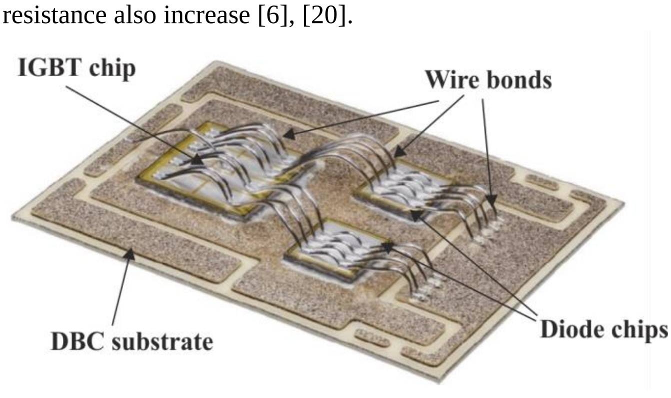


Figure 3: Typical wirebonding application [15].

B. *Ribbon bonding*

Ribbon is an alternative to bondwire for use in power electronic packaging to reduce electrical resistance and thus improve current handling capability. Compared with bondwire, the larger cross-section area of ribbon enables it to possess lower electrical resistance and thus carry larger current. Some wedge wire bonders can be accommodated to bond a ribbon. Additional benefits provided by the use of ribbon instead of wire include improved reliability due to a larger cross-section at the heel of the bond, less heel cracking due to the lower loop profile, less cratering because the bond force and ultrasonic energy are evenly distributed over a larger area and there is no wire sway because of the structural rigidity of flat ribbon [6], [24].

C. *Solder joints*

Solder alloys, which is mainly used to attach the die to the substrate, can work as an interconnect between the top die pad and the leadframe or substrate. The resistance between the source or gate and the top leadframe was significantly reduced compared to that of wirebonding. Additionally, the heat dissipation was dramatically improved due to an enlarged contact area between the leadframe and the gate and source terminal pads. However, reliability became a big concern due to the large coefficient thermal expansion (CTE) mismatch between the copper leadframe and the silicon die [4], [6], [8].

In solder alloys is one of basic properties homologous temperature, because this parameter determines to which temperature solder may be used. The principle is that maximum operating temperature should be expressed as a fraction of the metal’s melting point (in K). Seen from the perspective of operational strength, consideration of possible fatigue failure applies to the design of metal materials if the homologous temperature T_{hom} ($T_{hom} = T_{use}/T_{melting}$) is above 0.4. Most metals are used in the range of 0.3 – 0.5 but solders can be up to 0.87. To add a little more leniency we can use a ratio of 0.9 [24].

The use of a high temperature solder preform is proposed and demonstrated for use as a die-attach material in high power devices. Solder preforms alloys can be indium-contained, gold-contained, lead-free, fusible or standard tin-lead, as well as many others. Solder Preforms are used in a variety of applications that require precise amounts of solder. Preforms come in standard shapes such as squares, rectangles, washers and discs. Typical sizes range from 0.254 mm up to 50.8 mm. Smaller and larger sizes, as well as custom shapes, are also available. Dimensions can be held to tight tolerances to assure volume accuracy [23].

D. *Sintered silver*

The high junction temperature of SiC devices and the hard operating environment mean their packaging must withstand a high operating temperature. For example, some SiC power devices can work at up to 175°C or even 250°C, but at 250°C, most solders would melt. Moreover, when the working temperature is close to the melting temperature, the long-term reliability of the device is compromised. To meet high power and high temperature requirements, Semikron employed sintered silver as interconnect material because silver, with a melting point of 961°C, which is far beyond operating temperatures, is substantially cheaper than gold and palladium, is not susceptible to oxidation like other metals, and has significantly better electrical an thermal properties [7]. Moreover, sintered silver is more reliable than bondwire and solder alloy since no intermetallic compound is involved in the bonding microstructure. However, a high sintering temperature of 600°C and a long processing time become an obstacle to the introduction of sintered silver into manufacturing. To facilitate sintering, assisted-pressure is applied during processing. Figure 4. shows sintered silver interconnection in a power module carried out by Semikron [6], [7].

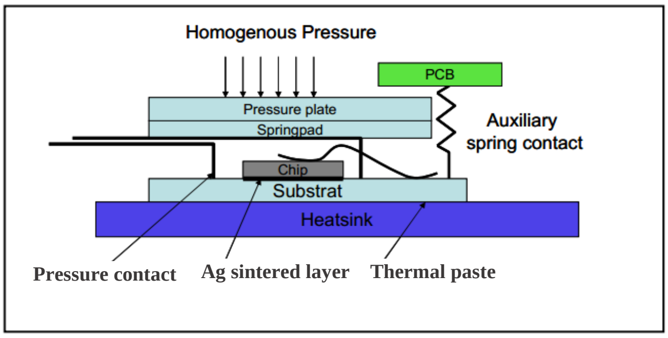


Figure 4: Cross section of sintered modules [7].

E. Press pack

To eliminate parasitic effects caused by bondwire and to relieve concerns about reliability of a solder interconnect or metal deposition, press-pack technology emerged as an effective interconnect method to not only handle high voltage and current but also to have improved reliability. In press-pack packaging, the contact to the die is made only through externally applied pressure, which eliminates the need for wire and substrate bonds, thus minimizing the stress and associated lifetime reduction factors for the die. Individual die gate terminals are contacted via sprung pin, which is commonly connected to the external gate terminal via a planar distribution board, which is carefully configured to ensure series impedance to the die and therefore good homogeneous switching. However, the expensive manufacturing cost involved with precision machining and planarization limits the press-pack to high power packaging applications [8], [22].

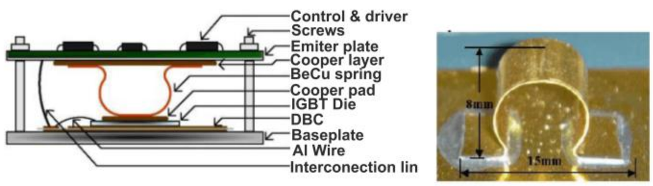


Figure 5: The cross section of the pressure contact cell [22].

IV. Conclusion

This paper offers an overview of the possibilities use of ceramic materials which ceramic substrates in the power electronic. Work describes the main advantages ceramic substrates such as DBC for this area. Furthermore, the paper describes other types substrates (organic substrates or substrates with metal core) and their limitation such as low use temperature, low thermal conductivity or high price. Paper describes the importance of knowledge the basic parameters of substrates. It is also very important to know the parameter of thermal conductivity, thermal expansion, used temperature as well as volume resistivity. DBC ceramic substrates are the base materials of future for both the construction and the interconnection techniques of electronic circuits. They will be employed as base material for electronic components with high values of power dissipation and demanding requirements concerning their thermal shock behavior as well as their failure rate, whenever normal printed circuit boards are no longer adequate. Because power modules with DBC bases have fewer layers, they have much lower thermal resistance values and because the expansion coefficients matches silicon, they have much better power cycling capabilities.

My future work will be focus on development of new joints, which will have high melting temperature and high thermal cycling resistance. I will be study they behavior these joints will be used in power electronic on the connection silicon chips with DBC substrates. One of aims of my future work is based on study heat distribution and his simulation in power modules and to study and analyze of termomechanical thermal cycling in various power electronics modulus and measuring of their behavior and reliability.

References

[1] Yunhui Mei, "Bonding 1200 V, 150 A IGBT Chips (13.5 mm × 13.5 mm) with DBC Substrate by Pressureless Sintering Nanosilver Paste for Power Electronic Packaging," University Tianjin China, 2014.

[2] Liu, Xingsheng, "Processing and Reliability Assessment of Solder Joint Interconnection for Power Chips," China, 2010.

[3] Matthew B. Squires, James A. Stickney, „Atom chips on direct bonded copper substrates," USA, 2010.

[4] Patrick Baginski, "Modules without a baseplate: A reliable and cost-effective solution," USA, 2011.

[5] Mitch Olszewski, "High-temperature high-power packaging techniques for her traction applications," Washington, 2007.

[6] Naili Yue, "Planar Packaging and Electrical Characterization of High Temperature SiC Power Electronic Devices," Virginia, 2008.

[7] C. Göbl, "Low temperature sinter technology Die attachment for power electronic applications," Nuremberg/Germany, 2010, ISBN 978-3-8007-3212-8

[8] Torrey hills technologies, "Review of Power Electronics Packaging and the Use of Copper Tungsten as Heat Sink," San Diego, 2010

[9] Roger corporation Curamic®, Ceramic substrates, Technical data sheet, 2015.

[10] PCB-pool®, Insulated Metallic Substrate PCB, Technical data sheet, 2011.

[11] Multi-CB®, Metal core circuit boards, Technical data sheet, 2015.

[12] VENTEC®, VT-42 Laminate/Prepreg, Technical data sheet, 2014.

[13] Brush Wellman, "A Qualitative Overview of the Use of Beryllium, Beryllium-Containing Alloys and Beryllium oxide Ceramic in Electrical and Electronic Equipment (EEE)," February 28, 2008.

[14] Michael Kettner, Patrick Biebersmith, "Aluminum Nitride vs. Beryllium Oxide for High Power Resistor Products," Florida, 2001.

[15] Murawa®, Aluminum Nitride Substrates, Technical data sheet, 2015.

[16] Martin Kachlík, "Příprava Al2A3 keramiky se zvýšenými mechanickými vlastnostmi," Brno, 2010.

[17] TYGRE®, "Production of a high value ceramic material: Silicon carbide," 2015

[18] Hnatko M., "Nitrid kremíka (Si3N4)," Ústav anorganickej chémie SAV 2006

[19] Douglas C. Hopkins, "Excerpt – Direct Bonded Copper," Buffalo, 2003

[20] Yunhui Mei, "Bonding 1200 V, 150 A IGBT Chips (13.5 mm × 13.5 mm) with DBC Substrate by Pressureless Sintering Nanosilver Paste for Power Electronic Packaging," University Tianjin China, 2014.

[21] Manfred Goetz, Bernd Lehmeier, "Silicon Nitride Substrates for Power Electronics," Germany, 2012

[22] Xiaoyu He, Xiangjun Zeng, "A Hybrid Integrated Power Electronic Module Based On Pressure Contact Technology," China, 2006, ISBN 0-7803-9716-9.

[23] Tunhøvd Haugan, "Characterization of the material properties of two FR4 printed circuit board laminates," ISBN 978-82-464-2322-7, 2014

[24] Indium corporation®, Technical documents, 2015

[25] Öko-Institute, "Temperature Load Measurements and Specification," 2010

[26] A. A. Wereszczak, "Low-Cost Direct Bonded Aluminum (DBA) Substrates," USA, 14 May, 2013

Technologies for the Next Generation of Wireless Communication Systems

[1]*Lukáš SENDREI (3^{rd} year),*
Supervisor: [2]*Stanislav MARCHEVSKÝ*

[1,2]Dept. of Electronics and Multimedia Communications, FEI TU of Košice, Slovak Republic

[1]lukas.sendrei@tuke.sk, [2]stanislav.marchevsky@tuke.sk

***Abstract*—The demand on high data rate, low latency, high spectral efficiency and low power consumption creates an ongoing pressure on the development of communication systems. Although cellular technologies like Long-term evolution and Long-term evolution - Advanced can face many of these demanding challenges the dramatic increase of mobile data transfer in the last years gives rise to design even more sophisticated methods. Multi-carrier modulation formats retain great potential to meet the requirements of the next generation 5G systems. The presented paper is an overview of the research activities in the field of multicarrier communication that have been done during the last term. It also summarizes the ongoing work of the research projects in which the author participate.**

***Keywords*—5G, GFDM, OFDM, wireless technologies**

I. Research activities

Wireless communication has become a key element of modern telecommunication industry. Recently, mobile technologies of the *4-th generation* denoted as Long-Term Evolution - Advanced (LTE-A) are deployed in many countries. These systems allow high data throughput thanks to the application of Orthogonal Frequency Division Multiplexing (OFDM) along with Multiple-Input Multiple-Output (MIMO) techniques. OFDM is widely adopted because of its favorable features like simple implementation based on Fast Fourier Transforms (FFT) and robustness against fading channels. However, the requirements of particular applications scenarios foreseen for 5G might not been coped by OFDM. Along standard *bitpipe* transmission, scenarios like *machine to machine communication* (MTM) and *Wireless Regional Area Network* (WRAN) are in the field of interest for 5G networks. The MTM communication requires extreme low power consumption which can cause serious impact on the synchronization process, thus it is not possible to preserve the orthogonality between individual sub-carriers. Although the cyclic prefix (CP) is a powerful tool to overcome the issues related to multipath fading it is responsible for low spectrum efficiency of OFDM while applying in WRAN. Especially, the high out-of-band (OOB) radiation of OFDM limits the utilization in opportunistic and dynamic spectrum access. Therefore, the focus of recent research efforts is dedicated to alternative physical layer technologies. One of the most discussed is Filter Bank Multi-Carrier (FBMC) [1]. FBMC uses well designed filter banks to shape the individual subcarriers, thus the OOB radiation is kept extremely low. Owing to the suitable filter design in the frequency-time domain, no CP is used, hence extremely high spectral efficiency is achieved. On the other hand, the application of MIMO is not straightforward and the equalization process for rapid time varying channels is more challenging. Generalized Frequency Division Multiplexing (GFDM), proposed in [2], is an another possible approach to meet the requirements on flexible modulation technique. Here, the data symbols are proceeded in frequency-time block manner and pulse shaping is performed per subcarrier. On the one hand, pulse shaping enables to control the OOB and Peak-to-Average Power Ratio (PAPR), on the other hand it causes self-interference, which needs to be compensated, e.g. on the receiver side using interference cancellation technique. Hence, all these properties are making GFDM an attractive choice for deployment in 5G communication systems.

Since GFDM is a sum of pulse shaped subcarriers it suffers from high signal envelope fluctuations, thus the nonlinear characteristics of the high power amplifier (HPA) causes major performance degradations. Several PAPR reduction techniques, to overcome the issues of nonlinearities, have been introduced in the literature. In the paper [3], the clipping technique has been investigated as a PAPR reduction scheme in GFDM. The results have shown that GFDM can in special case outperform OFDM. To the authors best knowledge, there is no paper dealing with nonlinear amplification in GFDM systems. Therefore we have investigated the impact of nonlinear amplification on GFDM systems and analyze the error probability performance. The simulation results show that the particular HPA model and its transmission parameters significantly influences the bit error rate (BER). All these and many other results will be published in upcoming international conferences and scientific journals.

II. Projects

A. The research of coexistence between broadband LTE networks and digital terrestrial TV broadcasting DVB-T/DVB-T2

The project deals with the research of coexistence between broadband LTE800 and digital DVB-T/-T2 terrestrial TV broadcasting in the Slovak Republic. Part of the project will be focused on research of mutual influence of radio communications systems that were invented only a few years ago and even some of their new additions are currently under development.

The main objectives of the project are:

- Identification of areas of the Slovak Republic, which may result in degradation of the reception of DVB-T/-T2 signal as a result of the introduction of LTE800 transmission.

- Establishment the functional public information system to identify the risk of interference of terrestrial DVB-T/-T2 system.
- Proposals on measures to prevent unwanted interactions between LTE800 and DVB-T/-T2 systems and solutions leading to elimination of problems in practical operation.

Up to the present no research in Slovakia in terms of compatibility of broadcasting DVB T/ T2 and LTE800 systems with a focus on assessing the impact on the population of the Slovak Republic was realized, which could be used in the design, implementation and real service of these networks, and therefore the project is clearly genuine.

- Definition of the resident groups and areas within Slovak Republic that will be affected by interference between LTE800 and DVB-T/-T2, which may result in the degradation of terrestrial TV reception.
- Establishment and introduction of functional information system on the internet designated for the general public, telecommunication and radio communication operators containing information of possible interference of terrestrial DVB T/ T2 in specific locations within Slovakia and processing of related documentation.
- The design and application of processes and policies leading to the suppression of interference between LTE800 and DVB-T/-T2 systems to eliminate adverse impacts of unwanted interference of the radio communication systems in particular cases of real operation.

The procedures of project realization will lead to the original application outcomes, including in particular:

During the last term the cooperation of this project resulted in a publication [4], which was presented on the conference *NoTeS 2014*.

B. Interactive Multiview Video Streaming for Supporting Education

The project is oriented to development and application of accessible software and hardware for interactive streaming of multimedia content, without necessity to save it in local data store. The aim of project is to elaborate and realize methods of access to various formats of multimedia content, such as video, sound record, and access to web cameras array with multiple visual angles (multi-view video streaming), as well. It is planned to design multimedia educational content for exploitation of such access. That all will be dedicated for supporting teaching in 1-st to 3-rd degrees of university education in the frame of telecommunication subjects as follows: switching technology, digital television, satellite technologies and services, etc. Access to learning materials mentioned above will serve naturally to support distance education. In this [5], a multi-view e-learning system for supporting remote education is introduced. Since video is the most natural way to perceive information and of interaction between people, it also has inherent capabilities when using in education environment. The system, we propose, should encourage students more use the opportunities offered by remote learning to foster their training on the university. To enable multi-view video (MVV) streaming a system with three cameras is used. A central server is used to capture the partial streams and it also serves as the control unit to manage the whole system. The student will have the choice to decide which particular part of the captured scene he/she wants to watch. Another option will be a guided stream, where the streams are switched automatically, based on the area where the presentation is currently focused. We believe that this model might contribute to the educational process.

C. Agent based modeling of the spectrum distribution in cognitive radio networks

The goal of the project is the design of the novel and efficient models of the spectrum sharing and trading mechanisms in the cognitive radio networks. In the paper [6], we propose an agent-based model for spectrum trading in the shared use model of dynamic spectrum access. Spectrum trading is employed using the single-unit sealed-bid first-price auction, which takes into the account risk caused by the imperfect spectrum sensing. Bidding strategies of the bidder are controlled by the reinforcement learning algorithm. Cooperative energy-based spectrum sensing is used as a spectrum sensing mechanism. Two different decision fusion strategies, which provide different levels of risk are discussed. The results demonstrate that in risky environment, total revenue and total payoff of the auctioneer and bidder respectively is higher, than in the case of system with lower level of risk. On the other hand, normalized revenue and payoff per a single auction round is higher in the case with lower level of risk. Moreover, the results have shown that the optimum sensing time for maximizing revenue and payoff is different.

III. Conclusion

This paper analyses the scientific work of the author what have been done during the last term. 5G wireless communication are in great interest of research activities in the last time. The important role of mulcirarrier technologies will have an important role also in the upcoming communication systems. Therefore, the focus in this field of study has already brought some interesting results and we can courageously state that it will bring more promising result and application in the following years

Acknowledgment

This work was supported by the Slovak Cultural and Educational Grant Agency (KEGA) under contract No. 062TUKE-4/2014.

References

[1] B. Farhang-Boroujeny, "OFDM Versus Filter Bank Multicarrier," *Signal Processing Magazine, IEEE*, vol. 28, no. 3, pp. 92–112, May 2011.

[2] G. Fettweis, M. Krondorf, and S. Bittner, "GFDM - Generalized Frequency Division Multiplexing," in *Vehicular Technology Conference, 2009. VTC Spring 2009. IEEE 69th*, April 2009, pp. 1–4.

[3] L. Sendrei, S. Marchevsky, N. Michailow, and G. Fettweis, "Iterative receiver for clipped GFDM signals," in *Radioelektronika (RADIOELEKTRONIKA), 2014 24th International Conference*, April 2014, pp. 1–4.

[4] M. Felix, L. Sendrei, and S. Marchevsky, "Výskum koexistencie širokopsmových sietí LTE a pozemského digitálneho TV vysielania DVB-T DVB-T2," in *Nová technika a služby v telekomunikáciách ČR a SR 2014 : Zborník 8. medzinárodnej konferencie*, Jun 2014, pp. 1–27.

[5] J. Valiska, L. Sendrei, M. Ludmila, and S. Marchevsky, "A Multiview E-learning System for Remote Education," in *Emerging eLearning Technologies and Applications (ICETA), 2013 IEEE 11th International Conference on*, Dec 2014, pp. 489–494.

[6] J. Pastircak, L. Sendrei, S. Marchevsky, and J. Gazda, "An agent-based model of the risk-based spectrum auction in the Cognitive Radio Networks," in *2014 1st International Conference on 5G for Ubiquitous Connectivity (5GU)*, Nov 2014, pp. 23–28.

The Assessment of Dielectric Parameters of XLPE Insulation Using Dielectric Relaxation Spectroscopy

[1]*Martin GERMAN-SOBEK (4th year)*
Supervisor: [2]Roman CIMBALA

[1,2]Dept. of Electric Power Engineering, FEI TU of Košice, Slovak Republic

[1]martin.german-sobek@tuke.sk, [2]roman.cimbala@tuke.sk

***Abstract*—This paper deals with the diagnostic of XLPE insulation. Many studies and experiments show an influence of ageing and moisture to degradation of XLPE insulation. It is subsequently reflected by changing of dielectric parameters of insulation. The dielectric parameters of XLPE cable samples were measured using dielectric relaxation spectroscopy method (DRS) in time and frequency domain. The results are compared and it was observed a change of parameters in consequence of additional ageing and influence of moisture.**

***Keywords*—aging, capacitance, dielectric relaxation spectroscopy, dissipation factor, moisture, polarization current, XLPE cable.**

I. Introduction

The power cables play an important role in the safety of the power load and reliable transmission of electricity and they are important devices in the power system. Therefore, the cable insulation system, but also insulation system of any high voltage equipment, is very important and sensitive part. Damage of insulation can lead to equipment failure and other disorders.

Nowadays, polymeric materials are widely used as insulation system of power cables. Cross-linked polyethylene (XLPE) is widely used as electrical insulation material for high-voltage distribution power cables. XLPE is characterized by excellent physical, chemical and electrical properties. The power cables may be exposed to high currents and voltages and they are critical parts of the transmission infrastructure. Therefore, it is expected their high resistance against possible failures.

The motivation for the solving these research is importance of modern operational diagnostic and systematic maintenance, which can ensure the longest possible operation of electrical device, with a maximum performance, quality and reliability of operation.

II. Aging of XLPE Insulation

The insulation degradation is inevitable during the operation and the failure rate increases with the service time [5]. During operation, the XLPE insulation of power cables may be exposed to high currents, overvoltage, mechanical and chemical stress and pollution of environment. Interaction various factors together may significantly speed up the degradation processes. The cables also are permanently exposed to thermal aging. It may cause change in dielectric parameters of cables and also irreversible damage of cable insulation. In most cases failures in the insulation are related to moisture and integral degradation of the insulation like water treeing in XLPE cables. The insulation failures may be also caused by lower dielectric strength due to aging processes or by internal defects in the insulation system. Many studies have been performed concerning water trees in XLPE based materials. [4], [5], [6]

It is necessary to assess degradation and insulating state of cables, since process of aging of insulation is the most acting on the quality of insulation and it is a phenomenon that is essentially cannot be affected. XLPE insulated cables for high voltage applications and XLPE cable ageing have been studied and investigated for nearly 45 years in order to evaluate a function of service stresses and aging time, and in order to improve dielectric performance of XLPE material. Many methods have been proposed to evaluate the properties of XLPE. One of the methods for detection of insulating state of XLPE insulation is measurements of dielectric parameters of insulation using the dielectric relaxation spectroscopy (DRS). This method is a non-destructive method. [2], [3]

III. Task Solved in Previous Year

My research is focused on measurements and investigation insulating state and quality of solid insulation material, specifically XLPE, using measurement of dielectric parameters by the method DRS in time and frequency domain. This method is widely applied in the characterization of ion-conducting solids and polymers and uses polarization as the response of the sample on a time-dependent electric field. The principle of this method is based on examination of molecular dynamics of polarized and polar materials. [1]

I evaluate the time response using measurement of the polarization and depolarization current. The total polarizing current flowing through the dielectric can be following assumption of existence the independent Debye's polarization processes and Maxwell-Wagner equivalent model, expressed as the sum of currents with exponentially decreasing amplitude I_{mi} with definite-time component τ_i. These components represent the polarization occurring in the material. Debye's polarization processes are exponentially decreasing in time. Number of members of expansion n

depends on the structure and homogeneity of the dielectric and on the time at which the polarizing action are measured [1]. The parameters I_{mi}, τ_i are correlated with the material properties. For analyze the quality of materials is an important analysis of these parameters and it is necessary to know these values and their development of previous measurements.

In the frequency domain, I measured the frequency dependence of dielectric dissipation loss factor *tanδ* and complex capacitance *C(ω)*. Using these parameters it is possible to evaluate the state and quality of insulating materials. [1]

The experimental measurements were performed on several samples of power cables that have been degraded - operationally unknown technical condition, or new - operationally unaged and undamaged. Lengths of the samples were different. Using the electrometer Keithley 617, and Keithley 6517B were recorded polarization and depolarization currents depending on time during the charging period 1000s. Using the precision LCR meter Agilent E4980A were recorded changes of capacitance and dielectric dissipation loss factor depending on frequency in range from 20 Hz to 2 MHz.

The samples were subjected to accelerated thermal aging at 90 °C in a laboratory air oven for a specified time and all measurements of dielectric parameters were repeated. The aging time always was doubled (24 h, 48 h, 96 h and so on) and the overall aging time was 1512 h. There also were carried out thermal aging of samples at the maximum operating temperature of 105 °C for 216 h. Other samples were exposed to moisture and were used three different solutions. The obtained data and time constants τ are continuously analyzed and compared, and it is monitored the change of insulation state due to aging and moisture.

Detailed results and conclusions are presented in my publications [7], [8], [9], [10], [11], [12], [13], some of which are in the SCOPUS database, or WoS.

I also cooperated on significant publication [14], [15]. I cooperate with my colleagues on various others experiments. In the present I also would like to publish my experiment result in journal which is referred by Current Contents.

IV. Conclusion

The dielectric properties and electrical parameters of insulation are dependent on a number of factors that affect to the insulation during operation. It is well known that aging of insulation influences on the parameters and quality of insulation. Using suitable measuring method can be evaluated change of properties of investigated insulation material. Further research could be directed to improving the properties of XLPE insulation for example by means of appropriate nanocomposites and addition agents, and also for the improvement of diagnostic methods.

The aim of my work is diagnostic of insulation materials using dielectric spectroscopy. In the near future I will evaluate and publish research results and my aim is to successfully complete graduate degree.

Acknowledgment

This paper was developed with support of operating program Research and development for the project: "Univerzitný vedecký park Technicom pre inovačné aplikácie s podporou znalostných technológií" (University Science Park Technicom for innovative applications with support of knowledge technologies), code ITMS: 26220220182, co-financed from European funds.

References

[1] Zaengl, W. S.: “Dielectric Spectroscopy in Time Domain for HPV Power Equipment, Part I: Theoretical Consideration,” IEE Transaction on Dielectrics Electrical Insulation, Vol. 19, No. 5, 2003, pp. 5-19.

[2] Jianying, L. et al.: The Effect of Accelerated Water Tree Ageing on the Properties of XLPE Cable Insula-tion. In: IEEE Transaction on Dielectrics and Electrical Insulation, Vol. 18, 2011, pp. 1562-1569

[3] Fei Liu et al.: Insulation ageing diagnosis of XLPE power cables under service conditions. In: Condition Monitoring and Diagnosis (CMD), 2012 International Conference, 2012, pp.647-650.

[4] Hoff, G., Kranz, H. G.: “Interpretation of Dielectric Response Measurement Data from Service Aged XLPE-Cables,” IEEE 7th International Conference on Solid Dielectrics, Eindhoven, June 25-29, 2001, pp. 381-384.

[5] Fothergrill, J. C., et al.: “The Measurement of Very Low Conductivity and Dielectric Loss in XLPE Cables,” IEE Transaction on Dielectrics Electrical Insulation, Vol. 15, No. 5, 2011, pp. 1544-1553.

[6] Shuvalov, M. et al.: Analysis of Water Trees in Power Cable Polymeric Insulation. In: Journal Applied Polymer Science, Vol.88, 2003, pp. 1543-1549.

[7] German-Sobek M., Cimbala R., Király J.: Change of dielectric parameters of XLPE cable due to thermal aging. In: EEA - Electrotehnica, Electronica, Automatica. Vol. 62, no. 3 (2014), p. 47-53. - ISSN 1582-5175

[8] Molnár L., German-Sobek M., Cimbala R.: Sledovanie vplyvu prívodných vodičov meracej aparatúry na meranie dielektrickej relaxačnej spektroskopie. In: Starnutie elektroizolačných systémov. Roč. 9, č. 1 (2014), s. 15-18. - ISSN 1337-0103

[9] German-Sobek M., et al.: Ageing assesment of XLE cable using dielectric relaxation spectroscopy. In: Current Problems of Maintenance of Electrical Equipment and Management. - Košice : TU, 2014 S. 41-47. - ISBN 978-80-553-1818-9

[10] Cimbala R., Király J., German-Sobek M., Pavlík M.: Dielectric Spectroscopy of Transformer Oil Based Ferrofluid from View of i(t) Characteristics at Initial stage of ageing. In: Electric Power Engineering 2014 : Proccedings of the 15th International Scientific Conference : May 12-14, 2014, Brno, Czech Republic. - Brno : Brno University of Technology, 2014 P. 467-472. - ISBN 978-1-4799-3806-3

[11] Cimbala R., Király J., German-Sobek M.: Ageing of Transformer Ferrofluid Oil as to Capacitance in Frequency Domain at Initial Stage of Ageing. In: EEA - Electrotehnica, Electronica, Automatica. Vol. 62, no. 3 (2014), p. 90-95. - ISSN 1582-5175

[12] German-Sobek M., Cimbala R., Király J.: Influence of ageing and water treeing to degradation of XLPE insulation. In: Starnutie elektroizolačných systémov. Roč. 8, č. 2 (2013), s. 26-30. - ISSN 1337-0103

[13] German-Sobek M. [et al.]: - The effect of thermal aging on dielectric parameters of XLPE cable insulation. In: Elektroenergetika 2013 : proceedings of the 7th International Scientific Symposium on Electrical Power Engineering : September 18-20, 2013, Stará Lesná, Slovak Republic. - Košice : TU, FEI, 2013 S. 359-362. - ISBN 978-80-553-1441-9

[14] Marton K. [et al.]: Magnetic fluid in ionizing electric field. In: Journal of Electrostatics. Vol. 71, no. 3 (2013), p. 467-470. - ISSN 0304-3886

[15] Cimbala R., et al.: Polarizačné procesy v magnetických kvapalinách. In: Chemické listy Vol. 109, no. 2(2015), p. 117-124. ISSN 0009-2770

The economics of solar power plant

[1]Vladimír GÁLL (2nd year)
Supervisor: [2]Alexander MÉSZÁROS

[1,2]Department of Electric Power Engineering, FEI TU of Košice, Slovak Republic

[1]vladimir.gall@student.tuke.sk, [2]alexander.meszaros@tuke.sk

***Abstract* — this work deals about solar cells as renewable source of energy. In other part this work is a sample calculation of economic efficiency of the most powerful photovoltaic power plant in the word.**

***Keywords* — economic efficiency, silicon cells, solar cells, solar power plant, Topaz solar farm,**

I. Introduction

In the past people started use fossil fuels like gas, crude oil and coal. They could be simple mined from surface and depth of planet Earth. In the present we combust the fossil fuels in big volumes. So there are many problems with mining, transport and finding new places with occurrence of fossil fuels. One approach is replacement of fossil fuels with renewable sources. As renewable sources of energy we can use solar energy, wind energy, hydro energy, biomass energy and geothermal energy. One option is using solar energy and build huge solar power plant for satisfy demand of electric energy. Cost efficiency and space efficiency for solar power plant are the determining factors for most investors [6] [6].

II. Monocrystalline Silicon Solar Cells

Solar cells made of monocrystalline silicon, also called single-crystalline silicon and are quite easily recognizable by an external coloring and uniform look. This indicates high clean silicon used in solar cell. Monocrystalline solar cells are made out of silicon ingots, which are cylindrical in shape. To optimize performance and lower costs of a single monocrystalline solar cell, four sides are cut out of the cylindrical ingots to make silicon wafers and that is something, what gives monocrystalline solar panels their characteristic look. The typical monocrystalline solar cell is a dark black color [5] [6].

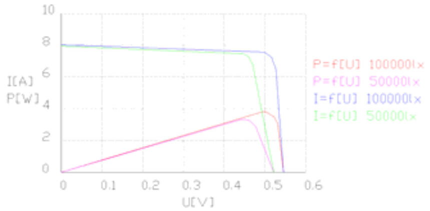


Fig. 1. Characteristics for monocrystalline cells

III. Polycrystalline Silicon Solar Cells

The first solar cells based on polycrystalline silicon, which also is known as polysilicon and multi-crystalline silicon. Raw silicon is melted and poured into a square mold, which is cooled. Then silicon is sliced into perfectly square wafers. Polycrystalline solar cells look perfectly rectangular without rounded edges. Polycrystalline solar cells are identifiable by its signature light or dark blue color [5] [6].

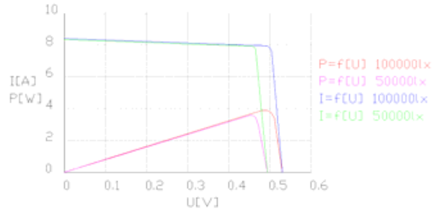


Fig. 2. Characteristics for polycrystalline cells

IV. String Ribbon Solar Cells

String Ribbon solar panels are also produced of polycrystalline silicon. String Ribbon is the name of a manufacturing technology that produces a form of polycrystalline silicon. Temperature resistant wires are pulled through molten silicon, which results in very thin silicon ribbons. Solar panels made with this technology looks similar to traditional polycrystalline solar panels [5] [6].

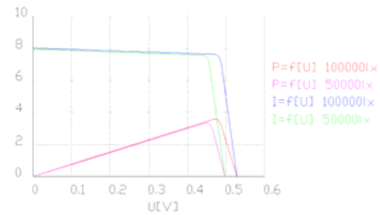


Fig. 3. Characteristics for string ribbon silicon solar cells

V. Amorphous Silicon Solar Cells

Amorphous silicon solar cells belong to the category of thin film solar cells, where one or several layers of photovoltaic material are deposited onto a substrate. The silicon material is not structured or crystalized on a molecular level, as many other types are of silicon based solar cells. This problem is partially solved by stacking several amorphous solar cells on top of each other, which increases their performance and makes them more space efficient [5] [6].

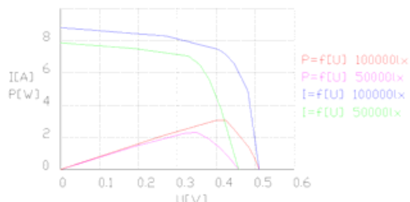


Fig. 4. Characteristics for amorphous silicon solar cells

VI. The most powerful photovoltaic power plant

The most powerful photovoltaic power plant in the world has power P_m is 550 MW. Time of using maximum τ is 1200 h/year, and annual production of electric energy A is 660 GWh. The amount of investment unit costs for plants using renewable energy sources have an influence mainly energy density in these resources, which results in high specific requirement of material and area. In this table are compared four types of solar panels for this power plant [6].

TABLE I

DIFFERENT TYPES OF SOLAR CELLS

Type of solar panels		Monocrystalline	Polycrystalline	Thin Film	String Ribbon
Initial costs per 1 kW	(€/kW)	2165,376	1369,413	1392,444	1700,064
Area per 1 kW	(m^2/kW)	7	8	15	10
Variable costs	(€/kWh)	0,008	0,008	0,008	0,008
Exteranlity costs	(€/kWh)	0,0004	0,000975	0,005	0,002125
Avarge effecinency	(-)	17	14	8	13
Iinitial costs for power plant	(€)	1190956800	753177150	765844200	935035200
Area for power plant	(km^2)	11,67	13,34	25	16,67
Cash flow per year	(€/year)	55916212	70858999,75	67759153	63734968

Annual production costs N_v for photovoltaic power plant can be calculated by next equation. These calculations are shown on the next figure [1] [4].

$$k_1 = k\ N_i \qquad k_2 = n_{po} + n_{ex} \qquad N_v = k_1 + k_2\ A \quad (1)$$

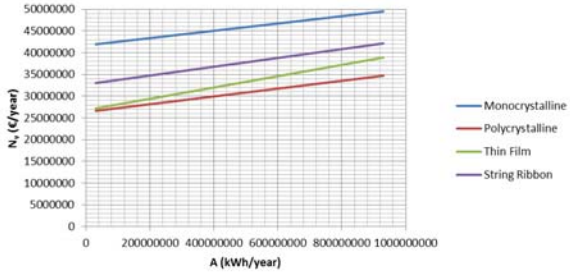


Fig. 5. Characteristics of costs per year

If annual production costs N_v are divided by annual production of energy A, we get equation for unit costs n_v. These calculations are shown on the next figure [1] [4].

$$n_v = \frac{k_1}{A} + k_2 \qquad (2)$$

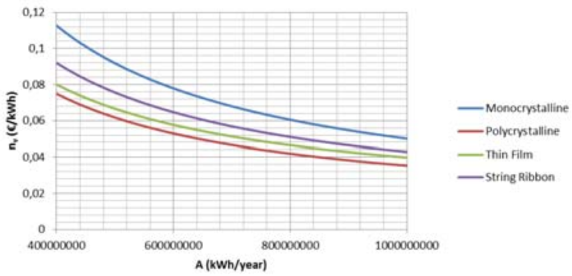


Fig. 6. Characteristics of unit costs

If annual production costs N_v are divided by power of plant P_m, we get equation for specific unit costs ${}_1N_v$. These calculations of equilibrium diagrams are shown on the next figure [1] [4].

$$k_3 = \frac{k_1}{P_m} = \frac{k\ N_i}{P_m} \qquad k_4 = k_2 = n_{po} + n_{ex} \qquad {}_1N_v = k_3 + k_4\ \tau \quad (3)$$

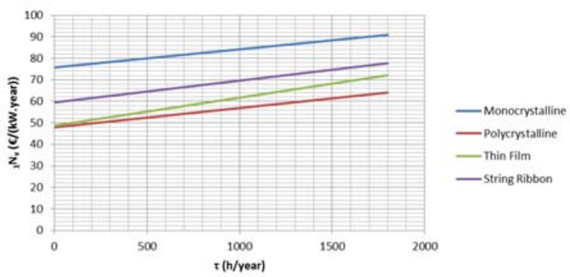


Fig. 7. Equilibrium diagram

Information about return cash flow is decided for building this power plant. The best results have solar panels made from polycrystalline cells. This is shown on the figures without and with time factor [1] [4].

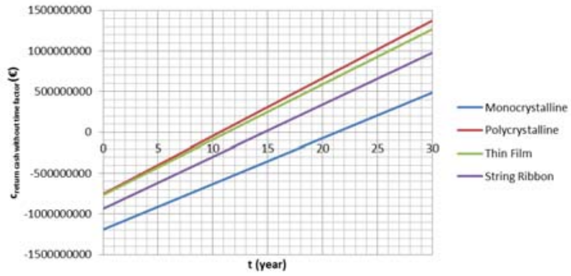


Fig. 8. Return cash flow without time factor

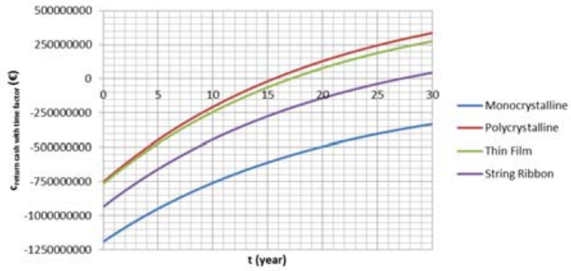


Fig. 9. Return cash flow with time factor

VII. Conclusion

Currently, these calculations simulate results only for solar power plant with photovoltaic panels. For the future calculations can be compared with other types of power plants. For the future problems with fossil fuels can be solved. This solution is economically effective but we need large area for this power plant.

Acknowledgment

This This work was supported by the Scientific Grant Agency of the Ministry of Education of Slovak Republic and the Slovak Academy of Sciences under the contract No. Vega 1/0388/13.

References

[1] Kolcun, M.; Beňa, Ľ.; Mészáros, A.; "Optimalizácia prevádzky elektrizačnej sústavy", Technical University of Košice, 2009, ISBN 978-80-553-0323-9

[2] Kolcun, M.; Chladný, V.; Mešter, M.; Cimbala, R.; Tkáč, J.; Hvizdoš, M.; Rusnák, J.; "Elektrárne", Technical University of Košice, 2006, ISBN 80-8073-704-5

[3] Dorf, C. R.; "Energy, resources and policy", Addinson Wesley Publishing Company Inc., 1978, ISBN 0-201-01673-7

[4] Kolcun, M.; Griger, V.; Beňa, Ľ.; Rusnák, J.; "Prevádzka elektrizačnej sústavy", Technical University of Košice, 2007, ISBN 978-80-8073-897-2

[5] Novák, P.; "Elektrické teplo", Technical University of Košice, 2010, ISBN 978-80-89284-48-1

[6] Pietriková, A.; Banský, J.; "Základy inžinierstva materiálov", Technical University of Košice, 2007, ISBN 80-8086-001-7

[7] Jaret, P.: "Electricity for Increasing energy efficiency", John Wiley & Sons, 1992

[8] Khatib, H.; "Economics of reliability in electrical power systems" Technicopy, 1978)

[9] Khatib, H.; and Munasinghe, M.; "Electricity, the environment and sustainable world development", 1992

[10] Mészáros, A.; Gáll, V.; "Global electrical power planning, investments and projects", 2014, Technical University of Košice, ISBN 978-80-553-1704-5

The end of a chapter in one man's life

[1]*Jozef LIPTÁK (4^{th} year),*
Supervisor: [2]*Ján ŠALIGA*

[1,2]Dept. of Electronics and Multimedia Communications, FEI TU of Košice, Slovak Republic

[1]jozef.liptak@tuke.sk, [2]jan.saliga@tuke.sk

***Abstract*—This paper is a summarizational article, that presents a summarization of my work in the past and last year of my post gradual study. I describe here assets of my work, and add some praise of me - based of my major publications. My main work in the last year was focused on study and proposal of new methods for decomposition of signal's into a set of exponentional functions and the behavioral studies of this decomposition. The other part of my work was focused on development of a library of functions for analog-to-digital converters testing, according to the existing standards and deploying this library on the web.**

***Keywords*—last year, Jozef Lipták, praise, summarization, work.**

I. Introduction

ADC's (analog digital converter) and DAC's (digital analog converter) are a major part of our technical World. Because ADC and DAC testing is a key operation that secures the accurate interpretation of measured data, developing new and faster testing methods is still a actual topic. Also new ADC architectures were presented - when assuming they typical behavior, development of optimal testing procedures is possible.

My other minor interests are focused on signal analysis. Many real-world signals are very complex, covered in high amounts of noise, or need to be decomposed in order to obtain some information about the source or to provide knowledge about degradation of the signal path through which is the signal propagating.

II. What was done before

We continued to cooperate with István Kollár from the Department of Measurement and Information Systems in Budapest. We made a exhaustive comparative study of ML (maximum likelihood) vs. LS (least squares) fitting for ADC testing. The resulting article was very successful and is indexed in Current Contents.

Work related to signal conditioning and analysis was done with cooperation with Slovak Academy of Sciences. This resulted to a software package and a measurement stand for automated respirometric measurements.

A newly introduced DAC measurement method was analyzed for its uncertainty. The results were published in the International Conference on Measurement.

We run a training stand for ADC study on our server. This was the result of a didactic oriented KEGA project. The web-page is accessible on the web and runs also some other electronic oriented training material and simulations of main electronic circuits (e.g. operational amplifier, stabilized power supply, etc.).

Another signal analysis work was done with cooperation with Domenico Luca Carnì. Experiences gathered within this cooperation resulted to a study that tried to classify capacitors on behalf of their discharging curve.

All of the above was presented in detail on the last year Scientific Conference of Young Researchers in my paper [5].

III. The actual work

A. Scientific and dissertation thesis related work

Exponential input stimulus signal is attractive for dynamic ADC tests because of the simplicity of the generating RC circuit. The estimation of the testing error requires measurement of the distortion of the stimulus signal. Despite the simplicity of the RC circuit there is a distortion in the exponential signal: distortion components are caused by dielectric imperfections of the capacitors in the discharging circuit. So we continued to analyze the properties of such exponential signals, their decomposition and the influence of noise and distortion to this decomposition.

This resulted to a Current Contents indexed paper [2] that presents a method for measurement of multiexponential signal components as an example of the more general task of signal decomposition where signal components are non-orthogonal. Such a complex signal is then represented as follows:

$$x_s(t) = A_1 e^{-B_1 t} + \sum_{i=2}^{L} A_i e^{-B_i t} + C + n(t), \qquad (1)$$

where $A_i \ll A_1, B_i \ll B_1, i = 2, 3, \ldots$.
While parameters A_1 and B_1 represent the basic exponential signal, the parameters A_i and B_i represent the distorting exponential components. Number L is the number of all exponential components assumed. The constant C in (1) describes the offset of the whole exponential signal. The distorted multiexponential signal is corrupted by additional (mainly thermal) noise of the analog components and by interferences from external sources ($n(t)$). The well known commonly used method for estimation of multiexponential signal is Prony's method. The paper presents new alternative estimation method based on maximum likelihood optimization. It was aimed on the comparison of both methods using simulation of taking into account a real reference waveform recorder and by measurement performed on the circuit with known distorting elements. The proposed method is also suitable to identify parasitic components representing dielectric absorption in the capacitors. This paper was a extended and exhausting study of a previous paper [4], in which was this concept firstly mentioned.

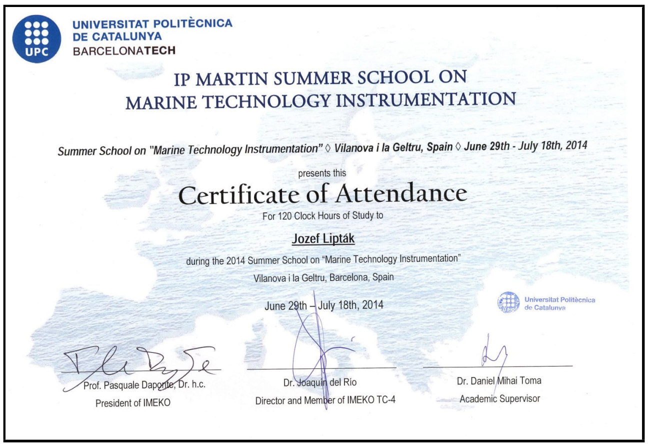


Fig. 1. Certificate of Attendance on IP MARTIN 2014

The second part of my work was to develop a library that allows very fast and effective development of software in LabVIEW® for ADC testing according to the standards. The library is accessible on http://meas-lab.fei.tuke.sk/ADC_test. The library consists of the following classes:

- **Static test**
- **Sinewave histogram processing**
  - Test conditions
  - Histogram test
- **Dynamic test**
  - Test in time domain
  - Test in spectral domain

As an example of application of the library, software demonstrating all common ADC test methods accessible across the Internet was developed. Here any user can perform simulated test, change ADC characteristics and test conditions, and can learn how these condition influence results of testing. The library was presented in [3].

B. Other (awards, certificates, etc.)

Right after the deadline for paper submision of the last SCYR conference I had the opportunity to participate on The Erasmus Intensive student meeting called Intercultural Knowledge Transfer in Engineering for a Sustainable Global ICT Community: SUSCOMTEC 2014 that took place at the Technical University of Sofia (Bulgaria). I was a member of the group that took the 1^{st} place for their project presentation.

I participated during the summer vacation also on a intensive program of MARTIN Summer School Marine Technology Instrumentation, organized by the SARTI centre of Technical university of Catalonia in Vilanova i la Geltrù (Spain). Lectures, exercises and projects of this summer school were mainly focused on the technology used in the Navy. All projects were presented in front of a committee and participation diplomas (see Fig. 1) were awarded. The projects were also published in the journal Instrumentation Viewpoint [1].

Last years PhD students publication scoring went very good for me. I ranked in the 8^{th} place at our faculty and in the 3^{rd} place at our department. My resulting score is in the Table I.

I prepared during the winter semester study material's for the Tempus Project "Technological Transfer Network".

My last but not least experiment was to write a paper using $\LaTeX\,2_\varepsilon$. You are reading it now. After this test I decided to write my dissertation thesis also in $\LaTeX\,2_\varepsilon$.

TABLE I
FINAL SCORE AFTER 3^{rd} YEAR OF STUDY

ADC in CC	AFC, AFD in SCOPUS and WoS	AFC, AFD non SCOPUS and WoS	**TOTAL**
8.0	1.2	2.0	**11.2**

According to https://hodnotenie.fei.tuke.sk and https://epc.lib.tuke.sk/

IV. Next steps

My main interest in the following days is given by my status – a student in his last year of doctoral study. I want to focus all my time to work on my dissertation thesis. All of the investigation and research work is done. No new areas will be examined. I need to finish just few experiments for data collection purposes. They are related to a ongoing bachelor thesis, so they will be done right in time to get the deadlines.

After the submission of mi dissertation thesis I will center on the preparation of the defense of this thesis and continue to work on ongoing projects at our department.

Acknowledgment

First of all, I would like to express my gratitude to my dissertation thesis supervisor, prof. Ing. Ján Šaliga, PhD.. He has been a constant source of encouragement and insight during my research and helped me with numerous problems and professional advancements. Also many thanks to prof. Ing. Linus Michaeli, DrSc., head member of our research group for his ideas and helping hand with all electronic related problems that anyone can imagine.

Special thanks go to the staff of the Department of Electronics and Multimedia Communications, who maintained a pleasant and flexible environment for my research. I would like to express special thanks to the department management for providing most of the funding for my research.

Finally, my greatest thanks go to my family members, for their infinite patience and care...

References

[1] V. Bouliou-Hello, J. Lipták, and M. O. Pascual, "Automatic Morse Transceiver," *Instrumentation Viewpoint*, no. 16, pp. 36–37, 2014, ISSN 1886-4864. [Online]. Available: http://upcommons.upc.edu/revistes/bitstream/2099/15158/1/Bouliou-Liptak-Oriol_36-37.pdf

[2] L. Michaeli, J. Šaliga, J. Lipták, M. Godla, and I. Kollár, "Measurement of distorted exponential signal components using maximum likelihood estimation," *Measurement: Journal of the International Measurement Confederation*, vol. 58, pp. 503–510, 2014, ISSN 0263-2241. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0263224114004047

[3] J. Šaliga, J. Lipták, V. Vitkovič, and L. Michaeli, "Adc test library in labview," in *20th IMEKO TC4 Symposium on Measurements of Electrical Quantities: Research on Electrical and Electronic Measurement for the Economic Upturn, Together with 18th TC4 International Workshop on ADC and DCA Modeling and Testing, IWADC 2014*, 2014, pp. 667–668, ISBN-14: 978-92-990073-2-7. [Online]. Available: http://www.imeko.org/publications/tc4-2014/IMEKO-TC4-2014-851.pdf

[4] L. Michaeli, J. Šaliga, M. Godla, and J. Lipták, "Measurement of dielectric absorption of capacitors by signal decomposition," in *20th IMEKO TC4 Symposium on Measurements of Electrical Quantities: Research on Electrical and Electronic Measurement for the Economic Upturn, Together with 18th TC4 International Workshop on ADC and DCA Modeling and Testing, IWADC 2014*, 2014, pp. 106–110, ISBN-14: 978-92-990073-2-7. [Online]. Available: http://www.imeko.org/publications/tc4-2014/IMEKO-TC4-2014-262.pdf

[5] J. Lipták, "Ubi es et quo vadis Jozef Lipták," in *SCYR 2014 : 14th Scientific Conference of Young Researchers : proceedings from conference*. FEI TUKE, May 2014, pp. 116–117, ISBN 978-80-553-1714-4. [Online]. Available: http://scyr.fei.tuke.sk/scyr-files/history/SCYR_2014_proceedings.pdf

The influence of accelerated ageing on the oil impregnated paper

[1]*Lukáš Lisoň (3rd year)*
Supervisor: [2]Iraida Kolcunová

[1,2]Dept. of Electrical power engineering, FEI TU of Košice, Slovak Republic

[1]lukas.lison@tuke.sk, [2]iraida.kolcunova@tuke.sk

***Abstract*— This paper offers information about research of oil-paper insulation system. Insulation system is one of the most important part of electrical devices. It is very important to understand electrical and dielectric properties of insulation system. The insulation system of many type of electrical equipment is mainly composed by the insulating oil and paper. Oil impregnated insulation paper is used in transformers for many years. Conventional mineral oil is some possibility of environmental pollution and fire with explosion. But natural ester insulating oils are non-toxic, more biodegradable and less flammable than a mineral oil. Therefore conventional mineral oil is being replaced with natural ester insulating oils. The next aim of this paper is research of influence of accelerated ageing on the oil impregnated paper.**

***Keywords*— mineral oil, natural ester, thermal stress, ageing.**

I. Introduction

In a liquid-filled transformer, the insulating liquid plays an important function by providing both the electrical insulation (in combination with a solid such as cellulose) and the means of transferring the thermal losses to the cooling system. The insulating liquid can also provide important and easily obtainable information for use in diagnosing the health of a transformer [1]. For the reliable operation of high voltage power transformers, it is essential that the cellulose insulation structures used in their construction are completely oil impregnated. The oil impregnation procedure is important to ensure that no cavities are left inside the cellulose insulation and thereby dangerous partial discharges are avoided [2]. For more than one hundred years, the majority of liquid-immersed transformers have been filled with mineral oil. The significant use of this petroleum-based product has been justified until now by its wide availability, its good properties, its good combination with cellulose and its low cost. However, with environmental issues now becoming extremely important, the use of a product with a high fire point temperature and high biodegradability is becoming extremely attractive [3]. Natural esters are produced from vegetable oils, which are manufactured from plant crops. They offer the advantage of a high fire-point as well as good biodegradability, but all types of natural esters suffer from not being as oxidation stable as other types of insulating liquids. Although natural ester fluids can be produced from a wide variety of crop oils, natural esters for electrical applications are most commonly produced from soya, rapeseed and sunflower oil. This is due to factors such as availability, cost and performance characteristics [3]. Therefore, my research is focused on electrical properties of natural esters and the option to use of these liquid dielectrics in electrical equipment.

II. Recent Research

Research of electrical and dielectric properties of oil paper insulation is the aim of many scientific collective. Properties such as breakdown voltage, dissipation factor, relative permittivity belong between the basic of diagnostic variables.

Current research of this area compares the electric and dielectric properties of mineral oil paper insulation and rapeseed oil paper insulation. The results [4] shown that rapeseed oil paper insulation has a higher permittivity as a mineral oil paper insulation. Dissipation factor has a similar behavior for a both insulation systems. Temperature is one of the factors which has strong influence on condition of oil paper insulation.

III. My Research

Because of insulation system of oil- paper is composite from liquid part and solid part, my current research is focused on these parts together.

Since last SCYR conference my research was focused only on oil-paper insulation system. Properties as relative permittivity, dissipation factor, and electric breakdown voltage were measured. The results [5] shown that, both insulation systems have similar dependence of dissipation factor on frequency. Dissipation factor exponentially decreasing as the frequency is increasing. Rapeseed oil paper insulation has a higher value of dissipation factor than mineral oil paper insulation. The breakdown voltage of natural esters is higher than breakdown voltage of mineral oil. The average value of breakdown voltage for rapeseed oil paper insulation is 19.3 kV and for mineral oil paper insulation is this value 18.3 kV. As well as the breakdown voltage, the electrical breakdown strength of rapeseed oil paper insulation has higher value for E_p than mineral oil paper insulation. The average value of E_p for rapeseed oil paper insulation is 53.61kV/mm and for mineral oil paper insulation is 50.84kV/mm. These results were published in Proceeding of the Faculty of Electrical

Engineering and Informatics of the Technical University of Košice, Electrical Engineering and Informatics V.

The aim of my next work was analyze of properties such as relative permittivity and dissipation factor of oil paper insulation depending on frequency and voltage. In this experiment were used two types of mineral oil and two types of natural esters with combination of paper. From the results [6] we can make following conclusions. Insulation paper impregnated by rapeseed oil has the highest value of relative permittivity in throughout frequency range. Insulation paper impregnated by sunflower oil and insulation paper impregnated by mineral oil Nynas-Lyra X have very similar waveforms of relative permittivity in frequency range from 10 Hz to 2 MHz. Waveforms of dissipation factor decreases exponentially for all specimens of oil paper insulation. The highest value of dissipation factor was measured at insulation paper impregnated by rapeseed oil. Significant difference of dissipation factor is among specimens in the frequency range from 10 Hz to 0.1 MHz. This difference is not such significant at higher frequencies, and at higher frequencies the dissipation factor starts increasing. Frequency has strong influence on changes of relative permittivity, whilst influence of applied voltage has not such significant. The same conclusions apply for dissipation factor. Voltage dependence of relative permittivity is constant or the changes are not such significant as at frequency. These results were presented on International conference of Current Problems of Maintenance of Electrical Equipment and Management, 10-12 September in High Tatras.

Temperature has great influence on condition of insulation system and it causes accelerated ageing of this system. Electrophysical properties as relative permittivity, AC breakdown voltage and polarization index of insulation oil paper before and after thermal ageing were measured. From the results we can make following conclusions. AC breakdown voltage of rapeseed oil paper is higher than AC breakdown voltage of mineral oil paper in each case. The value of relative permittivity is decreasing after each thermal stresses. Decreasing of relative permittivity is caused by evaporation of moisture from oil paper. Relative permittivity of mineral oil paper is increasing if voltage is increased in each of cases. The resistance of mineral oil paper insulation is higher than resistance of rapeseed oil paper insulation. The Results from this experiment will be published in Acta Electrotechnica et Informatica in current issue.

The last experiment describes the similar electrophysical properties as above, but there were used mineral oil, synthetic oil and natural ester as a one part of oil paper insulation. The temperature of accelerated ageing test was 90°C and there was chosen two different intervals of ageing. The first measurement was realized with new samples, the second one of measurement was realized after 500 hours ageing and the third one of measurement was realized after 750 hours ageing. Properties such as relative permittivity, dissipation factor and the breakdown voltage were described and analyzed. Conclusions from this experiment will be presented in 16th International Scientific Conference Electric Power Engineering (EPE) 20- 22 may 2015 in Kouty nad Desnou.

IV. FUTURE RESEARCH

The last experiment confirmed that the 750 hours of accelerated aging has not degradation effect on oil paper insulation. Degradation of insulation system causes that the properties as relative permittivity, dissipation factor and AC breakdown voltage are worse. This fact has not demonstrated in this experiment. The next research will be focused on investigating of these properties with longer time of accelerating ageing like 750 hours. The second direction of my future research will be devoted to mixture of mineral oils with natural esters which can improve anti-ageing properties of insulation system.

CONCLUSION

This article presented some electrical and dielectric properties of liquid insulators and combination insulating oil-paper. Results shown that natural esters can be used as substitute of traditional mineral oils obtained from the petroleum. Therefore, the research of natural ester oil paper insulation is very important.

ACKNOWLEDGMENT

"We support research activities in Slovakia/This project is being co-financed by the European Union".Paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182,supported by the Research & Development Operational Programme funded by the ERDF."

REFERENCES

[1] A. Dravin, Ch. Errier, P. Folliot.:"The use of natural ester fluids in transformers". Mat Post 07: Published online: www.matpost.org/.../MATPOST07_0036_paper.pdf.

[2] J. Dai, D.Z. Wang. : A Comparison of the Impregnation of Cellulose Insulation by Ester and Mineral oil. IEEE Transactions on Dielectrics and Electrical Insulation Vol.15, No. 2; April 2008.ISBN 1070-9878/08

[3] M. Rycroft, "Vegetable oil as insulating fluid for transformers," Transmission and Distribution. Published Online April 2014 ,pp. 37-40. http://www.ee.co.za/?s=Vegetable+oil+as+insulating+fluid+for+transformers.

[4] J. Li, S. Gryybowskí, Y. Sun and X. Chen, "Dielectric properties of rapeseed oil paper insulation." Annual Report Conference on Electrical Insulation and Dielectric Phenomena.2007.

[5] L. Lisoň, M. Kmec, I. Kolcunová : "Meranie elektrofyzikálnych vlastností izolácie olej-papier" In: Electrical Engineering and Informatics 5 : Proceedings of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice. - Košice: TU, 2014 S. 825-829. - ISBN 978-80-553-1704-5

[6] L. Lisoň, M. Kmec, I. Kolcunová, Z. Čonka : "Relative permitivitty and dissipation factor of oil paper insulation" In: Current Problems of Maintenance of Electrical Equipment and Management. - Košice : TU, 2014 S. 197-203. - ISBN 978-80-553-1818-9.

UWB radar measurement challenges

[1]*Daniel NOVÁK (2nd year)*
Supervisor: [2]Dušan KOCUR

[1,2]Dept. of Electronics and Multimedia Communications, FEI TU of Košice, Slovak Republic

[1]daniel.novak@tuke.sk, [2]dusan.kocur@tuke.sk

***Abstract*—Number of measurements realized by our department with UWB radars is increasing every academic year. During these measurements we are aware of multiple problems devoted to the measurement with a UWB radars. At this moment there is no simple solution to solve such problems. However, taking into account and avoiding of such situations is very effective in all aspects of measurements. In this paper, an overview of present known problems is given.**

***Keywords*—UWB radars, measurement, undesired effects, problems.**

INTRODUCTION

Ultra WideBand (UWB) sensors are lately very popular devices in several fields of applications. Such sensors can be used for human motion monitoring purpose e.g. human walking [1], human breathing [2] or human heartbeat detection [3]. In recent years, there are many interesting applications of UWB sensors arising such as building mapping using Synthetic Aperture Radar (SAR) [4], impedance spectroscopy for material characteristics investigation or radar configuration as Ground Penetrating Radar (GPR) for underground landmine detection [5].

UWB sensor is commonly known as UWB radar device. It consists of minimum one transmitting and one receiving antenna. Such radar emits one type of UWB signal. This signal is defined as an electromagnetic wave with certain spectral characteristics. In case of UWB signal, its value of absolute bandwidth B has to be at least 0,5 GHz. Type of emission signal can be e.g. set of impulses, frequency modulated signal, noise or pseudonoise respectively. At our department, 3 pseudonoise radars emitting M-sequence and one impulse radar are available.

During various measurements using these UWB radars we are experiencing several problems which need to be taken into account. In this paper, a quick overview of mentioned problems is given.

This paper is organized as follows. First section is discussing about multiple problems that can be experienced during measurements with UWB radars. Also, origin and utilization of some problems is given. In the second section, author's other academic activities for the last year are presented.

I. SIGNAL ACQUISITION PROBLEMS

Let's focus on the detection of human targets using UWB radar. Successive detection and localization of a human target strongly depends on the environment or surroundings. We suppose that a movement of the target is placed within *complex environment* [6]. This environment contains a lot of solid obstacles made of miscellaneous materials. Propagation of the electromagnetic wave through such surroundings causes this wave to scatter and/or diffract. Some part of energy is absorbed by material.

For successful detection of the target it is required to receive the signal transmitted from the radar and scattered from the target at a sufficient power level. Appearance of the target in complex environment causes lowering of the power level of the received signal. This situation might in some cases make detection almost impossible. In the next parts, some of the undesired factors related to such situation are mentioned. Many of them are commonly experienced during conventional measurement with UWB radar system.

A. Internal factors of the radar system

Some of these phenomenons can be solved by proper choice of the radar and antenna equipment or by setting the suitable parameters of a radar system respectively.

Antennas of the radar system – gain, directivity and radiation pattern have great influence on the amount of a signal that can be acquired from the environment. Also lower height and small mutual distance of the antennas are negative factors that have, especially for the targets situated at a greater distance from the antenna system, negative influence for detection and localization of such targets. In cases where possible is recommended place the antenna system to the sufficient height (more than 150 cm, less than 200 cm).

Radar sensor parameters – range, resolution, emitted power, measurement speed or frequency band of the radar signal. Especially for the through-the-wall measurements it is important to provide a signal operating in the baseband. Lower frequencies of the emitted signal are more suitable to penetrate through solid nonmetallic obstacle.

Speed and length of measurement – proper speed of measurement (data acquisition) can ensure detectability of moving target. Length of measurement is important especially when vital signs of the monitored target are detected.

B. External factors

Many of these issues cannot be eliminated easily. However, some of them can be suppressed by proper placing of the radar system. This category contains mainly properties of the monitored environment.

Character of the surroundings – number and position of walls, windows, bigger pieces of furniture and strong static reflectors. These were defined in [6] as objects that provide much better reflection of electromagnetic wave compared to human target. Partial solution of strong static reflectors removal consists in using of two-staged detector. This method named advanced background subtraction was presented in [7].
Character of the environment – moisture and dust level.
Obstacle parameters – electrical parameters such as permittivity, permeability, conductivity and geometrical parameters as corners, ceilings (horizontal walls) alcoves etc. Special parameter is homogeneity, which can be defined as difference of obstacle (wall) thickness, especially when this obstacle is a front wall (in front of antennas). In this case, a method called *wall effect compensation* can be used with advantage [8].
Current target position – when a target is situated close to the strong static reflectors or behind the metallic obstacle, localization errors may occur.
Number and mutual position of multiple targets – in case of multiple target appearance within the monitored area, a special situation may occur. Target situated closer to the antenna system may block the path of radiated electromagnetic wave. Solution for this situation is use of the method called *weak signal enhancement* [9].
Type of target's motion – generally, motion of the target that can be marked as "obvious" (walking, moving of limbs etc.) is easier to process than periodic motion caused by vital signs. Therefore, better conditions such as higher signal-to-noise ratio or no other additional motions are required for successive detection.
Interference and jamming of the radar signal – UWB signal is partially resistant to common interference with TV, radio or mobile communications signals. However, interference at the lower frequencies signals with concentrated power may effectively overlay UWB radar signal at the radar input. Example of this case can be local radio frequencies of security and emergency services.

Not considering effects and phenomenons mentioned above may lead to errors and defects in final estimation of target's position or track.

II. Other Author's Academic Activities

For purposes of measurement simplification, a modular construction RexRoth was obtained into our laboratory (Fig. 1). This construction can be used as an antenna stand and also as a radar equipment trolley.

Our laboratory also took part in action called "Researchers night 2014". It was an opportunity to present our equipment and work to a large audience.

During summer vacations me and my colleagues were participating at a summer school called "MARTIN Summer School Marine Technology Instrumentation" organized by SATRI Centre, Technical University of Catalonia, Barcelona, field office Villanova i la Geltrú. This school provided three weeks of intense courses focused of technology used in the Navy. Our two professors, prof. Michaeli and prof. Šaliga also participated to this school as a lecturers.

Output from this summer school is a journal Instrumentation Viewpoint with a submission from each group [10].

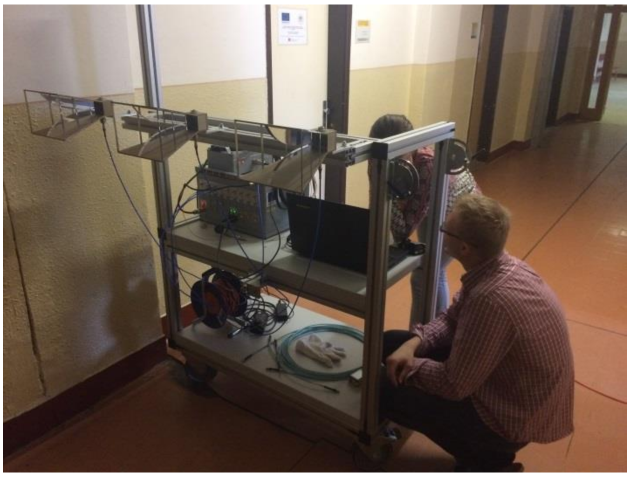


Fig. 1: Construction RexRoth in use

III. Conclusion

Summarizing of problems appearing during measurement is important by means of understanding and future work. Avoiding of just some undesired effects can clearly improve target detection probability and localization accuracy. For problems that cannot be avoided, existing methods should be improved or new methods could be implemented.

Acknowledgment

This work was supported by the Scientific Grant Agency of the Ministry of Education, science, research and sport of the Slovak Republic under the contract No.1/0563/13 and by the Slovak Research and Development Agency under the contract No. APVV-0404-12.

References

[1] Kazimir, P. et al. "A simple approach to through wall localization of persons moving in 3-dimensional space." *European Radar Conference (EuRAD), 2014 11th*. IEEE, 2014.
[2] Sachs, Jürgen, et al. "Remote vital sign detection for rescue, security, and medical care by ultra-wideband pseudo-noise radar." *Ad Hoc Networks* 13 (2014): 42-53.
[3] Jian, Qiuchi, et al. "Detection of breathing and heartbeat by using a simple UWB radar system." *Antennas and Propagation (EuCAP), 2014 8th European Conference on*. IEEE, 2014.
[4] de Wit, J. J. M., et al. "Multifunction radar concept for through-wall surveillance." *Radar Conference (RADAR), 2013 IEEE*. IEEE, 2013.
[5] Daniels, David J. *Ground penetrating radar*. John Wiley & Sons, Inc., 2005.
[6] Kocur, D., Novak, D. and Rovnakova, J. "Moving person tracking by UWB radar system in complex environment." *Intelligent Signal Processing (WISP), 2013 IEEE 8th International Symposium on*. IEEE, 2013.
[7] Novak, D., Zetik, R. and Kocur, D. "Defective localization of target in UWB radar applications." *Radioelektronika (RADIOELEKTRONIKA), 2014 24th International Conference*. IEEE, 2014.
[8] Rovnakova, J. and Kocur, D. "Compensation of wall effect for through wall tracking of moving targets." *Radioengineering* 18.2 (2009): 189.
[9] Rovnakova, J., and Kocur, D. "Weak signal enhancement in radar signal processing." *Radioelektronika (RADIOELEKTRONIKA), 2010 20th International Conference*. IEEE, 2010.
[10] Novak, D., Fura, V., Paillot, A. "Acoustic Channel Frequency Response" in *Instrumentation Viewpoint*. No. 16 (2014), p. 28-30. - ISSN 1886-4864, 2014

Vehicle power management

Milan BIROŠ [1] (1st year)
Supervisor: [2] František ĎUROVSKÝ

[1,2] Dept. of Electrical Engineering and Mechatronics, FEI TU of Košice, Slovak Republic

[1] milan.biros@tuke.sk, [2] frantisek.durovsky@tuke.sk

***Abstract*—The paper presents short introduction of power management in vehicles. It contains description of powertrain configurations in currently produced vehicles, their advantages and disadvantages and also power management approach to optimizing fuel economy and emissions reduction. Last section of this paper describes problems of currently used control strategies.**

***Keywords*—hybrid electric vehicles, power management.**

I. Introduction

In the present, almost all propulsion power in conventional vehicles is provided by spark-ignited or compression-ignited internal combustion engine. As it is well known, as the energy source for mentioned engines serves gasoline or diesel; both refined from fossil oil. In the 2008, the world oil reserves were 1,342 trillion barrels [1] and the daily consumption was about 85 million barrels [2]. Around 60% of the total oil consumption goes to transportation. The United States Energy Information Administration predicted that the world daily oil consumption would increase to 98,3 million barrels in 2015 and 118 million barrels in 2030 [3]. Burning of fossil fuels also contributes to the increase of carbon dioxide (CO2) in the atmosphere which conduces to greenhouse effect [4].

Those are the main reasons why European parliament, as an international regulator, forces vehicle producers to improve the propulsion systems efficiency and to lower emissions, produced by a passenger car, to 95 g CO2/km [5] in 2020.

Accordingly, new propulsion system configurations, such as hybrid vehicles (HEV), plug-in hybrid vehicles (PHEV) and pure electric vehicles (EV) are under development. To decrease amount of wasted energy, power management is used.

The main aim of power management is to control combustion engine and electric motor power demand, manage battery charging and discharging and energy regeneration.

This paper is dedicated to power management in vehicles. The contributions is organized as follows, section two describes the main items which have to be considered in power management controller design. The power management in conventional vehicles is described in section three and in hybrid electric vehicles in section four. In last section reader may find problems of currently used optimization methods in power management.

II. Vehicle Power Management

Power management in vehicles is very important for various objectives, such as better fuel economy, reduction of pollutant emissions, prolonging lifetime of power sources (e.g. the battery or fuel cell) and enhancing vehicle drivability and reliability. To accomplish these objectives, it is important to develop proper power management strategies with respect to:

- *Optimal engine operating point* – controlling engine operation at area with highest fuel economy and minimum emissions.
- *Optimal engine operating line* – controlling on the line constituted by optimal operating points related to different power demands.
- *Optimal engine operation region* – operating region on the torque-speed plane where the fuel economy of the engine is high.
- *Minimum engine dynamics* – regulating at proper speed avoiding fast fluctuations
- *Minimum engine speed* – turning off the engine below threshold level where engine economy is low
- *Minimum engine turn-on time* – avoiding of frequent engine turning on and off
- *Proper battery capacity* – keeping battery state of charge (SOC) at proper level for supplying power to acceleration and absorbing regenerative breaking power.
- *Safety battery voltage* – avoiding of over-voltage and under-voltage of the batteries.

III. Power Management in Conventional Vehicles

The power management in conventional vehicles may be divided into two areas:

- Powertrain torque-speed characteristic mapping
- Electric power management

Conventional internal combustion engine (ICE) is capable to deliver torque only at bounded speed range (e.g. from 800 to 7500RPM) and there is a narrow range which provides better fuel economy. Therefore, dual clutch transmission and automatic transmission were developed. Their task is to match the engine outputs (speed and torque) with vehicle demand while optimizing engine efficiency. However, both transmissions have finite amount of gear ratios what is very limiting for ICE operation optimization. Continuous variable transmission overcomes this problem by providing infinite, continuously changeable gear ratios. Such a transmission enables engine to run in optimum speed-torque area in whole speed range.

On the other hand, more and more mechanical systems of the vehicle are replaced by electric control (so-called drive-by-wire systems), for example electric steering, electromechanical breaking, active suspension, etc.

All these devices, among amount of other onboard electronic, such as infotainment systems, heated seats, heated windows, etc., have to be supplied with electric power provided by the battery and alternator. The alternator is coupled with the engine through the pulley, so loading of the alternator means also the loading of the engine. As in typical passenger car, electric power demand may be more than 3-5 kW, which is no longer negligible.

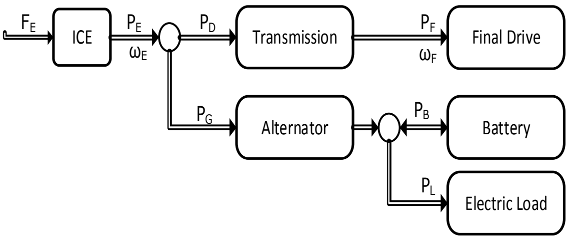


Fig. 1 Power flow in conventional vehicle

The power flow in the conventional vehicle is shown in Fig. 1. The fuel rate F_E is nonlinear function of engine speed ω_E and engine power P_E, i.e., $F_E=F(P_E,\omega_E)$. The whole engine power is split between the drivetrain and the alternator:

$$P_E = P_D + P_G$$

where: P_E is the engine power, P_D is the power of drivetrain and P_G is the power of the alternator.

The alternator converts mechanical energy P_G to electrical power where one portion, denoted as P_B, charges the battery and other portion, denoted as P_L, supplies electric loads. In some cases, the power to electric loads is provided by the battery, thus alternator load is reduced. Hence, F_E may be modeled as a function of P_D, P_B, P_L and ω_E.

P_D and P_L depend on drivers' demand, therefore, in the most cases; it is not possible to control them. ω_E needs to match vehicle speed by controlling the transmission gear ratios. Also magnitude and direction of P_B may be controlled depending on the state of charge (SOC) of the battery. Since only P_B and ω_E can be controlled, it is possible to describe fuel rate F_E as the nonlinear function of P_B and ω_E,

$$F_E = f(\mathrm{P}_B, \omega_E)$$

It follows that fuel economy may be improved by power management [10].

IV. Power Management of Hybrid Vehicles

Hybrid vehicles combine ICE with electric motor. There are several different drivetrain configurations of hybrid vehicles. The main configurations are

- Series hybrid
- Parallel hybrid
- Series – parallel hybrid
- Complex hybrid

Control strategies in HEV are more complicated than in conventional vehicles due to more freedom of power splitting between various power sources. Thanks to this, it is possible to maintain better fuel efficiency results as with conventional vehicles. Power management controller in HEV is designed to meet driver's power demand while optimizing fuel economy, reducing pollutant emission and maintain state of charge (SOC) of the battery [9].

In series hybrid, ICE is not directly connected to wheels so ICE serves only to drive generator to recharge traction battery (as shown on Fig 2.A). The main goal of power management in this case is to maintain SOC of the battery and control ICE to operate in optimal torque-speed area. Wheels are driven only by electric motor, which is also able to regenerate energy during braking. The main disadvantage of this configuration is that it requires more propulsion components (ICE, generator, motor). Motor has to be designed for the maximum power required by vehicle. For long distance trip, all three propulsion components (ICE, motor and generator) has to be rated at maximum power.

In parallel hybrid, an electric motor and ICE are connected to the transmission through a mechanical coupling device. In this case, vehicle may be propelled by electric motor, by ICE or by both. Mechanical coupling between motor and ICE can be configured to share common transmission, or use separate transmissions or even separate axles. Vehicle is primarily propelled by ICE and electric motor is in use when burst of power is demanded or during regenerative breaking as generator. That means that if more power is needed, control algorithm has to decide whether demand more power from ICE and let it operate in less efficiency torque-speed area or increase power using electric motor. On the other hand, when the ICE operates at given speed and above optimal speed-torque range it is possible to use electric motor to charge the battery depending on SOC. The main advantage of parallel hybrid against series is that parallel hybrid propulsion system needs only two propulsion components (ICE and electric motor). Also for short-distance trips ICE and electric motor may be rated at reduced power levels. For long-distance trips the engine may be rated at maximum power, while motor/generator may still be rated to half of maximum power.

Considering advantages of both, parallel and series hybrid configurations, manufacturers have developed series-parallel hybrid vehicles. In this configuration, ICE is mechanically coupled with electric motor like in parallel hybrid but is also connected with generator through selection device as in series hybrid. Typical representative of series-parallel hybrid is Toyota Prius where a small series element is placed in addition to primarily parallel HEV to charge the battery during prolonged waiting periods such as traffic lights or in traffic jam [6].

These HEV can operate using electric motor alone or with assistance of the ICE. As mode selection device, sets of clutches may be used to select which shaft is connected to ICE, whether the final drive shaft or the electric generator shaft. Another choice is to use power split device such as planetary gear train, which can split power of the ICE to final shaft and the electric generator. Control unit is in the charge of driving mode selection. This configuration allows power management control unit to maintain ICE in optimal operation area with high fuel efficiency.

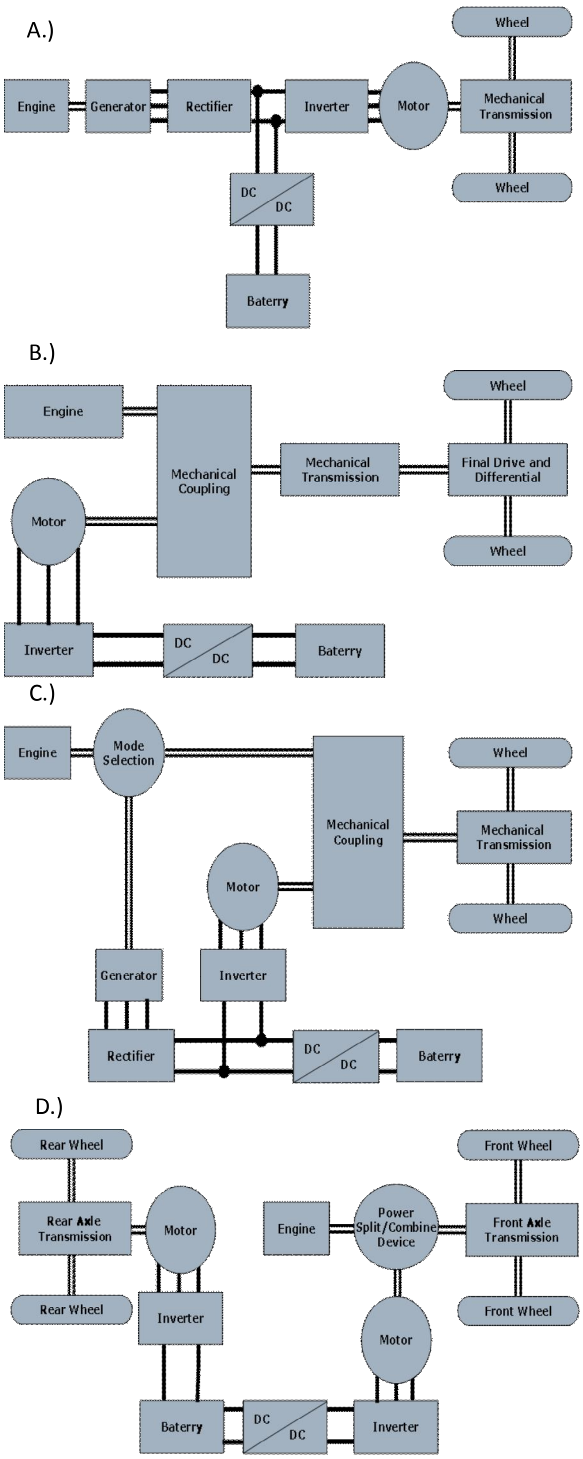


Fig. 2 Configurations of HEV (A. Series hybrid, B. Parallel Hybrid, C. Series-Parallel hybrid, D. Complex Hybrid)

There is also more configuration of HEV which cannot be classified in previous categories. In generally, they may be designated as complex hybrids. For example, in Fig. 2, there is dual-axle four-wheel drive vehicle in which rear axle is driven by electric motor while front axle is propelled by combination of ICE and electric motor. Complex hybrid is very similar to series-parallel. The main difference is that in the series-parallel hybrid, the generator allows only unidirectional power flow whereas the complex hybrid motor connected to power split/combine device allows bidirectional power flow. This bidirectional power flow allows operating in the three propulsion operating modes that are impossible in series-parallel hybrid. In complex hybrid rear- and front-axle are separately driven by electric motor or by hybrid powertrain. When low power is demanded, battery supplies front electric motor to drive front axle. In normal operation mode, ICE power is split to propel front axle and drive electric motor (as generator) to charge the battery. If the vehicle is running at heavy load, for example during acceleration, ICE with cooperation of front electric motor propel front axle and meanwhile rear electric motor drive rear axle. Both motors become generators during regenerative breaking [4, 9, 11].

Configuration of complex hybrid may differ in amount of electric motor or which axle is driven by ICE. For example, Toyota Highlander uses three electric motors and ICE to maximize fuel economy and improve drivability. Two electric motors are used to separately drive rear wheels what allows to use so-called electronic differential and one is used to propel front axle in cooperation with ICE.

V. Problems in Present Vehicle Power Management

So far, optimal control strategies, especially dynamic programming (DP) and artificial intelligent (i.e., fuzzy logic and neural networks) supervisory systems, have been successfully applied to vehicle power management.

In optimal control scheme for vehicles, optimization laws are based on cost functions of the vehicular systems [7]. The disadvantage is that optimal control is sensitive to variations of parameters and to measurement noises. It means that even small measurement inaccuracy may cause stability problem. For the optimization process, all the static and dynamic behaviors of the vehicle have to be taken into account. Sometimes, some assumptions may be used to simplify calculations due to complexity of given problem. But optimal solutions are obtained only under these assumptions so it may give worse results in real-world conditions. The discrete time event strategy is used in the optimal controller since it is simple and more robust. Depending on certain decision rules, system behaviors move from one state to another. Because discrete time event strategy uses binary (on/off) mode, only partial optimal solution may be obtained. The resolution of rules determines system performance.

The most popular artificial intelligent system approach, used in vehicle power management, is the fuzzy logic. Fuzzy logic control with a nonlinear structure, suitable for power split problem, gives more flexibility to optimization and brings more robustness than other methods [8]. On the other hand, the defuzzyfication process, if implemented, consumes a lot of computation time and memory. In addition, optimization of the defuzzyfication process may cause problems in fuzzy control.

Artificial neural network models are used for prediction of vehicles behavior, which covers fuel consumption and emission. In order to use neural networks in real-world conditions, a large diversity of training sets is needed. Implementing an effective artificial neural network requires also large storage and computation resources.

The wavelet technology can identify high- frequency transients from power demand of the driveline. With this technology, it is possible to control power split. Doing so it is possible to improve fuel efficiency and prolong life time of power sources as well. Its main disadvantage is the need of supercapacitor, which absorbs high power density during transient processes [12].

VI. Conclusion

The paper presented a survey of power management strategies which are used in the present vehicles. The common problem of all mentioned strategies is their implementation in real-world and real-time conditions. Although all of them are able to achieve good results in fuel efficiency and emissions producing during predefined driving cycle, it is difficult to achieve same results in real traffic, even with cooperation with some prediction methods to predict future road conditions. In the most cases, small prediction horizon and measurement inaccuracies decrease power management performance. To sum up, future vehicle power management requires new advanced control strategies, optimization strategies and better prediction algorithms to create more efficient real-time vehicular system.

Acknowledgment

This paper was developed with support of Research and development operating program for the project "University Science Park Technicom for innovative applications with support of knowledge technologies", code ITMS: 26220220182, cofinanced from European funds.

References

[1] Energy Information Administration, *Worldwide look at reserves and production*. (2008) Oil and Gas J 106:23-24

[2] Roberts Paul, *World oil*, http://ngm.nationalgeographic.com/2008/06/world-oil/roberts-text

[3] Energy Information Administration, *World oil consumption by region, reference case*, 1990-2030, www.eia.doe.gov/oiaf/ieo/pdf/ieoreftab_4.pdf

[4] Zhang Xi, Chris Mi, *Vehicle Power Management*, New York: Springer-Verlag, 2011

[5] REGULATION (EU) No 333/2014 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 11 March 2014, http://eur-*lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2014.103.01.0015.01.ENG*

[6] Knebusch M, Toyota Prius, 2004,Mach Des 76:111

[7] *Sciarretta A, Back M, Guzzella L, Optimal control of parallel hybrid electric vehicles, IEEE Trans Control Syst Technol* 12:352-363

[8] Augusto FA, Antenor PJ, Giorio S et al, Energy management fuzzy logic supervisory for electric vehicle power supplies system. *IEEE Trans Power Electron* 23:107-115

[9] CHAU, K. T.; WONG, Y. S. Overview of power management in hybrid electric vehicles. *Energy Conversion and Management*, 2002, 43.15: 1953-1968.

[10] John T. B. A. Kessels, Michiel Koot, Bram de Jager, Paul P. J. van den Bosch, N. (Edo) P. I. Aneke, and Daniel B. Kok, Energy Management for the Electric Powernet in Vehicles With a Conventional Drivetrain, *IEEE TRANSACTIONS ON CONTROL SYSTEMS TECHNOLOGY*, VOL. 15, NO. 3, MAY 2007,

[11] Langari, R., Jong-Seob Won, *Intelligent energy management agent for a parallel hybrid vehicle-part I: system architecture and design of the driving situation identification process,* IEEE Transactions on Vehicular Technology, (Volume:54 , Issue: 3)

[12] Salmasi, F.R., *Control Strategies for Hybrid Electric Vehicles: Evolution, Classification, Comparison, and Future Trends,* IEEE Transactions on Vehicular Technology, (Volume:56 , Issue: 5)

[13] O. Erdinc, B. Vural, M. Uzunoglu, *A wavelet-fuzzy logic based energy management strategy for a fuel cell/battery/ultra-capacitor hybrid vehicular power system,* Journal of Power Sources, Volume 194, Issue 1, 20 October 2009, Pages 369–380

Viscosity of Polymer Solutions Near Theta Temperature

[1]*Katarína PAULOVIČOVÁ (1^{st} year)*
Supervisor: [2]Jana TÓTHOVÁ

[1,2]Department of Physics, FEI TU of Košice, Slovak Republic

[1]paulovic@saske.sk, [2]jana.tothova@tuke.sk

***Abstract*—Viscometry has played an important historical role in advancing our knowledge of macromolecular structure and dynamics in solution. By measuring the solution viscosity, in principle, important polymer parameters, such as its relaxation times or the gyration and hydrodynamic radii can be determined. However, the interpretation of such experiments encounters serious difficulties with the treatment of the experimental data, but also because of fundamental problems with the theory of viscosity of polymer solutions. Particular problems reviewed in this contribution concern the so called theta temperature and the permeability of polymers to the solvent. Our first viscometry experiments and attempts to give their adequate description are also presented.**

***Keywords*—dilute polymer solutions, Huggins and Kraemer viscosity functions, theta temperature, viscosity.**

I. Introduction

Viscosity, light scattering, small-angle X-ray scattering, and osmotic pressure, are the main types of measurement performed in solutions. From these techniques viscometric measurements have played crucial role for determination of the conformational properties of polymer chains and subsequently, properties of polymers in solutions [1]–[3]. From historical point of view, the anomalously large viscosities of dilute polymer solutions were a key component of the arguments in advancing the macromolecular hypothesis. Theoretical analyses of the viscosity of impermeable particles enabled the application of viscometric measurements to determine the native molecular shapes of proteins in solution. Determination of intrinsic viscosities remains a widely used technique to obtain information about macromolecular structure in solution. The concentration dependence of polymer solution viscosity is influenced by thermodynamically driven changes in molecular conformation, as well as intermolecular interactions, which may be direct (hard-core, van der Waals, dipolar, electrostatic, hydrogen bonding) or indirect (hydrodynamic). Understanding the role of these interactions and their relationship to polymer structure is basic to controlled processing of polymers in the solution state [3]. In this way the starting point is to investigate the polymer solutions near their so called theta temperature, when the polymers behave like ideal chains. It allows using rigorous mathematical and physical methods to obtain “universal” polymer characteristics, which can be then improved for concrete chemical composition and structure for specific polymeric systems.

II. The Theta Temperature

Any macromolecule assumes a conformation in solution which is directed by the balance between the strengths of interaction of the polymer segments between themselves and with the solvent molecules. This balance is generically referred to as solvent quality. The interaction between two monomers (repeated segments of polymer) depends on the type of chain and on the solvent too. It is possible to express the potential energy of this interaction $u(r)$, as a function of the distance r between the monomers. Qualitatively, main features of $u(r)$ are quite common for all types of molecules and monomers: if r is small, $u(r)$ is positive and very large. This is because the monomers cannot penetrate into each other. As r becomes larger, monomers usually start to attract each other. The volume taken up by each monomer is automatically excluded from that available to any other one (hence the phrase “excluded volume”). To bring two monomers together, as close as r, some work has to be done. This work is stored in $u(r)$. It is done against the solvent molecules, as they need to be squeezed out of the way. Hence, the potential energy $u(r)$ represents the effective interaction of monomers through the solvent. It should depend, therefore, on the contents and state of the solvent, as well as on the temperature. The segment density n is very low so we can write for internal energy U of segment interaction

$$U = V k_B T n^2 B + n^3 C + ..., \tag{1}$$

where V is the volume of the coil, and B and C are expansion (or virial) coefficients. They are fully determined by the form of $u(r)$ and the temperature T. Obviously, the first term in (1) stands for the binary interactions. The most important values of r are those where $u(r) > 0$. So the internal energy of the coil (as well as the second virial coefficient B) is positive. In contrast, at lower T it is the “attractive” part of $u(r)$, where $u(r) < 0$, that gives the biggest contribution. So the internal energy of the coil U and B are both negative. We say that we are dealing with a good solvent, and in the latter case with a bad one. If in a solvent the segments of polymer chains tend to repel each other, the polymer will dissolve. Conversely, if the segments attract each other, the polymer chains stick together and precipitate out rather than dissolve. The quality of a solvent (i.e., whether it is good or bad) may change with its contents or with temperature. Hence, there has to be a special point where the second virial coefficient goes through zero: $B = 0$. It is usually called the theta-point (or Flory theta temperature T_Θ). At the T_Θ, attraction and repulsion between the segments completely cancel out, and the behavior of the polymer becomes ideal [4] forming random coils. The polymer parameters are then easily calculated and compared to experimentally measured quantities. Due to this much work has been devoted to the determination of T_θ. However, different methods could lead to different T_θ [5]. Moreover, often reliable values of T_θ are not known, since they are outside the range of the temperatures at which the experiments on polymer solutions are conducted, or even the temperatures when the polymer chains exist [6]–[8]. In such cases the interpretation of the experiments and the determination of the polymer parameters become difficult and ambiguous. To make a comparison between the theory and experiment, T_θ should be well defined.

Much effort has been devoted to find ways of careful interpretation of the viscometric experiments on polymer solutions and to determine the main phenomenological parameters of the polymer. The most popular ones come from the Rouse and Zimm bead-spring models [2], [4], [9]. A generalization of these theories includes the hydrodynamic interactions (HI) between the beads without the assumption of the impermeability of the polymer coil with respect to the solvent [10]. Contrary to the usual approach [2] to the interpretation of experimental data, the polymer dynamics is not considered in its non-draining limit. Only for infinitely strong HI the coil behaves as an impenetrable body. In general, the theory should contain the Rouse and Zimm models as limiting cases of small (Rouse) and large (Zimm) HI [9]–[12]. The strength of the HI is related to the draining parameter h, which appeared already in the famous work by Kirkwood and Riseman [13]. Its results for the intrinsic viscosity were corrected in [9], [10]. To describe the experiments, usually the draining effects are ignored. We have demonstrated that the "non-universal" theory with nonzero h much better describes the obtained experimental data.

III. Viscosity measurements near theta point

In our experiments we investigated Poly(2-ethyl-2-oxazoline) (also known as PEOX), the polymer, for which water is assumed to be nearly a theta solvent at room temperature (≈ 25°C) [14]. Its intrinsic viscosity scales with the molecular mass M as $\sim M^{0.56}$, with the exponent close to 0.6, as predicted for theta conditions by the Mark-Houwink equation for very high M polymers adopting ideally an Gaussian form conformation [15]. PEOX is a nonionic, synthetic water- and organic-soluble polymer, which is biocompatible, heat stable, and blends well with other polymers. Due to these properties it finds many applications, including the use in thermo-sensitive materials, as sensors, in drug delivery systems, and others [16]–[19]. Despite a number of papers devoted to the physico-chemical studies of PEOX, to our knowledge, the detailed comparison between the experiments and existing theories of the dynamics of dilute polymer solutions is absent.

PEOX in our experiments was a commercial product of Sigma-Aldrich Co. (Germany) without other modifications. The weight-average molecular weight was 500,000. The polymer was dissolved in deionized water. The apparatus used was an automated microviscometer (AMVn), combined with a density meter (DMA 4500 M), both from Anton Paar, Austria. The measurement principle of AMVn is the falling ball system consisting of the detection of a ball´s drop time in a diagonally mounted glass capillary filled with sample. The shear rate is influenced by changing the inclination angle of the capillary. The measurement system configuration with the capillary and ball diameters 1.6 mm and 1.5 mm, respectively, was used. It is suitable for the measuring range from 0.3 to about 10 mPa·s with reproducibility better than 0.5%. The temperature is measured with an uncertainty 0.05°C. The experimental data were treated with the software OriginPro (OriginLab, USA).

Since we were interested in the behavior of the PEOX solution near T_θ, the measurements were carried out at 20, 25, and 30°C. The polymer concentration c changed from 0.5 to 4 mg/ml, well below the concentration $1/[\eta]$ (for the intrinsic viscosity $[\eta]$ as determined in [15]). The sample thus can be considered as a dilute solution and up to the critical concentration $c \approx 2$ mg/ml (corresponding to the first permanent contacts between macromolecules in solution) even as an extremely dilute one [20]. In the treatment of the experimental data the effect of overlapping of different coils thus plays a minor role.

The experimental results are shown in Figs. 1 and 2. Let us denote the viscosity of the solution as η and that of the solvent as η_0. The relative viscosity, $\eta_{rel} = \eta/\eta_0$, showed no dependence on the shear rate, so that the solution can be treated as a Newtonian fluid. The results for η_{rel} for different temperatures and concentrations are presented in Fig. 1. The curves are second order polynomial fits $\eta_{rel} = 1 + B_1c + B_2c^2$. The standard errors of the coefficients are shown in the parenthesis: 1 (< 3×10^{-3}), B_1 (< 3×10^{-3} ml/mg), and B_2 (< 7×10^{-4} ml^2/mg^2). The fits match the curves obtained by simple linking of the experimental points.

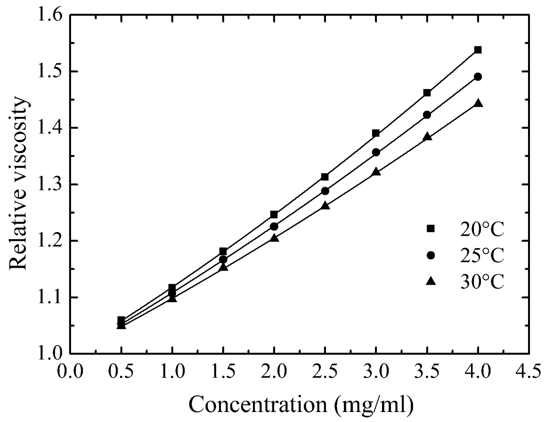


Fig. 1. Relative viscosity of the PEOX 500,000 aqueous solutions at different concentrations and temperatures.

IV. Treatment of viscometric data

At low concentrations the Huggins equation is used to describe the dependence on the polymer concentration c [1],

$$\eta_{red} = \eta_{sp}/c = (\eta_{rel} - 1)/c = [\eta](1 + k_H[\eta]c + ...). \qquad (2)$$

Here, η_{red} and η_{sp} are the reduced and specific viscosity, respectively, $[\eta]$ is the intrinsic viscosity, and k_H is the Huggins coefficient. Alternatively, the Kraemer expression for the viscosity is used,

$$\ln\eta_{rel} = c[\eta]\,(1 + k_K[\eta]c + \ldots), \tag{3}$$

where k_K is Kraemer´s constant. At low c approximately the equation $k_K \approx ½ - k_H$ should be obeyed. Figure 2 shows the dependences of the viscosity functions η_{sp}/c and $c^{-1}\ln\eta_{rel}$ on the concentration for three experimental temperatures.

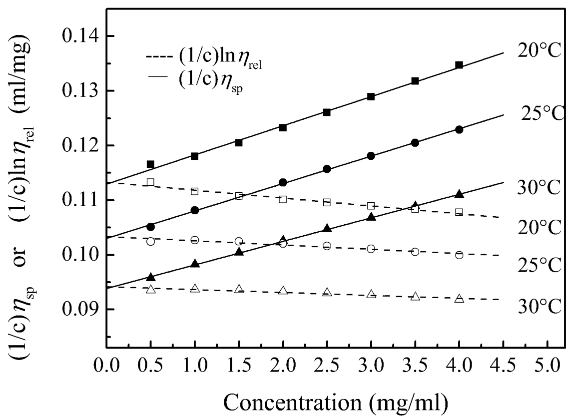


Fig. 2. Huggins and Kraemer viscosity functions from dilute to extremely dilute PEOX solutions.

TABLE I
INTRINSIC VISCOSITIES AND HUGGINS AND KRAEMER CONSTANTS OF PEOX SOLUTIONS

T(°C)	k_H	k_K	$k_H + k_K$	$[\eta]_H$ (ml/mg)	$[\eta]_K$ (ml/mg)
20	0.418 (0.418)	0.112 (0.113)	0.530 (0.531)	0.113 (0.113)	0.113 (0.113)
25	0.472 (0.453)	0.073 (0.087)	0.45 (0.540)	0.103 (0.103)	0.103 (0.104)
30	0.489 (0.464)	0.060 (0.078)	0.549 (0.542)	0.094 (0.094)	0.094 (0.095)

The dependences on c are well fitted by linear polynomials $A + Bc$, with the errors in the determination of the intercept A being less than 3.5×10^{-3} times smaller than A. The slopes of the lines, B, are determined with the errors 3×10^{-2} and 8×10^{-2} times smaller than B in the Huggins and Kraemer functions, respectively. The quantities A and B were used to calculate the intrinsic viscosities $[\eta]_H$ and $[\eta]_K$, and the constants k_H and k_K for different temperatures, as shown in Table 1. In parenthesis, the values are given that were obtained without fitting the relative viscosity by the second order polynomial. For every experimental value of the viscosity the Huggins and Kraemer functions were calculated and only then the linear fits were used to get the intrinsic viscosities and the constants $[\eta]_H$ and $[\eta]_K$. It is seen that the two approaches give close results lying within the experimental errors.

In the interpretation of the viscometry experiments on dilute polymer solutions usually the “universal” Flory theory of steady-state processes [1], which is consistent with the theory of Zimm [2] and that of Kirkwood and Riseman [13] in its non-draining limit, have been widely used for a long time. However, several experiments indicate a non-universality of these theories, possibly caused by draining effects. The first clear observation of the draining effect for long polymer chains, using static and dynamic light scattering experiments and viscometry, was reported in [21]. In the calculation of the viscosity of polymer liquids these effects can be accounted for based on the theory [9] of the long-time relaxation of the internal modes of flexible polymers. In this theory the polymer solution is described as a porous medium permeable to the solvent flow. This flow is governed by the linearized Navier-Stokes equation with an additional term $-\kappa^2\eta_0\vec{\upsilon}$ (the average value of the force acting on the liquid in an element of volume dV, provided that the average number of polymers in solution per dV is c), where $1/\kappa^2$, η_0, and $\vec{\upsilon}$ are the solvent permeability, viscosity and velocity, respectively. It holds $\kappa^2\eta_0 = c(t)f$, where f is the friction factor on one polymer chain that can be determined from the Einstein relation for the diffusion coefficient of the polymer coil, $D = k_BT/f$. The polymer is assumed to consist of $N >> 1$ beads connected by elastic forces that are given by the Gaussian equilibrium distribution of the beads [2]. The motion of the solvent created by the motion of beads is much faster than the motion of the coils, which determines the changes of $c(t)$. Thus, these changes around the equilibrium value c, which are due to the motion of the coils, are neglected. We do not a priori assume the validity of a concrete, Rouse or Zimm, model of the polymer dynamics. Only the strength of the HI determines which type of the polymer behavior is dominant. With the increase of c the Zimm contribution to the observable quantities (such as the coil diffusion coefficient or the viscosity of the solution) decreases and the polymer tends to behave (as distinct from the previous theories [2]) in correspondence with the Rouse model. Now the diffusion and the relaxation of the polymer internal modes depend on c. The diffusion coefficient is a sum of the Zimm and Rouse contributions (independent on c), $D(c) = D_Z(c) + D_R$, where D_R and $D_Z(0)$ are the well-known Rouse and Zimm limits for the diffusion coefficients [2].

In our experiments, $[\eta]$ near the assumed theta point at 25°C changes around 0.1 ml/mg, see Table 1. This is in good agreement with the result obtained in [14]. However, there is no evidence that at $T = 25$°C the solvent is exactly at the θ condition. Since at T_θ the exponent in the Mark-Houwink equation should be 0.6 giving a larger $[\eta]$, based on the data of Table 1, we rather expect that the theta temperature should be lower than 25°C. This is supported also by the theory [9]. Let us compare the theoretical value for the Huggins coefficient with the one determined from the experimental data. In the region of the expected theta temperature, k_H changes from 0.418 to 0.489. According to the theory [9], its value 0.3275 in the non-draining (Zimm) limit is notably lower. This indicates that the studied polymer coils are not perfectly impermeable to the solvent and behave partially as Rouse polymers. We have calculated, see Fig. 3, that between the limiting Rouse ($k_H = 0$) and Zimm value $k_H(h)$ represents a peak with the maximum $k_H = 0.417$ at $h \approx 2.92$. This maximal k_H is very close to the experimentally determined value 0.418 at $T = 20$°C. It thus suggests that the real theta temperature of the aqueous PEOX solution is 20°C rather than 25°C.

If the polymer coil is regarded as a sphere of volume V, from the familiar Einstein formula for the viscosity of suspensions, $\eta - \eta_0 = (5/2)\eta_0\phi$ [22] (ϕ is the volume fraction of polymer chains in solution), the intrinsic viscosity and the hydrodynamic (viscometric) radius are expressed as [11]

$$[\eta] = 2.5\frac{VN_A}{M}, \quad R_H = \left(\frac{3[\eta]M}{10\pi N_A}\right)^{1/3} \tag{4}$$

The equation gives R_H = 20.8 nm, close to the Zimm value 19.4 nm at $h \to \infty$. Within the theory [9] $\pi^{1/2}R_G/R_H = 8/3 + 2^{1/2}/h$, so that one obtains a lower value R_H = 16.4 nm.

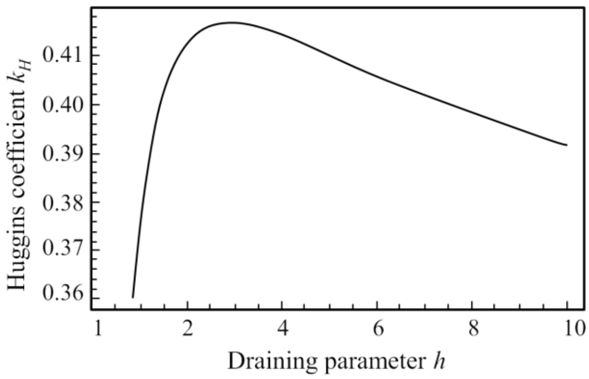


Fig. 3. Theoretical dependence of the Huggins coefficient on the draining parameter *h*.

Finally, let us assume that the shear viscosity of the studied solution is a thermally activated process. Using the Arrhenius equation to describe the temperature dependence of the viscosity [23], we have $\eta = A\exp(B/T)$. Here, T is the absolute temperature, $B = B_0 + \alpha(M)c$, and $\ln(A/A_0) = \beta(M)c$, with A_0 and B_0 being the values for the solvent. By comparison of this formula with the representation of viscosity at $c \to 0$ in Eq. 1, $\eta = \eta_0(1 + [\eta]c + \ldots)$, one finds $[\eta] \approx \beta + \alpha/T$. The experimental results for the intrinsic viscosity presented in Table 1 are excellently described by this equation with the constants α = 162.7 K ml/mg and β = –0.443 ml/mg.

In conclusion, contrary to the "universal" Flory theory of steady-state transport processes in dilute polymer solutions [1], the polymer characteristics, such as the diffusion coefficient of the coils or their relaxation times of internal modes, seem to depend on the draining parameter connected with the permeability of the coils with respect to the solvent and the strength of the hydrodynamic interactions between the polymer segments. Although several experiments indicated non-universality in the polymer behavior, the polymers are usually considered in the non-draining limit, which corresponds to the infinitely large draining parameter h. We have calculated the Huggins coefficient and compared it to its experimentally determined value from dilute PEOX solutions. We have found that for h corresponding to the maximum of the peak in the theoretically predicted function $k_H(h)$ the experimental value of k_H exactly agrees with the experiment. This happens at the temperature T = 20°C. Since the used theory is built for the theta condition, we propose that the theta temperature for PEOX is 20°C rather than 25°C known from the previous studies.

In our opinion, further investigations are needed to solve this controversy. This can be done, e.g., by studying the dependence of the intrinsic viscosity on the polymer mass at different temperatures. This will be the subject of our future study.

ACKNOWLEDGMENT

This work was supported by the Slovak Ministry of Education Agency for the Structural Funds of the EU within the projects 26220120097, 26220220182, and by the Grant VEGA 1/0348/15.

REFERENCES

[1] P. J. Flory, *Principle of Polymer Chemistry*. Ithaca: Cornell Univ. Press, 1953.
[2] M. Doi, S. F. Edwards, *The Theory of Polymer Dynamics*. New York: Oxford Univ. Press Inc., 1986.
[3] L. A. Utracki, A. M. Jamieson, Eds., *Polymer Physics: From Suspensions to Nanocomposites and Beyond*. New York: John Wiley & Sons, Inc., 2010.
[4] A. Yu. Grosberg, A. R. Khokhlov, *Statistical Physics of Macromolecules in Polymers and Complex Materials*. Woodbury: American Institute of Physics Press, 1994.
[5] K. Kamide, S. Matsuda, "A method for determining the Flory theta temperature and entropy parameter of single- or multicomponent-polymer–single solvent systems from the critical solution point data", *Polymer J*. vol. 16, no. 11, 1984, pp. 825-837.
[6] H. Yan, Y. Yan, P. Zhu, H. Li, Q. Zhu, C. Fan, "Studies on the viscosity of polymer solutions at low concentrations", *Eur. Polymer J*. vol. 41, no. 2, 2005, pp. 329-341.
[7] J. Tothova, V. Lisy, "Relaxation times of flexible polymer chains in solution from non-conventional viscosity measurements", *The Open Macromol. J.* vol. 4, Special Issue, 2010, pp. 26-31.
[8] K. Bechekh, N. Ghaouar, "Rheological properties of Polyethylene Glycol (PEG 35000): An interpretation of a negative intrinsic viscosity and a high Huggins coefficient value", *J. Macromol. Sci. B* vol. 53, no. 3, 2014, pp. 391-397.
[9] V. Lisy, J. Tothova, A. V. Zatovsky, "The joint Rouse-Zimm theory of the dynamics of polymers in dilute solutions", *Cond. Mat. Phys*. vol. 9, no. 1, 2006, pp. 95-102.
[10] V. Lisy, J. Tothova, A. V. Zatovsky, "Long-time dynamics of Rouse-Zimm polymers in dilute solutions with hydrodynamic memory", *J. Chem. Phys.* vol. 121, pp. 21, 2004, pp. 10699-10706.
[11] R. G. Larson, "The rheology of dilute solutions of flexible polymers: Progress and Problems", *J. Rheol.* vol. 49, no. 1, 2005, pp. 1-70.
[12] H. Yamakawa, "*Modern Theory of Polymer Solutions*", in Harper´s Chemistry Series,The Electronic Edition. New York: Harper and Row, 2001.
[13] J. G. Kirkwood, J. Riseman, "The intrinsic viscosities and diffusion constants of flexible macromolecules in solution", *J. Chem. Phys.* vol. 16, no. 6, 1948, pp. 565-573.
[14] T. T. Chiu, B. P. Thill, W. J. Fairchok, "Poly (2-ethyl-2-oxazoline): A new water-and organic soluble adhesive", in Water-Soluble Polymers, *Advances in Chemistry* vol. 213, Chapter 23, 1986, pp. 425-433.
[15] V. Aseyev, H. Tenhu, F. M. Winnik, "Non-ionic thermoresponsive polymers in water", *Adv. Polym. Sci.* vol. 242, 2011, pp. 29-89.
[16] Y. Han, Z, He, A. Schulz, T. K. Bronich, R. Jordan, R. Luxenhofer, A. V. Kabanov, "Synergistic combinations of multiple chemotherapeutic agents in high capacity Poly(2-oxazoline) micelles", *Mol. Pharmaceutics* vol. 9, no. 8, 2012, pp. 2302-2313.
[17] H. Schlaad, R. Hoogenboom, "Poly(2-oxazoline)s and related pseudo-polypeptides", *Macromol. Rapid. Commun*. vol. 33, no. 19, 2012, pp. 1599-1599.
[18] N. Adams, U. S. Schubert, "Poly(2-oxazolines) in biological and biomedical application Contexts", *Adv. Drug Delivery Reviews* vol. 59, no. 15, 2007, pp. 1504-1520.
[19] J. C. M. E. Bender, R. Hoogenboom, P. A. A. van Vliet, "Drug delivery system comprising Polyoxazoline and a bioactive agen", *U.S. Patent* No. US 8,642,080 B2, 2014.
[20] A. Dondos, C. Tsitsilianis, G. Staikos, "Viscometric study of aggregation phenomena in polymer dilute solutions and determination of the critical concentration c**", *Polymer* vol. 30, no. 9, 1989, pp. 1690-1694.
[21] T. Konishi, T. Yoshizaki, H. Yamakawa, "On the "Universal Constants" ρ and Φ of flexible polymers", *Macromolecules* vol. 24, no. 20, 1991, pp. 5614-5622.
[22] A. Einstein, "Eine neue Bestimmung der Moleküldimensionen", *Ann. Physik* vol. 19, Suppl. , 1906, pp. 289-306.
[23] W. R. Moore, "Viscosity-temperature relationships for dilute solutions of high polymers", *Nature* vol. 191, 23 Sept. 1961, pp. 1292-1293.

Section:
Informatics & Telecommunications

A cloud-based multi-robot system

[1]*Tomáš CÁDRIK (2nd year)*
Supervisor: [2]Marián MACH

[1,2]Dept. of Cybernetics and Artificial Intelligence, FEI TU of Košice, Slovak Republic

[1]tomas.cadrik@tuke.sk, [2]marian.mach@tuke.sk

***Abstract*—Cloud computing and cloud robotics are getting more and more popular. Because of this we chose to create a cloud-based platform for helping robotic agents to solve tasks in an environment. We want to use the advantages of cloud in robotics and to design and create methods that will increase the effectiveness of the system. The system itself and research done are described in this paper.**

***Keywords*—cloud, cloud robotics, learning, multi-robot system, rules.**

I. Introduction

Cloud robotics is a relatively new field in robotics. This area is the fusion of cloud computing and robotics. The motivation for this topic came from two existing cloud robotic systems. The first one is called RoboEarth [1]. This system communicates with agents using the Robotic operating system. It has a database for storing knowledge and information about the environment. It can communicate and control multiple types of robotic agents. The manufactural information of each robotic platform is added to this cloud platform. Another cloud robotics system that inspired my work is called DaVinci [2]. This system can use parallelization advantages of the cloud. The goal of my work is to create an alternative to those mentioned systems, where also new methods for increasing the effectiveness of how the system will work and how robotic agents will react to the environment be added.

II. Analysis

The goal consists of two subgoals. The technological subgoal is the usage of a cloud-based platform for the support of robots. For this purpose, we do not want to create a new cloud platform, but utilize an existing one. For now we choose Azure provided by Microsoft [3], which has all needed features that a cloud platform needs to have.

For the scientific goals of our work, we want to base our work on two existing methods. One of the scientific goals is a modification of a method that will dynamically react to the change of the number of inputs or outputs. The first method we want to utilize uses the MF-ARTMAP neural network. We have chosen this method because of its feature of dealing with the problem where the possible number of outputs increases. The second method is the learning classifier system [4] which is a population algorithm using reinforcement learning mechanisms and genetic operators. We chose this method because individuals can have different number of inputs and also can be used when the number of possible actions increases, but for now it cannot effectively deal with the problem where the relation between inputs and outputs is suddenly changed. We also want to create a method for extracting rules from a behavior encoded in another form, not in rules corresponding condition-action patterns. There were some mechanisms for this as in [5], but those methods are not universal enough (they do not work for multiple algorithms).

III. Summary of solved and unsolved problems in the cloud-based multi-robot system

We proposed the architecture of the cloud-based multi-robot system that was the first time described in ATP journal [6]. This scheme is shown in Fig. 1. This system will be used as the support for robots and will be used for sparing resources of the robot and sharing data among robots.

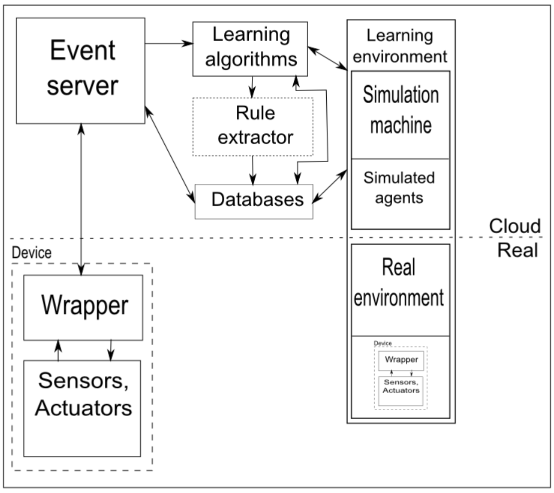


Fig. 1: Architecture of the cloud-based multi-robot system modified from [6].

Each part of this system was detailed described in a series published in the ATP journal. The first part described the functionality of the system in general and three blocks of the system: Event server, devices, and rule extractor [6]. The second part of this series described the AI brick approach and the block of learning algorithms [7]. The last part of the series was about the blocks for learning environment and databases [8]. One part in the previous series in ATP journal was also focused on the cloud-based architecture for supporting multiple robotic platforms. It was mainly focused on communication between devices and the cloud service through the Simple object access protocol technology using the Windows Communication Foundation technology from Microsoft.

We were also focusing on designing a mechanism for extracting rules from an output produced by a learning algorithm. For now we focused only on the rule extraction from the Zeroth-level classifier system (ZCS). The first important method we published was [9] on the Mendel conference in Brno. The scheme of this rule extractor is shown in Fig. 2.

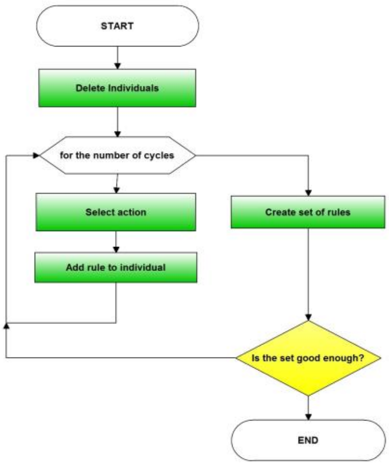


Fig. 2: Diagram of a rule extractor from [9]

We trained a ZCS and then extracted rules from it. Then we used the ZCS and a simple production system with the extracted rules for controlling an agent in the environment. The production system had better results as the classifier system that was learning continuously from the environment. The drawback was that this method is not universal enough. It is possible to use this method only if we want to extract rules from learning algorithms that are population based, and the rules are coded in the form of individuals.

The second method for rule extraction was published on the conference SCIS & ISIS 2014 in Japan. We named it "Interference of waves based usage of an optimization algorithm" [10]. The method uses multiple optimization algorithms (in this case evolutionary algorithms) for extracting rules from an output of a learning algorithm. It uses the combination of sensory inputs and actions used by the output from learning/controlling the agent within the environment. These conditions and actions are used as the input to the extraction method. At the beginning, the method evolves only one rule and when the evolution process ends all the methods start evolving the second rule, where the first one remains the same. This incremental process goes on until all evolutionary algorithms will not have the fitness values in the actual cycle less or equal to the fitness in the previous learning period. The best solution from the entire set of individuals is the solution.

In addition, we have created a cloud service for storing data recorded from the Kinect sensor (skeleton). This service recalculates those movements so that the NAO robot can use those movements. We also created a wrapper that can connect to that cloud service and download the action that was selected through the website.

The paper [10] has been chosen as one of the 45 papers, the extended version of which can be submitted for publication in one of the five journals indexed in the Scopus database. Another significant success was that we contributed to a proposal of a cloud project for robotic devices. On the basis of this proposal, we have obtained a grant for resources on Azure cloud to the value of 40,000 dollars.

IV. Future Work

Next we will test the rule extraction method on a more complex task as the animat problem (agent who needs to go from the start position to the position with food). We want to test this method not only with the ZCS but also in connection with another learning mechanism. The plan is to test it for extracting rules from neural networks. We will test the particle swarm optimization algorithm instead of the evolutionary based method because particle swarm optimization is built primarily for numerical optimization (neural networks works with float values instead of Boolean or integer values).

In the close future, the goal is the improvement of the wrapper system, so the wrapper does not need to be specialized for one type of the task as it is now. Another technological aim is to select an existing simulator and make a mechanism where the cloud services will be able to communicate with the simulation core of that simulator.

V. Conclusion

This paper starts with a brief introduction to the field of cloud robotics. Two systems: RoboEarth and DaVinci framework are mentioned as the motivation why we choose this topic for our research. The paper also describes the analysis of facts which serve as the starting point for our work. Next section describes research and publications. The last section is a list of future work tasks.

Acknowledgment

Research supported by the National Research and Development Project Grant 1/0667/12 "Incremental Learning Methods for Intelligent Systems" 2012-2015.

References

[1] "RoboEarth." [Online]. Available: http://roboearth.org/.
[2] R. Arumugam, V. R. Enti, K. Baskaran, and A. S. Kumar, "DAvinCi: A cloud computing framework for service robots," in 2010 IEEE International Conference on Robotics and Automation, 2010, pp. 3084–3089.
[3] "Microsoft Azure." [Online]. Available: http://azure.microsoft.com/en-us/.
[4] O. Sigaud and S. W. Wilson, "Learning Classifier Systems: A Survey," Soft Comput. - A Fusion Found. Methodol. Appl., vol. 11, no. 11, pp. 1065 – 1078, 2007.
[5] G. Liu and L. Yang, "An improved neural network and its applicable study," in 2010 International Conference on Intelligent Computation Technology and Automation, ICICTA 2010, 2010, vol. 1, pp. 567–570.
[6] Cádrik, T. – Ondo, J. – Mach, M. – Sinčák, P.: Základná architektúra cloudového prostredia na podporu multirobotických systémov (1). In: ATP Journal,Roč. 22, č. 1 (2015), s. 37-38 – ISSN 1335-2237.
[7] Cádrik, T. – Ondo, J. – Mach, M. – Sinčák, P.: Základná architektúra cloudového prostredia na podporu multirobotických systémov (2). In: ATP Journal, Roč. 22, č. 2 (2015), s. 39-41 – ISSN 1335-2237
[8] Cádrik, T. – Ondo, J. – Mach, M. – Sinčák, P.: Základná architektúra cloudového prostredia na podporu multirobotických systémov (3). In: ATP Journal, Roč. 22, č. 3 (2015), s. 46-47 – ISSN 1335-2237.
[9] T. Cádrik and M. Mach, "Extracting rules from ZCS evolutionary classifier system for the purpose of future usage in robotic systems," in MENDEL 2014: 20th International Conference on Soft Computing, 2014, pp. 393–396.
[10] T. Cádrik and M. Mach, "Interference of Waves Based Usage of an Optimization Algorithm for Finding Rules in an Agent System," in SCIS and ISIS 2014: proceedings, 2014, pp. 1068–1072.

A Final Look Back at Underactuated Systems

[1]*Slávka JADLOVSKÁ (4th year),*
Supervisor: [2]Ján SARNOVSKÝ

[1,2]Dept. of Cybernetics and Artificial Intelligence, FEEI TU of Košice, Slovak Republic

[1]slavka.jadlovska@tuke.sk, [2]jan.sarnovsky@tuke.sk

***Abstract*— The presented paper describes the state of our research in the area of modeling and optimal control of underactuated mechanical systems. A particular focus is placed on the results obtained during the past year. Relevant open problems are briefly outlined to stimulate further research on the subject.**

***Keywords*—underactuated mechanical systems, Lagrangian mechanics, optimal control techniques, hybrid systems theory**

I. Introduction

Underactuated systems, defined as *nonlinear mechanical systems with fewer control inputs than degrees of freedom*, appear in a broad range of applications including robotics, aerospace, marine and locomotive systems. The ultimate motivation behind the research into underactuation is the ability to control nonlinear systems without complete control authority by exploiting their *natural dynamics*, which is similar to how biological systems execute motions involving a loss of instantaneous control authority [1]. Underactuated devices are therefore expected to be more efficient, simpler and more reliable than their fully actuated alternatives with the drawback that the control of underactuated devices is more complex theoretically. Strong structural properties which facilitate control design for fully actuated systems (feedback linearizability, passivity, linear parametrizability) are usually lost in underactuated systems, while at the same time undesirable properties (higher relative degree, nonminimum phase behavior) tend to appear [2].

The presented PhD thesis, titled *Modeling and Optimal Control of Nonlinear Underactuated Dynamical Systems*, is the first doctoral thesis at the DCAI-FEEI to comprehensively explore the topic of underactuated systems. As formulated in the Feb. 2013 proposal, the general objective of the thesis is to identify and pursue open research problems occurring at the mutual overlaps between three principal areas: mathematical modeling of underactuated mechanical systems, optimal control of nonlinear systems subject to constraints, and hybrid systems theory, notably switching control structures. This paper details which aspects of this objective have been accomplished and to what extent, and describes the algorithms and supporting tools which were developed to achieve selected tasks.

II. A Generalized Approach to Modeling and Control of Underactuated Systems

Benchmark underactuated systems, such as inverted pendulum systems (IPSs), two-link planar robots *Acrobot* and *Pendubot*, *the Inertia Wheel Pendulum* or the *convey-crane* system create complex low-order nonlinear dynamics which enables us to gain insight into the principles of modeling and control of advanced, higher-order underactuated systems. We started to develop a set of algorithms which derive the Euler-Lagrange equations of motion for a *selected benchmark multi-body underactuated system*. We specifically introduced the concept of a *generalized* (*n-link*) *inverted pendulum system*, which allows us to treat an arbitrary system of interconnected inverted pendula as a particular instance of the system of *n* pendula attached to a given stabilizing base, such as a cart or a rotary arm. Physical derivation and MATLAB implementation of general procedures of automated model generation for classical and rotary IPSs were first presented in [3] together with generated motion equations of example systems and verification of their validity. For classical IPSs, more generalizations were soon introduced into the procedure, which by now covers all possible combinations of initial assumptions for the pendulum reference position and direction of rotation, and considers various shapes of weight load attached to the last pendulum link.

Optimal control algorithms based on the quadratic functional minimization were implemented and verified in a variety of control structures. Initial case studies which covered modeling and stabilizing control of the classical double and the rotary single IPS were published in [4][5]. Simulation experiments were conducted using pre-prepared blocks from a *Simulink* block library which was developed as a software framework for the analysis and control of IPSs [6]. The extended version [7] of [5] presented a complex problem of swinging up and subsequent stabilization of a rotary single IPS in the unstable position, which involved a hybrid setup of a swing-up and balancing controller in a *switching control structure*. The potential of nonlinear control techniques was first explored in [7] which examined the *state-dependent Riccati equation* (SDRE) algorithm and generalized it to be used with an *n-link* IPS. It was next verified on a rotary single IPS that the SDRE-based controller, which solves a separate optimal control problem at each time step using original nonlinear system dynamics, gradually outperforms the conventional LQR algorithm as the state vector moves away from the equilibrium.

The above-mentioned contributions are the result of the Vega project implementation *Dynamic Hybrid Architectures of the Multi-agent Network Control Systems* (No. 1/0286/11), completed in Jan. 2015, and in part of KEGA project implementation *CyberLabTrainSystem–Demonstrator and Trainer of Information-Control Systems* (No. 021TUKE-4/2012).

III. Recent Advances in Modeling and Control of Underactuated Systems

The past year's efforts have been predominantly directed towards the completion of the PhD thesis. The interim results of the thesis have so far been summarized several times, starting with an award-winning survey paper on the main achieved results in modeling and control of underactuated systems by worldwide research groups [9], and a total of three chapters of a monograph [10], which respectively focus on the results obtained in the three main areas of our research by the end of 2013. Another summary paper [11] gave a report on the previous year's results. Finally, a poster presentation was prepared for the competition for PhD students held on Nov. 10, 2014 as part of the *Science and Technology Week* national event – mine being one of the two PhD theses-in-making nominated to represent FEEI. In addition to the overall motivation, each of the three main areas of the future thesis was characterized by corresponding theoretical contributions, developed software tools, and example simulation results.

It is our ultimate goal to complete a readily available collection of mathematical/simulation models of underactuated mechanical systems to serve as a testbed model basis in simulation experiments exploring their properties and testing control strategies. As a follow-up to our previous work, algorithms for automated model generation with an optional reference link position have recently been developed for other typical underactuated systems, such as the *Acrobot*, *Pendubot* and *inertia wheel pendulum*, rotary IPSs, and the 3D convey-crane system. Although these will first occur in the thesis itself together with their detailed classification, our earlier work on the expanded model generation procedure for classical IPS was recently published as a Springer book chapter [12].

We have further explored control problems of under-actuated systems which require the use of *switching control structures*. The hybrid control setup composed of a swing-up and balancing controller was tested for several *pendulum-like* benchmark systems (IPSs, *Acrobot/Pendubot*), where the swing-up was performed via energy-based methods according to Lyapunov, or via partial feedback linearization techniques. Since linear quadratic optimal control was confirmed as a reliable control technique for a class of underactuated systems we went on to analyze *model predictive control* (MPC) algorithms with respect to their suitability for stabilizing control of benchmark underactuated systems.

Hybrid systems theory provides a framework for modeling and control of systems characterized by an interaction between continuous and discrete dynamics, as was the case of phase transitions in a particle movement system [13]. In particular, hybrid MPC is suitable for systems defined by switched linear dynamics subject to constraints on state/input variables.

The obtained results have provided us a starting point for the research of advanced underactuated systems, notably *biped robots and the mechanism of robot walking*. Walking robots are often based on benchmark underactuated systems and manage to exploit their dynamics to create naturally-looking gait. Hybrid models and switching control structures are useful if event-based dynamics of legged locomotion is considered.

IV. Conclusion and Plans for Future Research

This paper summarizes the results that have been achieved in the area of underactuated mechanical systems in the final stages of the PhD program. The principal aim of our research was to identify and overcome the critical aspects of control algorithm design caused by the disadvantageous physical properties of underactuated systems. In order to contribute to the modeling and control education at the DCAI-FEEI TU, underactuated systems are being integrated into the research/teaching activities of the *Center of Modern Control Techniques and Industrial Informatics* at the DCAI. In addition to the developed library of simulation models, laboratory models of classical and rotary inverted pendulum are being prepared to enable control algorithm verification.

Acknowledgment

This contribution was supported by the Research and Development Operational Program for the "University Science Park Technicom for innovative applications with knowledge technology support" project (ITMS code 26220220182), co-financed by the ERDF – 100%.

References

[1] R. Tedrake, Underactuated Robotics: Learning, Planning and Control for Efficient and Agile Machines. Course Notes for MIT 6.8., Cambridge: Massachusetts Institute of Technology (MIT), 2009.

[2] M. W. Spong, "Underactuated Mechanical Systems: Control Problems in Robotics and Automation", in *Lecture Notes in Control and Information Sciences*, vol. 230, 1998, pp. 135-150

[3] S. Jadlovská, J. Sarnovský, "Modelling of Classical and Rotary Inverted Pendulum Systems – a Generalized Approach", in *Journal of Electrical Engineering*, vol. 64, no. 1, 2013, pp. 12–19, ISSN 1335-3632.

[4] S. Jadlovská, J. Sarnovský, "Classical Double Inverted Pendulum - a Complex Overview of a System", *Proc. of the IEEE 10th Jubilee Int. Symp. on Applied Machine Intelligence and Informatics – SAMI 2012*, Herľany, Slovakia, pp. 103-108, ISBN: 978-1-4577-0195-5

[5] S. Jadlovská, J. Sarnovský, "A Complex Overview of the Rotary Single Inverted Pendulum System", *Proc. of the 9th Int. Conf. – ELEKTRO 2012*, Žilina - Rajecké Teplice, Slovakia, pp. 305-310, ISBN: 978-1-4673-1178-6.

[6] S. Jadlovská, J. Sarnovský, "An Extended Simulink Library for Inverted Pendula Modeling and Simulation", *Proc. of the 19th Int. Scient. Conf. Tech. Computing 2011*, Prague, Czech Rep, ISBN 978-8070807941.

[7] S. Jadlovská, J. Sarnovský, "A Complex Overview of Modeling and Control of the Rotary Single Inverted Pendulum System", in *Advances in Electrical and Electronic Engineering*, vol. 11, no. 2, 2013, pp. 73–85, ISSN 1336-1376.

[8] S. Jadlovská, J. Sarnovský, "Application of the State-Dependent Riccati Equation Method in Nonlinear Control Design for Inverted Pendulum Systems", *Proc. of the IEEE 11th Int. Symp.on Intelligent Systems and Informatics - SISY 2013*, Subotica, Serbia, 2013, pp. 209-214, ISBN: 978-1-4799-0303-0.

[9] S. Jadlovská, "Modeling and Optimal Control of Nonlinear Underactuated Mechanical Systems – a Survey", *Proc. of the 13th Scientific Conference of Young Researchers – SCYR 2013*, Herľany, Slovakia, pp. 32-35, ISBN: 978-80-553-1422-8.

[10] A. Jadlovská, S. Jadlovská, *Modern Methods of Modeling and Control of Nonlinear Systems* [Moderné metódy modelovania a riadenia nelineárnych systémov]. Košice: Elfa, 2013, 279 p., ISBN: 978-80-8086-228-2, ch. 1-3.

[11] S. Jadlovská, "Underactuated Mechanical Systems: Recent Findings and Future Challenges", *Proc. of the 14th Scientific Conference of Young Researchers – SCYR 2014*, Herľany, Slovakia, pp. 275-276, ISBN: 978-80-553-1714-4.

[12] S. Jadlovská, J. Sarnovský, J. Vojtek, D. Voščel, "Advanced Generalized Modeling of Classical Inverted Pendulum Systems", in *Advances in Intelligent Systems and Computing : Emergent Trends in Robotics and Intelligent Systems*, vol. 316, Switzerland : Springer, 2014, pp. 255-264, ISBN 978-3-319-10782-0, ISSN 2194-5357

[13] S. Jadlovská, A. Jadlovská, P. Rabatin, M. Šviková, "Simulation of Particle Movement in a Two-Dimensional Model System Using MATLAB", *Proc. of the Int. Scient. Conf. Technical Computing 2014*, Bratislava, Slovakia, ISBN 978-80-7080-898-6

Advanced Architectures Distributed Systems for the Implementation of Neural Networks

[1]Marek ČOPJAK (2st year)
Supervisor: [2]Martin TOMÁŠEK

[1,2]Dept. of Computers and Informatics, FEI TU of Košice, Slovak Republic

[1]marek.copjak@tuke.sk, [2]martin.tomasek@tuke.sk

Abstract— **This article focuses on the description of prior research on the use of GPGPU technology for the distribution of neural networks and on my progress in the previous year in a distributed architecture for the implementation of the neural network and plan future research in this area.**

***Keywords*—Neural network, GPGPU, CUDA, OpenCL, distribution.**

I. INTRODUCTION

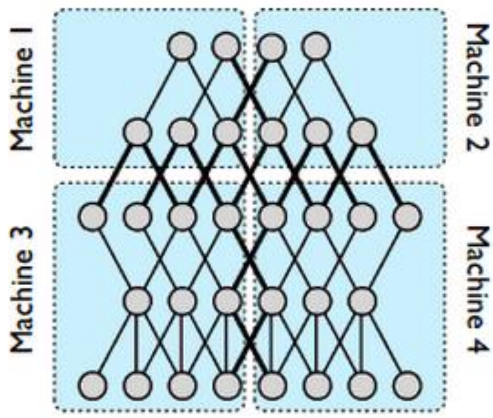
Neural networks are becoming increasingly popular in application areas where the emphasis is focused on obtaining results in real time. Although there are different implementations of neural networks available to sequentially operating computer systems, a large number of these patterns requires an enormous amount of time for training and the life stage of the neural network, where the network is large. Artificial neural networks are massively working parallel systems with a large number of simple computational elements. Therefore, it is natural to try to implement such a system of parallel computing on. As one approach to working parallel to implement a neural network is used GPGPU technology, which nowadays provides high computing power for working parallel computations. The aim of the thesis is to design a distributed system architecture for neural networks using GPGPU. The distribution of data to the neural network computing system will take care of the algorithm, which ensures efficient power distribution and thus achieves a reduction of time spent.

II. CURRENT STATE OF THE ART IN NEURAL NETWORK IMPLEMENTATIONS

In recent years, increasing the data in the commercial and academic sphere has gained great dimensions. As the consequence, the examination of the possibilities of parallelism and distribution of machine learning was tried by several authors [1], [2], [3], [4], [5], [6], [7]. Most of this research has been focused on the convex linear models where the distribution gradient calculation is the first essential step.

In this area it has been managed to achieve a reduction claims to be synchronized with the review of late updates convex problems [2], [7]. Other groups working in parallel with this research address the problems of fine gradients (sparse gradients) (issues where only a tiny fraction of the coordinates of the vector gradient is zero for this example) investigated the induction gradients, stochastic process, which does not require mutexes (lock memory) on shared memory architectures (eg. single machine) [8], [9].

Jeffrey Dean et. al. deals with the idea that combines the best of both approaches, allowing the use of cluster, Asynchronous Counting gradients, but without it, the problem is either convex or fine [10]. In the context of deep learning, most work has been focused on training relatively small models of individual computers (eg., Theano [11]). Application for extension for deep learning involves the use of GPU for training deck large number of small models and subsequent averaging the prediction [12], or by modifying standard deep network for their parallelism [13]. The examination is based on scaling deep learning techniques in the direction of learning large models containing several billion parameters, but without this leading to a reduced form model. In specific cases, where one layer controls calculations, some authors consider the distribution calculations in a single layer and replication calculations in other layers [8].

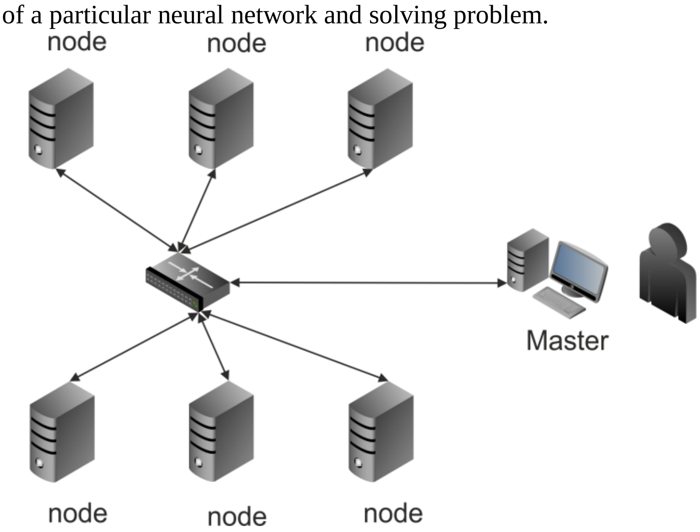


In general, where the use of a number of layers of the model is computationally intensive process, it is necessary to parallelize this model [14]. To the calculation be successful, the model parallelism must be properly joined and combined with distributed optimization techniques, which uses data parallelism. It were considered large numbers of existing computational tools to use for a given problem, MapReduce [15] and GraphLab [16]. Finally, an in-depth examination was that Mapro -Duce designed for parallel data processing is not suitable for iterative calculations associated with deep learning networks, while GraphLab intended for general (unstructured) graph calculations will use the computing power available in structured graphs, usually located in deep networks.

III. Proposal of distributed architecture for neural network implementation

The aim of the chapter is a description of the dissertation proposal distributed architecture using the CPU or the use of high performance GPGPU technologies of parallel data processing. Large-scale neural networks require large computing power to be able to process data in real time. The proposed distributed architecture can enhance the performance and closer to the processing of data in real time. The proposed model of distributed architecture shows Distribute different neural network system computing nodes that will be available. The algorithm of proposed architecture takes care of the distribution of data and uniform load nodes and post-processing partial results into a single unit in the master.

Due to the distribution of the neural network, there is a need to create a design file system for labeling each neuron. This file system will take care about where computing nodes are and how neurons will communicate with each other (exchange data). Individual results of neurons depend on the deployment of a particular neural network and solving problem.



IV. Future work

Given the current stage of the analysis of the problems of neural networks and their implementation on a distributed architecture using GPGPU technology seems completely unexplored. Further work should focus on the analysis of existing opportunities for cooperation of these sectors and to provide a more optimal solution by proposing a model of distributed architecture, which makes use of existing knowledge for efficient data processing in neural networks. The result is a chance to move closer to the reduction of time-consuming calculations.

Future work aims are divided into three different steps:

- Design and implementation of distributed architectures for neural networks
- Design of the optimization algorithms for uniform load nodes of distributed architecture for neural networks
- Benchmark to measure the performance of distributed architectures for neural network implementation and measurement.

V. Conclusion

From the analysis of the current state of the problem of neural networks and GPGPU technology, was designed model of distributed architecture for neural networks. Current trends and innovations constantly confirmed, that without parallel computing , difficult and demanding tasks cannot be solved in reasonable time. The options of distributed computing power of neural networks using GPGPU and deployed on such new technologies were examined.

Acknowledgment

This work was supported by the Slovak Research and Development Agency under the contract No. APVV-0008-10 and project KEGA 008TUKE-4/2013: Microlearning environment for education of information security specialists.

References

[1] SHI, Q., PETTERSON, J., DROR, G., LANGFORD, J., SMOLA, A., STREHL, A., VISHWANATHAN, V., 2009. Hash kernels. In AISTATS, 2009.
[2] LANGFORD, J., SMOLA, A., ZINKEVICH, M., 2009. *Slow learners are fast*. In NIPS, 2009.
[3] MANN, G., McDONALD, R., MOHRI, M., SILBERMAN, N., WALKER, D., 2009. *Efficient large-scale distributed training of conditional maximum entropy models.* In NIPS, 2009.
[4] McDONALD, R., HALL, K., MANN, G., 2010. *Distributed training strategies for the structured perceptron.* In NAACL, 2010.
[5] ZINKEVICH, M., WEIMER, M., SMOLA, A., LI, L., 2010. *Parallelized stochastic gradient descent.* In NIPS, 2010.
[6] AGARWAL, A., CHAPELLE, O., DUDIK, M., LANGFORD, J., 2011. *A reliable effective terascale linear learning system.*. In AISTATS, 2011.
[7] AGARWAL, A., DUCHI, J., 2011. *Distributed delayed stochastic optimization.* In NIPS, 2011.
[8] BENGIO, Y., DUCHARME, R., VINCENT, P., JAUVIN, C., 2003. *A neural probabilistic language model.* Journal of Machine Learning Research, 3:1137–1155, 2003.
[9] NIU, F., RETCH, B., RE, C., WIRGHT, S. J., 2011. *A lock-free approach to parallelizing stochastic gradient descent.* In NIPS, 2011.
[10] DEAN, J., CORRADO, S. G., MONGA, R. *Large Scale Distributed Deep Networks [online].* Google Inc., Mountain View, CA, *[online]*: https://docs.google.com/viewer?url=http\%3A\%2F\%2 Fresearch.google.com\%2Fpubs\%2Farchive\%2F40565.pdf.
[11] BERGSTRA, J., BREULEUX, O., BASTIEN, F., LAMBLIN, P., PASCANU, R., DESJARDINS, G., TURIAN, J., WARDE-FARLEY, D., BENGIO, Y., 2010. *Theano: a CPU and GPU math expression compiler.* In SciPy, 2010.
[12] CIRESAN, D., MEIER, U., SCHMIDUBER, J., 2012. *Multi-column deep neural networks for image classification..* Technical report, IDSIA, 2012.
[13] DENG, L., YU, D., PLATT, J., 2012. *Scalable stacking and learning for building deep architectures.* In ICASSP 2012.
[14] KRIZHEVSKY, A., 2009. *Learning multiple layers of features from tiny images.* Technical report, U. Toronto, 2009.
[15] DEAN, J., Ghemawat, S., 2008. *Map-Reduce: simplified data processing on large clusters.* CACM, 2008.
[16] LOW, Y., GONZALEZ, J., KYROLA, A., BICKSON, D., GUESTRIN, C., HELLERSTEIN, J., 2012. *Distributed GraphLab: A framework for machine learning in the cloud.* In VLDB, 2012.

Advancement in Fault Detection and Estimation Methods in Fault Tolerant Control Systems

[1]*Pavol LIŠČINSKÝ (3rd year),*
Supervisor: [2]*Anna FILASOVÁ*

[1,2]Dept. of Cybernetics and Artificial Intelligence, FEI TU of Košice, Slovak Republic

[1]pavol.liscinsky@tuke.sk, [2]anna.filasova@tuke.sk

***Abstract*—In this paper, an enhanced approach to fault estimation is described. Its modification allows simple design of adaptive fault observer in case of singularity of standard linear matrix inequalities. Although estimation of actuator faults is part of the fault detection and isolation, its results give means for compensation of a influence of the fault in the system. Further, an design procedure for dynamic output control for non-linear systems is outlined.**

***Keywords*—fault estimation, fault detection, output control, inverse LQ problem**

I. Introduction

Active fault tolerant control systems (AFTCS) provide a possibility to maintain a function of the system that is under influence of a fault. Common sense approach uses hardware redundancy, higher frequency of maintenance activities, more robust system components. All these paths to obtain better performance of the controlled system will be hardly ever replaced by AFTCS, however combination of both approaches offers significant advantages. AFTCS use analytical redundancy of the system, that is dependent upon hardware redundancy, although in some cases not specifically. Use of analytical redundancy still demands existence of a component or set of components able to perform at least remedial actions in the system. Interesting advantages mentioned in [1] could be achieved if prerequisites of AFTCS would be taken into consideration during the system design phase.

Traditionally, processes of AFTCS can be divided in two separate parts. First one is fault detection and identification (FDI) and deals with detection of influence of the fault occurred in the system, then localization of the faulty component. In the end, FDI should be able to provide useful information about the state of the system, that is usable in decision process, whether it is viable to maintain the system operation or what remedial actions should be taken. Second part of AFTCS processes provides mentioned remedial action, that will allow the system to continue in operation or to modify operational goals. This can be achieved using available resources, both analytical and hardware redundancy, and selecting pre-computed control law or synthesizing a new one in real-time. It worth to note that some authors in this field use slightly different nomenclature, however that is irrelevant, given the content is more or less same.

Place for application of AFTCS is apparent in safety-critical, cost-critical or volume-critical systems [2]. In previous years AFTCS principles were largely applied in wind turbines [3].

II. Initial status

Parity space principles, observer-based methods and parameter identification techniques applied in fault detection and identification were studied extensively in the past. Different approach for fault detection has been applied using LQ control properties [4], where residual function is based on evaluation of the LQ performance index. Application of the scheme is described in [5]. The approach allows use of the scheme for fault detection in non-linear systems described by Takagi-Sugeno (TS) models. Useful property of the scheme is ability to use arbitrary state space control instead of LQ control, that could be expected because of the evaluation of the LQ performance index. Limitation of LQ control is avoided solving inverse LQ problem, where parametric matrices $\boldsymbol{Q}$, $\boldsymbol{S}$, and $\boldsymbol{R}$ are not given, but computed so the resulting LQ control is equal to given arbitrary state space control represented by control law $\boldsymbol{K}$.

The scheme provides simple algorithm with high sensitivity to additive sensor and actuator faults, that are detected almost immediately after its occurrence. Application of the mentioned scheme for fault detection based on evaluation of the LQ performance index for case of non-linear systems led during previous year to development of dynamic output control for such systems [6].

III. Description of solved tasks

A. Dynamic output control of non-linear systems

In practice, it is usually impossible to measure all states of the system and only the observable outputs are available for static or dynamic control. Described method of control law synthesis changes synthesis into a linear matrix inequality (LMI) optimization problem. Using bounds on the system properties allows to change possibly non-convex conditions into solvable convex problem.

A non-linear dynamic system described by TS model with dynamic output controller takes form

$$\dot{\boldsymbol{q}}^{\circ}(t) = \sum_{i=1}^{s}\sum_{j=1}^{s} \boldsymbol{h}_i(\boldsymbol{\theta}(t))\boldsymbol{h}_j(\boldsymbol{\theta}(t))\boldsymbol{A}^{\circ}_{cij}\boldsymbol{q}^{\circ}(t), \tag{1}$$

$$\boldsymbol{y} \circ (t) = \boldsymbol{I}^{\circ}\boldsymbol{C}^{\circ}\boldsymbol{q}^{\circ}(t), \tag{2}$$

where

$$\dot{\boldsymbol{q}}^{\circ T}(t) = \begin{bmatrix}\boldsymbol{q}^T(t) & \boldsymbol{p}^T(t)\end{bmatrix} \tag{3}$$

$$\boldsymbol{A}^{\circ}_{cij} = \begin{bmatrix}\boldsymbol{A}_i + \boldsymbol{B}_i\boldsymbol{N}_i\boldsymbol{C} & \boldsymbol{N}_i\boldsymbol{M}_j \\ \boldsymbol{L}_j\boldsymbol{C} & \boldsymbol{N}_j\end{bmatrix}, \tag{4}$$

$$\boldsymbol{I}^{\circ} = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{I}_m \end{bmatrix}, \quad \boldsymbol{C}^{\circ} = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{I}_n \\ \boldsymbol{C} & \boldsymbol{0} \end{bmatrix} \tag{5}$$

$\boldsymbol{q}(t) \in \mathbb{R}^n$, $\boldsymbol{u}(t) \in \mathbb{R}^r$, $\boldsymbol{y}(t) \in \mathbb{R}^m$ are vectors of the state, input and output variables and $\boldsymbol{A}_i(t) \in \mathbb{R}^{n\times n}$, $\boldsymbol{B}_i(t) \in \mathbb{R}^{n\times r}$, $\boldsymbol{C}(t) \in \mathbb{R}^{m\times n}$ are constant matrices, $\boldsymbol{h}_i(\boldsymbol{\theta}), i = 1, 2, \ldots, s$ is set of normal membership functions, $\boldsymbol{\theta}$ is the vector of the premise variables.

The task of the problem was to develop design scheme for dynamic output controller

$$\dot{\boldsymbol{p}}(t) = \sum_{j=1}^{s} \boldsymbol{h}_j(\boldsymbol{\theta}(t))(\boldsymbol{J}_j p + \boldsymbol{L}_j \boldsymbol{y}(t)), \tag{6}$$

$$\boldsymbol{u}(t) = \sum_{j=1}^{s} \boldsymbol{h}_j(\boldsymbol{\theta}(t))(\boldsymbol{M}_j \boldsymbol{p} + \boldsymbol{N}_j \boldsymbol{y}(t)), \tag{7}$$

$\boldsymbol{p}(t) \in \mathbb{R}^n$ is the vector of controller state variables, the set of parameter matrices $\boldsymbol{K}_j^{\circ} \in \mathbb{R}^{(n+r)\times(n+m)}$

$$\boldsymbol{K}_j^{\circ} = \begin{bmatrix} \boldsymbol{J}_j & \boldsymbol{L}_j \\ \boldsymbol{M}_j & \boldsymbol{N}_j \end{bmatrix} \tag{8}$$

is describing the desired control law able to stabilize the given system, if there exists a positive definite matrix $\boldsymbol{Q}^{\circ} \in \mathbb{R}^{2n\times 2n}$, symmetric matrices $\boldsymbol{X}_{ij} \in \mathbb{R}^{2n\times 2n}$, a regular matrix $\boldsymbol{H}^{\circ} \in \mathbb{R}^{(n+m)\times(n+m)}$ and matrices $\boldsymbol{Y}_j^{\circ} \in \mathbb{R}^{(n+r)\times(n+m)}$, $\boldsymbol{X}_{ij}^{\circ} \in \mathbb{R}^{2n\times 2n}$, such that following conditions set in form of LMI and linear matrix equalities (LME)

$$\boldsymbol{X}_{ij}^{\circ} = \boldsymbol{X}_{ji}^{\circ}, \quad \begin{bmatrix} \boldsymbol{X}_{11}^{\circ} & \boldsymbol{X}_{12}^{\circ} & \cdots & \boldsymbol{X}_{1s}^{\circ} \\ \boldsymbol{X}_{21}^{\circ} & \boldsymbol{X}_{22}^{\circ} & \cdots & \boldsymbol{X}_{2s}^{\circ} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{X}_{s1}^{\circ} & \boldsymbol{X}_{s2}^{\circ} & \cdots & \boldsymbol{X}_{ss}^{\circ} \end{bmatrix} > 0, \tag{9}$$

$$\boldsymbol{C}^{\circ}\boldsymbol{Q}^{\circ} = \boldsymbol{H}^{\circ}\boldsymbol{C}^{\circ}, \quad \boldsymbol{Q}^{\circ} = \boldsymbol{Q}^{\circ T} > 0, \tag{10}$$

$$\boldsymbol{A}_i^{\circ}\boldsymbol{Q}^{\circ} + \boldsymbol{Q}^{\circ}\boldsymbol{A}_i^{\circ} + \boldsymbol{B}_i^{\circ}\boldsymbol{Y}_j^{\circ}\boldsymbol{C}^{\circ} + \boldsymbol{C}^{\circ T}\boldsymbol{Y}_j^{\circ T}\boldsymbol{B}_i^{\circ T} + \boldsymbol{X}_{ij}^{\circ T} \tag{11}$$

hold for $\boldsymbol{h}_i(\boldsymbol{\theta})\boldsymbol{h}_j(\boldsymbol{\theta}) \neq 0, i, j, \ldots, s$. Then the control law gain matrices are given as

$$\boldsymbol{K}_j^{\circ} = \boldsymbol{Y}_j^{\circ}(\boldsymbol{H}^{\circ})^{-1}, \quad j = 1, 2, \ldots, s. \tag{12}$$

Proves of above mentioned statements and further description of the scheme can be found in [6].

B. Fault estimation based on adaptive observers

An interesting approach to actuator fault estimation for linear continuous time systems has been presented in [7] and provides useful and easily implementable tool in process of FDI. Modified design conditions were presented in [8]. Additive actuator fault signals estimated by an adaptive fault observer can be potentially used in combination with the control law to compensate the influence of the fault. In both cases, conditions for the stability of the error dynamics were given in terms of LMIs. Considered are single actuator faults in the multi-input linear systems with unknown time of occurrence .

Adaptive fault observer is stable if for positive $\delta \in \mathbb{R}$, there exist matrices $\boldsymbol{P}$, $\boldsymbol{Q} \in \mathbb{R}^{n\times n}$, $\boldsymbol{H} \in \mathbb{R}^{p\times m}$, $\boldsymbol{Y} \in \mathbb{R}^{n\times m}$ and a positive scalar $\gamma \in \mathbb{R}$ such that following conditions hold

$$\boldsymbol{P} = \boldsymbol{P}^T > 0, \quad \boldsymbol{Q} = \boldsymbol{Q}^T > 0, \quad \gamma > 0, \tag{13}$$

$$\begin{bmatrix} \boldsymbol{QA}+\boldsymbol{A}^T\boldsymbol{Q}-\boldsymbol{YC}-\boldsymbol{C}^T\boldsymbol{Y}^T+\boldsymbol{C}^T\boldsymbol{C} & * & * \\ \boldsymbol{P}-\boldsymbol{Q}+\delta\boldsymbol{QA}-\delta\boldsymbol{YC} & -2\delta\boldsymbol{Q} & * \\ \boldsymbol{F}^T\boldsymbol{Q} & \boldsymbol{0} & -\gamma\boldsymbol{I}_p \end{bmatrix} < 0, \tag{14}$$

$$\boldsymbol{C}^T\boldsymbol{H} = \delta\boldsymbol{QF}. \tag{15}$$

Then observer matirx $\boldsymbol{J}$ is given by the relation

$$\boldsymbol{J} = \boldsymbol{Q}^{-1}\boldsymbol{Y}. \tag{16}$$

Advantage of the structure (14) in comparison to the similar structure presented in [7] is in possibility to find a feasible solution in case that the fault input matrix $\boldsymbol{F}$ would cause the unmodified structure to become singular, thus the scheme would not provide a usable solution.

Further description of the schemes together with system model, fault description, proofs and illustrative examples can be found in [8] and [7].

IV. Next steps

Set of strict conditions required for control law in some cases can limit application of previously proposed fault detection schemes. Better understanding or development of more suitable control schemes will hopefully provide means necessary for application of AFTCS with dynamic controllers.

Fault estimation based on adaptive observers provide useful information about the fault in the system. This allows to possibly compensate the influence of the fault in the system or can be used in different control reconfiguration schemes. Up until this point, the fault has been always directly compensated by simple subtraction of the estimated fault from the control input. Interesting results could be obtained from combination of given fault estimation and control reconfiguration schemes usable in case when a simple compensation of fault is not possible.

Further steps in the research will lead to application of fault estimation for decentralised system and to implementation of reconfigured control using the already obtained results.

Acknowledgment

The work presented in the paper was supported by VEGA, the Grant Agency of the Ministry of Esucation and the Academy of Science of Slovak Republic, under Grant No. 1/0348/14. This support is very gratefully acknowledged.

References

[1] M. Blanke and J. Schröder, *Diagnosis and fault-tolerant control*. Springer, 2006, vol. 691.

[2] M. Mahmoud, J. Jiang, and Y. Zhang, *Active fault tolerant control systems: stochastic analysis and synthesis*. Springer Science & Business Media, 2003, vol. 287.

[3] N. Luo, Y. Vidal, and L. Acho, *Wind Turbine Control and Monitoring*. Springer, 2014.

[4] D. Krokavec and A. Filasová, "Fault detection based on linear quadratic control performances," in *Diagnostics of Processes and Systems DSP 2011 : 10th International Science and Technology Conference*, Zamość, Poland, 2011, pp. 52–56.

[5] V. Serbak and P. Liscinsky, "Fault detection based on linear quadratic control performance in nonlinear systems," in *Control Conference (ICCC), 2014 15th International Carpathian*, May 2014, pp. 517–521.

[6] D. Krokavec, A. Filasova, and P. Liscinsky, "Dynamic output control of nonlinear systems described by takagi-sugeno models," in *Control Applications (CCA), 2014 IEEE Conference on*, Oct 2014, pp. 959–964.

[7] D. Krokavec, A. Filasová, V. Serbák, and P. Liščinskỳ, "An enhanced approach to actuator fault estimation design for linear continuous-time systems," in *Journal of Physics: Conference Series*, vol. 570, no. 7. IOP Publishing, 2014, p. 072002.

[8] V. Serbák and P. Liščinský, "Additive faults estimation based on adaptive observers for linear systems," 2015, paper submitted for a conference ICCC2015.

An Overview of the Advanced Language Modeling for the Slovak Continuous Speech Recognition

[1]*Daniel ZLACKÝ (3rd year)*
Supervisor: [2]Anton ČIŽMÁR

[1,2]Dept. of Electronics and Multimedia Communications, FEI TU of Košice, Slovak Republic

[1]daniel.zlacky@tuke.sk, [2]anton.cizmar@tuke.sk

***Abstract*—The quality of stochastic language models in large vocabulary continuous speech recognition mostly depend on an sufficient amount of the text data and on the methods and techniques that are used for estimation of these language models. This paper summarizes the work of student for the previous year of the post gradual study aiming on comparison of several advanced language modeling techniques for the Slovak continuous speech recognition that were published in relevant journals in recent years. Maximum entropy, power law discounting, hierarchical Pitman-Yor process, variable-order Kneser-Ney smoothed models and recurrent neural networks language models were compared according to their model size, perplexity, speech recognition performance and overall complexity of their usage in the real conditions of Slovak speech recognition system.**

***Keywords*—continuous speech recognition, language model, model perplexity, word error rate.**

I. Introduction

In general, stochastic language models for large vocabulary continuous speech recognition (LVCSR) need to accurately estimate the probability P(W) for a given word sequence $W = \{w_1, w_2, \ldots, w_n\}$, called history ($h$). This probability can be defined as follows:

$$P(W) = \prod_{i=1}^{N} P(w|h) = \prod_{i=1}^{N} P(w_i|w_1 w_2 \ldots w_{i-1}) \qquad (1)$$

where $P(w_i|w_1, w_2, \ldots, w_{i-1})$ is a conditional probability and a word w_i is given by the sequence of words $(w_1, w_2, \ldots, w_{i-1})$. Real systems are based on the several previous words, which leads to an *n*-gram language models. Language models are usually evaluated by two standard measures: perplexity (PPL) and word error rate (WER).

This paper is focused on evaluation of language models that were published in the last year in real conditions using Slovak automatic dictation and transcription system [1].

This paper is organized as follows. Section II briefly overview several advanced modeling technique. In the Section III Slovak LVSCR system is mentioned. Experimental results are summarized at the final section.

II. Advanced Language Modeling Techniques

The next subsections briefly introduce several advanced statistical technique that have never been tested in the Slovak LVCSR system. The detailed description can be found in accompanying literature.

A. Maximum Entropy Language Model

A maximum entropy language modeling (MaxEnt) toolkit, as an extension to SRILM toolkit, with using the hierarchical feature technique and scalable optimization algorithm that can handle large amount of training data, was first introduced in [2]. The results showed that the MaxEnt models are comparable in terms of PPL and WER to the convenient *n*-gram language models.

B. Hierarchical Pitman-Yor Process Language Model

A hierarchical Bayesian language model based on Pitman-Yor process (HPYLM) for meeting speech recognition, where the language models can be created from large amount of training data using the advantages or parallel computation was firstly mentioned in [3-4], where PPL and WER performed better results than the *n*-gram language models.

C. Power Law Discounting Language Model

A power law discounting language model (PLDLM) uses the power law distribution over word tokens that removes the main disadvantages of time expensive computation process in HPYLM. In [5], PLDLM obtains comparative results to HPYLM and convenient *n*-gram language model in two meeting speech transcription tasks.

D. Variable-Order Kneser-Ney Smoothed Language Model

In [6], authors used the Morfessor unsupervised segmentation tool and the state-of-the-art Kneser-Ney smoothing algorithm to create variable-order Kneser-Ney smoothed language models (VariKN) that are suitable for modeling the under-resourced languages, which rapidly decreased WER results in their LVCSR system [6].

E. Recurrent Neural Network Language Model

Neural network language model have been continuously reported to perform well amongst other language modeling technique. The best results were mostly obtained by recurrent neural network based language model (RNNLM). In [7], a new RNNLM is used in speech recognition task. The experimental results showed that RNNLM are superior to standard *n*-gram techniques, except their computational complexity.

TABLE I
THE OVERALL RESULTS OF THE SLOVAK LVCSR SYSTEM

LM technique	Size [GB]	Number of n-grams		PPL_{T1}		WER_{T1} [%]		PPL_{T2}		WER_{T2} [%]	
		2g	3g	2g	3g	2g	3g	2g	3g	2g	3g
WBSLM	5.61	81,768,292	100,222,797	106.42	60.63	**7.62**	**6.55**	458.75	309.91	**13.05**	**11.68**
MaxEnt	13.50	81,273,825	328,933,339	108.86	58.62	7.89	6.89	490.12	303.46	13.68	12.79
HPYLM	5.62	81,768,398	100,222,888	110.72	61.66	7.95	7.41	474.61	291.24	13.53	12.60
PLDLM	14.23	81,768,292	352,087,702	109.54	59.15	8.13	7.11	468.80	289.75	14.27	13.14
VariKN	14.20	81,768,310	351,167,374	112.24	57.47	9.59	8.32	505.30	289.70	17.35	16.15
$RNNLM_s$	12.53	383,474,264	55,539,632	138,71	98.50	8,36	7.74	637,31	514.52	14,66	14.80
$RNNLM_d$	10.33	305,849,388	57,065,940	139,20	98.48	8,37	7.80	635,16	512.44	14,60	13.97
WBS+ME	13.73	81,768,291	336,315,612	**105.07**	57.21	7.64	6.64	458.75	286.41	13.19	12.08
WBS+HPYLM	5.62	81,768,397	100,223,121	105.75	58.10	7.68	6.73	455.97	281.86	13.22	11.93
WBS+PLDLM	14.23	81,768,292	352,087,702	105.40	55.88	7.84	6.81	**449.77**	275.47	13.61	12.46
WBS+VariKN	14.21	81,768,310	351,460,657	106.31	**54.09**	9.90	7.97	465.65	**270.20**	17.54	15.75
WBS + $RNNLM_s$	16.70	431,743,528	137,763,540	106.07	59.89	7.63	6.59	461,60	305.78	13,32	12.07
WBS + $RNNLM_d$	14.47	354,357,102	138,435,360	106.07	59.88	7.63	6.62	459,02	303.80	13,26	12.01

III. LARGE VOCABULARY CONTINUOUS SPEECH RECOGNITION SYSTEM SETUP

The created language models that will be used in our mentioned experiments in the next chapter, were tested by LVCSR system that consist from 4 state triphone acoustic model based on hidden Markov models using HTK toolkit. Julius decoder based on two-pass strategy was also used. Language models have been created using the SRILM, MaxEnt, HPYLM, PLDLM, VariKN and RNNLM toolkits. The partial language models were trained on 24 independent domain-oriented models (2.1 billion tokens in 120 million sentences) and merged together using linear interpolation. The dictionary consist from 325 555 unique words. The first test dataset (T1) consist from 3 426 sentences and 41 820 words and the second test dataset (T2) contains 132 156 words in 12 163 sentences that were not part of the training data.

RNNLM cannot be directly used in our LVCSR system because they are not finite states machines and cannot be saved in appropriate *.arpa* format. We used them as a generative model to generate a new 24 text data corpora. $RNNLM_s$ have generated around 100 million tokens for each corpora. $RNNLM_d$ have generated around the same amount of token as each of the 24 corpora.

IV. EXPERIMENTAL RESULTS

As we can see in the Table I., we have trained 13 different bigram and trigram language models. The baseline language model smoothed by the Witten-Bell back-off algorithm (WBSLM), which is successfully used in our LVCSR system and other five advanced models are mentioned in the first part of the table. The second part illustrates experimental results of interpolated models with WBSLM using linear interpolation.

The results showed that the number of bigrams are almost the same in every case except RNNLM, where was generated a huge amount of new bigrams in contrast to the number of trigrams, where was generated the smallest amount.

The PPL of bigrams models were only decreased when the advanced models were interpolated with WBSLM. Otherwise the PPL of trigram models were decreased almost in every case in both test data sets. The lowest PPL were obtained by the interpolation of WBSLM and VariKN trigram models.

The best recognition results we observed with the convenient WBSLM language models in both test sets. Bigram and trigram models created with the advanced techniques achieved worse recognition results in word error rates as the baseline language models in both performance evaluation scenarios. Although interpolated models achieved at least better recognition results than the convenient language models, but still worse than the baseline model. The worst results were observed if the variable-order Kneser-Ney smoothed language models were used.

We proved that the convenient *n*-gram language models are still the optimal choice for our large vocabulary continuous speech recognition system even if the MaxEnt, HPYLM or RNNLM can achieve comparable results, but their memory requirements due to the increasing number of *n*-grams are not very suitable in case of real-time continuous speech recognition and automatic transcription. The detailed comparison of these modeling technique can be found in [8].

ACKNOWLEDGMENT

The research presented in this paper was supported by the Ministry of Education, Science, Research and Sport of the Slovak Republic under the project VEGA 1/0075/15 (50%) and the Research and Development Operational Program funded by the ERDF under the projects ITMS-26220220182 (50%).

REFERENCES

[1] M. Rusko [et al.], "Slovak automatic transcription and dictation system for judicial domain," In: *Human Language Technologies as a Challenge for Computer Science and Linguistics: 5th Language & Technology Conference*, Poznan, Poland, 2013, pp. 365-369.

[2] T. Alumäe, M. Kurimo, "Efficient estimation of maximum entropy language models with n-gram features: An SRILM extension," In: *Proc. of INTERSPEECH'10*, Chiba, Japan, 2010, pp. 1820-1823.

[3] S. Huang, S. Renals, "Hierarchical Pitman-Yor language models for ASR in meetings," In: *Proc. of IEEE Workshop on Automatic Speech Recognition & Understanding*, ASRU'07, Kyoto, Japan, 2007, pp. 124-129.

[4] S. Huang, S. Renals, "A parallel training algorithm for hierarchical Pitman-Yor process language models," In: *Proc. of INTERSPEECH'09*, Brighton, UK, 2009, pp. 2695-2698.

[5] S. Huang, S. Renals, "Power law discounting for n-gram language models," In: *Proc. of IEEE Int. Conf. on Acoustic, Speech, and Signal Processing*, ICASSP'10, Dallas, Texas, USA, 2010, pp. 5178-5181.

[6] V. Siivola, M. Creutz, M. Kurimo, "Morfessor and VariKN machine learning tools for speech and language technology," In: *Proc. Of INTERSPEECH'07*, Antwerp, Belgium, 2007, pp. 1549-1552.

[7] T. Mikolov, M. Karafiat, L. Burget, J. Cernocky, S. Khundanpur, "Recurrent neural network based language model," In: *Proc. of INTERSPEECH'10*, Chiba, Japan, 2010.

[8] D. Zlacky, J. Stas, J. Juhar, A. Cizmar, "Evaluation of advanced language modeling techniques for the Slovak LVCSR," In: Proc. of 25 International Conference Radioelektronika, Pardubice, Czech republic, 2015. (to be published)

Architectural Prototype for an Autonomic Software System – Requirements

[1]*Veronika SZABÓOVÁ (3st year)*
Supervisor: [2]Zdeněk HAVLICE

[1,2]Dept. of Computers and Informatics, FEI TU of Košice, Slovak Republic

[1]veronika.szaboova@tuke.sk, [2]zdenek.havlice@tuke.sk

***Abstract* — This paper deals with the requirements model for an autonomous computing system. Requirements, why? The inputs for logical architecture design are the domain model, system context, dictionary and the requirements model, which one is the basis of architecture design. We focus this paper on the autonomous manager and present the steps of its logical architecture design.**

***Keywords*— autonomic system, requirement, architecture**

I. Introduction

Autonomous systems are systems that are capable of self-care. Within the IT sector, usage is oriented on automation of monotonous human work consisting of small and easy steps[1][2]. The most important advantage is continuous operation without system shutdown or reboot due to maintenance. Note that this does not mean elimination of human interaction. We say a system is autonomous in the case it fulfills the requirements stated in [3][8]:

- Self-monitoring of internal states.
- Self-awareness via internally stored aims and goals.
- Self-healig in emergency situations, without effecting running operations.
- Configurable by users, who can add and/or change the stored aims and goals of the system. This interface should provide monitoring reports as well.

The basic model includes the four attributes as ability to self-configuration, self-healing, self-evolving and self-protection[1][4][3], which is included with the autoreflexive software architecture (ARSA) [6][10].

The aim of this paper is to present requirements necessary for creation of an architectural prototype based on the methodology described in [11].

II. Related work

The concept of autonomous element is standalone software that implements all four attributes of an autonomous system as defined in the previous section. Each element contains sensors with which scans provided the resources and the tools for its subsequent modification if necessary. It can evaluate the situation and adjust their behavior based on external as well as internal factors[5][7]. The behavior of the element therefore depends on what is perceived or not perceived at that time, or in the past and is tailored to meet all the requirements of their knowledge. Although it was regarded as an independent, it is still restricted to enter the destination and any change must be made to inform the administrator who has the authority to possible external interference [5]. One element can itself constitute the entire autonomous system, but it is usually only a part of a larger autonomous system, which consists of a number of autonomic elements but also other elements. These elements have independent properties, need not be dynamically adapted and are independent of the potential of the situation.

The structure of autonomous elements is built on the principles of architecture model introduced by IBM in [5]:

- Managed resource
- Effectors
- Sensors
- Autononic manager

Managed resource is a component that belongs either to the program or computer equipment. This may include a server, database, storage unit or other entity. Each resource provides an interface to autonomous manager called touch points. There are two types of touch points: sensors and effectors. The sensors are used to monitor the controlled resources. Continuously acquire information provided by autonomous manager to analyze and plan the necessary changes[9]. Effectors then implement these changes in the managed resource. Touch points allow encapsulation of resources making it possible to be self-accessed with the help of their own interfaces.

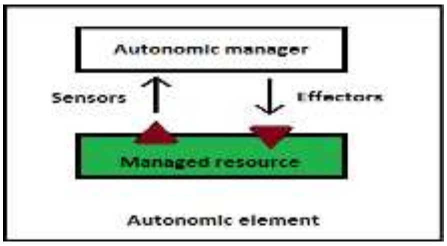


Figure 1 Autonomic element structure.

III. Requirements and logical architecture design

The architecture of the autonomous system will be designed as a centralized, autonomous and scalable system. In case of extension of terminal equipment system, resp. adding new components does not modify the original architecture solutions.

The main features of the proposed architecture are mainly:

- Flexible architecture of the autonomous system will be oriented onto web services
- Clearly defined roles and levels of the different components of the autonomic system
- Separation of components on the basis of clearly defined functionality and interfaces
- Selection of all values obtained in the process of autonomous system for the purpose of evaluation and follow-on system configuration
- Automatic error diagnostics, evaluation and adequate correction of errors by taking measures to prevent their occurrence
- Sync multiple processes to ensure optimization and adaptation to the new conditions on the fly, without affecting ongoing operations

A. Requirements for autonomous element of the autonomous system

The autonomous element will be designed with given a predefined targets and the level of knowledge. Tracking the source control will be realized by means of sensors that will continuously collect information for the purpose of providing autonomic manager on the analysis and implementation of appropriate changes made through effectors on a proposal autonomic manager.

The sensors are designed to identify those parameters source control that are essential for efficient operation of the exclusion of unnecessary data that would slow down the entire system and not by the progressive improvement of autonomy.

Implementation of effectors will be synchronized together with sensors to the extent that there is no improper and inadequate resources proposed change in control of autonomous managers.

System administration will be ensured through autonomous manager-based adaptation changes obtained by systematic monitoring data in combination with the overall knowledge of the system.

B. Requirements for the management of autonomous system

Autonomous management system will be ensured through intelligent control loop, which will play a key role not only in collecting information from the system, but also in creating a decision and, if necessary, the adjustment of the system.

Architecture autonomous system will be designed in such a way as to ensure mutual interactivity of the components of intelligent control loop and autonomic manager, points of contact and resource control.

Autonomic manager with the assistance of sensors will be capable of monitoring resource control. After subsequent analysis of the information obtained and comparing them with the overall knowledge about the system, if necessary, change plans, the execution to take place through resource control effectors.

IV. Conclusion

Simpler and probably more effective option for obtaining regular primary messages will be direct conversion of information in managed resource. This functionality can be provided by using the tool "The Common Base Event API" intended for the Eclipse development environment [1].

For creating blank common base event interface is used "ICommonBaseEvent interface" tool provided by "The Common Base Event API".

To define the source component that has the problem it is necessary to use "IComonentIdentificator" interface.

Completion report of the problem, and the importance of implementation time follows based on the data obtained via monitoring.

After creating and filling out reports is necessary to ensure communication with separate managers. The source is assigned after initialization control interface, thus touch points through which the assignment of a particular manager implements.

Proposed communication via the contact points and the controlled device is shown in Figure 2.

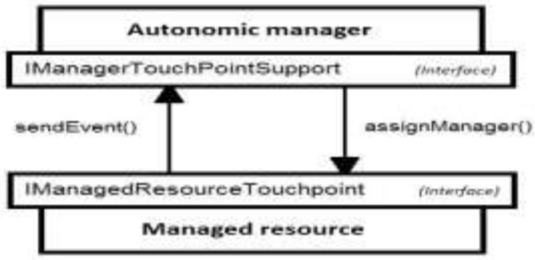


Figure 2 Proposed messaging between manager and resource

References

[1] Bart Jacob, Richard Lanyon-Hogg, Devaprasad K Nadgir, Amr F Yassin. A Practical Guide to the IBM Autonomic Computing Toolkit [online]. An IBM Redbooks publication. ISBN-13 9780738498058

[2] Vaughn Rokosz. Lotus Developer Domain: Autonomic Computing: It's about making smarter systems. In: web5.notes.net, 2 June 2003 [online]

[3] Manish Parashar. Autonomic Computing: Concepts, Infrastructure, and Applications. CRC Press (December 15, 2006), 568 pages, ISBN-13: 978-0880117487

[4] Russell, D.M.; Maglio, P.P.; Dordick, R.; Neti, C., "Dealing with ghosts: Managing the user experience of autonomic computing," IBM Systems Journal , vol.42, no.1, pp.177,188, 2003 doi: 10.1147/sj.421.0177

[5] Philippe Lalanda, Julie A. McCann, Ada Diaconescu. Autonomic Computing Principles, Design and Implementation. Springer; 2013 edition (May 28, 2013), 288 pages, ISBN-13: 978-1447150060

[6] Z. Havlice. Auto-Reflexive Software Architecture with Layer of Knowledge Based on UML Models. In: International Review on Computers and Software (IRECOS). Vol. 8, no. 8 (2013), p. 1814-1821. - ISSN 1828-6003

[7] Veronika Szabóová, Csaba Szabó. Quo vadis, self-* software? In: 9th International Conference on Applied Informatics : Eger, January 29-February 1, 2014. - [Eger : Eszterházy Károly College], Vol.2. 2015 P.317-325. doi:10.14794/ICAI.9.2014.2.317

[8] Veronika Szabóová. Hypotheses on a self-healing system architecture. In: SCYR 2014 : 14th Scientific Conference of Young Researchers : proceedings from conference : May 20th, 2014, Herľany, Slovakia. - Košice : TU, 2014 S. 191-192. - ISBN 978-80-553-1714-4

[9] Csaba Szabó, Veronika Szabóová. On Representation and Usage of Requirements in Self-* Systems. In: 10th Joint Conference on Mathematics and Computer Science : Cluj-Napoca, May 21-25, 2014. - [Cluj-Napoca : Babeş-Bolyai University], 2014 P. 1. Available online: http://www.cs.ubbcluj.ro/~macs/2014/abstracts.php

[10] Dominika Hurajtová, Zdeněk Havlice, Veronika Szabóová. Knowledge-Layer Integration into Information System. In: Computer Science and Technology Research Survey : Volume 7. - Košice : TU, 2014 S. 62-67. - ISBN 978-80-553-1857-8

[11] EELES, P. – CRIPPS, P.: Architektura softwaru, Computer Press, Brno, 2011. – ISBN978-80-251-3036-0. 328 s.

Augmented reality system for multiuser

[1]Martin Varga (4st year)
Supervisor: [2]Branislav SOBOTA

[1,2]Dept. of Computers and Informatics, FEI TU of Košice, Slovak Republic

[1]martin.varga@tuke.sk, [2]branislav.sobota@tuke.sk

***Abstract*—This paper presents the problem solving of augmented reality (AR) system for more users. Paper is divided into three parts. The first part describes introduction of augmented reality. Second part describes summarization of the research task in AR. The last part presents concept of a solving the problem for more users using head-mounted displays for visualization of virtual scene in the real environment.**

***Keywords*—Augmented reality, mixed reality, modularization, visualization, communication protocol**

I. Introduction

Augmented reality (AR) is computer-generated data integration with the real world, which among others can be done with computer graphics rendering on a real-time footage. AR can be used for many things, such as displaying a mobile directions to head-up display, in the medical field, the AR may help doctors to insert information on a patient's medical record (such as x-ray result from the patients), or to reconstruct the historic buildings as reality which can be seen at real time [4], [6], [10], [11]. AR in the architecture merges virtual designs with real construction sites, and enables new types of interactions that enhance the design process [12]. Behzadan et al. developed a hardware and software framework for visualization of construction processes (e.g., machinery placement) for construction sites [13].

II. Dividing and Summarization of AR

According to the method how virtual objects are aligned with real scene image there are two systems use:

- Marker systems – special markers are used in a real scene (article [7]).
- Markerless systems – additional information for real environment, for example image (e.g. photo (semi-markerless system (article [3]))), face recognition, GPS data, inertial and electromagnetic tracking devices (articles [1], [4], [5]), etc.

Based on how a user sees mixed reality there can be two types of systems:

- Optical see-through systems where the user sees real world directly and computer generated objects are added to this view. This category of systems usually works with semi-transparent displays (articles [2], [6], [8], [9]).
- Video see-through where captured real world image with added virtual objects is displayed to the user. This is usually realized via camera – display system [3].

The goal is to create augmented reality system for more users using head-mounted displays for visualization of virtual scene in the real environment. The main significance is in interaction with one virtual world.

III. Solving the Problem for More Users

The basic concept (Fig.1) of an augmented reality system in our proposal consists of following most important module functions:

- Observer module – for direct interaction between system and users
- Storage module – to store current scene data
- Dynamic module – out dynamic case (e.g. scripted movement of objects in scene)
- Synchronization module - for synchronization of other modules
- Communication module - is a form of server application, which implements communication protocol,

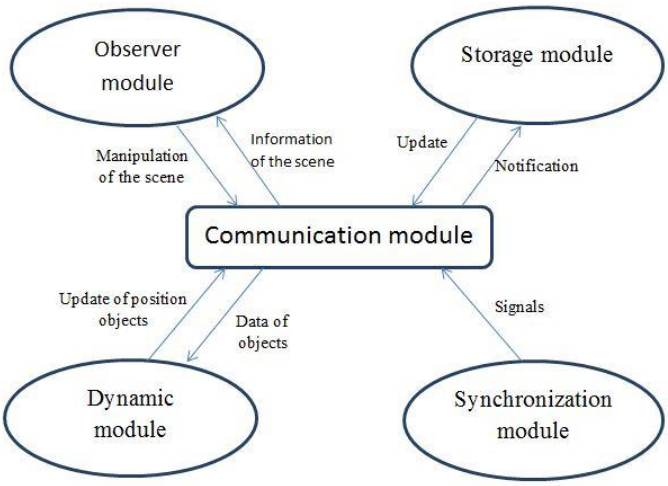


Fig. 1. The basic concept of modules for an augmented reality system.

Users directly interact with AR using observer module. This module is divided on input and output modules. Main task of output module is to visualize data with use of data helmet or monitor and in case of mixed reality also to combine virtual and real worlds. Input modules provide entry data from sensors to correct point of observer view (it receives tracking data of position and orientation of observer) and entry data from data glove. Output module (for visualization) can use more users and each user can see a virtual scene in the real environment.

Figure 2 shows diagram of relations between modules in terms of specific data exchanged. Communication module is not displayed, but all modules used this module to communicate with system. As the figure shows, central module of system is storage module, which provides scene data to other modules and they modify data in storage.

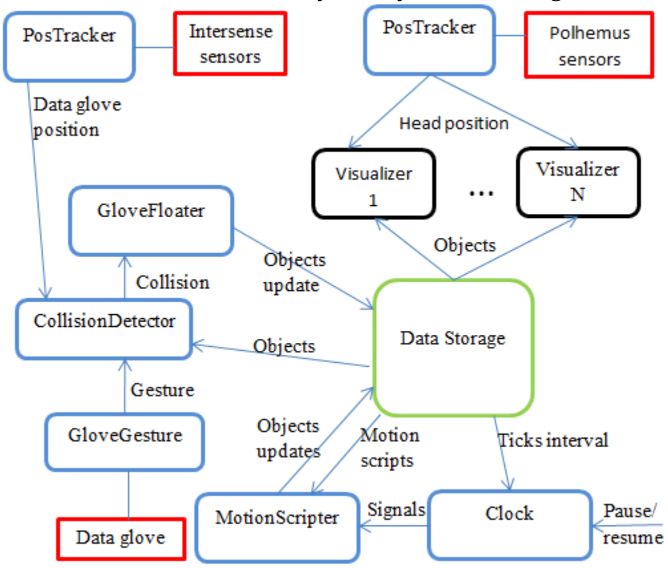


Fig. 2. Relations between system modules with a specific data exchanged. Red components are an input hardware, green component is data storage, black components show visualizers for more users and blue components are other modules of mixed system.

Visualizer also allows calibrating position of virtual world in real world. The transformation M used to create the correct view on scene can be expressed using formula 1(x - matrix multiplication):

$$M = M_1 \text{ x } M_2 \text{ x } M_3 \qquad (1)$$

The M_1 in formula corresponds with the initial transformation of a scene (identity), M_2 is transformation caused by calibration and M_3 is application of virtual (left or right) camera's position and orientation.

The Fig.3 illustrates users with head-mounted displays which can manipulate with a virtual object using data gloves.

Fig. 3. The users with head-mounted display can manipulate with a virtual object using data gloves (illustration).

IV. Conclusion

The proposed system for more users with head-mounted displays, presented in the paper, was constructed at the DCI FEEI TU of Kosice (Department of computers and informatics, Faculty of electrical engineering and informatics, Technical university of Kosice) in the LIRKIS (Laboratory of Intelligent Interfaces of Communication and Information Systems). Advantage of this solution is that more users can see same virtual scene and they can change position of 3D objects in the virtual scene. Users can interact with the virtual scene using hand gestures. The next advantage is that augmented reality system is composed of modules, and new features can be added with new modules. Further research will focus on the creation of additional modules with new features of the system (e.g. input data from gyroscope, GPS, Kinect , implementation of animation for Bezier curve, etc.).

Acknowledgment

This work was supported by the KEGA grant no. 083TUKE-4/2015 “Virtual-reality technologies in the process of handicapped persons education”.

References

[1] VARGA, Martin: Markerless augmented reality using head-mounted display. 2012. - 1 elektronický optický disk (CD-ROM). In: POSTER 2012: 16th International Student Conference on Electrical Engineering: May 17, 2012, Prague, Czech republic. - Prague: Czech Technical University in Prague, 2012 P. 1-5. - ISBN 978-80-01-05043-9

[2] VARGA, Martin - MITRO, Lukáš: Head-mounted display. 2014. In: Developments in Virtual Reality Laboratory for Factory of the Future 2013-2014. - Miskolc: Bay Zoltán Nonprofit, 2014 P. 67-70. - ISBN 978-963-88122-4-7

[3] VARGA, Martin: Markerless augmented reality using SURF method. 2012. - 1 elektronický optický disk (CD-ROM). In: SCYR 2012: proceedings from conference : 12th Scientific Conference of Young Researchers : May 15th, 2012, Herľany, Slovakia. - Košice : TU, 2012 S. 173-176. - ISBN 978-80-553-0943-9

[4] M. Varga, Z. Dudláková, Markerless augmented reality using head tracking device. In.: POSTER 2013 17th International Student Conference on Electrical Engineering: May 16, 2013, Prague. - Prague: Czech Technical University in Prague, 2013 P. 1-5. - ISBN 978-80-01-05242-6

[5] M. Varga, P. Ivančák , Markerless augmented reality using electromagnetic tracking device, SCYR 2013, Proceedings from conference : 13th Scientific Conference of Young Researchers : May 14th, 2013, Herľany, Slovakia. - Košice : TU, 2013 S. 156-159. - ISBN 978-80-553-1422-8

[6] VARGA, Martin - SOBOTA, Branislav - HROZEK, František - KOREČKO, Štefan: Augmented Reality with Interactive Interfaces. 2014. In: Studia Universitatis Babes-Bolyai : Series Informatica. Vol. 59, no. 2 (2014), p. 21-33. - ISSN 2065-9601

[7] VARGA, Martin: Ovládanie počítačového systému handikepovanými osobami. 2014. In: Aplikácia moderných informačných technológií pri vzdelávaní hendikepovaných osôb : zborník z workshopu : Prešov, 14.11.2014. - Košice : TU, 2014 S. 30-34. - ISBN 978-80-553-1868-4

[8] HOVANEC, Michal - PAČAIOVÁ, Hana - HROZEK, František - VARGA, Martin: Proactive Ergonomics Based on Digitalization Using 3D Scanning and Workplace Modeling in Texnomatix Jack with Augmented Reality.2014. In: Our Sea, International Journal of Maritime Science and Technology. Vol. 61, no. 1-2 (2014), p. 22-26. - ISSN 0469-6255

[9] M. Hovanec, M. Varga, et al., Inovatívne trendy a vízie v ergonómii využitím rozšírenej a virtuálnej reality. In: Aktuálne otázky bezpečnosti práce : 25. medzinárodná konferencia : Štrbské Pleso - Vysoké Tatry, 06.-08. 11.2012. - Košice : TU, 2012 S. 1-7. - ISBN 978-80-553-1113-5
[10] B. SOBOTA, et al.: Riešenie úloh spracovania rozsiahlych grafických údajov v prostredí paralelných počítačových systémov. 2012, Editačné stredisko TU, Košice, ISBN 978-80-553-0864-7
[11] Y. Ariyana, A.I. Wuryabdari, "Basic 3D interaction techniques in Augmented Reality," System Engineering and Technology (ICSET), 2012 International Conference on , vol., no., pp.1,6, 11-12 Sept. 2012
[12] Wang, X. and M.A. Schnabel, Mixed Reality in Architecture, Design, and Construction. 2008: Springer-Verlag New York Inc.
[13] Behzadan, A.H. and V.R. Kamat, Enabling smooth and scalable dynamic 3d visualization of discrete-event construction simulations in outdoor augmented reality. Proceedings of the 39th conference on Winter simulation (WSC 07), 2007: p. 2168-2176.

Automatic Acoustic Data Processing for ASR Systems

[1]*Tomáš KOCTÚR (1st year),*
Supervisor: [2]*Jozef JUHÁR*

[1,2]Dept. of Electronics and Multimedia Communications, FEI TU of Košice, Slovak Republic

[1]tomas.koctur@tuke.sk, [2]jozef.juhar@tuke.sk

***Abstract*—Automatic speech recognition (ASR) quality is a huge topic and improvements in this field are necessary. Acoustic model (AM) is a base for hidden Markov model based ASR systems. It is modeled from manually transcribed speech data which are not easy to obtain. In this article unsupervised AM training will be proposed. Unsupervised system needs just a small seed AM which is iteratively retrained with new untranscribed data that are widely available. Such system concepts and theoretical basis for this problematic is presented in this article.**

***Keywords*—Automatic speech recognition, acoustic modeling, unsupervised modeling, hidden Markov models, speed signal parametrization.**

I. Introduction

Quality improvement of automatic speech recognition systems (ASR) is needed nowadays. In the core of many ASR systems is acoustic model used, but these models was made from many hours of manually transcribed speech data. Manual speech data transcription is tedious and not very cost effective. The fact that several hundreds hours of transcribed speech is necessary for only fair quality of acoustic model. Automatic self-training system for acoustic modeling is widely know as unsupervised or lightly-supervised acoustic modeling. These systems are found to be useful for under-resourced languages [15]. In the core of both systems is seed acoustic model used which was made from several hours of transcribed speech. Lightly supervised acoustic modeling use loose transcriptions which are not aligned with speech or have errors in transcribed text. Speech is being recognized by ASR with seed acoustic model and resulting hypotheses are aligned with loose transcriptions. Aligned utterances are used for retraining of the seed acoustic model. This process can be repeated until alignment of the hypotheses of loose transcribed audio data is not significant any more. This method is not cost effective because the loose transcriptions needs to be obtained.

There are many methods for unsupervised acoustic model training as described in [14]. One of the methods use two independent ASR systems with two seed acoustic models. Untranscribed speech data is being recognized by those ASR systems and their hypotheses results are compared and aligned. The probability that two independent ASR systems make the same error is very small and aligned hypothesis can be considered as correct. Aligned speech data are used for retraining of seed acoustic model and it can be done iteratively as presented by lightly supervised system. The benefit of this method is that easy obtainable speech data can be used for acoustic model training.

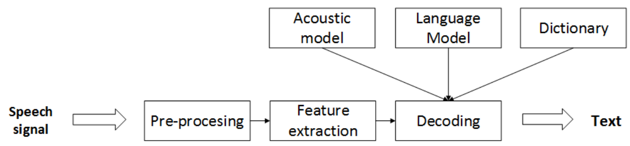


Fig. 1. Simplified flow diagram of HMM based automatic speech recognition.

A model of an unsupervised training system, theoretical basis for features and hidden Markov model for acoustic decoding and modeling will be proposed in this paper. [1] [2] [3] [14].

II. HMM based automatic speech recognition

Basic concept of ASR system based on HMM is shown on Fig. 1. There are 3 main blocks where spoken utterance is pre-processed, acoustic signal features extracted and at last feature vector is decoded supported by acoustic and language model and dictionary. Pre-processing and Feature extraction can be merged into "signal processing front-end" [13].

A. *Signal processing front-end*

Purpose of signal processing front-end is to transform signal from acoustic representation of speech to parameters which mostly describe the signal. Parameters should:

- be perceptually significant, which means that they should represents characteristics of speech signal which si most similar to human ear perception characteristics.
- be speaker, communication channel or environment independent.
- contain information about spectral changes in signal over time.
- be normalized and not redundant.
- be computed effectively by current computers.

1) Digital filtration: Analog to digital conversion is necessary and essential for pre-processing of audio signal. One of the most common pre-processing technique is digital filtration of speech signal. The purpose is to adjust speech signal in such way that boosts selected parts of frequency spectrum of the signal, for example raise signal-to-noise ratio etc. Typical filtration of speech signal in many ASR systems is pre-emphasis filter which emphases frequency components approximately 20 dB/decade. Such filtration reduces natural attenuation and perceptually emphasizes more important parts of speech spectrum [10].

Pre-emphasis is realized by finite impulse response (FIR) filter with transfer function:

$$H_{pre}(z) = \sum_{k=0}^{N_{pre}} a_{pre}(k) z^{-k} \tag{1}$$

More common is a transfer function with just one coefficient a_{pre} with typical range from -1 to -0.4.

$$H_{pre}(z) = 1 + a_{pre} z^{-1} \tag{2}$$

2) Temporal speech signal processing: Analog to digital conversion and pre-emphasis filtration is processed continually but all further operations are done sequentially frame-by-frame. Frame in this context is a sequence of digital audio samples which represents signal in a time period from 10 to 20 mili-seconds. Signal is considered as stationary on such small periods of time. This range is divided from articular organ movement speed and shape transformation frequency of articular tract during speech production.

Both frame edges are overlapping with previous or following respectively. Frame overlap is usually from 33%-66%. The problem with rapid increase or decrease of signal on on frame edges is solved by a window functions applied on frame. Signal samples in frames are weighted by windows function. Most commonly is Hamming window function used but there are more of them available like Hann, Blackman, Kaiser.

3) Spectral analysis: The goal of spectral envelope estimation is to get parameters which represent vocal tract transfer functions. There are many methods used such as frequency bank based methods, linear prediction algorithms, discrete Fourier transformation (DFT) and very popular cepstral coefficient calculation.

Gain and vocal tract impulse response characteristics separation was proposed by homomorphic filtration (HMF) signal decomposition. HMF is very efficient and became one of most used for speech signal spectral analysis. Result of HMF is a set of parameters representing spectral envelope. These parameters are called cepstral coefficients because they are obtained from cepstral transformation. Logarithmic cepstral transformation is widely used by ASR systems.

$$\hat{x}(m) = \frac{1}{2\pi j} \oint_C log[\sum_{n=-\infty}^{\infty} x(n) z^{-n}] z^{m-1} dz \tag{3}$$
$$-\infty < m < \infty$$

where $x(m)$ is complex spectrum which is a mirror of input image in quefrency domain. Complex cepstrum can be effectively computed by Fourier (DFT) , Harley, sine or cosine transformations. In modern ARS systems,DFT method is used. Almost all ASR systems use non-linear frequency scale instead of linear. Usually they use Mel frequency scale. Mel-frequency cepstrum coefficients (MFCC) are often used, which differ only in frequency scale with cesptrum coefficients proposed earlier.

$$\hat{c}_{mel}(m) = \frac{2}{N} \sum_{i=1}^{N_{cb}} \hat{X}_{cb}(k_i) \cos(\frac{2\pi}{N} m k_i), \tag{4}$$
$$m = 1, 2, ..., L,$$

4) Normalization of spectral parameters: Normalization of spectral parameters is a method for gaining smoother shape of spectral envelope minimizing unequal loudness influence in speech signal. Benefit of this method is adaptability for speaker/microphone change or noise influence in communication channel. MFCC parameters may be normalized by cepstral mean normalization [11] [16].

5) Spectral parameters derivation: Disadvantage of temporal spectral analysis is that it does not represent longer dynamic changes in spectrum. Usually auxiliary parameters are added to feature vector which should describe time-varying spectrum changes of speech signal. The first and higher order time derivation of spectral estimation parameters are called delta or delta-delta parameters. There are many methods for delta parameters calculation, but normally they may boost parameter noise. To reduce such noise, we need to work with weighted or filtered temporal spectral estimations.

$$\dot{x}(n) = \frac{d}{dt} x(n) \approx \sum_{m=-N_d}^{N_d} m x(n+m) \tag{5}$$

In regression analysis in Eq. 5 parameter noise is reduced. In this derivation are subtractions between symmetrically spaced samples from current sample in time n. For computation of current sample are N_d samples used of both directions. Delta or Delta-Delta parameters should reflect dynamic changes is speech spectrum on longer time frame. They are computed from multiple frames. From equation above is clear that N states for number of frame and m is frame offset in forward or backward direction [12].

X. Huang et. al found in their research [5] that optimal number of parameters for feature vector is 39. 13 MFCC, 13 delta (Eq. 6) and 13 delta-delta (Eq. 7) coefficients are optimal and efficient for speech recognition and decoding. They experimentally found that 16 MFCC does not improve word error rate (WER) of ASR system. The third order delta coefficients also didn't improve WER. Combination of 13 MFCC, 13 delta and 13 delta-delta coefficients are considered as the most effective feature vectors.

$$\triangle c_k = c_{k+2} - c_{k-2} \tag{6}$$

$$\triangle\triangle c_k = (c_{k+3} - c_{k-1}) - (c_{k+1} - c_{k-3}) \tag{7}$$

B. Statistical modeling for ASR systems

Current ASR systems are based on hidden Markov models. HMM represents statistical model which emits a symbol sequence. Speech is a non-stationary process but in HMM modeling it may be assumed that on short time interval acoustic representation is unchanged. Left-to-right HMM is used in almost all ASR HMM based systems. HHM with non emitting states on both sided and emitting symbols between them are usually used [6] [7].

1) Acoustic modeling: Hidden Markov model is defined as double stochastic process (Q, O) where two related time sequences of random variables are generated. Stochastic process $Q = q_1^N = \{q_i\}_{i=1,2,...,N}$ is a finite sequence of model states. It is the underlying Markov first order string which is related to time varying changes of speech signal. This process can't be directly observed therefore is so-called "hidden". Stochastic process $O = o_1^T = \{o_t\}_{t=1,2,...,T}$ is emitted random variable sequence of HMM output.

The transition NxN matrix $A = \{a_{i,j}\}_{i,j=1,2,....,N}$ consist of discrete probabilities $a_{i,j}$ of transition from one state to another where:

$$a_{i,j} \triangleq P(q_t = j | q_{t-1} = i) \tag{8}$$

Hidden Markov string $Q = q_1^N = \{q_i\}_{i=1,2,...,N}$ is hidden Markov model for observed sequence $Q = q_1^N = \{q_i\}_{i=1,2,...,N}$ only if for every t is o_t defined and Eq. 9 is true.

$$b(o_t | o_1^{t-1}, q_1^t) = b(o_t | q_t) = b_{q_t}(o_t) = b_j(o) \tag{9}$$

where $b_j(o)$ is probability density for observation o in j state with time observation-independence assumption. Then HMM can be defined by set of parameters

$$\lambda = (A, B) \tag{10}$$

where B is vector $B = \{b_j\}_{j=1,2,...,N}$ of observation o probability densities in state j. HMM definition in general contain vector $\pi = \{\pi_i\}_{i=1,2,...,N}$ which describes probability of model starting state. In that case HMM would be defined as $\lambda = (A, B, \pi)$. As proposed earlier HMM may have non-emitting starting symbol thus hidden vector Q with state q_0 and q_{T+1} is extended. Then HMM can be defined Eq. 10. Hidden Markov process can be defined as:

$$G = (Q, O, \lambda) \tag{11}$$

Probability of symbol emission in specific state is set by probability density function (PDF) which shape is modeled in HMM training process. PDF can be discrete (DHMM) or continuous (CHMM). If two or more states shares PDF function of one state then HMM is called semi-continuous SCHMM. DHMM uses vector quantization where codebook is added to system. Those systems are not very often used by ASR system due accumulation of quantization distortion in training process. ASR systems adopted SHMM method with continuous Gauss PDF functions.

$$N(o, \mu, \Sigma) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \; exp(-\frac{1}{2}(o - \mu)' \; \Sigma^{-1} \; (o - \mu)) \tag{12}$$

where n is observation vector size, μ is mean vector Σ is covariance matrix. In practice non-diagonal elements are set to zero to reduce number of available model parameters.

Parameter vector's PDF has usually complicated shape therefore their modeling with simple Gaussian function with diagonal covariance matrix is not sufficient, especially for robust ASR systems. To cover acoustic space represented by feature vectors are Gaussian mixtures used. Emission probability can be described by equation:

$$b_j(o_t) = \sum_{m=1}^{M} c_{jm} N(o_t; \mu_{jm}, \Sigma_{jm}) \tag{13}$$

where M is number of PDF mixtures and c_{jm} is weight of m^{th} PDF mixture. To determine HMM parameters is only coarse parameters estimation used in the first step. Fine values can be obtained for example with Baum Welch re-estimation algorithm.

Size of HMM means how many states HMM has. It's topology describes which states transitions are allowed. In DHMM is codebook also HMM size parameter. CHMM size parameter is number of components in distribution mixture. If is phoneme used as basic element of ASR system then

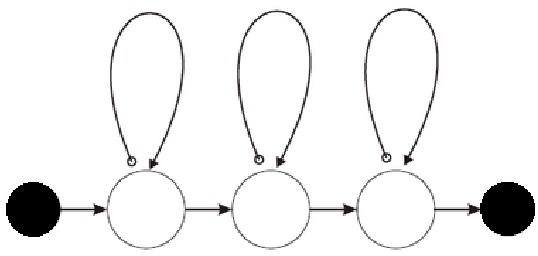


Fig. 2. HMM topology for phoneme modeling.

HMM uses often 3 - 5 states. Words level HMM uses more states. Less than 3 states per phoneme could not map all co-articulation phonemes. Example for phoneme HMM topology is shown on Fig. 2. Model has 2 non-emitting states on both ends and 3 emitting between them [8] [9].

2) Decoding: Decoding is final step of speech recognition process where is found such word sequence $\hat{W}$ which satisfies Eq. 14.

$$\hat{W} = \arg\min_{W \in V} P(O|W) \; P(W) \tag{14}$$

where $P(O|W)$ is represented by acoustic model and $P(W)$ is represented by language model. To this problem are two method proposed.

- decoding with deep lookup;
- decoding with wide lookup;

Deep lookup decoding is search for the most likely hypothesis by sequential way from the beginning of sentence to its end. For this method can be stack or A* decoders used.

Wide decoders do parallel search for hypothesis like Viterbi algorithm [4]. If $\phi_j(t)$ represents maximum probability that observed sequence of speech vectors o_1 to o_t and model was in time t in state j then probability can be calculated as:

$$\phi_j(t) = \max_i \{\phi_i \; (t-1) a_{ij}\} \; b_j \; (o_t) \tag{15}$$

when starting conditions $\phi_1(1) = 1$ and $\phi_j(1) = a_{1j} b_j(o_1)$ for $1 < j < N$ is satisfied. Maximum probability is calculated by solving Eq. 16.

$$\phi_N(T) = \max_i \{\phi_i \; (T) a_{iN}\} \tag{16}$$

Multiplying very small numbers could add inaccuracy to calculation, therefore it is logarithmic scale used. Viterbi algorithm can be interpreted as the best path lookup over graph where x-axis represents observation vectors and y-axis states of HMM.

$$\psi_j(t) = \max_i \{\psi_i \; (t-1) + \log(a_{ij})\} + \log(b_j(o_t)) \tag{17}$$

Speech recognition with CHMM based ASR requires some changes in proposed theory of Viterbi algorithm. So-called Viterbi token passing algorithm is based on that object (token) exists and remain its logarithmic value $\psi_j(t)$ during state transition (moving over states) in time t. This value represents partial probability comparison of observation o_1 to ot and HMM model with restriction that model is in time t in state j. Basically Viterbi token passing have 2 steps.

- Token copy in state i is placed to all j states (only to those where is transition from state i to j allowed). New logarithmic probability is calulated as $\log[a_{ij}] + \log[b_j(o_t)]$.
- All tokens in state j are compared and that with the highest probability remains and other are deleted.

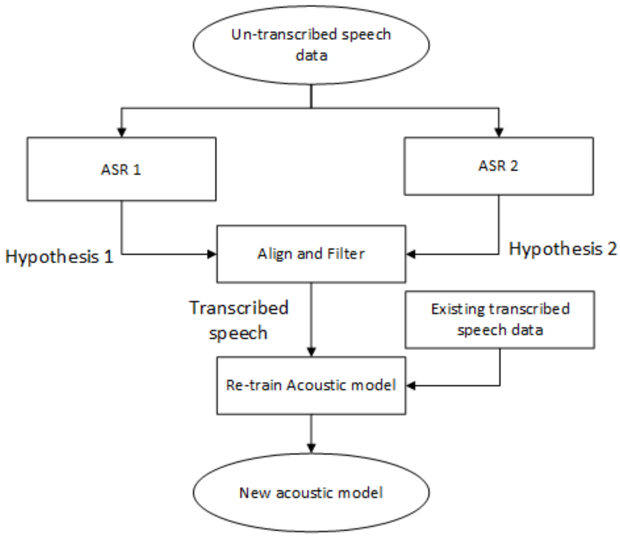


Fig. 3. HMM based automatic speech recognition flow diagram.

III. Future work and motivation

Slovak language in speech recognition field is not under-resourced language any more. Manually transcribed data of several hundreds hours are available for acoustic modeling. Moreover ASR systems in present time can produce satisfying result on some circumstances e.g. when speech is prepared or carefully read text. for example speech recognition in noisy background is difficult.

Recognition quality depends on dictionaries, acoustic and language model. Training acoustic models on many speech data may have significant quality improvement. Speech data have to be transcribed which require lot of effort (and cost) to obtain. Another point is that if actual model modeled with hundreds hours of transcribed speech data is retrained with small amount of speech data then recognition improvement wouldn't be very significant. Retraining such acoustic model would need many times more of speech data to reduce satisfying amount of WER.

Unsupervised acoustic modeling as proposed in introduction would deal with this problem. Speech data in sufficient audio quality is available nowadays. TV or radio broadcast, lectures, audio books even Slovak parliament sessions are easy to obtain. Model of unsupervised acoustic modeling with such data without transcription is proposed on Fig. 3.

Unsupervised acoustic model training may be valuable for under resourced languages and could have significant improvement in WER. H. Cucu successfully tested that seed AM made of several hours may be retrained iteratively with untranscribed speech data resulting better recognition quality on Romanian Language.

There are many methods for unsupervised training like method with just one ASR system with selecting hypothesis based on thresholded recognition score. Output sequence which passes threshold is considered as correct and model is retrained with it. Another method with two independent ASR systems as shown on Fig. 3 will be designed and build with HTK tools. Further experiments will be done with different configurations.

In some cases changing of dictionaries, language and acoustic models prior recognition depended on speaker, gender, domain and environment could increase quality of speech recognition in some cases. Acoustic model may be created (or retrained) with unsupervised acoustic modeling.

Of course that not only acoustic model is responsible for errors in recognition. Language models could be adapted for domain, and speaker as well. New words are used in Slovak language almost on weekly basis, e.g. names, places, events or foreign words. These words should be added to dictionary because ASR system can recognize only those words which are in dictionary. These topics and grapheme modeling will be studied in future.

IV. Conclusion

This article describes selected methods for feature extraction, speech signal parametrization and acoustic model training and decoding as theoretical basis for further work with unsupervised acoustic model training. Result of authors future work should be working dual ASR unsupervised acoustic model training system and further experiments on them.

Acknowledgment

The research presented in this paper was supported by the Ministry of Education, Science, Research and Sport of the Slovak Republic under the project VEGA 1/0075/15 (50%) and the Research and Development Operational Program funded by the ERDF under the project ITMS-26220220182 (50%).

References

[1] H. Cucu, A. Buzo, L. Besacier, C. Burileanu, *SMT-based ASR Domain Adaptation Methods for Under-Resourced Languages: Application to Romanian in Speech Communication*, Vol. 56 Special Issue on Processing Under-Resourced Languages, 2014, pp. 195-212.
[2] H. Cucu, L. Besacier, C. Burileanu, *Text spotting In Large Speech Databases For Under-Resourced Languages*, in Proc. Int. Conf Speech Technology and Human-Computer Dialogue (SpeD), Cluj-Napoca, Romania, 2013, pp. 77-82.
[3] G. Zavaliagkos, T. Colthurst, *Utilizing Untranscribed Training Data to Improve Performance in DARPA Broadcast News Transcription and Understanding Workshop*, Lansdowne, USA, 1998, pp. 301-305.
[4] J. Psutka, L. Müller, J. Matoušek, V. Radová, *Mluvíme s počítačem česky*, Academia Praha, 2006.
[5] X. Huang, A. Acero, H. W. Hon, *Spoken Language Processing: A Guide to Theory, Algorithm, and System Developement*, New Jersey. Prentice Hal 2001.
[6] L. R. Rabiner, B. H. Juang, *Fundamentals of Speech Recognition*, Prentice Hall: Signal Processing Series, 1993.
[7] J. R Deller, J. G. Proakis, J. H. L. Hansen, *Discrete-time Processing of Speech Signals*, Englewood Cliffs, N. J. USA. Prentice Hall, 1993.
[8] S. Lihan, *Automatické rozpoznávanie reči pomocou skrytých Markovových modelov*, Ph.D. Dissertation, FEI TU Košice, 2004.
[9] J. Juhár, S. Ondáš, A. Čižmár, M. Rusko. G. Rozinaj, R. Jarina, *evelopment of Slovak GALAXY/VoiceXML Based Spoken Language Dialogue System to Retrieve Information from the Internet*, In: Proc. of the 9th Intl. Conf. on Spoken Language Processing, INTERSPEECH - ICSLP, pp. 485-488, Sep. 2006.
[10] J. R. Deller, J. G. Proakis, J. H. L. Hansen, *Discrete-time Processing of Speech Signals*, Englewood Cliffs, N. J., USA, Prentice Hall, 1993.
[11] A. V. Oppenheim, R. W. Schafer, *Discrete-time Signal Processing*,Englewood Cliffs, NJ, USA, Prentice-Hall, 1989.
[12] E. L. Bocchieri and G. R. Doddington, *Frame specific statistical Features for Speaker independent Speech Recognition*, IEEE Trans. on ASSP. Vol. 34, No. 4, pp. 755-764, Aug. 1986.
[13] J. Juhár et al., *Rečové technológie v telekomunikačných a informačných systémoch*, Equlibria, Košice, 2011, pp. 41-124 ISBN 978-80-89284-75-7.
[14] H. Cucu, A. Buzo, C. Burileanu, *Unsupervised acoustic model training using multiple seed ASR systems*, in proc. of SLTU-2014, St. Petersburg, Russia, 2014.
[15] M. Heck, *Unsupervised Acoustic Model Training for Simultaneous Lecture Translation in Incremental and Batch Mode*, Diploma Thesis, Dept. of Informatics, Karlsruhe Institute of Technology, Dec. 2012.
[16] P. S. Moharir, *Deconvolution by Cepstral Transformation*, Journal Inst. Electronics and Telecom. Eng., Vol. 28, pp. 515-532, 10 1982

Big Data Analysis for Creation of Added Value in Organizations

[1]*Miroslava MUCHOVÁ (1st year)*
Supervisor: [2]Ján PARALIČ

[1,2]Dept. of Cybernetics and Artificial Intelligence, FEI TU of Košice, Slovak Republic

[1]miroslava.muchova@tuke.sk, [2]jan.paralic@tuke.sk

***Abstract*— Timely and cost effective big data analysis is key factor of success in many companies. Companies have huge amount of structured and unstructured data in which they want to: search, look up relations and extract important information. Big data allows the companies to gain insight and make better decisions, so they gain a competitive advantage on the market. The aim of this paper is to explain and highlight the big data impact in organizations and also explain how big data helps to creating additional value, because big data does not only allow us to rise the value of goods and services, but also make processes simpler and more effective. We want also show how big data allows the companies to understand customers' behavior.**

***Keywords*— Big data, data analytics, value added, customer relationships**

I. Introduction

Nowadays, data analysis is having growing importance. With rising usage of computers in organizations, huge amount of data has been generated. Data analysis is important because of ability to change the data into information and knowledge useful for the business. This analysis allows e. g. segmentation of customers, i.e. dividing customers into groups based on their similar characteristics; offer the customers goods and services which satisfy their needs; predict who will respond to the supply of certain good; forecast occurrence of an insurance event etc. Organizations and companies use data for decision support task in order to maximize their profit, make processes more effective and simpler, increase the effectiveness, and minimize frauds and costs. All of these tasks help to maximize the value of organization.

Section 2 discusses the basic description and characteristics of big data, 3Vs and other characteristics. At the end of this section, general architecture of the big data analysis is mentioned.

Section 3 provides an overview of how organizations are using big data in organizations, basic characteristics of business intelligence and business model. Customer relations and appropriate big data analysis tasks are also described. The last part of Section 3 focuses on increasing the value added and value of enterprise with use of big data.

II. Big Data

In recent years, big data has become the key source of competiveness on the market. Big data analysis is important mainly for companies with huge customer network. If company wants to analyze what is happening inside as well as outside of it, it needs huge volumes of detailed data. With data sample, an analysis can be conducted. Based on the analysis results, data models are created and using prediction business is optimized. This is just one of examples of usage the data. However, every data analysis has a common objective – making better decisions.

The term "big data analytics" is used to refer to the process of examining large amounts of data in an effort to uncover hidden patterns or unknown correlations [1]. Another definition states that big data analytics is process of using advanced analytic methods to analyze big data files [2].

The business and non-business subjects have to realize that importance of information is growing. Information has never been as easily accessible as nowadays. Except of that, they are spreading very quickly.

A. 3Vs of Big Data

Big data is often characterized by means of 3V: volume, velocity, variety. Many researchers and experts add also other characteristics such as variability, veracity and complexity.

Audio records, pictures, photographs, tweets and facebook' statuses, call center conversations, text documents represent unstructured data which significantly contribute to rising amount of data. Volume describes data sets which are notable big (measured from terabytes to zettabytes) [2]. The most important task is the scalability of data. The biggest challenge is both storing and fast processing of huge volumes of data. In last decades, lot of researchers focus on optimization of algorithms and accelerating the process of data analysis [3].

The term "velocity" means the speed of generating data or the speed of data generation and processing [1]. Data changes daily, and the objective is to maximize their value. There are many types of the data sources which are dynamical, e.g. nearly real-time big data sources providing changing information about stock markets including supply and demand prices, etc [2].

Another aspect of big data processing is the variability, i.e. type of data format. Data occur in structured (such as tables, numerical values from database) and unstructured form (such as e-mails, videos, audios, text documents etc.) [2]. Unstructured data cannot be easily fitted into some data model, so they do not have formally defined structure [3]. By analyzing the data, it structure may be deduced. Data also exists in semi-structured forms organized in semantic entities [4]. This type of data is represented by XML data, data decoded in markup languages (HTML, SVG)…

Variability refers to inconsistencies in the data stream, inconstancy in data analysis in time [5]. Especially, variability is important in sentiment analysis tasks.

Data is worthless unless it is accurate. We can state that quality of an analysis depends on the veracity of data sources [2] which is also an aspect of big data.

Complexity is used to describe that data comes from different sources and integrating these sources and transformation of this data is challenging and difficult task [2] for many companies.

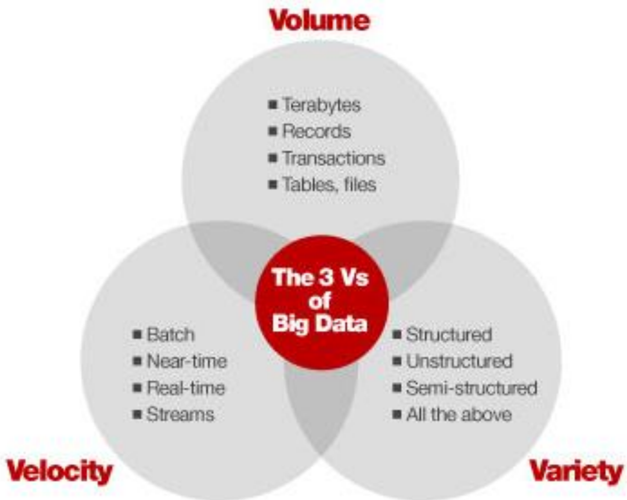


Fig. 1 3Vs of Big Data [2]

B. *Architecture of Big Data Analytics*

In the Fig. 2, general architecture of big data analytics is depicted [6]. It consists of the following three parts: multi-source big data collecting, distributed big data storing and intra/inter big data processing.

The first block on Fig. 2 stresses 3V characteristics of big data [7]. The goal of the next block is to organize the data effectively, and the third one shows the needs to analyze structured and unstructured data in order to maximize their value.

The volume of generated data rises faster than the budgets in companies. This discrepancy between spending and requirements on data storing increases the difficulty of effective data storing and processing.

Data storing is computationally-intensive science which conducts the tasks in distributed computing systems [6]. The problem is the lack of space for storing such huge volume of data. Some data coming from financial and medical science had to be removed because of lack of storing space [7]. The priority of knowledge discovery process in big data is its accessibility and availability [10]. The stores need to provide quick and as simple as possible access to data. The data storing requires storing platforms which are distributed, scalable, elastic and fault-tolerant [12]. Data warehouses represent one of ways how data can be stored.

Many technologies used in big data processing are outdated and do not manage to process large amounts of data, so if we want to capture the value of big data new techniques and technologies for analysis need to be developed [8]. Big data is stored in distributed way and should be process in parallel way [9]. Different applications require different techniques and optimization.

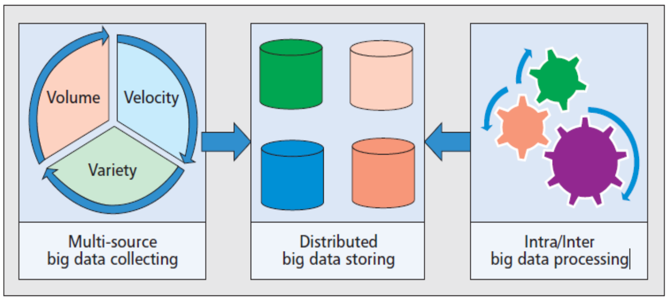


Fig. 2 Architecture of Big Data Analytics [11]

III. BIG DATA AND ORGANIZATIONS

Companies and organizations realize the potential value of their data which allows understanding their customers. Using these analyses, companies can make a great progress in different areas and support their decision making. In addition, big data provide a place for development of real-time applications and services working with electronic data in order to present the value to the customer. The objective of each organization usually are to maximize its value, rise profit and effectiveness, make its business processes more simple and effective. With big data era, the problem of integrating of information coming from new data sources to own enterprise systems appears [3]. In addition, the question of legitimate usage of data and customer privacy arises. Therefore, it is necessary that the companies must use mechanism providing data security before starting big data project [4].

Companies aiming to effectively use big data have to answer the following questions [5]:

- What if we will not be able to manage large volumes of data because data are too big and too varied?
- Which data has to be stored?
- Which data has to be analyzed?
- How to find out which data points are really important?
- How can the data be used to best advantage?

Companies that use their data reasonably, can understand the behavior of their customers in the past and use their actual (real-time) data in decision process, so they are able to maximize their productivity and profit. The profit may be from 5 to 6% higher than the profit of competitive companies not using the data effectively [11]. Companies that are able to process and evaluate these large volumes of data have significant advantages [10] in comparison with companies lagging in this field.

In the Fig. 3, there are results of a survey conducted by TDWI [11]. The respondents answered the question whether the organization they work in executes advanced analytics against big data.

As can be seen, 74% of respondents have practiced some kind of data analysis. 40% of respondents practice some form of advanced analytics, but not with big data. Data analysis with and application of big data is performed by 34% of asked analysts. 23% of respondents do not practice any form of advanced analytics.

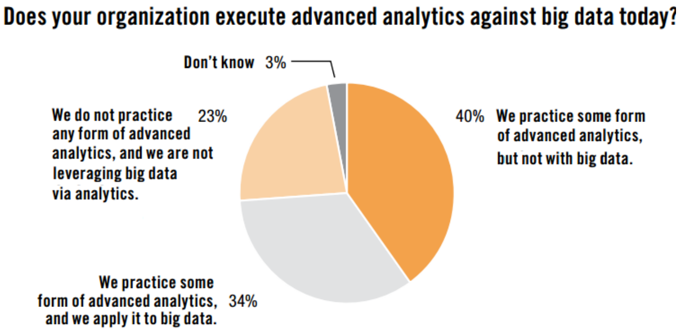


Fig. 3 Big Data analytics [10]

A. Business intelligence

Business intelligence tools help the organizations to understand their internal and external environment [12]. They help to collect, normalize, analyze, presentation and interpretation of business (transaction) data [13]. The primary objective of such systems is supporting processes, decisions and planning in many areas of business

B. Business model

In data mining, model means the structure modeling relations among features of cases covered by available data. Models gained as the results of data mining have been designed to support the decision. A conflict between model accuracy and the feasibility of a simple interpretation occurs very often [14]. More complex models often provide more accurate results, but it is usually difficult to clarify the principles of these models to other people, which can make the implementation more difficult.

The business model is designed for a wide range of formal and informal descriptions within the business. Its feature is to define and classify the attitude of company for the business. It represents the main characteristics of certain company and covers overall strategy for the organizations (depicted in the Fig. 4), determines how the business creates, distributes, manages and improves different forms of value, including economic, social, cultural and technological value [13]. Based on the analysis results, company may monetize the data and create new business model. Based on the business model, an analysis of data applications is carried out [15].

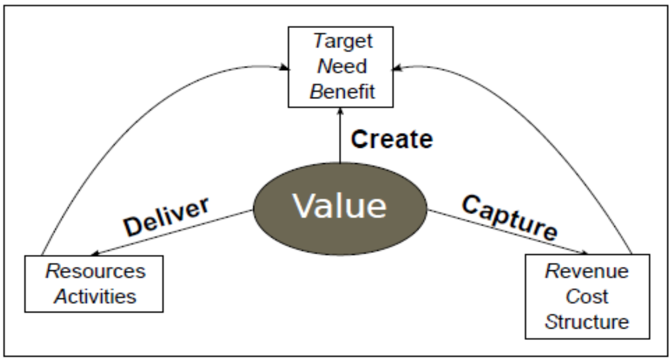


Fig. 4 Business model [16]

C. Customer relationships

Using big data, companies are able to segment their customers, forecast the reaction of customers on goods supply, resulting in reduction of spending on marketing [15]. Big data analyses allow creation of models which can determine whether the customer will be interested in buying also other product of a certain company. Big data also "contains" information needed to create model determining and describing the characteristics of leaving customers [16]. Customer relationships topic describes how big data "reduces" the gap between customers and companies and companies can better satisfy their customer's needs [17]. Therefore, companies have to improve and innovate theirs services and look for new approaches to managing customer relations.

In examining customers' behavior, it is necessary to find correlations that may be indicators of causal relationships. These can be used in marketing to improve target marketing [16]. When considering unstructured data, it is useful to know how to extract interesting links and relations among data.

Using big data analysis and continuous monitoring and by analysis of transaction data coming from e-shop, we can create model describing the shopping behaviors [15]. This model let the company to improve their target advertising and savings on spending enable maximization of company's value [17]. Both customer and company are satisfied, customer gets what he asked for and company maximized its profit.

D. Value added

The term "value" refers to an indicator of business performance . The objective of companies is to optimize and maximize this performance [8]. Value added comes from employers' salaries, as well as interests earned and product profit [17].

The organization increases its value, when value created by its own activities is higher than costs spent on conduction of these activities [18].

Companies using big data for analyses and making decision are reducing the cost and the risk of wrong decisions. Data may prove whether the company was working like its analyses and predictions showed [3]. This s the highest value because it firstly validates its forecasts and predictions and after validation, decisions are taken.

In most of organizations, there are countless amounts of data which do not need to be stored only in databases [17]. A lot of interesting and possibly useful data and information can be found in company's book-keeping entries, e-mail communications etc. Finding interesting and important relations is often very difficult not only due to size of the data but also due to lack of structure [19].

Companies operating in the manufacturing sector use their data mainly for improving their information technologies and automation. According to several studies [20], there is a possibility of reducing operating costs in almost 50% across all sectors of manufacturing with usage of big data potential.

Long-term usage of big data leads to increased productivity and reduced costs. Manufacturing companies have a lot of data which can be analyzed and turned into the value, but they are not able to use their potential and gain a value [20].

Predictive analytics is a field in data mining which deals with the prediction of future probabilities, trends, risk forecast, and segmentation etc [19]. Predictive analytics enables improving key metrics of a company, which makes higher profit and faster growth [20].

IV. FUTURE WORK

The aim of my thesis is the analysis of big data in order to extract information from them, which organizations helped to increase the added value. By analyzing big data company would have to understand and know not only their own business, customers and suppliers, but also the relationship between market and any aspects relevant to a given segment in which the company operates.

Using various analytical tools help us to become familiar with the huge amount of information, thereby facilitating decision making in the company to accelerate the process etc. Selected tool to set up and prepare for advanced data analysis and design and implement applications to work with an analytical tool for the end user.

V. Conclusion

Nowadays, big data is coming to the center of attention and there is no doubt that our society has entered the era of big data. Big data is considered to be phenomena changing the future.

Big data comes from all sources and affect our lives. If the company is able to view this data correctly, it will be able to find answers to almost any questions.

Analyzing big data helps the organization to gain deeper understanding of their own businesses and what is the most important, it enables forecasting of customers' behavior and creates value added for the company itself. The aim of analyzing current and historical data is identification of potential threats and opportunities in organizations' environments.

Companies use big data for analytical purposes, decision support and big data also enable the development of big data applications and big data services using a huge amount of electronic data in order to present the value of the customer.

Acknowledgment

The work presented in this paper was supported by the Slovak Cultural and Educational Grant Agency of the Ministry of Education, Science, Research and Sport of the Slovak Republic under grant No. 025TUKE-4/2015.

References

[1] Shang Weiyi, “Assisting developers of Big Data Analytics Applications when deploying on Hadoop clouds.“ Software Engineering (ICSE), 35th International Conference on. San Francisco: IEEE, 2013. pp 402 – 411, 18-26 May 2013

[2] X.L. Dong, D. Srivastava, “Big data integration.” Data Engineering (ICDE), 2013 IEEE 29th International Conference on. Brisbane: IEEE, 2013. pp. 1245-1248. 8.-12. April 2013

[3] J. Zhang, M. L. Huang, “5Ws Model for Big Data Analysis and Visualization”. Computational Science and Engineering (CSE), 2013 IEEE 16th International Conference on. Sydney: IEEE, 2013. pp. 1021-1028, 3.-5.Dec.2013

[4] S.Singh, N. Singh,”Big Data Analytics.” Communication, Information & Computing Technology (ICCICT), 2012 International Conference on. Mumbai: IEEE, 2012. pp. 1-4. 19.-20.Oct. 2012

[5] A. Katal, M. Wazid, R.H.Goudar, “Big data: Issues, chalenges, tools and Good practices.“ Contemporary Computing (IC3), 2013 Sixth International Conference on. Noida: IEEE, 2013. pp. 404-409, 8.-10. Aug. 2013

[6] D. Loshin, “Business Problems Suited to Big Data Analytics.“ Big Data Analytics, 2013. pp. 11-19.

[7] A. Balar, N. Malviya, S.Prasad “Forecasting consumer behavior with innovative value proposition for organizations using big data analytics.“ Computational Intelligence and Computing Research (ICCIC), 2013 IEEE International Conference on. Enathi: IEEE, 2013. pp. 1-4, 26-28. Dec. 2013

[8] K. Krishnan, “Big Data Driving Business Value”, Data Warehousing in the Age of Big Data, 2013. pp. 101-123

[9] G.Lafuente, “The big data security challenge.” Network Security: Elsevier BV, 2015. pp. 12-14

[10] Barton and David Court, “Making advanced analytics work for you,” Harvard Business Review, October 2012, Volume 90, Number 10, pp. 78–83.

[11] P.Russom, “Big Data Analytics”. TDWI Best practices report. 2011, http://public.dhe.ibm.com/common/ssi/ecm/en/iml14293usen/IML14293USEN.PDF

[12] L. Rongxing, H. Zhu, L. Ximeng, “Toward efficient and privacy-preserving computing in big data era.”Network, IEEE, 2014. pp. 46-50, 24. July 2014

[13] Y. Shi, X. LU, “The Role of Business Intelligence in Business Performance Management.” Information Management, Innovation Management and Industrial Engineering (ICIII), 2010 International Conference on. Kunming: IEEE, 2010. pp. 184-186. 26.-28.Nov. 2010

[14] M. Terek, A. Horníková, V. Labudová, “Hĺbková analýza údajov.” Iura Edition, 2010, ISBN 978-80-8078-336-5

[15] R. Winter, O. Marjanovic, B.H. Wixom, “Introduction to Business Analytics, Business Intelligence, and Big Data Minitrack.” System Sciences (HICSS), 2013 46th Hawaii International Conference on. Wailea: IEEE, 2013. pp. 3767. 7.-10. Jan. 2013

[16] F. C. P. Muhtaroglu, S. Demir, M. Obali, “Business model canvas perspective on big data applications.“ Big Data, 2013 IEEE International Conference on. Silicon Valley: IEEE. pp. 32 – 37, 6-9. Oct 2013

[17] C.L. Philip Chen, Ch. Y. Zhang, “Data-intensive applications, challenges, techniques and technologies: A survey on Big Data” , Information Sciences. 2014. pp. 314-347. 10. Aug. 2014

[18] K. Kambatla, G. Kollias, V. Kumar, “Trends in big data analytics”, Journal of Parallel and Distributed Computing. 2014. pp. 2561-2573. July 2014

[19] N.Kshetri, “Big data’s impact on privacy, security and consumer welfare.“ Telecommunications Policy, 2014. pp. 1134-1145.

[20] M.S.B. Prasad, S.H. Sastry, “Big data and predictive analytics in ERP systems for automating decision making process.” Software Engineering and Service Science (ICSESS), 2014 5th IEEE International Conference on. Beijing: IEEE, 2014. pp. 259-262, 27.-29.June 2014 6933558

Big Data Tool Selection: Criteria and Benchmarking

[1]*Eva TURŇOVÁ (2nd year)*
Supervisor: [2]Ján PARALIČ

[1,2]Dept. of Cybernetics and Artificial Intelligence, FEI TU of Košice, Slovak Republic

[1]eva.turnova@tuke.sk, [2]jan.paralic@tuke.sk

***Abstract*— Big data processing and analysis tools offer the features which cannot be provided by traditional data processing tools. Nowadays, users may choose from wide range of tools according to their preferences, but these preferences may cause inappropriate tool selection and process of following data processing gets complicated. Due to missing complex view on selection of the analysis tool, there is still a huge potential for next research, which is addressed also by my work and sketched in this paper.**

***Keywords*—big data, tool selection, big data benchmarks.**

I. Introduction

Big data era comes with a lot of challenges. There is a need for creation of new approaches to big data analysis, new algorithms, new tools etc. Except these, there is also another interesting big data challenge – selection of appropriate tool. In this field, almost none research have been conducted. Only few authors discussed the suitability of big data analysis tool for different tasks. Results of this research are described in [1].

The main objective of our research is to find out critical factors influencing the suitability of a tool or an approach to big data analysis and create a decision support model. The model should include several criteria as well as results of big data systems benchmarks.

II. Selecting an Approach to Big Data Processing

There are only few models describing the suitability of certain approach to big data processing. The research potential and importance of this topic is outlined in [2]. According to one of principles of knowledge discovery process in big data sets, one tool is not suitable for each task [3].

Model summarizing existing works in this area is proposed in [1]. This model is focused on the big data strategies, i.e. approach which needs to be used. According to this model, there are four basic approaches or ways how to store and analyse big data. These approaches include relational databases (RDBMS, mainly in parallel form), big data as a service (BDaaS) provided by third party, distributed file systems (DFS) with MapReduce, and a hybrid approach. Every one of these strategies needs a different level of integrating certain big data technology into organisations' environment.

Authors of above mentioned model identified eight key contingency factors and grouped them into following groups: strategic factors, resource factors, and operating environment factors. Strategic factors include the relevance of big data analysis in terms of the business objectives, investment (sponsorship) level into big data technologies and urgency of analysis needed to be done. Resource factors describe the availability of human and financial resources. The availability and absorptive capacity of personnel are also referred as the resource factors. The third group focuses on factors influencing operating environment of an organisation. It includes two factors – routines of tasks (analytical tasks and queries) and issue of data privacy which can be limiting factor of putting the data into external cloud.

Model based on these factors is the only existing and usable model which can be used as a decision support tool. The main disadvantage of this model is its focus only on one aspect – organisation's environment and its characteristics. That makes this model incomplete because selection of the tool is more complex problem and it should include more factors.

III. Decision Support Model Proposal

As mentioned in Section 2, selection of appropriate tool is a complex problem and suitability of a tool needs to be examined from different points of view. Factors described in previous section can be referred as organisational factors due to their direct relation to the organisation's environment. These factors allow selecting the strategy, i.e. type or group of tools which are the most suitable in certain the organisation.

We propose a model illustrated in Fig. 2. This model contains two groups of criteria added – technological and financial factors. These groups directly influence the suitability of certain tool.

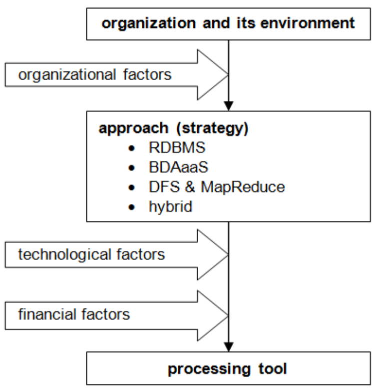


Fig. 1. Factors influencing selection of big data processing tool

Technological factors include characteristics describing technological aspects of big data and its processing. These criteria are e.g. the amount of data to be processed and analysed, the average size of records, the need of real-time processing, the structure of available data, the desired features according to CAP theorem etc. The most of technological factors can be mapped to 3Vs. These factors' impact may be supported by results of benchmarks which are described in Section 4.

Other very important aspect to consider is financial aspect. There are a lot of different financial measures expressing how much the investment costs and what value it creates for organisation. Since most of big data tools are cloud-based, we proposed adopted versions of TCO (total cost of ownership) and ROI (return on investment) measures in [4] for measuring cost and investment effectiveness.

IV. Big Data Benchmarks

Big data tools deal with big data processing problems and challenges in different ways. Evaluation and comparison of performance of each system (or tool) is the main reason of developing and creating new benchmarks [5]. The creation of new benchmarks is very actual topic both in industrial and research community. As big data is still new and developing research area, also creating benchmarks brings new interesting challenges.

The most of current state-of-the-art benchmarks were designed for specific types of systems, e.g. HiBench [6] and GridMix [7] for Hadoop clusters, Yahoo! Cloud Serving Benchmark (YCSB) [8] for Cassandra, HBase, PNUTS (now known as Sherpa) and shared MySQL implementation, TPC-DS [9], BigBench [10] for comparison of DBMS and MapReduce systems, LinkBench [11] for systems working with graph data, BigDataBench [12] for comparison of different processing architectures, and many others.

Based on knowledge from practice and existing benchmarks, benchmark should support bigger data and workload diversity, but workload suit should not be too wide in order to avoid time-consuming computations [5][13].

A. *Benchmarking process for big data tools*

Benchmarking process of big data systems consists of five steps illustrated in the Fig. 2.

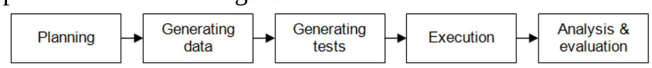


Fig. 2. Benchmarking process

Planning

The first step of benchmarking process is planning when the objective of benchmarking, application domain, and performance evaluation metrics are selected [5].

Generating data

For running different workloads, we need data. Benchmarks use synthetically generated or real data. Each of these two alternatives brings some problems. Getting real data is difficult because data owner is usually not willing to let third party get an access to the data. This data is also not as flexible as synthetic one. Therefore, some researchers recommend using synthetic data [5]. But there rises a question whether synthetic data are "real" enough and whether the results obtained using synthetic data will correspond to the results of real data. Therefore other authors recommend the use of real data [14].

When using synthetic data, the data generator has to be able to generate different amount of data as input of typical workloads, e.g. text processing algorithms, graph data systems etc. Data generators should also support setting the velocity of data generating in order to determine whether the system is suitable for processing data streams or not. Generators should also allow generating different data structures – structured data, unstructured and semi-structured data and generate data specific to various sources (tables, text, data streams, graphs, etc.)

Generate tests

Successful results corresponding to the reality is meaningful only when applying an application-specific benchmarking test [15]. Task from application domain needs to be abstracted to their basic functionality such as specific read, write and drop operations. One workload may consist of several operations.

Typical workloads include read, write, rewrite and updates, sorting algorithms, PageRank, WordCount and other.

Execution

Execution phase includes execution of selected workloads (tasks) with use of selected data. Execution of formulated tasks runs as a series of experiments done by several tools or on different architectures.

Analysis and evaluation

In the last step, the results of benchmark experiments are analyzed and evaluated. Successful and effective benchmarks give realistic and accurate results, which enable to make good decisions regarding tool to be used for particular big data analytic task.

V. Conclusion and Future Work

By analysing existing approaches to big data tool selection and studying theoretical background of this topic, we have stated these tasks:

1. identify the most important criteria in each group of factors mentioned in Section 3;
2. create big data benchmark and run series of experiments and examine whether tool features correlate with benchmark results;
3. create data generator for benchmarks with ability to simulate real world data.

Successful fulfilling these tasks may contribute to current state of research in this area and help to cover uncovered research gap.

Acknowledgment

This work was supported by the Slovak Grant Agency of Ministry of Education and Academy of Science of the Slovak Republic under grant No. 1/1147/12.

References

[1] K. Ebner, K; T. Buhnen; N. Urbach, "Think big with big data: Identifying suitable big data strategies in corporate environments," in *System Sciences (HICCS), 2014 47^{th} Hawaii Int. Conf. on,* pp. 3748 - 3757.

[2] E. Turňová; "Introduction to big data," in *SCYR 2014: 14th Scientific Conference of Young Researchers: proceedings from conference*. pp. 197 - 200

[3] E. Begoli; J. Horey, "Design principles for effective knowledge discovery from big data," in *Software Architecture (WICSA) and European Conference on Software Architecture (ECSA), 2012 Joint Working IEEE/IFIP Conf. on.* pp 215 - 218

[4] E. Turnova; A. Tusanova; J, Paralic, "Learning support tool for SMEs to measure their cloud investments," in *Emerging eLearning Technologies and Applications (ICETA), 2013 IEEE 11th Int. Conf. on* , pp.393 – 398

[5] R. Han; L. Xiaoyi; X. Jiangtao, "On big data benchmarking," in *Big Data Benchmarks, Performance Optimization, and Emerging Hadrware Lecture Notes in Computer Science*, vol. 8807. Springer International Publishing, 2014, pp. 3 – 18

[6] S. Huang; J. Huang; J. Dai, "The HiBench benchmark suite: Characterization of the MapReduce-based data analysis," in *Data Engineering Workshops (ICDEW), 2010 IEEE 26th Int. Conf.* on, pp. 41 - 5

[7] GridMix documetation

[8] B. F. Cooper; A. Silberstein; E. Tam; R. Ramakrishan; R. Sears, "Benchmarking cloud serving systems with YCSB," in *SoCC 2010 Proc. of the 1st ACM symposium on Cloud computing*, pp. 143 - 154

[9] R. Othayoth; M. Poess, "The Making of TCP-DS," in *Proc. of the 32nd Int. Conf. on Very Large Data Bases*, pp. 1049 – 1058

[10] A. Ghazal et al., "Big Bench: towards an industry standard benchmark for big data analytics," in *Proceedings of the 2013 ACM SIGMOD Int. Conf. on Management of Data*, pp. 1197 - 1208

[11] T. G. Armstrong; V. Ponnekanti; D. Borthakur; M. Callaghan, "LinkBench: a database benchmark based on the Facebook social graph," in *Proc. of the 2013 ACM SIGMOD Int. Conf. on Management of Data*, pp. 1185 – 1196

[12] L. Wang et al., "BigDataBench: A big data benchmark suit from internet services," in *High Performance Computer Architecture (HPCA), 2014 IEEE 20th Int. Symposium on*, pp. 488 – 499

[13] C. Baru; M. Bhandarkar; R. Nambiar; M. Poess; T. Rabl, "Setting the direction for big data benchmark standards," in *Selected Topics in Performance Evaluation and Benchmarking, LNCS*, vol. 7755. Springer, 2013, pp. 197 – 208

[14] X. Wen et al., "A characterization of big data benchmarks," in *Big Data, 2013 IEEE Int. Conf.* , pp.118 – 125

[15] Y. C. Tay, "Data Generation for Application-Specific Benchmarking"

Blind Image Steganalysis

[1]Vladimír HAJDUK (1st year)
Supervisor: [2]Dušan LEVICKÝ

[1,2]Dept. of Electronics and Multimedia Communications, FEI TU of Košice, Slovak Republic

[1]vladimir.hajduk@tuke.sk, [2]dusan.levicky@tuke.sk

Abstract—**This paper gives an actual status of knowledge in the field of blind image steganalysis. Steganalysis is a scientific discipline utilized for detecting the subliminal channels established for secret communication among two participants. On the other hand steganography aims to creating those channels. Steganalysis uses the characteristic features which are changed after embedding the secret information into unsuspicious data by steganography. This work focuses on the methods which use image as transferring media for secret message. Discussed methods of steganalysis extract the features from different domains to create universal, blind detecting system. The classifier is an important part of every blind steganalytic method so there are mentioned different individual and ensemble techniques used nowadays.**

Keywords—**blind steganalysis, classification, ensemble classification, steganalysis, steganography.**

I. Introduction

The term *steganography* consists of two Greek words: *steganós* – hidden and *grapheín* – to write. So the aim of that scientific field is to hide a secret message in the background of any transmitted multimedia data, originally in the text form. Related discipline to steganography is *steganalysis* that is utilized to break a system of steganography (in other words its purpose is to detect a presence of the hidden message). Steganalysis is being developed side by side with the systems of steganography. Steganography is the part of wider area which is called *information hiding*. As it is illustrated in the Fig. 1, information hiding, in accordance with the purpose and method of use, is divided into several groups [1].

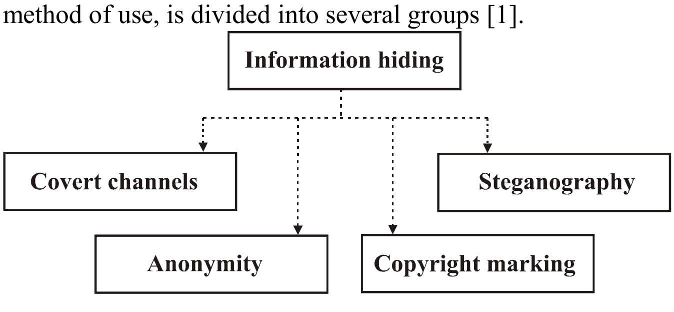


Fig. 1. Classification of an information hiding.

Covert channels represent the use of a communication channel for the purpose for which it was not intended.

Anonymity is the method which is used to hide original source of the information.

In the steganography information is hidden in order that only the recipient knows where to look for the hidden message.

Copyright marking forms a group of methods intended for user authentication or authorization data source. This may include watermarks or user authentication using biometrics (e.g. fingerprints).

A. Image Steganography

The main goal of steganography is to transfer the secret message in a background of unsuspicious data (cover data). Any type of multimedia information, such as video, audio, still images, and text, can be possible cover data. Steganography discussed in this paper utilizes still images as cover data. Such branch is called *image steganography*. According domain where embedding process is performed, image steganalysis can be divided into two main classes: *spatial domain image steganography* and *frequency domain image steganography*. The former works with the images without necessity of their transformation to another domain, whereas the second one embeds message to the image transformed to a frequency domain. The most commonly used transformations are discrete cosine transformation (DCT) and discrete wavelet transformation (DWT). In both cases LSB (Last Significant Bit) is the most popular steganographic method. It deals with the substitution of LSB bits of certain code words by bits of the secret message.

B. Image Steganalysis

Steganalysis is a scientific discipline primarily utilized to detect the subliminal channels established by steganography.

Steganalysis focused on a detection of the secret message potentially embedded in the still images, or simply method used to detection of the image steganography is likewise called *image steganalysis*.

When steganalytic technique is adapted to the steganographic method and characteristics which are influenced, this technique can achieve higher efficiency in the process of detection. Such a system of steganalysis is called *targeted steganalysis*. On the other hand there is a *blind steganalysis*. It has no information about used steganographic method. Blind steganalysis usually extracts more statistical features in the spatial and transform domain for detection more than one steganographic tool. Even, it is appropriate to detect new no-well-known algorithms, too. Both targeted and blind steganalysis extract features in training and testing phase of the process as well. There are two approaches of blind steganalysis: first one defines cover and their appropriate stego images which contain as many steganographic algorithms as possible in the training phase of the system, while the other one uses only cover images to train a model. Such approach is known as *one-class learning*. In this method there is no necessity of any other training the model to detection of new no-well-known steganographic algorithms as well [2].

II. Blind Image Steganalysis

Model of blind steganalysis is shown in the Fig. 2.

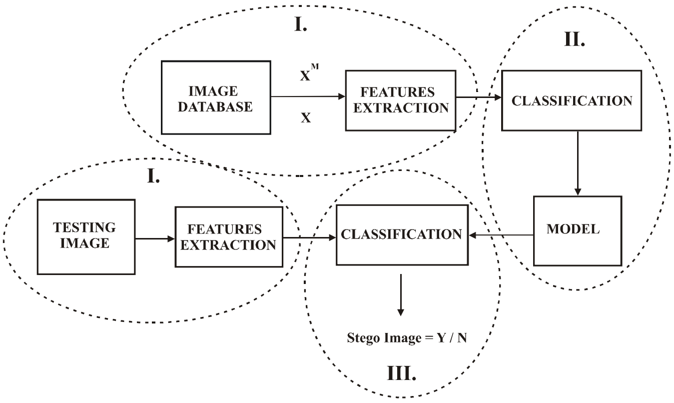


Fig. 2. System of the blind image steganalysis. (I) Phase of the features extraction. (II) Training phase. (III) Phase of the testing. Input image is tested by the trained model.

The first phase (marked as I.) is the phase of extraction the characteristic features from an image. All the images are from an image database that consists of cover images X and appropriate stego images X^M. The testing image is the image that we want to know whether it contains embedded message or not. Features are extracted in the next block with characteristic name. Phase II. is called training phase and it works with features obtained in the block mentioned before. Classifier marks extracted vectors of features as stego or cover according class where they belong to. There is a method of machine learning (more accurately supervised learning) utilized to train a model based on the input arguments. In the testing phase (III.) there is the model used to determine whether testing image contain secret message or not.

A. Features Extraction

Steganalysis aims to the distinction between cover and stego objects. In this order steganalysis searching for the statistical features that are affected by steganographic algorithm.

Basic methods of blind steganalysis can be divided by domain where the features are extracted. When process of extraction is performed in the spatial domain, *Binary similarity measures* (BSM) is considered. Steganalysis of the features obtained from wavelet domain is known as *Wavelet-based steganalysis* (WBS) and when features are extracted from the DCT domain (more accurately from JPEG domain), we discuss about the *Feature-based steganalysis* (FBS).

1) Binary similarity measures

There is a supposition that steganographic algorithms leave statistical artifacts between neighboring bit planes of an image in the spatial domain. It means that if LSB method is used, cover and stego images will be differ between seventh and eighth bit plain. In [3] are shown all the similarity measures chosen by Avcibas that are performed on cover and stego images. The results (extracted features) are subsequently utilized to construct a classifier. As Avcibas published, BSM has higher efficiency of detecting LSB method whereas Farid´s WBS is proved as superior when the steganographic algorithms F5 and Outguess are analyzed.

2) Wavelet-based steganalysis

WBS represents a different approach for the features extraction. Authors in [4] assert that most of steganalitic methods are focused on first-order statistics, i.e. histogram of DCT coefficients; however the simple measurement could keep them intact. That fact caused a proposition of the higher-order statistics which are extracted from the each sub-band after transformation of an image to the wavelet domain by QMF filters (Quadratic Mirror Filters). These statistics consist of mean, variance, skewness, and kurtosis. The second part of the identical statistics is calculated from an error obtained by linear prediction of the each sub-band´s coefficients.

3) Feature-based steganalysis

FBS, in some papers called as *Feature calibrated-based steganalysis*, calculates statistics from the both JPEG and spatial domain. The set of features chosen by Pevný and Fridrich was first appeared in [5]. The method uses a calibration to estimate the cover image [6].

Process of callibration is ilustrated in the Fig. 3. First, JPEG image J_1 is decoded into the spatial domain. Then is cropped by 4 pixels in each direction and re-encoded with the same quality factor q_f to obtain a callibrated image J_2. The features are extracted from an input (I_1) and callibrated image (I_2). Both sets are substracted each other to create the set of features entering a classifier ($I_3 = I_1 – I_2$).

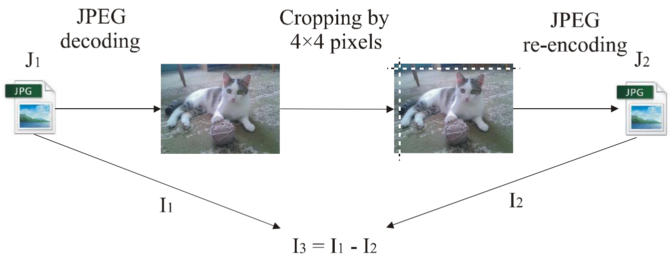


Fig. 3. Process of the calibration. Input image is decoded into spatial domain, cropped by 4 pixels in each direction and re-encoded to make a calibrated image.

Use of the callibration decreases a range of feature values. The consequence is reflected in significantly shorter time of the training phase.

B. Classification

Classification represents the main role of steganalitic system. It works in both training and testing phase of blind steganalytic system as well. Its role is to put a testing object to the appropriate class. This separation is performed by classifier.

1) Individual Classifiers

Among the most popular individual classifiers belong [7]:

- *Support Vector Machines* (SVM)
- *Fisher linear discriminant analysis* (FLD)
- *quadratic discriminant analysis* (QDA)
- *ordinary least-squares regression* (OLS)
- *Bayesian belief networks* (BBNs)
- *naive Bayes classifiers* (NBCs)

From the upper mentioned list, the most efficient are SVM and FLD.

a) Support Vector Machines

Support Vector Machines (SVM), proposed by Vapnik [8], is method of machine learning which is used to classify linear separated or nonlinear separated problems. Based on the input data, SVM computes the parameters of a separated hyperplane to classify data to the appropriate class. Problem of the training a model is to find an optimal border by which are cover and stego characteristic features divided. In the Fig. 4 is illustrated linear separated problem by SVM classifier.

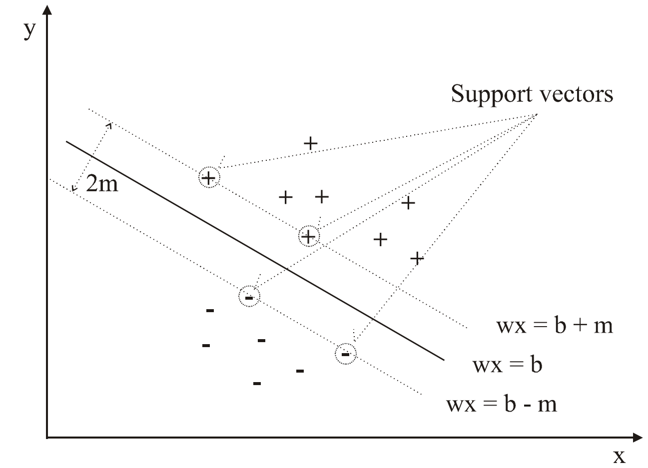


Fig. 4. Linear separated problem classified by SVM.

Optimal separated hyperplane is defined in (1) [9], where x is an input vector, w – a vector of weighting coefficients, and b – an offset.

$$wx = b \tag{1}$$

Hyperplane is situated in the middle of range 2m which is given by support vectors.

Case described above is determined for the linear separated problems. If the problem is nonlinear separated, input vector is transformed to the space with more dimensions. It is achieved using a kernel function (see Fig. 5) [10].

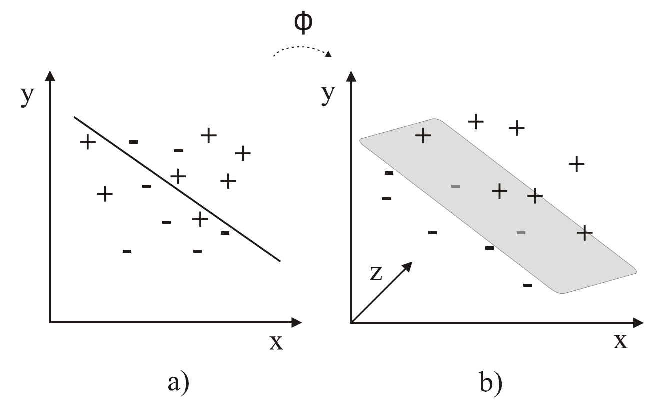


Fig. 5. (a) Nonlinear separated problem. (b) Transformation into the multidimensional space.

Now, the classifier searches for optimal separated plane in multidimensional space. Separated hyperplane in multi-dimensional space is defined as

$$w\Phi(x) = b, \tag{2}$$

where: $\Phi(x)$ is the transformation of vector x to multidimensional space by the kernel function.

b) Fisher linear discriminant analysis

Fisher linear discriminant analysis (FLD) [26] is a classic statistical method determined to classify linear separated and nonlinear separated problems. FLD represents a projection of the samples from n-dimensional space to a line in order to a better classification the input data.

In the Fig. 6a) is illustrated an ideal separation in two-dimensional space, whereas projection in Fig. 6b) represents a non-ideal separation.

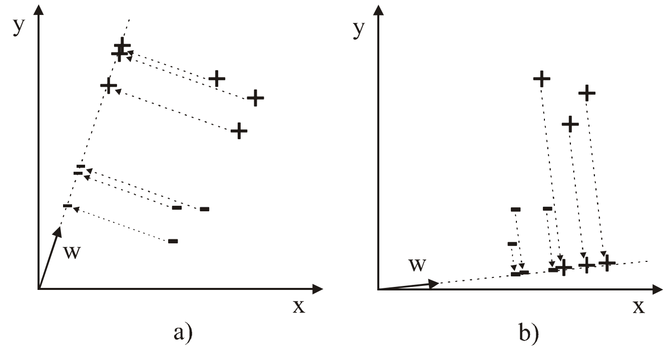


Fig. 6. (a) Ideal-separated projection of samples by FLD classifier. (b) Non-ideal-separated projection of the same samples.

Vector w defines a line (in general, hyperplane) which is utilized to definite classification the input data. In the case of nonlinear separated problem input samples are transformed into a multidimensional space by the kernel function to make linear separated problem as well as in the case of SVM.

2) Ensemble Classifiers

Another method of classification in steganalysis is based on joining more individual classifiers to the one functional aggregate. Classifier made using this technology is called Ensemble classifier. Some techniques of combined models are:

- *Bootstrap Aggregation*
- *Boosting*

As we can see in the Fig. 7, *Bootstrap aggregating* (*bagging*) is a technique where Ensemble classifier consists of many base learners B_L.

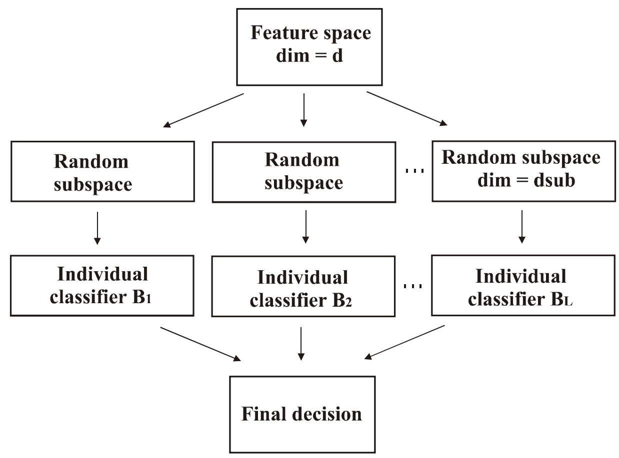


Fig. 7. Scheme of Ensemble classification. Random subspaces are obtained from feature space dim and randomly and uniformly divided into each d_{sub}.

Each base learner is an individual classifier trained with (uniformly) randomly selected subspace d_{sub} of the feature space dim of input images. In phase of the testing each base learner produces final decision whether testing subset of features belongs to the cover or stego class. Final decision is achieved by aggregating of each minority decision [11].

The other technique is *boosting*. Boosting combines many weak classifiers to make strong high-accurate one. The earliest boosting framework is AdaBoost. AdaBoost, in compare with bagging, trains the individual classifiers sequentially and each next one with those samples that have been difficult to classify by previous one. The final decision is achieved by majority decision of individual minority likewise in the bagging. This method gives high accurate and effective training time [11].

III. Conclusion and Proposal for The Next Steps of my Scientific Research

In this paper, there was performed short review of the blind image steganalysis. Steganography, as part of the wider area information hiding, is used to secret communication in order to avoid detection of a presence of the communication itself. This property of the steganography is attractive to criminals mainly what encourage to developing effective methods of the steganalysis. Blind image steganalysis method is most popular nowadays because extracts more statistical features in both spatial and transform domain for the detection of more than one steganographic tool. Blind techniques differ in a domain where the characteristic features are extracted. Next variance is in a use of the classifier. Ensemble classification is very effective tool. It brings primarily shorter time of training phase with comparable efficiency with individual classifiers.

In the next period of time I am going to focus on the sets of extracted parameters in DCT and DWT domain to find more characteristics which the different steganography tools changes. Moreover the method of steganography is able to do returnable censoring of an image, so I would like to aim on this using of that discipline, too.

Finally I want to thank to my supervisor Prof. Ing. Dušan Levický, CSc. for his useful advices and leadership.

References

[1] BUCHANAN, B.: “Security Tips”. Edinburgh: Napier University. Available on internet: <http://buchananweb.co.uk/security04.aspx>.
[2] COX, I. J. – et al.: “Digital Watermarking and Steganography”. 2nd ed., USA: Burlington: Morgan Kaufmann Publishing, 2008. ISBN: 978-0-12-372585-1
[3] AVCIBAS, I.; Memon, N.; Sankur, B., "Image steganalysis with binary similarity measures," Image Processing. 2002. Proceedings. 2002 International Conference on , vol.3, no., pp.645,648 vol.3, 24-28 June 2002.
[4] LYU, S. – FARID, H.: “Detecting hidden messages using higher-order statistics and support vector machines”, inProc. 5th Int. Workshop on Information Hiding 2002.
[5] PEVNÝ, T. – FRIDRICH, J.: “Merging Markov and DCT Features for Multi-Class JPEG Steganalysis”, In E. J. Delp and P. W. Wong, editors, Proceedings SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX, volume 6505, pages 3 1–3 4, San Jose, CA, Jan 29–Feb 1, 2007
[6] FRIDRICH, Jessica.: “Feature-based steganalysis for jpeg images and its implications for future design of steganographic schemes”, in Proc. 6th Information Hiding Workshop, Toronto 2004.
[7] BÖHME, Rainer: “Advanced Statistical Steganalysis”. Dresden: Springer 2010. ISBN 978-3-642-14312-0
[8] VAPNIK, V.: “The Nature of Statistical Learning Theory.” New York: Springer 1995.
[9] BENNETT, K. P. – BLUE, J. A.: “A Support Vector Machine Approach to Decision Trees,” Neural Net-works Proceedings, 1998. IEEE World Congress on Computational Intelligence. The 1998 IEEE Interna-tional Joint Conference on, vol. 3, May 1988, pp.2396,2401 vol.3, 4-9.
[10] DE BIE, T. – CHRISTIANINI, N.: “Kernel Methods for Exploratory Pattern Analysis: a Demonstration on Text Data.” Structural, Syntactic, and Statistical Pattern Reconstruction Lecture Notes in Computer Science, vol. 3138, 2004, pp. 16-29.
[11] KODOVSKÝ, J. – FRIDRICH, J. – HOLUB, V.: “Ensemble Classifiers for Steganalysis of Digital Media”. IEEE Transactions on Information Forensics and Security, Vol. 7, No. 2, pp. 432-444, April 2012.

Categorical model of the multiplicative fragment of predicate linear logic

[1]*Emília DEMETEROVÁ (3rd year),*
Supervisor: [2]*Valerie NOVITZKÁ*

[1,2]Dept. of Computers and Informatics, FEI TU of Košice, Slovak Republic

[1]emilia.demeterova@tuke.sk, [2]valerie.novitzka@tuke.sk

***Abstract*—Linear logic is able to describe real processes used in computer science. By its we can specify dynamics, non determinism, consecutive processes and important resources as memory and time on syntactic level. One can verify specified properties by its deduction system. By extending propositional linear logic with predicates and quantifiers the expression power of linear logic grows. In this paper we define predicate linear logic and construct a categorical model of predicate linear logic as a symmetric monoidal closed category.**

***Keywords*—Predicate linear logic, symmetric monoidal closed category.**

I. INTRODUCTION

Linear logic was introduced by Girard [1]. Linear logic is the only logic which is able to describe the dynamic of processes, internal or external non determinism, consecutive processes manage with resources on a syntactic level.

Propositional linear logic is often used for describing program systems [2], [3], their behavior [4] and its extension with modal operators enables the modeling of knowledge achievement [5]. The expressing power of propositional linear logic is insufficient for describing properties of some objects and relations between them. Predicates together with quantifiers provide predicate linear logic. Predicates expresses properties or relations of calculations. Quantifiers are able to specify a group of objects for which some feature or relation is valid.

We introduced types into linear logic because we need them to express predicate linear logic. We formulated the semantics of types by symmetric monoidal closed category [6] as

the *interpretation function for object* is

$$i : \mathscr{T} \to \mathscr{C}_{obj}$$

and it assigns an object in $\mathscr{C}_{obj}$ to every type from $\mathscr{T}$ as follows:

$$\begin{array}{lcl@{\qquad}lcl} i(\sigma) & = & [\![\sigma]\!] & i(\sigma \otimes \tau) & = & [\![\sigma]\!] \otimes [\![\tau]\!] \\ i(I) & = & [\![I]\!] & i(\sigma \multimap \tau) & = & Hom([\![\sigma]\!], [\![\tau]\!]) \end{array}$$

The *interpretation function for morphism*

$$j : \mathscr{TERM} \to \mathscr{C}_{morp}.$$

assigns a morphism in $\mathscr{C}_{morp}$ for every typed linear term from a set $\mathscr{TERM}$.

$$j : \tau \vdash t : \sigma \to [\![t]\!] : [\![\sigma]\!] \to [\![\tau]\!].$$

Semantics of the linear types is defined as a pair of functions

$$(i, j).$$

II. MULTIPLICATIVE FRAGMENT OF PREDICATE LINEAR LOGIC

The multiplicative fragment of predicate linear logic contains the multiplicative connectives with their neutral elements from the propositional linear logic. A linear formula φ has a form defined by the following BNF rule:

$$\begin{array}{rcl} \varphi & = & p \mid 1 \mid \bot \mid \varphi \otimes \varphi \mid \varphi \parr \varphi \mid \varphi \multimap \varphi \mid \varphi^{\bot} \mid !\varphi \mid ?\varphi \mid \\ & & \mid P(t_1, ..., t_n) \mid \forall x \varphi \mid \exists x \varphi \end{array}$$

The multiplicative fragment of predicate linear logic [7] is modeled in symmetric monoidal closed category.

III. CATEGORICAL MODEL OF THE MULTIPLICATIVE FRAGMENT OF PREDICATE LINEAR LOGIC

We constructed the semantics of linear type theory as symmetric monoidal closed category. Next we constructed the categorical model of the multiplicative fragment of predicate linear logic in symmetric monoidal closed category, too [8].

Elementary sentence is interpreted as a basic type, an object in category $\mathscr{C}$.

Neutral element

$$1 \equiv I$$

is interpreted as terminal object I of the category $\mathscr{C}$. The neutral element $\bot$ is dual to 1

$$\bot^{\bot} \equiv 1,$$

therefore from the properties of category $\bot$ is interpreted as initial object of $\mathscr{C}$.

For the interpretation of negation we use the following equivalence:

$$[\![\varphi^{\bot}]\!] \equiv [\![\varphi \multimap \bot]\!]$$

According to [9] we interpret every sequent

$$\varphi_1, ..., \varphi_n \vdash \psi$$

as morphism

$$[\![\varphi_1]\!] \otimes ... \otimes [\![\varphi_n]\!] \to [\![\psi]\!]$$

in $\mathscr{C}$, where $[\![\varphi]\!]$ expresses object, a representation of the formulae φ in $\mathscr{C}$.

Connectives are interpreted as morphisms in category $\mathscr{C}$ as follows:

$$\begin{array}{c} [\![\varphi]\!] \otimes [\![\psi]\!] \to [\![\varphi \otimes \psi]\!] \\ [\![\varphi]\!] \parr [\![\psi]\!] \to [\![\varphi \parr \psi]\!] \\ [\![\varphi]\!] \multimap [\![\psi]\!] \to [\![\psi^{\varphi}]\!] \end{array}$$

Because the symmetric monoidal closed category is cartesian closed category, the existence of objects $[\![\varphi \otimes \psi]\!]$, $[\![\varphi \invamp \psi]\!]$ and exponential object $[\![\psi^{\varphi}]\!]$ arises from its properties [10].

An unary predicate $P(t)$ is a property of the value of a term t of type σ. Because $[\![\sigma]\!]$ is an object in $\mathscr{C}$, the interpretation of the predicate $P(t)$

$$[\![P(t)]\!] \subseteq [\![\sigma]\!]$$

is a subset of the object $[\![\sigma]\!]$ in the $\mathscr{C}$.

The interpretation of the n-ary predicate $P(t_1, \ldots, t_n)$ is

$$[\![P(t_1, \ldots, t_n)]\!] \subseteq [\![\sigma]\!] \otimes \ldots \otimes [\![\tau]\!]$$

is a subset of the product of objects $[\![\sigma]\!] \otimes \ldots \otimes [\![\tau]\!]$ where a term t_1 has a type σ and a term t_n has a type τ.

We define the semantics of modal operators *of course !* and *why not?* and quantifiers by adjoint endofunctors. The definition of functors and endofunctors is cited in [11].

Adjunction means that there exists exact correspondence between morphisms $\varphi \rightarrow G(\psi)$ and $\psi \rightarrow F(\varphi)$, i.e. the Homsets

$$Hom(F(\varphi), \psi) \cong Hom(\varphi, G(\psi)). \tag{1}$$

are isomorphic. Adjunction can be illustrated also by the following commuting diagram.

$$\begin{tikzcd} \varphi \arrow[r, "F"] \arrow[d, dashed, "f"] & F(\varphi) \arrow[d, dashed, "g"] \\ G(\psi) & \psi \arrow[l, "G"'] \end{tikzcd}$$

The property (1) is useful in defining semantics of modal operator *of course !* as follows. Let F and G be a pair of adjoint endofunctors

$$F \dashv G$$

in $\mathscr{C}$. We define this modal operator as a composition

$$[\![!]\!] : G \circ F$$

such that for any object $[\![\varphi]\!]$

$$(G(F([\![\varphi]\!]))) = [\![!\varphi]\!]$$

it returns an object isomorphic with $[\![\varphi]\!]$, i.e. We can model unexhaustible resource $[\![\varphi]\!]$ by composition of adjoint functors as it is illustrated in figure 1.

The modal operator *why not ?* is dual to modal operator *of course !*

$$([\![?\varphi]\!])^{\perp} \equiv [\![!\varphi]\!]$$

To interpret quantifiers we use adjoint functors, too.

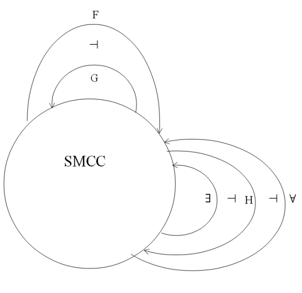


Fig. 1. Model of predicate linear logic

Let $[\![P(t)]\!] \subseteq [\![\sigma]\!]$ be an interpretation of the unary predicate symbol, where $[\![\sigma]\!]$ is an object in $\mathscr{C}$. We consider a variable $y : \tau$ and we construct a predicate $P(t, y)$ interpreted as

$$[\![P(t, y)]\!] \subseteq [\![\sigma]\!] \times [\![\tau]\!],$$

where y has no free occurrence in t. We construct an auxiliary endofunctor H as follows:

$$\mathrm{H} : \mathcal{P}([\![\sigma]\!]) \rightarrow \mathcal{P}([\![\sigma]\!] \times [\![\tau]\!]),$$

where $\mathcal{P}([\![\sigma]\!])$ is a power set over $[\![\sigma]\!]$.

We define a left adjoint functor of H that is interpretation of existential quantifier $\exists$

$$[\![\exists]\!] \dashv H$$

as

$$[\![\exists]\!] : \mathcal{P}([\![\sigma]\!] \times [\![\tau]\!]) \rightarrow \mathcal{P}([\![\sigma]\!]),$$

that for quantified formula $\exists y.P(t, y)$ returns a value of type τ (if it exists) satisfying predicate P(t,y):

$$[\![\exists y.P(t, y)]\!] = \{[\![t]\!] \in [\![\sigma]\!] |\ \text{exists a value in } [\![\tau]\!] \models [\![P(t, y)]\!]\}.$$

Because of duality between existential and universal quantifiers, we interpret universal quantifier as a right adjoint to the auxiliary functor H

$$H \dashv [\![\forall]\!].$$

Interpretation of quantifiers by adjoint functors is illustrated in the figure 1.

In the following text we explain how deduction rules and proofs can be interpreted by morphisms in our model. The identity rule

$$\frac{}{\varphi \vdash \varphi}\ (\mathrm{id})$$

is interpreted as identical morphism:

$$id_{[\![\varphi]\!]} : [\![\varphi]\!] \rightarrow [\![\varphi]\!].$$

The proofs

$$\frac{\begin{matrix}\tau_1 \\ \vdots\end{matrix}}{\Gamma \vdash \varphi, \Delta} \qquad \qquad \frac{\begin{matrix}\tau_2 \\ \vdots\end{matrix}}{\Gamma' \vdash \psi, \Delta'}$$

are interpreted as morphisms

$$f : [\![\Gamma]\!] \rightarrow [\![\varphi \invamp \Delta]\!] \qquad g : [\![\Gamma']\!] \rightarrow [\![\psi \invamp \Delta']\!]$$

in category $\mathscr{C}$.The proof of the multiplicative conjunction

$$\frac{\dfrac{\begin{matrix}\tau_1 \\ \vdots\end{matrix}}{\Gamma \vdash \varphi, \Delta} \quad \dfrac{\begin{matrix}\tau_2 \\ \vdots\end{matrix}}{\Gamma' \vdash \psi, \Delta'}}{\Gamma, \Gamma' \vdash \varphi \otimes \psi, \Delta, \Delta'}\ (\otimes_{\mathrm{R}})$$

is interpreted as morphism

$$[\![\Gamma]\!] \otimes [\![\Gamma']\!] \xrightarrow{\quad f \otimes g \quad} [\![\varphi \otimes \psi \invamp \Delta \invamp \Delta']\!]$$

in category $\mathscr{C}$.

Our model of predicate linear logic is constructed as a symmetrical monoidal closed category of types together with appropriate adjoint functors for modal operators and quantifiers. A proof of a formula is modeled as a finite path of category morphism.

IV. Conclusion

Linear logic is the most appropriate logical system for computer science because of its dynamic nature, expressing causality, non determinism and handling resources. In this paper we defined the multiplicative fragment of predicate linear logic and constructed categorical model based on symmetric monoidal closed category. In this category types are represented as objects and they enable direct connection with computing. We used adjoint endofunctors for modeling the modal operators expressing non exhaustibility of resouces and for modeling quantifiers. In our further research we will either specify contracts and dependencies between components in modeling component based systems, or in modeling observable behavior of such systems on the base of coalgebras.

References

[1] J.-Y. Girard, "Linear Logic," *Theoretical Computer Science*, vol. 50, pp. 1–102, 1987.

[2] D. Mihályi, V. Novitzká, and V. Slodičák, "Categorical models of logical systems in the mathematical theory of programming," *P. U. M. A*, vol. 17, no. 3–4, pp. 367–378, 2006.

[3] D. Mihályi and V. Novitzká, "Resource-oriented Programming Based on Linear Logic," *Acta Polytechnica Hungarica*, no. 4(2), pp. 157–166, 2007.

[4] D. Mihályi, V. Novitzká, P. Prazňák, and P. Popovec, "Network routing modelled by game semantics," *Studia Universitatis Babes-Bolyai, Informatica*, vol. 57, no. 4, pp. 19–29, 2012.

[5] D. Mihályi and V. Novitzká, "Towards the knowledge in coalgebraic model of IDS," *Computing and Informatics*, vol. 33, no. 1, pp. 61–78, 2014.

[6] S. Abramsky, "Computational interpretations of linear logic," *Technical Report 90/20, Department of Computing, Imperial College*, pp. 1–15, 1990.

[7] D. Mihályi and V. Novitzká, "What about Linear Logic in Computer Science?" *Acta Polytechnica Hungarica*, vol. 10, no. 4, 2013.

[8] E. Demeterová, D. Mihályi, and V. Novitzká, "A categorical model of predicate linear logic," *Journal of Applied Mathematics and Computational Mechanics, Scientific Research of the Institute of Mathematics and Computer Science*, vol. 14, 2015.

[9] P.-A. Melliès, "Categorical Semantics of Linear Logic," *Panoramas et Syntheses 27, Citeseer*, 2009.

[10] M. Barr and C. Wells, "Category theory for computing science," *Prentice Hall International Ltd.*, 1990.

[11] V. Novitzká and V. Slodičák, *Kategorické štruktúry a ich aplikácie v informatike*. Equilibria, 978-80-89284-67-2, 2010.

Code Obfuscation Problem

[1]Ján Hurtuk (2nd year)

Supervisor: [2]Liberios Vokorokos

[1,2] Dept. of Electronics of Computers and Informatics, FEI TU of Košice, Slovak Republic
jan.hurtuk@tuke.sk, liberios.vokorokos@tuke.sk

Abstract— **Nowadays, in world of modern technologies and technology society, information can be the most valuable commodity on the market and so gathering it, buy or sell it becomes the most important part of doing business. Because of the prizing, the more sophisticated ways of gathering and processing of the information are still being developed, various kinds of more sophisticated and faster systems are being created and more complex technologies are being made, which clearly opens various paths for various criminal menaces. To be able to react appropriate and fast on this threads and protect valuable property of its own, to protect the data and be able to find an appropriate solution for protecting them and for preventing them from damage or any other misusing, man should be aware of the processes behind the whole cycle and be able to define the whole menace scope given to the general situation and scenario.**

Keywords- obfuscation, garbage, permutation, structure

I. Introduction to Malware Problem

There exists many ways how to look at and how to understand term Malware and malicious software. Malware or malicious software is code sequence used to corrupt or sabotage computer operations, operations on specific sensitive and vulnerable parts of computer systems or various kinds of devices or to get access to sensitive information or gain sensitive data in order to achieve using rights to private devices or computer systems. Early states of malware like software were meant as prank to do simple damage, which was easy to detect and remove or as learning tool for gaining principles of first simple computer systems working. Nowadays, mainly because of the financial sector policy and virtual money transfers and sensitive easy to money change information floating, malware became more sophisticated, aiming the most sensitive spots of the whole process. Tons of new concepts of already existing malware versions of are released in daily manner and the brand new malware families are still developed. In order to detect and categorize them and to be able to do prevention, sophisticated detection and analyzing process is needed. It is although limited by various code hiding techniques as obfuscation, code encryption and using polymorphic concept [1].

II. Code Obfuscation Techniques

From a practical side of the problem, code obfuscation technique ground lies in making the desired code sequence as “unintelligible” as possible, which means hiding its signature. This kind of practice has come to use lately with the widespread of high-level programming languages such as .net and java. As far as higher-level programming languages are in discuss, the byte code format produced at the compilation time, still contains the whole code sequence core information. Because of that, it is possible to trace back the original source code sequence and additionally the underlying algorithms as well. Code obfuscation is made for thwarting decompilation process in order to protect the original code meaning and source. It is important to notice that code sequence obfuscation can be operational on various layers of the code sequence operational life cycle, the program data and control flow. A detailed taxonomy of obfuscation techniques can be found in [7]. Below are mentioned the usual obfuscation approaches that are particularly used by metamorphic viruses for example:

- ***data flow obfuscation*** (instruction substitution, instruction permutation, *dead code* or called *garbage code* insertion, variable substitution);
- ***control flow obfuscation*** (changing the control flow).

A. Data flow obfuscation

Instruction substitution the core of it lies in replacing a chosen instruction block with another instruction block or blocks while keeping the same code semantics.

Instruction permutation the core of it lies in modification process of the instruction execution order in the given sequence while keeping the same semantics.

Variable substitution the core of it lies in either changing the used registers or in modifying the extent of the used variables by using local or global variables instead of registers at the binary code level.

Dead code insertion this approach is more sophisticated and using it requires more knowledge about the the whole given changed entity. The core of it lies in introducing and additionally adding so called useless pieces of code into the sequence code entity. In a similar approach, it can be structured from more complex instruction sequences that will never be executed. Table 1 shows some rules, which can be applied as “dead code insertion”. The left column shows instructions which are semantically equivalent to the `NOP` (No Operation) instruction. The exact meaning of the given rules is shown in the right column.

Rules	Meaning
`ADD Reg,0`	`Reg ← Reg + 0`
`MOV Reg,Reg`	`Reg ← Reg`
`OR Reg,0`	`Reg ← Reg | 0`
`AND Reg,-1`	`Reg ← Reg & -1`

Table 1 Dead code approach Rules Meaning

B. The control flow obfuscation

Its principles lie in modifying and changing the execution flow of the code sequence by inserting conditional or unconditional branching sets of instructions, while preserving the program result. Figure 1 shows an example of such modification approach by inserting unconditional branching instructions. The original sequence made of sequential instructions can be replaced by the two other sequences whose execution is modified to be executed non-sequentially. In both cases the whole semantics of the entity remains the same.

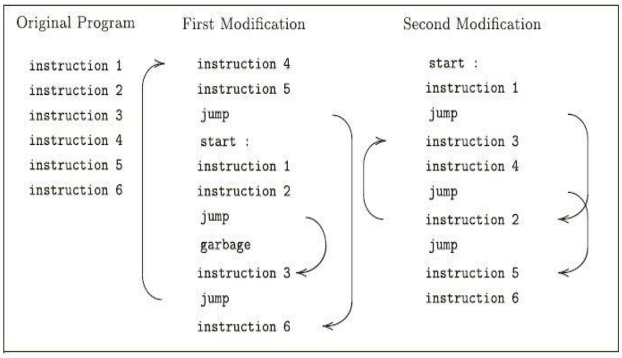


Figure 1 Control flow obfuscation example

C. The main abfuscation development approaches

According to Collberg s work [3] as reference it is important to define a metric commonly used to evaluate the efficiency of obfuscators. It can be defined as follows:

- ***the potential*** which serves to evaluate the understanding of the obfuscated code complexity during analysis run by human (human-driven analysis);
- ***the resilience*** which is meant as measure of the of the inverse operation (deobfuscation) complexity by using of the automatic tools;
- ***the cost***, which defines the price it needs to be paid in terms of computing time and memory space required for the analysis[9].

It is clearly visible that in terms of the malicious software, the main field of interest would be the resilience. In matter of fact, resilience can informally be addressed as the difficulty of detection the malicious software code sequence. The main approach study presents the using of the opaque predicates in terms of increasing the resilience of an obfuscated program. Predicate *P* can be considered as opaque if it has a property *q* known to the obfuscator algorithm but which is hard for a deobfuscator to deduce and reproduce.

D. Transformation potency

When define potency of obfuscating transformation, the first measure must be defined, to equalize, the readability of code sequence P1 in comparison to code sequence P2. Measuring program sequence complexity is quite old problem, which have been investigated through [3], in software theory. Many different systems and measuring techniques have been created. Some of them though cannot be used due to their using of high level languages.

Measure	Description
Length of program	number of instructions + number of arguments
Nesting level	number of nested conditions
Data flow	number of references to local variables

Table 2 Measuring techniques of program complexity

Definition 1 *For given complexity measure, E(P) potency of obfuscation transformation T in relation to obfuscated program sequence P, Π (T,P) is defined as:*

$$\Pi(T,P) = \frac{E(T(P))}{E(P)} - 1$$

It is given, that T means strong obfuscating transformation when *Π (T,P)>>0* for selected group of complexity measuring techniques. When considering machine code measuring techniques from mentioned in Table 2, they can be defined as:

1. Length measuring E_L, which specifies concrete length of code sequence P, containing N instructions, which gives it the value based on this condition:

$$E_L(T,P) = \sum_{i=1}^{N} c_i \quad \text{dla} \quad c_i = \begin{cases} 0 & \text{when instruction } i \text{ has no arguments} \\ 0.5 & \text{when instruction } i \text{ has one argument} \\ 1 & \text{when instruction } i \text{ has two arguments} \end{cases}$$

2. Depth measuring E_d, has typical value of integer, which describes the nesting level of conditional branches. It can be computed by different ways similar as this algorithm which core is defined as follows:
   If instruction can is defined as conditional branch, find its jump depth, proceed until the end of the examined code sequence and then return the maximum of the founded depths.
3. Flow measuring E_f has typical rational values and counts and describes the average number of references to local memory in program basic code block. The basic block can be described as code sequence between the two nodes of control flow graph:

$$E_F(T,P) = \frac{1}{M}\sum_{i=1}^{M} a_i$$

where M is defined as the basic number of code sequence blocks and a_i is defined as number of references in local memory in block i.

All three measuring values should be influenced by the obfuscation process in the same way. The final value can be computed as follows [11,12,113,14]:

$$\Pi_A(T, P) = \frac{1}{3}\left(\frac{E_L(T(P))}{E_L(P)} + \frac{E_D(T(P))}{E_D(P)} + \frac{E_F(T(P))}{E_F(P)}\right) - 1$$

E. Obfuscation limits brief overview

It is a very well-known and can be proven that the bulletproof obfuscation is impossible to achieve. The semantics of the given code sequence cannot be perfectly hidden [10,11]. In other words, however good obfuscation technique or combination of them, given code semantics can always be recovered. But, the semantics extraction cannot be achieved by automated process (Rice theorem [12]). Because of that, some detection experimental approaches try to built an model of the created program by using code compiling optimizations [13,14]. This creation process is performed in two basic steps:

- control flow graph of the examined sequence is built. This graph is a model of the possible program execution flows,
- data flows are analyzed and simplified. In case of finding the more obstacles in the process, it is repeated simplify it.

III. Conclusion

Information about the techniques and processes described in previous chapters are becoming more and more useful because of the growing cyber criminality trends and though the necessity of well educated specialists, who can successfully understand the security issues and adopt measures to stop or even prevent the given systems from being crashed and information being corrupt or even stolen and misused.

Acknowledgment

This work was supported by the Slovak Research and Development Agency under the contract No. APVV-0008-10 and project KEGA 008TUKE-4/2013: Microlearning environment for education of information security specialists.

References

[1] Bartosz Zoltak (2004), VMPC One-Way Function and Stream Cipher", *Fast Software Encryption, FSE 2004*, Lecture Notes in Computer Science 3017, Springer-Verlag, pp. 210–225, doi:10.1007/978-3-540-25937-4_14, ISBN 3-540-22171-9

[2] Souradyuti Paul, Bart Preneel. "Analysis of Non-fortuitous Predictive States of the RC4 Keystream Generator" (PDF). Indocrypt 2003. pp. 52 – 67.

[3] Collberg, C., Thomborson, C., Low, D.: A taxonomy of obfuscating transformations. Technical Report 148, Univerity ofAuckland, New Zealand (1997)

[4] Peter Szor, Zmist opportunities, Virus Bulletin Conference, pp. 6-7,March 2001.

[5] Issues in the Passive Approach of Network Traffic Monitoring / Adrián Pekár ... [et al.] - 2013. In: INES 2013 : IEEE 17th International Conference on Intelligent Engineering Systems : proceedings : June 19-21, 2013, Costa Rica. - Budapest : IEEE, 2013 P. 327-332. - ISBN 978-1-4799-0828-8

[6] Beaucamps, P., Filiol, E.: On the possibility of practically obfuscating programs—towards a unifed perspective of code protection. J. Comput. Virol. (WTCV'06 Special Issue, Bonfante, G., Marion, J.-Y. eds), **2**(4) (2006)

[7] Barak, B., Goldreich, O., Impagliazzo, R., Rudich, S., Sahai, A.,Vadhan, S., Yang, K.: On the (im)possibility of obfuscating programs. In: Crypto '01, LNCS No.2139, pp. 1–18 (2001)

[8] Rice, H.G.: Classes of recurively enumerable sets and their decision problems. Transactions of the American Mathematical Society, pp. 358–366 (1953)

[9] Aho, A.V., Sethi, R., Ullman, J.D.: Compilers: Principles, Techniques and, Tools. Addison-Wesley, Reading, MA (1986)

[10] A Multicore Architecture Focused on Accelerating Computer Vision Computations / Liberios Vokorokos ... [et al.] - 2013. In: Acta Polytechnica Hungarica. Vol. 10, no. 5 (2013), p. 29-43. - ISSN 1785-8860

[11] Gregory Wroblewski :General method of program code obfuscation Wroclaw 2002, available on: https://tuts4you.com/download.php?view.2908

[12] M.H. Halstead, Elements of software science, Elsevier North-Holland, 1977

[13] Liberios Vokorokos, Anton Baláž, Branislav Madoš Intrusion detection architecture utilizing graphics processors .In: Acta Informatica Pragensia. Vol. 1, no. 1 (2012), p. 50-59. - ISSN 1805-4951

[14] Liberios Vokorokos, Anton Baláž, Branislav Madoš, Ján Radušovský Cloud file system In: INES 2013 : IEEE 17th International Conference on Intelligent Engineering Systems : proceedings : June 19-21, 2013, Costa Rica. - Piscataway : IEEE, 2013 P. 311-315. - ISBN 978-1-4799-0828-8

Component design for mobile application in smart home

[1]Ján TIMÁR, [2]Michal KOVALČÍK

[1,2] Dept. of Computers and Informatics, Technical University of Košice, Letná 9, 042 00 Košice, Slovak Republic

[1]jan.timar@student.tuke.sk, [2]michal.kovalcik@tuke.sk

***Abstract*— New homes and appliances often contains functions for interconnection with others appliances. With creating infrastructure connected appliances is necessary for easy and comfortable using unify user control interface. This article is dedicated to the unification design for all appliances in homes with mobile app. Mobile app brings more ways for controle smart homes.**

***Keywords*—mobile app, design, iOS, iPhone, smart home, arduino.**

I. INTRODUCTION

Recently lot of devices changed in smart. It is devices with own operating system and various functions as remote control. Homes with smart devices called smart home. Smart home is complicated system. Depending on how it is implemented, in publication [1] Richard Harpper divided households into five groups: Homes which contain intelligent objects, Homes which contain intelligent, communicating objects, Connected homes, Learning homes, Attentive homes. In addition to the first two group are all connected devices in the home connected to network. By adding a control unit such as the Arduino, used in the project [2].

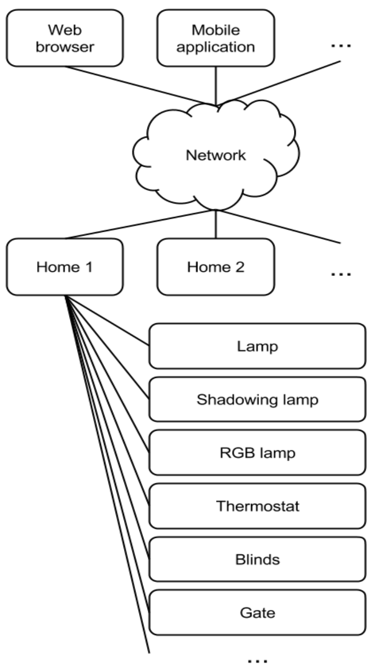


Fig. 1. Schematic description of controlling smart homes connected to network.

Home is easy connect to Internet and then controlled with any connected device[Fig. 1.]. With web server used in project [3.] is easy control home from any PC. Basic smartphone ability is connect to network either through a wireless connection or mobile network. With this ability is easy controlled all devices in home from one device in pocket. In application stores are many apps of same type. The user decides between them on the basis of their design and ease of use.

Fig. 2. Smart home devices in list.

II. SMART DEVICES

The user will learn most easily controlled application if already with a similar design have ever met in the past. For this reason is important standardize the design of appliances for control. User does not need to learn to control appliance separately. To system can be add new devices without change on client app. For this reason was appliances in smart home organized into groups by how many states can take. We created Two stateful devices group, More stateful devices group and Devices with floating point stateful group. All devices are displayed in a simple list [Fig 2.], as it is assumed that each user has experiences with list component. Devices are split to group by rooms for better user orientation. For more better way how find devices which can be complicated if app will be use in company with lot of connected devices can be use search bar. When search bar is not empty, in list will be shown appliances only with suitable name or type.

A. *Two stateful devices*

In two stateful group is appliances usually with on/off state or open/closed state. For control can be use button or two stateful segment. This states are represented in Arduino with bit values according to logical attribution. The appliance has value set on 1, meaning open or turn on. This can be showed with white or green background color on the line in the table. The value of 0, meaning turn off, locked is showed with background with transparent color or with a certain degree of translucency. White backlight can be replaced by red and green backlighting. For standardize user interface is best way use universal meaning colors like white for turned on and transparent for turned off.

B. *More stateful devices*

Arduino control unit can read and write values on analog and digital pins. Digital pin can have two states 1 and 0. On analog pins can be set value between 0 and 255. Devices with more as two values is harder standardize. Most more stateful devices is shadowing lamp, RGB lamp and devices worked with intermediate state such as blinds. RGB lamp is composed of three successive values in a row. Take values from 0:0:0 to 255:255:255 where first three digit affect red color, next three digit green color and last three digit affect blue color. When color change, values is sended in a row to control unit. To control RGB lamps was elected special component [Fig. 3. left] showing all colors. User can with one touch in the color spectrum easily select color. For clean design is not necessary more components, actual selected RGB color is set as background color. When user changed color, component background color is change in simple animation, for notify user about succesful action. Shadowing lamp has values from 0 to 255. Is control like in RGB lamp where spectrum enable colors is changed for suitable colors for shadowing lamp. Other components like blinds was easily created with added next text label for show actual state. This intermediate state like when blinds is moving to other state is only for some short time and don't need own component. Is necessary only for information user about some action is happening with appliance. In this situation can be text label replaced by image represent state. This can not be universal in some special case like when we create new appliance. We will be need update images database on mobile app.

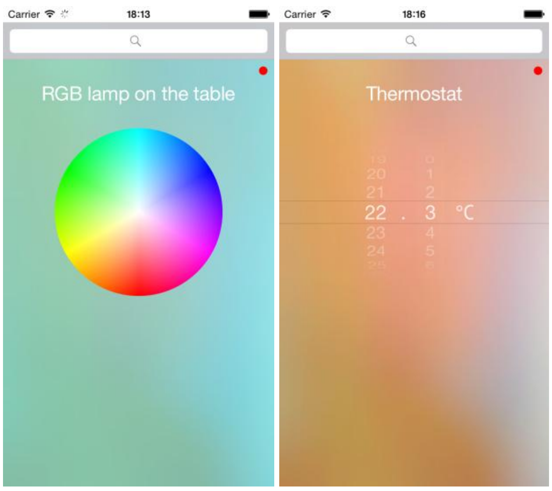


Fig. 3. Component for control RGB lamp (left) and component for control devices with floating point (right)

C. *Devices with floating point values*

Arduino device can receive values with floating point via Post communication. This values are used in devices like thermostat. In conventional thermostat user normally set new value with two buttons for increase and decrease value with constant step. When user wants to increase the temperature must click on buttons until the temperature on the thermostat does not adjust. Mobile devices with touch screens, allows user change value of simply with swipe finger up or down on screen. With split screen for two parts for control values before the comma and decimal value [Fig 3. right] can user change value easily, more precisely and quickly. With this way of thermostat is user possible enter value with one or two swipes across the screen. For example user with conventional thermostat with step 0,5°C must click 13 times for change value from 16,5°C to 23.0°C. Next devices use floating point is alarm with settable sensitivity or humidifier.

III. Conclusion

With the development of new smart appliances will make our home more and more smart. Design and controlling this devices will be more important factor for choose which will be use not only for smart appliances but too for way how we will be choose for the control. Standardize, easily control and clean design is best way how attract user for choose mobile app for use. With standardized devices we are able create API for create and add devices to big existing systems without many or any changes.

Acknowledgment

Paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDFF. We support research activities in Slovakia/This project is being co-financed by the European Union.

References

[1] HARPER, R.: Inside the Smart Home: Softcover 2003. 34-35.pp. : ISBN-13: 978-1852336882

[2] KURILEC, M. - KOVALČÍK, M. - FECIĽAK, P. - JAKAB, F.: Control module for smart house based on Arduino and its modules: Electrical Engineering and Informatics 5 : Proceedings of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice. - Košice : TU, 2014 S. 602-605. - ISBN 978-80-553-1704-5

[3] HUSÁR, M. - KOVALČÍK, M. - FECIĽAK, P. - JAKAB, F.: Web application of smart home: Electrical Engineering and Informatics 5 : Proceedings of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice. - Košice : TU, 2014 S. 597-601. - ISBN 978-80-553-1704-5

Concern Annotations in Program Comprehension

[1]*Milan NOSÁL' (4th year),*
Supervisor: [2]Jaroslav PORUBÄN

[1,2]Dept. of Computers and Informatics, FEEI TU of Košice, Slovak Republic

[1]milan.nosal@tuke.sk, [2]jaroslav.poruban@tuke.sk

***Abstract*—This paper presents a proposal of using source code annotations to record design decisions and explicit semantic properties directly in the code. The proposal builds upon several works that use design and semantic annotations (together concern annotations), however, it focuses solely on their usage for the purposes of program comprehension. We introduce the problem and shortly discuss our experimental study that provides evidence that concern annotations can be reused between developers, thus providing basis for using concern annotations for program comprehension.**

***Keywords*—Source code annotations, concern annotations, program comprehension, experimental study.**

I. Introduction

Since the beginning of the computers' era the programming process evolved to software engineering due to the growing complexity of software systems. Multiple different research fields have risen to support software development – software development processes, programming languages, programming tool support, etc. One of the fields that were result of the software system complexity growth is the *program comprehension* field.

An important phase of software development cycle is maintenance and evolution. In this phase there may be many changes in teams and the code written by one developer has to be comprehended and updated by another. In the development process, developer communicates with customers to get all the requirements for the software system. These requirements define the problem domain. But when it comes to implementation of the system, there is an abstraction gap between problem and solution domain [1]. Program comprehension research field aims to aid the process of code comprehension – the process of recreating the mental model describing the program's source code and mapping the problem and the solution domain.

II. Explicit Documentation Support

In our work we focus on explicit documentation support. The comprehension problem is many times elevated by the lack of communication of problem understanding. During the implementation phase programmers create their own mental model of the problem they solve and they apply it in the implementation. Unfortunately, this mental model is many times lost, or inadequate, because programmers find writing internal documentation (internal documentation explains the source code) too demanding with very low return value. But the internal documentation (comments) plays an important role in program comprehension.

We examine usage of annotations for recording design decisions (design annotations) and code semantic properties (semantic annotations) explicitly in the source code. These annotations will be called concern annotations in the rest of the paper (because they record concerns of the code). The goal of the work is to examine whether concern annotations can be beneficial for the process of program comprehension.

III. Program Comprehension and Annotations

My work is inspired by works that use source code annotations (attribute-oriented programming) to record decisions in the source code. Existing research in comprehension annotations uses annotations primarily as a configuration input for various tools. E.g., Hedin [2] uses annotation-like language extensions to record semantic concerns of the code that are later used by a recommendation system to detect possible design pattern instances. Sabo et al. [3] uses design decisions expressed by annotations to preserve design patterns in the code. Annotations are used by a tool to check whether the recorded design pattern instances did not break. Again, the design decisions and semantic properties recorded by annotations are mainly used by tools for some automation. The same applies to the semantic annotations dicussed by Cachopo [4] or by Kiczales et al. [5] in context of aspect-oriented programming (AOP). They too focus rather on annotation processing for the purposes of AOP configuration than on annotations' potential for program comprehension.

Two main comprehension problems that can be addressed by concern annotations and annotations-based tool support is the problem of feature/concept location and the problem of comprehension of relations between program elements. *Feature (or concept) location problem* [6] is a problem of locating a feature or a concept implementation in the source code. As a result of crosscutting concerns, a concept/feature implementation can be scattered in code [7]. If the given feature or concept implementation is logged in the code by concern annotations, its location can be easily found by finding all usages of the given annotation. The *problem of comprehension of relations between program elements* [8] is a problem of understanding (and locating) which program elements are related to the one currently inspected. Providing a fast way to locate and navigate to all related program elements of a given program element could make developers' work more effective [9]. A concern annotation expresses a relation between a set of program elements that are annotated by it.

We performed a study in which we hypothesized that *concern annotations can be reused between people*. That means that two different people could share the same mental model, at least partially. Thus concern annotations created and considered proper by one person would be considered proper and significant by someone else, too. To examine this phenomenon we designed a study in which participants annotated a project with concern annotations. Afterwards we inspected the degree of mutual coverage (agreement) of annotations describing the same concern. Accepting this hypothesis is a precondition for further research considering concern annotations. Without possibility of their reuse they cannot be anyhow beneficial for program comprehension.

The study was performed with 7 participants. All of them have master of science degree in computer science and have experience with programming in Java. The code base under inspection was the source code of the *EasyNotes* project[1]. All the participants started with a clean (not annotated) source base and were told to use annotations to comment the code. They did not cooperate nor discussed their mental models. After performing the study we have inspected the code and we found 110 annotation types for concerns total, and 598 concern annotations annotating the program (5.44 annotations for an annotation type in average). To examine the agreement between participants, we identified separate concerns expressed by annotations and analysed how often did multiple developers recognized the same concern as significant. To understand the agreement better, we defined *effective agreement*, where effective agreement is a percentage of effective annotations from all annotations in all considered projects. The average effective agreement in the study was 17.48%. When we did not consider the concerns that were not shared, we got 36.55%. *The higher the effective agreement, the better is the concerns reuse, since not only the same concern was recognized, but it was also mapped to the same program elements in the source code*.

Annotating study provided data that support our hypothesis that the *concern annotations can be reused between people*. Of course, not each concern is interesting to all the people. However, considering that 26 of 46 concerns were recognized by at least 2 participants, we can conclude that concern annotations can be reused.

IV. Concern Annotations Support by Metaprogramming

To promote usage of concern annotations to document the source code we proposed two approaches based on metaprogramming that use concern annotations to aid program comprehension. First approach provides concern-oriented projections of the source code. The basic idea behind concern-oriented projections is to provide editable views that present together program elements manifesting some common property (design decision, semantic property, etc.). Early proposal of the projections was published in [10]. In [11] we present a prototype implementation of projections and we also discuss different types of projections depending on how the tool identifies concerns of the code. Concern-oriented source code projections challenge standard file-based view of the source code.

The second approach uses a correspondence between documentation phrases and concern annotations. A documentation phrase is a set of documentation fragments that all document the same semantic or design property that is shared by the documented program elements. The same way a concern annotation annotates program elements that share common semantic property or design decision. The proposal of annotation-based documentation phrases was published in [12]. A more detailed presentation along with a case study was published in [13].

V. Conclusion and Future Work

Presented work is a follow-up of works in field of design annotations and semantic annotations. Using a experimental study we examined their potential for the purposes of program comprehension. In section IV we provided an overview of our works that use metaprogramming techniques (IDE modification and generative programming) and concern annotations to aid program comprehension.

Acknowledgment

This work was supported by VEGA Grant No. 1/0341/13 "Principles and Methods of Automated Abstraction of Computer Languages and Software Development Based on the Semantic Enrichment Caused by Communication".

References

[1] J. Greenfield and K. Short, “Software Factories: Assembling Applications with Patterns, Models, Frameworks and Tools,” in *Companion of the 18th Annual ACM SIGPLAN Conference on Object-oriented Programming, Systems, Languages, and Applications*, ser. OOPSLA ’03. New York, NY, USA: ACM, 2003, pp. 16–27.

[2] G. Hedin, “Language Support for Design Patterns Using Attribute Extension,” in *Object-Oriented Technology*, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 1998, vol. 1357, pp. 137–140.

[3] M. Sabo and J. Porubän, “Preserving Design Patterns using Source Code Annotations,” *Journal of Computer Science and Control Systems*, vol. 2, no. 1, pp. 53–56, 2009.

[4] J. a. Cachopo, “Separation of concerns through semantic annotations,” in *Companion of the 17th annual ACM SIGPLAN conference on Object-oriented programming, systems, languages, and applications*, ser. OOPSLA ’02. New York, NY, USA: ACM, 2002, pp. 2–3.

[5] G. Kiczales and M. Mezini, “Separation of Concerns with Procedures, Annotations, Advice and Pointcuts,” in *Proceedings of the 19th European Conference on Object-Oriented Programming*, ser. ECOOP’05. Berlin, Heidelberg: Springer-Verlag, 2005, pp. 195–213.

[6] V. Rajlich and N. Wilde, “The Role of Concepts in Program Comprehension,” in *Proceedings of the 10th International Workshop on Program Comprehension*, ser. IWPC ’02. Washington, DC, USA: IEEE Computer Society, 2002, pp. 271–.

[7] M. Petrenko, V. Rajlich, and R. Vanciu, “Partial Domain Comprehension in Software Evolution and Maintenance,” in *Proceedings of the 16th IEEE International Conference on Program Comprehension*, ser. ICPC 2008, June 2008, pp. 13–22.

[8] C. Parnin and C. Gorg, “Building Usage Contexts During Program Comprehension,” in *Proceedings of the 2006 IEEE 14th International Conference on Program Comprehension*, ser. ICPC 2006, 2006, pp. 13–22.

[9] D. Ratiu and L. Heinemann, “Utilizing Web Search Engines for Program Analysis,” in *Proceedings of the 2010 IEEE 18th International Conference on Program Comprehension*, ser. ICPC ’10. Washington, DC, USA: IEEE Computer Society, 2010, pp. 94–103.

[10] M. Nosáľ, J. Porubän, and M. Nosáľ, “Concern-oriented source code projections,” in *Proceedings of the 2013 Federated Conference on Computer Science and Information Systems*, ser. FedCSIS 2013, Sept 2013, pp. 1541–1544.

[11] J. Porubän and M. Nosáľ, “Leveraging Program Comprehension with Concern-oriented Source Code Projections,” in *Proceedings of the 3rd Symposium on Languages, Applications and Technologies*, ser. OpenAccess Series in Informatics (OASIcs), vol. 38. Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2014, pp. 35–50.

[12] M. Nosáľ and J. Porubän, “Software documentation through source code annotations,” in *Informatics 2013 : Proceedings of the Twelfth International Conference*, 2013, pp. 180–185.

[13] ——, “Reusable software documentation with phrase annotations,” *Central European Journal of Computer Science*, vol. 4, no. 4, pp. 242–258, 2014.

[1] https://code.google.com/p/easy-notes/, revision **r6**

Content creation for the virtual reality world

[1]Dušan JANOVSKÝ (2st year)
Supervisor: [2]Branislav SOBOTA

[1,2]Dept. of Computers and Informatics, FEI TU of Košice, Slovak Republic

[1]janovsky.dusan@gmail.com, [2]branislav.sobota@tuke.sk

***Abstract*— Computer graphics and virtual reality has become a part of the twenty-first century. Slowly but surely it found its way into almost every aspect of our everyday life. Graphics editor is special software designed for creating 2D and 3D graphic content. The objective of each software designer is to create the most natural and the most intuitive interface without restrictions in functionality. This paper aims to compare geometry representations and modeling techniques of graphics editors which are often use for content creation for virtual reality worlds.**

***Keywords* – content creation, virtual world, modeling techniques, model creation, geometry representation.**

I. Introduction

The methods for creating content of virtual world are different processes of developing a mathematical representation of any three-dimensional (3D) object or rather its surface. The final products are call 3D models and because they are based on math the computer can very easy display, work, manipulate and transforms these objects. Trend in increasing computing power resulted in huge development of 3D computer graphics.

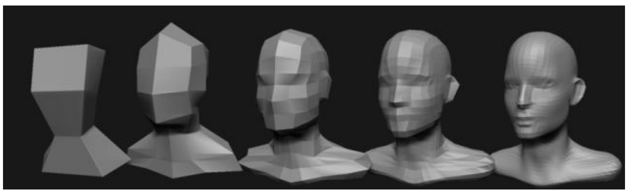
II. Representations of Geometry

Important part in developing the content of the virtual world is to understand the different types of representations of geometry and modeling techniques to help you select one which additionally satisfies the initial conditions for attaining the desired results. They are something like languages for describing geometry in way that computers can handle.

The point cloud represents a set of data points in some coordinate system. Point clouds may be created by 3D scanners and those clouds represent the set of points that the device has measured [1]. Works aimed on point cloud processing are [1][2][3]. Range imaging is the name for a collection of techniques that are used to produce a 2D image showing the distance to points in a scene from a specific point, normally associated with some type of sensor device [5][6]. The resulting image, the range image, has pixel values that correspond to the distance. Range imaging VLSI sensor which gives a range image of a scene with high robustness and high accuracy at a video frame rate is described in [4]. Polygon soup is a group of unorganized triangles, with generally no relationship whatsoever are a geometry storage format in a 3D modeling package, such as 3D Studio Max, Blender, or Maya. Floating polygon soup for applications like 3DTV and FTV (Free Viewpoint Television) is presented in [7]. Fast algorithm to approximate the swept volume (SV) boundary of arbitrary polygon soup models is presented in [8].

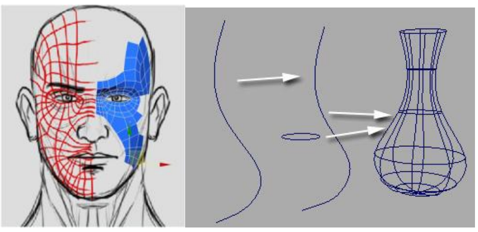
Binary space partitioning (BSP) is a method for recursively subdividing a space into sets by planes. This subdivision gives rise to a representation of objects within the space by means of a tree data structure. BSP trees have proven their utility in 3D modeling, graphics and image processing, and their tree structure allows efficient algorithms to be developed that are compact and numerically robust [9]. A voxel represents a value on a regular grid in three-dimensional space. Similar with pixels in a bitmap, voxels themselves do not typically have their position explicitly encoded along with their values. Voxels are good at representing regularly sampled spaces that are non-homogeneously filled [10][11]. Constructive solid geometry (CSG) is a technique used in solid modeling. Allow to create a complex surface or object by using Boolean operations. CSG presents a model or surface that appears visually complex, but is actually far more than combined objects. Using 3D texture mapping and frame buffer pixel operations, the algorithm can interactively generate a binary volume of the CSG model [12]. Modeling method based on CSG are presented in [13][14]. Sweep process is represented as set moving through space which may trace or sweep out volume. That volume may be represented by the set and its trajectory. Current research has shown several approximations of three dimensional shapes moving across one parameter, and even multi-parameter motions [15][16][17].

Box/Subdivision modeling is the most popular technique. Modeling starts with a primitive (usually a box or cube) and begins adding detail by dividing the box into pieces and extending faces of the box to gradually create the more detailed and complex form [18][19].



Box/Subdivision modeling

A spline is a curve in 3D space defined by at least two control points. The most common splines in 3D geometry are NURBS. A cage of splines is created to form a "skeleton" of the object you want to create. The software can then create a patch of polygons to extend between two splines, forming a 3D skin around the shape. Spline modeling is not used very often these days for character creation, due to how long it takes to create good models. The models that are produced usually aren't useful for animation without a lot of modification [20][21]. Edge/Contour it is advanced technique of poly modeling perhaps very precise technique. Often starts out with a single quad (4 points object) and extrude an edge of the quad, creating a second quad attached to the first. While poly modeling is not as fast as box modeling, it requires less tweaking of the mesh to get it "just right," and you can plan out the topology for animation ahead of time. Digital sculpting, also known as Sculpt Modeling or 3D Sculpting, is the use of software that offers tools to push, pull, smooth, grab, pinch or otherwise manipulate a digital object as if it were made of a real-life substance such as clay.



Edge/Contour and NURBS/Spline modeling

Procedural modeling is term for a set of techniques in computer graphics to create 3D models from numbers of rules. L-Systems, fractals, and generative modeling are procedural modeling techniques since they apply algorithms for producing scenes. The set of rules may either be embedded into the algorithm, configurable by parameters, or the set of rules is separate from the evaluation engine. Procedural models often exhibit database amplification, meaning that large scenes can be generated from a much smaller amount of rules [22][23].

III. Conclusion

This paper aims on techniques of geometry representations and creations. All of those methods find their usage in different fields of practice. Polygon soup can help save memory compared to the equivalent polygon mesh. For instance particle simulations can reach into the millions of polygon, causing large disk space and read/write overhead. At the other hand BSP tree allows spatial information about the objects in a scene that is useful in rendering, such as their ordering with respect to a viewer at a given location, to be accessed rapidly. Voxels are frequently used in the visualization and analysis of medical and scientific data. Procedural modeling has relation to CSG which can also be performed on polygonal meshes, and may or may not be procedural and/or parametric. Sweep representation is important in the context of applications such as detecting the material removed from a cutter as it moves along a specified trajectory. Box modeling is popular method for creating the basic shape of the model. Once practiced, the technique is very quick to get acceptable results. The downside is that the technique requires a lot of rotation of the model along the way of creation. Box modeling is useful as a way to create organic models, like characters but also hard objects like buildings, however precise curved shapes may be far more difficult to create. NURBS and Beziers curves are perhaps the most effective and precise technique in poly modeling however they are not the easiest to get started with. Splines are extremely useful when creating hard objects like car bodies, airplanes, furniture and smooth curved objects with no separately animating parts.

Poly modelers often combine these techniques to create either organic or hard objects. Digital sculpting as still a relatively new method, but it has become very popular in the few years. Method can often introduce details to meshes that would otherwise have been difficult or impossible to create using traditional 3D modeling techniques. This makes it preferable for achieving photorealistic result. But count of polygons for this technique is of the order of tens of thousands. For that reason is primarily used in high poly organic modeling (movies). Procedural modeling is often applied when it would be too cumbersome to create a 3D model using generic 3D modelers, or when more specialized tools are required. This is often the case for plants, architecture or landscapes. As we can see all these methods and representations has relation to practical usage.

Our previous work was directed to comparison hardware and software interfaces of virtual reality editors. This paper was about different approaches of virtual model creation. Our future work will be directed for issues how to implement different representations and techniques to VR editor which should be able create content for virtual world.

Acknowledgment

This work is supported by KEGA grant project No. 083TUKE-4/2015: Virtual-reality technologies in the process of handicapped persons education.

References

[1] Fengxia Li; Rong Tang; Chen Liu; Haikun Yu, "A method for object reconstruction based on point-cloud data via 3D scanning," Audio Language and Image Processing (ICALIP), 2010 International Conference on , vol., no., pp.302,306, 23-25 Nov. 2010 doi: 10.1109/ICALIP.2010.5684622

[2] Ying Wang; Ewert, D.; Schilberg, D.; Jeschke, S., "Edge extraction by merging 3D point cloud and 2D image data," Emerging Technologies for a Smarter World (CEWIT), 2013 10th International Conference and Expo on , vol., no., pp.1,6, 21-22 Oct. 2013 doi: 10.1109/CEWIT.2013.6713743

[3] Zhen Wang; Liqiang Zhang; Tian Fang; Mathiopoulos, P.T.; Xiaohua Tong; Huamin Qu; Zhiqiang Xiao; Fang Li; Dong Chen, "A Multiscale and Hierarchical Feature Extraction Method for Terrestrial Laser Scanning Point Cloud Classification," Geoscience and Remote Sensing, IEEE Transactions on , vol.53, no.5, pp.2409,2425, May 2015 doi: 10.1109/TGRS.2014.2359951

[4] Yokoyama, A.; Sato, K.; Yoshigahara, T.; Inokuchi, S., "Realtime range imaging using adjustment-free photo-VLSI-silicon range finder," Intelligent Robots and Systems '94. 'Advanced Robotic Systems and the Real World', IROS '94. Proceedings of the IEEE/RSJ/GI International Conference on , vol.3, no., pp.1751,1758 vol.3, 12-16 Sep 1994 doi: 10.1109/IROS.1994.407621

[5] Huei-Yung Lin; Jui-Wen Huang, "Extended Dynamic Range imaging: A spatial down-sampling approach," Systems, Man and Cybernetics (SMC), 2014 IEEE International Conference on , vol., no., pp.1771,1776, 5-8 Oct. 2014 doi: 10.1109/SMC.2014.6974173

[6] Nayar, S.K.; Mitsunaga, T., "High dynamic range imaging: spatially varying pixel exposures," Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on , vol.1, no., pp.472,479 vol.1, 2000 doi: 10.1109/CVPR.2000.855857

[7] Colleu, T.; Morin, L.; Pateux, S.; Labit, C., "Floating polygon soup," Digital Signal Processing (DSP), 2011 17th International Conference on , vol., no., pp.1,8, 6-8 July 2011 doi: 10.1109/ICDSP.2011.6005017

[8] Himmelstein, J.C.; Ferre, E.; Laumond, J.-P., "Swept Volume Approximation of Polygon Soups," Automation Science and Engineering, IEEE Transactions on , vol.7, no.1, pp.177,183, Jan. 2010 doi: 10.1109/TASE.2008.2010553

[9] Reed, M.; Allen, P.K.; Abrams, S., "CAD model acquisition using BSP trees," Intelligent Robots and Systems 95. 'Human Robot Interaction and Cooperative Robots', Proceedings. 1995 IEEE/RSJ International Conference on , vol.2, no., pp.335,340 vol.2, 5-9 Aug 1995 doi: 10.1109/IROS.1995.526238

[10] Steinbach, E.; Girod, B.; Eisert, P.; Betz, A., "3-D object reconstruction using spatially extended voxels and multi-hypothesis voxel coloring," Pattern Recognition, 2000. Proceedings. 15th International Conference on , vol.1, no., pp.774,777 vol.1, 2000 doi: 10.1109/ICPR.2000.905504

[11] Prakash, E.C., "Voxel tube algorithm for object oriented virtual navigation in medical volumes," TENCON 2000. Proceedings , vol.1, no., pp.252,255 vol.1, 2000 doi: 10.1109/TENCON.2000.893581

[12] Liao, Duoduo; Shiaofen Fang, "Fast CSG Voxelization by Frame Buffer Pixel Mapping," Volume Visualization, 2000. VV 2000. IEEE Symposium on , vol., no., pp.43,48, 9-10 Oct. 2000 doi: 10.1109/VV.2000.10005

[13] Yan Yan; Lin Kunhui, "3D Visual Design for Mobile Search Result on 3G Mobile Phone," Intelligent Computation Technology and Automation (ICICTA), 2010 International Conference on , vol.1, no., pp.12,16, 11-12 May 2010 doi: 10.1109/ICICTA.2010.489

[14] Zhang, N.; Huamin Qu; Kaufman, A., "CSG operations on point models with implicit connectivity," Computer Graphics International 2005 , vol., no., pp.87,93, 22-24 June 2005 doi: 10.1109/CGI.2005.1500380

[15] Bo Shu; Xianjie Qiu; Zhaoqi Wang, "Image based modeling via plane sweep based surface growing," Shape Modeling and Applications, 2009. SMI 2009. IEEE International Conference on , vol., no., pp.15,20, 26-28 June 2009 doi: 10.1109/SMI.2009.5170158

[16] Wu, J.H.; Tay, A.A.O.; Yeo, K.S.; Lim, T.B., "A three-dimensional modeling of wire sweep incorporating resin cure," Components, Packaging, and Manufacturing Technology, Part B: Advanced Packaging, IEEE Transactions on , vol.21, no.1, pp.65,72, Feb 1998 doi: 10.1109/96.659508

[17] Kretschmer, J.; Godenschwager, C.; Preim, B.; Stamminger, M., "Interactive Patient-Specific Vascular Modeling with Sweep Surfaces," Visualization and Computer Graphics, IEEE Transactions on , vol.19, no.12, pp.2828,2837, Dec. 2013 doi: 10.1109/TVCG.2013.169

[18] Zhuo shi; Min Zhou; Ke Yu, "Subdivision modeling method in net shell design," Consumer Electronics, Communications and Networks (CECNet), 2011 International Conference on , vol., no., pp.2691,2694, 16-18 April 2011 doi: 10.1109/CECNET.2011.5768761

[19] Akleman, E.; Chen, J.; Srinivasan, V., "A new paradigm for changing topology during subdivision modeling," Computer Graphics and Applications, 2000. Proceedings. The Eighth Pacific Conference on , vol., no., pp.192,201, 2000 doi: 10.1109/PCCGA.2000.883941

[20] Zi-Liang Liu; Juan Yang, "NURBS modeling for higher-order basis functions," Microwave Conference, 2009. APMC 2009. Asia Pacific , vol., no., pp.76,79, 7-10 Dec. 2009 doi: 10.1109/APMC.2009.5385328

[21] Besl, P., "Hybrid modeling for manufacturing using NURBS, polygons, and 3D scanner data," Circuits and Systems, 1998. ISCAS '98. Proceedings of the 1998 IEEE International Symposium on , vol.5, no., pp.484,487 vol.5, 31 May-3 Jun 1998 doi: 10.1109/ISCAS.1998.694538

[22] Watson, B.; Muller, P.; Wonka, P.; Sexton, C.; Veryovka, O.; Fuller, A., "Procedural Urban Modeling in Practice," Computer Graphics and Applications, IEEE , vol.28, no.3, pp.18,26, May-June 2008 doi: 10.1109/MCG.2008.58

[23] Li Gang; Shi Guangshun, "Procedural Modeling of Urban Road Network," Information Technology and Applications (IFITA), 2010 International Forum on , vol.1, no., pp.75,79, 16-18 July 2010 doi: 10.1109/IFITA.2010.122

Control module for creating panoramic pictures through DSLR

[1]*Igor ŠČUR*, [2]*Michal KOVALČÍK*
Supervisor: [3]*Peter FECIĽAK*

[1,2,3]Department of Computers and Informatics, FEI TU of Kosice, Slovak Republic

[1] igor.scur@student.tuke.sk, [3]peter.fecilak@tuke.sk, [2]michal.kovalcik@tuke.sk

***Abstract*— In this paper we present you a device, which is able to take a panoramic photography and subsequently it enables us to see it. We can also find in this paper example of current devices that capture a panoramic photography and differences between them and proposed device. This paper describes a various functions of this device and its usage in practice. At the end, we are dealing with extensions and improvements of the whole functionality of the proposed control module and its possible usage in practice.**

***Keywords*— panorama, DSLR, Raspberry Pi, stepper motor**

I. Introduction

From the ancient times people have tried to capture some important moments of their lives, whether using mural images, or later, via paintings. The big break came in 1840, when Alexander Wolcott constructed the very first camera in the world [1]. Since then cameras came a long way during which they changed their shape, size, methods of capturing and functionalities. Their development proceeds further and currently the photography became an unavoidable part of our daily life. Nowadays, the main task of photography is not only capture people but also the environment, in which they are. To accomplish this need the panoramic camera was created.

Panoramic photographs are special kind of photographs that captures a greater angle of view when compared to classic photography. They are used, for example, in the capturing of buildings, whether historical or new-buildings and various natural sceneries. The main difference between the classic and panoramic photography is aspect ratio of the photography. Classic photography has aspect ratio 1.33:1, while in panoramic photography it is usually 3:1 [2]. Panoramic photographs can be classified into three main types: cylindrical, spherical and planar panoramas [3]. Currently there exist a big number of devices that are able to capture panoramic photographs by various means for capturing. These devices may use several cameras that are located inside the sphere and panoramic photography is created by linking the photographs, which was created by every camera at the same time. Other solutions use only one camera that progressively take a photo of the whole environment. However, these devices are expensive and in the most cases they are only used for taking of the panoramic photography, while no attention to is given to the following display.

Proposed device uses only one digital single-lens reflex (DSLR) camera that during progressive rotation captures the whole room and then connects photographs into the panoramic photography.

II. Design of the Device

Design of the device consists of three parts. The first part is design of the mechanical part of the device. The main part of the mechanical part is tripod, on which is placed the panoramic head that allowed automatic capturing of panoramic photography. Panoramic head is constructed according to several criteria to meet all requirements. The first requirement is the material that has to be strong enough to hold the weight of the camera and light enough to not exceed the load of the tripod.

Another aspect that we need to consider is size that allows proper setting of the nodal point of the lens. Nodal point is always in the axis of the lens and is located in the area of the front the lens of the objective. Lens is rotating as close as possible around the nodal point. This can eliminate parallax errors, which may require a lot of retouching in the final image of panorama. Parallax error in panoramic photography can be described in the easiest way as the fault of the geometry of the picture that results due the rotation of camera during panoramic capturing around the wrong point. In practice, this error can be seen especially as deformation of close objects near the lens, creating non-continuous line of the final panoramic photography [4].

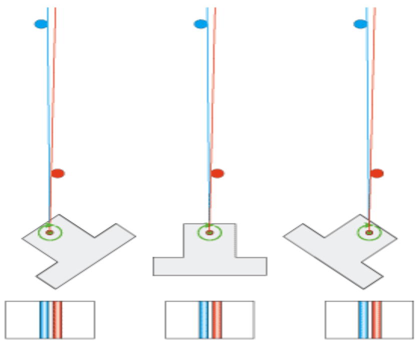


Fig. 1 The rotational axis of the node point [5]

Fig. 1 illustrated that if the rotational axis is on the node point, the gap remains constant. In the case that rotation axis is in front of the nodal point (Fig. 2), the rotation away from the object increases the width of the gap and rotation camera towards the object reduces the width of the gap and vice versa, if the rotation axis is behind of the nodal point (Fig. 2), rotating of camera away from object reduces the width of the gap and rotating camera towards near object increases width of the gap.

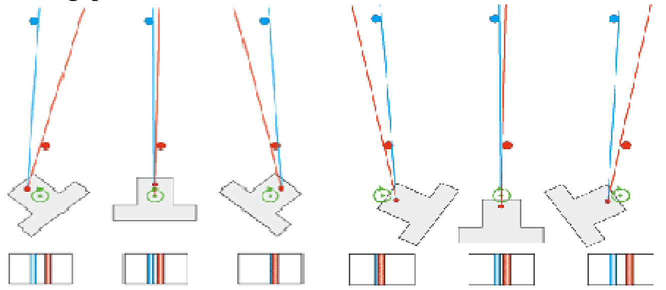


Fig. 2 Rotation axis in front and behind of nodal point [5]

Due to the rotation of the two axes, the panoramic head consists of three parts.

A. Mechanical component design

First part (1) is attached to the tripod. The second part (2) that enables horizontal rotation is in the bottom part of the first part. The third part (3), which enables vertical rotation of the camera, is also connected to the first part (1) in its upper part. The stepper motor and gear wheels are properly placed on the parts 2 and 3 and their task is to rotate with this two parts. In the Fig. 3 we can see simplistic graphical model of the device.

In case of using different cameras and lenses it is necessary to include the possibility of adapting panoramic head accordingly, following the nodal point of the lens. Nodal point is at each lens and each focal length located elsewhere for any particular lens. Because of possibility of setting of this point, is in the all parts of panoramic head carved groove that allows camera to move properly.

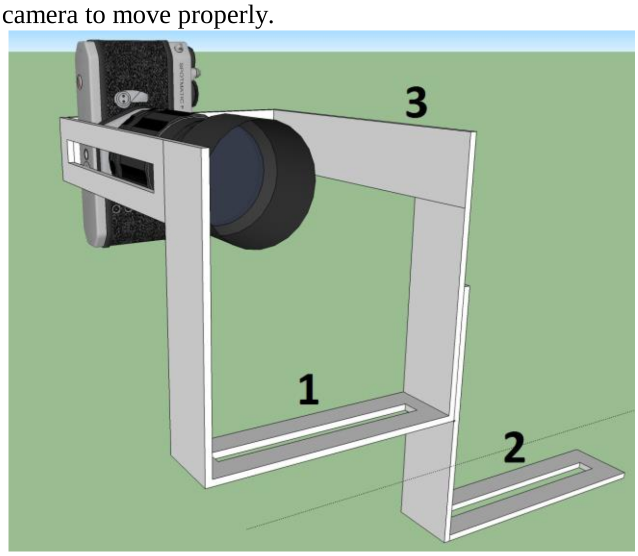


Fig. 3 Design of the mechanical part

B. Electronic part design

The second part of proposed device is proposal of its electronic part. Electronic part of the project is used for control of the rotation of the panoramic head and capturing of photographs. The main component of the electronic part is Raspberry Pi, which controls stepping motors thereby turns panoramic head. It also controls the shutter trigger of DSLR camera. To control stepper motor through GPIO pins of Raspberry Pi Darlington Transistor array is used: ULN2804A that amplifies the signal from RSPI for stepper motors. In the Fig. 4 is shown a scheme of the connection of the electronic parts of the device.

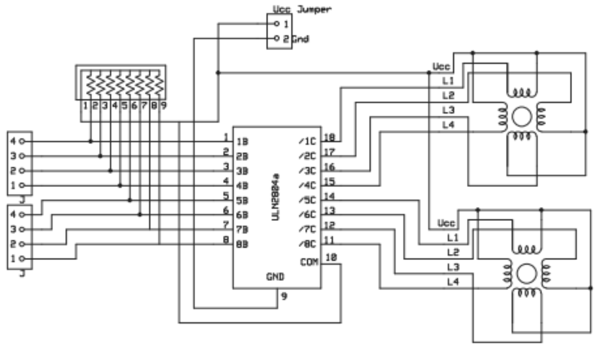


Fig. 4 Scheme of the connection of the electronic parts of the device

C. Software design

In the last part is proposed the software part of device. The whole process of creation of the panoramic photography is controlled through the website, which is running on Raspberry Pi. This website allows user to access the panoramic photographs taken previously and also particular views of the photographs that create the panoramic photograph, user can also create another panoramic photography. When creating a panoramic photograph, the parameters of camera such as ISO, shutter speed, aperture and angle of view of used lens may be set.

Capturing itself includes three steps that are logically repeating. The very first step is capturing of the photography, which is followed by downloading of photography from the memory card on the Raspberry Pi. The second step is horizontal rotation of the head of the tripod by horizontal angle of view. These two steps are repeating until the tripod head reaches its initial position. In the last, third, step the third part of the tripod head is rotated by vertical angle of view upwards and continues with progressive variation of first and second step.

The script written in the Python language is used to activate stepper motor that rotates the tripod head. USB protocol PTP enables activation of shutter and parameters of the camera. Once the capturing of the whole room is finished, all the photographs are connected using an algorithm into the one panoramic photograph, which is then possible to access through the website.

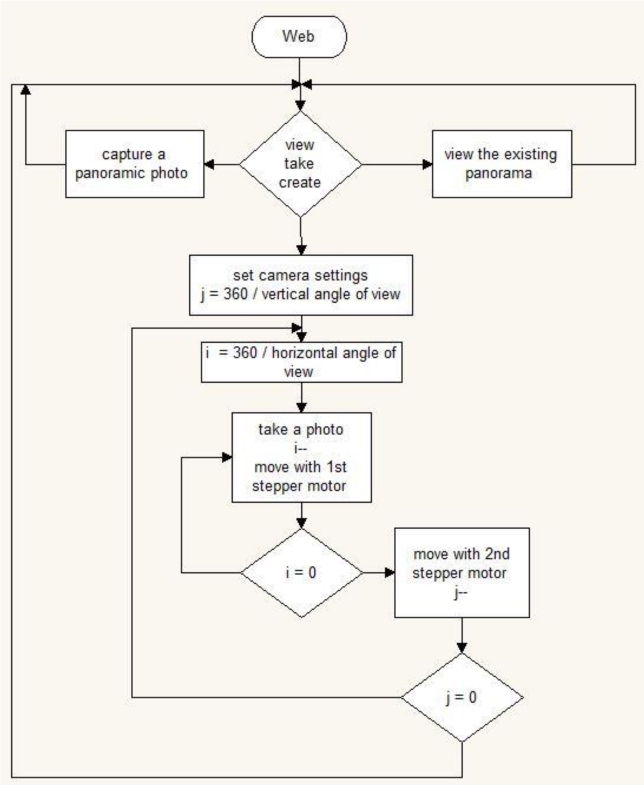


Fig. 5 Flowchart of the software part of the device

III. Solution

Proposed device was designed and allows rotation in two directions. Gear, composed of three gear wheels, is used in the vertical rotation to multiply the power of rotation of stepper motor. Gear wheel that allows horizontal rotation was created by 3D printer and is attached to the tripod. In Fig. 6 can be seen the actual solution, i.e. fully functional device.

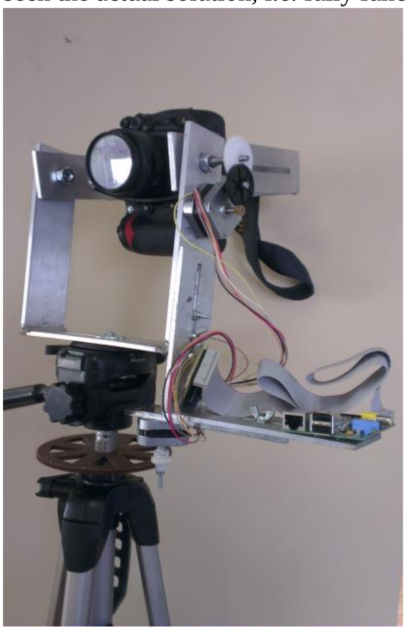


Fig. 6 Designed device

Website was designed and created in a way to allow user to select whether to capture 360 degrees (cylindrical), or 360 * 180 degrees (spherical) panoramic photography. It is also possible to set own settings of camera parameters and angle in horizontal and vertical plane, which is subsequently captured. In the Fig. 7 can be seen the output, i.e. panoramic photography obtained using this device.


Fig. 7 Panoramic photography created by designed device

IV. Conclusion

Following the outputs of the testing procedure using the final solution was found that the gear used for vertical rotation is not sufficient. Plastic gear wheels used in this gear are not strong enough and it is necessary to replace them by the metal gear wheels. The fully manual control mode of the head rotation and remote control of the camera are to be further implemented. Another important enhancements is utilization of stronger stepper motors to ensure smoother rotation of panoramic tripod head. Deployment of additional power source, i.e. external battery, is one of the possible extensions. Such extension is to allow utilization of device in the absence of electrical current. Another functionality we have been considering is to enable capturing of HDR photographs. Proposed device can be used in the creation of virtual reviews of historical monuments, or buildings intended for sale by estate companies.

References

[1] HistoricCamera.com, 2002 – 2013. http://www.historiccamera.com/history1/photo_history1843.html
[2] Pavlík, Gabriel. Ako vytvoriť panoramatickú fotografiu, 2011-2013. http://www.panoramyslovenska.sk/Panoramaticka-fotografia-Cast-prva-vybavenie-a-priprava-na-fotenie-panoram.html.
[3] Šrom, David. Virtuální navigátor ve městě Pardubice.. 1.10.2010. http://dspace.upce.cz/handle/10195/37473?mode=full
[4] Fizkej, Jan. Tvorba fotopanoramatické mapy.. 2007. http://www.gis.zcu.cz/studium/dp/2007/Fikejz__Tvorba_fotopanoramaticke_mapy__BP.pdf
[5] Kreunen, Ben. Determining the nodal point of a Lens., 24.02.2003. http://archive.bigben.id.au/tutorials/360/photo/nodal.html

Data acquisition and remote control with SOA and CPS

[1]*Tomáš LOJKA (2st year)*
Supervisor: [2]Iveta Zolotová

[1,2] Department of Cybernetics and Artificial Intelligence , FEI TU of Košice, Slovak Republic

[1]tomas.lojka@tuke.sk, [2]iveta.zolotova@tuke.sk

***Abstract*— This work deals with data acquisition, industrial remote control, related problems and my proposed solutions in the field of industry. Number of data and devices are increasing every day. Increasing number of data and devices is causing problems related with remote control. Intention of this work is to solve these problems with designed architecture, services and created algorithms. The architecture is based on Serviced Oriented Architecture (SOA), algorithms are inspired from machine learning and data mining. Wished results of this work will improve remote control and are based on better data acquisition and integration of operational and informational technologies in industry.**

***Keywords*— SCADA, CPS, SOA, remote control.**

I. Introduction

IT is one of the fastest emerging field in the world and it leads to new revolution approaches in industry in connection with automation and control. The new industrial approach's name is Industrial revolution 4.0. Citation of president and chief operating officer in Rockwell, Don H. Davis, Jr. describes principle of the revolution, who said:

"*The driving force behind productivity today isn't working faster, or working cheaper. It's working together*".

I am concentrating on industry, exactly on remote control and data acquisition, where I defined problems according research and commercial publications. I picked these: improving interactive interoperability, connection between technological layer, connection to the technological layer, better quality of monitoring, devices and data management over much higher amount of devices. According mentioned problems, I am focusing on:

1. Managing such amount of data to improve collecting, integrating, controlling, storing and visualizing in industry.
2. Decrease energy consumption and network overload with such big amount of sensors and actuators.
3. Unified and non-fixed connection to technological layer.

Defined problems are in relation with SCADA (Supervisory Control and Data Acquisition) system and remote control [1]. SCADA should take care about how to operate with amounts of devices for monitor and control. Another problem appears how to real-time represent data for controllers and humans.

I divided this paper into summarization of already known facts, solved tasks and results. The last chapter represents proposed results and next steps.

II. Initial status

My application field is monitoring and control system in large area. There are problems to use wire communication with sensors and actuators in large area systems. Battery or other power sources play key role in sensors and actuators. Therefore, sensors and actuators use wireless communication. The better mobility is an advantage in using wireless communication, too.

Sensors and actuators represent low power and computational units. It exists solution how to spare energy on nodes. For example protocols (WiseMAC, TRMA), which reduce energy consumption thanks the idle states, scheduled communication, by defining main nodes, which collect messages from other nodes in specified sequences or events or other protocols [2]. These solutions are not evolving meaning of embedded systems and smartness into communication.

III. Solved tasks in previous year

The completely defined solution for defined problems consists of [3]:

1. CPS-s (Cyber-physical systems) – in my work, represent low power devices, which connect physical and digital world.
2. Gateway - represents gateway to technological processes (represented by CPS).
3. Clients – represent HMI and MES system.

A. CPS communication reduction

I concentrated on creating base layer for monitoring technological processes in the previous year.

I have already designed, realized and done experiments only with the first point (CPS-s). In the first part, my main goal was to design architecture and create a method for data acquisition. Intention was to design solution, which will reduce amount of communication messages and decrease energy consumption. The reduction should decrease amount of messages, but not decrease information about monitored environment. Finally, the intention is to reach better remote control and data acquisition describe in [4],[5].

In my **first** step, I used two methods, which help to find information in measured raw of data.

First motivation came from Shannon entropy. Shannon entropy can detect change of entropy in raw of sensor data. Tsallis defined better representation of entropy. His entropy does not use logarithm. Tsallis entropy saved six processor's cycles for one loop in my implementation. I did experiments only in laboratory conditions. Finally, in production area, we can reached 23% time reduction according performance results. Used Tsallis entropy was enough responsive to dynamic processes like sensing position. I reached better sensitivity to unexpected changes or higher probability changes. I realized implementation in Java and tested on Android with gyroscope in laboratory conditions. I described this method and solution in [6].

In contrast to previous solution, I realized alternative solution, SPRT (Sequential Probability Ratio Test). SPRT detects unexpected changes in sensor's data raw. SPRT operates with two hypothesis and allows global view on monitored process. I have published these results, but publication has not been registered yet.

Energy is critical for low power devices. Wireless communication is defined like a big energy consumer. Therefore, amount of messages and capacity of messages have big influence on power consumption [6]. All sensors is based on the SOA and its detailed architecture is described in my phd thesis proposal.

B. Conclusion Entropy vs. SPRT

In conclusion, I picked Tsallis entropy. Tsallis gives better results in sensitivity and better scalability of monitored data. Better scalability enables to do much more with detected changes.

C. Definition of CPS - sensor

According reached result, I designed and build sensor node. First was designed data processing method.

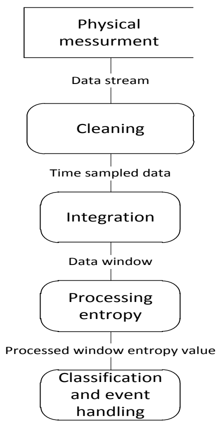


Fig. 1. Simplified processing method implemented in device for data sensing

Method was partially motivated from data mining. The whole method I described in [3]. I designed an architecture for sensor node and communication interface for data acquisition in SCADA system. Data are integrated in gateway, which enables communication between low power devices and server. I have already defined principles for data management and functionalities in gateway. Gateway creates a gate to technological processes, but analyze data, manage devices with help of machine learning. The gateway is based on SOA. I picked K-NN like a first preprocessing method in gateway. K-NN helps to classify devices. I designed another method, which will use classified these devices for data and device management. Finally, it will help to realized unified connection to technological layer thru amount of low power devices, partially solved in [7]. Complex results have not been published yet here.

IV. Future work

I want to finish gateway realization and implement already designed methods. Then, I will verified these methods for data and device management. In the next step, I will design and realize unified connection to the technological processes, what should improve remote monitoring and control. In other words, I want to realize a unified connection between OT (Operational Technologies) and IT (Information Technologies). This connection will help to collect data from increasing number of devices. I will create sample for HMI and MES system, which I will use for gateway's and CPS's methods verification. According specified criteria, I will find other improvements in used methods. I want to use this solution in the meaning of SOA [8], to imrpove integrity in industrial conditions in data acquisition and remote control.

V. Conclusion

Last year, I designed, implemented and verified communication reduction in CPS. According requirements, I designed SOA inspired architecture for CPS and data acquisition. I designed gateway for interconnection between OT and IT with help of machine learning, a unified connection for client (MES, ERP) and I realized gateway partially.

Acknowledgment

This publication is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF (30%) and by grant KEGA - 001TUKE-4/2015 (70%).

References

[1] T. Lojka, M. Miškuf, I. Zolotová, "Service oriented architecture for remote machine control in ICS", in: *SAMI 2014: IEEE 12th International Symposium on Applied Machine Intelligence and* Informatics, 2014, Herľany, Slovakia. - Danvers: IEEE, 2014 S. 327-330. - ISBN 978-1-4799-3441-6

[2] T. Lojka, "Sensor networks autonomy", in: *SCYR 2014: 14th Scientific Conference of Young Researchers*, 2014, Herľany, Slovakia. - Košice : TU, 2014 S. 239-242. - ISBN 978-80-553-1714-4

[3] T. Lojka, I. Zolotová, "Distributed sensor network - data stream mining and architecture, in: *Advances in Information Science and Applications - Volume 1*: Proceedings of 18th International Conference on Computers (part of CSCC `14): 2014, Santorini Island, Greece. P. 98-103. - ISBN 978-1-61804-236-1 - ISSN 1790-5109

[4] K. Tafsi, T. Lojka, I. Zolotová, "Man-Machine interfaces for mobile devices in ICT", in: *EEI 5*: Proceedings of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice. - Košice : TU, 2014 S. 581-584. - ISBN 978-80-553-1704-5 (in Slovak)

[5] T. Lojka, et al., "Communication Engine in Human-Machine Alarm Interface System", in: *Advances in Intelligent Systems and Computing : Emergent Trends in Robotics and Intelligent Systems*. - Switzerland : Springer, 2014 Vol. 316 (2015), p. 129-136. - ISBN 978-3-319-10782-0 - ISSN 2194-5357

[6] I. Zolotová, T. Lojka, "Online data stream mining in distributed sensor network, in: *WSEAS Transactions on Circuits and Systems*. Vol. 13 (2014), p. 412-421. - ISSN 1109-2734

[7] T. Lojka, I. Zolotová, "Improvement of Human-Plant Interactivity via Industrial Cloud-Based Supervisory Control and Data Acquisition System", in: *Advances in Production Management Systems*. Vol. 440, no. Part 3 (2014), p. 83-90. - ISSN 1868-4238

[8] I. Zolotová, L. Laciňák, T. Lojka, "Architecture for a universal mobile communication module, in: *SAMI 2013: IEEE 11th International Symposium on Applied Machine Intelligence and Informatics:* proceedings : 2013, Herľany, Slovakia. - Budapest : IEEE, 2013 S. 61-64. - ISBN 978-1-4673-5926-9

Data Analyses with Use of Business Intelligence and Big Data with Focus on Industrial Systems

[1]*Martin MIŠKUF (1st year)*
Supervisor: [2]Iveta Zolotová

[1,2]Dept. of Cybernetics and Artificial Intelligence, FEI TU of Košice, Slovak Republic

[1]martin.miskuf@tuke.sk, [2]iveta.zolotova@tuke.sk

***Abstract*— This article deals about problems connected with processing huge amounts of data created in all science and engineering domains, including industry. Thanks to fast development and decreasing prices of networks, storages we can now monitor, transfer and save more data than ever before. In this increasing trends of data sources may be hidden new potentially useful information so new techniques for these purposes was made. With this progress are also connected Industrial revolution 4.0 and Internet of Thing. Nowadays Big Data is the one of the hottest terms connected with processing of large-volume, complex, growing data sets with various sources. In this article I´ll focus on growing amounts of data with focus on industrial systems and how can be Big Data and other ideas possibly used.**

***Keywords*— Business intelligence, Big Data, data analyses, industry systems.**

I. Introduction

Data is all around us. We are living in the world where information technologies are used almost everywhere. With this trend come various advantages and disadvantages. One of this moot news is growing amounts of data of various types. Some analysis says that 90% of the data was created in the last two years. [1][2]

Rapidly generated data thanks to decreasing prices of networks, mobile devices, sensors, storages and others. All devices we can connect to the internet, so data can be stored and used later. We can see it in business, industry, health care etc. where technologies are used almost everywhere. Imagine just mobile phones. In these devices are other sensors collecting data like GPS, accelerometers, light sensors, speakers etc. from which can be all measurements collected and potentially used to answer business, health and other questions.

II. Changes from past years

In the past years, emergence of promising tools such as Enterprise Systems provided companies with solutions to improve their productivity and service quality. Today's competitive nature of world enforces industry companies to implement more recent technologies to secure their position among competitors. [19]

Recent technological advances in the field of communication and computer science have provided cost-effective solutions for companies to acquire and transfer gigantic amount of data from their fleet of assets. Consequently. Handling these huge sets of data is not easily achievable, therefore supporting "Big Data" is one of the most recent topics in the world industry. [18]

III. Increasing amounts of data

There are many predictions about amounts of data in our world, one of them, made by IBM, was mentioned in introduction of this paper.

Fig. 1. Amounts are increasing in orders of magnitude [2]

These predictions may have different numbers but they have one thing in common. Amounts are increasing in orders of magnitude. This fact we can divide in to three main periods:

- **Data created by employees** – in the first period data was generated and accumulated by workers. In other words, employees of companies was entering data into computer systems.
- **Data created by users** – with the development of internet things has changed, so now internet users can generate more data than employees and our amount of data are increasing. Think about pages like Facebook, YouTube, Instagram or Flickr.
- **Data created by machines** – now is even a third level in this progression, because now machines are accumulating data. The buildings and all of our cities are full of the monitors that are monitoring humidity and temperatures, electricity usage, there are smart meters in our homes. And that once machines are accumulating data that orders of magnitude higher than users

These big volumes of data are the secondary consequences of new technologies that we can find in every field of our life. Some of this big steps that are connected with industry are Industrial Revolution 4.0 and Internet of Things.

A. Industrial Revolution 4.0

The Industrial Revolution is a concept and a development that has fundamentally changed our society and economy. On the next picture Fig. 2. Industrial Revolution comes through four stages until now. From the steam machine to intelligent cyber-physical systems. [6] That means more data created during production.

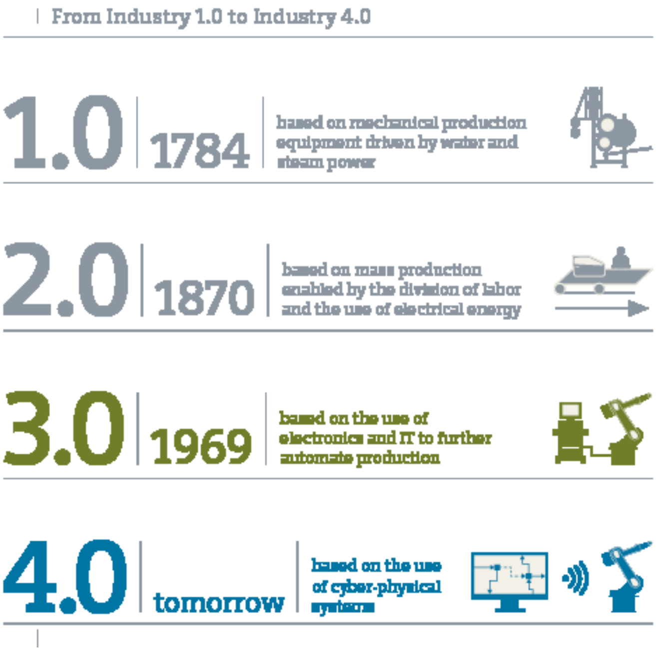


Fig. 2. Different stages of Industrial revolution [6]

B. Internet of Things (IoT)

This term indicates network which consist of physical objects or "things" embedded with electronics, software, sensors and connectivity to enable it to achieve greater value and service by communication with other connected devices and exchanging data with the manufacturer, operator. Each thing is uniquely identifiable through its embedded computing system but is able to interoperate within the existing Internet infrastructure. [7]

In the next decade it is estimated that there will be some 25 billion "intelligent systems", internetconnected devices that analyze information collected in real time. It is larger than the combined number of PCs, phones and tablets. [21]

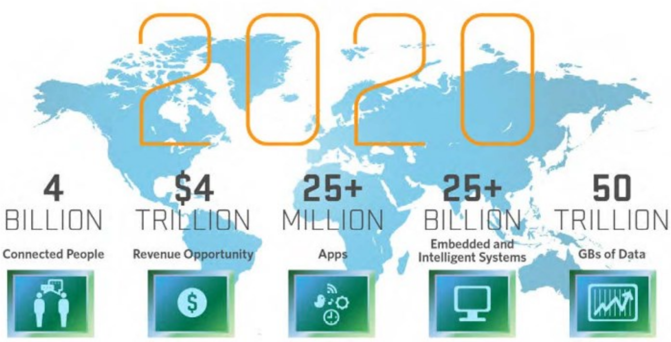


Fig. 3. Intelligent systems in the next decade [21]

IV. PROCESSING AND ANALYZING OF DATA

The fact that we can collect data from numerous sources and store them is only part of success. This data are usually stored in data warehouses. With this stage problems come. Data has various formats. Data warehouses must be capable to handle, store and effectively manage all kinds of incoming data that will be used in later analyses.

If we have data in our databases, we have to find relations among data to understand them. For these purposes was made few techniques which have one thing in common, it is finding information among data. It labels as data analyzing or data mining. Data analysis is a process of inspecting, cleaning, transforming, and modeling data with the goal of discovering useful information, suggesting conclusions, and supporting decision-making. [13] [14]

Two terms that are one of the most connected with data analysis are Business intelligence and Big Data. I choose them because they will be field of research in my next study. On the next picture is graph with search rates of these terms during last years.

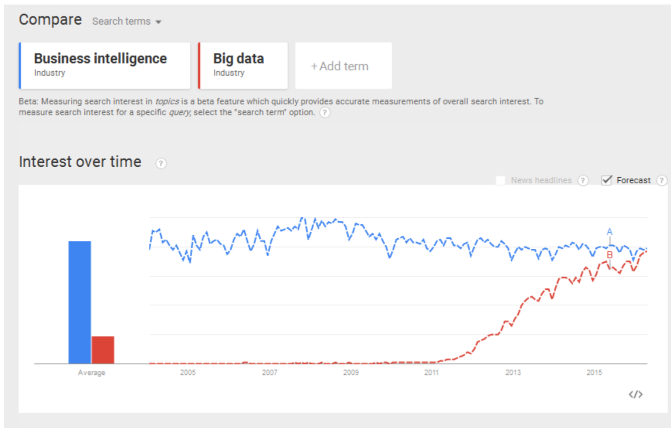


Fig. 4. Google Trends - search rates for terms Business intelligence and Big Data in past years [6]

As we can see that Business intelligence was more frequented term. But now with growing amounts of data are coming new approaches. Big Data have the same idea as Business Intelligence but they are oriented to handle large-volume, heterogeneous, autonomous sources with distributed and decentralized control, and seeks to explore complex and evolving relationships among data.

V. BUSINESS INTELLIGENCE

As Business Intelligence (BI) we understand all techniques that are needed for distributing right information, collected from internal and external sources, for right people and in right time. This idea is old and has really long history.

Richard Millar Devens used term Business Intelligence in 1865 to describe how banker gained profit by receiving and acting upon information collected from his environment. This idea is today still at the heart of BI. [9]

A. Definition

Business intelligence is defined as set of techniques and tools, which allows transformation of raw data into meaningful and useful information for business analysis purposes. BI technologies must be capable of handling large amounts of unstructured data from various sources to help identify, develop and otherwise create new strategic business opportunities. BI technologies provide historical, current and predictive views of business operations. [11]

B. Architecture

Data collected from various sources is transformed into data warehouse by ETL procedures (Extraction, Transformation, Loading). This is used as data source for OLAP (Online analytical processing). OLAP tools enable users to analyze multidimensional data interactively from multiple perspectives. [12]

C. Outputs from BI

Information are presented in form that is easy to understand. With numerous different functions, users can create static or dynamic reports and analysis. These outputs can be distributed by mail or visible on internet on the BI portal.

VI. Big Data

The term "Big Data" was firstly used Roger Magoulas in 2005. The most fundamental challenge for Big Data applications is to explore the large volumes of data and extract useful information or knowledge for future actions. [5]

Big Data is a new idea from last few years, younger than Business Intelligence but they have much in common. The main difference between them is in volumes and diversity of data they are work with. Information extracted from data has value only if is delivered in time. This causes problems with real time analyses of big volumes data. Computers used for these purposes have to be powerful.

Fig. 5. Latest technologies and other terms connected with Big Data [15]

According to HACE Theorem Big Data starts with large-volume, heterogeneous, autonomous sources with distributed and decentralized control, and seeks to explore complex and evolving relationships among data. [14]

If we are talking about data, there are some definitions to help to define data, which is used in Big Data cases. We are talking about "V-s".

A. The 4 "V":

- Volume is about size of collected data. It is amount of collected data and number of data sources.
- Velocity means speed which data is produced. Usually we are collecting data from various sources in different rates. It is also rate at which data is collected.
- Variety comes with different types of captured data. Data are usually collected from sensors and they can be simple meters measuring one value or more specific devices able to capture, process and transfer more complex measurements. The most common types of data are documents, pictures and videos.
- Veracity is the fourth "V" and it means the level of reliability with certain kinds of data.
- Value is as the fifth "V". We can find information in data, this information labeled as gold and the value is limited by our ability to mine or extract gold from the data.

VII. Problems connected with data analyses

Amounts of data are increasing and if companies want to analyze them, they have to solve various problems. One of the biggest problems are connected with huge volumes of data and real time analyses. These problems involves everything from networks, powerful computers through new approaches like Hadoop and much more.

A. Problems connected with performance of systems

First group formed by problems connected with systems that are processing data. In this major group there scientist are solving questions about.

- Faster networks. When we speak about Big Data, we are speaking in Petabytes and more than that. Imagine that you want to transfer 1PB via network, it will take more than one week.
- New architectures like grid-computing, clouds. With Big Data comes need for powerful computers that allow real or near real time analyses.
- New types of data storages were we could effectively store structured and unstructured data without any increased needs on performance.
- Security and privacy problems. If Big Data are working with some confidential data there can be restrictions about places where data can be stored.

B. Problems connected with analyzes an visualizations of processed data

Second group is about find better ways for processing data, finding new useful information and presenting them. In this group can be defined problems from many fields.

- Data mining problems connected with mining in big piles of data.
- Mathematical problems connected with statistics and making predictions.
- Finding new ways how to present information in better ways and different devices. We could see this trend in past 10 years where computers were replaced by mobile devices.

These were some of the most common solved and unsolved problems connected with Big Data. If we focus on industrial systems, we can define more specific tasks for future research and development in next years.

Current industrial evolution is guiding industry toward maximum leverage. Companies with more futuristic vision will have better opportunity of being significantly successful and profitable in recent future. For these purposes are the importance of intelligent prognostics and health management in industry for retaining production and service excellence. [16]

VIII. Big Data and big opportunities

There are multiple uses for Big Data in every industry. Some of the latest solutions are in healthcare, banking, insurance, energy, etc. Solutions offer many possibilities, they can analyze larger volumes of data than was previously possible to drive more precise answers, to analyzing data in motion to capture opportunities that were previously lost. Big Data and analytics platform will enable your organization to tackle complex problems that previously could not be solved. [3]

IBM made ones of the latest noticeable solutions. Watson and Bluemix, which can be used for building applications.

A. IBM Watson

We can Watson describe as a cognitive technology that processes information more like a human than a computer. It can understand natural language and can gets smarter by being taught from users and by learning from prior interactions. [4]

Watson is a question answering computing system that IBM built to apply advanced natural language processing, information retrieval, knowledge representation, automated reasoning, and machine learning technologies to the field of open domain question answering.[22]

In 2011 Watson competed on Jeopardy! against former winners and received the first place prize of $1 million. After two years in February 2013, IBM announced that Watson software system's first commercial application would be for utilization management decisions in lung cancer treatment. Watson is used as chef in cognitive cooking. Now it is possible build solutions with Watson services.

B. IBM Bluemix

IBM's Open Cloud Architecture implementation based on the Cloud Foundry project. Cloud Foundry is an open source platform as a service (PaaS) that lets you quickly create and deploy applications on the cloud.[23]

IX. Future Work

In my future work, I want to focus on cooperation with IBM. In Kosice they solving problems connected with IoT, where I can contribute with data analyses, and they have other plans with Watson.

X. Conclusion

During my previous study, I was working with companies like U.S. Steel and IBM in Business intelligence field, which is widely used in many companies. Predictions say that amounts and diversity of data are rapidly increasing so for analytical purposes have to be new techniques used. I think that Big Data is field where can be done lot of research in next 5 years. Big Data are used to extract maximum information from every possible data source all around us. Therefore, we can say that Big Data require a "big minds" to consolidate data for maximum values. [20]

Acknowledgment

This publication is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF (30%) and by grant KEGA - 001TUKE-4/2015 (70%).

References

[1] IBM, *"What Is Big Data"*, http://www-01.ibm.com/software/data/bigdata/what-is-big-data.html, IBM, 2012, Retrieved March 2015.

[2] IBM, *"Analytics - Analyze all of your data for a fast ROI"*, http://www.ibm.com/analytics/in/en/what-is-smarter-analytics/innovate-with-analytics-tools.html, Retrieved March 2015.

[3] IBM, *"Big Data in action"*, http://www-01.ibm.com/software/data/bigdata/industry.html, Retrieved March 2015.

[4] IBM, "IBM WATSON - *What is Watson?"*, http://www.ibm.com/smarterplanet/us/en/ibmwatson/what-is-watson.html, Retrieved March 2015.

[5] A. Rajaraman and J. Ullman, *"Mining of Massive Data Sets"*, Cambridge Univ. Press, 2011.

[6] Siemens, *"On the road to the Fourth Industrial Revolution"*, http://www.siemens.com/annual/13/en/company-report/report-industry-solutions/strategic-context/, Retrieved March 2015

[7] J. Höller, V. Tsiatsis, C. Mulligan, S. Karnouskos, S. Avesand, D. Boyle, *From "Machine-to-Machine to the Internet of Things*", Introduction to a New Age of Intelligence, Elsevier, 2014, ISBN 978-0-12-407684-6.

[8] H. Moller, *"The elephant and the blind men"*, http://cdn2.hubspot.net/hub/134568/file-1208368053-jpg/6-blind-men-hans.jpg, Retrieved March 2015.

[9] M. D. Richard, *"Cyclopaedia of Commercial and Business Anecdotes; Comprising Interesting Reminiscences and Facts"*, D. Appleton and company, p. 210.

[10] H. P. Luhn, *"A Business Intelligence System"*, IBM Journal 2 (4): 314, doi:10.1147/rd.24.0314.

[11] R. Olivia, *"Business Intelligence Success Factors: Tools for Aligning Your Business in the Global Economy"*, Hoboken, 2009, N.J: Wiley & Sons, ISBN 978-0-470-39240-9.)

[12] *"OLAP Council White Paper*", OLAP Council, 1997.

[13] C. O'Neil, R. Schutt , *"Doing Data Science"*, O'Reilly, ISBN 978-1-449-35865-5.

[14] W. Xindong; Z. Xingquan; W. Gong-Qing; D. Wei, *"Data mining with big data"*, Knowledge and Data Engineering, IEEE Transactions on , vol.26, no.1, pp.97,107, Jan. 2014

[15] Forbes, *"Billions Of Reasons To Get Ready For Big Data*", http://www.forbes.com/sites/oracle/2012/12/13/billions-of-reasons-to-get-ready-for-big-data/, Retrieved March 2015

[16] J. Lee; B. Bagheri, H. Kao, *"Recent Advances and Trends of Cyber-Physical Systems and Big Data Analytics in Industrial Informatics"*, IEEE Int. Conference on Industrial Informatics (INDIN), 2014.

[17] J. Lee; L. Edzel, B. Bagheri, H. Kao, *"Recent advances and trends in predictive manufacturing systems in big data environment"*, NSF Industry/University Cooperative Research Center on Intelligent Maintenance Systems (IMS), University of Cincinnati, United States, 2013

[18] E. Dumbill, *"Big Data is Rocket Fuel"*, Big Data, 2013.

[19] L. Da Xu, *"Enterprise systems: state-of-the-art and future trends"*, Industrial Informatics, 2011.

[20] A. Jacobs, *"The Pathologies of Big Data"*, Comm. ACM, vol. 52, no. 8, pp. 36-44, 2009.

[21] V. Turner, C. MacGillivray, J. Gaw, R. Y. Clarke, M. Morales, B. Kraus, *"IDC FutureScape: Worldwide Internet of Things 2015 Predictions"*, IDC FutureScape, 2014

[22] IBM Research, *"The DeepQA Research Team"*, http://researcher.watson.ibm.com/researcher/view_group.php?id=2099, Retrieved April 2015.

[23] IBM, *"What is IBM Bluemix?"*, http://www.ibm.com/developerworks/cloud/library/cl-bluemixfoundry/, Retrieved April 2015.

Design of Decision Support System in Data Analysis Area

[1]*Peter MICHALIK (3rd year)*
Supervisor: [2]Iveta ZOLOTOVÁ

[1,2]Dept. of Cybernetics and Artificial Intelligence, FEI TU of Košice, Slovak Republic

[1]peter.michalik.2@tuke.sk, [2]iveta.zolotova@tuke.sk

***Abstract*—This paper describes the results of my work in recent years during my doctoral studies, but the emphasis is mainly on previous year. Part of the paper is a brief introduction to the topic, my motivation to solve the problems of selected area as well as references to the articles that were written during the study. I would also mention my plans for further papers in the forthcoming period as well as other activities related to my studies. I am talking about the passing of summer school or working in the various projects that are more or less related to the topic of my dissertation.**

***Keywords*—Data Analysis, DSS, Knowledge Discovery in Databases, Business Intelligence.**

I. INTRODUCTION

The field of data analysis is currently one of the most attractive areas of IT´s world. Reason is a very simple. Amount of data available to different types of organizations, businesses or any other entities is constantly increasing. This data often differ according to the type of the domain in which it originated. It results that we currently have the market a full of different analytical platforms, data-mining-tools and business intelligence technologies that serve to organizations to mine useful information and knowledge from the data storing in their data warehouses [1] [2]. The problem is that these organizations often do not know what analytical platform they need or what tool for knowledge discovery is thus suitable for them. Frequently they just have no idea in what formats are keeping their data, volume and thus whether it is appropriate for them to consider, for example, the Big Data technologies[3][4][5].

The topic of my thesis is the design of decision support system, which should simplify and streamline the selection of appropriate analytical platform based on the input parameters, choosing by different organizations. Design of that system is necessary to divide into several stages of research. It should be noticed that not all phases are scientific research. We talk about the application research in the case of implementation of that system.

II. STAGE RESEARCH

The aim of my research is to design decision support system for the selection of appropriate analytical environment. Such a system or methodology should be implemented in a suitable web application. This will ensure that individual entities will have easier orientation in the selection. In the first steps, I assume that the behavior of organizations in the selection can be categorized into two groups. There are organizations that do not know what criteria are important to them. On the other hand, organizations preferring their rank (it is their personal weight of the criterion) are not very exceptional. Based on these facts, a series of mathematical models and algorithms were designed. These models will be the basis for my decision support system. However, it should be noticed that design of these models and algorithms is only one of the successive stages of research. Figure 1 shows all five research and application stages to fulfill my explorations.

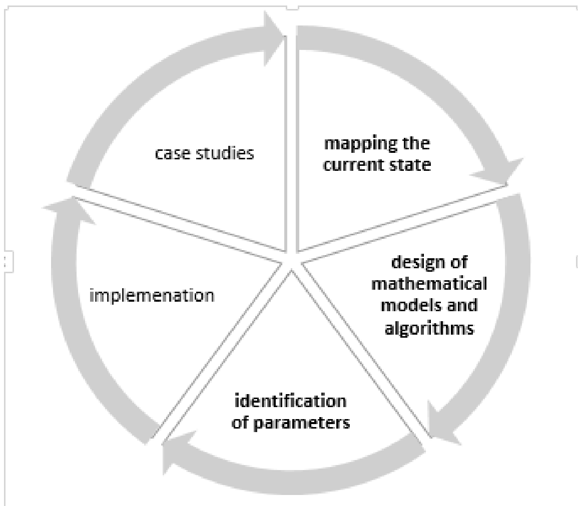


Figure 1. Stage research

There are phases highlighted in bold. These stages have been done already. Mapping the current state was a very time consuming and took about 1.5 years. Design phase of mathematical models and the identification of parameters entering the model is the outcome of my work in the previous period.

III. SOLVED TASKS IN PREVIOS YEAR

A. Design of mathematical models and algorithms

As mentioned above, I have been working in previous year on the design of mathematical models and algorithms with the identification of parameters that would come into these models. I designed three basic mathematical models.

- *Model with the criteria selection without priorities of the organization.*
- *Model with the criteria selection rated by experts.*
- *Model with the criteria selection rated by experts and own preferences.*

A detailed description of these models and operating principle of the proposed algorithm is currently part of the paper to be submitted to the 14th International Conference on Perspectives in Business Informatics Research in Estonia.

B. Identification of parameters

The third stage of my research is the identification of parameters entering to proposed models. That stage is also important part of the entire system. It is not very easy to design a mathematical model working with *n* parameters, but identification of relevant criteria for choosing a good analytical platform is more complex. This task is the result of a comprehensive analysis of current trends and the current state of the art. It is the most demanding phase from the terms of time.

There are several ways to identify criteria entering into my proposed decision support system. One option is to talk with experts and people who deal with this issue. We could use a questionnaire for this purpose, but it is not very practical solution because every expert would have to identify criteria and give them weights (due to the second and third model). Therefore, I decided for an experimental and also bit risky way. Criteria will be identified from set a 200 scientific and commercial articles. My decision follows from fact that the authors of these papers are ultimately also experts of people who are dealing with this topic. Each time when an attribute describing some technology appears, this attribute will be recorded as an input parameter – criterion for selection. This will yield *n* input parameters and also we are able to give weights to them, what is more interesting.

IV. PLANS FOR THE FURTHER STUDY

Plans for the future can be quite easily deduced from Figure 1. Since the first three stages have been done, the stage of implementation is my aim to the next period of my study. That stage encompasses the implementation of the proposed algorithms and models in an appropriate and pleasant appearance. I decided that my DSS will be in the form of a web application. Part of the web application that will serve as a DSS will also categorizing the different input parameters (decision criteria). Thus, this application will, inter alia, highly informative nature.

The last important stage of research is to verify whether the DSS really alleviates decisions the organizations and whether recommended output is really best for the organization. This verification will be done using the method of case studies, which is related to my other activities in our department. Among other things, I am contributing to the analysis of data from the monitoring environment of T-Systems Company. That my activity is one of the case studies. I work with two analytical platforms – R and RapidMiner in analyzing the data. I would also mention that I attended a summer school in Madrid to improve my abilities in data-mining just in R in last summer.

V. CONCLUSION

Finally, I would like to notice that the proposed DSS is a combination of scientific and application activities and a verification desired effect is very difficult. It is related to the amount of time that is required for such verification. Organizations need time to use DSS and also after the selection of analytical platforms to have a lot of time to analyze their data. After that period, they can say whether the selected analytical tool for them was really the most satisfying solution. And even that is not quite surely, as this field of IT is rapidly changing. What is today for this type of data suitable and fastest, tomorrow may no longer be true. In any case, I would like to make this system as a first step in the research of DSS in the field of data analysis.

ACKNOWLEDGMENT

This paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF (30 %) also by grant KEGA-001TUKE-4/2015 (70 %).

REFERENCES

[1] Adriána Šántová, Peter Michalik, Iveta Zolotová - “Analýza veľkých dát použitím technológie IBM InfoSphere BigInsights” *In: Electrical Engineering and Informatics 5 : Proceedings of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice*. - Košice : TU, 2014 S. 875-879. - ISBN 978-80-553-1704-5

[2] Gabriela Konkoľová, Peter Michalik, Iveta Zolotová – “Procesná analýza výrobnej firmy s využitím vybraných technológií” In: *Electrical Engineering and Informatics 5 : Proceedings of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice*. - Košice : TU, 2014 S. 880-885. - ISBN 978-80-553-1704-5

[3] Daniela Judičáková, Peter Michalik, Iveta Zolotová –“Tvorba reportovacej vrstvy v prostredí Microsoft SQL Server 2012*”In: Electrical Engineering and Informatics 5 : Proceedings of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice*. - Košice : TU, 2014 S. 886-890. - ISBN 978-80-553-1704-5

[4] Katarína Jeňová, Peter Michalik, Iveta Zolotová – “Úlohy objavovania znalostí v prostredí technológie Microsoft SQL Server 2012” *In: Electrical Engineering and Informatics 5 : Proceedings of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice.* - Košice : TU, 2014 S. 891-895. - ISBN 978-80-553-1704-5

[5] Pavol Hajastek, Peter Michalik, Iveta Zolotová - “Analýza meteorologických dát s využitím nástroja QlikView”In: *Electrical Engineering and Informatics 5 : Proceedings of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice.* - Košice : TU, 2014 S. 896-890. - ISBN 978-80-553-1704-5

Design of Modern Natural User Interfaces

[1]Róbert PEŤKA (2nd year)
Supervisor: [2]Branislav SOBOTA

[1,2]Department of Computers and Informatics, FEI TU of Košice, Slovak Republic

[1]robert.petka@gmail.com, [2]branislav.sobota@tuke.sk

***Abstract*—Virtual reality becomes an everyday part of our lives. Important part of virtual reality is human-computer interaction. With advancement of technology, there are interesting new ways opened to control virtual reality systems in human natural manner. Such interfaces are commonly found interesting and intuitive by users. Despite the undeniable advantages of these systems, it is yet not possible to say that we have found considerable applications in enterprise environment. This paper provides insight into the development of systems based on virtual reality with regard to the possibility of effective deployment in such enterprise environments.**

***Keywords* — virtual reality, human-computer interaction, 3D interface, natural user interface, gesture, fatigue, enterprise environment**

I. Introduction

Thanks to sophisticated implementations and more reasonably priced solutions, virtual reality technologies become part of a growing number of research institutions and households. Within the household, it is logical that advanced interfaces are used mainly for playing computer games that are critical mass of interest and financial income therefore, long-term driver of technological progress of ancillary hardware. Despite the undeniable advantages that such advanced 3D solutions can provide, it is not a common practice to use them within a professional environment. Dominant technologies in this field are yet traditional and well established inputs and outputs such as keyboard, mouse and 2D monitor.

Most of advanced virtual reality interfaces are based on control by movements of the user`s body. This may be rather difficult task in the long term usage. Gaming is relatively time-limited activity, which usually takes from several minutes to few hours in extreme situations. Because player spends controlling the virtual world disproportionately less time than a professional user within their typical working hours, this problem is not that much critical. Players can leave the game at any time and rest if they feel physical fatigue without the risk of potentially dangerous consequences. Moreover, in the context of gaming physical activity is welcomed as it can be understood as relax in a way, which is in many cases similar to serious sports. On the other hand in a professional environment, there must be emphasis on efficiency and long-term sustainable performance.

II. Design of Modern Natural User Interfaces

The term "natural user interface" (NUI) is an emerging computer interaction methodology which focuses on human abilities such as touch, vision, voice, motion and higher cognitive functions such as expression, perception and recall. A natural user interface or NUI seeks to harness the power of a much wider breadth of communication modalities which leverage skills people gain through traditional physical interaction.

Much in the same way the graphical user interface (GUI) was a leap forward for computer users from command line interfaces, natural user interfaces in all of their various forms will become a common way we interact with computers. The ability for computers and human beings to interact in diverse and robust ways, tailored to the abilities and needs of an individual user, will release us from the current constraints of computing allowing for complex interaction with digital objects in our physical world. [1]

The new generation of NUI can greatly profit from user interface techniques that are adapted the way people perceive and interact with the real world. Despite the fact that almost all studies on the design and implementation of advanced 3D interfaces have been implemented in the context of virtual reality systems, their deployment is actually much lower than would be expected [2]. Although the researchers demonstrated 3D interface with a high degree of functionality and usability, most VR applications in the real world has very simple user interface.

One of the main conditions for the active deployment of 3D interfaces based on VR technologies is their effective utilization when interacting with VR system. The most common act is the selection of the object over which the user intends to perform a given action [3][4][5][6]. Typical issues within the 3D interface include double vision, spatial disorientation and eye accommodation problem due to change in its movement within the created virtual stereoscopic scenes and real world objects.

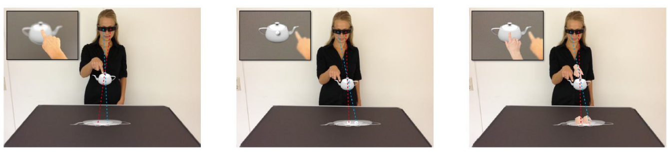


Fig. 1. Problems of object selection in 3D space

Appropriate adaptation of selection algorithm based for example on the distance of the virtual object from the observer ultimately increases the efficiency of a given VR system while reducing the burden on the user visual system.

User fatigue factor is crucial for the usage in an enterprise environment. Problems such as hand fatigue and poor gesture identification are primarily caused by irrationality between input defined by gestures and mapping of these gestures within the software [7]. The mobility of the arm and the fingers must be taken into account when defining gestures in a way they do not cause unnecessary high muscle tone.

When designing and implementing interactive mixed and virtual reality environments, one should explore various sources of inspirations such as narration (storytelling) and transitional environments. Such metaphors could empower users with special abilities and help to familiarize themselves with interactive system's capabilities, behavior and limitations [8]. Many developers do not fully explore and deploy the sensorimotor possibilities of the human body, partly because of methodological and knowledge limitations. 3D interaction techniques design approach should consider the full potential of the human body. Such analysis can be instrumental in designing new or alternative multi-sensory and potentially full body interfaces [9]. When analyzing human potential, an assistive technology design perspective can aid: it can offer a stringent test environment to uncover issues and provides a different view on design by looking at human potential [10]. A number of issues come to mind when analyzing gaps in 3DUI design, though the full breadth and depth of issues is still not yet fully understood. Assistive technology has touched upon (3D) user interface design in numerous cases, including the usage of speech recognition, gestures, eye-tracking, and brain-computer interfaces. As such, it can have an extended impact on 3DUI design in an additional way by offering a different design perspective, methodology and set of user needs.

Recent advent of natural user interfaces based on motion and gesture detection, for example the Microsoft Kinect, has outrun software testing research. This leaves a rapidly growing domain of software ranging from entertainment to medical applications without suitable test automation techniques. To address this issue one can use a technique that automatically tests such applications by generating test cases based on statistical model trained on a corpus of common gestures [11].

III. Conclusion

There are many interesting areas suitable for research in the field of virtual reality systems. Despite the great progress in advanced 3D virtual reality interfaces, many of them are designed for short-term use and experience. 3DUI design ergonomics is still an issue often handled insufficiently. Partly grown out of a strong computer science community with less experience in ergonomics, the design of physical devices as well as the motion patterns afforded does not necessarily reflect ergonomics best practice. Longer duration usage is therefore extremely tiring for the users. This is ultimately affecting their usage in our everyday life.

Main design trend of systems utilizing virtual reality technologies in the past was based on tightening the user to the system environment. This was primarily due to the state of currently available technologies in particular time period. However, progress in the field of virtual reality technologies closely tied to users nowadays, enables us to address this problem from the opposite direction. Modern virtual reality systems need and can be designed in the context of the user located in some specific, highly ergonomic environment. User and his capabilities are closely linked and defined by such environment. We should not force user to adapt to the system, but rather force the system to adapt to user`s needs and capabilities in order to solve defined tasks.

Focus of our future research is on effective utilization of virtual reality systems based on modern advanced 3D interfaces with respect to the user and environment they are deployed into. One of the main goals is wider acceptance of such systems in an enterprise environment.

Acknowledgment

This work is supported by KEGA grant project No. 083TUKE-4/2015: Virtual-reality technologies in the process of handicapped persons education.

References

[1] Weiyuan Liu, "Natural user interface- next mainstream product user interface," Computer-Aided Industrial Design & Conceptual Design (CAIDCD), 2010 IEEE 11th International Conference on , vol.1, no., pp.203,205, 17-19 Nov. 2010

[2] Bowman, D.A.; Coquillart, S.; Froehlich, B.; Hirose, M.; Kitamura, Y.; Kiyokawa, K.; Stuerzlinger, W., "3D User Interfaces: New Directions and Perspectives," Computer Graphics and Applications, IEEE , vol.28, no.6, pp.20,36, Nov.-Dec. 2008

[3] Bruder, G.; Steinicke, F.; Sturzlinger, W., "Effects of visual conflicts on 3D selection task performance in stereoscopic display environments," 3D User Interfaces (3DUI), 2013 IEEE Symposium on , vol., no., pp.115,118, 16-17 March 2013

[4] Lubos, P.; Bruder, G.; Steinicke, F., "Analysis of direct selection in head-mounted display environments," 3D User Interfaces (3DUI), 2014 IEEE Symposium on , vol., no., pp.11,18, 29-30 March 2014

[5] Kopper, R.; Bacim, F.; Bowman, D.A., "Rapid and accurate 3D selection by progressive refinement," 3D User Interfaces (3DUI), 2011 IEEE Symposium on , vol., no., pp.67,74, 19-20 March 2011

[6] Teather, R.J.; Stuerzlinger, W., "Pointing at 3D targets in a stereo head-tracked virtual environment," 3D User Interfaces (3DUI), 2011 IEEE Symposium on , vol., no., pp.87,94, 19-20 March 2011

[7] Xiaodan Jiao; Hui Deng; Feng Wang, "An Investigation of Two-Handed Manipulation and Related Techniques in Multi-touch Interaction," Machine Vision and Human-Machine Interface (MVHI), 2010 International Conference on , vol., no., pp.565,568, 24-25 April 2010

[8] Mostafa, A.E.; Sharlin, E.; Sousa, M.C., "Poster: Superhumans: A 3DUI design metaphor," 3D User Interfaces (3DUI), 2014 IEEE Symposium on , vol., no., pp.143,144, 29-30 March 2014

[9] Kruijff, E., "Human-potential driven design of 3D user interfaces," Artificial Reality and Telexistence (ICAT), 2013 23rd International Conference on , vol., no., pp.129,136, 11-13 Dec. 2013

[10] Kruijff, E., "3DUI design inspired by assistive technology," Virtual and Augmented Assistive Technology (VAAT), 2014 2nd Workshop on , vol., no., pp.7,11, 30-30 March 2014

[11] Hunt, C.J.; Brown, G.; Fraser, G., "Automatic Testing of Natural User Interfaces," Software Testing, Verification and Validation (ICST), 2014 IEEE Seventh International Conference on , vol., no., pp.123,132, March 31 2014-April 4 2014

Detection of Emotion Words through Web-based Application

[1]*Martina TARHANIČOVÁ (3rd year)*
Supervisor: [2]Kristína MACHOVÁ

[1,2]Dept. of Cybernetics and Artificial Intelligence, FEI TU of Košice, Slovak Republic

[1]martina.tarhanicova@tuke.sk, [2]kristina.machova@tuke.sk

***Abstract*— Emotion detection is one of the actual research topics. Despite a lot of research is done in facial expression recognition we have decided to focus on detection of emotion in text. We have begun with simple standalone application to see how precise results we will obtain. Seeing that results were average, we proposed improvements which were implemented in our web-based application. We can say that this paper offers the solution to creating dictionaries. Such dictionaries can be used beyond the scope of this work. In conjunction with other sources of information, it can create a complex system for an autonomous behavior of robots.**

***Keywords*— crowdsourcing, emotion detection, human-robot interaction, sentiment analysis**

I. Introduction

Analyzing sentiment and detection of emotion in a text is still popular task these days. As was mentioned in [1], authors focus on using emotion detection in human-robot interaction. Human-robot interaction is the study of interaction dynamics between humans and robots, a multidisciplinary field that includes engineering (electrical, mechanical, industrial and design), computer science (human-computer interaction, artificial intelligence, robotics, natural language understanding and computer vision), social sciences (psychology, cognitive science, communications, anthropology and human factors) and the humanities (ethics and philosophy).

According to [2] author, is not aware of any system in robotics which allows a robot to use information about emotion in a text to act accordingly. Generally experiments in human-robot interaction are conducted using the Wizard of Oz technique, which means that the robots are not acting autonomously, but they are tele-operated by a human [3].

Our goal is to move from robots that do not interact with humans, are programmed to do a specific task and are unable to learn new things from humans to robots that interact with humans and are able to acquire new knowledge and use it productively.

II. Initial status

There are four major approaches to detecting emotions in a text: corpus-based methods, machine learning methods, knowledge-based methods and hybrid methods. In this paper, we will focus on corpus-based methods.

As for emotions, three major directions in affect computing are recognized: categorical/discrete, dimensional and appraisal-based approaches. The categorical and dimensional approaches are the most commonly used models for automatic analysis and prediction of affect in the continuous input. In this paper, we are using categorical approach. To be more specific, we are using Eckman emotion model that identifies the six emotion: happiness, sadness, anger, disgust, surprise and fear.

As for language, both, Slovak and English texts were used for experiments. As will be explained later, we created two application from which the first one used texts in Slovak and the second application is independent of the selection of language.

III. Current work

From our scheme presented in [1] we are currently working on "DATA ACQUISITION" and "CLOUD" blocks. Data acquisition block consists of application and learning algorithm. Cloud block consists of data processing and emotion detection. At the moment name of the second block is not reflecting the reality because we choose not to use cloud but rather server.

We made two application that their main goal was to correctly identify emotion in text.

A. First Application

We created application Fig. 1 to detect emotion in the short text about sport [4]. The database of 100 sports news from *www.sport.sk* website has been created to test our classification method. Firstly two experts were asked to categorize text to emotion category. For automated detection of emotion, we used emotion dictionary created by students in our university. The dictionary consists of 19 000 Slovak words. Each word is characterized by its part of speech, polarity (in the range -3 to 3) and emotion.

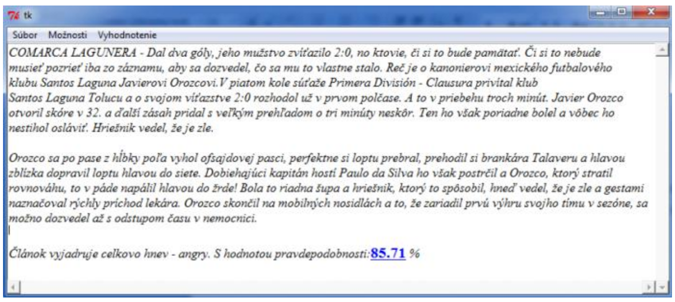


Fig. 1 Example of evaluated text

We obtained the result shown in Table 1. According to the table, we can say that our application agreed with experts mostly in emotion disgust, anger, and happiness. On the other hand, emotions fear, surprise and especially sadness were not categorized with high reliability. Also, experts between themselves did not unambiguously agreed on emotion categories. For more details on obtained results see [5].

We deduced that due to two reasons the accuracy was low. Firstly the size of the dictionary was not sufficient. Dictionary did not contain words required to categorize sports news. Both experts are individuals with their own perception of words. From this fact, we can deduce the need for personalization. By personalization, we mean that we all are individuals and cannot generalize triggers for specific emotions (one thing may be funny for one but sad for another).

To solve mentioned problems we came up with the second application.

Table 1 Results of categorizing text according to emotion from application and both experts.

	Disgust	Anger	Happiness	Sadness	Fear	Surprise
Application	6	5	41	5	2	33
Expert #1	3	9	44	11	8	25
Expert #2	2	5	44	18	3	28

B. Second Application

We have two ways of annotating words for further natural language processing research. One is hiring an expert to manually annotate (label) words. However, this approach is time-consuming, expensive and tedious. That presents a second method attracting audiences recently, crowdsourcing. We break the problem into smaller chunks (sub-problems) and let the crowd solve it. It's more efficient, often more accurate and quicker. Now the question is how to further implement crowdsourcing. Crowdsourcing takes many forms that require different motivational methods to achieve the end goal of annotation.

In our second application, we decided to use crowdsourcing as a main way of creating dictionaries. We created a web-based application, shown in Fig. 2, which allows anybody to input any text. At the moment, we have around 150 Aesop Fable. After saving fable in our system, everybody can access the story. Every word in every story can be afterwards marked by anyone for its polarity in range between 5 (to be strongly positive) to -5 (to be strongly negative) and emotion (six categories: happiness, sadness, anger, disgust, surprise and fear.

To solve the problem with personalization users can use this application after they register. This will prevent any unwanted inputs, same as deleting texts and marking unreasonable words. Registered users have also option to download the dictionary in JSON a CSV format.

IV. FUTURE PLANS

Future plans are the implementation of part of speech tagger and implementation of machine learning algorithms for evaluating emotion for given text.

Part of speech tagger can indicate nouns, adjectives, verbs, and interjections. These words are usually charged with emotion. Marking such words can help users not to overlook them.

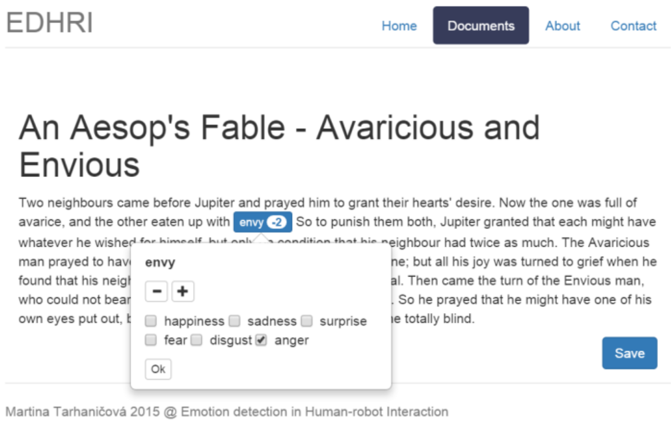


Fig. 2 Marking words with appertaining emotion and polarity

After sufficient evaluating of emotion in a text, we could be able to implement this solution on the robot. As we mentioned above, we are currently detecting emotion in Aesop fables. In one of our experiment, we plan to give children robot that will read stories. While reading the Aesop fable the robot will be adapting its voice and movement to the emotion in the text. We are expecting that children's attention will be higher in comparison with robot reading the fable without any expression of emotion.

V. CONCLUSION

Presented work introduced a new way of creating a dictionary for further analysis of emotion and polarity of words. Firstly standalone application was made for detection of emotion. The results were insufficient, therefore, there was need to reevaluate the way of making a dictionary. The decision to make crowdsourcing application was made. Every user of our application is able to make besides one general dictionary his own dictionary. This leads to the creation of personalized dictionaries. On top of that second application is web-based therefore it is always online and there is no need for additional installation.

ACKNOWLEDGMENT

The paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF and research supported by the National Research and Development Project Grant 1/0667/12 "Incremental Learning Methods for Intelligent Systems" 2012-2015.

REFERENCES

[1] M. Tarhanicova, K. Machova, and P. Sinčák, "Computers Capable of Distinguishing Emotions in Text," in *Emergent Trends in Robotics and Intelligent Systems*, vol. 316, Springer International Publishing, 2014, pp. 57–64.

[2] M. Tarhaničová, "Sentiment Analysis in Human-robot Interaction," Technical University of Kosice, 2014.

[3] G. Magyar, "Social Learning for Personalization in Human-Robot Interaction," Technical University of Kosice, 2015.

[4] P. Kováč, K. Machová, and M. Tarhaničová, "Aplikácia slovníkového prístupu pri analýze sentimentu," in *Electrical Engineering and Informatics 5*, TU, 2014, pp. 606–612.

[5] P. Kováč, "Aplikácia slovníkového prístupu pri analýze sentimentu," Technical University of Kosice, 2014.

Device-to-device routing protocol in multihop communication

[1]*Martin MATIS*

[1]Dept. of Electronics and Multimedia Communications, FEI TU of Košice, Slovak Republic

[1]martin.matis@tuke.sk

***Abstract*—In this paper are described possible solutions for communication in areas without infrastructure based on multihop transfer of message. Maintenance of communication is very important in those areas, because natural disaster or target attack can destroy operative infrastructure. After these incidents, by rescue operations, is very helpful, when many mobile devices as smartphones, tablets or laptops are possible to use for transfer messages, pictures or voices in those unexpected and critical situations about actual state in the epicentre of the disaster.**

There may be used a multihop communication in which are proposed many new approaches and routing solutions, how to create and maintain the communication. After those analyses we want to create a routing protocol, which will combine advantages of Mobile Ad-hoc network (MANET) and Delay/Distribution Tolerant Network (DTN) routing solutions. A pattern for MANET routing will be a DSR routing and for DTN routing will be a single copy transfer based on social behaviour between nodes. We expect that this type of routing will provide a transfer in areas without infrastructure with well-connected areas and/or sparse areas with independent level of movement and velocity.

***Keywords* — Device-to-device routing protocol, MANET routing, DTN routing, intermitted communication, multihop communication, communication without infrastructure, mobility model.**

I. INTRODUCTION

Communication and transfer of information are one of the most important things between people. It is not a problem in our society to use global infrastructure network or radio links network, which is provided with some organizations or administrators. For this transfer of information is mostly used equipment of daily use, such as smartphones, laptops or tablets.

This system of communication is great till infrastructure is correct and functioning. In the situation like natural disaster, target attack and then rescue operation when all infrastructure on the small area, in terms of tens of kilometres, is destroyed, people are unconnected in this area and outside from it. In this case is impossible to use available devices of daily use. Multihop communication can solve the problem, how to keep communication in areas without infrastructure with the same devices, which lost connection in traditional infrastructure network. For example, two ways are available for the transmission of information (messages, pictures, videos, voices, and act.), using a) MANET routing and b) opportunistic or DTN routing. It depends on many factors, for example velocity and density of mobile nodes. Which method is better for the use depends on these factors, because both of them have many advantages and disadvantages [3], [5], [7], [9], [11], [13].

We want to propose a routing protocol, which will contain advantages of both routing methods and will ensure transfer of information between devices without infrastructure and with constantly changing connectivity between devices. It allows the creation of several islands with good connection inside and worse connection outside of them. It is expected that the delivery will be more successful in terms of higher probability of delivery, but with a higher delay.

II. COMMUNICATION WITHOUT INFRASTRUCTURE

Communication between people is matter of course in our society during everyday life. There are many supported devices which can transfer data by supported applications to everyone. The problem begins, when the infrastructure is unusable for various reasons. For example the provider stops to support the services, the infrastructure is destroyed by natural disaster or a group of people is in an area with no cell coverage. There are many possibilities how to communicate in those areas. The easiest way is using a radio transmitter. However, not all the people own or carry it all the time. A better solution is utilization of devices which are used daily by people and which can use standard technologies and wireless interfaces, such as WiFi, Bluetooth, NFC etc. These devices will be able to create other type of network based on multihop communication. Therefore it can and it knows to create a path from source to destination using only other nodes in the neighbourhood based on many criterions which are specified by the routing protocols. This network is composed from mobile nodes and the type of movement is very important for the selection of optimal routing solution, which will provide finding a path and delivering a message [1], [10], [11].

III. MULTIHOP COMMUNICATION

A. Movement and velocity

Mobile nodes have some patterns of movement. We can identify whether the node was walking, running around or went in the vehicle based on information about movement and velocity.

The movement can be a) random, b) forward learned or

c) influenced by social behaviour. Movement and velocity can form a network, which will be dense or sparse. A well-connected network or many well-connected separate sub networks (islands) without mutual radio range as well as individual devices without neighbours can be created there.

We can know which type of routing will be the optimal solution for successful transmission based on this movement and velocity information [9], [10].

B. Mobile Ad-Hoc Network (MANET)

MANET is a collection of mobile nodes, which forms a temporary and dynamic network without aid of centralized administration or standard support devices regularly available, such as conventional networks. These nodes generally have a limited transmission range and so each node seeks the assistance of its neighbouring nodes in forwarding packets and hence the nodes in MANET can act as both routers and hosts. Thus the node may forward packets between other nodes as well as run user applications. These types of networks are suitable for situations in which either no fixed infrastructure exists or deploying network is not possible. Mobile nodes will often be battery powered, which limits the capacity of CPU, memory and bandwidth. This will require network functions that are resource effective. MANETs have found many applications in various fields like military, emergency, conferencing and sensor networks. Each of these application areas has their specific requirements for routing protocols.

The unique feature of these protocols is their ability to trace routes in spite of a dynamic topology. In the simplest scenarios, nodes may be able to communicate directly with each other. However, MANET has to also support communication between nodes that are only indirectly connected by a series of wireless hops (multihops) through other nodes.

In general, MANET is a network where every node is potentially a router and every node is potentially mobile. MANET is not a traditional wired network because of presence of wireless communication and mobility, which requires that the routing protocols used in MANET should be based on new and different principles [1], [2], [8].

MANET is very practical solution for an area without infrastructure with many advantages like: a) no infrastructure and lower cost, b) mobility, c) decentralized and robust, d) easy to build paths and spontaneous infrastructure; but on the other side many disadvantages like: a) higher error rate, b) lower data rate, c) dynamic topology and scalability, d) security, e) high velocity [14].

Routing solutions for MANET

The main idea for routing protocols in MANET is to correctly and effectively estimate the path between a pair of nodes in the network. The routing protocols are created in view of the QoS parameters [1].

MANET network is dynamically changing network which always creates new connections between nodes and loses the old connections. A problem is arisen, when standard routing protocols fail and therefore begin to create a new variety of routing protocols for MANET networks. For propose of MANET routing protocol should take into account a) the environment in which it will be largely deployed, b) the bandwidth required for the run, c) the velocity of movement, d) where it will be implemented and e) including some QoS parameters, such as energy consumption, exploiting width bandwidth, computational efficiency, reliability and safety. Routing protocols for MANET can be categorized into tree basic groups [2]:

- *Reactive (Source-initiated)* – it represents a class of routing protocols where the route is created only when the source requests the route to the destination. The route is created through a procedure of route discovery which involves flooding the network with route request packets (RREQ) from start node to the destination, till it will be found. Once the route is formed or multiple routes are obtained to the destination, the route discovery process comes to the end. A route maintenance procedure maintains the continuity of the route in the time by the source;
- *Proactive (Table-driven)* – these protocols always maintain up-to-date information about routes from each node to every other node in the network. Routing information is stored in the routing table of each node and route updates are propagated throughout the network to keep the routing information as actual as possible. Most of the routing protocols are not suitable for highly dynamic networks due to extra control overhead, which is generated to keep the routing tables consistent and fresh for each node in the network;
- *Hybrid* – this routing scheme combine elements of reactive and proactive protocols. The general idea is the area where the connections change relative slowly and are more amenable to table driven routing, while areas with high mobility are more appropriate for source-initiated approaches.

Another division of routing protocols can be: singlepaths or multipaths. For MANET routing protocols were created much more detailed divisions and many categories based on their property and utilization like: Location-aware, Hierarchical, Multicast, Geographical Multicast or Power-aware.

C. Delay/Distribution tolerant network (DTN)

DTNs [4] have the unique feature of intermittent connectivity, which makes routing quite different from other wireless networks. For example, since an end-to-end connection is hard to keep in MANETs, a store-carry-forward solution is used to deliver the packets to the destination.

Nodes in the DTN choose to store or to forward their data to neighbouring nodes. If there is no neighbouring node in its communication range, the node stores the data and waits until a neighbouring node appears. The node sends the stored data to neighbouring node after finding it. On the other hand, if there are some neighbouring nodes, the node sends the data to some of these neighbouring nodes according to the applied routing method. DTN routing methods do not find the path to the destination, in the simplest term they send data to the neighbour nodes and hope that the neighbour nodes deliver them to destination or bring them closer to the destination. The sophisticated method of sending the data is based on computing of probability whereby the neighbour node is selected as optimal node which can meet the destination node. This method of decisions is used in cases, when the message can be forwarded, or can be stored and wait for better opportunity.

It is good solution how to keep transfer of message in areas without infrastructure and it has many advantages, but also disadvantages, like MANET routing approaches. As the main DTN advantages can be presented: a) transfer method store-carry-forward, b) no need of end-to-end paths, c) high mobility improves success of delivery of message. On the other hand, to DTN belong some main disadvantages, such as: a) increasing time of delivery, b) flooding of network with copies of messages, c) allocation of more resource on devices [4], [5], [9].

Forwarding solutions in DTN

DTN can help to solve a problem, which arise in MANET. Communication without end-to-end connection is helpful in disconnected network. Every routing method has other transfer of data, which is categorized into three groups [10]:

- *Deterministic* (scheduled) – much context information between nodes is known;
- *Force-message* – transmission is provided by special nodes (Message ferry node - MF), which are moving on the known routes, MF picks up, stores, carries and forwards messages;
- *Opportunistic* – it uses every opportunity to transfer a message. It is type of epidemic sending of data.

This transfer of messages is provided by routing protocols, which can be divided into three groups [1], [3]:

- *Redundant* – family of redundant routing protocols creates many copies of message through transfer of data and finds the destination. Those methods improve probability of delivery, but paralyze a network and waste resources (bandwidth, battery, storage, etc.);
- *Utility based* – routing protocols, which belong to this group, take into account, if the transfer to neighbour node is useful or no and evaluate probability of delivery by neighbour based on measurable parameters (battery life, storage capacity, number of contacts with destination, social behaviour between nodes). This group of routing protocols reduces number of transmitted copies of message to one or two what can increase time of delivery, but can reduce probability of delivery;
- *Combination of redundant and utility based* – this group of routing protocols provides delivering of message to destination with combination of both previous methods because of higher probability of delivery and lower time of delivery.

D. Device-to-device communication (D2D)

This type of communication referred as D2D, which combines advantages of MANET and DTN routing, was proposed and begun to be implemented because of a) great natural disaster, which can destroy infrastructure network and/or b) rescue operations in areas without infrastructure and c) improving the delivery of message in every velocity and topology situation and finally d) the transfer can be cheaper. There are used store-carry-forward methods in combination with MANET transfer solutions for successfully delivery of message with a higher delay [3], [7], [10], [11].

Some of D2D routing method was proposed and continuous improved based on combination of both methods, because MANET routing is well functioning in network with low velocity and high density and for DTN network is better a network with higher velocity and sparse density of mobile devices.

Some of the main representatives are [7]:

- *Context-aware Adaptive Routing* – it integrates choosing of nodes and temporary storing of data with proactive MANET routing solutions. Proactive routing is adapted to exchange context information between devices, which can than calculate a probability of delivery [20];
- *Adaptive routing schema* – source node makes a decision, which routing method (for MANET-AODV or for DTN-spray&wait) will use based on accumulated testing packets. Only source node makes the decision and nexthop cannot change this decision about routing type [15],[18];
- *DT-DYMO* – probabilistic model is inserted to DYMO proactive routing protocol based on counters of meeting between nodes. Process of finding a path is modified to choose potential custodian node, which can store a message.
- *HYMAD* – it combines MANET and DTN routing solutions in terms of transfer data using a) MANET, when the network is well-connected and end-to-end paths exist and b) DTN, when the network is divided into many smaller subnetworks without end-to-end paths using spray&wait routing [6], [18];
- *Store and Forward BATMAN* – this routing protocol adds to MANET routing protocol a DTN feature store and forward, which is activated when no paths to destination exist;
- *Storage Aware Routing* – makes a decision between storing and forwarding of packet. OLSR routing method is used to find a path from source to destination. High delay is accepted for this new routing method as standard OLSR routing solution [17];
- *AODV-DTN* – one routing protocol from family of D2D routing methods, which supports bundle protocol (something similar as usage of TCP/IP, but it transfers one big block of message, which can wait on some node, more specification is in RFC 5050[19]). AODV [15], especially finding packets for new path to destination, are changed for the purpose of ascertaining, which supports bundle protocol. If source node supports a bundle protocol and no path by AODV exists, than the transfer will be changed to bundle transfer;
- *DTS-OLSR* – it is D2D routing method, which uses overlay network based on proactive routing called OLSR, which sends periodic update. Only nodes with bundle layer can be used for D2D transfer. The node without bundle has to send a message to other node with bundle layer immediately [16];

- *Robust replication routing* – it estimates a delay and based on it routing protocol chooses MANET part of routing or multipath routing.

IV. Propose of D2D Routing

We want to develop a Combined MANET-DTN routing protocol after analyses of the actual routing methods in mobile network without infrastructure, which is composed from MANET approach - modified reactive protocol DSR and DTN approach based on single copy sending of message based on social behaviour between nodes. The main idea of our proposal is shown in Fig1. The path will be found by modified DSR routing for actual topology in timeslot t_1. Topology will be divided into more subnetworks and our protocol ensures transfer the rest of data to Virtual source (V_s) by partial usable path in t_2. V_s will try or wait for appropriate conditions for the search of a new DSR path in t_3. When V_s, as a new source, will find the path to D, than it will use this path to finish the transfer.

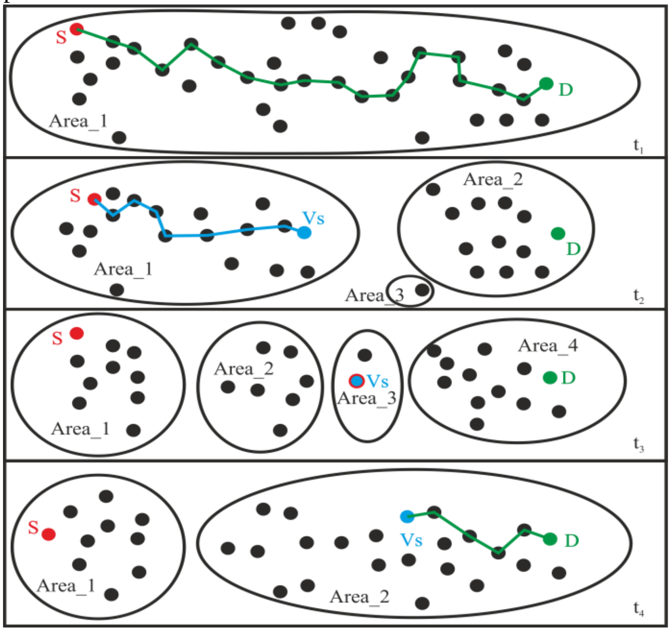


Fig. 1 The main idea of Combined MANET-DTN routing protocol

A. DSR modification

Source (S) finds a path using modified DSR routing based originally on RFC 4728 [12]. Route Request Packet (RREQ) is distributed as broadcast to any node and finds a destination (D). After this finding the S expects Route Replay Packet (RREP) with valid path or paths. We expand transfer information in RREQ about social information, which the S owns about D. The probability of delivery by every node, to which a RREQ arrives, is calculated based on social information. It will be helpful for next usage, when MANET part of routing - modified DSR will crash and DTN mechanism ensures the delivering, but with higher delay.

B. Social information

Calculation of probability of delivery is important because of selection of nexthop, respectively a new virtual source, which will act as a new source. This calculation is performed based on personal and social information (where the person lives and works, how to travel, where he spends a free time) about people, because every device has an owner [13].

V. Conclusion

We want to create new routing protocol for network without infrastructure independent on velocity and connectivity. We want to create a simulation tool, which confirms our solution and compares other routing methods usable for disconnected or occasionally connected network with different level of density and velocity. We want to create some models of mobility for nodes because of implemented social behaviour.

Acknowledgment

The research presented in this paper was supported by the Ministry of Education, Science, Research and Sport of the Slovak Republic under the project VEGA 1/0075/15.

References

[1] S. Basagni, et al., eds. Mobile Ad Hoc Networking: The Cutting Edge Directions. Vol. 35. John Wiley & Sons, 2013.
[2] A. Boukerche, et al. "Routing protocols in ad hoc networks: A survey." Computer Networks 55.13 (2011): 3032-3080.
[3] Ch. Raffelsberger, H. Hellwagner. "Combined Mobile Ad-hoc and Delay/Disruption-tolerant Routing.", 13th International Conference, ADHOC-NOW 2014, Benidorm, Spain, June 2014 Proceedings: 1-14
[4] V. Cerf, et al. "Delay-tolerant networking architecture." RFC4838, April (2007).
[5] K Fall, "A delay -tolerant network architecture for challenged internets." Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications. ACM, 2003.
[6] J. Whitbeck, V. Conan, "HYMAD: Hybrid DTN-MANET routing for dense and highly dynamic wireless networks," World of Wireless, Mobile and Multimedia Networks & Workshops, 2009. WoWMoM 2009. IEEE International Symposium on a , vol., no., pp.1,7, June 2009
[7] C. Raffelsberger, H. Hellwagner, "A hybrid MANET-DTN routing scheme for emergency response scenarios," Pervasive Computing and Communications Workshops (PERCOM Workshops), 2013 IEEE International Conference on , vol., no., pp.505,510, March 2013
[8] A. Boukerche, ed. Algorithms and protocols for wireless, mobile Ad Hoc networks. Vol. 77. John Wiley & Sons, 2008.
[9] V. Athanasios, at al.: Routing in Opportunistic Networks.Springer, 2013.
[10] Ch. Mayer, O. P. Waldhorst. "Offloading infrastructure using Delay Tolerant Networks and assurance of delivery." Wireless Days (WD), 2011 IFIP. IEEE, 2011.
[11] H. Nishiyama, N. Kato, "Relay-by-smartphone: realizing multihop device-to-device communications," Communications Magazine, IEEE , vol.52, no.4, pp.56,65, April 2014
[12] D. Johnson, Y. Hu, D. Maltz. The dynamic source routing protocol (DSR) for mobile ad hoc networks for IPv4. Vol. 260. RFC 4728, 2007.
[13] A. N. Hoang, S. Giordano, A. Puiatti, "Probabilistic Routing Protocol for Intermittently Connected Mobile Ad hoc Network (PROPICMAN)," World of Wireless, Mobile and Multimedia Networks, 2007. WoWMoM 2007. IEEE International Symposium on a , vol., no., pp.1,6, June 2007
[14] T. P. Singh, D. Vikrant, S. Maheshwari. "ADHOC NETWORKS: AN ANALYTICAL OVERVIEW." January 2012.
[15] D. R. Samir, M. B. Elizabeth, E. P. Charles. "Ad hoc on-demand distance vector (AODV) routing." (2003).
[16] T. Clausen, P. Jacquet: "Optimized Link State Routing Protocol (OLSR)." (2003).
[17] L. Sung-Ju, W. Su, M. Gerla. "On-demand multicast routing protocol in multihop wireless mobile networks." Mobile Networks and Applications 7.6 (2002): 441-453.
[18] T. Spyropoulos, P. Konstantinos, S. R. Cauligi: "Single-copy routing in intermittently connected mobile networks." Sensor and Ad Hoc Communications and Networks, 2004. IEEE SECON 2004. 2004 First Annual IEEE Communications Society Conference on. IEEE, 2004.
[19] L. K. Scott, B. Scott. "Bundle protocol specification." (2007).
[20] M. Musolesi, S. Hailes, C. Mascolo. Adaptive routing for intermittently connected mobile ad hoc networks. In Proc. WoWMoM '05. Pp. 183–189. IEEE, 2005.

Diagnostic and Backup System of an Aircraft Turbojet Engine

[1]Ladislav NYULÁSZI (2st year)
Supervisor: [2]Ladislav MADARÁSZ

[1,2]Dept. of Cybernetics and Artificial Intelligence, FEI TU of Košice, Slovak Republic

[1]ladislav.nyulaszi@tuke.sk, [2]ladislav.madarasz@tuke.sk

***Abstract*—The paper summarizes the work and results obtained in the previous year of the PhD study. The research aims to increase the reliability and safety of an aircraft turbojet engine by diagnostic/backup system. This article describes the process of proposal and implementation of the diagnostic and backup system, which uses majority method (diagnostic method) and mathematical models (backups) calculated by methods of experimental identification. As an object of research was chosen the small turbojet engine ISTC-21V located in the Laboratory of Intelligent Control Systems of Jet Engines (LIRS LM).**

***Keywords*—Backup, diagnostics, experimental identification, small turbojet engine, voting method.**

I. Introduction

The increase in the complexity of aircrafts (their structure contains a greater number of elements) also leads to increased failure probability of an aircraft components. Consequently, any potential fault of a turbojet engine could cause even more extensive damages, therefore the issue of reliability and safety come to the fore [1]. With these concepts are also related activities diagnostics and backup [2]. The main aim of a diagnostic/backup system of an aircraft turbojet engine, and also my motivation for solving this issue, is to ensure accurate fault detection and isolation, prevent expensive damage to engine parts, avoid failures that can lead to catastrophe and reduce costs of maintenance.

The article deals with the application of the diagnostic and backup system of a small turbojet engine, which contains the proposal of the structure and its implementation in the chosen software environment. The main focus is oriented on a real-time diagnostics which function is to detect sensor errors and isolate them, so they do not affect control.

II. Initial State

In the previous SCYR article [3], we described integrated system health management for aircraft turbojet engine [4]. It should be able to ensure decrease of the probability of catastrophic failure, detection of errors and their isolation, prediction of future failure, etc. All these goals can be achieved by three health management techniques, which are:

- Diagnostics and backup
- Prognostics
- Life extending control

As the first, we decided to focus on diagnostic and backup methods.

III. Solved Tasks in the Previous Period

In the following section are described tasks solved in the last year of my PhD study.

A. Proposal of Diagnostic Modules

The overall diagnostic/backup system [5] consists of diagnostic modules for each of the selected parameters of the small turbojet engine ISTC-21V [6]. We have created a proposal of diagnostic modules, where inputs represent signals from sensors. Those either directly enter the selected diagnostic method (majority method) [7] or they are used to obtain modeled values of parameter through models acquired by methods of experimental identification [8]. The output of the majority method enters the control circuit. As an example, the proposed architecture of diagnostic module for parameter p_{2C} (pressure of air) is shown in Fig. 1.

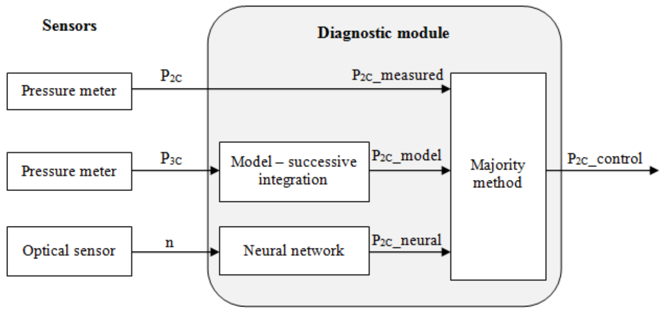


Fig. 1. The structure of the diagnostic module for parameter p_{2C} [5]

B. Creation of Experimental Models

Experimental identification represents the process leading to the compilation of a mathematical model of a certain system by using measured data [8]. We used three methods of experimental identification (method of successive integration, area method and artificial neural network) to create models of ISTC-21V engine's parameters. These will serve as backups which can replace sensor errors with correct data. We have calculated experimental models using programming environment MATLAB and implemented them in Simulink. For example, the scheme of the neural model of parameter p_{2C} is shown in Fig. 2.

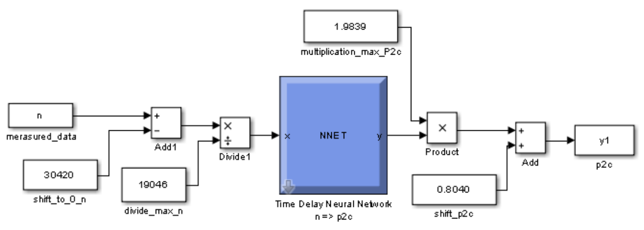


Fig. 2. The scheme of the neural model of parameter p_{2C} in Simulink

C. Implementation of the Diagnostic Method

From different diagnostic methods, we have chosen majority (voting) method [7] as the best for the purpose of diagnostics of the ISTC-21V engine. There are three signals, which enter the majority method. One input is measured data (I_1) of the selected parameter and other two are model values (I_2, I_3) of this parameter (backups) calculated through experimental models. Each pair of these signals is compared in a block of pair comparison on the basis of permissible fault values. After this comparison, average value I_P is calculated from the signal values that do not overstep the maximum allowed error size. Then I_P is transferred through the function of $qi(\varepsilon_i)$ (maximum allowed deviation from error residuum ε_i), which decides if input provides correct data or not (signalization V_i). In the end, arithmetic mean I_C is computed from errorless input values and enters the control circuit.

Based on the described process, we designed and implemented the voting method in LabVIEW environment. It communicates with experimental models created in Simulink through OPC interface.

D. Testing of the Diagnostic and Backup System

We have tested the functionality of the implemented system for real-time diagnostics and backup in the Laboratory of Intelligent Control Systems of Jet Engines (LIRS LM) during the operation of the small turbojet engine ISTC-21V. Tests were run during the errorless engine operation and also when a two types of faults occur (failure of individual inputs and presence of random input values) [5].

Results have shown that the diagnostic/backup system is working properly and can eliminate these types of failures. Accuracy of the implemented system from the view of created experimental models for individual engine parameters (temperatures T_{1C}, T_{2C}, T_{3C}, pressures p_{2C}, p_{3C}, fuel flow supply Q_{Pal} and speed of engine's compressor n) is shown in Table I.

IV. Conclusion and Future Work

In this article, we have shortly described the results of the work done in the previous year. The main objective was to increase reliability and safety of the small turbojet engine ISTC-21V through the creation and implementation of the diagnostic/backup system. Based on the tests performed is shown that system is able to detect sensor faults, isolate them and replace with correct values. As can be seen in Table I, on the basis of selected parameter, the accuracy is from 1 to 5 %.

Future research will be focused on improvement of the proposed diagnostic/backup system by decreasing amount of error. It will also be important to ensure that the system is able to distinguish between sensor fault and engine component fault.

Further research will focus on *prognostics* for prediction of future failures and on *life extending control* for engine parts damage reduction as other techniques of the integrated system health management [3], [4].

TABLE I
Comparison of the Created Experimental Models Based on the Amount of Mean Absolute Error (MAE)

Parameters	Experimental model	MAE	MAE [%]
T_{2C} [°C]	Neural network	5.0788	3.4026
	Succesive integration	5.0359	3.3008
T_{3C} [°C]	Neural network	11.7036	1.2377
	Succesive integration	19.9546	2.1120
T_{4C} [°C]	Neural network	7.4042	1.4834
	Succesive integration	8.9987	1.8058
p_{2C} [At]	Neural network	0.1046	4.5225
	Succesive integration	0.0425	1.8685
p_{3C} [At]	Neural network	0.0770	5.0378
	Succesive integration	0.0687	4.3637
Q_{Pal} [l/min]	Neural network	0.0255	2.5638
	Area method	0.0307	3.0958
n [ot/min]	Neural network	742.0916	1.6726
	Succesive integration	596.4056	1.3331

Acknowledgment

The work presented in this paper was supported by KEGA under Grant No. 014TUKE-4/2015 – "*Digitalization, virtualization and testing of a small turbojet engine and its elements using stands for modern applied lecturing*".

References

[1] R. Isermann, "Fault-Diagnosis Systems: An Introduction from Fault Detection to Fault Tolerance," Springer – Verlag Berlin Heidelberg, 2006, 475 pp., ISBN 3-540-24112-4.
[2] L. Madarász, et al., "Situational Control Modeling and Diagnostics of Large Scale Systems," in Toward Intelligent Engineering and Information Technology, ISBN 978-3-642-03736-8, ISSN 1860-969X, 2009, Springer – Verlag Berlin Heidelberg, pp. 153-164.
[3] L. Nyulászi, "Health Management Techniques for Aircraft Turbojet Engine," in SCYR 2014: Proceedings from conference: *14th Scientic Conference of Young Researchers*, 2014, pp. 187-190.
[4] D. L. Simon, S. Garg, G. W. Hunter, T.-H. Guo, K. J. Semega, "Sensor Needs for Control and Health Management of Intelligence Aircraft Engines," NASA/TM—2004-213202, ARL-TR-3251, 2004.
[5] L. Nyulászi, L. Madarász, R. Andoga, V. Gašpar, "Redundant Backup and Diagnostic System of MPM-20 Engine," in CINTI 2013: *14th IEEE International Symposium on Computational Intelligence and Informatics*: Proceedings: November 19-21, 2013, Budapest. - Piscataway: IEEE, 2013, PP. 43-47, ISBN 978-1-4799-0195-1.
[6] Főzo, L., et al., „Description of an intelligent small turbo-compresor engine with variable exhaust nozzle,". in SAMI 2015: *13th IEEE International Symposium on Applied Machine Intelligence and Informatics*:, 2015, pp. 157-160. ISBN 978-1-4799-8220-2.
[7] L. Madarász, et al., "Komplexný výskum efektívnosti a inovácia technológie skúšok malého prúdového motora (4). In: AT&P Journal, vol. 20, no. 12, 2013, pp. 49 – 52. ISSN 1335-2237.
[8] L. Nyulászi, L. Madarász, "Experimental Identification of the Small Turbojet Engine MPM-20," in CINTI 2014: *15th IEEE International Symposium on Computational Intelligence and Informatics*: Proceedings: November 19-21, 2014, Budapest. - Piscataway: IEEE, 2014, pp. 497-501, ISBN 978-1-4799-5338-7.

Diagnostic tool for movement of industrial robot

[1]*Peter PAPCUN (4th year)*
Supervisor: [2]Ján JADLOVSKÝ

[1,2]Dept. of Cybernetics and Artificial Intelligence, FEI TU of Košice, Slovak Republic

[1]peter.papcun@tuke.sk, [2]jan.jadlovsky@tuke.sk

Abstract — **Contribution describes progress on dissertation research. Title of dissertation is Modeling and optimization of robotic and technological production lines. Dissertation deals with optimizing effector speed of an industrial robot to a required accuracy. The resulting system input will by parametric equation of trajectory, required speed or accuracy and output will be movement of robot's effector according to specified requirements.**

Keywords — **industrial robot, diagnostic, speed, accuracy.**

I. INTRODUCTION

Article is dedicated for dissertation. This dissertation is about accuracy and speed of industrial robot's effector on specified trajectories. These trajectories will be specified to robot by parametric equation. After specifying trajectory, industrial robot will move on that trajectory with required speed and accuracy. During movement on specified trajectory, robot will not use own program for movement, but program developed by dissertation, which this article describes.

II. INITIAL STATE

During the diploma thesis [1] processing, two problems were found. Diploma thesis was not solving these problems. The first problem was speed of industrial robot's effector. Despite the fact that the speed was set to a specific value, the robot was not moving with this speed. This problem solved publication Optimizing industry robot for maximum speed with high accuracy [2]. The next problem was small option of movement trajectories. Programmer of industrial robots can mostly use only three types of trajectories and these trajectories are ellipse, straight line, point-to-point, specifically. When diploma thesis was being solved, mentioned trajectories was not sufficient. That was reason for created functions for added movement described in publication Control of robot integrated in flexible production line [3]. A small selection of trajectories and necessary of functions for new trajectories leaded to next research. This research described contribution Industrial robot optimization for required accuracy and speed [4]. This contribution designed function where input is parametric equation of curve (trajectory) for industrial robot movement.

Next aim was to associate mentioned results to complex system, which is described in thesis for dissertation exam [6]. This complex system is containing:

1. Functions to robot movement along a defined trajectory using the parametric equation of the curve (trajectory),
2. Optimize the speed of the robot's effector by set the iteration step [4], [6],
3. Diagnostic tool for diagnosis speed and accuracy.

This system is tested on an industrial robot MELFA RV-2SDB and on industrial robots of laboratory model with name Robotized workplace. These models are DCAI (Dept. of Cybernetics and Artificial Intelligence) models.

An important task was calculated the direct and inverse kinematics (mathematical model) to these industrial robots. Direct and inverse kinematics was calculated in publication Mathematical model of robot Melfa RV-2SDB [7]. When mathematical model was being created, then university textbooks [8], [9] and [10] were helpful.

When the mathematical model of industrial robots was known, then it was necessary to create a mentioned diagnostic tool. The first and basic design of tool was discussed in [2], [5] and [6]. This diagnostic tool is near at the end of realization and so far it has not been written for its any publication.

Control and diagnostic of industrial robot are distributed therefore was used methods described in [11].

III. DESCRIPTION OF THE TASKS SOLVED IN PREVIOUS YEAR

University Science Park Technicom was processed during last year.

Diagnostic tool can be tested only with industrial robot. Therefore has to be programmed robot controller. Various functions of robot movement were programmed through parametric expression of curves (trajectories), which was designed in [4]. These functions were programmed in programming language Melfa Basic V in integrated development environment RT Toolbox 2. Diagnostic tool can be tested after programming these functions along with basic control of speed and accuracy of robot's effector. Designed and programmed basic algorithms of control of speed and accuracy were tested with diagnostic tool together.

Diagnostic tool is programmed in programming language C# in development environment Visual Studio. This application monitors the actual speed and acceleration of the robot, calculated from the current position of the robot's effector. Application receives information about the current position of robot's effector from robot controller through RS-232 interface. To application can be specified parametric equation of curve. When parametric equation and current effector location are known, then it can be calculated position error. This error can be seen in application on graph and in real time as number. The application also monitors the speed error, because the desired speed is known.

On defense of thesis for dissertation exam and on some conferences was recommended to add secondary sensors at movement diagnostic of robot's effector.

Secondary sensor of movement was chosen black-and-white (grayscale) camera Guppy F-503-B. This camera with new algorithm of image recognition in diagnostic tool recognizes position of effector and plots this position to application. Effector 3D position was plotted in orthogonal projection by diagnostic tool (position of the observer was not parallel to any basic plane – XY, XZ, YZ), but after the addition of image recognition algorithm was decided that the plotting position in space will be the view from the top (position of the observer is parallel to basic plane XY).

Effector position is plotted three times into application of diagnostic tool. The first time application is plotting required position (3x3 pixels, blue color), than position detected by sensors of servomotor (1 pixel, red color) and last position detected by camera with algorithm of image recognition (1 pixel, green color). Part of diagnostic tool window is shown on figure Fig. 1 with plotted positions.

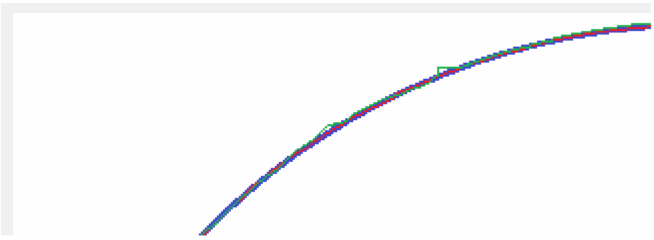


Fig. 1. Part of diagnostic tool window with 3D plotting of effector position

Diagnostic tool works in real time, therefore it was necessary to design algorithm of image recognition to require the least processor's (CPU's) time.

Picture from camera is inserted to RAM memory in a row directly. Access to information about brightness is realized through pointers. Image thresholding is omitted in the resulting solution, because mark on robot's effector has specific interval of brightness and is in shape of rectangle. Mark's first pixel is found by searching in vector (not in the array, because pointers is used) on both sides. Center of mark is calculated on the basis of these two information (the first pixels on both sides), and this point (center) is plotted to application of diagnostic tool. Knowledge of the subject Computer Vision and university textbook [13] for this subject helped in this procedure.

Then it is necessary to synchronize the plotted position of the camera and the servomotor's sensors, this synchronization is made by transformation spatial point to planar point with formulas of this transformation. Calibration is needed at this synchronization. Camera position has to be known in coordinate system of robot and their common coordinate system has to be moved to appropriate place. Mentioned transformation is applied from a theory described in the university textbook [12].

Optimization tools were programmed during last year, which was mentioned in [5]:

1. optimization for required speed,
2. maximum speed optimization for required accuracy,
3. speed optimization for required accuracy and speed.

A detailed description of these three types of optimization is in contribution [4] and [6].

IV. Conclusion

Last year, contribution [5] mentioned three phases of research continuation. These days, the first phase is completed and work continues on the second and the third phases, parallel.

The second phase consists of the following tasks:

- verification of created functions on real model,
- creation and description of methodology for optimizing robot speed in relation to accuracy,
- verification of methodology in terms of laboratory models of the department.

The third phase processes these results into a dissertation.

Acknowledgment

This work has been supported by the Research and Development Operational Program for project: University Science Park Technicom for innovative applications with knowledge technology support, ITMS code 26220220182, co-financed by the ERDF.

References

[1] P. Papcun, *Riadenie robota integrovaného v pružnej výrobnej linke*, Diploma Thesis, 2011, Dept. of Cybernetics and Artificial Intelligence, FEI TU of Košice, Slovak Republic, Košice, 2011, p. 95.

[2] P. Papcun, J. Jadlovský, *Optimizing industry robot for maximum speed with high accuracy*, in Procedia Engineering: MMaMS 2012: Modelling of Mechanical and Mechatronics Systems 2012: November 6^{th}-8^{th} 2012, Zemplínska Šírava, Slovak Republic, Košice, 2012, No. 48 (2012), p. 533-542, ISSN 1877-7058.

[3] P. Papcun, M. Čopík, J. Ilkovič, *Riadenie robota integrovaného v pružnom výrobnom systéme*, in Electroscope – online journal for electrical engineering, Czech Republic, Plzeň, 2012, No. II, p. 9, ISSN 1802-4564.

[4] P. Papcun, *Industrial robot optimization for required accuracy and speed*, in SCYR 2013: 13th Scientific Conference of Young Researchers of Faculty of Electrical Engineering and Informatics Technical University of Košice: proc. Košice: FEI TU, Slovak Republic, Košice, 2013, p.144-147, ISBN 978-80-553-1422-8.

[5] P. Papcun, Industrial robot trajectories, speed and accuracy, in SCYR 2014: 14th Scientific Conference of Young Researchers of Faculty of Electrical Engineering and Informatics Technical University of Košice: proc. Košice: FEI TU, Slovak Republic, Košice, 2014, p.195-196, ISBN 978-80-553-1714-4.

[6] P. Papcun, *Modelovanie a optimalizácia rýchlosti priemyselných robotov*, Thesis for Dissertation Exam, 2013, Dept. of Cybernetics and Artificial Intelligence, FEI TU of Košice, Slovak Republic, Košice, 2013, p. 80.

[7] P. Papcun, *Mathematical model of robot Melfa RV-2SDB*, in SETINAIR: Symposium on emergent trends in artificial intelligence & robotics: September 15^{th}-17^{th} 2013, TU of Košice, Slovak Republic, Košice, 2015, ISBN 978-3-319-10782-0

[8] D. Krokavec, *Spracovanie údajov v robotike*, University textbook, Slovak Republic, Košice, 1985, p. 260, ISBN 85-629-85.

[9] V. Mostýn, V. Krys, *Mechatronika průmyslových robotů*, University textbook, Czech Republic, Ostrava, 2012, p. 201, ISBN 978-80-248-2610-3.

[10] P. Božek, O. Barborák, Ľ. Naščák, V. Štollmann, *Špecializované robotické systémy*, University textbook, Slovak Republic, Bratislava, 2011, p. 224, ISBN 978-80-904766-8-4.

[11] J. Jadlovský, M. Čopík, P. Papcun, *Distribuované systémy riadenia*, University textbook, Slovak Republic, Košice, 2013, p. 216, ISBN 978-80-8086-227-5.

[12] D. Krokavec, *Inteligentné senzorické systémy*, University textbook, Slovak Republic, Elfa, Košice, 1992, p. 96, ISBN 80-7099-210-7.

[13] M. Bundzel, I. Zolotová, *Počítačové videnie v praxi*, University textbook, Slovak Republic, Elfa, Košice, 2013, p. 90, ISBN 978-80-8086-225-1.

Diagnostics of Sensors and Actuators Within Distributed Control System

[1]*Michal KOPČÍK(2nd year),*
Supervisor: [2]Ján JADLOVSKÝ

[1,2]Dept. of Cybernetics and Artificial Intelligence, FEI TU of Košice, Slovak Republic

[1]michal.kopcik@tuke.sk, [2]jan.jadlovsky@tuke.sk

***Abstract*—This article deals with the structure of diagnostics of sensors and actuators within distributed control systems. Proposed structure consists of several layers which can be separated and processed by multiple control systems. This makes diagnostics more transparent and gives a way to share computational load between control systems.**

***Keywords*—Mobile robot, diagnostics, sensor, distributed control system.**

I. Introduction

Diagnostics of dynamic systems is very current topic, whether in system with increased reliability where the error of the sensor can cause big financial loses or a threat to life. Diagnostics is also required in system with poor maintenance access such as satellites, research rovers, etc.

Diagnostics is wide concept which includes detection, localization, identification and isolation of the error, alternatively reconfiguration of the control system.

Diagnostics as a whole should make the system more resistant to errors and in the case of occurrence of the non critical error, the system should be able to continue functioning without losing nature of its function. For proper function there need to be some sort of redundancy in the form of redundant sensors or analytical redundancy.

Since the final form of diagnostics will be implemented on mobile robots which are equipped by microcontroller(s) with limited computational power, the diagnostics should be modular to provide the tool for distribution of processing load between multiple control systems.

II. Initial state

In previous work there were introduced few ways of fusing data from multiple sensors, especially focusing on Multisensor data fusion (MSFDF) using Extended Kalman filter (EKF).[1]. There were also developed prototypes of mobile robots (MR) like MR for MiroSot competition [2] and mobile robot ALFRED [3]

III. Task solved in previous year

Diagnostics of sensor system is quite difficult task, which consists of multiple steps. Based on the analysis the current state of diagnostics, there was created diagnostic system divided into several logic parts (layers), whose together form a pyramid shown in Fig.1 where every layer can be implemented using several different ways. Block diagram of sample diagnostics system with interconnections between layers is shown in Fig. 2.

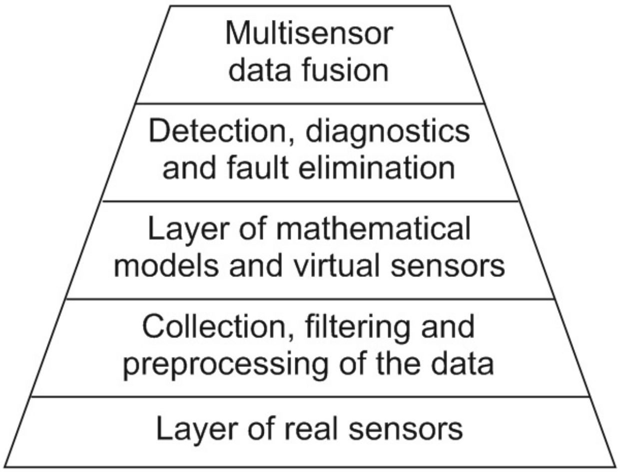


Fig. 1. Proposed structure of diagnostics within distributed control systems.

A. Layer of physical sensors

On this layer there are real (physical) sensors, whose output is analogue signal in the form of voltage or current, or digital signal transfered using communication interfaces (I2C, SPI, UART, CAN, Ethernet, etc.). This layer also includes scaling of the range of sensor to the range of analog to digital (AD) converter. Sensors within mobile robotics can be divided into three sections:

Sensors of internal state are sensing units directly associated with internal states of mobile robot like voltage of the battery, current flowing trough motor coils, linear or angular velocity of the wheels, etc.

Position sensors are used to determine current position and angle of the robot. Position sensor may be subdivided into two categories namely absolute (GPS, accelerometer, compass) and relative (accelerometer, rotary sensors, gyroscope) position sensors.

External environment sensors are measuring units environment outside the mobile robot (distance sensors, environment quality sensors, cameras, etc.).

B. Collection, filtering and preprocessing of the data

Raw data from the sensors often need to be preprocessed, so the control system will know to work with them.In this case the preprocessing of the data is considered filtering (smoothing) and transformation of the data from engineering to physical units.

This layer is based on hardware and software. Hardware part consists of communication buses, hardware filters and AD converters and software part includes software filtering and transformation of the data. Some of the sensor and communication bus errors can be detected on this layer

C. Layer of models and virtual sensors

On this layer there are mathematical models of the system as kinematic and dynamic model, as well as models describing relationships between measured and state values of the system. Mobile robots are described by kinematic and dynamic model.

Based on complexity, dynamic model may include the weight of MR, moment of inertia of MR and wheels (including gears in gearbox), frictions, etc. [4].

Kinematic model describes the relationship between linear or angular velocities of wheels (tracks, legs, etc.) and the position and angle of the MR in space. Kinematic model with association with wheel speed sensors can be used to calculate relative position of the mobile robot. Basic kinematic model can be enhanced to include parameters line slippage between wheels (tracks) and surface.

On this level there are also virtual sensors, which take readings from real sensors and calculate the outputs using some system models [5].

D. Detection, diagnostics and elimination of errors

For diagnostics and to improve reliability of the system, it is advantageous to have multiple sensors sensing the same value, ideally with sensors based on different physical sensing method.

Detection of the errors is done using residues which are next used to determine exact error of the sensor or actuator. Residue is indicator of an error derived from the difference between the results from sensors and models of the system. Analytical methods which uses residues as error indicators are generally called as a method of analytical redundancy.

Diagnostics - its most basic problem is to differentiate between unnatural or faulty state of the sensor and between non standard state of the system. In real dynamic systems such a mobile robots it is difficult task, because dynamics of the system and outer influence may hide actual fault of the sensor. Accurate diagnostics of the sensor fault may obstruct the fact, that multiple sensor faults can be described by similar symptoms [6].

Fault elimination is the process when the fault data from corrupted sensors are eliminated to ensure that incorrect data wouldn't affect the result of measured unit. In the case when the data from specific sensor are used as a feedback for control system, it is necessary to modify (reconfigure) the control system to get the feedback data from another real or virtual sensor.

E. Multisensor data fusion

This layer serves to join data from multiple sources (real or virtual sensors) to create one more robust and precise result than by using just one sensor. For application of MSDF there are several ways, for more informations see [1] or [7].

The proper function of MSDF depends on few conditions, namely the noise in data from sensors should be uncorrelated white noise, median of difference between measured and real values should be close to zero and data from faulty sensors must be eliminated before entering MSDF.

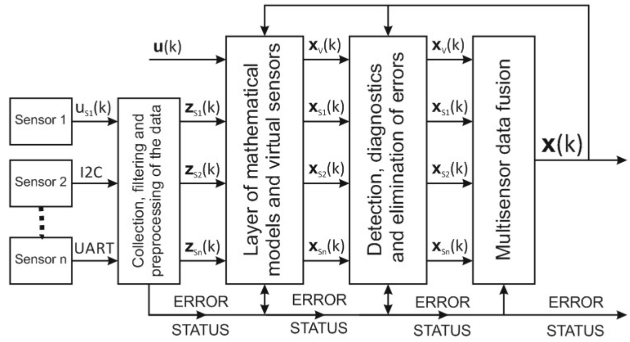


Fig. 2. Block diagram of sample diagnostics.

F. Development and construction of new mobile robots

During last year were developed and constructed several mobile robots for robotic soccer category MiroSot V2 based on the prototype from the past [2]. These MR are more accurate, reliable, lighter, compact and durable than prototype. There was also developed and constructed tracked mobile robot TrackBot which is controlled by Arduino board with custom made sensor and communication shield.

IV. Future Research

Future research will be focused on creating of the algorithms for every layer of proposed diagnostics system. After this phase will follow testing of the created algorithms on the real models of mobile robots, which will consists of three parts:

1) Off-line testing of algorithms and creating of mathematical model based on measured data.
2) On-line testing of algorithms using mobile robot as a remote agent.
3) On-line testing of algorithms with their execution on mobile robot's control system.

Testing of the algorithms of diagnostics will be performed on the mobile robots for robotic soccer category MiroSot V2 and on tracked MR TrackBot.

Acknowledgment

This work has been supported by the Research and Development Operational Program for project: University Science Park Technicom for innovative applications with knowledge technology support, ITMS code 26220220182, co-financed by the ERDF.

References

[1] M.Kopcik, "Multisensor data fusion for differential wheeled mobile robots," in *14th Scientific Conference of Young Researchers of Faculty of Electrical Engineering and Informatics Technical University of Košice*, Herľany, Slovakia, 2014.

[2] M. Kopcik and R. Bielek, "Construction and operating system of robosoccer agents," in *13th Scientific Conference of Young Researchers : May 14th, 2013, Herlany, Slovakia*, May 2013, pp. 80–84.

[3] J. Jadlovský and M. Kopčík, "Basic motion control of differential-wheeled mobile robot alfred," in *Emergent Trends in Robotics and Inteligent Systems*, vol. 316, 2015, pp. 73–80.

[4] J. Šembera and F. Šolc, "Modelování a řízení mobilního robotu s diferenciálním podvozkem," in *ATP Journal PLUS*, 2007.

[5] L. Liu, S. Kuo, and M. Zhou, "Virtual sensing techniques and their applications," in *Networking, Sensing and Control, 2009. ICNSC '09. International Conference on*, March 2009, pp. 31–36.

[6] D. Krokavec and A. Filasová, *Diagnostika dynamických systémov*. Elfa, 2007, no. 1.

[7] B. Siciliano and O. Khatib, *Springer Handbook of Robotics*. Springer, 2008, no. 1. [Online]. Available: http://www.springer.com/978-3-540-23957-4

Elise - Semantic Search & Transcription Correction

[1]*Stanislav Dvorščák (6st year),*
Supervisor: [2]*Kristína Machová*

[1,2]Dept. of Cybernetics and Artificial Intelligence, FEI TU of Košice, Slovak Republic

[1]stanislav-dvorscak@solumiss.eu, [2]kristina.machova@tuke.sk

***Abstract*—The paper summarizes the research realized during the last year. It provides the information about the Elise - Semantic Network determined for searching purposes and its performance optimization done by white noise modulation. It also provides the information about the customization of the Elise Semantic Network for a transcription errors correction.**

***Keywords*—misspelling support, voice transcriptions, information rebuild, white-noise**

I. Introduction

The court of Slovak Republic uses the transcription of the court hearing. Those transcriptions contain different misspelling errors, semantic errors and other types of errors.

Current solutions for this domain use different ways based on a vocabulary continues speech recognition. Those solutions are mainly based on a DBN – deep belief network, a HMM – hidden Markov model, or hybrid systems like a DBN-HMM – context dependent solutions [5].

We are developing Elise – Semantic Search Web Agent. This project works with natural language and it is build up a semantic network model. It supports misspelling processing natively. Currently a founded documents are the output of the semantic network. It works with words, phrases, sentences and paragraphs as building blocks, and it represents good base for such kind of a problem solving [6].

II. Semantic Network

A core's implementation is realized via the semantic network. This semantic network is customized in several directions.

In the figure 1. an example of our semantic network is illustrated. It is a digraph of symbols as vertices and time delay relations as edges.

A difference of our semantic network opposite to common semantic network is in the relations. In the common semantic network relations have a meaning. And they hold the information (for us, this kind of the information is represented again as the symbols). The relations hold only the information about a time delay.

A formal semantic network formula is: N = (S, C), where S is the set of symbols, objects (nodes) and the C is the set of connections (edges). Because, it is possible to have multiple connections to the same symbol, the aggregation function have to be defined – f_a.

Symbol definition $S =< id, f_a, x >$

id: identity of the symbol

f_a: the aggregation function AND(weighted distance, ...), OR(max)

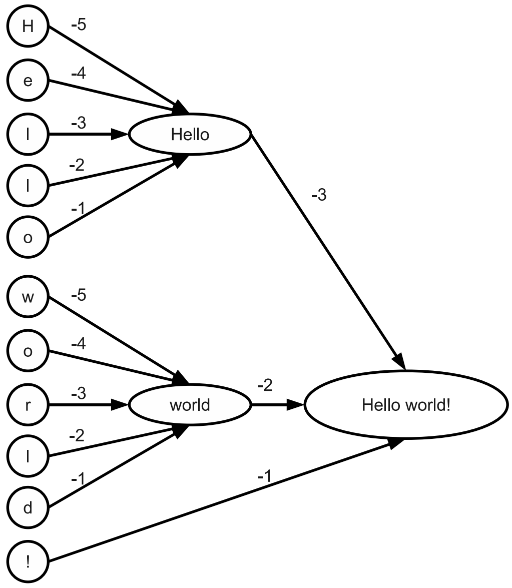


Fig. 1. Semantic Network

x: strength of the signal/activity $< 0, 1 >$,

The connection definition is:

$C =< s_{input}, s_{output}, \tau_{delay}, \lambda_{probability} >$

s_{input}: input symbol

s_{output}: output symbol

$\tau_{delay} = 1...n$: time delay of signal transition, where $t_{delay} > 0$

$\lambda_{probability}$: probability that next connection will be fired.

III. Separate timing / White Noise Modulation

Because of a performance optimization a White Noise Modulation was introduced the last year. We decide to use stochastic principles for inference.

A random signal, which has a constant power spectral density is called white noise [1]. With other words any distribution of values having a zero mean value [2] can be considered as the white noise.

The main properties for us are:

1) The mean value is zero
2) The random signal / random distributions of values
3) Constant power spectral density

Because the mean value is zero, the amplitude of the modulated information will be not affected, if the white noise which is used will not correlate with the modulated information. This is ensured by the second condition, which says that the values are random, and thus the probability of correlation depends on the randomness of white noise. Uniform power spectral density ensures that the result will not be affected even cross-reference symbols activation happened. E.g: those references where a same term/symbol is activated in two different contexts [4].

IV. Transcription Errors Correction

As the Elise was not originally designed for this kind of problem, we have introduced next modules:

1) Decision maker
2) Mirroring
3) Difference detection
4) Focusing
5) Error detection & Synthesizing

A. Decision Module

The decision module is used to change some characteristics to change strategy which drives how are the symbols fired.

In a simplified form, each decision has own input and output symbol. When output symbol for a decision is fired, a same symbol is fired also on input layer. E.g. when symbol for correction decision is fired. Elise starts to synthesize provided input, and will write it on the output.

B. Mirroring module

During the learning phase, Elise learns to mirror the input. When we talking about mirroring, we mind that output symbols should be fired according to fired input symbols to keep a same order of activating, and to keep also a same semantic. The same meaning means, that only similar symbols can be fired.

That it is almost the same as it is in case of the original semantic network of the Elise. Because it has symbols in the middle layer, which do the same job, and those symbols are used to drive a selection of founded documents. But for this case, we change a structure of the semantic network, that each input term / symbol has also counterpart symbol in the output layer.

C. Difference Detection

The difference detection module is composed from the symbols, which calculate difference between predicted symbols (symbols on the output layer) and the symbols which are in the original focused text.

Difference detection has own symbols on the output layer, which are used to change a behavior of a focusing and their are talking us how big difference is in a current focused place.

For difference detection was used coded symbols. It is special kind of symbols in the Elise network. Which are not created by learning process, but have connection to fixed symbols. And their activation is driven by program code.

D. Focusing

The focusing module is composed from the symbols, which have the connections to the symbols from the ”Difference Detection module”, and the symbols which drives a focusing mechanism.

The focusing mechanism controls a size and a place of a focused area. ”It is kind of driving how fast the Elise reads the input text”. It means, if it can move to the next word, or it should return back, or it needs to be on a same place more longer.

With other words focusing mechanism controls, which input symbols will be activated, and it also drives the difference calculation.

E. Error detection & Synthesizing

When the focused area is small and the difference for the focused area is big. It means that the focused area contains an error, which need to be corrected, than an error is founded, and the decision for a correction is activated.

When the Elise is switched to the process for an error correction. Then the focused area will be synthesized. Synthesizing is based on a special kind of the symbols, which according to the activated input symbols activate the output symbols in the way, which is similar to the natural language.

V. Conclusion

Elise in the numbers, it is a massive semantic network with:

- cca. 2 000 000 – base terms
- cca. 150 000 000 – instances of base terms
- cca. 1 000 000 000 – connections between terms

The Elise semantic search is in the progress, because of some issues. And the plan is to fix the remaining issues and to deploy semantic search into a production.

Transcriptions error correction - it is in the progress we are preparing real testing data, and we will see how to correct proposal solution.

References

[1] Carter Mancini, Bruce Ron Op Amps for Everyone. Texas Instruments (2009) ISBN 0080949487 p. 10-11

[2] Jeffrey A. Fessler, On Transformations of Random Vectors, Technical report 314, Dept. of Electrical Engineering and Computer Science, Univ. of Michigan (1998)

[3] Krokavec, D. – Filasová, A., Optimal Stochastic Systems, 2002, Elfa, ISBN 80-89066-52-6

[4] Dvorščák, S., Machová, K., Paralelná Inferencia Sémantickej Siete, WIKT 2014 November 20.-21. Smolenice, Slovakia ISBN 978-80-227-4267-2

[5] G. Dahl and Dong Yu and Li Deng and Alex Acero Large Vocabulary Continuous Speech Recognition With Context-Dependent DBN-HMMS Proc. ICASSP, 2011 Prague

[6] Dvorščák, S., Machová, K., Intelligent Search Machine with the Semantic Support Proc. of the 9th annual conference Znalosti 2010, Jindřichuv Hradec, Czech Republic, February 3-5, 2010, Publisher: VŠE v Praze, Praha, 2010, 199-202, ISBN 978-80-245-1636-3.

[7] Frank Amthor, Neuroscience For Dummies, 2011, For Dummies, ISBN 978-1118086865

Evaluating efficiency of experimental identification using an information system

[1]Vladimír GAŠPAR (4th year)
Supervisor: [2]Ladislav MADARÁSZ

[1,2]Dept. of Cybernetics and Artificial Intelligence, FEI TU of Košice, Slovak Republic

[1]vladimir.gaspar@tuke.sk, [2]ladislav.madarasz@tuke.sk

***Abstract*—This paper presents outcomes of numerous research studies published and presented during previous four years. The research conducted in the field of efficiency has proven that methodologies and methods used to quantify qualitative parameters are usually mathematically difficult or time consuming to be implemented and used in research practice. However, using the combination of simple figure or merit evaluation methods and key performance indicators, it is possible to easily automate the quality evaluation process. We seized upon this premise and applied respective methods in order to create a design of an information system, suitable for evaluating any qualitative parameter in scientific or research practice. The final outcome is a functioning prototype of a scientific research information system (SRIS), presented in this paper.**

***Keywords*—efficiency, laboratory of intelligent control systems of aircraft engines, SRIS, turbojet engine iSTC-20v.**

I. Introduction

Efficiency of identification is one of the factors in the applied research that is difficult to evaluate, not only because of its qualitative character, but also because of amount of information that are involved in the efficiency composition. Different analytical methods may be applied to carry out evaluations in manner of creating models of efficiency, or in some cases, even optimization models. These approaches are difficult for implementation in practice, because the formal description and runtime of a model in practice requires at least a simulation software, which is running the evaluations over experimental data [14]. Thus, our motivation originated from the idea to omit simulation software and mathematical overhead and find the simplest possible approach [15]. A longer but simpler approach is to utilize key performance indicators that easily quantify any qualitative parameter, which is conformant with SMART criteria [13]. They may be used in any information system or program firstly to plan and then to evaluate certain qualitative criteria. The character of applied research requires planning, data collection, analysis, evaluation and discovery of new knowledge [12], which is an ideal opportunity to use key performance indicators as a carrier of the quality information since planning until the knowledge discovery. The entire practical research has been adapted to fit the needs of the laboratory of intelligent control systems of jet engines, and the research object iSTC-20v in particular, which is a small turbojet engine [1].

II. Proof of concept

Although there might be doubts whether it is necessary to create a unique information system for a task of efficiency evaluation, an information system is a solution of several other problems that rise in the efficiency evaluation problem field in general.

As we already stated in the introduction, the efficiency evaluation in experimental identification is directly linked with physical testing that results in collections of different relevant or irrelevant data [15]. The second problem arises from the need to compare the expected and the achieved results. The expectations are represented by experimental plans. Besides their hierarchy, experimental plans contain the time parameter of research tasks, as well as the measurable criteria of expected results (valuated key performance indicators) [6][10]. Another problem that is connected with the evaluation of quality is, taking its components into account by means of their involvement in the composition of quality. For example, the reliability is one of the factors that is present in identification of technical systems and is emphasized as the most important factor of its operation. However, the reliability does not only rely on measured data but also on other technical data, obtained by observation of the system or its environment during its operational or stationary state. These observations are represented as log entries (operational, service, error) and allow the evaluation of reliability and safety, depending on the type of obtained technical data [8].

Previously enumerated factors emphasized the need for integrating all necessary tasks into a single platform, a scientific research information system SRIS [8].

III. Recent work

During previous years, we proposed a combination of possible methods [7] - [11], [17], tested them in a real environment (LIRS LM) and obtained sufficient knowledge about requirements of such an integrating platform [11]. We have done several research tasks as a proof of concept [2][4][5][6][14] in order to scientifically prove the ability of chosen methods to fulfill the needs of their practical implementation. In [7], we proposed the methodology of experiments planning and defined a basis of the conceptual character of the SRIS operation. Also the possibilities of various numerical methods in data analysis were studied, in order to discover valuable functions in analysis of thermodynamic data. Studies about the basic data representation and possibilities of supporting the (tacit) knowledge discovery were conducted. Other studies involved possibilities of evaluating quality by key performance indicators, measurement analysis from the statistical point of view, future possibilities of efficiency, reliability and quality modeling for purposes of their enhancement in the operation of the iSTC-20v. Implementation and development studies about the SRIS concept in particular were published in [7] [10]. As a full summary, we published the entire conceptual structure of a SRIS system with its background and fundamental ideas in [8] and a simple overview with future possibilities in [9].

IV. Implementation and development

The particular implementation development and deployment were carried out in conditions of the laboratory of intelligent control systems of aircraft engines (LIRS LM). Needs and basic processes of the laboratory were mapped to the proposed architecture and the architecture was adapted to be logically compatible with these internal processes. The mutual compatibility was double checked with the research team in LIRS LM. Functionalities of the system have either risen from the task of efficiency evaluation or the process structures in LIRS LM. The final concept included the following essential logical modules and functions:

- Experiment planning
  - Strategic planning
  - Tactical planning
  - Operative planning
    - Key performance indicators assignment
- Measurement processing
  - Measurement parsing
  - Structured storage of measurements
- Laboratory logs
  - Operational log
  - Service log
  - Error log
- Quality evaluation
  - Partial evaluations for quality components
  - Partial evaluations for planning levels
- Analysis
  - Numerical and statistical analysis (description)
  - Measurement playback and overview
  - Visual analysis of measurements

Due to other requirements of the LIRS LM laboratory and possible perspective future requirements, the following modules and functions were proposed:

- Environmental settings
  - Basic enumerable components
  - Key performance indicators definition
  - User definition
  - Errors definition
- Electronic library
  - Semiautomatic publication retrieval
  - Informational efficiency evaluation
- Real-time data transfer to remote locations [16] [3].

The development of the system was done using Asp.NET MVC framework and deployed at the laboratory server system. The prototype of the SRIS system is currently being filled with relevant data and will be used in the future for research agenda as well as the quality evaluation of the experimental identification.

V. Conclusion

In the past four years we have proposed, consulted, developed and deployed a SRIS information system as a solution for relatively different problems, with aim to enable evaluation of quality in experimental identification and as a support for knowledge discovery in the identification process. Contents of this paper resulted from the finished dissertation thesis, which has been submitted in the beginning of April 2015.

Acknowledgment

The work presented in this paper has been funded and supported by project of the culture-educational grant agent of ministry of education in Slovak republic, no. 014TUKE-4/2015 “Digitalization, virtualization and testing of a small turbojet engine and its elements using stands for modern applied lecturing”. This support is greatly acknowledged. I would also like to thank prof. Ing. Tobiáš Lazar, DrSc. (supervisor specialist) and doc. Ing. Rudolf Andoga, PhD. (consultant) for their support and leadership in the research process.

References

[1] 014TUKE-4/2015: Digitalizácia, virtualizácia a testovanie malého prúdového motora a jeho častí pomocou stendov pre potreby modernej aplikovanej výuky (Digitalization, virtualization and testing of a small turbojet engine and its elements using stands for the purpose of modern applied lecturing).

[2] Černej, M., Madarász, L., Gašpar, V.: Ekonomické aspekty identifikácie malého prúdového motora (Economical aspects of identification of a small turbojet engine) - 2012. In: Electrical Engineering and Informatics 3 : proceeding of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice. - Košice : FEI TU, 2012 pp. 52-58. - ISBN 978-80-553-0890-6 .

[3] Fiľko, M.: Možnosti prenosu laboratórnych dát s nízkou latenciou (Possibilities of low latency laboratory data transfer). Bakalárska práca, KPI FEI TU Košice. Vedúci: Dr. h.c. prof. Ing. Ladislav Madarász, PhD., Konzultant: Ing. Vladimír Gašpar. 50pp. 2014.

[4] Gašpar, V., et al.: Replacement impacts of MPM-20 turbojet engine using measured data – 2012. 6pp. In: proceedings of 14th Conference of Doctoral Students ELITECH’ 12 Bratislava, Slovak Republic. ISBN 978-80-227-3705-0.

[5] Gašpar, V., Karoľ, T., Havrilová, C.: Ideal MPM-20 jet engine running process with zero degrees of freedom - 2012. - 1 elektronický optický disk (CD-ROM). In: SCYR 2012 : Proceedings from conference : 12th Scientific Conference of Young Researchers : May 15th, 2012, Herľany, Slovakia. - Košice : TU, 2012 S. 158-161. - ISBN 978-80-553-0943-9.

[6] Gašpar, V., Madarász, L., Andoga, R., Főző, L., Judičák, J.: Operational constraints definition of a turbojet engine MPM 20. In: Proceedings of SAMI 2013 (11th IEEE International Symposium on Applied Machine Intelligence and Informatics). Herľany, Slovakia. Jan. 31-Feb. 2, 2013. ISBN 978-1-4673-5927-6, pp.6.

[7] Gašpar, V., Madarász, L., Andoga, R., Glodová, I.: On Scientific Research Information System. In: CINTI 2013: 14th IEEE International Symposium on Computational Intelligence and Informatics. November 19-21, 2013. Budapest, Hungary. pp. 75-78. ISBN 978-1-4799-0195-1.

[8] Gašpar, V., Madarász, L., Andoga, R.: Scientific Research Information System as a Solution for Assessing the Efficiency of Applied Research. In: Fodor, J., Fullér, R (eds.), Advances in Soft Computing. Intelligent Robotics and Control, Topics in Intelligent Engineering and Informatics 8., Spinger International Publishing 2014. pp. 273 – 293. ISBN 978-3-319-05945-7., ISSN 2193-9411.

[9] Gašpar, V., Madarász, L., Lazar, T., Andoga, R., Főző, L.: Komplexný výskum efektívnosti a inovácia technológie skúšok malého prúdového motora (8). (Complex research of efficiency and inovation of the

experimenting technology for a small turbojet engine (8)). In: ATP Journal, vol. 21, no. 4, 2014, pp. 52-55, ISSN 1335-2237, Slovak Republic.
[10] Gašpar, V., Madarász, L.: Information system implementation concept for qualitative parameters of a technical system. In: Novus Scientia 2013 : 12th international scientific conference of doctoral candidates at mechanical engineering faculties and universities. April, 10, 2013, Košice, Slovak Republic. 5 pp., ISBN 978-80-553-1381-8.
[11] Gašpar, V., Madarász, L.: Purpose-driven design of a KDD enabled application for experimental data analysis. In: ICCC 2013: IEEE 9th International conference on Computational Cybernetics. July 8-10, 2013. Tihany, Balaton, Hungary. 2013. pp. 257-261. ISBN 978-1-4799-0060-2.
[12] Gašpar, V.: Role of knowledge discovery in experimental identification of complex systems and its effectiveness. In: Poster 2013 : 17th International Student Conference on Electrical Engineering. May 16, 2013. Prague, Czech Republic. ČVUT. 5pp., ISBN 978-80-01-05242-6.
[13] Gašpar, V.: Utilization of key performance indicators in assessment of operational efficiency. In: SCYR 2014 : 14th Scientific Conference of Young Researchers., May, 20, 2014, pp. 283-285, ISBN 978-80-553-1714-4.
[14] Lazar, T., Madarász, L., Andoga, R., Gašpar, V.: Obzor a potenciál efektívnosti laboratória LIRS LM (Overview and pottential of efficiency in LIRS LM laboratory). In: Aeronautika 2013: International scientific conference, April, 25-26, 2013, Košice, Slovak Republic, pp. 178-181., ISBN 978-80-553-1416-7.
[15] Lazar, T., Madarász, L., Gašpar, V.: Procesná analýza odhadu efektívnosti identifikácie MPM s inteligentným riadením. (Estimation process analysis of identification efficiency of small turbojet engine with intelligent control), elfa, s.r.o., Košice. 2013. 160pp., ISBN 978-80-8086-200-8.
[16] Madarász, L., Gašpar, V., Rudas, I., Andoga, R., Gašpar, Ľ.: Proposal of Dissemination and Broadcasting of Laboratory Data within Small Time Latencies. In: Acta Mechanica Slovaca. Vol. 17, no. 3, pp. 26-32, 2013, ISSN 1335-2393, Slovak Republic.
[17] Madarász, L., Lazar, T., Gašpar, V., Andoga, R.: Perspectives in Evaluating Quality of Complex Technical Systems Using an Information System. In: MSC 2014 : The 2014 IEEE Multi-Conference on Systems and Control : Antibes, France, 8-10 October 2014. pp. 1918-1923., ISBN 978-1-4799-7405-4.

Evolution principles in formal languages

[1]*Michal SIČÁK (1st year),*
Supervisor: [2]*Ján KOLLÁR*

[1,2]Dept. of Computers and Informatics, FEI TU of Košice, Slovak Republic

[1]michal.sicak@tuke.sk, [2]jan.kollar@tuke.sk

***Abstract*—Evolution as a process of gradual and incremental quality improvement is used in various areas of computer language engineering. Basic principles of evolution are discussed and ways of applications are presented. Genetic programming is presented as a way to improve effectiveness of languages. Grammatical evolution as a successor to genetic programming is presented and shown on example. Use of evolution is presented in context-free grammar inference from a language without formal definition. Engineering discipline of grammarware is explained and similar approach, the grammar-based systems are explored.**

***Keywords*—language evolution, formal grammars, grammarware, genetic programming, grammatical evolution**

I. Introduction

Evolution is a process of incremental quality improvement. It occurs in nature where it is governed by complex laws of physics, biology and other natural laws. Evolution by natural selection is an automated process. Explanation and evidence of evolution by natural selection is presented by Dawkins in [1]. Other procedures, like selective breeding, is governed by human mind with foresight of future achievement. An example of that process can be selective breeding of dogs. From a common ancestor, wolf, mankind was able to create huge variety of dogs, from large Irish wolfhounds to diminutive chihuahuas. Other examples of natural or artificial evolution are presented by Dawkins in [2]. So to conclude, evolution can either be directly governed or an automatic process.

In computer science, evolution can represent historical changes in various languages. I am using term evolution in other sense, the one that is much more close to natural evolution. It is an automated process primarily used for domain specific languages (DSL) creation [3], [4]. Such languages are tailor suited for various needs. Needs that may change over time and therefore an automated process of constant improvement is useful feature for convenience. Völter in [5] argues, that examination of ties between modeling and programming can be solution to that problem. He proposes creation of modular languages that programmers of DSL languages can benefit from.

II. Symbolization and Conceptualization

The main driving force of my research is automatic evolution of human-computer languages. The idea is to create some sort of automaton, that will accept input and process it into a grammar. More input it gets, the more powerful the grammar becomes. The input must be some meaningful information, like an image or voice. The meaning itself is provided by a lecturer, a human operating on that automaton. If the grammar is strong enough, it will recognize similar patterns and will know their meaning. Process on which this automaton works is called formal processing of informal meaning by abstract interpretation, and it is an active research of my supervisor [6].

Meaning is a concept that we humans understand, but computer is a machine that blindly obeys orders written in instructions of its own language. It can process symbols, and it can process them really efficiently and fast. But this is still just a process predefined by some architect, a creator of computer code. Structures like words, or music bear information that can be symbolized in order to be processed by a computer. But to extract meaning, a process called conceptualization is initiated. Philosophical aspects of conceptualization and symbolization are discussed by Gärdenfors in [7].

III. Formal grammars

Grammar is a formalism that describes the structure of language. In the computer science, we can represent them as a 4-tuple:

$$G = (\Sigma, N, P, S) \tag{1}$$

Where:

- Σ - Finite set of input alphabet, i.e. the symbols actually appearing in sentences of language defined by grammar G.
- N - Finite set of nonterminal symbols. Such symbols do not appear in resulting language and they need to be transformed into terminal by rules.
- P - Finite set of rules in form: $(\Sigma \cup N)^* N (\Sigma \cup N)^* \rightarrow (\Sigma \cup N)^*$, where $*$ is Kleene star, signing zero to n repetitions and $\cup$ is union.
- S - Starting symbol of grammar, $S \in N$.

A. Grammarware and grammar based systems

Term grammarware was coined by Klint et al. in [9]. It comprises grammars and all grammar-dependent software, where term grammar is meant in wide aspect, like context-free grammars, XML schemas, class dictionaries, tree grammars and other. Their proposition is that grammars are widely used formalisms, yet the lack of engineering tools and procedures is evident.

The main idea of their work is a proposition to create new engineering discipline of grammarware. They describe various uses of grammars and state, that grammars are essentially everywhere. A lot of ad hoc tools and procedures are in existence, where majority of those seemingly different problems can be solved with use of same processes, if grammars are extracted and worked upon.

In a similar fashion, Mernik et al. in [10] argue that a lot of current existing software systems contain grammars in some way. Yet the main problem is, that the grammar is not formalized. They define term "Grammar based systems". It is any system that uses grammars, or grammatically produced sentences to solve problems outside the domain of programming language definition. The important part is, that such system has to be well structured. In their publication, they offer various examples of grammar based systems, that are not obviously tied with grammars, yet they concur to definition of grammar-based systems. Such examples include:

- grammatical approach to software engineering,
- adaptive programming,
- evolutionary programming,
- neural networks
- data representation.

They conclusion is one similar to Klint et al. in [9], that identification of grammar-based systems and their popularization may lead to improving the software and domain specific languages development and may create standardized tools and processes that will be applicable to various fields of computer science.

IV. Evolution of languages

Computer languages do evolve. Favre in [11] explores process of change of various computer languages over time. He argues, that languages are software too, and by acknowledging that fact, various problems with migration can be avoided. Evolution in his meaning is however not the type of evolution I will work on, as I have already explained in Introduction of this paper.

Process that used similar processes as biological evolution by natural selection is called genetic programming. Explained in depth by Koza in [12].

The main idea of genetic programming lies in changing the program according to needs of current situation. Where in nature, generally we have survival of the fittest. Or more specifically, those who are more likely to reproduce will reproduce, and they successful genes will spread. Hence the next generation of better replicators is created. In genetic programming, a fitness function must be created. Such function tests and selects the best behaving program or a set of programs. Genetic programming then "breeds" from the successful program population a next generation with some random mutations in code. And, again, fitness function selects the best program. In the end, we can see, that evolved program fits the problem nearly perfectly.

Evolution in this case is done on code itself, where code is put inside a tree structure and mutations are done by adding, removing or reordering branches. The problem of this approach is, that the process is bound to specific language. In the work of Koza it was Lisp.

A. Grammatical evolution

Different approach, more closely tied with biological evolution was presented by O'Neill and Ryan in [13]. They contributed to genetic programming work and created a process they coined as grammatical evolution. [14], [15], [16], [17], [18]

Grammatical evolution performs evolution not on the program itself, but on genetic string. This, similar to the DNA, is the essential part that is changed gradually by the evolution. The resulting program itself is constructed from this string on the basis of its grammar. We usually have some grammar defined in a meaningful form, such as Bakchus-Naur form (BNF) or other representation of CFG. This approach eliminates the problem of specific language in genetic programming, since it can be used on arbitrary formal language.

Consider a simple expression language written in BNF:

$$Expr \rightarrow Expr\ Op\ Expr \mid Var \quad (2)$$

$$Op \quad \rightarrow + \mid - \mid \times \mid \div \quad (3)$$

$$Var \quad \rightarrow n \mid v \quad (4)$$

where $n, v, +, -, \times, \div \in \Sigma$, n is number constant and v is variable, both are in their own respect regular languages. $Expr, Op, Var \in N$; $(2), (3), (4) \in P$ and $Expr = S$. Metasymbol $\mid$ represents choice, it is called a sum operator. Genetic string consists of numbers, each can determine the choice from a single BNF rule. The mod function is used in the process of determination. Take rule (2) for example, it has two choices. Number of choices will be signed as m, therefore a number from genetic string, let us mark it as x is processed accordingly:

$$r = x \mod m \quad (5)$$

Result r is the choice from BNF rule that is selected for program creation. This process is called genotype to phenotype mapping [13].

Process of grammatical evolution for genetic string $200, 25, 8, 58, 105, 353$ and grammar defined by rules mentioned earlier is depicted in Tab. I. We can see a resulting sentence of language unfolding as the genetic string is processed. Each non-terminal symbol is being converted by use of its production rule and selection number r.

TABLE I
Example of grammatical evolution

x	rule	r	result form
200	(2)	0	$Expr\ Op\ Expr$
25	(2)	1	$Var\ Op\ Expr$
8	(4)	0	$n\ Op\ Expr$
58	(3)	2	$n \times\ Expr$
105	(2)	1	$n \times\ Var$
353	(4)	1	$n \times v$

In order to evolve useful programs, again, a fitness function must be devised, that will test each program in current generation and select the best one. But this process differs from genetic programming, because only genetic string is mutated. A random member of it is changed. Since O'Neill et al. were using function mod, single codon (element in genetic string) can point to the same rule even with different value, gained by mutation. This also occurs in biological evolution, where random mutations do not have direct effect on phenotype. Those mutated genes themselves can spread and create phenomena called genetic drift.

Same process can be applied to mutate grammars themselves, therefore creating new languages. Process of grammatical evolution was used in research done by Kollár and

Pietriková in [19]. They were applying genetic strings on the definition of EBNF itself, so they were not creating new programs, not even new languages but new metalanguages. In their work, they argue, that we can see a program as either executable or data form, where both are semantically equivalent. From one form we may get another with help of genetic strings, hence we find ourselves inside an evolutionary circle, where no semantic change occur. By enrichment or attenuation of semantics, the circle become spiral where next generation may have different level of semantics.

Then, they explore the Extended Backus Naur Form (EBNF) and possibility of mixing the meta-level (language constructing level) with execution level. Hence, genetic string contains not only parts for constructing the language, but also data that are processed by the program being constructed. An example of evolution spiral is presented, where EBNF has been transformed from a data from to an executable form and back.

Further extensions are possible, as shown, when from weaker form of EBNF, the WBNF they evolved EBNF while using principles from evolutionary spiral. They argue, that unlike Völter in [5] who says that there are differences between programs and models, in their conception, there is no such difference, because each generated shape of a target program at any meta-level is in fact model, if it's abstracted, and a program, if it can be generated.

B. Grammar inference and language learning

So far I have discussed ways to create new languages by means of evolution from defined grammar. The evolution algorithms can be applied on inverse approach. Consider an existing domain specific language that does not possess defined formal grammar. It may have been created by domain specialist with little knowledge in the field of formal grammars to define one. Yet, there are ways to infer the grammar from samples of language like that. Hrnčič et al. in [20] present an algorithm that infers CFG from a set of code samples of an existing language.

Grammar inference usually works both on positive and negative samples, where the former are sentences of terminal symbols that belong to target language and the latter don't. Their algorithm, which they called MAGIc (Memetic algorithm for grammar inference) works on initial positive samples of a target language to infer CFGs, which are non-ambiguous and have type LR(1). MAGIc is composed of following steps:

- initialization
- improve and evaluate
- mutation
- generalization
- selection

Initialization is a step performed only once. It constructs initial grammar population based purely on input samples. For every sample, one grammar is generated, based on repetitions in its internal structure. After this step, the main evolutionary cycle is stared, and runs until the finishing condition is met, in this case it is the fixed count of algorithm iterations.

Improving and evaluation step is the first part of evolutionary cycle. So far, we have a set of grammars, that parse at least one input sample, and may or may not parse other. For each grammar, the samples are divided between positive and negative. In case that a grammar from a current population fails to parse a sample, improvements are made to it, so it can parse more positive samples. Those improvements include adding new or extending existing production rules. The process of improving is determined by application of linux diff command onto the samples. It returns add, replace or delete differences. Grammars are modified accordingly and are put to the existing population set.

Mutation is next step inside a cycle. In this process, some grammars from the set are randomly mutated on the level of BNF, where rules may be put inside option (zero or one), iteration^{+} (one to many) and iteration* (zero to many) operators. Successfully mutated grammars are then put inside grammar population.

Generalization is an important step that generalize grammar rules, but not to extreme extents, since overly generalized grammar is not optimal solution, as it can accept unnecessary wide range of samples. New grammars that passed the fitness test are put into grammar population.

Selection is the last step of evolutionary cycle. It evaluates all grammars in population according to fitness function, i.e. their strength to input samples. The best are selected for next cycle. When the condition of maximum repetitions is met, the resulting grammars should be the inferred grammars of input language.

They test their algorithm in comparison to another existing CFG inference processes, and find out that using evolution is a very improving aspect. They also show the inference process on an existing domain specific language for graphical tree drawing. They show the inferred grammar that parses all samples, hence they formalized a language that was created without formal grammar in mind.

V. Conclusion and future work

Use of evolution principles in computer language engineering has proved to be a huge asset to this field [21]. Since the beginning of this approach almost 30 years ago, there have been many problems solved and even more new topics found, as it is usual process in proper science. The use of genetic programing performed on the programs themselves has moved its way for better process, grammatical evolution. Mimicking the biological evolution by natural selection has proved to be vital step in the field of formal language evolutions opening new possibilities in a field of genetic programming and DSL creation.

Grammatical evolution solves many problems that arise in use of generative programming. It solved the problem of target language, where no longer it is needed to possess language specific algorithms, since it is possible to evolve programs of arbitrary language [13]. Consider a problem whose symptoms are known to us but internal details are concealed. In case like that, a fitness function that tests the symptoms can be constructed and grammatical evolution performed on arbitrary language can find optimal solution.

Evolution process can be applied in inverse problematic. If we have a language that was created by someone oblivious to formal grammars, but still it is well formed, algorithms that can infer formal grammar from such language have been devised. Some of them, that do not use evolution principles are less efficient in their results than those that apply the processes of random mutation and survival of the fittest [20].

The fact that domain specific languages are being created by domain specialists without proper grammar definition leads to existence of many ad-hoc tools and processes. The main idea behind grammarware [9] is elimination of this problem. The main benefits of this field of engineering is that by proper definition of grammars we can categorize and unify various software artifacts that may seemingly have nothing in common. Recent discoveries show, that many artifacts do converge in grammatical sense, as shown by Lämmel and Zaytsev in [22]. The field of grammarware is still in active research and still a lot of questions are unanswered. For example the problem of proper grammarware tools creation, testing utilities development, debugging and more.

Grammar-based systems offer alternative to view on how to use grammars in software engineering. As the creators of this term argue [10], in contrast to grammarware, grammar-based systems offer wider variety of grammar applications to be researched and then reused in form of computer languages enrichment.

My research's further direction will be determined by evolutionary principles discussed in this paper. The process of change and mutation can be applied to languages obtained by information processing i.e. symbolization. By performing this process, known as conceptualization, we may obtain grammatical machine that will be able to recognize known patterns and even be capable of recognition of previously unknown patterns that bear similarities to knowledge base already stored within. Another inverse operation can be performed: the information generation, where based on its internal knowledge, the machine can generate random artifacts. Only the human teacher can determine, whether such creations have meaning, and if they have, they are absorbed into ever expanding knowledge base, therefore evolving better and more capable grammatical machine.

Acknowledgment

This work was supported by project VEGA 1/0341/13 *Principles and methods of automated abstraction of computer languages and software development based on the semantic enrichment caused by communication.*

References

[1] R. Dawkins, *The blind watchmaker: Why the evidence of evolution reveals a universe without design*. WW Norton & Company, 1996.
[2] ——, *Climbing mount improbable*. WW Norton & Company, 1997.
[3] A. Van Deursen, P. Klint, and J. Visser, "Domain-specific languages," *Centrum voor Wiskunde en Informatika*, vol. 5, p. 12, 2000.
[4] M. Mernik, J. Heering, and A. M. Sloane, "When and how to develop domain-specific languages," *ACM Comput. Surv.*, vol. 37, no. 4, pp. 316–344, Dec. 2005. [Online]. Available: http://doi.acm.org/10.1145/1118890.1118892
[5] M. Völter, "From programming to modeling-and back again," *Software, IEEE*, vol. 28, no. 6, pp. 20–25, 2011.
[6] J. Kollár, "Formal processing of informal meaning by abstract interpretation," *Smart Digital Futures 2014*, vol. 262, p. 122, 2014.
[7] P. Gärdenfors, "Symbolic, conceptual and subconceptual representations," in *Human and Machine Perception*. Springer, 1997, pp. 255–270.
[8] N. Chomsky, "On certain formal properties of grammars," *Information and control*, vol. 2, no. 2, pp. 137–167, 1959.
[9] P. Klint, R. Lämmel, and C. Verhoef, "Toward an engineering discipline for grammarware," *ACM Transactions on Software Engineering and Methodology (TOSEM)*, vol. 14, no. 3, pp. 331–380, 2005.
[10] M. Mernik, M. Crepinsek, T. Kosar, D. Rebernak, and V. Zumer, "Grammar-based systems: Definition and examples." *Informatica (Slovenia)*, vol. 28, no. 3, pp. 245–255, 2004.
[11] J.-M. Favre, "Languages evolve too! changing the software time scale," in *Principles of Software Evolution, Eighth International Workshop on*. IEEE, 2005, pp. 33–42.
[12] J. R. Koza, *Genetic programming: on the programming of computers by means of natural selection*. MIT press, 1992, vol. 1.
[13] M. O'Neill and C. Ryan, *Grammatical evolution: evolutionary automatic programming in an arbitrary language*. Springer Science & Business Media, 2003, vol. 4.
[14] M. O'Neill, E. Hemberg, C. Gilligan, E. Bartley, J. McDermott, and A. Brabazon, "Geva: grammatical evolution in java," *ACM SIGEVOlution*, vol. 3, no. 2, pp. 17–22, 2008.
[15] M. ONeill and A. Brabazon, "Recent adventures in grammatical evolution," in *Proceedings of CMS*, 2005, pp. 245–253.
[16] I. Dempsey, "Grammatical evolution in dynamic environments," Ph.D. dissertation, University College Dublin Ireland, 2007.
[17] I. Dempsey, M. O'Neill, and A. Brabazon, "Constant creation in grammatical evolution," *International Journal of Innovative Computing and Applications*, vol. 1, no. 1, pp. 23–38, 2007.
[18] E. Hemberg, C. Gilligan, M. ONeill, and A. Brabazon, "A grammatical genetic programming approach to modularity in genetic algorithms," in *Genetic Programming*. Springer, 2007, pp. 1–11.
[19] J. Kollár and E. Pietriková, "Genetic evolution of programs," *Central European Journal of Computer Science*, vol. 4, no. 3, pp. 160–170, 2014.
[20] D. Hrnčič, M. Mernik, B. R. Bryant, and F. Javed, "A memetic grammar inference algorithm for language learning," *Applied Soft Computing*, vol. 12, no. 3, pp. 1006–1020, 2012.
[21] R. I. Mckay, N. X. Hoai, P. A. Whigham, Y. Shan, and M. O'Neill, "Grammar-based genetic programming: a survey," *Genetic Programming and Evolvable Machines*, vol. 11, no. 3-4, pp. 365–396, 2010.
[22] R. Lämmel and V. Zaytsev, "An introduction to grammar convergence," in *Integrated Formal Methods*. Springer, 2009, pp. 246–260.

Evolving Reactive Micromanagement Controller for Real-Time Strategy Games

[1]*Martin ČERTICKÝ*, [2]*Michal ČERTICKÝ*

[1]Dept. of Cybernetics and Artificial Intelligence, FEI TU of Košice, Slovak Republic
[2]Faculty of Electrical Engineering, Czech Technical University in Prague, Czech Republic

[1]martin.certicky@student.tuke.sk, [2]michal.certicky@agents.fel.cvut.cz

***Abstract*—Real-Time Strategy (RTS) games are a genre of video games representing an interesting, well-defined adversarial domain for Artificial Intelligence (AI) research. One of many subproblems that RTS players need to solve is the micromanagement of individual units (simple agents carrying out player's commands) during combat. Numerous multi-agent reactive control mechanisms have already been developed to maximize the combat efficiency of controlled units. Majority of these mechanisms make use of numeric parameters that need to be fine-tuned in order to achieve desired behavior. Due to a large number of these parameters, assigning them manually is inconvenient and training them by machine learning methods usually takes a long time (search space is too large). To reduce the search space and accelerate the training, we propose a simple reactive controller with only eight parameters. We implement it for a classic RTS game StarCraft: Brood War and train the parameters using genetic algorithms. Our experiments demonstrate an impressive combat performance after only a small number of generations.**

***Keywords*—Genetic Algorithms, Evolution, RTS, micromanagement, StarCraft.**

I. Introduction

Real-time Strategy (RTS) games, as a genre of video games in which players manage economic and strategic tasks by gathering resources and building bases, increase their military power by researching new technologies and training units, and lead them into battle against their opponent(s), serve as an interesting domain for Artificial Intelligence (AI) research. They represent a well-defined, complex adversarial systems [1] which pose a number of interesting AI challenges in the areas of planning, dealing with uncertainty, domain knowledge exploitation, task decomposition and, most relevant to the scope of this article, spatial reasoning and machine learning [2].

Research in the area of RTS game AI is usually classified according three levels of abstraction: high-level strategy, middle-level tactics and low-level reactive unit control (referred to as "micromanagement" by RTS players). Reactive control, as a challenge addressed by this paper, aims at maximizing the effectiveness of individual units of different types in combat by moving them on the battlefield and selecting the attack targets in response to terrain and the activity of opponent's units.

Decision-making for reactive unit control can in general take centralized or decentralized (distributed) approach. Centralized approaches try to control the whole group of units at once by searching a game tree. For example, Churchill et al. [3] presented a variant of alpha-beta search and Wang et al. [4] employed a Monte-Carlo planning approach to the problem of micromanagement. Unfortunately, centralized solutions are usually too slow due to large search space and can only be applied to situations with low unit counts.

Decentralized approaches are much more common, because they scale better to situations with higher numbers of controlled units. The search space is significantly reduced by making the control decisions for each unit independently of others. Decentralized solutions, which often take advantage of potential fields or influence maps, have one thing in common each unit is controlled by a relatively simple controller algorithm. The unit controller is a state machine that reacts solely to current game state without performing any kind of lookahead search.

The controller always comes with several parameters that need to be fine-tuned in advance to achieve a desired effective behavior. There has been a significant amount of work using machine learning techniques like Reinforcement Learning, Bayesian Modelling or Genetic Algorithms (GA) to train the parameters of underlying controller algorithm automatically – multiple examples of research in this area can be found in a survey paper by Ontanón et al. [2]. However, a general drawback of decentralized micromanagement techniques is the high number of controller's parameters, leaving us with too many possible value assignments [2] (problem of dimensionality). For example, Liu et al. [5] used GA to train the values of 14 different parameters (chromosome encoded as a 60-bit string), Ponsen [6] used 20 genes to train 20 parameters for combat-related actions and Sandberg el al. [7] trained 21 parameters of their micromanagement controller.

To address the problem of dimensionality, we propose a controller with only eight parameters, as described in Section II. We train these parameters using the GA with the population of 32 individuals, roulette wheel selection method and uniform cross-over method, according to detailed description in Section III. Finally, the performance of the trained controller is discussed in Section IV.

We implemented and tested our solution in a classic RTS game *StarCraft: Brood War* by Blizzard Entertainment, which was accessed via application programming interfaces BWMirror[1] and BWAPI[2].

II. Unit Controller

We propose a simple unit controller, implemented as a finite state machine with as few adjustable parameters as

[1]http://github.com/vjurenka/BWMirror/
[2]http://github.com/bwapi/bwapi/

possible to address the problem of dimensionality – low number of parameters to fine-tune reduces the assignment search space and makes the training faster. The controller is called individually for each unit on each logical frame of the game (approximately 23.81 times per real-world second on the *"Fastest"* game speed). On each frame, a unit is either left untouched or is assigned a new command. There are two types of commands that can be assigned by the controller: `move(Position pos)` or `attack(Unit enemy)`. See the flow chart describing the controller's implementation in Figure 1.

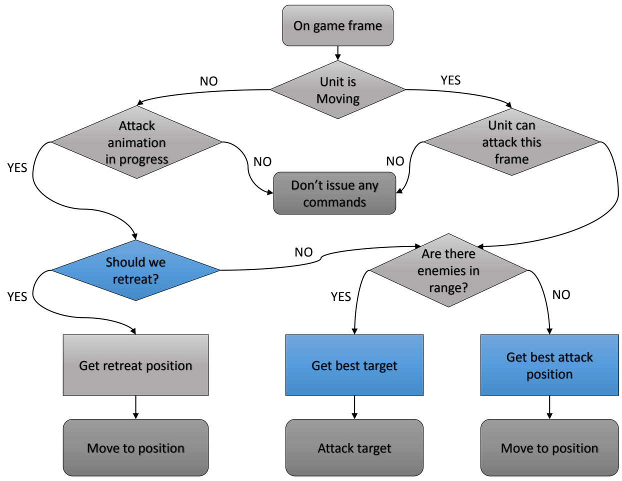


Fig. 1. Flow chart describing the reactive unit controller. Parts of the controller using parameters trained by GA are highlighted by the blue color.

First, the controller checks if a given unit is currently executing a `move` command. If it is, it checks if the unit is able to attack at this frame (units in StarCraft must wait a specified number of frames, called *"weapon cooldown"*, between individual attacks). If the unit is able to attack and there are some enemy units (targets) in its attack range, the controller selects the best target and issues an `attack` command (target selection is explained in subsection II-A). If there are no targets in range, the best position from its surrounding area is selected (subsection II-B) and the unit receives a `move` command sending it there. If the unit is already executing an *attack* command on the current frame but it's not in the middle of attack animation (attack animation should not be interrupted), it has two options: continue attacking or retreat to a safer position[3]. The decision to retreat is parametrized and described in subsection II-C, but the selection of retreat position is hard-coded – the controller simply computes the retreat vector directed away from the biggest enemy threats.

Following three subsections describe all the functions containing the parameters that were fine-tuned by the genetic algorithm. The total number of parameters is eight and each of them has a value from interval $[0, 1]$.

A. Selecting an Attack Target

The following function is used to select the most appropriate target from all the enemy units in our unit's attack range. The function scores enemy units and outputs the one with the highest score to be used as an argument for the `attack` command.

$$Score = (D \cdot p_1) - (HP \cdot p_2) + (L \cdot p_3)$$

where D is the damage of a given enemy unit (multiplied by 3 to increase its significance without normalising the value) and HP is the sum of its remaining Hit Points and Shields. The L variable equals 100 if the sum of unit's remaining Hit Points and Shields is lower than our unit's damage. Otherwise, the L variable equals 0. Each of the variables in this function is multiplied by an evolved parameter (p_1, p_2, p_3).

B. Selecting an Attack Position

Next function selects the most appropriate position for a unit to move to when it intends to attack enemy targets but has no targets in range. This function computes the score of a collection of walk tiles surrounding our unit (see Figure 2) and returns the position of the one with the highest score. This position is used as an argument for the `move` command.

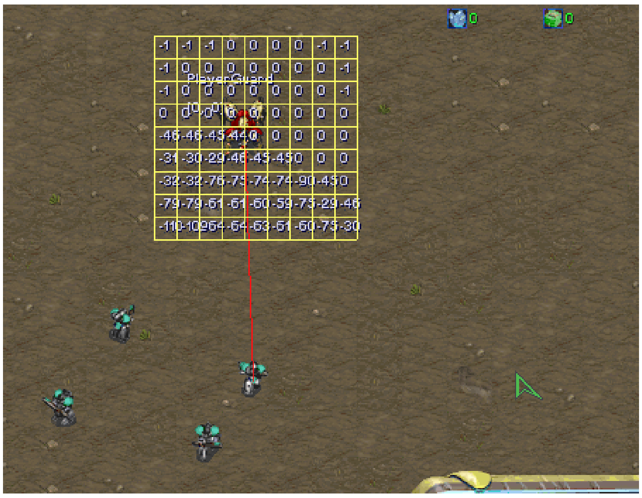


Fig. 2. Walk tiles around our unit with the assigned scores. Walk tiles in range of multiple enemy units have the lowest score.

We use the following equation to score a walk tile. Again, the one with the highest score is used as a function's output.

$$Score = D_f \cdot p_4 - D_t \cdot p_5 - L \cdot p_6 - C \cdot p_7$$

where D_f is the sum of the damage our unit is able to inflict from the examined tile and D_t is the sum of the damage that the enemy units are able to deal to our unit if it is moved to that tile. The L variable is similar to the one from the previous method – it equals 100 if the sum of the damage of enemy units having range on examined tile is higher than remaining Hit Points and Shields of our unit (equals 0 if not). Variable C equals 50 if the path to the examined position crosses the path of a moving friendly unit or if that path collides with any non-moving unit (equals 0 if not). Again, all the variables are multiplied by evolved parameters (p_4-p_7).

C. Deciding to Retreat

The last parametrized function returns `True` if our unit should retreat to a safer location. The retreat decision is made if the following holds:

[3] We intentionally do not let units switch between the targets in their range because such switch causes a long pauses between the attacks in StarCraft. This, in general, negatively impacts the performance.

$$D \geq HP \cdot p_0$$

where D is sum of damage of all units currently attacking our unit and HP is the sum of remaining Hit Points and Shields of our unit. The HP variable is multiplied by an evolved parameter p_0.

III. Evolving the Parameters

It is known that evolutionary optimization by simulating fights can be easily adapted to any parameter-dependent micromanagement control model [2]. To train the parameters of our controller, we used three different scenarios – each scenario had our AI player face a different type of commonly used enemy units (more details can be found in the following section). Enemy units were controlled by standard built-in StarCraft AI script. The controller was trained using the population of 32 individuals during 15 generations.

A. Encoding

Every individual is represented by a vector of 8 real numbers from the interval $[0, 1]$, each representing a single parameter used in our controller.

B. Evaluation

The game ends if one of the players loses all the units or if the game hits the predefined time limit. After that, we evaluate the current individual using the sum of remaining Hit Points and Shields of all its surviving units and substract the Hit Points and Shields of surviving enemy units. The final fitness of an individual is averaged over 5 games to lower the randomness caused by the StarCraft mechanics and by varying behavior of the enemy AI script.

C. Selection Mechanism

We use a roulette-wheel selection method in our GA. Since negative fitness values are not acceptable for this method, we apply windowing to ensure the positive values during the selection. Exactly half of the population is selected for the parent role (while each individual may be selected multiple times). Slight form of elitism is also applied on the populations – the best member of each population is guaranteed to become a parent at least once.

D. Genetic Operators

1) Cross-over: The GA uses a uniform cross-over with the mixing ratio of 0.5. The offspring has half of the genes of each parent on average, as the cross-over points are chosen randomly throughout the gene.

2) Mutation: We apply a 10% chance of uniform mutation for each individual carrying out to the next generation. The mutation operator is applied after cross-over of the population is done to prevent mutating the individuals prematurely and possibly losing well-scored offsprings.

IV. Performance

In the training scenarios, our AI player controlled a squad of *"Dragoon"* type units (widely used StarCraft unit with an average attack range). Three scenarios with different types of enemy units were chosen for the training due to difference in desired behavior depending on enemy unit's attack range. When facing melee (close-combat) or short-ranged units, the Dragoons have a major advantage if they learn to apply so-called *"kiting"* behavior (a hit-and-retreat technique commonly used in RTS games). On the other hand, too frequent retreating while facing enemy units with the same or longer attack range might lead to decrease in the performance.

The performance was compared to standard built-in StarCraft AI script (the variant used in "custom game" mode). We computed the fitness of the built-in AI using the same fitness function and averaged the value over 5 games.

A. Dragoons vs. Zealots scenario

Zealots belong among the most commonly used melee units in StarCraft. After the initial generation, the majority of individuals adopted similar values of first parameter (handling the retreat decisions) and showed an elusive kiting behavior patterns during their games. After only a few generations, the results of the controller were fairly close to absolute optimum (see Figure 3). The controller was outperforming the native AI massively.

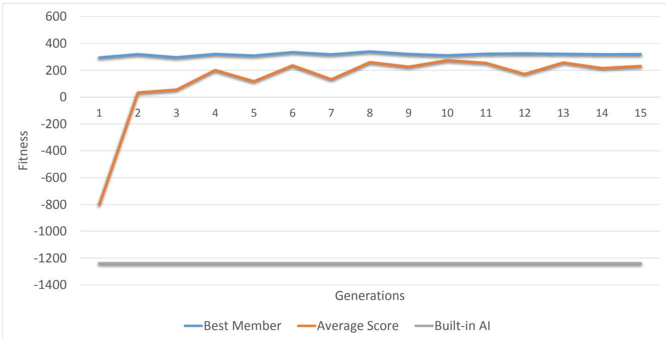


Fig. 3. Dragoons vs. Zealots scenario: Fitness of the best individual (blue line) and the average fitness of the population (red line) over time (15 generations) compared to the fitness of built-in AI (grey line).

B. Dragoons vs. Marines scenario

Marines were chosen for the second scenario as the cheap and common short-ranged unit. The convergence of the GA was not as fast as in first scenario (see Figure 4), yet the behavior of the AI player ended up being quite similar. The units became elusive, taking advantage of their superior range and kiting the enemy units. It should be noted that with the improving performance of the AI player, the games became longer. This was due to dragoons kiting the enemy units, attacking only if they were in safe distance.

C. Dragoons vs. Dragoons (mirror match scenario)

The last experiment was performed on a mirror match scenario, where the AI player faced a set of units identical to its own. While the first two scenarios resulted in an observable behavioral difference compared to native AI, this scenario was not able to produce the results as impressive as the previous

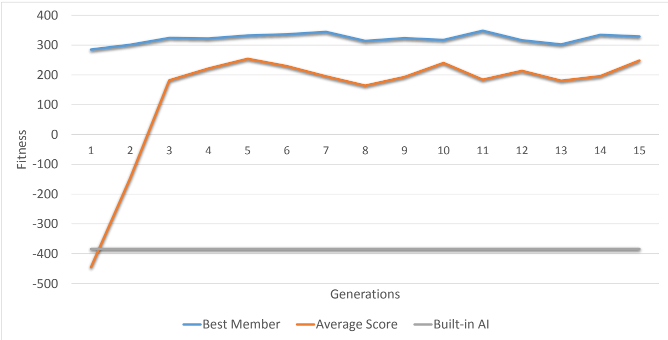


Fig. 4. Dragoons vs. Marines scenario: Fitness of the best individual (blue line) and the average fitness of the population (red line) over time (15 generations) compared to the fitness of built-in AI (grey line).

two. Kiting the enemy army with same number of units and the same attack range is ineffective – even moving the units may result in worse scores. Our individuals gradually learned to reduce their movement and trained only the target selection. After a few generations the AI player developed a behavior similar to that of built-in AI scripts, which is also reflected in the comparable fitness values (see Figure 5).

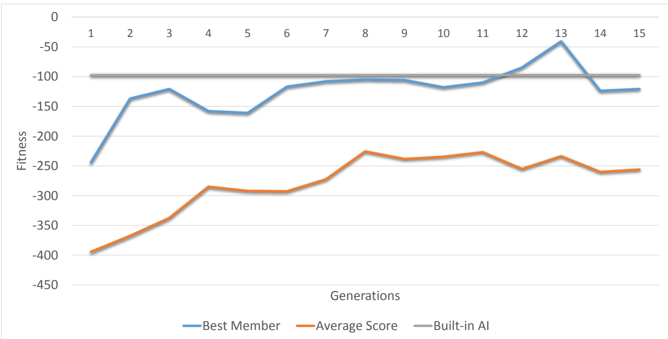


Fig. 5. Dragoons vs. Dragoons scenario: Fitness of the best individual (blue line) and the average fitness of the population (red line) over time (15 generations) compared to the fitness of built-in AI (grey line).

In two out of three scenarios, our AI player surpassed the built-in AI's performance significantly. The controller was performing best while facing melee or short-ranged units. When facing long-ranged units (the mirror scenario), the controller was performing slightly worse than the native AI.

V. Conclusion

In this paper, we addressed the dimensionality problem, typical for current parametrized RTS micromanagement solutions. The problem is caused by the large number of numeric parameters of the control mechanisms, that need to be fine-tuned in order to achieve a desired effective behavior.

To reduce the search space and accelerate the training of these parameters, we proposed a simple reactive controller with only eight parameters. We implemented it for a classic RTS game StarCraft: Brood War and trained the parameters using genetic algorithm with roulette-wheel selection and uniform cross-over methods.

Thanks to the reduction of search space, less than 10 generations with the population size of 32 proved sufficient for GA to converge in every experimental scenario.

Compared to a built-in StarCraft AI, the performance proved to be exceptionally good in two out of three experimental scenarios (in the third one, it was comparable to that of the built-in AI). The results depended on the type of enemy units presented in each scenario. The hit-and-retreat behavior of the AI player that emerged in two out of three scenarios was extremely effective against melee or short-range units. It proved to be ineffective against long-range enemy units and was suppressed after a few generations to be replaced with more static behavior.

The proposed controller has proven to be effective in typical StarCraft combat situations only after a short training by genetic algorithm.

References

[1] M. Buro, "Call for AI research in RTS games," in *Proceedings of the 4th Workshop on Challenges in Game AI*, 2004, pp. 139–142.

[2] S. Ontanón, G. Synnaeve, A. Uriarte, F. Richoux, D. Churchill, and M. Preuss, "A survey of real-time strategy game AI research and competition in StarCraft," *IEEE Transactions on Computational Intelligence and AI in Games*, vol. 5, pp. 293–311, 2013.

[3] D. Churchill, A. Saffidine, and M. Buro, "Fast heuristic search for rts game combat scenarios." in *AIIDE*, 2012.

[4] Z. Wang, K. Q. Nguyen, R. Thawonmas, and F. Rinaldo, "Monte-carlo planning for unit control in starcraft," in *2012 IEEE 1st global conference on consumer electronics (GCCE). IEEE, New York*, 2012, pp. 263–264.

[5] S. Liu, S. J. Louis, and C. Ballinger, "Evolving effective micro behaviors in rts game," in *Computational Intelligence and Games (CIG), 2014 IEEE Conference on*. IEEE, 2014, pp. 1–8.

[6] M. Ponsen, "Improving adaptive game ai with evolutionary learning," Ph.D. dissertation, Citeseer, 2004.

[7] T. W. Sandberg and J. Togelius, "Evolutionary multi-agent potential field based ai approach for ssc scenarios in rts games," Ph.D. dissertation, Masters thesis, University of Copenhagen, 2011.

Extraction of Human Body Parameters from Static Image

[1]*Ondrej Kainz (2nd year)*
Supervisor: [2]František JAKAB

[1,2]Dept. of Computers and Informatics, FEI TU of Košice, Slovak Republic

[1]ondrej.kainz@tuke.sk, [2]frantisek.jakab@tuke.sk

***Abstract*—In this paper is presented a novel approach for human body parameters estimation from static 2D image. We utilize known and constant parameters of human body and human face. Derived from known proportions and lengths of constant size we aim to develop a model for estimation of such parameters. Known algorithm for human detection called histogram of oriented gradients (HOG) is utilized. Overall framework is expected to provide outputs that should closely correlate with real measurements. Outputs may be further utilized for the statistics purposes, as a complement to human identification or in other areas of industry.**

***Keywords*—anthropometry, distortion, human body, proportions.**

I. Introduction

Direct measurement is always the most accurate and precise, however certain measurements or parameters may be also extracted by non-direct means, e.g. photography or video. Estimation of parameters from photography or further referred as 2D static image is complex process that requires several steps in order to provide value that is to converge to real number. One of such means that is to help in determining the parameters are proportions of human body. We build this on fact that human body has specific limits, as to height, length of bones etc. Further several of these values, when in relation one to another produce constant proportions in the larger group of population. Such relations may then be used for the estimation of the real size, e.g. as a reference value may serve distance between eyes or several weak references that are to serve as one robust.

II. State of the Art: Body Proportions

This part is dedicated to research in the field of body proportions related to the field of computer science. Outputs of these researches are to be further used in our research. In [1] authors refer to model of human body as a skeleton model. Research carried out by these authors was focused on orientation of body in space, in order to do that skeleton model extraction was necessary. Thus, authors first extracted five feature points of head and limbs, followed construction of edges that is connecting feature points (vertices). Anthropometric data was used for model development, usually 3D model. In research by [2] used model utilizing such data collected as a part of Civilian American and European Surface Anthropometry Resource Project (CAESAR). Data was collected from 64 white markers places on subject, resulting data was combined and surface was reconstructed. Another research by [3] was focused on re-parametrization of 3D body models based on statistical body model. Review of body modelling techniques is presented in [4], these authors also carried out study on fast deformation techniques for human body models generation, as an input were used anthropometric parameters. Application of 3D body models are described by [5], examples are: crash simulation, motion analysis, workplace assessment, entertainment or motion notation and understanding.

When it comes to body proportion it is comparison of length of one body part to another. In book of body measurements [6] are presented several of these proportions, e.g.: height of sitting person to height of standing person. Further is provided an example of just mentioned proportions to yield correlation of 44.71 to 45.45 percent.

As authors in [7] state such proportion may be used as an indicator associated with risk of fatness, heart diseases, diabetes or some types of cancer. Another indications may be quality of the environment during infancy since between birth and puberty legs grow faster. Authors further point out two way of measuring length of leg: iliac height, where the length is measured from iliac crest to floor and another is subischial leg length, where measurement is carried out as a difference between stature and sitting height. Thigh length, is third approach to leg measurement "between the proximal end of the greater trochanter and the distal lateral femoral condyle". The last two approaches are however difficult to measure. The last one is knee height measure between anterior surface of the thigh and floor. As it is noticeable from just stated there are many approaches toward measurement of body parts, therefore based on testing the best one should be chosen. Perhaps an interesting fact is stated by authors regarding the mean values of height: Pygmies of Africa 145 cm and 136 cm for men and women respectively, while 184 and 171 cm for Dutch of Europe. Another factor to be taken into account is that body shape is based on genetic basis, however human groups that live in the same environment for several generations manifest no significant difference.

Following the body proportions the study by [8] presents anthropometry data of several body dimensions - together nine, from three nationalities North American (NN), Japanese (J) and Hong Kong (HK). For these dimensions see Fig. 1 where:

A - Vertical grip/reach; B - Stature/head height; C - Shoulder height; D - Elbow height; E - Eye height; F - Forward grip/reach; G – Knuckle height; H - Knee height; I - Waist height.

Complete list and values is available in [8]. Based on these data we estimated the smallest error when compared ratio of dimensions of two measurements to be $A : B$ and $B : E$. As an example: B of 184.4 cm for NN and 175 cm for J, E of 172.7 cm for NN and 138.2 cm for J, yield ratio $B : E$ of 184.4 : 172.2 to be 1.0677 for NN and 175 : 163.5 to be 1.0703, estimated error between ratios is 0.26 percent. Such comparison was made for men:men, women:women and men:women.

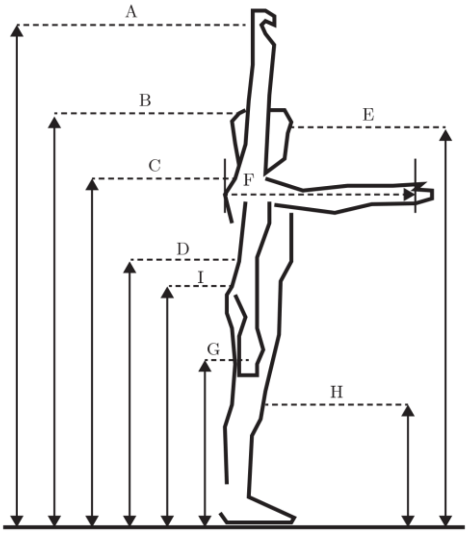


Fig. 1. Anthropometric measurements. Based on: [8].

III. Parameters Estimation

In order to extract parameters from the image there are certain requirements for the image to have - proper quality and resolution and person visible on image as much as possible. Note that only person facing the camera is considered in this phase of a project. See Fig. 2 for the following flow: once the image is loaded into the system, correction of the lens distortion is to be carried out, i.e. pincushion, barrel, and also non-lens distortion - perspective. Detection of person in the image is realized using robust feature descriptor referred to as histogram of oriented gradients or for short HOG. To detect and extract the specific limbs of person is to be used part-based model detection, this will later allow mapping of pre-defined 3D skeleton model of human person to be mapped to static 2D scene and also allow first approximation of parameters, e.g. height or limb length. To enhance the accuracy of parameters second approximation, based on the extraction of proportions of specific human body features is introduced and further deployed together with anthropometrical regression, i.e. the best proportions are used based on prior regression analysis. As an additional indicator may serve known objects that are present in the image - they are detected, classified and mapped. However using of such indicator has to deal with several issues - estimation of camera parameters, vanishing line and vanishing point, to achieve this, proper conditions have to be met. For each step of just described techniques, semi-automatic and automatic approaches are to be considered.

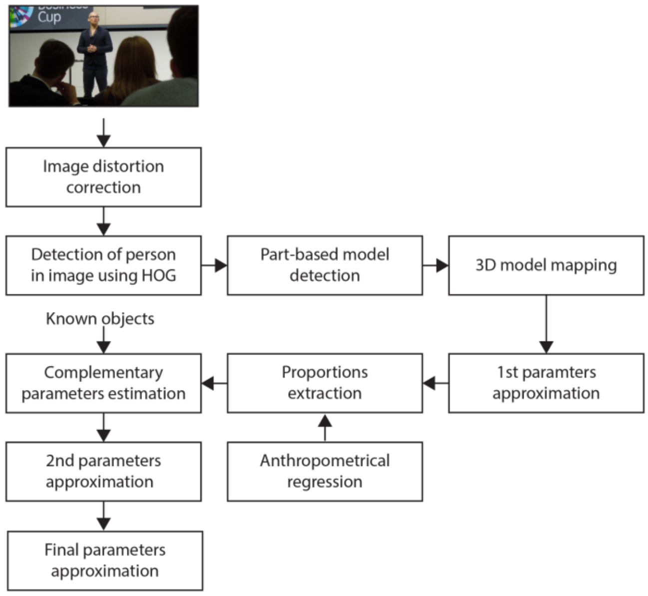


Fig. 2. Internal system processes for parameters estimation

IV. Conclusion

In this paper was presented novel framework for visual human body parameters estimation from static 2D image. Knowledge of well-known parameters that are analogous for each individual allow to carry out such estimation and further complementary approaches are to be used for enhanced approximation to real values.

Acknowledgment

We support research activities in Slovakia/This project is being co-financed by the European Union. Paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF.

References

[1] W. Wu, A. Hao, W. Wang, Y. Wang, *A PCA-Based Automated Method for Determination of Human Body Orientation*. In: Intelligence Computation an Evolutionary Computation, ISBN 978-3-642-31655-5, 2013, vol. 180

[2] B. Allen, B. Curless, Z. Popović, *The space of human body shapes: reconstruction and parameterization from range scans*. In: ACM Transactions on Graphics, ISBN 978-1-58113-709-5, 2003, vol. 22, issue 3

[3] C. Rupprecht, O. Pauly, C. Theobalt,. S. Ilic, *3D Semantic Parameterization for Human Shape Modeling: Application to 3D Animation*. In: International Conference on 3D Vision - 3DV 2013, 2013

[4] M. Kasap, N. Magnenat-Thalmann, *Parameterized human body model for real-time applications*, In: Proceeding Proceedings of International Conference on Cyberworlds, ISBN 978-0-7695-3005-2, 2007

[5] B.L. Webber, C.B. Phillips, N.I. Badler, *Simulating Humans: Computer Graphics, Animation, and Control*. Philadelphia : Oxford University Press, 1993. 288 p. ISBN 978-0-19536-086-8

[6] S.P. Singh, P. Mehta, *Human Body Measurements: Concepts And Applications*. New Delhi : PHI Learning Private Ltd., 2009. 469 p. ISBN 978-81-203-3863-0

[7] B. Bogin, M.I. Varela-Silva, *Leg Length, Body Proportion, and Health: A Review with a Note on Beauty*. In: International Journal of Environmental Research and Public Health, ISSN 1660-4601, 2010, vol. 7

[8] M. Tooley, *Design Engineering Manual*. Burlington : Elsevier Inc., 2010. 624 p. ISBN 978-1-85617-838-9

Improving performance of intrusion detection using prediction analysis

[1]*Michal ENNERT (4th year)*
Supervisor: [2]*Liberios VOKOROKOS*

[1,2]Dept. of Computers and Informatics, FEI TU of Košice, Slovak Republic

[1]michal.ennert@tuke.sk, [2]liberios.vokorokos@tuke.sk

***Abstract*— Intrusion detection is enormously developing field of informatics. This paper provides summarization of the work and a survey of actual research trends in intrusion detection. It presents a review about intrusion detection systems (IDS) with ways to achieve higher performance with usage of general purpose computing on graphics processors and effective analysis in detection engine. The most common technique is pattern matching algorithm used in detection engine.**

***Keywords*—Intrusion detection, high performance computing, pattern matching algorithm**

I. Introduction

Computer security plays an important role in the present. Majority of malwares and computer attacks is using for their expansion and purpose computer networks. Within the cost of information processing and Internet accessibility, organizations are becoming increasingly vulnerable to potential cyber threats such as network intrusions. A large scale of security systems is being used as a protection against them. There exists a need to provide secure and safe transactions through the use of Intrusion Detection Systems (IDS). Attackers can access the system from Internet or from inside as authorized users attempting to gain and misuse non-authorized privileges. Intrusion detection system (IDS) plays a role of the last defence line in overall system security. Although it doesn't replace direct access control devices it is successful for monitoring and violation detection of network traffic attempts or anomaly activities. Thereby it provides important information for countermeasures in time [1].

Network intrusion detection systems are faced with the challenge of identifying diverse attacks, in extremely high speed networks. For this reason, they must operate at multi-Gigabit speeds, while performing highly-complex per-packet and per-flow data processing. It is increasingly challenging to prevent the attack attempts at the edge of the Internet.

While many high-performance IDS employ dedicated network processors or special memory to meet the demanding performance requirements, it often increases the cost and limits functional flexibility. In contrast, existing software-based IDS stacks fail to achieve a high throughput despite modern hardware innovations such as multicore CPUs, multicore GPUs, and 10 Gbps network cards that support multiple hardware queues [2].

II. Background and Related Work

Intrusion detection systems are based on two fundamental approaches: the detection of anomalous behaviour as it deviates from normal behaviour, and misuse detection by monitoring those "signatures" of those known malicious attacks and system vulnerabilities. IDS produces alerts, messages or reports for administrators. There are three steps in the process of intrusion detection:

- monitoring and analysing traffic,
- identifying abnormal activities,
- assessing severity and raising alarm.

The potential goals of IDS usage include detection and prevention attacks, detection of policy violations, enforcement of use and connection policies and collection of evidence [3].

A. IDS performance issues

Intrusion detection systems must operate at multi-Gigabit speeds, while performing highly-complex per-packet and per-flow data processing. Thus, they (mainly signature-matching) can have two types of performance limitations [4]:

- CPU-bound limitations,
- I/O-bound limitations.

This is caused by the overhead of reading packets from the network interface card and high computing load in detection engine.

B. Research trends

Detection mechanism network of IDS systems uses the most system features and in some cases it can occupy up to 75% of modern IDS systems' performance. The main operation of detection engine is a comparison of incoming network packets to a set of signatures of known attacks. The comparison of packets to signatures is a string-matching operation which is CPU-intensive. Research into improving the runtime of the packet comparison falls into three categories:

- improvements in the string-matching algorithm in detection engine of IDS,
- utilizing packet-header characteristics to optimize comparison,
- usage of specialized hardware to improve IDS performance (GPU, FPGA,...).

The first two approaches are complementary and have been incorporated into most software-based IDSs, while the third approach has been widely deployed in router technology [1].

III. Summarization of the Research

Research is focused on predictive analysis and reaction during string-matching process in detection mechanism. Multiple-string-matching process uses some well-known algorithms like Boyer-Moore [5] or Wu-Mamber [6]. However, Aho-Corasick [7] algorithm (AC) provides high performance for the worst and the most complex cases. The algorithm excels in its robustness and in detecting various attacks, which is why the majority and widely deployed in today's IDS detection mechanisms [8]. The algorithm is based on tries defined by deterministic finite-state automata as it is shown on Fig. 1.

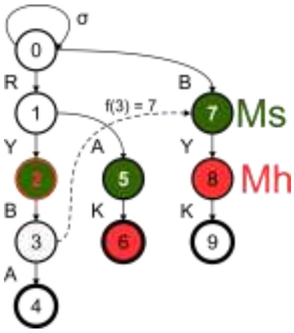


Fig. 1 Aho-Corasick algorithm used in predictive intrusion analysis with example patterns: RYBA, RAK, BYK

Process of predictive analysis is based on AC algorithm where we can predict (with certain probability) intrusion before finalization string-matching process. As it show on Fig. 2, AC provides an O(n) linear complexity, regardless of the number of patterns. We can prioritize different kinds of intrusion as it is shown on Fig. 2.

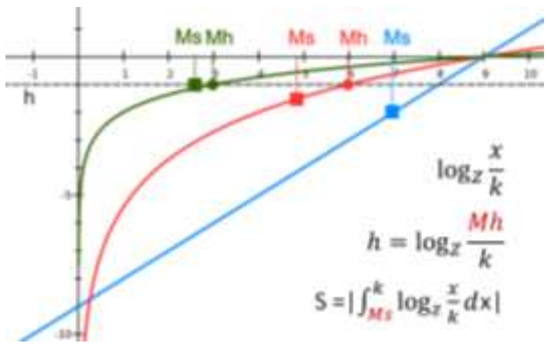


Fig. 2 Definition of intrusion detection priority Z and markers computation of predictive intrusion analysis Mh and Ms with pattern length k.

Intrusion detection priority leads to a focus on the most dangerous attacks and getting ability to prevent and prepare system to execute adequate reaction. In standard process algorithm results input string as possible attack when DFA reach final state. Our approach reaches possible attack with certain probability earlier with computation of markers (Ms, Mh). Computation of markers may be acquired by two ways (see in Fig. 2):

- defining level value h and gaining marker Mh,
- defining predicted intrusion area S and gaining marker Ms.

According to AC algorithm, pattern has finite number of states defined by length k. If computation state x reach Ms or Mh value, IDS results predictive alert.

This approach we applied on the proposed IDS architecture [9] consists of number (at least two) of standalone IDS stations which use the computing power of the graphics hardware (Fig.3). Cooperation between the IDS stations is covered by the central control node – mentor [10].

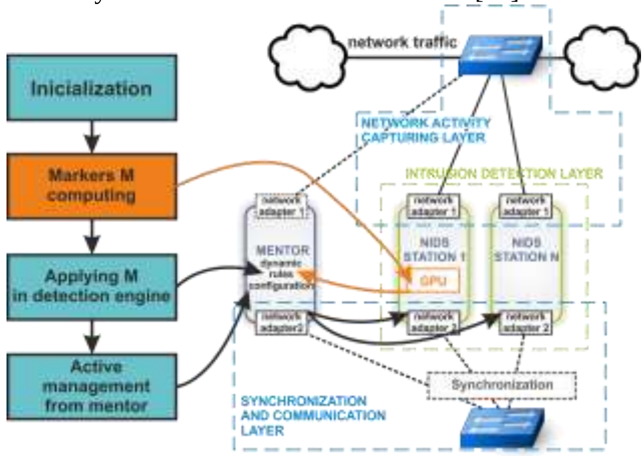


Fig. 3 Final application of predictive intrusion analysis on proposed IDS architecture.

IV. Future Work

Predictive intrusion analysis brings some issues that need to be solved:

- define evaluation intrusion parameter Z according to existed case-studies, related research etc.,
- setting common markers for similar kinds of intrusion,
- consider false functions in AC algorithm.

Acknowledgment

This work was supported by the Slovak Research and Development Agency under the contract No. APVV-0008-10 and KEGA 008TUKE-4/2013 Microlearning environment for education of information security specialists.

References

[1] A Survey of Parallel Intrusion Detection on Graphical Processors : Liberios Vokorokos ... [et al.]. 2013. Central European Journal of Computer Science [online]. Vol. 4, no. 4 (2014), p. 222-230. ISSN 2081-9935.

[2] M. Jamshed, J. Lee, S. Moon, I. Yun, D. Kim, S. Lee, Y. Yi, and K. Park. ”Kargus: a highly-scalable software-based intrusion detection system”, In Proceedings of the ACM Conference on Computer and Communications Security (CCS), 2012

[3] L. Vokorokos, A. Baláž, M. Chovanec, “Distributed detection System of Security Intrusion Based on Partially Ordered Events and Patterns,” Studies in Computational Intelligence 243, Budapest, September 2009.

[4] N. Jacob, C. Brodley, “Offloading IDS Computation to the GPU,” Computer Science Department, Tufts University, Medford, 2006.

[5] Robert S. Boyer, and J Strother Moore. A Fast String Searching Algorithm. Communications of the ACM, 20:762–772, October 1977.

[6] San Wu, and Udi Manber. A Fast Algorithm for Multi-Pattern Searching. Technical report, 1994.

[7] A. V. Aho and M. J. Corasick. Efficient String Matching: An Aid to Bibliographic Search. Communications of the ACM, 18:333–340, June 1975.

[8] G. Vasiliadis, M. Polychronakis, and S. Ioannidis. MIDeA: A Multi-Parallel Intrusion Detection Architecture. In Proceedings of the ACM Conference on Computer and Communications Security (CCS), 2011.

[9] A Distributed Network Intrusion Detection System Architecture Based on Computer Stations Using GPGPU : L. Vokorokos ... [et al.] 2013. In: INES 2013 : IEEE 17th International Conference on Intelligent Engineering Systems : proceedings : June 19-21, 2013, Costa Rica. - Budapest : IEEE, 2013 P. 323-326. - ISBN 978-1-4799-0828-8

[10] A control node for intrusion detection systems management / Liberios Vokorokos [et al.] - 2014. In: Acta Electrotechnica et Informatica. Roč. 14, č. 3 (2014), s. 28-31. - ISSN 1335-8243

Influence of performance monitoring on operation efficiency and problem root cause evaluation in productive IT environment

[1]*Matej KOSTROŠ (1st year)*
Supervisor: [2]František JAKAB; Consultant: [3]Jozef Janitor

[1,2,3]Dept. of Computers and Informatics, FEI TU of Košice, Slovak Republic

[1]matej.kostros@tuke.sk, [2]frantisek.jakab@cnl.sk, [3]jozef.janitor@cnl.sk

Abstract— **In this paper is elaborated the problematics of information systems and their first level maintenance in operational management. It discuses the background of such activities starting with basic theoretical concepts, followed by obstacles which may arise during the operation in large heterogeneous multi node systems, next are considered system availability problems and at the end applications and operating system monitoring fundamentals. In practice lots of today's operating systems can be supervised by a software designed for monitoring various events, collecting performance metrics and if desired taking corrective action when needed.**

In second part of the paper are presented usual techniques used mainly in practice for dealing with repetitive problems in operation. With the connection to theoretical background this paper will create basis for author's next research for Big Data analysis in corporation environment for increasing quality of provided services by improving efficiency of whole process using automation.

Keywords— **Performance, monitoring, system management, data analysis, heterogeneous systems, automation, maintenance, root cause evaluation**

I. Introduction

Current operating systems allow to install specialized software for analyzing and processing application and kernel related tasks and activities [1]. From system management point of view there is a little difference whether the system is virtualized or not, when we look at it as a monitored device. Main point of such tools is to observe, collect and if needed control the behavior of the system for better service delivery to the end user. There was performed a range of studies aimed to heterogeneous systems and their monitoring and system management [2]. This paper is therefore focused on a specific problematics of such systems mainly from operator's point of view and his/her efficiency when working with tools aiding such related tasks. One of the most important duties is to control availability of service. High availability is often demanded by a customer due to a high cost of outages and downtimes, therefore lots of features are added by developers to increase and enhance availability of software products [3]. Network management nor system management alone is sufficient for overseeing availability, but only their combination can bring wanted result [5]. On the other hand, in practice various license and contract related obstacles may limit the control over operations. With the need of bigger infrastructure also the demand for the restructuralization grows to better utilize limited staff and minimize possible human errors. Therefore the need for automation of known errors is one of the most wanted features. Efficiency of such tools can be evaluated by increasing mean time between failures and minimizing mean time to repair for specific tasks [6]. Thus, there arises need for improving existing and creating new tools for automation in operations environment.

II. Problems present in real life productive operations environment

Productive operations environment in data related enterprise can be regarded as a highly dynamic system, with deterministic as well as stochastic behavior [8]. This view of the operations will allow us to identify internal and external variables affecting the operational work-flow. External variables are those independent variables which are not influenced by anything else in the operations [9]. External or “exogenous” variables can also be defined as those that perturb the data center into its characteristic behavior determined by its structure [10]. It is very important to understand the structure of operations in data center which includes its attributes and deep elements.

In practice there is a certain difference between desired attributes of operation and the real ones. For instance, operator which is dealing with immediate problem cannot always intensively focus on a single problem, when there is a time schedule defined by service or operation level agreement and often the basic issue which can be handled on the first level support is forwarded or routed to the higher one. Here we understand the first level support according the multi-tiered technical support concept [11] as a initial point of contact with a problem. So their task is to collect needed information and determine the customer's issue by analyzing the symptoms and figuring out what the underlying problem may be causing issue. Reality is sadly different. First level operators deals only with basic analysis, mostly only correct routing to solving group of higher level support teams almost without any investigation, issue correlation or corrective actions whatsoever, even if there are prepared tools relevant to the issue. This often causes prolonging the mean time to repair and thus deteriorate KPI of solving groups by SLA or OLA breach.

Operators of IT resources are continuously becoming service providers. In this role operator needs new concepts and tools to provide service according SLA or OLA. To guarantee the quality of service, such as response time, availability, the operator needs to control the whole process of IT service life cycle. In the past IT service providers, operators, mostly depended on their intuition and partially experience. Knowledge gained from experience mostly controlled the way how the infrastructure components like servers, networks elements, or other infrastructure nodes were maintained [12]. This situation has changed mainly because of the upgrading to more decentralized and dynamic computing service platforms like cloud, virtual systems, etc. It was necessary to change, because obviously running such complex and heterogeneous system requires standardization for operation services to maintain certain quality of service demanded by customer who in turn is willing to pay an agreed price. These standards were focused in many standards, but one of the most used is ISO 20000 [13].

Tools for aiding system management operations, brings many advantages to operators, mainly concentrating important information into smaller number of interfaces. In practice, first level support operator is often challenged in manner of his concentration, ratio of configuration items to the personal often very high. This will increase the probability of interruption of thought process. For such cases the selection of appropriate tool is very important.

III. Performance monitoring through HP Operations and Performance Agent

Continuous performance monitoring is crucial to detect software aging and thus justify reason to enable performance tuning by software or hardware way. Software aging is phenomenon which implies that its state is degrading with time. This degradation is visible mainly in performance declination or hang/crash events. It has been observed, that this degradation results, particularly from operating system resources exhaustion. To prevent outages and downtimes from these degradation processes, we need effective performance and system monitoring tool which will collect data of runtime performance parameters in order to forecast software aging and when it reaches its critical levels [16].

Company HP provides tool aids and software for complete support of system management in big enterprises including performance metric collection and automatic alert generation. These tools create sufficient background for operators to handle issues presented in previous chapter. According their philosophy today's IT professionals must do more with less, therefore system contains basic monitoring, performance metering and alerts which send events to unified interface where operators can handle problems directly by using tools, or by incident ticket creation forward issue to higher level support team [14].

Effective support team needs to have correct data in reasonable time window. In practice, when system is not configured properly, operators have to deal with delayed events, and thus under time pressure are more probable to make mistakes. Therefore data collecting infrastructure must be reliable and able to process requests from whole environment. Operations Manager includes both agent based and agent less monitoring to ensure higher flexibility in monitoring infrastructure [Fig. 1].

Server includes a graphical interface that represents a single point from which operators can see, monitor and check both physical and virtual systems, networks, applications, storage, middleware. This console contains filters and role based views, which better structure event flow for operators. Built in reporting and performance tools allows bottlenecks analysis on one or more systems, thus streamlining operational processes and enhancing efficiency [15].

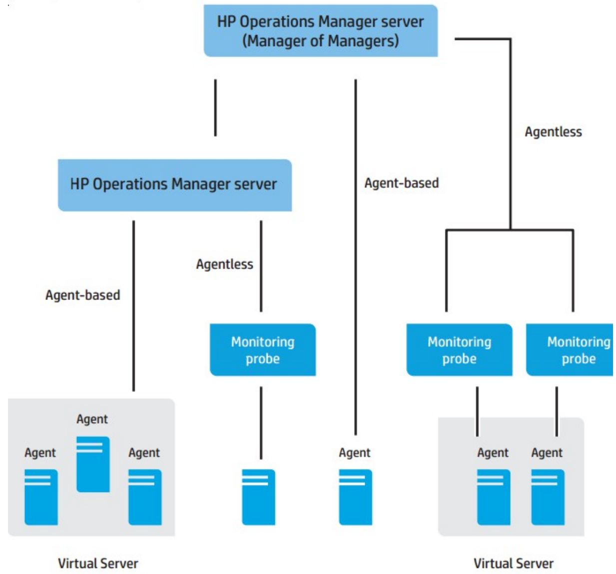


Fig. 1 HP Operations Manager infrastructure [14]

In practice operators equipped with these tools are encouraged to create their own tools, or at least create requests to engineering teams for automation common problem solving. However workload is mostly the reason, that they are not interested in such activities. Here we want to continue with a research, to find out, why is the situation such and to improve the automation by increasing efficiency by delivering more suitable tools to operators.

IV. Root cause evaluation for performance bottlenecks using descriptive statistics

In production environment there is need to know in time when system performance is decreasing. This need is due to nature of some contracts where end user is paying for consumed performance units instead of regular fee. It have advantages that user doesn't pay when system is not used, but on the other hand, user is complaining, if system utilizes more as expected.

The typical methodology for evaluation, systems requires reference system which can be used as a model for comparison. Usually testing processes are repeated numerous times under different conditions including load levels, system kernel settings, network configuration, etc [17]. From these data can be derived as presented in [17] that the response times of the system are proportional to the load. Therefore metrics related to system load will be most relevant for root cause analysis of system slowing.

HP Performance agent is collecting hundreds of metrics [18], which provides great source of data, but very important is correlating metrics with actual problems. Due to a mentioned time density of events and broad expertise to analyze them, the aim is to send event in as correlated manner as possible so the experience stress to operators in minimized [19].

In production environment we have performed analysis of approximately two and half hundred systems. We have detected handful of nodes utilizing CPU for more than 5 hours at more then 95% of CPU time. In [Fig. 2] is showed simple graph which encompass 7 day interval when probes were active and collecting data, every event occurrence means, that at least one core of CPU was utilized over 95%

threshold for more than 5 hours.

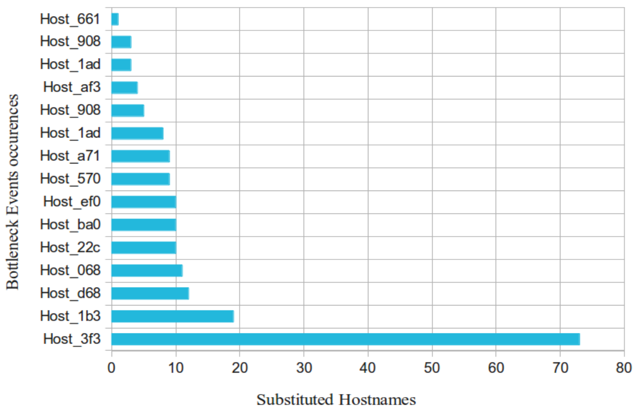


Fig 2. Graph of analyzed systems where at least one occurrence of testing event was present. Note: Host names are substituted due to a security reasons

Only fifteen systems were utilized over specified thresholds. Decision about thresholds settings was based on historical data, so regular backups and maintenance processes are not selected into analysis.

By means of collecting, grouping and analyzing known data we have performed basic descriptive statistics to find out the root cause of the most utilized system [20]. With help of HP Performance Manager we obtained more detailed performance data and compared times of mentioned five hour peaks with times of scheduled actions, which were referred as the root cause. Among them were for example HP DataProtector backup processes, Oracle Database processes, SAP application processes, etc.

Most of incident tickets which were created from these events were resolved as temporary issue caused by backup processes. Therefore we firstly compared times of backup with CPU peaks, and the results was that they are completely uncorrelated, therefore with high certainty could be stated, that they were not responsible for causing a CPU bottleneck. This procedure was performed also manually on the rest of processes. The final findings of root cause were SAP application processes which consumed high amount of CPU time.

Lessons learned from this analysis was, that if operator forwarded these events to solving group responsible for SAP application, and they would consolidate processes to other cluster nodes, probably end user would not experience slow response times at all.

V. Conclusion

We have explored the basic problems of operator problems in productive environment. Findings are that design of automated root cause detector, based on the analysis from last chapter, with cooperation with various automated corrective tools in monitoring software suite would help with delivering better services by decreasing mean times to repair and increasing mean times between failures, which are most common reason of breached operation and service level agreements.

Acknowledgment

This paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF. We support research activities in Slovakia/This project is being co-financed by the European Union.

References

[1] Wu Qingbo; Wang Chunguang; Tan Yusong, "System Monitoring and Controlling Mechanism Based on Hypervisor," Parallel and Distributed Processing with Applications, 2009 IEEE International Symposium on , vol., no., pp.549,554, 10-12 Aug. 2009

[2] Bup-Ki Min; Hyeon Soo Kim; Seunghak Kuk; Sung Woon Park; Chumsu Kim, "Dynamic reconfigurable integrated management and monitoring system for heterogeneous distributed computing systems," Signal Processing and Information Technology (ISSPIT), 2012 IEEE International Symposium on , vol., no., pp.000037,000042, 12-15 Dec. 2012

[3] Heindel, L.E.; Schlieber, K.J., "Total systems management in open architecture provisioning systems," Local Computer Networks, 1994. Proceedings., 19th Conference on , vol., no., pp.411,416, 1994

[4] Moran, P.; Gaffney, P.; Melody, J.; Condon, M.; Hayden, M., "System availability monitoring," Reliability, IEEE Transactions on , vol.39, no.4, pp.480,485, Oct 1990

[5] Haberkorn, M.; Trivedi, T., "Availability Monitor for a Software Based System," High Assurance Systems Engineering Symposium, 2007. HASE '07. 10th IEEE , vol., no., pp.321,328, 14-16 Nov. 2007

[6] Schneeweiss, W.G., "Computing Failure Frequency, MTBF & MTTR via Mixed Products of Availabilities and Unavailabilities," Reliability, IEEE Transactions on , vol.R-30, no.4, pp.362,363, Oct. 1981

[7] Gubarev, V.V.; Chernov, V.U., "The adaptive automated systems of statement and problem solving of a data analysis," Science and Technology, 2001. KORUS '01. Proceedings. The Fifth Russian-Korean International Symposium on , vol.1, no., pp.140,143 vol.1, 26 Jun-3 Jul 2001

[8] Ayman, N.; Hossain, H.; Faramarz, D., "A System-based Approach for Defining IT Operations Value Proposition," Management Science and Engineering, 2006. ICMSE '06. 2006 International Conference on , vol., no., pp.1767,1772, 5-7 Oct. 2006

[9] A. Diamantopoulos and J. Siguaw, Introducing LISREL, Sage publications, 2000.

[10] H. Kurstedt, It's about systems thinking leading to effective actions, INCOSE tutorial, April 23, 2005.

[11] Walker, Gary (2001). IT Problem Management (Harris Kern's Enterprise Computing Institute Series). Upper Saddle River: Prentice Hall. pp. 85–113. ISBN 0-13-030770-X

[12] Abeck, S.; Mayerl, C., "Modeling IT operations to derive provider-accepted management tools," Integrated Network Management, 1999. Distributed Management for the Networked Millennium. Proceedings of the Sixth IFIP/IEEE International Symposium on , vol., no., pp.843,856, 1999

[13] Randy A. Steinberg (2011). ITIL Service Operation. The Stationery Office. ISBN 978-0113313075

[14] HP Operations Manager Software Data sheet, available at: http://www8.hp.com/h20195/V2/GetPDF.aspx/4AA1-6028ENW.pdf

[15] HP Operationas Manager Java GUI Operator's Guide, available at: https://softwaresupport.hp.com/group/softwaresupport/search-result/-/facetsearch/document/KM00960401

[16] Jian Xu; Manwu Xu, "A Performance Monitoring Tool for Predicting Degradation in Distributed Systems," Web Information Systems and Mining, 2009. WISM 2009. International Conference on , vol., no., pp.669,673, 7-8 Nov. 2009

[17] Dror G. Feitelson, "Experimental Analysis of the Root Causes of Performance Evaluation Results: A Backfilling Case Study", IEEE Transactions on Parallel & Distributed Systems, vol.16, no. 2, pp. 175-182, February 2005, doi:10.1109/TPDS.2005.18

[18] Dictionary of Performance Metrics for the Performance Collection Component (All platforms), available at: https://softwaresupport.hp.com/group/softwaresupport/search-result/-/facetsearch/document/KM1020019

[19] Carjeu, I.G.; Shorrock, T.; Seeger, M., "Clustering IT Events around Common Root Causes," Services Computing (SCC), 2014 IEEE International Conference on , vol., no., pp.749,757, June 27 2014-July 2 2014

[20] Vergura, S.; Acciani, G.; Amoruso, V.; Patrono, G.E.; Vacca, F., "Descriptive and Inferential Statistics for Supervising and Monitoring the Operation of PV Plants," Industrial Electronics, IEEE Transactions on , vol.56, no.11, pp.4456,4464, Nov. 2009

Intelligent space for robotics

[1]Jakub HVIZDOŠ (1st year)
Supervisor: [2]Peter SINČÁK

[1,2]Dept. of Cybernetics and Artificial Intelligence, FEI TU of Košice, Slovak Republic

[1]jakub.hvizdos@tuke.sk, [2]peter.sincak@tuke.sk

***Abstract*—This works aims to provide an overview of various intelligent spaces and their applications in the context of robotics. As intelligent spaces vary greatly a categorization according to the sensors they utilize was used for improved clarity.**

***Keywords*—intelligent space, navigation, sensors, robotics**

I. Introduction

By intelligent space (iSpace) we understand environments with embedded systems, information and communication technologies that create an interactive environment (**Figure 1**). The purpose of intelligent space is to combine computational technologies to ease or enhance the activities that take place in them.

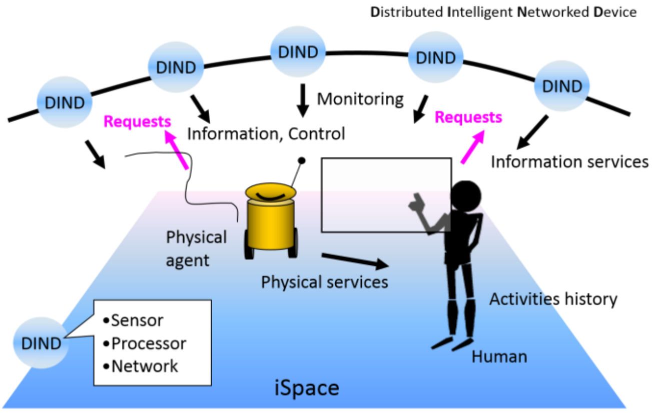


Figure 1 – Intelligent space [6]

This requires a combination of sensors and actuators. In the context of robotics actuators are provided by a robot or robots while the space itself provides them with additional information from the sensors embedded in the environment that would otherwise be difficult or impossible to obtain. Shape and sensor layout of the intelligent space depends on the specific place where it is constructed. However, sensors and the information they provide tend to be similar and can thus be used to categorize various examples of intelligent spaces.

II. Camera based iSpace

Since vision is one of our primary senses, understanding and working with visual data is very intuitive for us. Therefore even intelligent spaces that do not rely primarily on vision tend to use it to some degree. In the case that an intelligent space does primarily utilize vision a wide variety of information can be collected from vision alone.

One such example is the intelligent space described in Intelligent Space for Human Centered Robotics [1] (**Figure 2**).

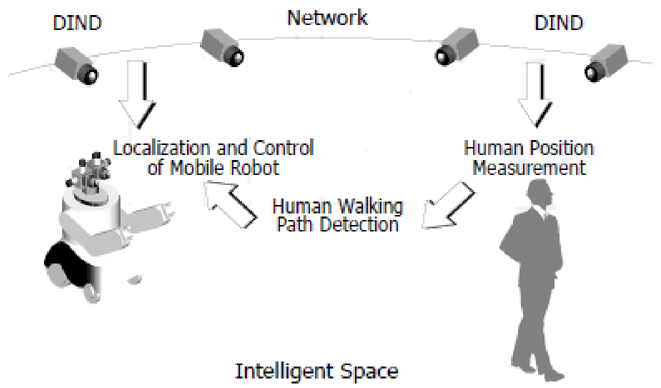


Figure 2 – Human following application [1]

Intelligent devices are distributed throughout the whole of the space. These intelligent devices have sensing, processing and networking functions, and are referred to as distributed intelligent networked devices (DINDs). One DIND consists of a CCD camera for acquiring space information, and a processing computer which has the functions of data processing and network interfacing. These devices observe the positions and behavior of both humans and robots coexisting in the intelligent space forming a multiple camera multiple object tracking system. Based on the accumulated information, the environment as a system is able to understand the intention of human beings (**Figure 3**). A hybrid tracking algorithm was created, which included MeanShift and Kalman filter for object tracking in one DIND. To track target objects in a wide area and to control mobile robots based on the environmental measurements, cooperation of the DINDs, effective communication and role assignment are required. As an application of the intelligent space for human centered robotic system a human-following mobile robot was introduced. This human-following robot based its movements on the measurements of the intelligent space with multiple DIND and a virtual spring model.

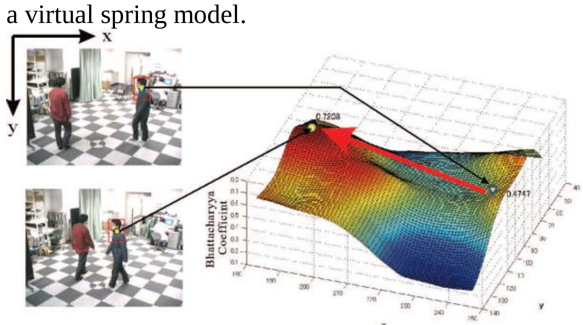


Figure 3 – Acquired measurements [1]

Another example is the intelligent space from Design Policy of Intelligent Space [7]. Here the images captured by the cameras were processed with a background separation algorithm for the purpose of locating moving objects. To increase the robustness of the system color information was used to help locate human hands and heads. The positions of heads indicate traversable places in the intelligent space and were then used to generate a 3D map that the robots could use for navigation (**Figure 4**).

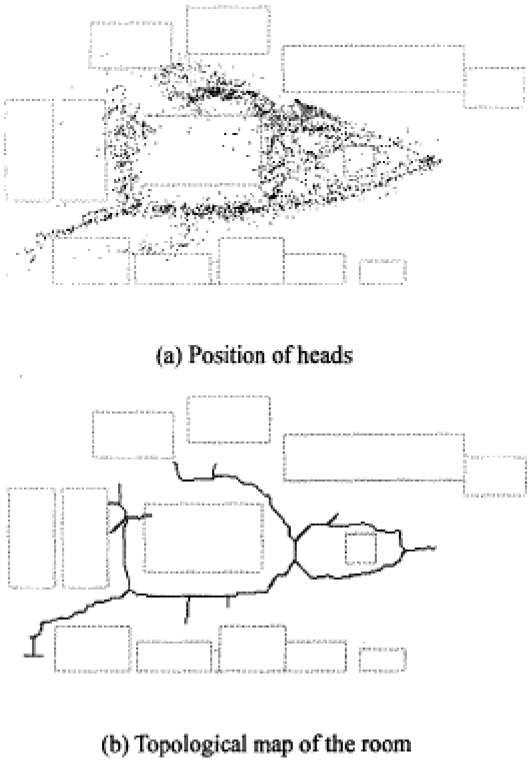


Figure 4 – Map generation [7]

The last example of the primarily camera utilizing intelligent space is Self-Identification of Distributed Intelligent Networked Device in Intelligent Space [10]. In this case several cameras placed at various angles were used to track the position of a robot or human in space (**Figure 5**).

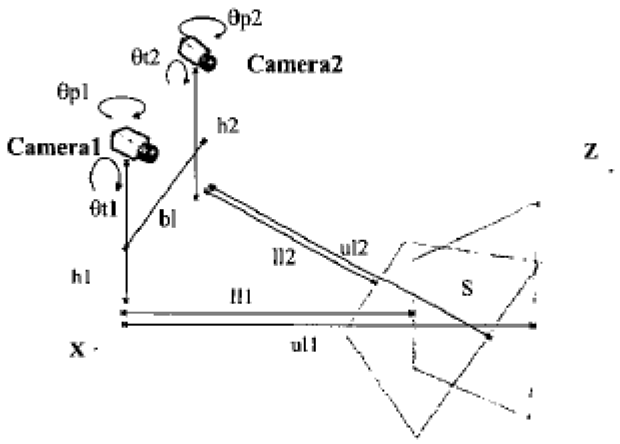


Figure 5 – Camera positioning [10]

The positions of the cameras in the intelligent space were known and were specifically set up so that certain areas of their vision would overlap (**Figure 6**). This allowed the final position to be determined.

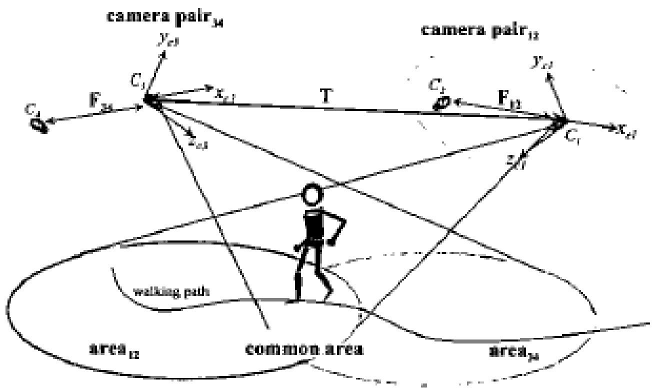


Figure 6 – Visual overlap [10]

Despite its availability and intuitive use by humans visual information does have some disadvantages. The main disadvantage being, that to be usable by machines, visual information needs to be process by some usually quite complicated method or methods. Also depending on the image quality this information may not be always reliable. These problems may be alleviated by using some of the following sensors instead or in addition to cameras.

III. Spatial sensor using iSpace

By spatial sensors we generally understand active sensors such as laser range-finder or Kinect sensors. Their advantage over cameras is that they automatically acquire spatial depth information without having to use complicated image processing algorithms.

The first example is described in Acting in Intelligent Space [13]. Here distributed laser range finders were used to map the environment (**Figure 7**).

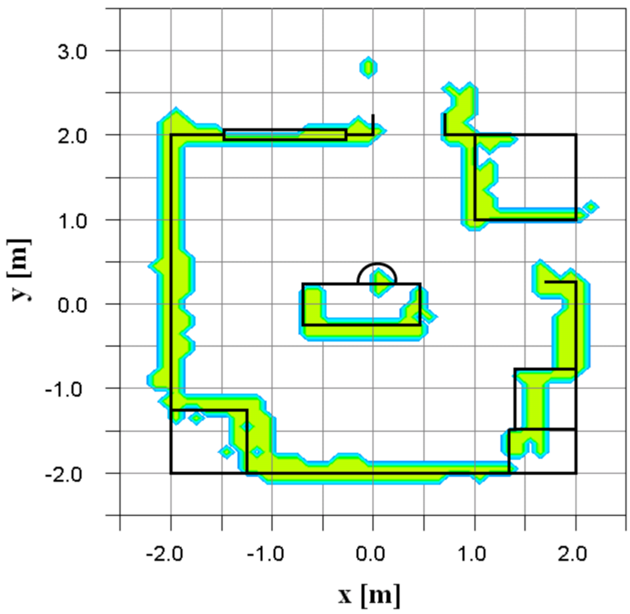


Figure 7 – Mapping result [13]

The acquired information was then processed clustering and data association algorithms and further with Kalman filter to track objects and people in the intelligent space. This way a robot was able to path its way between the obstacles towards a human user.

The second example is described in the fifth chapter of Human Robot Interaction in Social Robotics [2]. In this case a network of laser range-finders was used. Data form this network was then used by a system for simultaneous position and body orientation tracking of multiple people. This system utilized methods of particle filtering and motion analysis to determine orientation and walking direction of multiple people (**Figure 8**).

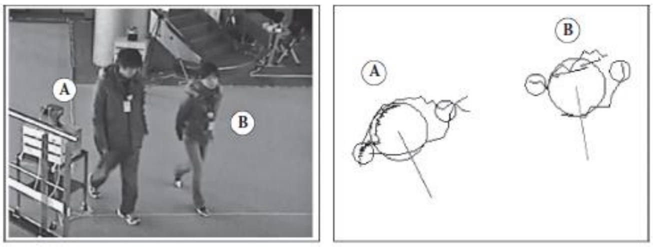


Figure 8 – Laser based tracking [2]

IV. RFID SENSORS

Radio-frequency identification (RFID) uses electromagnetic fields to automatically identify and track special tags. Such system consists of an active sensor (reader) and a passive chip (tag) (**Figure 9**).

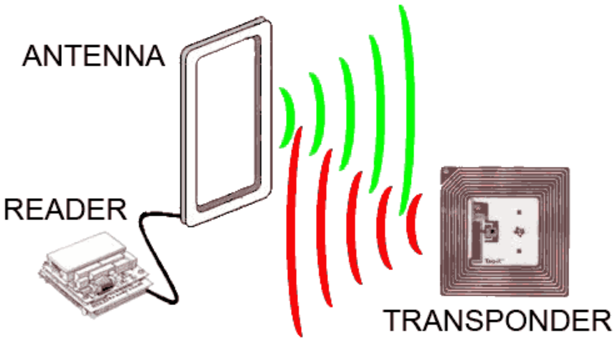


Figure 9 – RFID system [16]

In Building an Intelligent Home Space for Service Robot [17] RFID tags (**Figure 10**) were distributed across the intelligent space. By reading these tags a robot was able to determine its position in intelligent space.

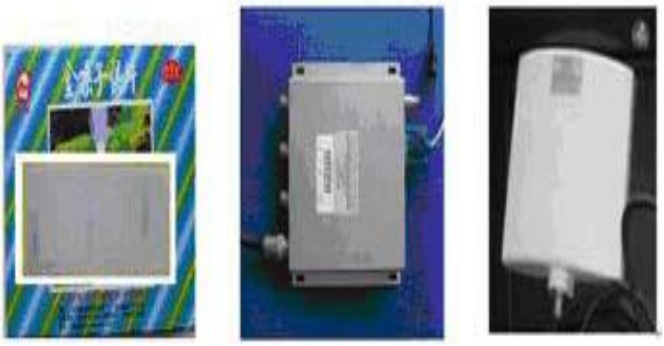


Figure 10 – Examples of RFID tag, reader and antenna [17]

QR codes (**Figure 11**) were then utilized to mark specific objects which the robot had to then identify by using its cameras and attempt to grasp.

Figure 11 - QR codes [17]

Another example is described in the second chapter of Human Robot Interaction in Social Robotics [2]. Here museum visitors wore RFID tags and the information obtained from these tags was then used to direct the robots' behavior. The robots could interact with the visitors using gestures or guide them towards various exhibits (**Figure 12**).

Figure 12 - Human-robot interaction [2]

V. OTHER SENSORS

Although previously mentioned sensors tend to be more common, there are no restrictions on what kinds of sensors an intelligent space can use.
For example chapter five of Human Robot Interaction in Social Robotics [2] describes the use of floor sensors to track multiple people in a crowd (**Figure 13**). Combined with the aforementioned RFID technology, movements of specific people could be tracked.

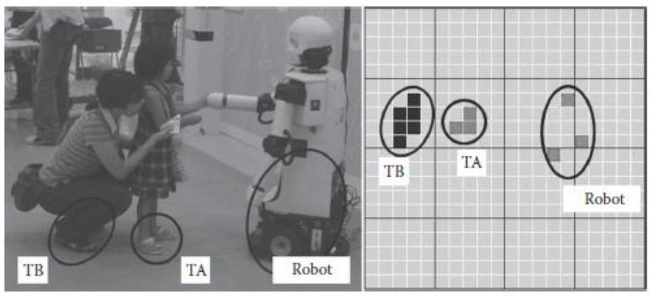


Figure 13 - Tracking using floor sensors [2]

Data form the floor sensors was then used to analyze group dynamics as well. That is identifying groups of people that travel together, the direction in which they are moving and whether they are clustering around the robot and paying attention to it.

VI. Conclusion

There are no set when deciding what sort of sensors should an intelligent space utilize. Specific technologies should be chosen based on the location and intended function of the intelligent space. However, for increased robustness of acquired data various different sensors provide an advantage. Therefore for our future work we like to create our own iSpace that would utilize various kinds of sensors to gather information. This information will then be used to help augment robotic experiments. We believe that by providing additional information to robots we can achieve greater variety and robustness of our experiments.

Acknowledgment

Research supported by the National Research and Development Project Grant 1/0667/12 “Incremental Learning Methods for Intelligent Systems” 2012-2015.

References

[1] Morioka, Kazuyuki, Hideki Hashimoto, and Joo-Ho Lee. *Intelligent Space for Human Centered Robotics*. INTECH Open Access Publisher, 2008.

[2] Kanda, Takayuki, and Hiroshi Ishiguro. *Human-robot Interaction in Social Robotics*. CRC Press, 2012.

[3] Ishiguro, Hiroshi, et al. "Robovie: an interactive humanoid robot." *Industrial robot: An international journal* 28.6 (2001): 498-504.

[4] Burgard, Wolfram, et al. "The interactive museum tour-guide robot." *AAAI/IAAI*. 1998.

[5] Montemerlo, D., Sebastian Thrun, and William Whittaker. "Conditional particle filters for simultaneous mobile robot localization and people-tracking." *Robotics and Automation, 2002. Proceedings. ICRA'02. IEEE International Conference on*. Vol. 1. IEEE, 2002.

[6] Human-system laboratory, http://www.mech.chuo-u.ac.jp/~hslab/index.php?ResearchList_en

[7] Lee, Joo-Ho, Noriaki Ando, and Hideki Hashimoto. "Design policy of intelligent space." *Systems, Man, and Cybernetics, 1999. IEEE SMC'99 Conference Proceedings. 1999 IEEE International Conference on*. Vol. 3. IEEE, 1999.

[8] Guido Appenzeller, Joo-Ho Lee, Hideki Hashimoto, “Building Topological Maps by Looking at People: An Example of Cooperation between Intelligent Spaces and Robots”, *IEEE/RSJ International Conference on Intelligent Robots and Systems.* 1997.

[9] C. Wren, A. Azabayejani, T. Darrel, and A. Pentland, “Pfinder: Real-time tracking of the human body”, *in International Conference on Automatic Face and Gesture Recognition*, pp. 51-56, Oct. 1996.

[10] Hashimoto, Hideki, Joo-Ho Lee, and Noriaki Ando. "Self-identification of distributed intelligent networked device in intelligent space." *Robotics and Automation, 2003. Proceedings. ICRA'03. IEEE International Conference on*. Vol. 3. IEEE, 2003.

[11] Hashimoto, Hideki. "Intelligent Interactive Space based for Robots."*Proceedings of the 2000 International Symposium on Mechatronics and Intelligent Mechanical System for*. Vol. 21. 2000.

[12] Hashimoto, Hashimoto. "Intelligent space-How to make spaces intelligent by using DIND?." *Systems, Man and Cybernetics, 2002 IEEE International Conference on*. Vol. 1. IEEE, 2002.

[13] Brscic, D., Takeshi Sasaki, and Hideki Hashimoto. "Acting in intelligent space-mobile robot control based on sensors distributed in space." *Advanced intelligent mechatronics, 2007 IEEE/ASME international conference on*. IEEE, 2007.

[14] Krumm, John, et al. "Multi-camera multi-person tracking for easyliving." *Visual Surveillance, 2000. Proceedings. Third IEEE International Workshop on*. IEEE, 2000.

[15] Morioka, Kazuyuki, et al. "Seamless object tracking in distributed vision sensor network." *SICE 2004 Annual Conference*. Vol. 2. IEEE, 2004.

[16] Radio frequency identification, http://animalmigration.org/RFID/index.htm

[17] Fei, Lu, et al. "Building an intelligent home space for service robot based on multi-pattern information model and wireless sensor networks." *Intelligent Control and Automation* 2012 (2012).

[18] Y. H. Xue, G. H. Tian, R. K. Li and H. T. Jiang, “A New Object Search and Recognition Method Based on Artificial Object Mark in Complex Indoor Environment,” *The* 8*th World Congress on Intelligent Control and Automation*, Jinan, 7-9 July 2010, pp. 6648-6653.

[19] D. Smith and S. Singh, “Approaches to Multisensor Data Fusion in Target Tracking: A Survey,” *IEEE Transactions on Knowledge and Data Engineering*, Vol. 18, No. 12, 2006, pp. 1696-1710. doi:10.1109/TKDE.2006.183

[20] H. S. Carvalho, W. B. Heinzelman, A. L. Murphy, *et al.*, “ A General Data Fusion Architecture,” *Proceedings of the IEEE International Conference on Information Fusion*, Queensland, 8-11 July 2003, pp. 1465-1472.

Mining Rules from Database to Support Decision

Alexandra LUKÁČOVÁ (4th year)
Supervisor: Ján PARALIČ

Dept. of Cybernetics and Artificial Intelligence, FEI TU of Košice, Slovak Republic

alexandra.lukacova@tuke.sk, jan.paralic@tuke.sk

***Abstract*—In this paper we present the review of case studies focused on finding patterns in the data. The primary purpose of these case studies was to provide as much insight in given data as possible. Consequently, in combination with the view of the domain expert, the patterns were more relevant to the user. We want to emphasize, that correctly applied data mining methods can be helpful supporting tool in order to provide for experts simple understandable diagnostic patterns. We used decision trees, logistic regression and association rules mining for different purposes, e.g. also for selecting of most important attributes from predefined groups of attributes. All obtained results were evaluated by human expert in order to confirm correctness and validity of the discovered knowledge.**

***Keywords*— frequent patterns, data mining, decision trees, association rules**

I. INTRODUCTION

Although data are everywhere, it is not unusual that the domain experts do not know how to make right analyses and get some valuable information from them. Knowledge discovery (KDD - Knowledge Discovery in Databases) aims to extract knowledge from data [1] and the goal of the field of exploratory data mining is to provide a domain expert as much insight in given data as possible.

As it is written in [2] within exploratory data mining, pattern mining aims to enable the discovery of patterns from data. A pattern is a description of some structure that occurs locally in the data. This approach has advantage in interpretable representations, is understood by human experts and can thus provide explanations. Methods extracting patterns in form of decision trees have proved to be very successful and effective, they provide a classification structure and can be easily transformed also into form of decision rules. This is in contrast to classifiers like neural network models, which may provide nice classification results, but as a kind of black box. An example of such a rule might be that "There is 98% probability to have mild cognitive impairment if patient is male and has high Body Mass Index." The disadvantage is in amounts of patterns that are found, out of which many are redundant and background knowledge of the domain expert is not taken into account. Most of the time manual filtering of the results is necessary and data analyst needs be both a domain expert and a data mining expert [2]. It makes the job extremely time consuming.

In our data mining projects we follow the Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology. It typically consists of six phases - Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation, Deployment [3].

II. THE ACTUAL RESULTS

In [4]–[6] we focused on the possibility to use selected methods from machine learning theory to provide the answers on specified medical questions. The case study used data, that were collected in a family practice located in an urban area of the town of Osijek, the north-eastern part of Croatia. The aim was to determine whether the patient is suffering from metabolic syndrome, a well-known cluster of cardiovascular risk factors, components of which include central obesity, impaired glucose tolerance, hypertension, increased serum triglycerides and decreased HDL-cholesterol.

In general, use of data mining in medical data processing represents an interesting way how to discover hidden relations, new combinations of variables or boundaries for their values that can be further used as a decision support system for medical experts, doctors or researchers. Early diagnostics of the diagnosis can prevent the occurrence of any other diseases based on mainly lifestyle changes.

By applying data mining approach on the dataset prepared in a way that collected parameters from many aspects describe the health status of patients, we wanted to find out whether there are some important extensions to the classical definition of the diagnosis, to add value to clinical reasoning and to map novel variables and pathways that can be used to direct future research.

In our experiments we used two alternative instances of algorithm C4.5: J48 implemented in Weka data mining tool and C5.0 provided by SPSS data mining software. In both cases we have tested different parameters and their values to find the combination with good precision and optimal decision ability. Because of the small number of input records we have instead of the traditional division into training and test sample used 10-fold cross validation. Finally, we joined the best results from both methodological approaches to provide effective supporting mechanism for the diagnosis decision process. We obtained many interesting rules which can be used to test their practical usefulness on real-life data, or as an introduction for planning population-based research.

Afterwards we searched for their optimal cut-off values as follows. For finding the optimal cut-off points *c*, which best distinguish diseased and healthy patients, we used the measure called Youden index. Its advantage is in offering the best result with respect to the maximum overall correct classification by maximizing the sum of sensitivity and specificity. We considered the cut off values only in the case, when importance of particular variable was statistically significant (i.e. $p<0.05$). We used student's unpaired t-test to for this purpose. Experiments were performed by using R software with installed package "OptimalCutpoints" [7].

In [8] use of data mining represents an interesting way how to optimize the air traffic, to prevent a possible dangerous situation with negative consequences or to improve the customer services based on collected historical data. In our case we were focused on accidents during air traffic with the aim to evaluate the reasons of their occurrence with possibility to generate an easily understandable and cost effective prediction model. Decision trees were as in previous cases used to generate rules.

Evaluation of prediction model contained not only accuracy metric, but also recall and specificity, as they provide relevant view in case of imbalanced datasets. Based on this fact, we applied some balancing on training set. We selected random oversampling method (positive records were replicated) in order to improve the conditions for model generation, as undersampling and cost sensitive learning did not bring better results.

In [9] an objective was to identify the pattern of co-morbid disorders in older people with mild cognitive impairment diagnosis, which may increase risk of later progressing to dementia. Relevance of this paper is emphasized by the fact that this diagnosis is a potentially serious condition. Prevention planning would be especially important for the reason that available medications are not able to stop cognitive impairment progression.

The important step dealt with necessary discretization procedure since algorithms for associations rule mining require discrete attributes. We applied two alternative ways. Fist one included the ChiMerge algorithm that uses the χ^2 statistic [10] and takes the class of the example into considerations when constructing intervals. We used package implemented in R [11]. ChiMerge algorithm has the tendency to construct too many intervals, so we also tried for comparison the K-Means clustering algorithm, implemented in R [12], [13]. Our experiments with K-Means algorithms resulted into value 3 as a good input number of clusters. We used these two alternatives as a replacement to the basic method of discretization with fixed length window, with which some relevant patterns can be missed if split is made by badly placed boundary points. Consequently we focused on association rules mining which was added by combination of logistic regression modelling and exploratory analysis. As expected, larger number of rules was generated. This paper represented the first step for establishing collaboration between data mining research groups and application domain.

III. Conclusion

Applying a range of data mining algorithms to data sets could improve the quality of the predictions in regards of accuracy but also efficiency and robustness. It is also important to actively involve the user in the exploratory data mining process in order to discover more interesting results [2]. A practical problem is that with support and confidence often too many association rules are produced. Although there are other measures that can exclude some redundant rules, the most important goal is to develop methods for learning, which uses task specific interestingness. It could provide the helpful decision support, which may result to faster diagnosis.

Interesting view on this problematic bring papers [14] and [15], where human in the loop decision support system (HIL-DSS) is developed. They give an example with a scenario where human experts use DSS to make decisions but not always follow them, i.e. the humans use expertise on top of the DSS to make the final decision. This could be a way to eliminate not only redundant, but also uninteresting rules.

Acknowledgment

This publication is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF.

References

[1] J. Paralič, *Objavovanie znalostí v databázach*. Košice: Elfa, 2003.

[2] M. van Leeuwen, "Interactive Data Exploration Using Pattern Mining," in *Interactive Knowledge Discovery and Data Mining in Biomedical Informatics*, A. Holzinger and I. Jurisica, Eds. Springer Berlin Heidelberg, 2014, pp. 169–182.

[3] P. Chapman, J. Clinton, R. Kerber, T. Khabaza, R. Wirth, T. Reinartz, and C. Shearer, "CRISP-DM 1.0 - Step-by-step data mining guide." Aug-2000.

[4] F. Babič, A. Lukáčová, and J. Paralič, "Diagnostika metabolického syndrómu ako riadený proces dolovania v dátach," in *WIKT 2014*, Smolenice, Slovakia, 2014, pp. 29–34.

[5] F. Babič, L. Majnarić, A. Lukáčová, J. Paralič, and A. Holzinger, "On Patient's Characteristics Extraction for Metabolic Syndrome Diagnosis: Predictive Modelling Based on Machine Learning," in *Information Technology in Bio- and Medical Informatics*, M. Bursa, S. Khuri, and M. E. Renda, Eds. Springer International Publishing, 2014, pp. 118–132.

[6] F. Babič, L. Majnaric, A. Lukáčová, J. Paralič, and A. Holzinger, "Decision rules to Support Metabolic Syndrome Diagnostics," 2015, unpublished.

[7] M. Lopey-Raton and M. X. Rodriguez-Alvarez, "Package 'OptimalCutpoints'."19-Sep-2013.

[8] F. Babič, A. Lukáčová, and J. Paralič, "Descriptive and Predictive Analyses of Data Representing Aviation Accidents," in *New Research in Multimedia and Internet Systems*, A. Zgrzywa, K. Choroś, and A. Siemiński, Eds. Springer International Publishing, 2015, pp. 181–190.

[9] L. Majnarić, A. Lukáčová, F. Babič, J. Paralič, Z. Paraličová, P. Chouhan, and A. Holzinger, "Pattern's Identification for Mild Cognitive Impairment Diagnosis," 2015, unpublished.

[10] R. Kerber, "ChiMerge: Discretization of Numeric Attributes," in *Proceedings of the Tenth National Conference on Artificial Intelligence*, San Jose, California, 1992, pp. 123–128.

[11] K. HyunJi, *R package: Data preprocessing, discretization for classification*. 2010.

[12] M. Hahsler, C. Buchta, K. Hornik, and C. Borgelt, *R package: Mining Association Rules and Frequent Itemsets*. 2014.

[13] M. Hahsler and S. Chelluboina, "arulesViz: Visualizing Association Rules and Frequent Itemsets." 2014.

[14] R. Pinto, T. Mettler, and M. Taisch, "Managing supplier delivery reliability risk under limited information: Foundations for a human-in-the-loop DSS," *Decis. Support Syst.*, vol. 54, no. 2, pp. 1076–1084, Jan. 2013.

[15] H. S. Subramania and V. R. Khare, "Pattern classification driven enhancements for human-in-the-loop decision support systems," *Decis. Support Syst.*, vol. 50, no. 2, pp. 460–468, Jan. 2011.

Modelling Methods, Control Design and Diagnostics Applications in the Distributed Control Systems

[1]*Matej ORAVEC (1st year)*
Supervisor: [2]Anna JADLOVSKÁ

[1,2]Dept. of Cybernetics and Artificial Intelligence, FEI TU of Košice, Slovak Republic

[1]matej.oravec@tuke.sk, [2]anna.jadlovska@tuke.sk

***Abstract*— This paper contains a review of modelling methods, control design and diagnostics methods to the systems with distributed architecture. Paper includes review of the worldwide research in modelling, control design and fault diagnostics of the simulation models and the real processes. It includes a lot of many approaches to modelling of the systems which are a part of the Distributed Control Systems. Also are presented possible methods of control design and fault detection in the levels of the Distributed Control Systems.**

***Keywords*—Distributed control system, modelling, experimental identification, control design, adaptive control, predictive control, fault diagnosis, assembly lines.**

I. Introduction

Distributed control systems (DCS) are complex systems built from many parts arranged in the levels from the lowest part, represented by real processes and models to the higher management part. Every level of the DCS can be described by some model. For the aims of my PhD. thesis are important models of the real processes. The model of the real process is a base of the control design. Implementation of the designed controller with the real process in correct control structure provides new quality of the system. However control provides the required quality of the DCS, sometimes may occurs some type of fault on the one or more real processes. Fault detection is research of fault diagnostics methods. All of these methods are usually used and implemented at the lower levels of the DCS [1], [2].

DCS provides a good opportunity for different kinds of modeling approaches, because this system includes continuous or discrete-event processes in continuous or discrete time form. There are many different methods of modelling and one of them is chosen by structure of the system. Different methods of system modelling are presented in [3] – [6].

Models of the real processes are used for next topic my PhD. thesis which is control and fault diagnosis design. Some interesting methods of the control design and diagnostics are presented in [7] – [16]. Usually control and fault diagnostics methods are interconnected, mainly in a real processes. Some of these methods can be implemented and verified on the simulation models and after successful verification used in real processes.

Modelling, control design and diagnostics implementation to the chosen simulations models are one of the main aims of my PhD. thesis. After successful testing and verification of control and diagnosis algorithms in simulation level will be these algorithms implemented to the real laboratory models.

All these facts presented in introduction are described in this paper.

II. DCS Architecture

DCS are created by many systems with the different manners. This fact causes different approaches to their modelling, control and diagnostics. Modeling, control and diagnostics are the aim of my PhD. thesis *Modern Methods of Modelling, Control and Diagnostics of Mechatronics Systems*. DCS provides a very good opportunity for study and application of different modelling, control and diagnostics algorithms. The DCS infrastructure of DCAI FEEI is presented in web page of Center of Modern Control Techniques and Industrial Informatics (*http://kyb.fei.tuke.sk/laben/infdsr.php*).

DCS architecture of DCAI, FEEI includes 5 levels, Fig. 1. The lowest level is the Technological level which includes real models, sensors and actuators. From many different models, for PhD. thesis are mainly important model of Ball and Plate, model of Hydraulics, model of Helicopter and Assembly lines (Flexible manufacturing System, Flexible Assembly Company).

The first level is Technological level of control and regulation. This level represents a set of sources for regulation and control based on programmable logic controllers (PLC) and technological PC's. Model sensors and actuators are connected to PLC's or technological PC's. For control of the Assembly lines and Hydraulic model are used PLC's. Helicopter model and Ball and Plate model are controlled by PC (interface between Helicopter model and PC is Lab Card; interface between Ball and Plate and PC is single-chip microcomputer).

The second level includes SCADA/HMI and simulation models of low level systems. SCADA/HMI supports supervisory control as well as the acquisition, collection and archivation of data from the low level processes. This level also includes simulation models which are very useful for modelling, control and diagnostics design for models in the lowest level of DCS.

The simulation models will be serving as the first step in control and diagnostics algorithms design. The partial aim of

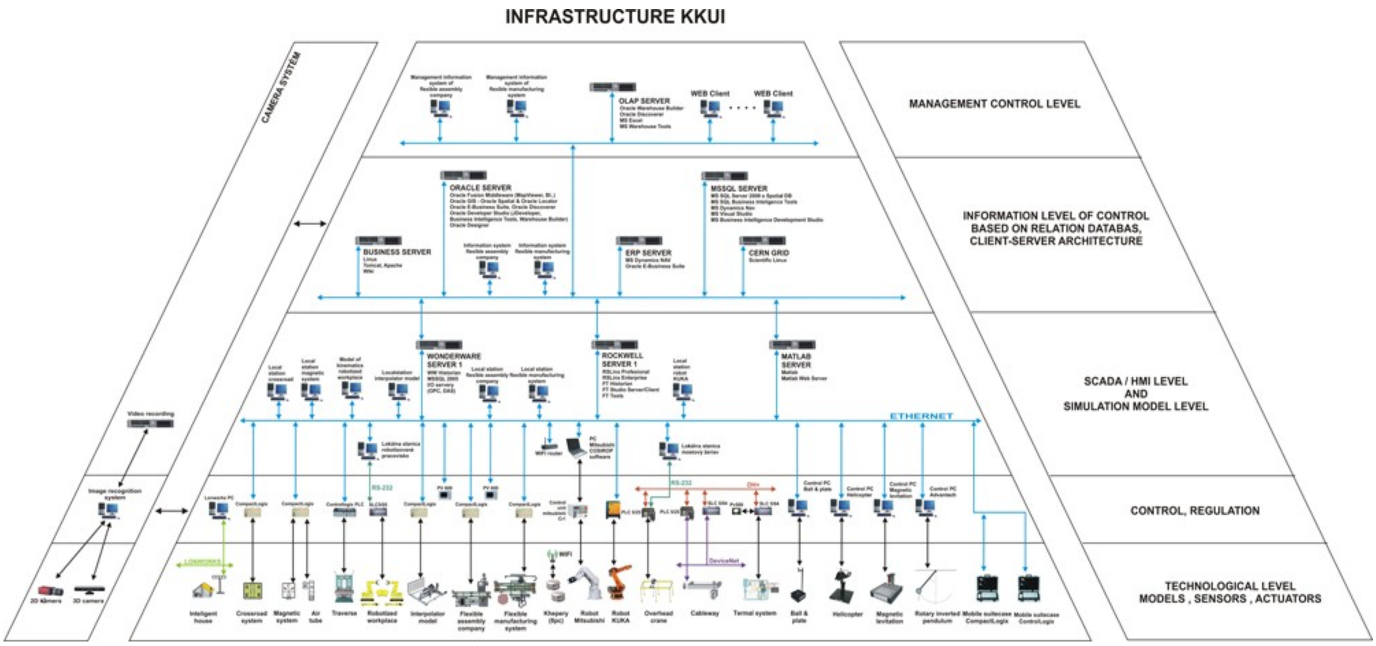


Fig. 1 Distributed Control System of DCAI, FEEI (http://kyb.fei.tuke.sk/laben/IMG/infsys.jpg)

my PhD. thesis is testing and verification of these algorithms on simulation models.

Next levels are Information level of control based on relational database system including MES and ERP/MRP. The last Management control level is based on multidimensional databases and OLAP technology. These levels are included in the paper for completeness of the DCS structure in this paper, but the main interests of this paper are in the Technological, Control and SCADA/HMI with simulation models level.

III. Systems Modelling at the Technological Level of DCS

Very important for aims of the PhD. thesis is choose suitable physical systems. These systems can be classified into [17]:

- continuous states
- discrete-event

systems and models of both systems can be realized in:

- continuous time
- discrete time

form.

The first category includes models of the systems with continuous states. These models can be modelling in continuous or discrete time form. For mathematical model derivation of these systems is possible to use theoretical modelling (analytical identification) or experimental modelling (experimental identification) [17], [18], [19].

Different kinds of the nonlinear systems models are shown in Fig. 2. Analytical and experimental identification includes many different methods of model derivation. From chosen models are in this category model of Ball and Plate, model of Hydraulics and model of Helicopter.

The analytical identification is usually based on physical laws. Application of this method is very exact, if structure and parameters of the real system are known. The results of the analytical identification are presented in [3], [4], [23]. Mathematical models of mechatronics systems can be implemented in the simulation software e.g. Matlab/Simulink.

Experimental identification using methods based on the Least Square method [19] or methods base on Neural Networks [5], [18]. For this methods are very important measured input and output signals from identified system. The experimental identification can be realized by System Identification and Neural Network toolboxes of Matlab/Simulink.

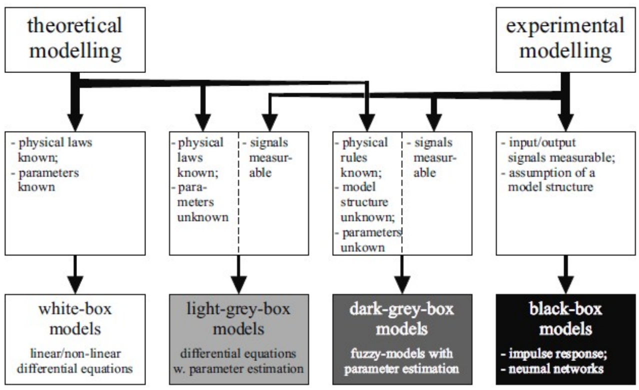


Fig. 2 Different kind of the nonlinear systems models [17]

The second category includes discrete-event systems and representative models are assembly lines [20]. Assembly lines can be modelling in continuous or discrete time form, too. Systems of this type are usually called Finite State Machines (FSM). For modelling of these systems are used state diagrams or tools like CPN Tools, Stateflow, SMI++ [6], etc.

Some processes with continuous manners can be transformed to the FSM model shape. In [7] is presented the Three Tank System in hybrid form. Continuous and discrete dynamics of the tanks are modeled separately. Continuous part is represented by the differential equations based on the law of the mass conservation in each tank. Discrete part is modeled by defining three binary variables and defining the logic statements associated with each binary variable. The three binary variables represent the three valves in the tank model with on/off states. The logic statements are defined depending on the valve positions.

IV. Control Design at the Control and Regulation Level of DCS

Control theory for systems with continues states includes many types of algorithms, but in PhD. thesis should be designed and verified mainly PID/PSD and LQ controller, adaptive controllers, predictive controllers and neural network controllers. Very important in the algorithm implementation is choice of the system sample period for correct computing and realization of the system control input.

The algorithms of the digital PID or LQ control was presented in many articles. These methods are very useful for simple implementation to PLC's or technological PC's control. Time of the control input computing is very quick and this reason is their main advantage.

The adaptive controllers are successful implemented to systems without knowledge about their mathematical model, what is their main benefit [8]. These algorithms can be implemented to PLC's or technological PC's, too.

Modern methods of processes control are algorithms based on prediction of the system state or neural networks. These algorithms are more complex than digital PID or LQ controllers, but there are many advantages of these algorithms. A predictive algorithm provides smoothness control input. Also, predictive algorithms are very useful for target of the reference trajectory tracking [23]. Neural networks provide independence from the linearization in the one equilibrium point, so they are more complex for the targets like a reference trajectory tracking, too.

Predictive control is based on prediction of the future system manners in the control structure. This prediction is derived from the model of the system. Predictive control algorithms can be divided in terms of many different manners, but mainly in terms of the used model of the dynamic system. Predictive control based on the ARX or CARIMA model form is Generalized Predictive Control (GPC) [9]. Also predictive control based on state space form is Model Predictive Control (MPC) [10]. Both algorithms used linear predictor [18], [21].

Another approach to dynamic system control using neural network. In [11] is used Feedforward Neural Network in control structure for computing of the nonlinear system control input. Another approach of using neural networks is in conjunction with predictive control algorithms. Neural network predicts future states of the system based on knowledge of current values of system inputs and outputs [12], [18].

FSM control consists of the specified conditions for transitions between states of the system. Control is usually realized by PLC's with algorithms in ladder logic form.

Design and implementation of these presented control algorithms will be a next part of my PhD. thesis. At first, the main aim will be their testing and verification on the simulation model. After successful results of the simulations, verified algorithms could be implemented and tested to the real laboratory models.

V. Diagnostics at the Technological, Control and SCADA/HMI with Simulation Models Level of DCS

Diagnostics is a recurrent thematic in research. Therefore is not surprising to find contributions from distinct areas: mathematics, system engineering, artificial intelligence, etc. As a general rule, different approaches can be classified as quantitative or qualitative depending on the nature of the measurements and/or observations being made over the system. In the industrial domain diagnosis is typically applied to an individual device or to the entire system [13].

The first three levels of DCS are suitable for implementation of diagnostics methods. Diagnostics algorithms are implemented like a part of the control structure with the real model which is in DCS represented by the Technological and Control level. This problematic contains many different methods and approaches. Very important for diagnostics methods are online input and output signals measurements. Similarly as in the previous section, very important is value of the sample period for online detection of the system faults. Type of the system and value of the sample period have influence to the choice of the diagnostics methods.

The SCADA/HMI level is usually used for visualization of low level processes and for their faults, too. Very important parts of this level are simulation models which describe low level processes by simulation software. Simulation models provide good opportunity for testing and validation of design control and diagnostics algorithms of dynamical systems.

In [17] is presented division of the fault diagnostics methods. From all of these methods is the aim of my PhD. thesis to choose right methods for selected laboratory models. Fig. 3 illustrated tree with fault detection diagnostic (FDD) methods.

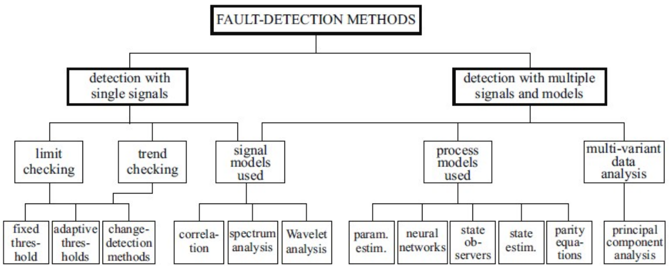


Fig. 3 Different kind of the nonlinear systems models [17]

For the chosen model in section II are useful methods based on the parameter estimation, state estimators or neural networks according to specified control algorithms in section III. These FDD methods are derived from the model of the system. Model of the system can be represented by transfer function, state space form or neural networks.

First of all for implementation of the FDD methods is tested and verified on simulation models. The evaluation of the simulation results decides about suitability of FDD method for implementation to chosen laboratory model or real process at the technological level.

At the technological level could be included systems with periodic manners like a motors or gearboxes. In [17] are systems with periodic manners called Signal models. Types of these systems are one of the possible topics in my PhD. thesis. For diagnostics of these systems are used algorithms based on the Fourier transformation, e.g. Fast Fourier Transformation (FFT) [17].

The FFT represents a periodic signal using a family of complex exponents with infinite time duration. Therefore, FFT is useful in identifying harmonic signals [14].

Another method for signal analysis is based on decomposing of the signal $x(t)$ into a family of functions which are translation and dilatation of a unique-valued function $\psi(t)$, to give the wavelet transform [15].

Signal Processing Toolbox of Matlab environment includes functions for Fourier transformation and Wavelet Toolbox is intended for wavelet transformations.

Last of the possible diagnostics systems are FSM models. For diagnostics of these types is possible to use Fault trees analysis (FTA). In operating safety and reliability science, the FTA is widely used tool to clear out the contributions of different parameters in an undesired event. A fault tree is defined as a graphical representation of the relationship between an undesired event (called a top event) and all its potential causes. The analysis proceeds in a "top–down" approach, starting with the top event (failure, ...) and determining all the causes that can lead to it. It determines how these top events can be caused by individual or combined lower level failures or events [16].

This method should be used for assembly lines diagnostics in my PhD. thesis. Another approach to using the FTA method is in application to hybrid systems. The answer to question if it is possible is one of aims of PhD. thesis.

VI. Conclusion

This paper provides a review of the methods of the modelling, control design and diagnostics applications. These methods may be used for implementation to the chosen simulation models from DCS, shown in Fig. 1. After their testing and verification these algorithms could be implemented to the real processes at the Technological level of the DCS. Main purpose of my PhD. thesis should be in choice of the modelling methods and in correct design of control and diagnostics algorithms. Another purpose should be in their correct combination and successful implementation in the simulation or technological level of DCS.

Acknowledgment

This contribution is the result of the Research and Development Operational Program for project: University Science Park Technicom for innovative applications with knowledge technology support, (ITMS code 26220220182), co-financed by the ERDF – 100%.

References

[1] P. Chochula et. al. "The ALICE Detector Control System," *IEEE Trans. On Nuclear Science*, Vol. 57, No.2 2010, pp. 472 - 478.

[2] K. Jo et. al. "Development of Autonomous Car – Part II: A Case Study on the Implementation of an Autonomous Driving System Based on Distributed Architecture," *IEEE Transaction on Industrial Electronics*, Vol. PP, Issue. 99, 2015, pp. 1 - 12.

[3] W. Bartelmus, "Mathematical Modelling and computer simulations as an aid to gearbox diagnostics," *Mechanical Systems and Signal Processing*, 2001, pp. 855 - 871.

[4] M. F. Ismail et. al. "A Control Performance of Linear Model and the MacPherson Model for Active Suspension System Using Composite Nonlinear Feedback," *IEEE International Conference on Control System, Computing and Engineering*, 2012, pp. 227 – 233.

[5] Y. Wang et. al. "Identification of Ball and Plate System Using Multiple Neural Network Models," *Identification of Ball and Plate System Using Multiple Neural Network Models*, 2012, pp. 229 – 233.

[6] B. Franek, C. Gaspar, "SMI++ object oriented framework for designing and implementing distributed control systems," *IEEE Nuclear Science Symposium Conference Record,* Vol. 3, 2004, pp. 1831 - 1835.

[7] C. Joseph et. al, "Hybrid modeling and discrete controller design of three-tank benchmark system," *2nd International Conference on Instrumentation Control and Automation (ICA)*, 2011, pp. 170 – 174.

[8] B. Andrievsky et. al., "Adaptive Control of 3DOF Motion for LAAS Helicopter Benchmark: Design and Experiments," in *Proc. Of the 2007 American Control Conference*, New York, 2007, pp. 3313 – 3317.

[9] J. Mikleš, M. Fikar, Process Modelling, Identification and Control, Slovak Technical University Bratislava, 2007.

[10] E. F. Camacho, Model Predictive Control, Universidad de Sevilla, 2007.

[11] J. Žilková et. al., "Nonlinear System Control Using Neural Networks," In: *Acta Polytechnica Hungarica (Journal of Applied Sciences)*, Vol. 9, No. 4, 2012, pp. 85 – 94.

[12] M. Lazar, O. Pastravanu, "A neural predictive controller for non-linear systems," In: *Mathematics and Computers in Simulation*, Vol. 60, Issues 3–5, 2002, pp. 315–324.

[13] J. Barata et. al., "Diagnosis using Service Oriented Architectures (SOA)," *5th IEEE International Conference on Industrial Informatics,* Vol. 2, 2007, pp. 1203 – 1208.

[14] G. Zhou et. al., "Gearbox fault diagnosis based on wavelet-ar model," In: *Proceedings of the Sixth International Conference on Machine Learning and Cybernetics*, 2007, pp. 1061 – 1065.

[15] W. J. Wang, P.D. McFadden, "Application of wavelets to gearbox vibrations signals for fault detection," In: *Journal of Sound and Vibration*, Vol.: 192, Issue 5, 1996, pp. 927–939.

[16] N. Y. Steiner et. al., "Application of Fault Tree Analysis to Fuel Cell diagnosis," *Developments of FuelCells – Conference 2011*, 2011, pp. 302 – 309.

[17] R. Isserman, Fault-Diagnosis Systems - An Introduction from Fault Detection to Fault Tolerance, TU Darmstadt, 2006.

[18] Š. Jajčišin, A. Jadlovská, Predictive algorithm design with using nonlinear models of physical systems, Elfa, Košice, 2013.

[19] K. Dolinský, A. Jadlovská, "Application of Results of Experimental Identification in Control of Laboratory Helicopter Model," In: *Advances in Electrical and Electronic Engineering, scientific reviewed Journal published in Czech Republic*, Vol. 9., Issue 4, 2011, pp.157-166.

[20] J. Ilkovič, P. Papcun, "Material flow in Flexible Assembly Company," In: *SCYR 2012: 12th Scientific Conference of Young Researchers of Faculty of Electrical Engineering and Informatics Technical University of Košice: proc.*, 2012, pp. 177-180.

[21] P. Jajčišin, A. Jadlovská, "Predictive Control Algorithms Verification on the Laboratory Helicopter Model," In: *Acta Polytechnica Hungarica (Journal of Applied Sciences)*, Vol. 9, No. 4, 2012, pp. 221 – 245.

[22] J. Jadlovský et. al., "External Access to Alice Controls Conditions Data," Journal of Physics: Conference Series (JPCS), Vol. 513, 2013, pp. 1 - 5.

[23] M. Oravec, A. Jadlovská, "Model Predictive Control of Ball and Plate Laboratory Model" In: *SAMI 2015: IEEE 13th International Symposium on Applied Machine Intelligence and Informatics: proceedings*, 2015, pp. 165 – 170.

Modular cloud-based system as a service for a smart agents: a system proposal

[1]Jaroslav ONDO (1st year)
Supervisor: [2]Peter SINČÁK

[1,2]Dept. of Cybernetics and Artificial Intelligence, FEI TU of Košice, Slovak Republic

[1]jaroslav.ondo@tuke.sk, [2]peter.sincak@tuke.sk

Abstract— **Rapid expansion of cloud technologies along with increasing speed of the Internet and spreading relatively cheap smart devices and sensors have created an environment for developing modern cloud-based applications. These applications are usable not only by humans but also by smart agents like robots, smartphones, smart vehicles, sensors etc. In this paper, we introduce this kind of cloud-based system. The proposed system consists of modules so-called AI bricks. It is able to solve a relatively difficult task in the cloud environment independently on smart agent's hardware and its capabilities. Moreover, cloud technologies used in the proposed system offer unification of different operating systems, programming languages, interfaces, by which smart agents are characterized. The only requirement on smart agents is the sufficient Internet connection. The consequence of that is expansion of agent's capabilities by abilities implemented like a cloud services.**

Keywords—**AI bricks, cloud, cloud robotics, object recognition, ROS.**

I. Introduction

In the year 2009, and only in SourceForge, there existed more than 500 free software projects related with robotics [1]. Examples of these projects implement: drivers for robotics devices and sensors; communication middleware; simulators and modeling tools of dynamic systems [2]. Unfortunately, none of these projects was developed as a cloud application. The reason is obvious - modern cloud technology is younger technology than robotics.

The basic idea of cloud computing has arisen in 1960's, but just development of the Internet and increasing its speed, launching of the Web2.0 in 2009 and at the beginning of the era of providing browser-based enterprise application as a services dramatically supported development of cloud technology in recent years [3].

The rapid development of cloud technology proves, that the only one year after launching the Web 2.0, cloud technology has become so attractive that was established new approach in robotics called cloud robotics [4]. The interconnection of robotics and the cloud technology brings many advantages, which will be discussed in chapter III.

The current problem is how to use the highly developed robotic software in cloud environment or how to benefit from cloud technology for developing new robotic software.

The primary goal of this paper is to provide an overview of the basic concept of modular cloud-based system as a service for smart agents. The aim of this system is to provide a possible solution for mentioned problems.

The chapter II briefly introduce proposed cloud-based system. In the chapter III, we will discuss each used method in the proposed system and their advantages and disadvantages. Chapter IV concludes this contribution by summarizing the state of the work and will provide possibilities of the future work.

II. Proposed cloud-based system overview

As a topic implies, the core technology of the described system is cloud computing. It means that every part of the system is developed as a cloud service in conformity with principles of cloud computing defined in [5], [6]. Since the proposed system is cloud-based and is usable not only by humans, but also for smart agents like robots or smartphones, we can classify proposed system into the field of cloud robotics.

Let us concretize the structure of the described system. It consists of software modules so-called Artificial Intelligence (AI) bricks. Each AI brick can be considered as a software library responsible for solving one particular specific task. For instance AI brick, which implements backpropagation learning for feed-forward neural networks, or AI brick for clustering, image processing etc. Since we want use advantages of cloud computing in robotics, many of AI bricks implements methods suitable for robots.

If AI bricks are software modules, then is possible employ any number of AI bricks for solving a more complex problem. E.g. right functionality of implemented system will be proved using the task of object recognition from static visual images. It requires a collaboration of AI bricks responsible for the image processing, for features extraction and brick for performing classification of objects by using MF ArtMap neural network [7], [8].

III. Used methods

In the proposed system, we will employ various technologies and methods. Many of them belongs to the field of AI. This section will deal with those methods, their advantages, disadvantages and also the needs of their implementation in the described system.

A. Cloud computing and cloud robotics

In recent years, along with the development of cloud technology, also exponentially rising its popularity. According to the European statistics bureau, in 2014 19% of all enterprises in the EU, used cloud computing [9]. Every year grows the number of companies whose use cloud computing for their business. However, the cloud computing has become popular also in research. An evidence of that is the emergence of a new approach, named cloud robotics [4].

Cloud robotics was defined in 2010 by James Kuffner. According to Kuffner's definition [4], cloud robotics is a new developing approach in robotics, which massively use cloud services, cloud data storages, and other internet technologies. These internet technologies allow massive usage of parallel computation and sharing a huge amount of data [10]. The main goal of cloud robotics is a development of infrastructure for robots, which allows communication between robots, recording, uploading, sharing and downloading knowledge and also solving the particular task in the cloud environment.

Using cloud technologies in robotics allows migrate computational difficult tasks from the robot's onboard computing to the cloud environment. So the entire computation is performed on the cloud site. The only one requirement for secure that is the full function Internet connection for robots.

By connecting robots to the cloud environment and allowing use them cloud services, is possible extend their capabilities. From that point of view, a smart agent is not more limited by own computational performance or by memory capacity. However, there are still tasks, which cannot be migrated to the cloud. These critical tasks must be executed in real time. For instance obstacle avoidance, robot's movement, stability etc. [11]

Although cloud robotics is relatively young field, till now was finished various successful projects. We can mention project DAvinCi [12], cloud-based project for robot navigation in the unknown environment [10] and object grasping using Google object recognition engine [13] or project RoboEarth [14].

Regarding to proposed system (See chapter II) we decided to use cloud technology mainly because of possibility of system scalability, 24/7 worldwide availability, data redundancy, agent's platform, operating system and programming language independency and of course simple system maintenance and update.

B. *Concept of AI bricks*

Following AI bricks concept mentioned in chapter II, we can generalize mentioned information. An AI brick is any piece of code which is able to solve one specific task and is programmed as general as possible to cover the widest possible range of applications.

In this paper, we consider AI brick as a software module, which usually implements a particular method of artificial intelligence or is suitable for implementing programs with any level of intelligence.

If the AI bricks should be programmed in conformity with cloud computing principles, must fulfill following requirements:

1. Each AI brick has to be implemented either as a software library attached to cloud service or as a stand-alone worker role. Note: In Microsoft Azure cloud is worker role equivalent to a back office.
2. Each AI brick has to completely solve its task. Task can solve itself or can collaborate with other AI bricks.
3. Each AI brick should have capability to communicate with others AI bricks or programs. It means that AI brick should be able to receive inputs from others AI bricks/programs and send them outputs.
4. Each AI brick must have exactly defined desired inputs and outputs.
5. Each AI brick must be scalable. This requirement means that the program should be implemented the way to run multiple instances on different virtual machines or use multiple threads running on various CPU cores.
6. Data stored in cloud storages must be synchronized between instances of the same AI brick deployed on different virtual servers.

It is necessary emphasize that AI brick concept creates a structure of software libraries or cloud worker roles deployed in the cloud environment. If AI brick is implemented as a software library, should be attached to every new cloud service, which uses a function of AI brick. (See Fig. 1)

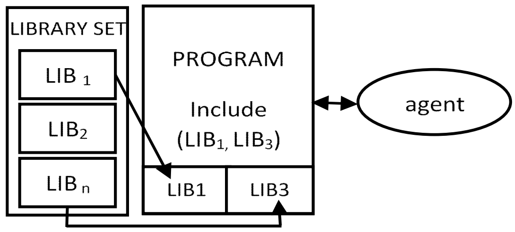


Fig. 1. AI bricks implemented as a set of software libraries with an example of attaching to a new program.

In the case of AI brick is implemented like a cloud worker role, must be referenced in every cloud service, which uses functions of AI brick. (See Fig. 2)

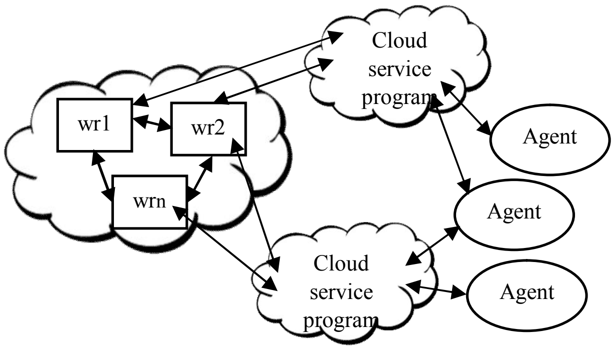


Fig. 2. AI bricks implemented like a set worker roles (wr) in the cloud service. The figure shows an example of communication between agents and cloud services, between cloud services and AI bricks implemented as worker roles and between AI bricks.

New cloud service usually implements more complex program, which uses functions and task's solutions provided by referenced AI bricks. New cloud service tends to support communication between implemented AI bricks and connected smart agents. (See "PROGRAM" block in Fig. 1 or "Cloud service program" blocks in Fig. 2)

Communication can be secured using Simple Object Access Protocol (SOAP) protocol, REST API, Windows Communication Foundation (WCF), Service Bus etc. [15]

Regarding to proposed system, we decided to employ AI bricks concept for implementing various different algorithms and methods. Main reasons are modularity, universality, easy maintenance, in the future possibility make experiments with different methods and simple change of use method.

C. Robotics Software Framework

In chapter I., were provided examples of robotic software stored on SourceForge. These software solutions are usually clustered into the bigger groups, marked as Robotic Software Frameworks (RSF). The RSFs make robotic software development much easier [16]. Usually, every newly developed piece of code can be added to the RSF and considered as a new tool. Till now was developed several RSFs. Most popular are Robotic Operating System (ROS) [16], MT-middleware, YARP, JDE+, OROCOS, Webots [2] and more.

The RSFs not only offer a set of tools for simpler development of robotic software, but they attempt to improve scalability, reusability and deployment of developed software [2]. The cloud technology attempt to improve basically the same characteristics. So the following question is emerged. Why not and mainly how to interconnect not cloud-based RSFs with cloud technology, specifically with cloud robotics?

If the RSF were implemented in the cloud environment, developers would get a powerful cloud robotic-oriented tool. We can consider two main ways of usage cloud-based RSF.

1. RSF as a communication middleware, which secure communication between implemented modules (AI bricks) and connected heterogeneous agents, sensors, actuators or hardware. RSF also provides the entire set of tools, simulators, compilers, debuggers etc.

   This way applications are built as modules attached to RSF. The only difference between RSF deployed locally and RSF deployed on the cloud is that RSF and also every new module is deployed in the cloud environment. The advantages of this approach are simple scalability, implemented by cloud service providers, unlimited computational power and storage capacity, modules and tools sharing with many different robots all over the word and more.
2. Usage of RSF as an AI brick. This way RSF is deployed in the cloud environment, but its communication middleware isn't used. In the cloud is developed own communication mean between cloud services, workers roles (AI bricks), RSF and connected heterogeneous agents.

   The goal of this approach is use preprogrammed tools offered by RSF as an AI bricks for solving the specific task in the cloud environment.

As a result of comparison RSFs in [2] and [17] was selected the Robotic operating system (ROS) for our next research. There are several reasons voting for ROS. We can mention big strong community, support of many different robotic platforms, programming language independency, a lot of preprogrammed drivers and tools like a simulators, possibility of scalability and capability of sharing ROS source code with another RSF. ROS has already been integrated with OpenRAVE, Orocos and Player RSFs [18].

Moreover, ROS offers downloadable preprogrammed packages for solving some of most frequent robotic tasks like a mapping unknown environment, robot teleoperation tools, robot navigation packages, basic grasping functionality, Kinect sensor driver and basic Kinect data processing etc.

D. Object recognition

We use the object recognition from the static images is a toy task for proving the right functionality of proposed system. Nevertheless, the object recognition task is not a trivial, so in this chapter, we will introduce an employed methodology.

In the fig. 3 is shown basic architecture AI bricks for object recognition.

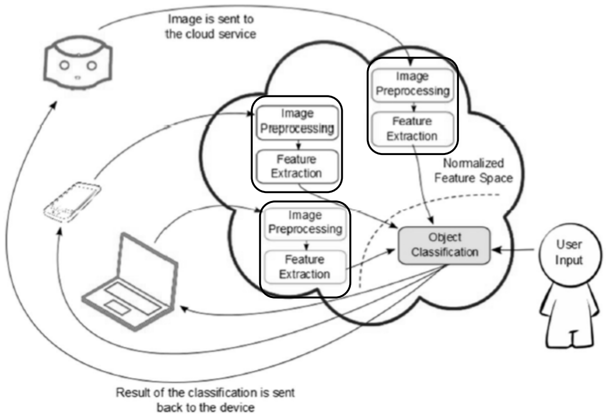


Fig. 3. The high-level architecture of the object recognition AI bricks as a service [19].

The requirements on connected devices (agents) are:

1. Capability capture image
2. The sufficient internet connection for sending captured images to appropriate AI brick.

In the shown architecture, the image processing block and feature extraction block create one worker role – AI brick, responsible for the extraction of features from the input image. In the fig. 3 are shown three times. It indicates the scalability of implemented AI bricks.

Image preprocessing includes image resize operation and conversion to a grayscale image. Block of feature extraction uses SIFT [20] and SURF [21] computer vision algorithm, which has capability find key points in the image and describes each key point by descriptor. Each descriptor is a vector of values – 128-dimensional vector in the case of SIFT algorithm or 64-dimensional vector in the case of SURF algorithm. The texture in the surrounding of each found key point is coded in the descriptor. This way is possible to gain a set of descriptors for input image.

Once the descriptors are extracted, they are propagated into the AI bricks responsible for classification. It means that this AI brick implements classifier. The classifier classifies the object on the picture described by previously extracted descriptors into one of known classes or create a new one if the object does not fit to none of known classes [19].

In our research was selected classifier neural networks for classifying purposes. More specifically ArtMap [22], [23] neural network subgroup. These networks are able to be trained using supervised learning. Finally, MF (membership function) ArtMap [7], [8] neural network was chosen as a classifier. This type of neural network combines the theory of fuzzy sets and ART theory. The consequence of this combination is structured output consisting of computed values of the membership function of every found fuzzy cluster of every known class for the input. This way, it is possible to compute how much the input belongs to every class [19].

IV. State of the Work

An overview of proposed cloud-based system, described in chapter II, is partially implemented. In this chapter, we will discuss the state of the work. We will analyze which parts of work are done and which not.

We started by studying cloud technologies and by programming cloud services deployed in Microsoft Azure cloud. The output is cloud service for image recognition from the static images. This system was implemented using the same principles as was described in point D, in previous chapter. Results were published in [19].

Next challenge was defining AI brick and requirements for implementation. It was discussed in point B, chapter III. In these chapter was mentioned that AI brick can be implemented as a software library, which is attached to every new cloud service project or can be implemented as a stand-alone worker role, which communicates with all others worker roles and connected cloud services. (See Fig. 1 and Fig. 2.) This principle we proved by implementing cloud service containing various worker roles (See Fig. 2). Each worker role is implemented as a library, which contains several functions. Using cloud storage queues, tables and blobs we implemented mean of communication between implemented worker roles, which allow call exact function from any worker role. At this time, we are finishing experiments and performance tests. The results will be published shortly.

In our research, we also focused on robotic software frameworks and robotic middleware. On the basis of comparison RSFs provided in [2] and [17] we selected Robotic operating system (ROS) as a suitable candidate for implementation to the cloud environment. In point C, chapter III we mentioned that exists two ways how to use RFS in the cloud. One of them is using RSF as a communication middleware, the second one is implement RSF as an AI brick. The second way was chosen. ROS was installed as a dedicated virtual Linux server in the Microsoft Azure cloud. However, each AI brick should have capability to communicate with others AI bricks or programs, but due to Microsoft azure cloud safety setting, this communication was prohibited. We solved this problem by installing OpenVPN server, side by side with ROS virtual server in the cloud environment. OpenVPN server allowed us to create VPN connection between devices communicating over the local network and ROS server in the cloud. This way we met all requirements on AI bricks and created fully function RSF AI brick on the cloud.

In our recent work, we finished basic overview of needed technologies and methods. We programmed a basic structure of each part of proposed system, mainly as a separate computer program. Our future work will be focused on integration partially developed parts of system with the aim to create the only one, integrated system. Moreover, each part of future integral system will be developing and enhancing by new functionalities and features.

V. Conclusion

Main goals of this paper are providing a basic overview of Ph.D. thesis topic and give appropriate theoretical background. This aim was met in chapter II, where we introduced the proposed system and consequently in chapter III we focused on the theory of cloud robotics, AI bricks and image recognition. We concluded this paper with discussion of work progress and with the sketch of future work. We believe that our work on proposed system will be beneficial for robotics.

Acknowledgement

Research supported by the National Research and Development Project Grant 1/0667/12 "Incremental Learning Methods for Intelligent Systems" 2012-2015.

References

[1] Xie, Z., & Miyazaki, K. (2009, October). „*Openization, Standardization and Diversification in the Case of Robotics Software Sector in Japan*," 603–606. Retrieved from http://hdl.handle.net/10119/8704

[2] Iñigo-Blasco, P., Diaz-del-Rio, F., Romero-Ternero, M. et al. (2012). Robotics software frameworks for multi-agent robotic systems development. *Robotics and Autonomous Systems*, *60*(6), 803–821. http://doi.org/10.1016/j.robot.2012.02.004

[3] Arif, M. (2009). A history of cloud computing. Retrieved March 15, 2015, from http://www.computerweekly.com/feature/A-history-of-cloud-computing

[4] J. J. Kuffner, "Cloud-enabled robots," in IEEE-RAS International Conference on Humanoid Robotics, 2010.

[5] Mell, P., & Grance, T. (2011). The NIST Definition of Cloud Computing Recommendations of the National Institute of Standards and Technology. *National Institute of Standards and Technology, Information Technology Laboratory*, *145*, 7. http://doi.org/10.1136/emj.2010.096966

[6] M. Armbrust, I. Stoica, M. Zaharia, et al, "A view of cloud computing," *Commun. ACM*, vol. 53, no. 4, p. 50, 2010.

[7] P. Sinčák, M. Hric, and J. Vaščák, "Membership Function-ARTMAP Neural Networks," *TASK Q.*, vol. 7, no. 1, pp. 43–52, 2003.

[8] P. Sincak, M. Hric, and J. Vascak, "Neural Network Classifiers based on Membership Function ARTMAP," in Systematic organisation of information in fuzzy systems, P. Melo-Pinto, H.-N. Teodorescu, and T. Fukuda, Eds. IOS Press, 2003, pp. 321–333.

[9] K. Giannakouris and M. Smihly, "*Cloud computing - statistics on the use by enterprises*," 2014. [Online]. Available: http://goo.gl/PDJ1lw. [Accessed: 15-Mar-2015].

[10] G. Hu, W. P. Tay, and Y. Wen, "Cloud robotics: Architecture, challenges and applications," *IEEE Netw.*, vol. 26, no. 3, pp. 21–28, 2012.

[11] D. Lorenčík, T. Cádrik, M. Mach, and P. Sinčák, "Cloudová robotika; Vplyv cloudového computingu na budúcnosť robotiky," *ATP JOURNAL*, vol. 1, no. 3, pp. 40–42, 2014.

[12] R. Arumugam, V. Enti, L. Bingbing, et al., "DAvinCi: A cloud computing framework for service robots," *Int. Conf. Robot. Autom.*, pp. 3084–3089, 2010.

[13] B. Kehoe, A. Matsukawa, S. Candido, J. Kuffner, and K. Goldberg, "Cloud-based robot grasping with the google object recognition engine," *Proc. - IEEE Int. Conf. Robot. Autom.*, pp. 4263–4270, 2013.

[14] R. Waibel, et al. "RoboEarth-A World Wide Web for Robots," *Robot. Autom. Mag. IEEE*, vol. 18, no. June, pp. 69–82, 2011.

[15] Unknown, "Interoperability Elements of a Cloud Platform." [Online]. Available: http://goo.gl/qOPrn.

[16] M. Quigley, et al., "ROS: an open-source Robot Operating System," *Icra*, vol. 3, no. Figure 1, p. 5, 2009.

[17] G. Magyar, P. Sinčák, and Z. Kriszán, "Comparison Study of Robotic Middleware for Robotic Applications," in *Emergent Trends in Robotics and Intelligent Systems*, 2015, pp. 121–128.

[18] T. Dirk, "ROS introduction," *ROS introduction*, 2014. [Online]. Available: http://wiki.ros.org/ROS/Introduction. [Accessed: 15-Mar-2015].

[19] D. Lorencik, J. Ondo, P. Sincak, and H. Wagatsuma, "Cloud-based Intelligence for Robots Cloud Based Framework for Cloud Robotics.", presented at 4th International Conference on Robot Intelligence Technology and Applications 2014, Beijing, China. 2014.

[20] D. G. Lowe, "Object recognition from local scale-invariant features," in Proceedings of the Seventh IEEE International Conference on Computer Vision, 1999, pp. 1150–1157 vol.2.

[21] H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded Up Robust Features," in European Conference on Computer Vision, 2006, pp. 404–417.

[22] G. A. Carpenter, "Default ARTMAP," Boston, 2003.

[23] N. Kopčo, P. Sinčák, and S. Kaleta, "ARTMAP Neural Networks for Multispectral Image Classification," J. Adv. Comput. Intell., vol. 4, no. 4, pp. 240–245, 2000.

Multi-objective optimization of modern assembly lines

[1]*Ján Čabala (1st year)*
Supervisor: [2]Ján Jadlovský

[1,2]Dept. of Cybernetics and Artificial Intelligence, FEI TU of Košice, Slovak Republic

[1]jan.cabala@tuke.sk, [2]jan.jadlovsky@tuke.sk

***Abstract*— This paper deals with various ways of using multi-objective (vector) optimization in assembly line optimization problem. Nowadays, there are two main approaches to solve multi-objective optimization tasks: artificial intelligence approach using mainly evolutionary algorithms; and mathematical approach based on detail mathematic description of optimized assembly line. This paper is summarizing knowledge of application possibilities of both approaches in assembly line optimizing. Also actual status of my dissertation thesis is mentioned, as well as list of solved and unsolved problems and future direction of my dissertation thesis.**

***Keywords*—assembly line, multi-objective optimization, objective function, constraints**

I. Introduction

Assembly line is a flow oriented production system consisting of productive units (workstations), which are aligned in a serial manner. The workpieces are transferred from one workstation to another via some kind of mechanical transport system (usually conveyor belt is used for this purpose).

Assembly lines were originally developed for cost efficient mass production of standardized products. First modern assembly line was designed by Henry Ford at the beginning of 20th century. This line was able to build Ford Model T in 93 minutes. However, since that time customer needs, as well as product requirements has changed dramatically. For example, German car producer, BMW, provides users with many additional features, which theoretically results in 10^{32} different models.

According to company requirements, designer is obligated to create assembly line, which is fulfilling customer's needs most. Assembly lines can be divided into many different groups considering various criteria, for example shape of assembly line. Some basic configurations of assembly lines can be found in Fig. 1.[1],[2]

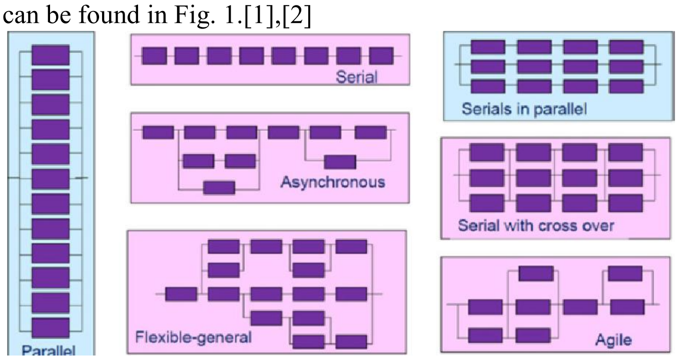


Fig. 1. Basic assembly line configurations

Assembly lines can produce only one simple model, or they can be adapted to produce number of different models. According to this criterion, assembly lines can be divided into simple, mixed and multi model assembly lines. Difference between them is displayed on Fig. 2.[3]

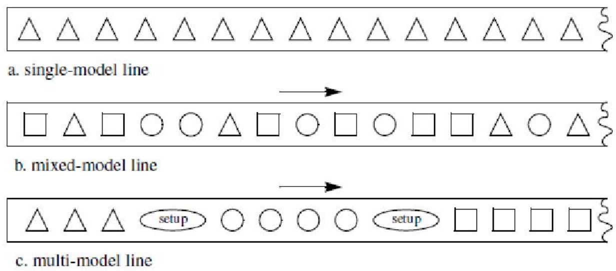


Fig. 2. Difference between single-model, mixed-model and multi-model assembly line

Some further information about mixed model assembly lines can be seen at [4] .Assembly lines and their workstations can be also divided into automatic, cooperative and manual. Other information about this topic and also about cooperation between humans and machines in assembly lines can be found in [5].

For any shape and any type of assembly line, there is always a need to use this line as efficiently as possible. Mentioned criteria, as well as many others, are influencing the way of access for optimization and balancing particular assembly line.

Assembly line balancing and optimization problem is represented by various optimization models aimed at supporting the decision maker in configuring efficient assembly system. Some problems and basic methods of assembly line balancing can be found in [6] and [3]. Ways of dealing with this problem are discussed in following chapters of this paper.

II. Multi-objective optimization

Optimization can be considered as choosing the best option from wide spectrum of alternatives. In our everyday life we are trying to make our work done spending as little time or as little energy as possible. Optimization methods are used to solve this problem most frequently, mostly because of their mathematical basis, which guarantees objectivity and accuracy of optimization process. Publications dealing with optimization problems and its solving are [7] and [8].

General optimization problem can be defined as minimization (maximization) of objective function

$$f = \{x_1, x_2, \dots, x_n\} \quad (1)$$

with respecting all constrains

$$g_i = \{x_1, x_2, \dots, x_n\}, for\ i = 1,2,\dots,n$$
$$x_j > 0, \qquad for\ j = 1,2,\dots,n. \quad (2)$$

For solving optimization problems, methods of mathematical programming are used. Depending on the type of objective function, these methods can be divided into:

- linear programming methods,
- non linear programming methods,
- integer programming methods,
- parameter programming methods,
- stochastic programming methods.

Multi-objective (vector) optimization is dealing with ways of optimizing the problem with multiple goals. It is used when it is necessary to accept more than one factor in finding the ideal solution of optimization task. This type of optimization was created for solving planning and organizing problems in manufacturing process. Nowadays it is used in many different areas (f.e.in dynamic management systems).

Vector optimization task is defined with relation to controlled system, which is described with *n*- dimensional vector $\boldsymbol{x} = (x_1, x_2, \dots, x_n x_1, x_2, \dots, x_n)$ with $x \in \{X\}$ representing *n* independent variables (decision variables). This system $\boldsymbol{x} \in \{X\}$is evaluated with k- dimensional vector functional

$$\boldsymbol{J}(\boldsymbol{x}) = (J_1(\boldsymbol{x}), J_2(\boldsymbol{x}), \dots, J_k(\boldsymbol{x})), \quad (3)$$

which elements are functions of vector $\boldsymbol{x}$, where *k* represents number of objective functions. This functional is optimized in subject to

$$g_i(x) \le 0, i = 1,2,\dots,m, \quad (4)$$

and

$$h_l(x) = 0, l = 1,2,\dots,e, \quad (5)$$

where *m* is number of inequality constraints and *e* is number of equality constraints. Solution of this task is $\boldsymbol{x}^* \in \{X\}$, which is optimal solution of functional (or also called objectives, criteria, payoff functions, cost functions or value functions) $J_1(\boldsymbol{x}), J_2(\boldsymbol{x}), \dots, J_k(\boldsymbol{x})$ according to chosen type of vector optimization. [9],[10].

Basically, there are two main approaches to solve vector optimization tasks:

- artificial intelligence approach,
- mathematical approach.

A. Artificial intelligence approach

Main feature of artificial intelligence used in multi-objective optimization are evolutionary algorithms. They are used mainly for generating so called Pareto optimal set of solutions (solutions that are not dominated by any other solution of particular problem). Detail definition of Pareto optimality can be found in [9].

Genetic algorithms are part of evolutionary algorithms, which are characterized by a population of solution candidates. Reproduction process enables to combine existing solutions and generating new possible solutions. Finally there is a natural solution which determines which individuals of current population will participate in the next one. Functional description of evolutionary algorithms can be seen in Fig. 3.[11]

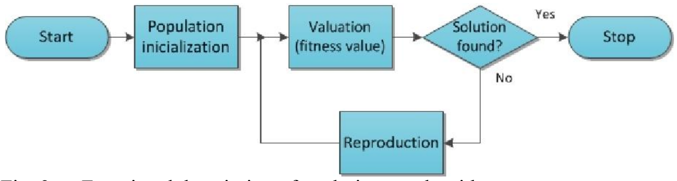


Fig. 3. Functional description of evolutionary algorithms

Some of evolutionary algorithms used for generating Pareto optimal set of solutions, but due to their working procedures, they often tend to stuck in good approximation and they do not guarantee identifying optimal trade-offs.

Artificial intelligence approach has wide application potential. For example, commonly used algorithm is ant colony algorithm. Its application in modeling and balancing time and space constrained assembly line can be found in [12]. Another application of ant colony algorithm is described in [13], where this algorithm is used for optimizing single model U-shaped assembly line.

There are many application possibilities for solving assembly line multi-objective optimization problem using genetic algorithms. Examples of their using in this area with detail description of their functionality can be found in [14].[15] is using genetic algorithm in dealing with assembly sequence planning problem. Solution for defining and two methods of pruning optimal Pareto set are mentioned in [16].

Variation of production rates and number of assembly line setups are optimized simultaneously with using multi-objective genetic algorithm approach in[17].

Using metaheuristic method of tabu search for solving simple assembly line balancing problem is described in [18].

B. Mathematical approach

From the 70's, when first vector optimization tasks were solved, numbers of methods dealing with this issue were created. All these methods assume that there is a definition of more than one objective function, as well as list of equality and inequality constraints defining feasible solution space. Mathematical methods dealing with this manner can be divided according to [7] into three main groups:

- methods defining the set of non improving elements ,
- compromising methods
- methods of hierarchical criteria sequence.

Methods defining the set of non improving elements

There is no hierarchy of criteria in this group of methods, every criterion is equally important. Task of vector optimization can be defined as trying to minimize the vector $\boldsymbol{J}(\boldsymbol{x}) - \boldsymbol{J}(\boldsymbol{x_\alpha})$, where $\boldsymbol{x_\alpha}$ is optimal values vector of variables $\boldsymbol{x}$ according to defined criteria.

a) Quadratic norm

The most common criterion used in this norm is minimal sum of quadratic variance of objective functions $J_\alpha(\boldsymbol{x})$ for random $\boldsymbol{x} \in \{X\}$ from objective function $J_\alpha(\boldsymbol{x_\alpha})$ for vector of ideal values of chosen criteria $\boldsymbol{x_\alpha} = (x_{1\alpha}, x_{2\alpha}, \dots, x_{n\alpha})$, $\alpha = 1,2,\dots,k$;

$$R(x) = \sum_{\alpha=1}^{k} (\boldsymbol{J_\alpha}(\boldsymbol{x}) - \boldsymbol{J_\alpha}(\boldsymbol{x_\alpha}))^2$$

$$opt\ R(x) = \min_{x \in \{X\}} \sum_{\alpha=1}^{k} (\boldsymbol{J_\alpha}(\boldsymbol{x}) - \boldsymbol{J_\alpha}(\boldsymbol{x_\alpha}))^2 \quad (6)$$

Values of this norm are usually divided by optimal value. Reason for this is non-dimensional solution value. Calculation formula then looks like:

$$R(x) = \sum_{\alpha=1}^{k} \frac{(\boldsymbol{J_\alpha(x)} - \boldsymbol{J_\alpha(x_\alpha)})^2}{\boldsymbol{J_\alpha(x_\alpha)}^2}$$

$$opt\ R(x) = \min_{x \in \{X\}} \sum_{\alpha=1}^{k} \frac{(\boldsymbol{J_\alpha(x)} - \boldsymbol{J_\alpha(x_\alpha)})^2}{\boldsymbol{J_\alpha(x_\alpha)}^2}. \tag{7}$$

b) Linear norm

It represents minimal sum of linear variance of objective functions $(\boldsymbol{J_\alpha(x)})$ from optimal values of objective functions $\boldsymbol{J_\alpha(x_\alpha)}$. Optimal value of functions $R(x)$ can be calculated using formulas:

$$R(x) = \left|\sum_{\alpha=1}^{k} (\boldsymbol{J_\alpha(x)} - \boldsymbol{J_\alpha(x_\alpha)}\right|,$$

$$opt\ R(x) = \min_{x \in \{X\}} \left(\sum_{\alpha=1}^{k} (\boldsymbol{J_\alpha(x)} - \boldsymbol{J_\alpha(x_\alpha)}\right). \tag{8}$$

c) Generalized norm

Objective functions of this norm is

$$R_L(x) = \sum_{\alpha=1}^{k} \left\{\left((\boldsymbol{J_\alpha(x)} - \boldsymbol{J_\alpha(x_\alpha)}\right)^L\right\}^{\frac{1}{L}}; L \geq 1. \tag{9}$$

For $L = 1$ this functions corresponds to linear norm, for $L = 2$ it corresponds quadratic norm and for $L = \infty$ it is

$$R_\infty(x) = \max_\beta\{(\boldsymbol{J_\alpha(x)} - \boldsymbol{J_\alpha(x_\alpha)}; = 1,2, \dots, n\}. \tag{10}$$

All these norms can be refined by multiplying objective function elements with appropriate coefficients. Searched point $\boldsymbol{x}^* \in \{X\}$ is called non improving in space $\{X\}$ regarding to functional $\boldsymbol{J(x)}$, if there is no point $\widetilde{\boldsymbol{x}}$ in this space, for which is valid $(\boldsymbol{J_\alpha}(\widetilde{\boldsymbol{x}}) \leq \boldsymbol{J_\alpha}(\boldsymbol{x}^*), \alpha = 1,2, \dots, k.$

<u>Compromising methods</u>

These methods are based on defining more strict constrains or adding another constrains to objective functions. Compromise means finding optimal solution by minimizing value of formula $\beta_1 J_1(\boldsymbol{x}) + \beta_2 J_2(\boldsymbol{x}) + \cdots + \beta_k J_k(\boldsymbol{x})$, where β_1 to β_k are carefully chosen importance coefficients. Their values are recommended to be chosen as follows: $\beta_1 = 1/J_{10}, \beta_2 = 1/J_{10}, \beta_k = 1/J_{k0}$, where J_{k0} are values of objective functions counted in optimization only via selected criterion.

These methods are used when it is possible to define importance of every criterion before starting optimization process, or when additional information about criterion importance is found out during optimization process. Main idea is in defining importance of every scalar criterion. This can affect the result of vector optimization (importance of scalar criterion will be labeled as λ). Most common methods from this group are optimization of weighted sum of scalar criteria and weighted sum of variance vector.

d) Optimization of weighted sum of scalar criteria

Function for this norm is defined by formula

$$opt\ R(x) = opt \sum_{\alpha=1}^{k} \boldsymbol{\lambda_\alpha J_\alpha(x)}. \tag{11}$$

Optimal solution is represented by maximum or minimum of weighted sum of scalar criteria, depending on character of problem being solved.

e) Weighted sum of variance vector

In this norm function $R(x)$ is represented by formula

$$R(x) = \sum_{\alpha=1}^{k} \boldsymbol{\lambda_\alpha} * (\boldsymbol{J_\alpha(x)} - \boldsymbol{J_\alpha(x_\alpha)}. \tag{12}$$

Optimal solution is always represented by minimum value of function $R(x)$, because this norm is using variation from ideal values.

$$opt\ R(x) = \min_{x \in \{X\}} \left(\sum_{\alpha=1}^{k} \boldsymbol{\lambda_\alpha} * (\boldsymbol{J_\alpha} - \boldsymbol{J_\alpha(x_\alpha)}\right). \tag{13}$$

It is also possible to use quadratic variant of this formula (weighted quadratic norm).[7]

<u>Methods of hierarchical criteria sequence</u>

In this qroup of methods, hierarchical sequence of criteria is created, so there is one superior criterion chosen from list of criteria, and all others are inferior to it.

Methods mentioned in this paper are not the only one used in solving multi-objective optimization problems, other methods can be found in [9] and [19].

III. Multi-objective assembly line optimization

First step in solving multi-objective optimization of assembly lines is creating a detail model of system. Assembly lines can be modeled by various means. During my work will be used Petri nets and Stateflow diagrams. Assembly lines modeling using Petri nets, as well as other possibilities for modeling assembly lines can be found in [20].

Created model must truly display real assembly line, for example follow precedence constraints of assembly line. This model can be also helpful in finding critical spots in production process, as well as defining optimization criteria:

- maximizing profit level, resp. reducing costs,
- maximizing reliability and safety,
- maximizing efficiency resp. minimizing overload,
- minimizing manual interventions,
- minimizing of production time.

Assembly line model can be also source for defining constraints for this multi-objective optimization tasks, f.e.:

- maximum level of costs,
- maximum time in one work shift,
- minimum level of needed sources etc.

All criteria and constraints should be defined by mathematical equations, representing objective functions. Once there are defined, a multi-objective optimization task is built and there is a place to find ways of solving it.

Focus of my dissertation thesis is to solve multi-objective optimization problem using mathematical approach. In my previous work, application able to solve vector optimization tasks was created in MATLAB. Present version of this application can be used for solving vector optimization problems with 2 objective functions. One of them can have either quadratic or linear form, the other one must be linear. Also all constraints have to be defined by linear equations or inequations. Solving sample tasks from production and economical sector using created application is listed in [21]. Algorithm solving mentioned vector optimization tasks is displayed in Fig. 4.

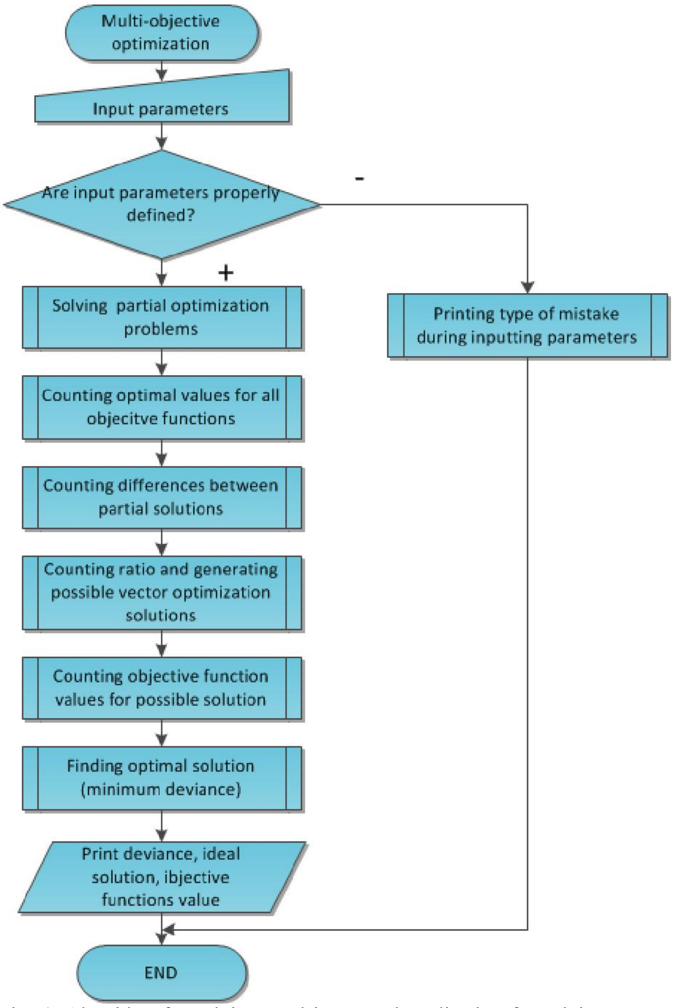


Fig. 4. Algorithm for solving used in created application for solving vector optimization problem

Application has also its own graphic user interface (GUI) designed for simple and intuitive inputting parameters, as well as for clear and easy-to-read displaying outputs. GUI of created application is shown in Fig. 5.

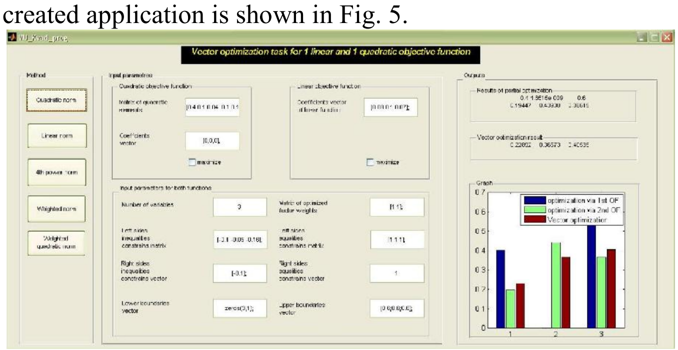


Fig. 5 Application created for solving vector optimization tasks

According to the fact, that this application can solve only limited number of tasks, my goal in my future work is to make it more complex and usable in solving optimization tasks described by other, more complicated objective functions and constraints.

Another task, which has to be solved in my future work on dissertation thesis is creation of detail model of optimized production line. For this purpose Stateflow diagrams will be used. Main advantage of this form of assembly line modeling is simple, but complex syntax able to model large variety of processes executed in production lines.

At the basis of this model, mathematical description of defined optimizing parameters will be created.

ACKNOWLEDGEMENT

This contribution is the result of the Research and Development Operational Program for project: University Science Park Technicom for innovative applications with knowledge technology support, (ITMS code 26220220182), co-financed by the ERDF.

REFERENCES

[1] BOYSEN, Nils; FLIEDNER, Malte; SCHOLL, Armin. Assembly line balancing: Which model to use when?. International Journal of Production Economics, 2008, 111.2: 509-528.

[2] HU, S. Jack, et al. Assembly system design and operations for product variety.CIRP Annals-Manufacturing Technology, 2011, 60.2: 715-733.

[3] BECKER, Christian; SCHOLL, Armin. A survey on problems and methods in generalized assembly line balancing. European journal of operational research, 2006, 168.3: 694-715.

[4] BOYSEN, Nils; FLIEDNER, Malte; SCHOLL, Armin. Sequencing mixed-model assembly lines: Survey, classification and model critique. European Journal of Operational Research, 2009, 192.2: 349-373.

[5] KRÜGER, J.; LIEN, T. K.; VERL, A. Cooperation of human and machines in assembly lines. CIRP Annals-Manufacturing Technology, 2009, 58.2: 628-646.

[6] BOYSEN, Nils; FLIEDNER, Malte; SCHOLL, Armin. A classification of assembly line balancing problems. European Journal of Operational Research, 2007, 183.2: 674-693.

[7] HRUBINA, Kamil - JADLOVSKÁ Anna – HREHOVÁ Stella: Algorithms of optimization methods with using programming systems. Vol.1 . Prešov – Košice, 2005.393 pages. ISBN 80-88941-31-8.

[8] ROSINOVÁ, Danica – DÚBRAVSKÁ Mária: Optimization. Vol.1, Bratislava: STU, 2008. 190 pages. - ISBN: 8022727952.

[9] MARLER, R. Timothy; ARORA, Jasbir S. Survey of multi-objective optimization methods for engineering. Structural and multidisciplinary optimization, 2004, 26.6: 369-395.

[10] WEISE, Thomas. Global optimization algorithms-theory and application. Self-Published,, 2009.

[11] ABRAHAM, Ajith; JAIN, Lakhmi. Evolutionary multiobjective optimization. Springer London, 2005.

[12] BAUTISTA, Joaquin; PEREIRA, Jordi. Ant algorithms for a time and space constrained assembly line balancing problem. European Journal of Operational Research, 2007, 177.3: 2016-2032.

[13] SABUNCUOGLU, Ihsan; EREL, Erdal; ALP, Arda. Ant colony optimization for the single model U-type assembly line balancing problem. International Journal of Production Economics, 2009, 120.2: 287-300

[14] LI, Hui; ZHANG, Qingfu. Multiobjective optimization problems with complicated Pareto sets, MOEA/D and NSGA-II. Evolutionary Computation, IEEE Transactions on, 2009, 13.2: 284-302.

[15] CHOI, Young-Keun; LEE, Dong Myung; CHO, Yeong Bin. An approach to multi-criteria assembly sequence planning using genetic algorithms. The International Journal of Advanced Manufacturing Technology, 2009, 42.1-2: 180-188.

[16] TABOADA, Heidi A., et al. Practical solutions for multi-objective optimization: An application to system reliability design problems. Reliability Engineering & System Safety, 2007, 92.3: 314-322.

[17] MANSOURI, S. Afshin. A multi-objective genetic algorithm for mixed-model sequencing on JIT assembly lines. European Journal of Operational Research, 2005, 167.3: 696-716.

[18] LAPIERRE, Sophie D.; RUIZ, Angel; SORIANO, Patrick. Balancing assembly lines with tabu search. European Journal of Operational Research, 2006, 168.3: 826-837.

[19] HUANG, Hong-Zhong; TIAN, Zhigang; ZUO, Ming J. Intelligent interactive multiobjective optimization method and its application to reliability optimization.IIE Transactions, 2005, 37.11: 983-993.

[20] ČOPÍK,Matej. (Supervisor: doc. Ing. Ján Jadlovský, CSc.):Methodics of diesgh and realization of production lines with focus on production time optimization: Dissertation thesis. Košice: KKUI FEI TUKE. 2014. 89 pages.

[21] ČABALA, Ján.;JADLOVSKÝ, Ján. Solving vector optimization problems using Optimization Toolbox, In: 22nd Annual Conference Proceedings of the International Scientific Conference - Technical Computing Bratislava 2014, 4th November 2014

Neural Networks for Speech Recognition

[1]*Dávid ČONKA (1st year),*
Supervisor: [2]Anton ČIŽMÁR

[1,2]Dept. of Electronics and Multimedia Communications, FEI TU of Košice, Slovak Republic

[1]david.conka@tuke.sk, [2]anton.cizmar@tuke.sk

***Abstract*—This work describes deep neural networks (DNNs). DNNs play an important role in the modern speech recognition systems. We introduce the theoretical characteristics of the NN. Primarily this paper is related to identification and division of different structures in DNNs. We write about feedforward and recurent neural network. The paper also describes basic function of Automatic Speech Recognition (ASR). The goal of the work is in using NN model for ASR.**

***Keywords*—deep neural network, feedforward network, recurrent network, perceptron, backpropagation**

I. Introduction

Speech is very important way for changing information. We can convey various kinds of useful information during communication. Speech signal has various features. It sounds different when we are happy or sad, friendly or angry. In the recent decades an improvement of speech processing has occured. This progress makes human life easier and more comfortable.

In section II. we describe basic functions of the Automatic Speech Recognition (ASR) system. The procedure shows some reasons which have to be done for good results in the ASR.

Section III. focuses on Neural Networks (NNs) and their basic characteristics. Networks which were developed in the biology are using in the modern speech recognition systems or for pattern recogniton. In this part we describe Deep NNs and we shortly explain what the deep means. We also describe Artificial Neural Networks (ANNs) and basic mechanism of ANNs. In the third subsection we speak about feed-forward neural networks and recurrent neural networks. We also show the figures where we can see the difference between these structures. Next subsection refers to Back-Propagation (BP) algorithm and shows practical example.

Fourth section refers about Kaldi Toolkit which will be used in the future for speech recognition in our laboratory. This toolkit is free available on the Internet and it is written in C++.

Next, fifth section is about related works. We write about some references of papers in this part. Those works include useful information using in the neural networks.

Conclusion part is situating at the bottom of this paper. There are summary of this paper and future plans too.

II. Automatic Speech Recognition

Speech signal used in the neural network as a training data is limited. New speakers, vocabularies, languages, dialects and some new conditions must be encountered. We must have a modern computation system, because training data has a high capacity and computation can take many hours or days. Neural network architecture is used in modern systems which can work faster. We can compare its performance with hidden Markov model systems based on ASR, but it depends on the application.

Automatic Speech Recognition (ASR) is very difficult process for several reasons:

1) limitation of speech signal imposed by the Nyquist interference criteria
2) redundant information in the acoustic signal,
3) signal degradation due to addititve and convolutional noise, for example enviroment noise, telephone bandwidth, device characteristics,
4) variabilities in pronunciation of words or phonemes, for example the region dialect,
5) coarticulation effects, context-dependent pronunciation of words or phonemes, and
6) using words which is out of the recognizer vocabulary.

ASR is a complex problem and it must be solved as well as possible. In the recent years, the best way to recognize the spoken speech was with using neural network (NN) algorithm. The NN we describe below [1], [2].

III. Neural Network

The term neural network was inspired by the biological neural network system, such as the brain, process information. The structure consists of a large number of highly connected processing elements called neurons. We can divide between input and output neurons. Input neurons get activated through sensors recognizing the enviroment, output neurons get activated weighted connections from previously active neurons.

One of the most common applications of NNs is in image processing. Some examples would be: identifying hand-written characters, matching a photograph of a person's face with a different photo in a database, performing data compression on an image with minimal loss of content. Other applications could be: voice recognition, RADAR signature analysis, stock market prediction. All of these problems involve large amounts of data, and complex relationships between the different parameters [3].

A. Deep Neural Networks

As the name refers, deep neural networks differ from conventional method in the structure. The basic method has only one hidden layer. When we speak about deep network we have to know what means the word deep. We have some questions: How many layers we need, if we want to have a

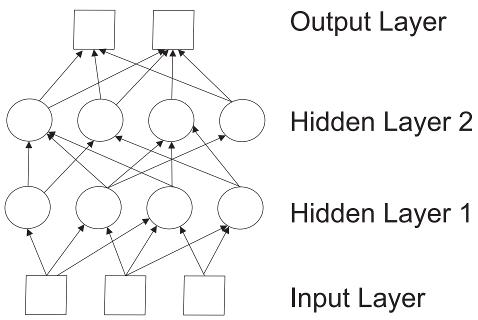


Fig. 1. An example of the deep neural network with two hidden layers.

deep network? Is two deep? According approximation theory ($2 > 1$), two is trully deep. Deep neural networks should have at least two hidden layers. We can have many hidden layers, but the large number of these layers can return worse results. Each layer has a negative feature for data signal as attenuation, noise, etc.. We have to train some system with different numbers of hidden layers and then we will find the best one [4], [5].

The term deep neural network (DNN) was originally introduced to mean multilayer perceptron (MLP) with many (often more than two) hidden layers. Later, it was extended to any NN with a deep structure. It plays an important role in the modern Speech Recognition Systems [6]. Fig. 1 depicts a DNN with four layers that include input layer, two hidden layers and output layer [7], [8], [9], [10].

B. Artificial Neural Network

Many laboratories focus their work to modern systems. Very popular approach for speech recognition is neural network. The neural network is a method, which has roots in the biology.

In the 1980s, a number of researchers began to apply ANN approaches to speech classification, and in particular to automatic speech recognition (ASR). ANNs, like people, learn by example. It is configured for a specific application, such as pattern recognition or data classification through a learning process. It consists of many biological features such as distributed computation mechanism, adaptivity (trainability), nonlinearity and simplicity as mentioned in [11].

The ANNs are based on the biological neural system and they have many nodes and connections. The nodes correspond to a neuron cell and the connections correspond to an axon in the real biological neural system. Nodes have one part for computation of basis function and the second part for computation of application function and they are mutually connected, see Fig. 2.

The basis function computes an input signal, which can be an input signal to the network or an output signal to another node, and computes the input signal to the activation function unit. The aplication function unit produces an output signal, which can be the final output signal of the network or an input signal for another node. An arrow indicates the direction of the information flow. We know two types of arrows: white and black arrow. The white arrow shows information flow in the backward direction and the black arrow marks the direction in the forward way.

Detailed computation of the neural networks is on the Fig. 3. There are N input connections x_n and one output connection

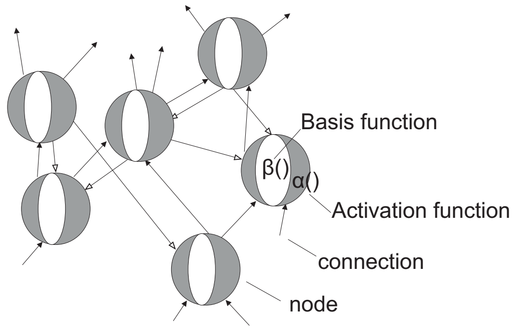


Fig. 2. Basic mechanism of the artificial neural network.

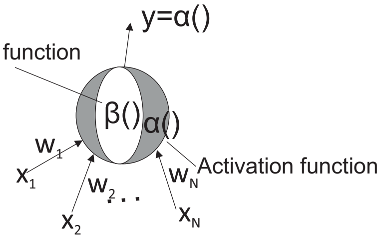


Fig. 3. Computation mechanism of the neural network.

y. Each arrow indicates $n-th$ connection with their own $n-th$ weight w_n. The input signal vector x_n consists from $x = (x_1, x_2, .., x_N)^T$, weight vector is $w = (w_1, w_2, .., w_N)^T$, where T is the vector transpose. Then, the basis function $\beta()$ computes an input signal to the activation function unit,

$$\beta(x) = \sum_{n=1}^{N} w_n x_n. \tag{1}$$

The activation function computes the final node output y as

$$y = \alpha(\beta(x)). \tag{2}$$

Actually, there are many possibilities for the computational selection of the basis function and of the activation function, and these selections are one of the main factors determining the mathematical characteristics of the ANN.

C. Structure

We can divide the structure in two separate types. One of them is feedforward neural network (FNN). Connectivity between the nodes (neurons) is made such as each neuron from one layer is connected to every neuron in the next layer. We define that signal flow only in one direction across the network. On the Fig. 4 we can see only black arrows. It means that information flow goes from the left side to the right side (from input to output). It uses one-way mode. Each neuron from the layer 1 receives the signals from the inputs of the network. Then, received signal from the layer 1 is sent to the next hop (layer 2). Signals from the layer 2 are transmitted to all of the neurons in the next layer (layer 3), etc.. Transceived signal from the layer 4 is an output signal. So, to use the network to solve a problem, we apply the input values to the inputs of the first layer, allow the signals to propagate through the network, and read the output values.

On the other side we have recurrent neural network. The output signal of each node is allowed to go back on the previous neural node. We can say that it uses two-way mode. As we can see on the Fig. 5, it consists of black and white

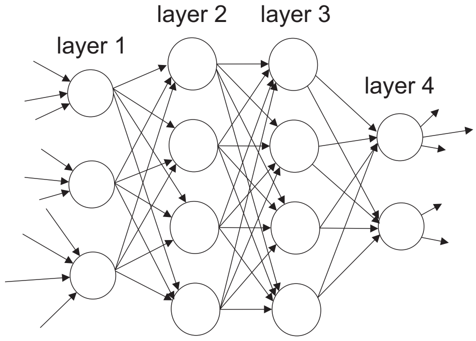


Fig. 4. Feedforward network.

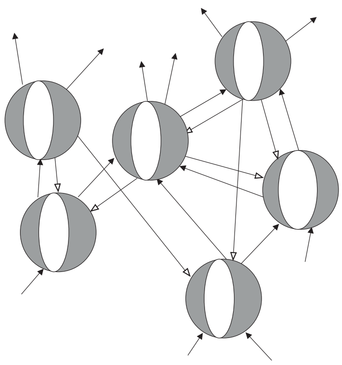


Fig. 5. Recurrent network.

arrows and it has bigger capability than feedforward network. The recurrent neural network is better candidate for data such as speech signal because of larger capacity [3], [12], [13], [14].

D. Backpropagation

This part rewievs the architecture of the backpropagation (BP) neural network (NN). BP is currently the most widely applied NN architecture. It was originally introduced by Bryson and Ho in 1969 and independently rediscovered by Werbos in 1974. This method uses Multiple-Layer Perception (MLP) as system framework and Back-Propagation Algorithm as training rule. Multiple-Layer in MLP model indicates that it is composed of many layers of neurons [15], [16].

Have an example. If the house shakes, we do not know either an unstoppable truck hits the house or there is an Earthquake. So it is hard to 'backpropagate' information signal. But other direction is easier: if there is an Earthquake, it is small chance that truck hits house, Fig. 6 [17].

IV. Kaldi Toolkit

This section shortly describes the design of Kaldi, open-source toolkit for speech recognition process. Kaldi is a toolkit for speech recognition written in C++ and licensed under the Apache License v2.0.

Kaldi is free available on the Internet (see http://kaldi.sourceforge.net/index.html). From the website we can download lot of files and examples. The toolkit can be compiled under the Unix system or Windows. Main goal of Kaldi is to have a modern and flexible code which is easy to understand, extend and modify. We know several ways of toolkits for design a recognition system. For example: HTK [18], Julius [19], Sphinx [20].

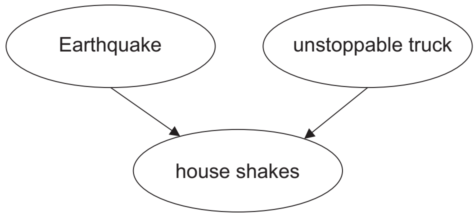


Fig. 6. Example of backpropagation model.

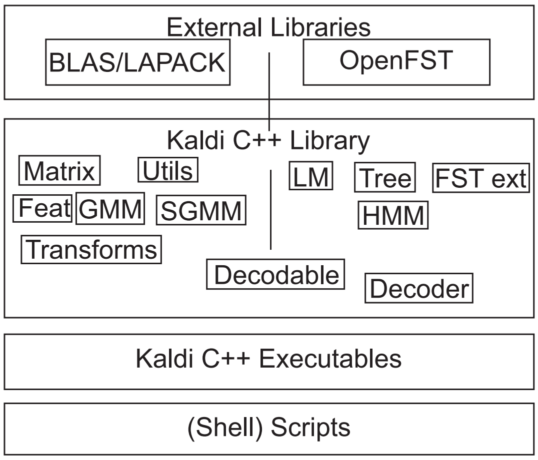


Fig. 7. Schematic overwiev of the Kaldi Toolkit.

In the Fig. 7 we show a scheme of the Kaldi toolkit. It consists of two external libraries which are free available too. OpenFST for the finite-state framework, and the next is numerical algebra libraries. It uses the standard Basic Linear Algebra Subroutines (BLAS) and Linear Algebra PACKage (LAPACK) routines for the latter. The library modules can be grouped into two parts, each depending on only one of the external libraries.

More descriptions about toolkit and about its functions we can find in the [21]. Our plans to the future is using Kaldi Toolkit for Slovak Speech Recognition.

V. Related Works

This section consists of some examples and practical applications. Many approaches are described below.

For testing cross-lingual knowledge transfer in DNNs, authors use the GlobalPhone corpus. They used 19 languages, 100 speakers per language and about 20 hours of audio material. As the main parameter for HMM-GMM systems they chose Word Error Rate (WER). For the 15 hrs corpus was WER around 15%, for 5 hrs of speech was WER 18.19% and the worst WER was in 1 hr of speech signal 21.53%. In the tandem setup (up to five layers deep) was WER for 15 hrs more than 21.2% in each language and the hybrid system had lower WER than tandem system (around 20%). More results are shown in [8].

Authors in [22] write about dimensionality reduction with Deep Neural Network. They describe hybrid model, context-dependent deep neural network hidden Markov models (CD-DNN-HMMs) and the Bottleneck features. The depth in this hybrid model is largely increased and the output nodes of the network are directly expanded from small number of monophone states to large number of context-dependent triphone tied-states.

The Bottleneck features are generated from neural networks, where one of the hidden layer has a very small number of hidden units. In this paper authors used a Mandarin LVCSR. They had 70 hours of speech from 1539 speakers. The results of dimensionality reduction and WER are written in [22].

In [5] authors report results of applying ANN. They used two databases where the first was TIMIT and the second one was corpus of Hungarian broadcast news. The Hungarian corpus consists of 8 TV channels and about 28 hours of recordings. They applied backpropagation training.

On the TIMIT database Mohamed at al. reported 22.3% error rate while Toth reported 21.8%. Hungarian news had 20.07% WER, but these results are not fully comparable. Because this two systems used different pruning beam widths. They used one to five hidden layers in the paper [5].

Paper [10] describes the time-delay neural network architecture. As input to the network it uses spectrograms, which are computed with a certain frame-rate. They are using the Tempo 2 Algorithm that is applied to phoneme classification. Bodenhause and Waibel designed an adaptive architecture of the optimal system. For training was used 783 phonemes and for testing was used 759 tokens. The network was initialized with random weights, random delays, constant widths of the input and one connection among two units. The weight had the best performance when we compare it with delays or widths. But their combination with weight can improve the performance. Authors trained and tested a variety of combinations of constant and adaptive parameters.

VI. Conclusion

In this work, the basic characteristics of the Neural Networks were introduced. This method consists of input, output and few hidden layers. It can has at least one hidden layer.

Deep NN must have two or more hidden layers. For instance some authors made an example using 1 to 5 hidden layers. When they used more hidden layers, the results were better. But we cannot use many hidden layers, because each layer has good but unwanted features too.

DNNs can divide in the feedforward neural network (FNN) and the second one is the recurrent neural network (RNN). RNN has a bigger capacity than a feedforward neural network because of forward and backward flow of information.

Deep neural networks (DNNs) are one of the most popular techniques in these days. It plays an important role in the modern recognition systems. There are lot of fuctions which can be used for solving many problems. Many laboratories are focus in these method.

Researchers and scientists are using DNN for pattern recognition, face recognition, speech emotion recognition, automatic speech recognition and many others. Neural networks increased the results and they are used in the modern systems.

NNs are becoming the standard in the world. In our speech laboratory we are using many approaches and various methods daily. We intend of transform our actual scripts for using with DNN.

Firstly we want to train some examples and find the best model for our data or we will looking for other (better) procedures. When we will train some model and the results will be comparable with our results from other techniques, we will be trying to increased it.

Our research team wants to investigate the NN in the next research. The visions to the future are in application of the Neural Networks for Large Vocabulary Continuous Speech Recognition (LVCSR) using Slovak Speech corpus.

Acknowledgment

The research presented in this paper was supported partially by Competence Center for Innovation Knowledge Technology of production systems in industry and services (ITMS project code 26220220155) (50%) and partially by VEGA 1/0075/15 (50%).

References

[1] N. Morgan and H. A. Bourlard, “Neural networks for statistical recognition of continuous speech,” *Proceedings of the IEEE*, vol. 83, no. 5, pp. 742–772, 1995.
[2] M. Gavrilescu, “Improved automatic speech recognition system using sparse decomposition by basis pursuit with deep rectifier neural networks and compressed sensing recomposition of speech signals,” in *Communications (COMM), 2014 10th International Conference on*. IEEE, 2014, pp. 1–6.
[3] P. McCollum, “An introduction to back propagation neural networks,” *The Newsletter of the Seattle Robotics Society*, 1998.
[4] J. Schmidhuber, “Deep learning in neural networks: An overview,” *Neural Networks*, vol. 61, pp. 85–117, 2015.
[5] L. Tóth and T. Grósz, “A comparison of deep neural network training methods for large vocabulary speech recognition,” in *Text, Speech, and Dialogue*. Springer, 2013, pp. 36–43.
[6] Y. Dong and D. Li, *Automatic Speech Recognition A Deep Learning Approach*. Springer, 2015.
[7] M. Katrák, “Klasifikácia foném reči neurónovou sieťou,” Ph.D. dissertation, Dept. of Electronics and Multimedia Communications, Technical University of Koice, Slovak Republic, 2010.
[8] P. Swietojanski, A. Ghoshal, and S. Renals, “Unsupervised cross-lingual knowledge transfer in dnn-based lvcsr.” in *SLT*, 2012, pp. 246–251.
[9] S. Thomas, M. L. Seltzer, K. Church, and H. Hermansky, “Deep neural network features and semi-supervised training for low resource speech recognition,” in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*. IEEE, 2013, pp. 6704–6708.
[10] U. Bodenhausen and A. Waibel, “Learning the architecture of neural networks for speech recognition,” in *Acoustics, Speech, and Signal Processing, 1991. ICASSP-91., 1991 International Conference on*. IEEE, 1991, pp. 117–120.
[11] S. Haykin and N. Network, “A comprehensive foundation,” *Neural Networks*, vol. 2, no. 2004, 2004.
[12] S. Katagiri, *Handbook of neural networks for speech processing*. Artech House, Inc., 2000.
[13] B. R. Smith, “Neural network enhancement of closed-loop controllers for ill-modeled systems with unknown nonlinearities,” Ph.D. dissertation, Virginia Polytechnic Institute and State University, 1997.
[14] S. J. Lee, K. C. Kim, H. Yoon, and J. W. Cho, “Application of fully recurrent neural networks for speech recognition,” in *Acoustics, Speech, and Signal Processing, 1991. ICASSP-91., 1991 International Conference on*. IEEE, 1991, pp. 77–80.
[15] R. Hecht-Nielsen, “Theory of the backpropagation neural network,” in *Neural Networks, 1989. IJCNN., International Joint Conference on*. IEEE, 1989, pp. 593–605.
[16] J. Wang, Z. Han, and S. Lung, “Speech emotion recognition system based on genetic algorithm and neural network,” in *Image Analysis and Signal Processing (IASP), 2011 International Conference on*. IEEE, 2011, pp. 578–582.
[17] G. Hinton, S. Osindero, and Y. Teh, *A Fast Learning Algorithm for Deep Belief Nets*. MIT Press, 2006, vol. 18.
[18] S. Young, G. Evermann, M. Gales *et al.*, “The HTK Book (for HTK version 3.4)(2009),” *Cambridge University Engineering Department. Inggris*, 1982.
[19] A. Lee, T. Kawahara, and K. Shikano, “Julius—an open source real-time large vocabulary recognition engine,” 2001.
[20] W. Walker, P. Lamere, P. Kwok, B. Raj, R. Singh, E. Gouvea, P. Wolf, and J. Woelfel, “Sphinx-4: A flexible open source framework for speech recognition,” 2004.
[21] D. Povey, A. Ghoshal, G. Boulianne, L. Burget, O. Glembek, N. Goel, M. Hannemann, P. Motlíček, Y. Qian, P. Schwarz *et al.*, “The kaldi speech recognition toolkit,” 2011.
[22] Y. Bao, H. Jiang, C. Liu, Y. Hu, and L. Dai, “Investigation on dimensionality reduction of concatenated features with deep neural network for lvcsr systems,” in *Signal Processing (ICSP), 2012 IEEE 11th International Conference on*, vol. 1. IEEE, 2012, pp. 562–566.

Object Tracking Based Model for Automatic Cut the Best Shot in Virtual Views of Live Streamed Video (May 2015)

[1]*Dávid CYMBALÁK (3rd year)*
Supervisor: [2]Frantisek JAKAB

[1,2]Dept. of Computers and Informatics, FEI TU of Košice, Slovak Republic

[1]david.cymbalak@cnl.sk, [2]frantisek.jakab@tuke.sk

***Abstract*—This summary paper deals with a approach for real-time object tracking from multiple virtual cropped views in 4K video eliminates the requirement to utilize multiple capturing devices or requirement to physically move the camera. System is based on TLD object tracking algorithm. Subsections of 4K video output from capturing device are utilized to create new "virtual copped views". Proposed model was experimentally implemented and tested in the laboratory conditions.**

***Keywords*—Automatic cut, Multi-cameras tracking, Live streaming, Object tracking.**

I. Introduction

Object tracking is widely covered area and focus of many studies. Survey carried out to estimate the performance of the latest algorithms may be found in [1]. Say we are interested in tracking of a lecturer or some presenter on a stage, first step is to provide some form of object tracking. In this way one user with one capturing device will be able to provide single tracking of an object.

When referring to 4K resolution we mean any video of size up to 4000 pixels per line. 4K is rather new technology, according to [2] was 4K support announced by YouTube in July of 2010. Though 4K is up to 4000 pixels per line [3] notes that most applications refer to 4K as a frame size with dimensions of 3840x2160. Authors further note that streaming of uncompressed 4K video is difficult due to high data bit-rate. However once some form of compression is used the communication is not that critical, this was proved by [4] who developed 4K bi-directional communication system using JPEG 2000. For brief introduction to 4K technology see [5].

From the variety of object tracking approaches we primarily focus on Tracking-Learning-Detection or for short TLD algorithm which was introduced by Zdenek Kalal as a part of his PhD thesis [6]. Idea of TLD was to design an algorithm for long-term tracking of the unknown objects, however detection of the objects with the known pattern is available as well (image with an known object of interest). Authors of research [7] extended TLD of external detector and thus enabled multi-object tracking at the same time. In [8] TLD was empirically demonstrated to be used in multi-camera environment, concluding that its operation is feasible, providing lightning conditions and orientation of cameras is kept analogous.In our research is TLD algorithm chosen to be used as a tool for the object tracking.

II. Automatic Cut Model Design

Based on our previous research [9] was created model and prototypes of solution which enhances video experience from specific live streams to distant audience with autonomous director that improves surveillance and tracking abilities through the various multi-camera systems. Solution evaluates the best view on the object in real-time and delivers only the best output. Multiple cameras, without manual operation, would track the speaker and automatically change to different viewing angles based on the movement. Our solution is designed to provide automated source switching based on actual position of the speaker without changing the physical placements of the cameras or without their rotation itself. The new approach of our auto switching best shot solution does not presume utilization of multiple physical cameras but only the one ultra-high definition camera through creation of a multiple virtual cropped views from the original image.

Frequently scene's conditions allow replacing the multiple physical cameras capturing such scene by one 4K camera. This method is usable for capturing various types of scenes with only one 4K camera, while being also adaptable in surveillance or production monitoring. Previous tracking of object of interest in video enabled us to define optimal virtual camera views in 4K based on object moving trajectory (Fig. 1) or establish the most common object of appearance. These virtual views could be of different sizes (360p, 480p, 720p, 1080p) and they can overlap or contain each other.

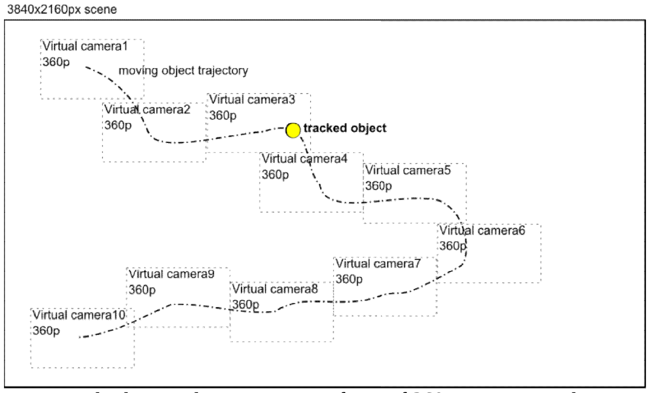


Fig. 1. Multiple virtual cameras view of size of 360p in one 4K video

Once the multiple views based on previous object appearances are defined or automatically formatted there is need to use an evaluation metric for each virtual view. This metric is calculated related to parameters of tracked object from sight of each virtual view. Metric is composed of positional element, dimensional element and reliability element:

$$M_n^t = (M_p^t + M_v^t) * M_d^t \qquad (1)$$

The dimensional element of metric is calculated from ratio of width of tracked object and width of virtual view. The reliability element is composed of the percentage expression of reliability of detection of the object being tracked using TLD algorithm. The positional element of metric is calculated from zonal division (Fig. 2) of each virtual view.

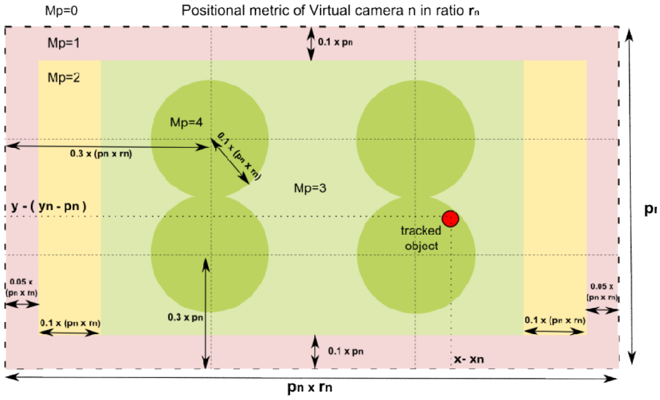


Fig. 2. Zonal division for calculating the positional metric for virtual view.

Experimental realization of system based on interconnecting the vMix system with 2x SDI 4K video capture card with RED 4K camera was implemented. This test environment allows us to create multiple various sized virtual views that are automatically switched to the output with FFmpeg encoder based on calculated evolutional metric in real time based on object tracking results from TLD (Fig. 3).

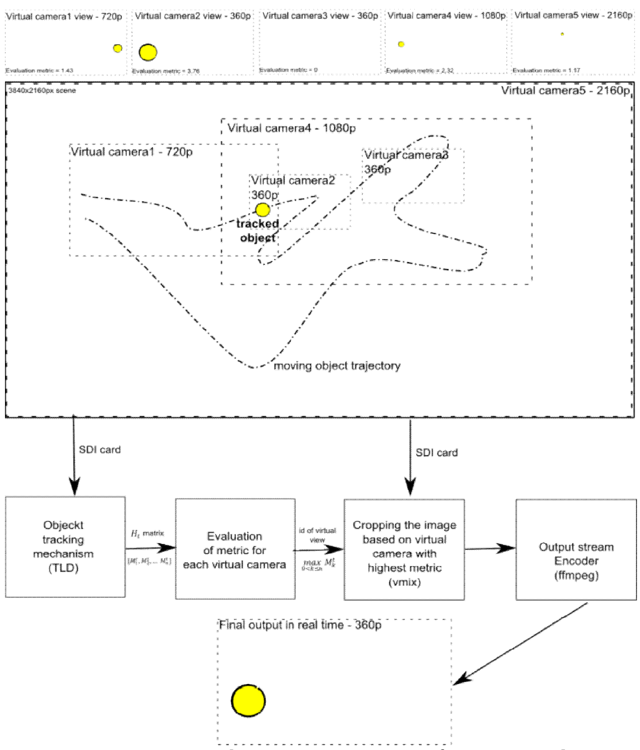


Fig. 3. Interconnection between components of experimental proposed solution

III. Conclusion and Next Steps

In this summary we presented a approach for automatic cut the best shot on the object of interest in a real-time utilizing 4K definition capturing device. Tracking of the object is utilized using TLD algorithm while selecting the best shot from the 4K video with virtual cropped view. This selection is automatic and based on the pre-learned pattern. As for the future work the plan test the system in the real world conditions. Related work could be found on another selected manuscripts created by author of this publication [10][11][12][13][14][15]

Acknowledgment

This Paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF.

References

[1] BRADSKI, G., & KAEHLER, A. (2008). Learning OpenCV: Computer vision with the OpenCV library. O'Reilly Media, Inc...
[2] Wu, Y., Lim, J. & Yang, M.-H. Online Object Tracking: A Benchmark. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2411 2418, 2013.
[3] Jarboe, G. YouTube and Video Marketing: An Hour a Day, 2nd Edition. Indianapolis: John Wiley & Sons, Inc., 2012.
[4] Shen, Y., Wu, L., Li, Y., Liu, S. & Wen, Q. Parallel Data Transfer Protocol. Enabling the New Era of Cloud Computing: Data Security, Transfer, and Management, pp. 238 255, 2014.
[5] Fujii, T., Kitamura M., Murooka T., Shirai D., Fujii T. & Takahara A. 4K & 2K multi-resolution video communication with 60 fps over IP networks using JPEG2000. International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS), pp. 167-170, 2009.
[6] Seeling, P. & Reisslein M. I. Want. Pixels. (Entering the age of 4k). IEEE Potentials, vol. 33, issue 6, pp. 27-30, 2014.
[7] Kalal, Z. Tracking Learning Detection. 2010. Cited: 18th December 2014. Available online: http://cmp.felk.cvut.cz/~matas/papers/kalal-2010-phd.pdf
[8] Piao, S., Berns, K. Multi-Object Tracking Based on Tracking-Learning-Detection Framework. Field and Assistive Robotics - Advances in Systems and Algorithms, pp. 74-87, 2014.
[9] Nebehay, G. Robust Object Tracking Based on Tracking-Learning-Detection. Cited: 23rd December 2014. Available online: http://gnebehay.github.io/OpenTLD/gnebehay_thesis_msc.pdf
[10] Cymbalák, D., Kainz, O., & Jakab, F. Extended Object Tracking and Stream Control Model Based on Predictive Evaluation Metric of Multiple-Angled Streams, International. Journal of Computer Theory and Engineering, Vol. 7, No. 5, pp 343-348, 2014
[11] CYMBALAK D., JAKAB F., MICHALKO M., Advanced multimedia solution for interactive teaching via whiteboards over IP network, Cyber Journals:

Multidisciplinary Journals in Science and Technology: Journal of Selected Areas in Software Engineering (JSSE), November Edition, 2012,s. 1-8, ISSN:1925-2676

[12] CYMBALAK D., PEKAR A.: Video streaming control based on actual location of tracked object in multi-cameras system, Poster 2013 : 17th International Student Conference on Electrical Engineering : May 16, 2013, Prague. - Praha : ČVUT, 2013 S. 1-6. - ISBN 978-80-01-05242-6

[13] CYMBALAK D., JAKAB F., MICHALKO M., Web portal based solution designed for delivery and management educational multimedia content with auto conversion ability, 19th Multimedia in Education : Conference Proceedings : June 13-14, 2013, Košice. Elfa, 2013 S. 29-32. - ISBN 978-80-8086-207-7

[14] CYMBALAK D., JAKAB F., MICHALKO M., Mechanism for video streaming control related to tracked speaker position ek, IEEE 17th International Conference on Intelligent Engineering Systems : proceedings : June 19-21, 2013, Costa Rica, 2013 P. 157-162. - ISBN 978-1-4799-0828-8

[15] CYMBALAK D., MICHALKO M.,JAKAB F.,Live broadcast of events with best shot on speaker using tracking methods on audience's mobile phones, ICETA 2013 : 11th IEEE International Conference on Emerging eLearning Technologies and Applications : proceedings : October 24-25, 2013, Stary Smokovec, Slovakia, : IEEE, 2013 P. 67-72. - ISBN 978-1-4799-2161-4

Object tracking in videoconferencing systems

[1]*Roman VÁPENÍK (1st year)*
Supervisor: [2]František JAKAB

[1,2]Dept. of Computer and Informatics, FEI TU of Košice, Slovak Republic

[1]roman.vapenik@tuke.sk, [2]frantisek.jakab@tuke.sk

***Abstract*—In this paper, three main topics – streaming technologies, videoconferencing systems and object tracking are presented. Brief overview of current high efficient video codecs is provided, followed by hardware, software and web-oriented videoconferencing solutions. Study are also presents techniques used in object tracking for an extraction of object information. In addition, the possibility of extending the videoconferencing systems with detection of facial expressions or parts of the body is introduced.**

***Keywords*—Object tracking, streaming technologies, videoconferencing system.**

I. Introduction

Nowadays videoconferencing systems have a large number of features. Except for transmitting video and audio between end users, they enable even more advanced options. Of course, these systems allow to compose messages, share files or own desktop on computer. In addition, often a common area is shared, allowing all participants to write or draw.

One possibility of expanding its features is implementation of object tracking mechanisms from video to the videoconferencing systems. Such a system could recognize the mood of the listener from facial expressions, differentiate whether to agree with something or not, or it would be possible to determine which answer shall be deemed correct.

II. Streaming Technologies

Streaming technologies, more specifically video codecs are very important for video data transport from one videoconferencing system to another. Comparison of three main video codecs is shown on Fig. 1.

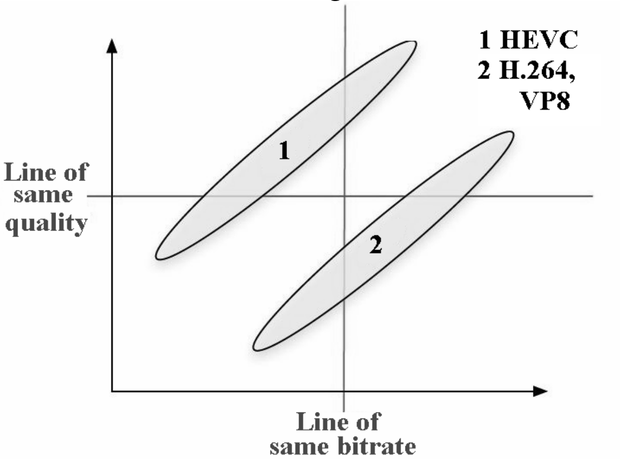


Fig. 1 Comparison of codecs - HEVC, H.264, VP8

A. Codec VP8

VP8 is an open source video compression format supported by a consortium of technology companies [1]. From the beginning of its development, the developers were focused to make overall design of codec which should comply with the following properties:

- low bandwidth requirement – VP8 is designed mainly for watchable quality video (~30dB in Peak Signal-to-Noise Ratio metric) to visually lossless quality (~45dB) [2],
- heterogeneous – the adaptability to each device from low-performance mobile phones to the most advanced desktop computers with multi-core processors [3],
- web video format – the adaptability of video to conditions which is the optimal for web browsers [3],
- hybrid transform with adaptive quantization and flexible reference frames – provides greater efficiency in processing video [2],
- efficient intra and inter prediction – improving the properties of both predictions from previous versions of the codec [2],
- high performance sub-pixel interpolation
- adaptive in-loop deblocking filtering
- frame level adaptive entropy coding
- parallel processing friendly data partitioning - joining units with similar entropy [3].

B. Codec H.264/AVC

As presented by [4] ***H.264*** is the standard for compression of video. It is a process of conversion digital video to format which takes less capacity when stored or transmitted. Video compression (or encoding video) is the basic technology for application like digital television, video on DVD, mobile television, videoconference and internet video streaming [5]. Standardization of video compression allow mutual co-operation for products from different manufacturers (e.g. encoders, decoders and storage media). The encoder converts video to compressed format and encoder converts video into a compressed format and decoder re-expands compressed video into uncompressed format [6].

H.264 video encoder implementing a prediction, transform and coding processes for the production of compressed H.264 stream. H.264 video decoder performs a complementary process of decoding, inverse transformation and reconstruction produced decoded video sequence [7].

The biggest advantage of H.264 compared to previous standard is its compression performance. Compared to standards such as MPEG-2 and MPEG-4 Visual an H.264 has:

- better video quality at the same bitrate compression,

- lower bitrate compression at the same image quality [8].

Standard DVD can, for example, store films with two-hour length with MPEG-2 format. Using H.264 codec, the length of the film may be doubled to four hours. Optionally the movie may be stored in higher quality than in the MPEG-2 [9].

In addition to improving the performance of compression, H.264 offers more flexibility in terms of compression and support of the transmission [8]. It is possible to choose from a wide range of compression tools, which is making it suitable for applications ranging from low-bitrate mobile data flow with a slight delay to the television with high definition on a professional level. The standard provides integrated support for the transfer and storage, including features that help minimize the impact of transmission errors [10].

C. Codec H.265/HEVC

High Efficient Video Coding - ***HEVC*** is a new standard for video compression which provides better performance than its successor H.264/AVC. Video source, consisting of a sequence of video frames is encoded or compressed by HEVC video encoder to form a compressed data stream. The compressed data stream is stored or transmitted. Video decoder "unpack" data stream and produces a sequence of decoded frames [11].

HEVC has the same basic structure as the previous MPEG-2 Video H.264/AVC, but it includes many partial enhancements:

- more flexible partitioning images from bigger to smaller parts,
- flexibility in prediction of size of transform blocks,
- more sophisticated interpolation and deblocking filter [12],
- more sophisticated prediction and signalization of motion vectors
- function for support efficient parallel processing [13].
- The result is standard for video coding which can allow better compression, at the cost of potentially increased processing performance.

III. Videoconferencing Systems

There are many videoconferencing systems providing basic or the advanced options of communication between two or more users. As an illustration in the following chapters are described some of these.

A. Software solutions

Skype is the most popular and best-known cross-platform software for Internet communication. The basic functions of Skype include text and voice communication and sending files. It allows video calls between two users or video conferencing between three or more users [14].

In addition to these basic services supports importing contacts from Microsoft Outlook and other email clients, web browsers integration with fast calling on published figures. It is also possible to buy premium features. SkypeOut allows calls to fixed and mobile networks worldwide, with SkypeIn service user is assigned a phone number to which you can attach to the telephone system, Voicemail provides voice mail and Skype SMS allows sending SMS messages [13]. The application is developed for a variety of platforms, is available for Windows, Linux and Mac OS for mobile devices running on Android, iOS, Windows, BlackBerry and Symbian, but also for Smart TV with support for Skype, for gaming or Xbox. Each user is identified by its Skype name. In the publication [14] author states that directory of users is decentralized and distributed among the nodes of the network and can therefore be easily adapted to a large number of user base without high costs. The main advantage of Skype is its ease of use and applicability almost anywhere.

B. Hardware solutions

TelePresence is one of the most popular hardware solutions. This system is designed to the best immediacy of communication. Cisco Telepresence is no ordinary conference solution. With this system it is possible to make a connection with other rooms worldwide. The rooms are designed in great detail, certification Cisco Systems guaranteeing compliance with strict standards as image and sound quality but also furniture, which is part of some larger Telepresence systems.

Cisco TelePresence systems use special cameras, which are located above the screen and provide optimum eye contact for all participants in the meeting [16]. Another part of the large integrated display provides live video with high resolution. Sound area is very good, thanks to the specially placed microphone and speaker in the room. Cisco TelePresence systems require approved type of lighting, acoustics, in some cases, integrated furniture in rooms. All these factors contribute to the quality that is required for each room Cisco TelePresence [16]. There are various equipped rooms, but the cost of these systems is too high, and few can afford it, which is the main drawback of this system

LifeSize is also a hardware solution, however for lower prices when compared to Cisco TelePresence. It offers a customizable solution usually consists of HD cameras, special microphone for high definition audio and hardware that process data. The entire system can be connected to a screen, projector or High Definition Television. It supports video communication in Full HD quality (1920x1080 pixels) with 30 frames per seconds (fps) or 60fps and HD quality (1280x720) with 60fps [17]. Devices also allow sharing documents, presentations and multimedia files. In addition to having a wide range of products that are easy to install and to operate.

LifeSize combines direct videoconferencing experience with many features and enables powerful, flexible and easy to use solution for communication. It does not provide such as authentic experience as Telepresence, but thanks to video and audio in HD, LifeSize products provide high-quality video conferencing solutions at comparatively lower price.

C. Web-oriented solutions

iWowWe is a web-oriented solution of videoconferencing system. It does not require installation of additional plug-ins or applications, it requires only webcam and internet connection. Everything happens in a Web browser. It allows up to 8 secure video streams, and in addition to video conferencing has many additional features such as sharing common areas (i.e. Whiteboard), where participants can draw, share various files like presentations, documents, videos and photos. It is also possible to edit documents, desktop sharing etc. The videoconference has capacity for up to 500 participants [18]. There is a web interface and also application for smartphones with Android operating system, which allows you to see and hear everything during the conference, open the shared documents, videos, and can also take part in the chat. The application is free, but the services themselves require a fee.

Webex is another web-oriented videoconferencing solution by world known company Cisco. According to the publication [19] Webex is operated "in the cloud", which means that the customer service is available 24 hours a day, 7 days a week, 365 days a year, without having to install of any device on the client's side and take care of its operation.

WebEx is an intuitive environment. On the right side we find the list of registered participants, chat and notes. The rest of the area can be filled with shared flat, presentation or sharing media files. Webex enables the sharing of some applications and web browser.

On the desktop, you can share a common area where conference participants can write or draw. Furthermore, they can also add documents, presentations or multimedia files. It also allows sharing of web content, either on your desktop, or as an application. The whole meeting can be recorded and saved directly to computer of user, but the disadvantage is that it is necessary to use the player from Cisco. The advantage is that Webex also supports mobile devices running on operating systems like iOS, Android, Windows and BlackBerry [19]. The number of users in the basic version is 8, and premium version is 100, of course, with a higher monthly fee.

The publication [20] states that ***Hangouts*** is a result of a combination of communication services, Google Talk, Google+ Messenger and Google+ Hangouts. It allows user to send and receive messages, pictures and other items, and even make free video calls and voice calls (calls between two clients or group calls).

Hangouts allows up to ten users simultaneously participate in videoconference with tools for conferencing and collaboration that goes beyond simple video. Anyone with a Google account can participate Hangouts sessions. Hangouts works on any browser whether the operating system is Windows or Mac. It is also available for users of smartphones and tablets.[21] Before first use, it is necessary to download easily installable plug-in the device.

Hangouts provides a set of tools for collaboration and conferencing for use in video chat interface. All users are able to view video from YouTube presentation on the main screen, share desktop and even edit a shared document on Google Drive [20].

Shortly after the launch Hangouts, Google allowed Hangouts On Air. After clicking the check box Hangouts On Air will broadcast the session to the attached YouTube channel where others can look up the live presentation. Hangouts On Air automatically creates a record of presentation and loads it on the YouTube channel [20].

As defined in [22], ***WebRTC*** (Web Real-Time Communication) is open and freely distributed project, which allows web browsers using real-time communication. Everything is made using relatively simple API JavaScript. All WebRTC components have been optimized to best serve this purpose [23].

WebRTC is a new standard that allows web browsers to communicate in real time using peer-to-peer. It provides a secure audio, video and data communications between HTML5 web browsers [23]. WebRTC offers programmers the ability to write web applications, rich media applications operating in real time without any need for plug-ins or download and install the application on your hard drive. It helps create a strong RTC platform, working between different browsers and between different platforms [24].

According to Johnston and Burnett [22] the biggest advantage of WebRTC is simplicity. With a few lines of JavaScript code the functionality of WebRTC can be utilized. However, initial configuration as a signaling, and search for a client, security settings, and many others is required when the WebRTC API is in use. In addition, some new layer protocols are necessary for full functionality of WebRTC [25]. However, most of the work is carried out by the three basic components of WebRTC [26]:

- MediaStream – access to input of the local device,
- RTCPeerConnection - setting the remote audio and video data,
- RTCDataChannel – data channel between end devices.

IV. Detection and Object Tracking

Methods for image detection using image segmentation divide the image into units called foreground in which the object and also section called the background can be found, background is thus unnecessary for further processing. The resulting foreground consists from group of elements in the image. In this group search element is detected using some kind of the model. Foreground can be extracted based on certain parameters such as color, shape or texture.

A. Extraction of objects on the basis of color

This type of extraction is carried out based on ***thresholding***. Thresholding is used to transfer an image with multiple levels of brightness in an image with two levels of brightness - black and white (Fig. 2). The point where brightness is greater than a predetermined threshold, the threshold (T), is associated with the brightness level 255 (or 1), the other point is assigned a value of 0. The resulting value of the threshold on the point can therefore be defined as g (x, y) = 1, where f (x, y) ≥ T or G (x, y) = 0, where f (x, y) <T [27]. The next step includes choosing the points with a value of 1 for the foreground and the value 0 as a background or vice versa. Thresholding can be performed with a global threshold value, which is valid for the entire image, or locally with a threshold value, which is adaptive in different parts of the image [28]. The threshold value can be defined constant or a method of automatic threshold or method of finding the optimal threshold (Otsu’s method).

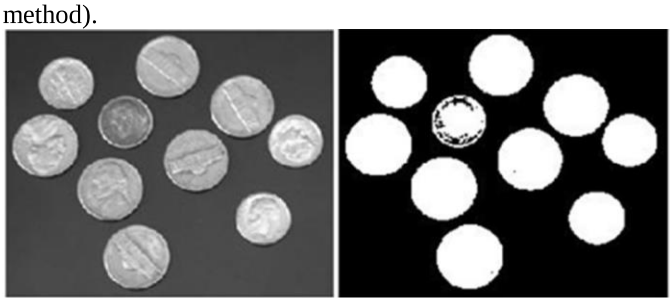


Fig. 2 Segmentation with thresholding [34]

B. Extraction of objects on the basis of shape

When light conditions is changing or there is lack of color differences from the background regarding the objects of interest, it is more efficient and reliable to implement detection based on proportional certain properties. Also, in case of implementation based on the extraction of color often resulting foreground object includes discontinuous points that lose form connected lines or shapes. These points can be connected by these methods together and reconstruct the original shape. In the extraction of objects based on shape it is first necessary to get the input image into shape with highlighted edges and lines associated with a clearly defined foreground and background in the form of using only two brightness levels (1,0). Subsequently it is necessary to transform the filtered image into a form suitable for finding the shape defined function. The first step - edge enhancement can be performed using Canny edge detection. The second step is the use of the Hough transform.

An algorithm called the ***Canny edge detection*** consists of 5 separate steps (Fig. 3): smoothing the image with noise removal, finding high levels of color changes, marking local maxima, double thresholding and finally eliminate insignificant unconnected edges [29].

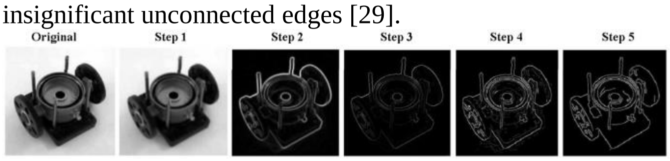


Fig. 3 Canny edge detection [29]

Hough transform has been developed for detection of analytically defined shapes such as lines, circles, ellipses etc. Theoretically it is possible to find any units that can be parameterized or express a function in the form f (v, c) = 0, where c is the vector of coordinates of vector parameters [29]. An example is the search of circle (Fig. 4), which contains three parameters: coordinates of the center and the radius. Input image to the Hough transform can be output image from Canny edge detector [30].

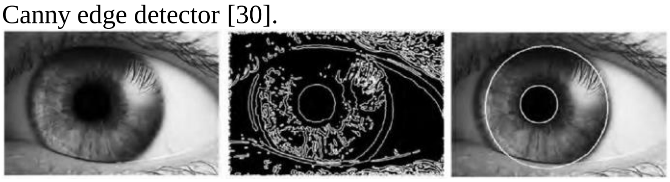


Fig. 4 Hough transform [35]

C. Extraction of objects on the basis of texture

Textures are visual patterns that have a certain homogeneity, which is not caused by the occurrence of only one color or one level of brightness in the image. Texture contains information on the structural arrangement of the surface of the object and its relation to the environment [32]. To obtain information about the textures, the most commonly used is ***discrete Fourier transform*** (DFT). It is the Fourier transform applied to a series of discrete complex numbers.

D. Current methods

As presented by [33] contemporary modern methods used for tracking objects in real time are: incremental visual tracking (IVT), variance ratio tracker (VRT), fragments-based tracker (FragT), online boosting tracker (BoostT), semi-supervised trackers (SemiT), extended semisupervised tracker (BeSemiT), L_1 tracker (L1T), multiple instance learning tracker (MIL), visual tracking decomposition algorithm (VTD) and track-learning-detection method (TLD). Reliability and performance of these algorithms can be verified by measuring the success of maintaining the detection of the reference object and by monitoring the wrong shift center position of the object.

V. Problems with Object Tracking in Videoconferencing Systems

In the introduction part was illustrated one of the possibilities for extending the video conferencing systems with ability of object tracking. The problem could occur especially in systems of peer-to-peer where the client requires a great deal of available resources. Personal computer might not be able to handle such processing of the objects in the video.

For example, as presented by [36], in WebRTC technology, if there is many-to-many conference type with Full Mesh topology (Fig. 5) the problem of resources depletion may occur. Data is distributed from each client to each. There is a problem with the procedural load for all participants.

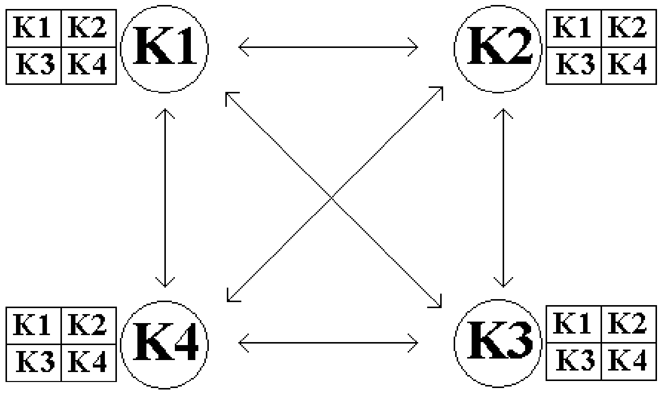


Fig. 5 Many-to-many conference type with Full Mesh topology in WebRTC technology [36]

Solution might be to distribute the processing of video to one or more machines dedicated to this purpose, with sufficient processing power.

VI. Conclusion

This paper introduces possibility of extending the videoconferencing systems with detection of facial expressions or parts of the body. The analysis and related researches carried out in the area of object tracking and videoconference systems shown possibilities of merging together both tracking and videoconference conclusion. However issues related to resources exhaustion are expected on the side of clients, due to this it might be feasible to shift computations to other resources. Bearing in mind just stated, future research is to focus on possibilities of implementation of specific object tracking techniques to particular videoconference systems and propose a solution for computational load distribution among other devices.

Acknowledgment

This Paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF.

References

[1] SEELING, Patrick, et al. Video network traffic and quality comparison of vp8 and h. 264 svc. In: Proceedings of the 3rd workshop on Mobile video delivery. ACM, 2010. p. 33-38.

[2] BANKOSKI, Jim; WILKINS, Paul; XU, Yaowu. Technical overview of VP8, an open source video codec for the web. In: ICME. 2011. p. 1-6.

[3] FELLER, Christian, et al. The VP8 video codec-overview and comparison to H. 264/AVC. In: Consumer Electronics-Berlin (ICCE-Berlin), 2011 IEEE International Conference on. IEEE, 2011. p. 57-61.
[4] KALVA, Hari. The H. 264 video coding standard. MultiMedia, IEEE, 2006, 13.4: p. 86-90.
[5] SCHWARZ, Heiko; MARPE, Detlev; WIEGAND, Thomas. Overview of the scalable video coding extension of the H. 264/AVC standard. Circuits and Systems for Video Technology, IEEE Transactions on, 2007, 17.9: p.1103-1120.
[6] WIEGAND, Thomas, et al. Overview of the H. 264/AVC video coding standard. Circuits and Systems for Video Technology, IEEE Transactions on, 2003, 13.7: p. 560-576.
[7] MALVAR, Henrique S., et al. Low-complexity transform and quantization in H. 264/AVC. Circuits and Systems for Video Technology, IEEE Transactions on, 2003, 13.7: p. 598-603.
[8] RICHARDSON, Iain E. H. 264 and MPEG-4 video compression: video coding for next-generation multimedia. Wiley. com, 2004.
[9] YIN, Peng, et al. Fast mode decision and motion estimation for JVT/H. 264. In: Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on. IEEE, 2003. p. III-853-6 vol. 2.
[10] VETRO, Anthony; WIEGAND, Thomas; SULLIVAN, Gary J. Overview of the stereo and multiview video coding extensions of the H. 264/MPEG-4 AVC standard. Proceedings of the IEEE, 2011, 99.4: p. 626-642.
[11] SULLIVAN, Gary J.; OHM, Jens-Rainer. Recent developments in standardization of high efficiency video coding (HEVC). In: SPIE Optical Engineering+ Applications. International Society for Optics and Photonics, 2010. p.77980V-77980V-7.
[12] RICHARDSON, Iain. HEVC: An introduction to High Efficient Video Coding
[13] UGUR, Kemal, et al. High performance, low complexity video coding and the emerging HEVC standard. Circuits and Systems for Video Technology, IEEE Transactions on, 2010, 20.12: p. 1688-1697.
[14] ROSSI, Dario; A Detailed Measurement of Skype Network Traffic. ENST ParisTech – INFRES, 2008.
[15] BASET, Salman A.; SCHULZRINNE, Henning. An analysis of the skype peer-to-peer internet telephony protocol. arXiv preprint cs/0412017, 2004.
[16] SZIGETI, Tim, et al. Cisco TelePresence Fundamentals. Cisco Press, 2009.
[17] PAN, Hui. Telecom Mergers & Acquisitions Monthly Newsletter November 2009. Information Gatekeepers Inc.
[18] LYNN, Raena, Video Conferencing? TSA iWowWe Review, 2015.
[19] Brusilovsky, Peter. "WebEx: Learning from Examples in a Programming Course." In WebNet, vol. 1, p. 124-129. 2001.
[20] BOLTON, R.: Google Hangouts Product Review. JCHLA/JABSC Vol. 34, 2013, p.39.
[21] XU, Yang, et al. Video telephony for end-consumers: measurement study of Google+, iChat, and Skype. In: Proceedings of the 2012 ACM conference on Internet measurement conference. ACM, 2012. p. 371-384.
[22] JOHNSTON, Alan B.; BURNETT, Daniel C. WebRTC: APIs and RTCWEB Protocols of the HTML5 Real-Time Web. Digital Codex LLC, 2012.
[23] GRIGORIK, Ilya; Secure Communication in WebRTC. High-Performance Browser Networking. O'Reilly Media, 2013
[24] BERGKVIST, A.; BURNETT, D.; JENNINGS, C. A. Narayanan," WebRTC 1.0: Real-time Communication Between Browsers. World Wide Web Consortium WD WD-webrtc-20120821, 2012.
[25] PERKINS, C.; WESTERLUND, M.; OTT, J. Web Real-Time Communication (WebRTC): Media Transport and Use of RTP. draft-ietf-rtcweb-rtp-usage-05, 2012.
[26] JOHNSTON, Alan; YOAKUM, John; SINGH, Kundan. Taking on webRTC in an enterprise. Communications Magazine, IEEE, 2013, 51.4: p. 48-54.
[27] DOBEŠ M.: Zpracování obrazu a algoritmy v C#. Praha : BEN, 2008, ISBN 978-80-7300-233-6.
[28] TOTH S.: Rozpoznavanie dopravnych znaciek a ioh pouzitie v mapovych aplikaciach, GIS Ostrava 2011
[29] MOESLUND T.: Canny Edge Detection, CSL783: Digital Image Processing, 2009
[30] DOBEŠ M.: Zpracování obrazu a algoritmy v C#. Praha : BEN, 2008, ISBN 978-80-7300-233-6.
[31] SHAH M. - JAVED O.: Automated Multi-Camera Surveillance, Algorithms and Practice, The international Series in Video Computing , Springer : New York, 2008, ISBN-10: 0387788808
[32] SMITH J. - CHANG S.: Automated binary texture feature sets for image retrieval , ASSP-96. Conference Proceedings., 1996 IEEE International Conference on. Vol. 4. IEEE, 1996.
[33] WANG Q. a kol.: An Experimental Comparison of Online Object Tracking Algorithms, Proceedings of SPIE: Image and Signal Processing Track, 2011
[34] KAMENSKISH M.: Segmentation using histogram thresholding, EXSE 529 Computer and Biological Vision, 2010
[35] BEBIS G.: General Hough Transform, CSL783: Digital Image Processing, 2009
[36] VÁPENÍK Roman.: Secured Web Oriented Videoconferencing System For Educational Purposes Using WebRTC Technology, ICETA conference, Starý Smokovec (Slovakia), 2014

On Reconfigurable Fault Tolerant Control Systems

[1]*Vladimír SERBÁK (4st year),*
Supervisor: [2]*Anna FILASOVÁ*

[1,2]Dept. of Cybernetics and Artificial Intelligence, FEI TU of Košice, Slovak Republic

[1]vladimir.serbak@tuke.sk, [2]anna.filasova@tuke.sk

***Abstract*—In this paper we briefly describe the state of our research in fault tolerant control systems, the solved design problems in the last year and the plans and aims to the future.**

***Keywords*—additive fault estimation, control reconfiguration, fault detection, fault tolerant control systems (FTCS), linear matrix inequalities**

I. Introduction

Operating conditions in modern engineering systems are still exposed to possibility of system failure. Any failure of the sensors, actuators or other system components can drastically significantly change the system behavior. Fault tolerant control (FTC) allows create a strategy to improve reliability of the whole system and so many techniques have been proposed especially for sensor and actuator failures with application to a wide range of engineering fields. In some cases, fault estimation strategies are needed only to carry on controlling the faulty system and, respecting this fact, many sophisticated modifications have been developed, e.g., sliding mode observers, neural network based approaches and adaptive observer technique. The standard approach to control reconfiguration discards the nominal controller from the control loop and replace it with a new controller so that controller parameters are re-tuned in fault conditions and in dependency on the remaining set of sensors and actuators to recover the performance of the fault-free control system [1]. By contrast, instead of adapting the controller to the faulty plant, the virtual approach keeps the nominal controller in the reconfigured closed-loop system and virtually adapt the faulty plant to the nominal controller in such a way that the virtual reconfiguration block together with the faulty plant imitates the fault-free plant [2], [3]. Designated to sensor faults the reconfiguration block is termed virtual sensor (VS), while in the case of actuator faults is named virtual actuator (VA).

II. The initial status

The primary objective of the PhD. thesis is focused on fault tolerant control systems (FTCS). In FTCS to estimate actuator faults the principles based on adaptive observers, we have provided an enhanced algorithm using H_∞ norm approach. Applied enhanced conditions in the scheme increase rapidity and develop a general framework for fast fault estimation in adaptive observer structures for linear deterministic systems. The approach utilizes the measurable input and output vector variables, design conditions are based on linear matrix inequality (LMI) technique giving an effective way to calculate the observer parameters [4].

For the industrial systems operating at acting conditions which are defined not only by given amplitude limitations but also by the ratios of certain process variables has been introduced in [5] the technique which makes use of the equality constraints formulation for the discrete-time linear systems, and was extensively applied in the reconfigurable control design tasks [6], [7]. Because the mentioned control algorithms use the static decoupling principle as a rule in setting the system working points, the prescribed values of equality constraints in the enforced mode are slithered in an additive offset. To become values of the equality constraints linearly dependent on the desired system output values, the discrete version of PI controller has been designed in the structure which ensures the system stability and sets up the required linear dependency of the equality constraints. Since the problem is generally singular, using standard form of the Lyapunov function and a symmetric positive definite slack matrix, the design conditions was proposed in the form of the enhanced Lyapunov inequality [8]. The next step was applying this technique in the reconfigurable control design tasks.

III. Current results

A. *On virtual observers in FTC*

The virtual observers adjusted for linear continuous-time systems under influence of sensor and actuator faults we have designed using linear matrix inequality (LMI) techniques and, extending the bounded real lemma approach that is used to regularize singular tasks in the synthesis VS, H_∞ norm principle we have proposed for design of VA with reflecting the unknown disturbance. Using virtual observers, the virtual elimination of faults influence on the system output is analyzed to obtain the minimum invasive control reconfiguration adapting the faulty plant to the nominal controller by hiding faults from the controller input point of view without the nominal controller redesign at a fault occurrence time. The virtual parts are formulated as autonomous algorithms that may be performed online starting with dependence on the detection subsystem fault localization time. The contribution was submitted to the *20th International Conference on Process Control 2015 in Slovak Republic, Štrbské Pleso.*

For the FTC we have designed VA also in structures with bi-proper dynamic output controllers (BP DOC). The considered FTC design problem is parameterized in such LMIs set with one additional LME which admit more freedom in guaranteeing quadratic performance of the output feedback control with BP DOC and VA respecting the disturbance acting on the system. Design conditions of the BP DOC controller existence manipulating the stability of the closed-loop systems imply the control structure which stabilize the system in the sense of Lyapunov and the controller design task is a solvable numerical problem. The additional benefit of the method is that controller uses minimum feedback information with respect to desired system output and the approach is flexible enough to allow the inclusion of unknown disturbance. The contribution was accepted to the *9th IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes 2015 in Paris*.

B. Control reconfiguration based on the PI constrained control algorithms

The method is based on discrete-time constrained PI control techniques with state constraints defined by linear equalities, where a single sensor fault is described by an equality constraint given on the state variable associated with the faulty sensor. Controller switching is taken into account since such different faulty system representations is known, and stabilizing controllers are pre-computed off-line. The design conditions are proposed in the form of the enhanced Lyapunov inequality. Design ideas and generalizations emphasize an advantage that the faulty sensor output can stay fixed at zero during the time the reconfigured PI controller is continuing its mission.

C. Actuator faults estimation

Taking into account that adaptive observers can be used in order to deal with only slowly varying additive system parameters, the considered faults are assumed to be additive in the adaptive observer structure and slowly-varying, which makes the design of the adaptive observer reasonable. Such adaptive observer structures can be described as follows

$$\dot{\boldsymbol{q}}_e(t) = \boldsymbol{A}\boldsymbol{q}_e(t) + \boldsymbol{B}\boldsymbol{u}(t) + \boldsymbol{F}\boldsymbol{f}_e(t) + \boldsymbol{J}(\boldsymbol{y}(t) - \boldsymbol{y}_e(t)), \quad (1)$$

$$\boldsymbol{y}_e(t) = \boldsymbol{C}\boldsymbol{q}_e(t), \quad (2)$$

where $\boldsymbol{q}(t) \in I\!R^n$, $\boldsymbol{q}_e(t) \in I\!R^n$, $\boldsymbol{u}(t) \in I\!R^r$, $\boldsymbol{y}(t) \in I\!R^m$, $\boldsymbol{y}_e(t) \in I\!R^m$, $\boldsymbol{f}(t) \in I\!R^p$, $\boldsymbol{f}_e(t) \in I\!R^p$ are vectors of the system, observer, input, output, estimated output, unknown fault, estimated fault, respectively and $\boldsymbol{A} \in I\!R^{n\times n}$, $\boldsymbol{B} \in I\!R^{n\times r}$, $\boldsymbol{C} \in I\!R^{m\times n}$ are matrices of the system dynamic, input and output, $\boldsymbol{F} \in I\!R^{n\times p}$ is the fault input matrix and $\boldsymbol{J} \in I\!R^{n\times m}$ is the estimator gain matrix.

The task is to design the matrix $\boldsymbol{J}$ in such a way that the observer dynamics matrix $\boldsymbol{A}_e = \boldsymbol{A} - \boldsymbol{J}\boldsymbol{C}$ is stable and $\boldsymbol{f}_e(t)$ approximates a slowly varying actuator fault $\boldsymbol{f}(t)$.

The state observer (1), (2) is combined with the law for the fault estimation updating of the form

$$\boldsymbol{f}_e(t) = \boldsymbol{G}\boldsymbol{H}^T\boldsymbol{e}_y(t), \quad \boldsymbol{e}_y(t) = \boldsymbol{y}(t) - \boldsymbol{y}_e(t), \quad (3)$$

where $\boldsymbol{H} \in I\!R^{m\times p}$ is the adaptation gain matrix and $\boldsymbol{G} = \boldsymbol{G}^T > 0$, $\boldsymbol{G} \in I\!R^{s\times s}$ is a learning weight matrix that has to be set interactively in the design step.

A disadvantage of the structure of the adaptation rule presented in [4] is that its feasibility is conditioned by the structure of the input fault matrix $\boldsymbol{F}$, which may cause the solution does not exist because this inequality becomes singular. We propose to solve the problem by using the adaptation rule of the form (3) when the inequality element $2\delta\boldsymbol{Q}$ is decoupled from $\delta\boldsymbol{F}\boldsymbol{T}\boldsymbol{Q}$.

In this case, the actuator fault estimator is stable if for given positive $\delta \in I\!R$ there exist symmetric positive definite matrices $\boldsymbol{P}$, $\boldsymbol{Q} \in I\!R^{n\times n}$, matrices $\boldsymbol{H} \in I\!R^{m\times p}$, $\boldsymbol{Y} \in I\!R^{n\times m}$ and a positive scalar $\gamma \in I\!R$ such that

$$\boldsymbol{P} = \boldsymbol{P}^T > 0, \quad \boldsymbol{Q} = \boldsymbol{Q}^T > 0, \quad \gamma > 0, \quad (4)$$

$$\begin{bmatrix} \boldsymbol{Q}\boldsymbol{A}+\boldsymbol{A}^T\boldsymbol{Q}-\boldsymbol{Y}\boldsymbol{C}-\boldsymbol{C}^T\boldsymbol{Y}^T+\boldsymbol{C}^T\boldsymbol{C} & * & * \\ \boldsymbol{P}-\boldsymbol{Q}+\delta\boldsymbol{Q}\boldsymbol{A}-\delta\boldsymbol{Y}\boldsymbol{C} & -2\delta\boldsymbol{Q} & * \\ \boldsymbol{F}^T\boldsymbol{Q} & \boldsymbol{0} & -\gamma\boldsymbol{I}_p \end{bmatrix} < 0, \quad (5)$$

$$\boldsymbol{C}^T\boldsymbol{H} = \delta\boldsymbol{Q}\boldsymbol{F}. \quad (6)$$

When the above conditions are affirmative the estimator gain matrix is given by the relation

$$\boldsymbol{J} = \boldsymbol{Q}^{-1}\boldsymbol{Y}. \quad (7)$$

If a single actuator fault represents loss of the associated actuator effectiveness, additive actuator fault signals, estimated by an adaptive fault observer, can be used in the fault tolerant control structure with the control law

$$\boldsymbol{u}(t) = -\boldsymbol{K}\boldsymbol{q}(t) - \boldsymbol{f}_e(t) + \boldsymbol{W}\boldsymbol{w}(t), \quad (8)$$

which can compensate the impact of faults, where $\boldsymbol{K} \in I\!R^{r\times n}$ is the control gain matrix, $\boldsymbol{w}(t) \in I\!R^m$ is desired closed-loop output signal vector and $\boldsymbol{W} \in I\!R^{m\times m}$ is the signal gain matrix. The contribution was accepted to the *16th International Carpathian Control Conference 2015 in Hungary*.

IV. Future Work

Our next goal in the field of FTC design is to investigate the problem of the reconfiguration of the continuous-time nonlinear systems using the Takagi-Sugeno dynamic VA in structures with Takagi-Sugeno dynamic output controller of order equal to the plant model order and using the Takagi-Sugeno VA extended to contain integrating state. The design procedure will be based on the solution of the set of LMI and LME that will ensure the closed-loop quadratic stability using Lyapunov approach, aggregating the fuzzy interactions among the subsystems.

The proposed LMI based methods present some design features and modifications where it was emphasized that the advantage offered by such approach is a collection of feasible algorithms with enough robustness to the sensor and actuator faults.

Acknowledgment

The work was supported by VEGA, the Grant Agency of the Ministry of Education and the Academy of Science of Slovak Republic under Grant No. 1/0348/14.

References

[1] A. Zolghadri, D. Henry, J. Cieslak, D. Efimov, and P. Goupil, *Fault Diagnosis and Fault-Tolerant Control and Guidance for Aerospace Vehicles. From Theory to Application*. London: Springer-Verlag London, 2014. ISBN 978-3-642-17627-2

[2] J. Richter, *Reconfigurable Control of Nonlinear Dynamical Systems. A Fault-Hiding Approach*. Heidelberg: Springer-Verlag Berlin, 2011. ISBN 978-3-642-17627-2

[3] T. Steffen, *Control Reconfiguration of Dynamical Systems. Linear Approaches and Structural Tests*. Heidelberg: Springer-Verlag Berlin, 2005. ISBN 978-3-540-25730-1

[4] D. Krokavec, A. Filasová, V. Serbák, and P. Liščinský, "An enhanced approach to actuator fault estimation design for linear continuous-time systems," *Journal of Physics: Conference Series*, vol. 570, 2014. doi: 10.1088/1742-6596/570/7/072002

[5] S. Ko and R. R. Bitmead, "Optimal control for linear systems with state equality constraints," *Automatica*, vol. 43, pp. 1573–1582, 2007.

[6] D. Krokavec and A. Filasova, "An active reconfiguration scheme for systems with single sensor faults," in *Prepr. 8th IFAC Symp. Fault Detection, Supervision and Safety of Technical Processes 2012*, Mexico City, 2012, pp. 624–629.

[7] D. Krokavec and A. Filasová, "Control reconfiguration based on the constrained lq control algorithms," in *Prepr. 7^{th} IFAC Symp. Fault Detection, Supervision and Safety of Technical Processes 2009*, Barcelona Spain, 2009, pp. 686–691.

[8] D. Krokavec, A. Filasová, and V. Serbák, "State control of discrete-time linear systems to be bound in state variables by equality constraints," *Journal of Physics: Conference Series*, vol. 570, 2014. doi: 10.1088/1742-6596/570/2/022005

Opinion analysis in conversational content focusing on dictionary approach

[1]*Martin Mikula (1st year),*
Supervisor: [2]Kristína Machová

[1,2]Department of cybernetics and AI, FEI, Technical University of Košice, Slovak Republic

[1]Martin.Mikula@tuke.sk, [2]Kristina.Machova@tuke.sk

***Abstract*—More and more people use Internet in these days. They communicate and exchange their opinions and life experiences. It is very difficult to read all of these discussions. Some of them can have more than hundred contributions. For this reason, the processing of these discussions uses methods of opinion mining and sentiment analysis. In this article, we describe both the meaning opinion mining and methods that are often used for opinion analysis. We divide methods into two main groups - machine-learning approaches and approaches based on lexicons. In last part we describe solved and unsolved problems and identify solvable and unsolvable problems for the next direction of our work.**

***Keywords*—opinion analysis, machine-learning approach, dictionary approach, opinion mining**

I. Introduction

Along with rising possibilities of the Web, the number of different communication lines between people increase. They communicate with each other and influence mutually their opinion. This is the reason why people have changed their buying habits in last few years. Previously people only went to shop and bought a new product. Now, they read few discussions and comments which include reviews and experiences from other users and then they buy a new product. For these new buyers it is important and interesting to know what other people think and what their opinion is. But these discussions are often very long and people haven't got any time to read them all. In this moment it is useful to use automated algorithm for opinion analysis.

The result of this effort are emerging fields such as opinion mining and opinion analysis. They are often used in the same way. They focus on polarity detection and recognition of different emotions. Both of these fields use techniques of data mining and natural language processing (NLP) for discovering information and opinion from the text. This textual data include opinionated texts such as blogs or review articles and comments of products. Associated numerical data such as aspect rating are also included [1], [2].

Mining sentiment of natural language is a big challenge. It requires understanding of explicit or implicit, regular and irregular and syntactical and semantic language rules. In NLP there are a lot of unsolved problems, e.g. negotiation handling, co-reference resolution, word intensification and word-sense disambiguation. Opinion analysis is very restricted by these problems because the system needs to understand positive and negative sentiment of each sentence [1].

Sentiment analysis consists of set of different tasks [3]:

- *subjectivity analysis* - (Does the conversation obtain subjective opinion?) depends on finding, whether the opinion is presented in text or not, but orientation of this polarity is not important
- *opinion classification* - (Is the conversation about subject positive or negative?) it identifies positive and negative expressions in text and classifies them into positive or negative sentiments
- *topic modeling* - (What is the object of this conversation?) the goal is to identify set of themes from collection of documents, researchers try to find documents that are relevant to all themes
- *authorship identification* - (Who is the opinion author?) it identifies person, who wrote selected contribution
- *opinion spam detection* - (Is the content of some discussion contribution informative?) - this task tries to remove text, which does not have any relevant information for reader

II. Process of opinion classification

Opinion classification involves these steps:

1) *word subjectivity* - we need only words that express opinion to subject
2) *word orientation* - we identify the type of word subjectivity
3) *strength of polarity* - we define the strength of word polarity

A. Word subjectivity

For classification we need only words with subjectivity. Subjective words express certain attitude about the subject. Subjective words are often adjectives. We also use verbs, nouns and adverbs. Finding words with subjectivity is one part of the text pre-processing. We divide these words into categories - words that are useful for classification and other words.

B. Word orientation (polarity)

In this step we determine the type of polarity for each subjective word. We often distinguish two types of polarity. Positive and negative polarity. In some cases we can use neutral polarity (when contribution doesn't contain opinion). There are words like *long* or *easy*. Their polarity depends on the domain in which they are used. In work [4] they present study of influence of multi-polarity words on domain adaptation. They detect this word in domain-specific reviews and open-domain corpus of tweets. Detection of correct connection word with domain improves opinion classification. Word orientation is linked with changing polarity. We use two types of negation:

- *switch negation* - changes polarity of word from positive to negative or negative to positive with the same strength
- *shift negation* - changes only strength of polarity, polarity can be changed to opposite but not with the equal strength or the same polarity with lower strength

It is interesting that the negative sentiment is perceived as more marked than positive sentiment, pragmatically and psychologically. In many cases, negation tends to be expressed in euphemistic way, which makes negation toi be more difficult to identify in general [5].

C. Strength of polarity

Some approaches, [6] and [5] work with strength of polarity. All words used in the text haven't got the same strength. Some words are more positive/negative than other. For example perfect is more positive than good. In these approaches authors created different degrees of polarity for more and less positive (or negative) words. Intensification is in connection with strength of polarity. Intensification means changing (increasing or decreasing) polarity for each word with another word called intensifier. We divide two types of intensification:

- in the first type is strength of polarity changed above one degree
- in the second type is strength of polarity changed above percentage of specific word

Second type of intensification allows the modulation of the strength of word polarity by the strength of intensifier. Intensifier with more strength (e.g. very) changes degree of word more than intensifier with less strength (e.g. little).

III. Opinion classification approaches

For opinion classification we utilize different types of approaches. In general, there are two basic types of approaches. The first approach is based on lexicons. In this approach we calculate the orientation of document from orientation of words. Words are saved in dictionaries. The second approach is based on machine-learning or statistical approach. This approach uses classification algorithms, mainly SVM (Support Vector Machine), Naive Bayes and KNN (K-Nearest Neighbors). Classifiers are trained on particular data sets. Thess classifiers reach high accuracy in detecting polarity in domain that they are trained to, their performance drops down when they are used in different domain. Other type of division is described as division of endogenous and exogenous methods [7]. Endogenous methods don't need any external information instead of information that are part of algorithm or sample data. Exogenous methods use external data such as dictionaries or created rules. In some works, methods are divided into approaches based on supervised and unsupervised learning.

A. Machine-learning approaches

Naive Bayes, SVM, KNN and Maximal Entropy are approaches that often use machine-learning approaches. Machine-learning algorithms are trained on large training corpus of opinionated annotated texts. They use features such as unigrams or bigrams with or without POS (part-of-speech) labels. The most successful is using basic unigrams [8], [9]. System doesn't learn only affective valence of keyword, but takes into computation the presence of other non-opinion words, punctuation and word frequencies. Classification tasks are divided into several sub-tasks:

1) Data pre-processing
2) Feature selection and/or feature reduction
3) Representation
4) Classification
5) Post processing

In feature selection and feature reduction we try to reduce the number of features in text for the next steps. The classification part finds the actual mapping between patterns and targets. Machine learning is a promising way for sentiment classification to reduce the annotation cost [10]. Machine-learning approaches achieve very good accuracy in domain in which they are trained but results decrease when used in other domain.

1) Naive Bayes Classification: Naive Bayes (NB) Classifier is the simplest Bayesian Classifier used in many machine-learning tasks. Although it is applied in many tasks, it requires independent of each example. Naive Bayes classifier is simple classifier based on probabilistic Bayes theorem. It is suited when the dimensionality of inputs are high.

It is approach of classification which assigns into classes by E.q.(1), where $C = \{c_1, c_2\}$ is a set of classes that consist of positive and negative messages and *d* is a given document.

$$c_{NB} = argmax_{c_j \in C} P(c_j|d) \quad (1)$$

Naive Bayes classifier uses Bayes rule Eq. (2):

$$P(c|d) = \frac{P(c)P(d|c)}{P(d)} \quad (2)$$

Where *P (d)* doesn't play role in selecting c_{NB}. For estimation of term *P (d|c)*, Naive Bayes decomposes if by assuming the *fi's* are independent to given *d's* class as in Eq. (3):

$$P_{NB}(c|d) = \frac{P(c)(\prod_{i=1}^{m} P(f_i|c)^{n_i(d)})}{P(d)} \quad (3)$$

Where *m* is the no features and fi is the feature vector. Training method consisting of relative frequency estimation *P(c)* and *P(fi|c)* [10]. Naive Bayes need these assumptions:

- words in text are independent
- doesn't count with sequence of word
- doesn't count with length of the message

We can use two types of implementation of Naive Bayes classifier - a Bernoulli model [11] and multinomial model [11]. The difference between these two models is that Bernoulli model is considering only in presence of word in message. Multinomial model counts also with the number of word occurrences in the text. In Table I is example of vector notation of text [12].

In [12] they used both types of NB classifier, Bernoulli and multinomial model and compare them. Multinomial modes obtained better results than Bernoulli model.

2) Maximum Entropy: Another type of machine-learning is Maximum entropy. It is a technique that has been proven effective in natural language processing. Sometimes, it reaches better results than Naive Bayes tan standard text classification. Its estimate of P(c|d) from the exponential as in Eq. (4):

$$P_{ME}(c|d) = \frac{1}{Z(d)} exp(\sum_i \lambda_{i,c} F_{i,c}(d,c)) \quad (4)$$

Z(d) is a normalization function. *Fi* and *c* are class functions for feature *fi* and class *c* as in Eq. (5):

$$F_{i,c}(d,c') = \begin{cases} 1 & n_i(d) > 0 \; and \; c' = c \\ 0 & \text{otherwise} \end{cases}$$

For instance, a particular feature/class function might fire if and only if the bigram “still hate” appears and the document’s sentiment is hypothesized to be negative. Maximum entropy as well as Naive Bayes makes no assumption about the relationship between words. It could perform better results when conditional assumptions did not meet in document [10].

3) Support Vector Machines: Support vector machine is very effective at traditional text classification and generally achieves better results than Naive Bayes. They search margins between two classes in contrast with Naive Bayes or Maximum entropy that use probabilistic approach. The main task of SVM is to find a maximum margin hyperplane, which is represented by vector. This method doesn’t separate one class from another, but typically, the separation is as large as possible. In work [13] Vapnik introduced algorithm support vector machine. He wanted to find the best possible surface to separate positive and negative samples in training corpus. The main task is to find a linear hyperplane, which separates data and where the margin is maximized. In figure 1 is illustrated problem where two classes are separated into two dimensions. There are many possible boundary lines to separate two classes. Each boundary has its margin. The reason of using SVM is that if we choose the one that maximizes the margin we prevent the wrong classification in the future.

4) Usage of machine-learning approaches: In many works researchers don’t use only one approach. They often use more of them and compare results between each of them. In work [12] they used SVM and NB to classify texts in Russian language. They used both types of NB classifier, then Bernoulli and multinomial model and compared all of those four methods. The best results achieved SVM, even with better results than both of Naive Bayes models (Bernoulli and multinomial model). For NB models multinomial modes obtained better results than Bernoulli model. In their work, they investigated the impact of different types of pre-processing and post-processing to final results. They applied bagging algorithm that had only a little positive effect to classification results. The same little positive influence had the usage of synonyms and lemmatization.

In work [14] they compared NB classifier with lexicon based approaches and approaches based on machine-learning (Maximum Entropy and Random Forests). NB reached the worst results for machine-learning approaches but comparable results such as lexicon approaches on their corpus. Maximum entropy achieved better results than both lexicon methods and Naive Bayes but worse results than Random forests

One part of work [15] is dedicated to comparison of different types of machine-learning approaches. They compared Naive Bayes, SVM and KNN on Arabic texts. In this work NB achieved the best results.

TABLE I
DIFFERENCE BETWEEN BERNOULLI MODEL AND MULTINOMIAL MODES

Model	Vector description
Bernoulli model	[0, 1, 0, 1, 1, 0, 1, 0, 0]
Multinomial model	[0, 2, 0, 3, 1, 0, 2, 0, 0]

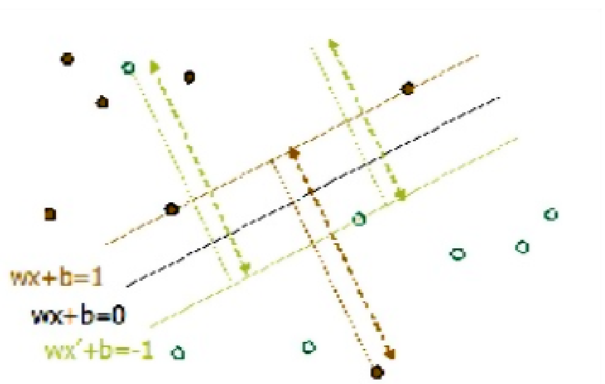


Fig. 1. Different boundary decisions are possible to separate two classes into two dimensions. Each boundary has an associated margin

B. Lexicon based approaches

Lexicon based approach uses opinion orientated words and phrases, which are saved in dictionaries. In sentiment lexicons, words are saved with their polarity and subjectivity. It helps to identity contextual polarity and subjectivity of text. Dictionaries can be created in many different ways.

First way, how to create a dictionary is to create it manually. When we create dictionary manually, we often use existing dictionaries. We can translate opinion dictionary from another language or merge several dictionaries into the big one. Manually created dictionary was used in work [5]. Dictionary in this work contains mainly adjectives, nouns, verbs and adverbs. These types of words have often emotional orientation. Advantage of this creation type is that all words in dictionary have opinion orientation and there is no need to check them after the dictionary creation. Disadvantage is that it takes a lot of time to create big dictionary for sentiment classification.

Another way is semi-automatically creation. This type of dictionary creation uses existing resources such as WordNet or SentiWordNet. SentiWordNet is a lexicon that was created for sentiment analyzing. It derived from WordNet by leveraging word relationships and word glosses. SentiWordNet contains set of opinion score for words with given meaning using semi-automated method where small bag of words are manually labeled. Ohama in his work [16] deals with the possibility to use SentiWordNet for opinion mining. They compared manually created dictionary with a method, which used the SentiWordNet. SentiWordNet obtained good results in comparison with manual lexicons.

Dictionary can be also created automatically via associations. Score for new word is calculated by using the frequency of the word’s proximity with respect to one or more seed words. Seed words are small group of words, which are strongly positive or negative. The association is usually calculated by Turney’s method for computing mutual information [5].

TABLE II
SIMPLE TEXT ORIENTATION BY NUMBER OF SENTIMENT WORDS

Polarity	Condition
Positive	$\sum pos > \sum neg$
Negative	$\sum pos < \sum neg$
Neutral	$\sum pos = \sum neg$

Advantages of automatically and semi-automatically dictionaries are that they can find many new words. For example, *unlistenable* was tagged as highly negative. This advantage was described in [17].

Basic lexicon approaches spot words with opinion in text and count them. Polarity of the input text is determined by number of different sentiment words in text as is shown in table II. $\sum pos$ and $\sum neg$ represent the number of opinion words in text. This type of analysis is presented in work [14].

Little modification of this approach includes strength of polarity for each word. We sum of strengths of polarities and in result we compare them. Other type of modification of this method includes identification of negation. In this approach we search words that change polarity of other words and in overall sentiment of whole sentence. Works [5], [6] include all of these modifications and add searching of words that change the strength of polarity (intensifiers). This approach is close to human processing of text. It works with different strength of polarity, identify negation (changing polarity) and intensification (increasing or decreasing strength of polarity).

IV. Future direction of work

In our work [18], [19], [20] we have created lexicon based approach in Slovak language. We have created manually annotated dictionary for opinion classification. This lexicon contains adjectives, verbs, nouns and adverbs, that are useful for classification. Every word in dictionary has assigned polarity (may be positive or negative) and strength of polarity from -3 (the most negative) to 3 (the most positive). We have also created algorithm, which processes these words from dictionary. Algorithm can use negation (switch) and intensification. It can work with intensify negation, negate intensification and process double negation and intensification. We use lemmatization in this algorithm. We make lemmas from any morphological form of adjectives. We remove prefixes and suffixes from word and change it with the same spelling.

In future work we want to extend our algorithm with lemmatization algorithm, which will process other word classes such as verbs and nouns. We want to add spam filter to algorithm and test if it will increase the quality of classification. We also want to work with negation and add shift negation, because switch negation is very inaccurate. We are working on method, which will combine lexicon approach and approach based on machine-learning.

References

[1] E. Cambria, B. Schuller, Y. Xia, and C. Havasi, "New avenues in opinion mining and sentiment analysis," *Intelligent Systems, IEEE*, vol. 28, no. 2, pp. 15–21, March 2013.

[2] H. D. Kim, K. Ganesan, P. Sondhi, and C. Zhai, "Comprehensive review of opinion summarization (survey)," 2011.

[3] K. Machová and L. Marhefka, "Opinion mining in conversational content within web discussions and commentaries," in *Availability, Reliability, and Security in Information Systems and HCI*, ser. Lecture Notes in Computer Science, A. Cuzzocrea, C. Kittl, D. Simos, E. Weippl, and L. Xu, Eds. Springer Berlin Heidelberg, 2013, vol. 8127, pp. 149–161. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-40511-2_11

[4] M. Marchand, R. Besançon, O. Mesnard, and A. Vilnat, "Domain adaptation for opinion mining: A study of multipolarity words," *JLCL*, vol. 29, no. 1, pp. 17–31, 2014. [Online]. Available: http://www.jlcl.org/2014_Heft1/2Marchand.pdf

[5] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede, "Lexicon-based methods for sentiment analysis," *Comput. Linguist.*, vol. 37, no. 2, pp. 267–307, Jun. 2011. [Online]. Available: http://dx.doi.org/10.1162/COLI_a_00049

[6] K. Machová and M. Krajč, "Klasifikácia názorov vo vláknových diskusiách na webe," *Znalosti 2011*, 2011.

[7] P. Koncz, *Aspektovo orientovaná analýza sentimentu*. Košice: Technická univerzita v Košiciach, 2012.

[8] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up?: Sentiment classification using machine learning techniques," in *Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing - Volume 10*, ser. EMNLP '02. Stroudsburg, PA, USA: Association for Computational Linguistics, 2002, pp. 79–86. [Online]. Available: http://dx.doi.org/10.3115/1118693.1118704

[9] F. Salvetti, C. Reichenbach, and S. Lewis, "Opinion polarity identification of movie reviews," in *Computing Attitude and Affect in Text: Theory and Applications*, ser. The Information Retrieval Series, J. Shanahan, Y. Qu, and J. Wiebe, Eds. Springer Netherlands, 2006, vol. 20, pp. 303–316. [Online]. Available: http://dx.doi.org/10.1007/1-4020-4102-0_23

[10] J. Khairnar and M. Kinikar, "Machine learning algorithms for opinion mining and sentiment classification," *International Journal of Scientific and Research Publications (IJSRP)*, vol. 3, 2013.

[11] C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*. New York, NY, USA: Cambridge University Press, 2008.

[12] N. Yussupova, D. Bogdanova, and M. Boyko, "Applying of sentiment analysis for texts in russian based on machine learning approach," in *IMMM 2012, The Second International Conference on Advances in Information Mining and Management*, 2012, pp. 8–14.

[13] V. N. Vapnik, *The Nature of Statistical Learning Theory*. New York, NY, USA: Springer-Verlag New York, Inc., 1995.

[14] B. Yuan, Y. Liu, H. Li, T. T. T. PHAN, G. Kausar, C. N. Sing-Bik, and W. Wahi, "Sentiment classification in chinese microblogs: Lexicon-based and learning-based approaches," *International Proceedings of Economics Development and Research (IPEDR)*, vol. 68, 2013.

[15] K. Khalifa and N. Omar, "A hybrid method using lexicon-based approach and naive bayes classifier for arabic opinion question answering," *J. Comput. Sci.*, vol. 10, 2014.

[16] B. Ohana and B. Tierney, "Opinion mining with sentiwordnet," *Knowledge Discovery Practices and Emerging Applications of Data Mining: Trends and New Domains: Trends and New Domains*, p. 266, 2010.

[17] L. Velikovich, S. Blair-Goldensohn, K. Hannan, and R. McDonald, "The viability of web-derived polarity lexicons," in *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, ser. HLT '10. Stroudsburg, PA, USA: Association for Computational Linguistics, 2010, pp. 777–785. [Online]. Available: http://dl.acm.org/citation.cfm?id=1857999.1858118

[18] M. Mikula and K. Machová, "Klasifikácia názorov dolovaných v konverzacnom obsahu," in *Electrical Engineering and Informatics V: Proceedings of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice*. Košice: FEI TU Košice, 2014, pp. 663–667.

[19] M. Mikula and K. Machová, "Klasifikácia názorov v konverzačnom obsahu," in *WIKT 2014 9th Workshop on Intelligent and Knowledge Oriented Technologies: Proceedings in Informatics and Information Technologies*. Bratislava: Nakladateľstvo STU Bratislava, 2014, pp. 3–8.

[20] M. Mikula and K. Machová, "Classification of opinion in conversational content," in *IEEE 13th International Symposium on Applied Machine Intelligence and Informatics: Proceedings*. Danvers: IEEE, 2015, pp. 227–231.

Orientation methods in space for solving anomalous behavior

[1]*Jaroslav LÁMER (3rd year)*
Supervisor: [2]František JAKAB

[1,2]Dept. of Computers and Informatics, FEI TU of Košice, Slovak Republic

[1]jaroslav.lamer@tuke.sk, [2]frantisek.jakab@tuke.sk

***Abstract*— Many actual researchers in computer vision are confronted with the main problem - solving anomalous behavior in bigger space. The main objective is therefore to determine the object position in workplace. We describing in this article the solution for orientation in space based on basic methods, inspired by human orientation principles. We have discovered in our previous research three basic methods. Based on them we have designed and evaluated two functional prototypes.**

***Keywords*—Orientation in space, optical deformations, sift, surf, optical flow, camera, OpenCV, computer vision.**

I. Introduction

Many projects and devices are facing troubles with insufficient sensors precision. Partial solution (enhancement) can be human-like camera sensors. These enable detection of objects, measurement of distances, positions and movements in space. In this way we would give human "eyes" to artificial intelligence systems, e.g. project presented in [1] is using camera in combat vehicle to determine persons in outdoor area, or project described in [2] is using camera to determine whether workspace robot is in. There are many other use-cases, such as orbital probes, personal vehicles, tactical forces control, work-flow management etc.

We carried out a research focused on human orientation in space; this was the fundamental requirement for later research.

II. Previous Research

We were inspired by human orientation senses. The experiment made included 10 respondents who had a ratio camera mounted on the head for documentation purposes. Each respondent is given one task and a tutorial with graphical map to solve it. Task consists from locating 5 different symbols (numbers from 1-5) in space and to mark them as "visited" in required order. Each marked symbol has indicated vector to next symbol depicted by arrow in given map. Once this was done, respondents must repeat the workflow without instructions and in correct order.

The results were divided into three groups, the first group consists of 3 respondents who do not use their own "virtual mind map" for task solving. After each marker they go back, look at the whole workspace, and locate the needed symbol (number). We refer to this principle as absolute positioning. The second group consists of 2 respondents. In order to locate the symbols they used mind-map using a kind of movement history and calculated actual position in work-space. We call this method relative positioning. The last respondents group consists of 5 members. Each of them used a hybrid method consisting from absolute and relative positioning. Hence we refer to it as hybrid positioning.

Based on Human inspired in-space orientation research we needed a big amount of state of the art algorithms which are perfectly optimized. OpenCV includes majority of required graphical algorithms like perspective transformation, optical pattern recognition, optical object tracking or optical flow determination [3].

III. Testing the Positioning Principles in Computer Vision

We have implemented and tested the first two researched methods of human orientation on single laptop computer with single camera input.

A. Absolute orientation principle

First, there was a requirement to evaluate the real image processing speed for each kind of in space-orientation, this is the key property for real usage. We have implemented first two human in-space orientation principles, as explained above.

For case of the absolute orientation we tested next three feature extraction methods: SIFT, SURF, ORB. These are based on feature extraction from pattern and input image, having the optical deformation invariant as their main advantage [4]. The testing machine handled 720p HD input video on Intel Celeron M 1.9 GHz with 1.5 GB memory (minimum requirements are specified as machine, which computes at least as 10 fps). Real-time processing speed using SURF method was rather poor. The best results reached was one frame per 0.7 second. Sift and ORB methods had better result however still not satisfactory. The first located required patterns in 0.5-0.6 second. Our results are in agreement with the research as presented in [5], where ORB has the best performance between tested methods – for our measurements: 0.4 seconds. The last tested method was hard-pattern recognition. The given pattern is recursively searched in input image [3]. Whet pattern matches, the position in image is returned. The main disadvantage is sensitivity to optical deformations [3]. Method has the poorest results from all tested algorithms – requires 1.2 second to evaluate one video-frame. These results are not applicable for real-time solutions.

Due to this we moved to relative orientation principle.

B. Relative orientation principle

Optical flow algorithm was used as computer vision tool for relative orientation. This algorithm is based on pixel neighborhood and pixel movements [6]. Each pixel has its own history of movements and its actual state (position). This exactly defines the actual position in space and perfectly matches our relative orientation principle.

The experiments having the same conditions as in absolute orientation principle evaluation and give very good results. The worst case of processing 60 random points in 720p HD video images gives up to 25 frames per second.

Based on these results we have designed the complex model for relative position estimation. The model structure is illustrated on next figure.

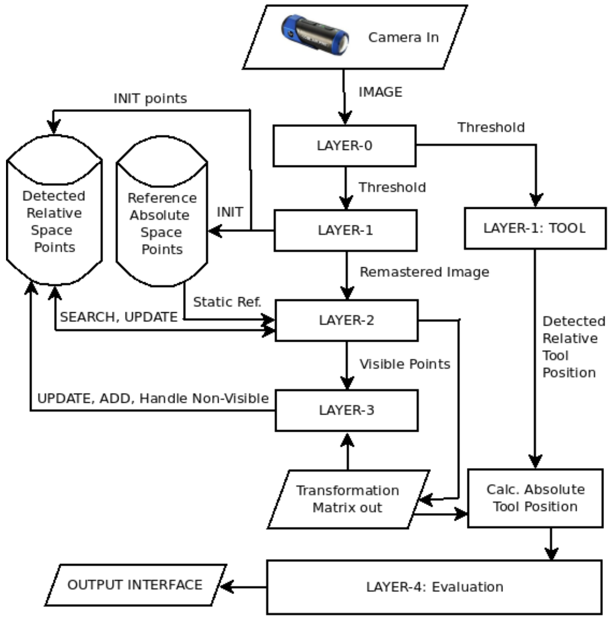


Fig. 1. Our new five layer model for relative orientation in space based on single video stream input in known environment.

Layer0: (HW and SW filter) this part of model consists from few filters for removal of unusable objects, colors and contours from input camera image.

Layer1: (Contours finder) this layer provides a contour finding functionalities. Sorts all the objects into classes depending on shapes, contours area and so on.

Layer2: (Optical Flow Core) is the core part of model and provides optical flow features. Optical flow algorithm here makes the key-role for recognition of all detected points.

Layer3: (Add on the flow and map non visible) When the point location process is carried out the new coordinates of visible markers are computed with this layer.

Layer1-TOOL: (Working tool, other objects positioning) This layer has the same functionality as Layer1, but it is used to determine the position of other objects, such as working tools.

Layer4: (Evaluation of tool position) this part is the last and manages absolute tool position. Here can be evaluated standard and non-standard situations. Layer is directly connected with outside systems through Output Interface.

We reached the best image results using camera with fisheye lenses capturing 60 frames per second. Evaluated was each n-th frame (real 4 - 15 fps). This principle minimizes camera image blurring and preserve sufficient frame rate.

All measurements are made on un-calibrated camera to show the worst case scenario. The next experiment is based on comparison of two points in static positions and their calculated virtual positions. In ideal case these two points have still the same distance to each other. As can be seen on Fig.2. the optical distortion has a high impact on computational precision. The point distance error was calculated using Eq. 1

$$err = (realDist/virtDist) * \sqrt{(x1-x2)^2 + (y1-y2)^2} \quad (1)$$

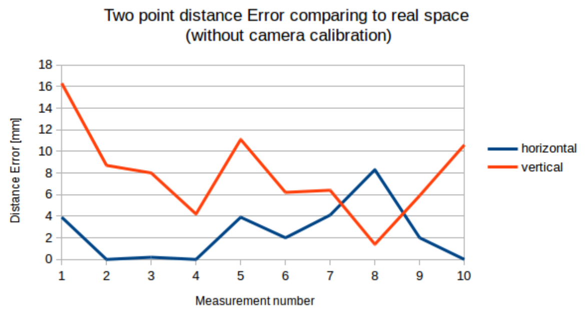


Fig. 2. Two point distance error calculated using Eq.1 for both image axis (horizontal, vertical). The main source of error is the camera lens distortion effect.

IV. Conclusion

We have designed and tested two main methods for orientation in space based on human orientation principles. Presented tests are illustrations of worst case, which can occur – this includes un-calibrated video camera with visible lens distortions. Relative orientation principle model outperforms the absolute principle orientation model in frame-rate and confidence of position estimation. The main disadvantage of relative method – loss of orientation can be solved in future using Hybrid orientation principle (combining absolute and relative principle together).

Acknowledgment

This publication is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF. We support research activities in Slovakia/This project is being co-financed by the European Union.

References

[1] SALGIAN, Garbis, et al. Combat ID-Combat Identification System. ARMY TANK AUTOMOTIVE RESEARCH DEVELOPMENT AND ENGINEERING CENTER WARREN MI, 2012.

[2] ZHANG, Dong Ping, et al. The Vision Orientation System Based on OpenCV of the Glass Curtain Wall Cleaning Robot. Applied Mechanics and Materials, 2013, 336: 1063-1067.

[3] BRADSKI, Gary; KAEHLER, Adrian. Learning OpenCV: Computer vision with the OpenCV library. O'reilly, 2008..

[4] PANG, Yanwei, et al. Fully affine invariant SURF for image matching. Neurocomputing, 2012, 85: 6-10.

[5] RUBLEE, Ethan, et al. ORB: an efficient alternative to SIFT or SURF. In: Computer Vision (ICCV), 2011 IEEE International Conference on. IEEE, 2011. p. 2564-2571.

[6] BROX, Thomas. Optical Flow. Computer Vision: A Reference Guide, 2014, 565-569.

Overview of Recent Progress in Speaker Adaptation for Slovak HMM-based Speech Synthesis

[1]*Martin SULÍR (3rd year)*
Supervisor: [2]Jozef JUHÁR

[1,2]Dept. of Electronics and Multimedia Communications, FEI TU of Košice, Slovak Republic

[1]martin.sulir@tuke.sk, [2]jozef.juhar@tuke.sk

***Abstract*— This paper presents an overview of recent progress in the speaker adaptation for Slovak statistical parametric speech synthesis based on the hidden Markov models together with a description of newly implemented systems with the help of this technique. Several new Slovak voices were trained with the help of prerecorded small databases together with the use of newly created large average voice model based on automatic speech recognition system's acoustic data. The results obtained so far show a high potential of this technique in case of Slovak language, because we manage to get a new intelligible and relatively natural sounding artificial speech with the small amount of data.**

***Keywords*—hidden Markov model; model adaptation; Slovak language; statistical parametric speech synthesis; text-to-speech**

I. INTRODUCTION

Speech synthesis is one of the most important part of the speech interaction with the computer. The main task of these systems is primarily making life easier for example to people with physical disabilities, such as blind people, or to totally ordinary people and facilitating their day to day operations [1]. Nowadays, the hidden Markov model based speech synthesis method represents one of the most progressive approach how to convert written text into sound, which ultimately sounds like human speech [2]. The progressiveness of this method is particularly involved in its high flexibility, where it allows to quite easily create the new voices with the help of adaptation, interpolation or, for example, using the technique of eigenvoice. The utilization of these techniques arise from using of hidden Markov models (HMM) which can be properly mathematically modified in order to obtain their desired modified versions. Models adaptation provides a relatively large range for customizing of models, where it uses a principle of using small adaptation database and large pre-trained average voice model to create a new voice which corresponds to the speaker who record the input adaptation database. In this case, the adapted voice quality is influenced by two major factors: the adaptation database size and its quality and the average voice model diversity. The main motivation for the solving of the research task is implementation and development of new Text-To-Speech (TTS) systems for Slovak language and thus advance not only the area of speech synthesis as such but also human-machine interaction in Slovak language.

II. AVERAGE VOICE-BASED SPEECH SYNTHESIS SYSTEM

The basic principle of HMM-based speech synthesis method is to use the context-dependent HMM models, which are trained from speech database, as generative models for speech synthesis process. Average Voice-based Speech Synthesis system (AVSS) can be divided into three basic parts [3]. The first one is the training part which consists of average models training. Its main task is the extraction of spectral and excitation parameters from speech database as well as the implementation of HMMs training. In the case of average models training, the speech database consists of multiple sub-databases, where each of them were recorded by one certain speaker. The second part of AVSS system is the adaptation step. Target speaker speech database together with its spectral and excitation parameters represent inputs to the adaptation process. The process of adaptation of average voice models is implemented in this part, where a lot of modifications of basic adaptation methods have been developed [4]. The synthesis part of HMM-based speech synthesis system consists of text analyzer, which convert given text into contextual label sequence and several blocks which are responsible for parameter generation from context dependent HMMs and duration models as well as excitation generation based on generated excitation parameters is used. Research in this area is aimed to the point, when it will be possible to use these voices in various spheres of life, without that they were somewhat limited and acts unnatural.

III. SPEAKER ADAPTATION IN SLOVAK HMM-BASED SPEECH SYNTHESIS

Experiments with speaker adaptation in Slovak HMM-based speech synthesis system arose from the work on the speaker dependent voices which are based on the principle of using one large input speech database which directly represents the parameters of output voice. Recently, the performance of the newly created Slovak speaker dependent HMM-based systems have been evaluated through the objective and subjective listening tests [5]. The evaluation results showed that each of these systems provide very good intelligibility of artificial speech at the output. Naturalness of synthetic speech achieves maximum possible quality, which is however limited by the current state of the used technology.

However, these results are encouraging since this is one of the first HMM speech synthesis system built for the Slovak language and many improvements are possible. The obtained results are also underlined by the fact, that the newly created Slovak speech synthesis systems have also been used in some practical applications and implementations. Firstly, they were adopted into multimodal interface for controlling functions of the modular robotic system, which can be used in difficult conditions such as rescue works, natural disasters, fires or decontamination [6]. These voices were also used as part of new version of *ZureTTS* system which is an initiative of Aholab Signal Processing Laboratory of University of the Basque Country to provide a personalized speech synthesizer to people with speech impairments and also to those who have completely lost their voice. The new version of *ZureTTS* system was undertaken by an international team of researchers during *eNTERFACE'14 ISCA Training School*, covering up to 8 languages: English, Spanish, Basque, Catalan, Galician, Chinese, German and also Slovak [7].

Speaker adaptation represents a technique which is directly related to speaker dependent HMM-based speech synthesis. Together in first experiments with this technique for Slovak language, two new Slovak systems were trained and evaluated [8]. These systems use previously developed modules for Slovak text analysis together with the proposed language dependent context clustering and Constrained Maximum Likelihood Linear Regression (CMLLR) combined with additional Maximum A Posteriori (MAP) adaptation were used for these experiments. Speech sub-databases that were necessary for average voice HMMs training arose from the acoustic data which are used for the speech recognition system in Slovak language. These acoustic data were algorithmically divided into isolated sentences and individual speakers. Thus formed average voice speech database include the sentences of seven male and seven female speakers, where each of them contain more than 450 sentences. Two small single speaker speech databases have been carefully recorded under studio conditions for the purposes of speaker adaptation for Slovak HMM-based speech synthesis, where each of them consist of 330 phonetically balanced sentences. The evaluation of newly created voices was performed by the subjective listening tests and the obtained results showed the potential of new voice creation with this technology for Slovak language. In many cases, they are comparable with the speaker dependent systems and this highlights their potential.

In the recent past, the influence of adaptation database size on the quality of HMM-based synthetic voice based on the large average voice model was also being solved [9]. In this case, the influence of various database size on the overall quality of HMM-based synthetic voices which were built with the help of large average voice model using average voice-based speech synthesis system together with AHO-coder vocoding technique was studied. Together, eight new voices of one male speaker were built by gradually adding new data into adaptation database while the quality of individual new voices were evaluated with the help of objective evaluation methods. The aim was to show the effect of adding more data into adaptation database to the overall quality as well as to define the threshold when the impact of the newly added data on the voice quality will be negligible in case of using a sufficiently large average voice model in adaptation procedure. The result of this work is a set of spectrum and fundamental frequency assessments which directly show the dependence of the voice quality on the amount of adaptation data together with the threshold definition. The above-mentioned tests showed that when large diverse average model is used it is possible to get quite good synthetic voice with the help of adaptation techniques with a small amount of data. The experiments also showed the existence of threshold beyond which the increasing size of the database cause only the quality oscillation. Acquired results underline the effectiveness of the model adaptation techniques as such especially when using the AHO-coder which represents one of the most advanced vocoder.

IV. Future Research

Future work will focus on the processing of the larger amount of speech data from the recognition system's database for average voice and on the gender dependent average voice models training. Another area of concern will be the adoption of a more sophisticated model for speech reconstruction and the inclusion of more prosodic properties in order to increase the naturalness of the produced speech. The text analysis module improvement will also form part of future work. Improvement of this module thus allows us to use these speech synthesis systems in different applications.

Acknowledgment

This publication is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF.

References

[1] P. Taylor, Text-to-Speech Synthesis. University of Cambridge, Cambridge: Cambridge University Press, 2009. p. 597. ISBN 978-0-521-89927-7

[2] K. Tokuda, Y. Nankaku, T. Toda, H. Zen, Y. Yamagishi, K. Oura, "Speech Synthesis based on Hidden Markov Models". In: Proc. IEEE, vol. 101, issue 5. 2013, pp. 1234-1252. ISNN 0018-9219

[3] J. Yamagishi, T. Kobayashi, "Average-Voice-Based Speech Synthesis Using HSMM-Based Speaker Adaptation and Adaptive Training". In: Transactions on Information and Systems archive, vol. E90-D, no. 2. 2007, pp. 533-543.

[4] J. Yamagishi, T. Kobayashi, Y. Nakano, K. Ogata, J. Isogai, "Analysis of speaker adaptation algorithms for HMM-based speech synthesis and a constrained SMAPLR adaptation algorithm". In: Transactions on Audio, Speech, and Language Processing 17. 2009, pp. 66-83.

[5] M. Sulír, J. Juhár, "Comparison of three vocoding techniques in subjective and objective evaluation tests of Slovak HMM-based TTS systems". 2015. (in review process)

[6] S. Ondáš, M. Pleva, M. Lojka, M. Sulír, J. Juhár, "Server-based speech technologies for mobile robotic applications". In: Journal of Electrical and Electronics Engineering. Vol. 6, no. 1, 2013, pp. 95-98. ISSN 1844-6035

[7] D. Erro, I. Hernaez, E. Navas, A. Alonso, H. Arzelus, I. Jauk, NQ Hy, C. Magarinos, R. Perez-Ramón, M. Sulír, X. Tian, X. Wang, J. Ye, "ZureTTS: Online Platform for obtaining Personalized Synthetic Voices". 2015. (in review process)

[8] M. Sulír, J. Juhár, "Speaker Adaptation for Slovak Statistical Parametric Speech Synthesis Based on Hidden Markov Models". In: Proceedings of Radioelektronika 2015. 2015. (will be published)

[9] M. Sulír, J. Juhár, "The Influence of Adaptation Database Size on the Quality of HMM-based Synthetic Voice Based on the Large Average Voice Model". 2015. (in review process)

Patterns for Structural Evolution of Context-free Grammars

[1]*Ivan HALUPKA (4th year),*
Supervisor: [2]*Ján KOLLÁR*

[1,2]Dept. of Computers and Informatics, FEI TU of Košice, Slovak Republic

[1]ivan.halupka@tuke.sk, [2]jan.kollar@tuke.sk

***Abstract*—Grammar evolution has been the topic of scientific interest for more than two decades. Despite of this, current state-of-the-practice in the domain of grammar evolution is characterized by a lack of comprehensive foundations and a lack of best practices. We addressed this issue by proposing two approaches to grammar evolution, namely mARTINICA for specification of evolutionary procedures on a high-level of abstraction, and pLERO for specification of evolutionary transformations on a low-level of abstraction. This paper presents summary of our research in the field, while it emphasizes most recently conducted research considering patterns of grammar evolution.**

***Keywords*—Grammar evolution, pLERO language, structural paterns, transformation operators**

I. Introduction

Grammar evolution is a process of changing a formal grammar driven by varying requirements related to grammar's purpose. Despite ongoing scientific research that spans over more than two decades, there is currently no well-established systematic approach to grammar evolution. In current practice, most grammars evolve in an ad-hoc manner [1], which makes execution of evolutionary procedures both difficult and error-prone.

Regarding an impact of an evolutionary transformation on a language that a grammar generates it is commonly distinguished between three types of evolutionary activities i.e. *grammar extension* increasing a set of sentences generated by a grammar, *grammar restriction* decreasing a set of sentences generated by a grammar, and *grammar refactoring* that has no impact on a generated language [2]. In our research, we mostly focus on refactoring transformations. These transformations play an essential role in adaptation of grammars recovered from various software artifacts, such as user interfaces [3].

Our research is strongly influenced by an evolutionary approach proposed by Ralf Lämmel [1] and Vadim Zaytsev [4]. They propose an approach in which an evolutionary procedure is specified as a sequence of instances of predefined transformation operators. In order to execute an evolutionary procedure, instances of transformation operators must be applied on a grammar in exact order of their specification, while each operator instance prescribes a certain precondition that must hold in order for a transformation to occur, and a certain postcondition that must hold in order for a transformation to be considered valid [1], [4]. Strong advantages of such approach in comparison with an ad-hoc approach to grammar evolution are traceability and predictability of evolutionary changes. However, this approach has three significant drawbacks:

- Reusability, since evolutionary scripts need to be adapted to be applicable on a different grammar that originally intended.
- Comprehensibility, since it is difficult to understand a non-trivial relation between individual operator instances and an overall purpose of an evolutionary procedure.
- Universality, since there is currently no suite of transformation operators applicable for specification of all evolutionary procedures [2].

We addressed the first two drawbacks by proposing an evolutionary approach in which evolutionary procedures are specified on a high-level of abstraction using grammar metrics, while sequences of instances of transformation operators are inferred in an automated fashion using evolutionary algorithm called *mARTINICA* (metrics Automated Refactoring Task-driven INcremental syntactIC Algorithm) [5]. Proposal of this approach was strongly influenced by [6], [7]. We further addressed third drawback by proposing a domain-specific language for specification of evolutionary transformations on a low-level of abstraction, meaning level of individual symbols contained within productions. We refer to this language as to *pLERO* (pattern Language of Extended Refactoring Operators) [8].

Both mARTINICA and pLERO are intended for performing of evolutionary transformations of context-free grammars whose productions are expressed in BNF notation.

II. pLERO language

Core specification entity of pLERO language is a *pattern*, which denotes a special kind of transformation operator that lacks a postcondition, and whose precondition is only implicitly specified and evaluated in a process of grammar transformation.

A pattern consists of a set of transformation rules and a set of declarations. Each transformation rule specifies a way in which productions that exhibit some structure should be transformed, while each declaration specifies certain property of a grammatical structure that is used in a process of grammar transformation. Each transformation rule consists of two parts, namely a predicate specifying a structure that a production must exhibit in order to be transformed by a transformation rule, and a transformation that specifies a way in which a structure described by the predicate should be transformed. Both predicate and transformation are specified using formalism of meta-production rules, where each meta-production rule specifies a structure of some class of productions, and it is possible to match productions against this structure. Moreover, it is possible to transform a production so that it exhibits structure prescribed by a meta-production, but only in case, when an implicit precondition specified by a predicate holds. Detailed description of pLERO language and a process of grammar transformation using patters can be found in [8].

In the last year, our research effort was focused on the following three aspects of pLERO language:

- Extension of expressive power of pLERO language [9]
- Comprehensive documentation of grammar patterns [10]
- Study of applicability of pLERO language [9], [10]

In [9] we have proposed extension of pLERO language that addresses the following aspects of a pattern's specification:

- Parameterization of patterns, since the originally published [8] specification of pLERO only included support for expressing parameterless grammar transformations.
- Matching of negative grammatical structures, meaning expressing preconditions that a grammar should not fulfill in order to be transformable by a pattern.
- Equivalence precondition for grammatical structures whose properties are expressed at multiple levels of abstraction.
- Iteration over structurally different productions in the context of a single transformation rule.

In [10] we have proposed a template for documentation of patterns consisting of informal description of an evolutionary problem addressed by a pattern, informal description of solution strategy specified by a pattern, description of nontrivial consequences of a pattern's application, formal specification of a pattern and an exemplary application of a pattern on a context-free grammar. Main goal of proposed approach to documentation of patterns is to facilitate comprehension and reuse of grammar patterns.

In [10] we have demonstrated the applicability of pLERO by creating pattern specifications of chosen transformation operators proposed in [1], [5]. In [9] we have conducted a case study in which we have created pLERO specifications of two transformation operators called unfold and pack, and we further compared these specifications to their implementations written in Java language. In practice, both of these operators are used in various evolutionary tasks. For instance, unfold operator is commonly used in restructuring of inferred grammars and in the process of grammar convergence. Pack is often used for reducing a length of grammars productions with the aim of improving grammar comprehension. In the case of unfold, we have found that specification of this operator in Java requires 37 language statements, while same specification in pLERO requires only 3 language statements. In the case of pack, we have found that specification of this operator in Java requires 18 language statements, while same specification on pLERO requires 4 language statements. These results present strong evidence, that an expressive power of pLERO in the domain of grammar transformation is significantly higher than the expressive power of general-purpose languages. This means that the development of pLERO language is justified in the terms of a tradeoff between language's generality and language's expressiveness.

III. Conclusion

In mARTINICA evolutionary procedures are specified using grammar metrics, which limits applicability of this approach exclusively to refactoring transformations. Reason for this is the weak-binding between a language generated by a grammar and grammar metrics. This limitation had a significant impact on a design of pLERO language, which was initially intended as a complementary tool for specification of transformation operators to be used in mARTINICA algorithm. In a recent time however, we have recognized the potential of pLERO to be used as a stand-alone tool of grammar transformation. Because of this, in future we would like to focus on extending the expressive power of pLERO in order to comprise evolutionary transformations that go beyond the scope of refactoring. This means, that we aim at extending pLERO of constructs that enable the introduction of new terminal symbols, and manipulation of the existing ones. Moreover, we would like to focus on a constructs that support composition of patterns, since currently pLERO does not support creation of complex evolutionary procedures from atomic ones. We believe, that solving these two research challenges, would serve as an appropriate basis for creation of complete theory of grammar evolution, which promises an increase in productivity of grammar development by a reuse of existing well-proven design knowledge.

We have been awarded the best research poster award at Poster 2012 conference, best design award at 2012 competition for best PhD work at Technical University of Košice, and best presentation award at SLATE 2013 conference.

Acknowledgment

This work was supported by project VEGA 1/0341/13 *Principles and methods of automated abstraction of computer languages and software development based on the semantic enrichment caused by communication.*

References

[1] R. Lämmel, "Grammar adaptation," in *FME 2001: Formal Methods for Increasing Software Productivity*, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2001, vol. 2021, pp. 550–570.

[2] P. Klint, R. Lämmel, and C. Verhoef, "Toward an engineering discipline for grammarware," *ACM Transactions on Software Engineering and Methodology*, vol. 14, no. 3, pp. 331–380, 2005.

[3] M. Bačíková, J. Porubän, and D. Lakatoš, "Defining Domain Language of Graphical User Interfaces," in *Slate 2013: 2nd Symposium on Languages, Applications and Technologies*, vol. 29. Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2013, pp. 187–202.

[4] V. Zaytsev, "Negotiated Grammar Evolution," *Special issue on Extreme Modeling of The Journal of Object Technology (JOT)*, vol. 13, no. 3, pp. 1:1–22, 2014.

[5] I. Halupka, J. Kollár, and E. Pietriková, "A task-driven grammar refactoring algorithm," *Acta Polytechnica*, vol. 52, no. 5, pp. 51–57, 2012.

[6] S. Chodarev, "Development of domain-specific languages based on generic syntax and functional composition," *Information Sciences and Technologies Bulletin of the ACM Slovakia*, vol. 4, no. 3, pp. 47–53, 2012.

[7] J. Kollár and E. Pietriková, "Genetic evolution of programs," *Central European Journal of Computer Science*, vol. 4, no. 3, pp. 160–170, 2014.

[8] J. Kollar, I. Halupka, S. Chodarev, and E. Pietrikova, "plero: Language for grammar refactoring patterns," in *Federated Conference on Computer Science and Information Systems (FedCSIS 2013)*, 2013, pp. 1503–1510.

[9] I. Halupka, "Dsl for grammar refactoring patterns," in *Central European Functional Programming School, 5th Summer School, CEFP 2013, Cluj-Napoca, Revised Selected Papers*, ser. Lecture Notes in Computer Science, 2015, vol. 8606, in print.

[10] I. Halupka and J. Kollár, "Catalog of grammar refactoring patterns," *Central European Journal of Computer Science*, vol. 4, no. 4, pp. 231–241, 2014.

Principles of shape recognition

[1]Milan SPIŠIAK (1st year),
Supervisor: [2]Ján KOLLÁR

[1,2]Dept. of Computers and Informatics, FEI TU of Košice, Slovak Republic

[1]milan.spisiak@tuke.sk, [2]jan.kollar@tuke.sk

***Abstract*—This paper is focus on shape recognition. We would describe varying type of approaches of shape recognition. Because we want to find new possibles of the shape recognition. We want to focus on properties and possibilities of these approaches. Extraction of these informations could provide a new base of system of recognition. Within our research we want to design a new grammar that provides new possibilities of the shape recognition. Probably it would be some type of shape grammar (now, this is only ours postulate).**

***Keywords*—recognition, grammars, shape recognition**

I. Introduction

Recently, many works deal with idea how to extract information from digital content such as: images, sounds, data of difference sensors, etc [1], [2]. Some works for this goal use:

- multi-agent interactions and complex operating procedures [3], [4]
- rule induction and multithread parsing [1]
- multi-resolution deformable part models [5]
- Self-Growing and Self-Organized Neural Gas (SGONG) network [6]
- Convolutional Neural Networks [7], [8]
- RBF Neural Networks [9]
- Context-Free Grammar [3]
- Stochastic Context-Free Grammar [10], [11], [12]
- Shape grammar meta-language [13], [14]

These methods are only drop in the ocean of possibilities. The goal is find functions that would allow computer recognition to get closer to recognition of human [10], [11], [8], [5], [15]. Just get closer to human recognition could allows using these functions in many sectors [12], [16], [17], [18], [19]. Varying types of recognition using difference type of approaches for achieved of this goal such as: approach based on neural networks using properties of human neurons. This approach has relative good results of recognition. But this approach is more like emulation of human recognition principles than the method that is design at computer environments.

Next approach could be grammars. This approach is closer to recently computer environments as previous approach. Although this approach is closer than neural networks but is not ideal too. Combination of both principles could allow a new level of shape recognition.

Goal of this paper is a review of possibilities and properties of currently shape recognition approaches.

II. Shape grammar meta-language

Shape grammar is grammar that manipulates with shapes to generate design. This grammar uses if-then rules of schemata to preforms changes in a design. Grammar shows the design process with using of rules set. This type of grammar is based on process of rule application in shape grammars. This process is composed of four phases [13], [14]:

1) rule selection
2) drawing state
3) matching conditions
4) application method

Matching conditions phases can be expanded to:

- parameter requirements
- transformation requirements
- contextual requirements

And this expansion expands process of rule application,too. The new version of rule application process is:

1) rule selection
2) drawing state
3) parameter requirements
4) transformation requirements
5) contextual requirements
6) application method

A. Rule selection

Rule selection of rule application process determinates which rule will be applied and which rule we can use anywhere in grammar. This part has two descriptors: Directive and Rule-set. Rule-set provides availability of rule set at the grammar derivation. This goal is achieved with using a parallel description. Parallel description includes rules which are called rule-set. Rule-set has a three control options: set-rule, add-rule, sub-rule. These options provide some operation above of rule-set, such as: add addition rules, remove rules. This descriptor allows the use of same rule again in varying grammar parts. Second rule descriptor is Directive. This descriptor determines what happens when current rule is successfully or unsuccessfully applied. Directive descriptor has two rules. One rule for each option. If rule was applied as successful, directive descriptor calls a rule for this scenario. If not, it was calls the rule for unsuccessful scenario. Directive descriptor allows to create a macro. Macro is composed of a one primary rule and one or more secondary rules [13], [14].

B. Drawing state

We need to use whole drawing area, but sometimes we need to apply rules only to specific drawing parts. Drawing state phase determines areas which we could use for rule application. In this phase we have two descriptors. First descriptor is label-filter. This descriptor provides filter out unnecessary sub-shapes in the drawing (each sub-shape which does not have the same labels as labels used in the rule schema is filter out). This descriptor we could use to filter out background information. Second descriptor is focus. Focus descriptor determines area or areas which we can use to find a sub-shape in drawing. This area is demarcated by enclosed polygons. These polygons are composed of labeled lines called: "focus". These lines we can alter or erase. If drawing has enclosed polygons, then all lines outside of this area are temporarily deleted. Subsequently we can apply rules inside demarcated area. If we want to apply rules to outside of the area, we must remove all focus lines as first[13], [14].

C. Parameters requirements

Parameter requirements phase corresponds with function g(). This phase provides values determination which are assign to parameters of the schema. This phase has not any descriptors. We can put restrictions to parameters. These restrictions influence looking for desired sub-shapes in the drawing [13], [14].

D. Transformation requirements

Transformation requirements phase determines transformations that are necessary to successful sub-shape match in the drawing. Just like phase of parameters requirements, this phase does not have any descriptors. The transformations requirements contains rotations, translations, reflections and scaling [13], [14].

E. Contextual requirements

Fifth phase of rule application process is contextual requirements. Usually matching conditions between drawing and schema are determined by the transformation function: t() and parameter function: g(). But shape grammar meta-language adds addition phase: contextual requirements. Phase of contextual requirements has descriptors that add additional constrains of matching. These constrains determinate which sub-shape of drawing are based on context of the sub-shape in the drawing. Shape grammars have one problem: matching conditions between drawing and schema work only on found shapes. We could define a conditions with use a set of compound rules. These sets of compound rules combine shapes and grammar of parallel description. Grammar of parallel description provides detection of any void shapes in the rectangle. Also we could use another techniques which are based on adding additional properties. Or we could use phase of contextual requirements. This phase uses visual properties of the sub-shape. This phase has two descriptors: zone and maxline. Maxline descriptors determines which line of the subshape would be a maximal line. Therefore maxline adds an additional constraint for purpose of finding maximal line in the sub-shape. Maximal line of the shape can not be smaller part of another line. Maxline descriptor we could use to determine whether is schema determinate or not. Schema is deterministic if has a finite number of sub-shapes. If schema has infinite number of sub-shapes or it is composed of one line then it is an non deterministic. Another approach of finding maximal line gets labeled points to endpoint of a line. If we want to find a maximal line in sub-shape, we must only look for a line with two endpoint labels. But this solution creates one problem: visual property is replaced with the labeled points, and this fact could cause that we choose line which has two labeled points but it is a part of larger line. Maxline descriptor fixes this problem with use of the visual information in the drawing. Second descriptor of contextual requirements is: Zone. Zone descriptor provides add an addition constraint. This constrain influences matching condition of a schema. Constrain is in the form of participate function. We could implement this descriptor as sequence of compound rules with a parallel description grammar. Example of forms these rules [13], [14]:

$$\text{A, } 0 \rightarrow \text{A', f(C,A,Z)}$$
$$\text{A', } 1 \rightarrow \text{B, } 0$$
$$\text{A', } 0 \rightarrow \text{A, } 0$$

Where:
C - is a drawing
A - is a schema
Z - is a zone

These computed rules have two parts. First part is a rule of shape and second part is a rule of parallel description grammar. First rule finds a shape A with use a parallel description of 0. Subsequently this shape is transformed to the shape A'. This transformation uses a result of the function f(). If it is a shape satisfactory then functions return 1 otherwise returns 0. If it has returned value 1 then rule two transform design. Otherwise third rule returns the drawing to original state [13], [14].

F. Application method

Last phase of rule application process is application method phase. This phase provides process of applying sub-shapes to drawing. This phase has one descriptor called: apply-mode. Apply-mode descriptor provides this process of applying sub-shapes to drawing. For this aim apply-mode descriptor has three control options: single, random and parallel. Single option allows to select one sub-shape for application. Second options is random. This option allows randomly select a sub-shape for application. Random selection can be provided with using user or conjunction a shape grammar interpreter and a computer randomly select of sub-shape. Last option is parallel. Parallel options applies the rule to all sub-shapes in the drawing [13], [14].

III. Artificial Neural Networks

An Artificial Neural Networks (ANNs) are based on mathematical model. This mathematical model was created to emulate a biological neural networks functions (for instance: a human neural networks). Basic unit of artificial neural networks is artificial neuron. Artificial neuron is mathematical model too. This model has three sets of rules [20], [21]:

- multiplication
- summation
- activation

Inputs of artificial neurons are weighed. This fact causes that each value of input is multiplied with weight. Each input value has another weight. In the second step, these results of multiplication and a bias are summed together. Result of sum is inserted to activation function (it is called transfer function too) [20], [21].

A. Artificial neuron

Design and functionalities of artificial neurons are derived from observation of the biological neurons. Functionalities of artificial neurons we can describe as mathematical functions. Artificial neuron is described as function [20], [21]:

$$\text{y(k) = F}\left(\sum_{i=0}^{m} w_i(k).x_i(k) + b\right)$$

Where:

- $x_i(k)$ is value of input in discrete time *k*
- $w_i(K)$ is value of weight in discrete time *k*
- b is bias
- F is activation function
- $y_i(k)$ is value of output in discrete time *k*

B. Artificial neural networks

When we combine two or more artificial neurons together, then we get a system that we call artificial neural networks. Single neuron has a little potential for real word problems solving. But if we add addition neurons then we create the artificial neural networks and potential of problems solving was raised. Interconnection between artificial neurons in the ANN is called graph, topology or architecture. Types of interconnection are many, but we can divide them into two basic kinds: feet-forward(FNN) and recurrent(RNN). Mainly different is that the feet-forward artificial neural networks allow flow of the information in one direction: from start to end, but recurrent artificial neural networks allow bilateral flow of the information [20], [21].

C. Feet-forward artificial neural networks

This type of artificial neural networks is based on flow of the information in one direction. This type has not limitations on layers number, number of connection between artificial neuron or type of activation function. Simple example of feet-forward artificial neural networks is shown below[20], [21]:

$$n_1 = F_1(w_1x_1 + b_1)$$
$$n_2 = F_2(w_2x_2 + b_2)$$
$$n_3 = F_2(w_2x_2 + b_2)$$
$$n_4 = F_3(w_3x_3 + b_3)$$
$$m_1 = F_4(q_1n_1 + q_2n_2 + b_4)$$
$$m_2 = F_5(q_3n_3 + q_4n_4 + b_5)$$
$$\text{y} = F_6(r_1m_1 + r_2m_2 + b_6)$$
$$\text{y} = F_6 \begin{bmatrix} r_1(F_4[q_1F_1[w_1x_1+b_1]+q_2F_2[w_2x_2+b_2]]+b_4)+\ldots \\ \cdots + r_2(F_5[g_3F_2[w_wx_w+b_2]+g_4F_3[w_3x_3+b_3]+b_5]) \\ +b_6 \end{bmatrix}$$

This example of FNN has three layout and six artificial neuron. And show how long can be these equations too.

D. Recurent artificial neural networks

This type of neural networks allows bilateral information flow . Others properties are same as feet-forward artificial neural networks. The bilateral flow is provided with use internal states of network. Many basic topologies of these artificial neural networks have fully recurrent each artificial neuron, that means each artificial neuron has connection to any others[20], [21].

E. Convolutional neural networks

Convolutional neural networks are one of type that can be used for recognition of complex events, for instance: human action. We can use 2D or 3D convolutional neural networks [7], [8].

2D convolutional neural networks preforms convolution on convolution layer for the purpose of extraction features from previous layer. Subsequently will be applied bias and results will be inserted to sigmoid function. Then value of a artificial neuron at position [x,y] in the i-th layer and j-th feature map is calculated according formula [8]:

$$v_{ij}^{xy} = tanh\left(b_{ij} + \sum_m \sum_{p=0}^{P_i-1} \sum_{q_0}^{Q_i-1} w_{ijm}^{pg} v_{(i-1)m}^{(x+p)(y+q)}\right)$$

Where:

- tanh() is the function of hyperbolic tangent
- b_{ij} is bais
- m is index of feature maps set, which is at the (i-1)-th layer
- w_{ijm}^{pg} is value at the position [p,q]
- P_i, Q_i are the height and width

3D convolutional neural networks are created by convolving a 3D kernel to a cube. This cube is formed with use multiple contiguous frames. Through this construction, feature maps of convolution layer are connected with previous layer, then value of a artificial neuron at position [x,y,z] in the j-th feature map and i-th layer is calculated according formula [8]:

$$v_{ij}^{xyz} = tanh\Bigg(b_{ij} + \sum_m \sum_{p=0}^{P_i-1} \sum_{q_0}^{Q_i-1} \sum_{r_0}^{R_i-1} w_{ijm}^{pgr} v_{(i-1)m}^{(x+p)(y+q)(z+r)}\Bigg)$$

Where:

- tanh() is the function of hyperbolic tangent
- b_{ij} is bais
- m is index of feature maps set, which is at the (i-1)-th layer
- w_{ijm}^{pgr} is value at the position [p,q,r]
- P_i, Q_i are the height and width
- R_i is the size of the kernel along the temporal dimension

IV. Conclusion

This paper is focus on approach of content recognition. We researched several possibilities in this segment. We started with grammars. Subsequently we focused on the shape grammars in more detail. We found a specific sub-grammar of the shape grammars: shape grammar meta-language (SGML) [13], [14].

Shape grammar meta-language provides rules and descriptors for working with shapes. SGML has seven descriptors: rule-set, directive, label-filer, focus, maxline, zone and apply-mode. These descriptors are used in the six phases of rule application process. Generally, SGML is created well. It has not significantly weaknesses. In our opinion we could use it as first stages of more complex system. SGML allows to define many properties of shape in the drawing. But SGML rely on relatively good image processing. This fact could be problem on the image with low resolution. Also we cannot define more abstract and complex event in the SGML such as: human gestures, human interaction,... In our opinion this is mainly weaknesses of the SGML. But we could join SGML and a context-free grammar or some artificial neural networks together and then we could get interesting results.

Next approach what we researched are artificial neural networks. Because we want to find possibilities of recognition, we started with general definitions and properties. We think that potential of large artificial neural networks is incredible. We found several approaches of recognition in the artificial neural networks such as: convolutional neural networks [7], [8] or RBF Neural Networks [9]. In the convolutional neural networks we can use principles based on 2D an 3D networks. Mainly different is in the process of use. In the 3D convolutional neural networks we can perform connection between frames of the video as next a layer of the network. This concept looks as powerful tool. In the our research we could use this concept for recognition of the complex events which are formed in over time. We believe that main weakness of convolutional neural networks and generally artificial neural networks is in the fact that ANNs are only emulation of the biological process. Although artificial neural networks do have relatively good results of recognition (probably ANNs have better results than any others approaches), but we think that exist functions that have better results than ANNs and they are not ANNs.

In the next research we want to find functions or approaches which would have a better results of content recognition. Now, we think that this system could be based on principles of the grammars and the artificial neural networks, or we could try to create system based on principles of self-organization. Possible approaches are many.

Acknowledgment

This work was supported by project VEGA 1/0341/13 “Principles and methods of automated abstraction of computer languages and software development based on the semantic enrichment caused by communication”.

References

[1] Z. Zhang, T. Tan, and K. Huang, “An extended grammar system for learning and recognizing complex visual events,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 33, no. 2, pp. 240–255, 2011.

[2] R. B. Girshick, P. F. Felzenszwalb, and D. A. Mcallester, “Object detection with grammar models,” in *Advances in Neural Information Processing Systems*, 2011, pp. 442–450.

[3] M. S. Ryoo and J. K. Aggarwal, “Recognition of composite human activities through context-free grammar based representation,” in *Computer vision and pattern recognition, 2006 ieee computer society conference on*, vol. 2. IEEE, 2006, pp. 1709–1718.

[4] N. T. Nguyen, D. Q. Phung, S. Venkatesh, and H. Bui, “Learning and detecting activities from movement trajectories using the hierarchical hidden markov model,” in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 2. IEEE, 2005, pp. 955–960.

[5] M. Pedersoli, A. Vedaldi, J. Gonzalez, and X. Roca, “A coarse-to-fine approach for fast deformable object detection,” *Pattern Recognition*, 2014.

[6] E. Stergiopoulou and N. Papamarkos, “Hand gesture recognition using a neural network shape fitting technique,” *Engineering Applications of Artificial Intelligence*, vol. 22, no. 8, pp. 1141–1158, 2009.

[7] S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back, “Face recognition: A convolutional neural-network approach,” *Neural Networks, IEEE Transactions on*, vol. 8, no. 1, pp. 98–113, 1997.

[8] S. Ji, W. Xu, M. Yang, and K. Yu, “3d convolutional neural networks for human action recognition,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 35, no. 1, pp. 221–231, 2013.

[9] M. J. Er, W. Chen, and S. Wu, “High-speed face recognition based on discrete cosine transform and rbf neural networks,” *Neural Networks, IEEE Transactions on*, vol. 16, no. 3, pp. 679–691, 2005.

[10] Q. Chen, N. D. Georganas, and E. M. Petriu, “Hand gesture recognition using haar-like features and a stochastic context-free grammar,” *Instrumentation and Measurement, IEEE Transactions on*, vol. 57, no. 8, pp. 1562–1571, 2008.

[11] R. Yamamoto, S. Sako, T. Nishimoto, and S. Sagayama, “On-line recognition of handwritten mathematical expressions based on stroke-based stochastic context-free grammar,” in *tenth International Workshop on Frontiers in Handwriting Recognition*. Suvisoft, 2006.

[12] Z. Sükösd, B. Knudsen, M. Værum, J. Kjems, and E. S. Andersen, “Multithreaded comparative rna secondary structure prediction using stochastic context-free grammars,” *BMC bioinformatics*, vol. 12, no. 1, p. 103, 2011.

[13] H. Liew, “Sgml: a meta-language for shape grammar,” Ph.D. dissertation, Massachusetts Institute of Technology, 2004.

[14] ——, “Sgml: A shape grammar meta-language,” in *in W Dokonal and U Hirschberg (eds), Digital Design, Proceedings of the 21st Conference on Education in Computer Aided Architectural Design in Europe*. Citeseer, 2003.

[15] P. Dollár, R. Appel, S. Belongie, and P. Perona, “Fast feature pyramids for object detection,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 36, no. 8, pp. 1532–1545, 2014.

[16] Y. Kato, H. Seki, and T. Kasami, “Rna pseudoknotted structure prediction using stochastic multiple context-free grammar,” *IPSJ Digital Courier*, vol. 2, pp. 655–664, 2006.

[17] M. Fanaswala and V. Krishnamurthy, “Detection of anomalous trajectory patterns in target tracking via stochastic context-free grammars and reciprocal process models,” *Selected Topics in Signal Processing, IEEE Journal of*, vol. 7, no. 1, pp. 76–90, 2013.

[18] C. Lu and R. R. Brooks, “Timing analysis in p2p botnet traffic using probabilistic context-free grammars,” in *Proceedings of the Eighth Annual Cyber Security and Information Intelligence Research Workshop*. ACM, 2013, p. 14.

[19] A. Wang, V. Krishnamurthy, and J. Araújo, “Syntactic inference for highway traffic analysis,” in *Information Fusion, 2009. FUSION'09. 12th International Conference on*. IEEE, 2009, pp. 1355–1362.

[20] A. Krenker, A. Kos, and J. Bešter, *Introduction to the artificial neural networks*. INTECH Open Access Publisher, 2011.

[21] B. Kröse, B. Krose, P. van der Smagt, and P. Smagt, “An introduction to neural networks,” 1993.

Program Comprehension: A Short Literature Review

[1]Matúš SULÍR (1st year),
Supervisor: [2]Jaroslav PORUBÄN

[1,2]Dept. of Computers and Informatics, FEI TU of Košice, Slovak Republic

[1]matus.sulir@tuke.sk, [2]jaroslav.poruban@tuke.sk

***Abstract*—Program comprehension is a study of how programmers understand existing programs. First, we delineate the research field and provide a brief overview of program comprehension theories. Then we present a literature review of program comprehension study results, techniques and tools – from overall comprehension through feature location to understanding the details and rationale behind the source code. Finally, future research directions are suggested.**

***Keywords*—debugging, feature location, mental model, program comprehension**

I. Introduction

According to Biggerstaff et al. [1], a person comprehends a program if one understands its structure, behavior and connection to the application domain. We can say that program comprehension describes how developers understand existing programs (whether or not written by themselves) in order to improve their functional or non-functional qualities.

A. Research Field Definition

To delineate the research field, let us put program comprehension in the context of a software development process (Fig. 1). As a developer reads the specification, a mental model forms in his or her head. A mental model is defined as a representation of a program in the programmer's mind [2]. This model is refined until the developer is able to implement parts of the application and create manually produced artifacts – traditionally the source code. The code is compiled, producing generated artifacts like a runnable application. To comprehend the program, the developer inspects the source code and debugs the produced application. This further refines the mental model and the development continues. Occasionally, the specification must be adapted because not all requirements were implemented in a desired way.

It is important to note that the order of the mentioned processes is not fixed and they often interleave or are performed in parallel. In many cases the programmer starts to familiarize with the program by inspecting generated artifacts – the running application. Finally, the diagram depicts only one specific developer's perspective and does not consider team aspects of software engineering.

B. Related Fields

Two most related research fields are software maintenance and reverse engineering. Maintenance is a much broader term which encompasses practically a whole software development process since it is difficult to distinguish between software creation and software modification.

The terms program comprehension and reverse engineering are often used interchangeably, although there is a slight difference. Reverse engineering focuses on techniques and tools for creation of higher-level abstraction documents from lower-level ones. Program comprehension research is also interested in the effect of using these tools. Therefore, reverse engineering techniques aid developers in program comprehension [3].

II. Theories and Methods

A cognitive model, also called a comprehension model, is a theory describing psychological processes involved in program comprehension which help to construct the mental model of a program [2]. There are multiple cognitive models described by researchers. Generally, we can divide them into top-down and bottom-up approaches.

Top-down program comprehension starts with the programmer's problem domain knowledge. The person forms hypotheses about the program and accepts or rejects them by inspecting the source code and other artifacts. The hypotheses are gradually refined into sub-hypotheses [4]. This type of comprehension is typically used when the program is familiar [5]. An example of a top-down cognitive model is the Brooks model [4].

Bottom-up comprehension is characterized by reading the source code line by line and incrementally grouping them to form higher-level abstractions [2]. An example is the Pennington model [6].

The fact that there are at least seven, relatively complex comprehension models described in literature (six in a review

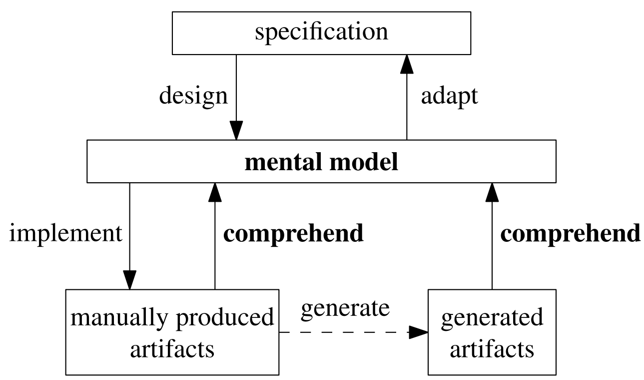


Fig. 1. Program comprehension in the context of a software development process. Solid arrows represent manual, human-performed processes; dashed arrows are automatic, computer-performed operations. The research area of program comprehension is marked in boldface.

[5] and one recent [7]) suggests us an idea that this number is not finite. Each person's cognition is quite different and it depends on a specific situation. It is therefore difficult to generalize the theories and design useful program comprehension tools directly after them. For this reason, many studies and experiments have been conducted in the field of program comprehension.

III. State of the Art

We will now explore the process of program comprehension from the most high-level perspective to detailed inspection, providing an overview of relevant techniques and tools available along with problems developers encounter.

A. Overall Comprehension

Architecture comprehension tools visualize the overall structure of a software system, its components and their relationship. Polymetric view tools like CodeCrawler [8] display code entities as nodes, relations between them as edges and arbitrary metrics as node shape attributes. For example, classes are displayed as rectangles with the width proportional to the method count and the height proportional to the line count. A disadvantage of such tools is the fact they usually focus just on simple quantitative metrics and perform only a static analysis of source code.

Metaphor-based visualizations transform code entities to 2D or 3D objects known from real life. For example, the famous city metaphor displays classes as buildings – like in the EvoSpaces tool [9]. An industrial application of such visualizations is questionable. A more promising approach is offered by ExplorViz [10], which combines them with so-called landscape view – a mixture of UML deployment and activity diagram produced by analyzing execution traces.

Domain analysis tools can be helpful for a new team member to become familiar with the system. The DEAL method [11] extracts a domain model from the GUI (graphical user interface) of a running application. However, manual user interaction is necessary and the method focuses only on a static structure of programs like relations between forms and their controls.

To gain the overview of domain knowledge contained in a program, it is possible to map program identifiers to an ontological dictionary like WordNet using subgraph homomorphism [12]. However, the approach is only semi-automatic and it is necessary to adjust the results manually. This is a result of representational defects like polysemy where one identifier in a program can represent multiple real-world concepts. Some types of defects can be detected and repaired, but the other are inherent to the nature of relationship between programs written in current general-purpose languages (GPLs) and the real world [13].

B. Location of Parts

As software becomes more complicated, it is impossible for a programmer to understand it whole. Therefore partial, or as-needed comprehension is necessary [14].

Each software consists of problem concepts and features (like a shopping cart or shipping method selection) and solution features (e.g., a web presentation and database persistence). One problem domain feature is scattered across multiple solution features. This is called feature delocalization [15]. One of the key objectives of program comprehension is to tackle this phenomenon which impedes code navigation by using feature location (or concept location) tools.

Textual code search is a common feature location approach. To find syntactical patterns, programmers often use complicated regular expressions and the results are unsatisfactory as programming languages are rarely regular. An AST (abstract syntax tree) querying is more viable. By using an intuitive "query-by-example" language [16], developers can search an AST without realizing it. However, the approach is not usable for more complicated patterns.

The MuTT tool [17] implements feature location by execution trace collection. While debugging the program, just before utilizing the feature of interest like adding an item to the shopping cart, the programmer clicks a button in the tool to start collecting the trace. After performing the desired action, the button is clicked again. A list of all methods executed between the two clicks is displayed in an IDE (integrated development editor). It must be manually filtered since it probably contains many auxiliary methods, not related to the tracked feature. Furthermore, to find multiple features, it is necessary to manually repeat the mentioned actions for each of them. The possibility of automation should be investigated.

After the features are located, it is desirable to mark their occurrences in the source code to avoid duplicate work in the future. This activity is called code labeling or concern tagging. In the previous example, we would mark all found methods with a tag "add to cart". An academic tool FLAT[3] [18] combines:

- full-text code search,
- feature location by execution trace collection using the MuTT tool,
- feature labeling with method-level granularity,
- and visualization of feature distribution across the source code files.

An another approach for feature location is differential code coverage [19]. If we run a program first selecting a feature of interest and then not, the difference of code coverage should be the source code of the feature. This approach suffers from similar limitations as execution trace collection.

It would be useful to combine execution trace collection of the MuTT tool with differential code coverage and assess the effectiveness of such approach. Furthermore, as bugs can be considered a kind of features (unwanted, of course), bug localization is a specific kind of feature localization. Differential code coverage could be used to localize the faults, too.

An interesting feature location tool is included directly in the core of an industrial IDE, NetBeans. During the runtime, it is possible to take so-called GUI Snapshot of the debugged program user interface and inspect the events to be performed after e.g., clicking a button or selecting a combo-box item.

C. Understanding the Details

The Theseus tool [20] shows how many times a particular method has been called during a web application execution. These numbers are displayed directly in the code editor and in real time. By clicking a given method, a retroactive log containing the argument values for each execution is shown. This allows for feature location along with a more thorough inspection of the program behavior. A similar tool [21], but for desktop Java applications, operates on individual source code lines instead of methods and thus focuses on detailed algorithm analysis.

Many difficult comprehension questions developers ask when debugging can be answered by asking reachability questions. A reachability question involves searching for all possible execution paths to find source code statements matching the given criteria [22]. However, developers' questions are often vague and difficult to formulate in mathematical terms. The overhead used to formulate such queries can exceed the benefits they provide.

As developers inspect the source code and associated artifacts using various techniques, they create mental models of the code which contain information far behind what is explicitly written. Despite there is a high demand for this implicit knowledge [23], it is not explicitly captured [24] or it is captured only in transient notes, not persisted across sessions [23]. One possible cause is that each developer's mental model is different and thus the personal notes created by one developer are not useful for an another one. This hypothesis should be tested by further research.

One form of temporary knowledge saving is the use of bookmarks in an IDE. To share such personal findings, collective bookmarks [25] can be used. In the study, they were found useful for marking the lines where to start with a given type of task, but not for detailed explanations.

D. Investigating the Rationale

Even if the code is thoroughly understood, programmers often ask why it was implemented this way [24]. One of the manually produced artifacts in Fig. 1 are version control system commits. Commit messages may contain invaluable information about intentions behind the source code.

Surprisingly, less developers actually read commit messages than write them [23]. A possible cause is that IDEs do not present them conveniently. Plugins like Deep Intellisense or Rationalizer [26] try to overcome this issue by displaying history information relevant to a given piece of code.

For version control analysis tools to be successful, each commit must contain code related to only one task. This is often not true in reality. A technique [27] can detect tangled changes before they are committed and suggest how they could be untangled. The disadvantage is a high ratio of false positives.

E. Making Programs Comprehensible

While analyzing existing software is useful, we must also learn lessons from it and design new systems to be more comprehensible.

There is an ongoing effort to create an intuitive programming language Quorum [28], designed after the results of empirical studies. However, it focuses too much on syntax of constructions (e.g., `repeat` instead of `for`) and copies the high-level design of semantics from the Java language.

Identifier naming is a significant factor of program comprehension. Method names can be evaluated for comprehensibility and more intuitive names can be suggested, using a database of commonly used n-grams [29].

Instead of using feature location approaches with often unsatisfactory results, there is a possibility to label classes and methods by their associated concerns using source code annotations immediately when writing the code. Then, source code projections [30] can be used to show classes and methods associated with a given concern in an IDE. While projections are perceived as useful, labeling the code manually is a time-consuming activity which does not directly affect the resulting program execution. Therefore, we can expect programmers to skip labeling or leave the annotations out of sync with the rest of the code, just as it often happens with comments.

Much work is done in the area of code clone detection and removal since there is a widespread opinion that duplicated code negatively affects comprehension. Surprisingly, a study [31] found out that fixing bugs affecting cloned code do not require more effort than fixing other bugs. Code with high regularity, which may be a result of cloning, is more easily comprehended by developers [32]. Furthermore, developers perceive code duplication in a broader sense than simple copy-pasting – for example, implementing the same method in multiple languages or in multiple version control branches [24].

IV. Future Directions

To design tools meeting the expectations of developers trying to understand complex systems, we must first perform experiments to study how they perform program comprehension using the tools available. Only then we can proceed to designing new ones. We plan to conduct qualitative studies designed to reveal the shortcomings of existing program comprehension approaches in specific contexts, produce hypotheses and then continue with quantitative experiments to accept or reject them.

A. Web and Mobile Applications

The majority of program comprehension studies and experiments are performed on a small or medium-sized project, usually self-contained and depending only on standard language and operating system API (application programming interface) and a few simple libraries, if any at all. The main area of interest are WIMP (windows, icons, menus, pointer) and command-line programs. Academic tools developed upon the results of these experiments are therefore usable only for projects subject to these limitations.

Nowadays, software projects are built using large frameworks. Systems have an immense number of dependencies and their behavior often depends on results of asynchronous calls to remote servers. They consist of multiple layers and having a web or mobile interface becomes a standard. Studying the comprehension of these systems is a promising research area.

B. Build Process Comprehension

Compilation and deployment processes of contemporary programs are sometimes more complex than the programs themselves. Especially when using generative programming and model-driven software development, it is important to know:

- what languages, frameworks, libraries and tools are used in the system,
- what are the dependencies between them,
- how do these artifacts interact to produce the intermediate results like object files and executables (if any)
- and how is the system configured to allow for the successful execution and presentation of results to a user.

C. Obtaining Terms from GUIs

In the source code, programmers are not technically obliged to use terms from the problem domain to name the identifiers. On the other hand, GUIs contain information shown directly to the end user an thus must contain domain terms. Using techniques like execution trace collection in combination with GUI ripping [33] could bring the terms back to the source code and at least partially substitute manual code labeling.

D. Language Composition

Many comprehension studies are performed on a program written in one general-purpose language. Today, language composition becomes a common way to construct programs – projects often use a mixture of severals GPLs and DSLs (domain-specific languages). The future research should investigate comprehension factors of using multiple languages in one program – not only from the syntactic aspect, but also the differences between individual paradigms and the effects of frequent switching from one paradigm to an another. It is important to know which combinations of languages and paradigms are more intuitive and closer to the mental model of the majority of programmers.

Acknowledgment

This work was supported by VEGA Grant No. 1/0341/13 Principles and methods of automated abstraction of computer languages and software development based on the semantic enrichment caused by communication.

References

[1] T. J. Biggerstaff, B. G. Mitbander, and D. Webster, “The concept assignment problem in program understanding,” in *Proceedings of the 15th International Conference on Software Engineering*, ser. ICSE ’93. Los Alamitos, CA, USA: IEEE Computer Society Press, 1993, pp. 482–498.

[2] M. Storey, “Theories, methods and tools in program comprehension: past, present and future,” in *Program Comprehension, 2005. IWPC 2005. Proceedings. 13th International Workshop on*, May 2005, pp. 181–191.

[3] H. A. Müller and H. M. Kienle, “A small primer on software reverse engineering,” University of Victoria, Tech. Rep., 2009.

[4] R. Brooks, “Using a behavioral theory of program comprehension in software engineering,” in *Proceedings of the 3rd International Conference on Software Engineering*, ser. ICSE ’78. Piscataway, NJ, USA: IEEE Press, 1978, pp. 196–201.

[5] A. von Mayrhauser and A. Vans, “Program comprehension during software maintenance and evolution,” *Computer*, vol. 28, no. 8, pp. 44–55, Aug 1995.

[6] N. Pennington, “Stimulus structures and mental representations in expert comprehension of computer programs,” *Cognitive Psychology*, vol. 19, no. 3, pp. 295 – 341, 1987.

[7] J. Belmonte, P. Dugerdil, and A. Agrawal, “A three-layer model of source code comprehension,” in *Proceedings of the 7th India Software Engineering Conference*, ser. ISEC ’14. New York, NY, USA: ACM, 2014, pp. 10:1–10:10.

[8] M. Lanza, S. Ducasse, H. Gall, and M. Pinzger, “CodeCrawler - an information visualization tool for program comprehension,” in *Software Engineering, 2005. ICSE 2005. Proceedings. 27th International Conference on*, May 2005, pp. 672–673.

[9] S. Alam and P. Dugerdil, “EvoSpaces visualization tool: Exploring software architecture in 3D,” in *Reverse Engineering, 2007. WCRE 2007. 14th Working Conference on*, Oct 2007, pp. 269–270.

[10] F. Fittkau, J. Waller, C. Wulf, and W. Hasselbring, “Live trace visualization for comprehending large software landscapes: The ExplorViz approach,” in *Software Visualization (VISSOFT), 2013 First IEEE Working Conference on*, Sept 2013, pp. 1–4.

[11] M. Bačíková, J. Porubän, and D. Lakatoš, “Defining domain language of graphical user interfaces,” in *2nd Symposium on Languages, Applications and Technologies*, ser. OpenAccess Series in Informatics (OASIcs), vol. 29. Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2013, pp. 187–202.

[12] D. Ratiu and F. Deissenboeck, “Programs are knowledge bases,” in *Program Comprehension, 2006. ICPC 2006. 14th IEEE International Conference on*, 2006, pp. 79–83.

[13] ——, “From reality to programs and (not quite) back again,” in *Program Comprehension, 2007. ICPC ’07. 15th IEEE International Conference on*, June 2007, pp. 91–102.

[14] V. Rajlich and N. Wilde, “A retrospective view on: The role of concepts in program comprehension,” in *Program Comprehension (ICPC), 2012 IEEE 20th International Conference on*, June 2012, pp. 12–13.

[15] J. Greenfield, K. Short, S. Cook, and S. Kent, *Software Factories: Assembling Applications with Patterns, Models, Frameworks, and Tools*. John Wiley & Sons, 2004.

[16] O. Panchenko, J. Karstens, H. Plattner, and A. Zeier, “Precise and scalable querying of syntactical source code patterns using sample code snippets and a database,” in *Program Comprehension (ICPC), 2011 IEEE 19th International Conference on*, June 2011, pp. 41–50.

[17] D. Liu and S. Xu, “MuTT: A multi-threaded tracer for Java programs,” in *Computer and Information Science, 2009. ICIS 2009. Eighth IEEE/ACIS International Conference on*, June 2009, pp. 949–954.

[18] T. Savage, M. Revelle, and D. Poshyvanyk, “FLAT^3: Feature location and textual tracing tool,” in *Proceedings of the 32nd ACM/IEEE International Conference on Software Engineering - Volume 2*, ser. ICSE ’10. New York, NY, USA: ACM, 2010, pp. 255–258.

[19] K. D. Sherwood and G. C. Murphy, “Reducing code navigation effort with differential code coverage,” Department of Computer Science, University of British Columbia, Tech. Rep., September 2008.

[20] T. Lieber, J. R. Brandt, and R. C. Miller, “Addressing misconceptions about code with always-on programming visualizations,” in *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems*, ser. CHI ’14. New York, NY, USA: ACM, 2014, pp. 2481–2490.

[21] T. Matsumura, T. Ishio, Y. Kashima, and K. Inoue, “Repeatedly-executed-method viewer for efficient visualization of execution paths and states in Java,” in *Proceedings of the 22nd International Conference on Program Comprehension*, ser. ICPC 2014. New York, NY, USA: ACM, 2014, pp. 253–257.

[22] T. D. LaToza and B. A. Myers, “Developers ask reachability questions,” in *Proceedings of the 32nd ACM/IEEE International Conference on Software Engineering - Volume 1*, ser. ICSE ’10. New York, NY, USA: ACM, 2010, pp. 185–194.

[23] W. Maalej, R. Tiarks, T. Roehm, and R. Koschke, “On the comprehension of program comprehension,” *ACM Trans. Softw. Eng. Methodol.*, vol. 23, no. 4, pp. 31:1–31:37, Sep. 2014.

[24] T. D. LaToza, G. Venolia, and R. DeLine, “Maintaining mental models: A study of developer work habits,” in *Proceedings of the 28th International Conference on Software Engineering*, ser. ICSE ’06. New York, NY, USA: ACM, 2006, pp. 492–501.

[25] A. Guzzi, L. Hattori, M. Lanza, M. Pinzger, and A. van Deursen, “Collective code bookmarks for program comprehension,” in *Program Comprehension (ICPC), 2011 IEEE 19th International Conference on*, June 2011, pp. 101–110.

[26] A. W. Bradley and G. C. Murphy, “Supporting software history exploration,” in *Proceedings of the 8th Working Conference on Mining Software Repositories*, ser. MSR ’11. New York, NY, USA: ACM, 2011, pp. 193–202.

[27] H. Kirinuki, Y. Higo, K. Hotta, and S. Kusumoto, “Hey! Are you committing tangled changes?” in *Proceedings of the 22nd International Conference on Program Comprehension*, ser. ICPC 2014. New York, NY, USA: ACM, 2014, pp. 262–265.

[28] A. Stefik and S. Siebert, “An empirical investigation into programming language syntax,” *Trans. Comput. Educ.*, vol. 13, no. 4, pp. 19:1–19:40, Nov. 2013.

[29] T. Suzuki, K. Sakamoto, F. Ishikawa, and S. Honiden, “An approach for evaluating and suggesting method names using n-gram models,” in *Proceedings of the 22nd International Conference on Program Comprehension*, ser. ICPC 2014. New York, NY, USA: ACM, 2014, pp. 271–274.

[30] J. Porubän and M. Nosáľ, “Leveraging program comprehension with concern-oriented source code projections,” in *3rd Symposium on Languages, Applications and Technologies*, ser. OpenAccess Series in Informatics (OASIcs), vol. 38. Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2014, pp. 35–50.

[31] F. Rahman, C. Bird, and P. Devanbu, “Clones: what is that smell?” *Empirical Software Engineering*, vol. 17, no. 4-5, pp. 503–530, 2012.

[32] A. Jbara and D. G. Feitelson, “On the effect of code regularity on comprehension,” in *Proceedings of the 22nd International Conference on Program Comprehension*, ser. ICPC 2014. New York, NY, USA: ACM, 2014, pp. 189–200.

[33] A. Memon, I. Banerjee, and A. Nagarajan, “GUI ripping: reverse engineering of graphical user interfaces for testing,” in *Reverse Engineering, 2003. WCRE 2003. Proceedings. 10th Working Conference on*, Nov 2003, pp. 260–269.

Recommendation techniques for off-line and on-line process mining tasks

[1]*Cecília HAVRILOVÁ (3th year)*
Supervisor: [2]Ján PARALIČ

[1,2]Dept. of Cybernetics and Artificial Intelligence, FEI TU of Košice, Slovak Republic

[1] cecilia.havrilova@tuke.sk, [2]jan.paralic@tuke.sk

***Abstract*—The aim of this paper is to present the area of on-line operational process support as a one of the current tasks of process mining. Whereas the scope of process mining is very wide, this paper deals with its specific subtasks of event logs analysis. Particular solutions are described through by realized experiments in Dutch Academic Hospital and Dutch Finance Institute. The area of recommendations is covered by basic case studies and experiments. Paper also considers the possibilities of the future research which are based on connection of process mining and recommender systems.**

***Keywords*— event logs, process mining, on-line process mining, recommendations**

I. Introduction

The ability to transform data into valuable information leads to significant improvements in management and decision making. Nowadays, organizations around the world store a multitude of data every day. Information systems daily log a large amount of events from real life. A lot of executed activities and events are mediated by some enterprise information system or other kind of software application, where they are recorded and kept. This enables to monitor and subsequently coordinate and simplify all behavior processes [1]. Events are used to characterize the dynamic behavior of a process in terms of identifiable, instantaneous actions, or decision making upon the next activity that will be performed. However, it is not only important to know how the processes were launched or how particular activities interacted. It is also important to know, how can be this information utilized in the future. Because the performing of actual process instances and executing particular activities cannot be retrospectively revised or changed, the techniques for on-line analysis of actually performed events come increasingly to the fore. These techniques allow the further influence of the running process in the future, by means of detecting deviations, making predictions or making recommendations. This paper describes the area of our general recommendations proposals (based on optimal process model) which can influence the process handling in the future. Recommender systems can be useful in improving decision making process in complex environments and enhancing decision ability for decision makers.

II. The State Of The Art

The key document titled Process Mining Manifesto (created by the IEEE Task Force on Process Mining) aims to promote the topic of process mining was published in October 2011. Since then, several significant documents in fields of process mining area were published, e.g. papers dealing with process model discovery or papers describing the particular methods and algorithms of process mining, written by Aalst, Weijters, Maruster [2], Weijters, Ribeiro [3] or Medeiros, Weijters, Aalst [4]. In 2011, Aalst [1] introduced the idea of on-line process mining. Afterward, basic concepts of on-line process mining techniques were defined in Aalst [5], Carmona and Gavalda [6]. In the context of recommendations, Conforti et. al proposed a recommendation system that supports process participants in taking risk-informed decisions, with the goal of reducing risks that may arise during process execution. Current approaches to develop recommendations, current trends and the most used methods and their applications are described in detail in publications of authors Jannach et al. [7] and Ricci et al. [8].

III. Solved Tasks

In the first phase of our research, we were concentrated on general description of process mining area. We proposed formal algorithm of Knowledge Discovery in Databases (KDD) in context of process mining [9]. After that, we started to deal with the analysis of relationships in the records of event logs [10], [11]. We also examined possibilities of graphical visualizations of process behavior via Fuzzy Miner algorithm [12].

In the next phase of the research, we dealt with general process mining tools – ProM and Disco. We designed own process mining tool (based on Fuzzy Miner algorithm) and implemented it into Matlab [13]. Then we realized some experiments based on process model discovery in different level of abstraction. Process models which we obtained were compared with models in other process mining tools and results were comparable. Compared with other process mining tools, our tool contains also explanations of process mining theory behind offered algorithms. This is the reason why it can be used in education process, where students can analyze different types of event logs and obtain necessary knowledge from process mining area. We also realized several experiments focused on off-line event logs analysis from different perspectives, e.g. organizational perspective, time perspective, control-flow perspective and case perspective.

We focused on organizational units' relationships and cooperation (via event logs analysis) in Deutsch Academic Hospital. Realized experiments were necessary for the optimal process model design. We also used the social network analysis and graphical visualizations of processes. Our ambition was to obtain significant results which could help various medical facilities (focused on gynecological diseases) in area of planning, security and disease prevention [14]. Other experiments were focused on complex process analysis in Dutch Financial Institute. Our research was oriented on analysis of credits and loans applications through process discovery. We discovered process model of credits and loans applications and analyzed records of particular applications. We also dealt with fraud detection and analysis of relationship between requested amount of money and duration of the application process.

All of realized experiments and studies have produced significant results but they were fixated on off-line event logs analysis. Nowadays, we design proposal for on-line event logs analysis (based on optimal process model) which takes account of recommendations to allow the further influence of the process in the future. The aim is to predict possible behavior in particular process instances. We have already realized experiments based on some recommendations, but these experiments were produced with non-process aware data. In all of these experiments we used methods of collaborative filtering based item recommendations (e.g. Bayesian personalized ranking matrix factorization, weighted regularized matrix factorization, user k-NN, item k-NN) and methods of item rating prediction (e.g. global average, user item baseline, user k-NN, item k-NN, Bi-polar slope one, biased matrix factorization, factor wise matrix factorization). As first, we analyzed data obtained from a sport e-shop, where we designed recommendations for all new and existing customers by explicit feedback. Our recommendations considered also other important factors such as seasonality, age of customers, gender, locality, etc. The results obtained with existing customers were obviously better than results obtained with new customers or customers who did not be logged in. Although deviations were permissible and acceptable for e-shop, we still continue with analysis and try to find the best recommendations. We also analyzed MovieLens data which contains information about users (user id, gender, age, and occupation), movies (movie id, title and genre) and ratings (user id, movie id, rating and timestamp). We used Rating matrix which contains film ratings from individual users on the scale from 1 to 5 and other methods to find top-N interesting movies for particular user, user-based collaborative filtering, item-based collaborative filtering and also random recommendations. For evaluation of results we used metrics focused on determination of accuracy errors of recommendations such as MAE (Mean Absolute Error), MAPE (Mean Absolute Percentage Error), RMSE (Root Mean Square Error) and Precision/Recall method. Based on these metrics, we evaluated the user–based collaborative filtering as the best and the most precise method for these recommendations. Conversely, the random recommendation was uncovered as the worst method for recommendations in this case. These results can help us in the next phase of our research, where we want to apply obtained knowledge to analyze process aware data and create set of recommendations for them.

IV. Conclusion And Future Work

Realized experiments are necessary for the process model design and optimization and also for analyzing of different methods of recommendations. In the next phase of our research, we aim to create recommendations in different parts of business processes and influence running processes in the future for the purpose of process optimization.

Acknowledgment

This work was supported by the Slovak Research and Development Agency under the contract No. SK-CZ-2013-0062.

References

[1] Aalst, W.M.P. van der, Process Mining: Discovery, Conformance and Enhancement of Business Processes, 2011: Springer Verlag.

[2] Aalst, W.M.P van der; Weijters, A.J.M.M.; Maruster, L., „Workflow Mining: Discovering Process Models from Event Logs,“ rev. IEEE Transactions on Knowledge and Data Engineering, 2004.

[3] Weijters, A.J.M.M.; Ribeiro, J.T.S., „Flexible Heuristics Miner (FHM),“ Eindhoven University of Technology, Eindhoven, 2010.

[4] Medeiros, A.K.A.; Weijters, A.J.M.M.; Aalst, W.M.P. van der, „Using Genetic Algorithms to Mine Process Models: Representation, Operators and Results,“ WP 124, Eindhoven University of Technology, 2004.

[5] Aalst, W.M.P. van der, „Business Process Simulation Revised,“ rev. Enterprise and Organizational Modeling and Simulation, volume 63 of Lecture Notes in Business Information Processing, Springer-Verlag, Berlin, 2010

[6] Carmona, J.; Gavalda, R., „Online Techniques for Dealing with Concept Drift in Process Mining,“ rev. 11th International Symposium, IDA 2012.

[7] Jannach, D., et al. Recommender systems: an introduction. Cambridge University Press, 2010. ISBN 978-0-521-49336-9

[8] Ricci, F., Rokach, L., Shapira, B.: Introduction to recommender systems handbook. In: Recommender Systems Handbook. Springer US, 2011. p. 1-35. ISBN 978-0-387-85819-7

[9] Havrilová, C.: Concept definition for process discovery purposes. In: SCYR 2013: Proceedings from conference : 13th Scientific Conference of Young Researchers : May 14th, 2013, Herľany, Slovakia. - Košice : TU, 2013 S. 77-80. - ISBN 978-80-553-1422-8

[10] Havrilová, C.; Paralič, J., “Identification of relationships and linkages in event logs” (Identifikácia vzťahov a väzieb v záznamoch udalostí) rev. Datakon Znalosti 2013 : sborník konferencí, Ostrava, 2013.

[11] Havrilová, C.; Paralič, J., “Identification of process bottlenecks through by analysis of event logs” (Identifikácia kritických miest procesov prostredníctvom analýzy záznamov udalostí) rev. Electrical Engineering and Informatics 4 : Proceedings of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice, 2013.

[12] Havrilová, C., „Process mining : Events and traces visualization proposal,“ rev. Poster 2013 : 17th International Student Conference on Electrical Engineering, Praha, 2013.

[13] Oščippovský, L.; Havrilová, C.; Paralič, J.,: Podpora vybraných fáz dolovania procesov v Matlabe; in Electrical Engineering and Informatics 5 : Proceedings of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice, 2014. - Košice : TU, 2014 S. 506-509. - ISBN 978-80-553-1704-5

[14] Havrilová, C.; Paralič, J.,: Analýza procesných dát z pohľadu organizačnej perspektívy; in Electrical Engineering and Informatics 5 : Proceedings of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice, 2014. - Košice : TU, 2014 S. 502-505. - ISBN 978-80-553-1704-5

Security of digital assets and anonymity of the owner

[1]*Ján ŠTOFA (4rd year)*
Supervisor: [2]Iveta ZOLOTOVÁ

[1,2]Dept. of Cybernetics and Artificial Intelligence, FEI TU of Košice, Slovak Republic

[1]jan.stofa@tuke.sk, [2]iveta.zolotova@tuke.sk

***Abstract*—the present contribution offers an overview of the actual work in the last year of PhD. study. It deals with secure future assets of the digital world. It considers the question of "What will happened in the virtual world after your "*passing away*" from the real world?". In this issue we look through the table of the current offered options of solutions and value to customers from selected research groups. Step by step we reveal a depth of the problem and its solution we conceive in a new dimension. Security, privacy and maximum of anonymity are the terms of highest priority. Maintaining the integrity ultimately we address through the prism of Shamir's algorithm and a refined sophisticated and automated system of digital will.**

***Keywords*— Shamir's algorithm, digital assets, anonymity, security**

I. Introduction and Motivation

The real world operates under the strict laws in force. If a person does not have a prepared last will and does not provide his heirs with the share of heritage, in this case all ownership subject to the just mentioned law. Today's modern legislation but somehow forgot on virtual life, which gradually has become our common part of life. These two worlds are interlinked. Often happens that people are completely clamped to the digital world. When a person "leaves" from the real world, the virtual world does not know about it. The life continues without this man. Nobody ensures your digital persona that won´t live on without you.

To that person will continue coming emails, Facebook messages, job offers, and remainder of subscriptions... These accounts are still active and digital assets remain without control of its owner. Above this issue I intended and worked along with five other young scientists whom I invited to my team during two-week summer school in Trento (Italy) on topic „ *Security & Privacy in Digital Life* " [3].

Table 1 Overview of competitors

Compe-titors	Estab-lished date	Value to customers (Strengths&Weaknesses)
Best Bequest	2009	Cloud based safety deposit box; Create special messages; Separate account trustees, Account activators, Media Heirs; non anonymous
Netarius	2009	Managing Social Networks, preparing last e-mails, goodbye messages on forums and blogs; leaving your all your corresponding access data and passwords; one master key which can be stolen or lost; non anonymous and secure
Legacy locker	2009	PasswordBox; world's first digital life manager; Passing the access to the online accounts to friends and loved ones; offers an unparalleled range of password manager features such as: instant 1-Click logins, secure sharing and multi-device syncing; allows users to securely store, retrieve and share passwords and other personal data anytime, anywhere on any device; non anonymous
Everplan	2011	Complete archive of everything your loved ones will need; Securely store wills, passwords, funeral wishes; uses bank-level AES-256 encryption, SSL using 2048-bit certificates and other industry-leading security technology to guard data; non anonymous

II. Overview of Competitors

Currently, in virtual world, doesn´t exist ensuring of "inheritance" of individuals' digital assets, which would be easily accessible to everyone. Research groups and companies dealing with similar issues address the matter only in general terms. Security does not have a significant position. Anonymity of client in their solution is unknown concept. Table No.1 gives an overview of several companies offering limited ways of inheritance in the digital world [6].

III. Solving the Issue

This issue can be called a "black hole" of the virtual world and of the current legislation. In order to minimize the threat posed by the "black hole", we implemented a plan offering a solution to this problem. The aim of this plan is to offer clients the opportunity to have in control all their digital assets and include them in their digital will. A new aspect to be taken into account is the maximum ensure of privacy and obtain maximum anonymity of clients and their digital assets. A key element for the solution of this problem is the application of Shamir's algorithm [4].

It is a form of secret sharing, where a secret S is divided into n parts of data $S_1, \ldots, S_n$ in such a way that:

1. Knowledge of any k or more S_i pieces makes easily computable.

2. Knowledge of any $k-1$ or fewer S_i pieces leaves S completely undetermined (in the sense that all its possible values are equally likely).

This scheme is called (k,n) threshold scheme. If $k=n$ then all participants are required to reconstruct the secret. The system relies on the idea that you can fit a unique polynomial of degree $(t-1)$ to any set of t points that lie on the polynomial.

Privacy: having access to any $k-1$ shares from $\{s_1,\ldots,s_n\}$ gives no information about the value of s, i.e., the probability distribution of $k-1$ shares is independent of s.

Shamir´s algorithm can be applied by using the principle of approximation either by interpolation (Lagrange and Newton) or by using the method of least squares [1], [2].

Lagrange's form of the interpolating polynomial is expressed by [5]:

$$P_n(x) = y_0 l_0(x) + y_1 l_1(x) + \cdots + y_n l_n(x) = \sum_{i=0}^{n} y_i l_i(x) \tag{1}$$

,where $l_i(x)$ is called fundamental polynomials defined by:

$$l_i(x) = \frac{(x-x_0)(x-x_1)\ldots(x-x_{i-1})(x-x_{i+1})\ldots(x-x_n)}{(x_i-x_0)(x_i-x_1)\ldots(x_i-x_{i-1})(x_i-x_{i+1})\ldots(x_i-x_n)} \tag{2}$$

Newton's form of the interpolating polynomial is expressed by:

$$\begin{aligned} Pn(x) &= a_0 + a_1(x-x_0) + a_2(x-x_0)(x-x_1) + \cdots \\ &+ a_k(x-x_0)(x-x_1)\ldots(x-x_{k-1}) \\ &+ a_N(x-x_0)(x-x_1)\ldots(x-x_{N-1}) \end{aligned} \tag{3}$$

The Method of Least Squares is expressed by [7]:

$$\varepsilon = \sum_i (Y_i - \hat{Y}_i)^2 = \sum_i [Y_i - (a + bX_i)]^2 \tag{4}$$

Currently are under testing models of encryption via these methods. An important attribute is the speed at which the algorithm is able to encrypt and decrypt a key document. Scheme no.1 simply describes the principle of encryption applied to the issues examined.

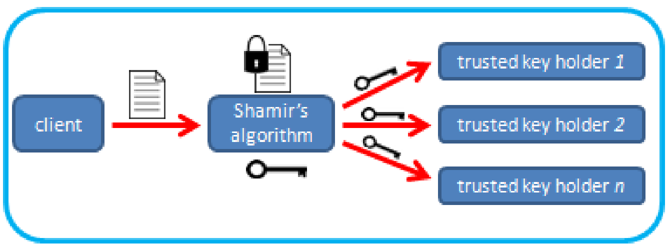


Figure 1 Encryption process

Scheme no. 2 simply shows a method for decrypting used to solve given problems.

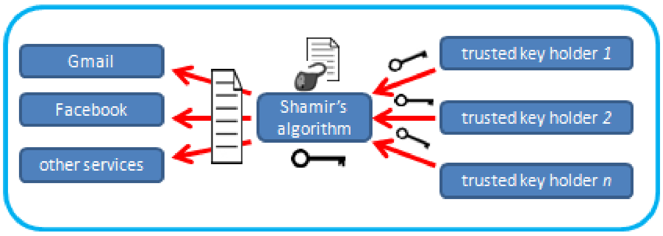


Figure 2 Decryption process

Just the idea of this project was awarded by important Italian investors and evaluated under the highest category of start-up projects as "*Best entrepreneur team*" during above mentioned summer school [3]. This appreciation has motivated me so much that I decided to develop this idea and to go more in details and offer a comprehensive solution to these issues that I have described above.

IV. FUTURE WORK

The big challenge was to create architecture for fully automated system. The objective of this system is to ensure maximum privacy and anonymity of the client without entering any other person in the process of inheritance of digital assets. Although it seems to be the utopian vision at present, but in the near future will have the strong position.

Full functionality of the system is blocked currently by the lack of legislation in the virtual world as well as by the lack of a direct connection with the competent authority. We believe that soon will be possible, based on a semi-automated system, to implement the heritage of digital assets also possibly in cooperation with the executors of last wills in the real world as a possible future start-up.

ACKNOWLEDGMENT

This publication is supported by the KEGA grant of the Ministry of education, science, research and sport of Slovak Republic No. 001TUKE-4/2015 project (100%).

REFERENCES

[1] BUŠA, J., PIRČ, V., SCHRÖTTER, Š.: Numerické metódy, pravdepodobnosť a matematická štatistika FEI TU Košice, 2006, ISBN 80-8073-632-4

[2] BEIMEL, A.: Coding and Cryptology - Secret-Sharing Schemes: A Survey, Department of Computer Science, Ben-Gurion University, Beer-Sheva, Israel. Lecture Notes in Computer Science Volume 6639, 2011, pp 11-46, 978-3-642-20901-7

[3] Summer school in Trento [on-line] [cit 2015.3.5]. Available online: <http://www.eitictlabs.eu/news-events/news/article/security-privacy-in-digital-life-summer-school-in-trento/#allView>

[4] Foundations and properties of Shamir's secret sharing scheme Research Seminar in Cryptography [on-line] [cit 2015.3.5]. Available online: <research.cyber.ee/~peeter/teaching/.../bogdanov.pdf>

[5] Interpolation [on-line] [cit 2015.3.5]. Available online: <fstroj.uniza.sk/kam/orsansky/pdf/interpolacie.pdf>

[6] Designing and Implementing an Effective Information Security Program: Protecting the Data Assets of Individuals, Small and Large Businesses [on-line] [cit 2015.3.5]. Available online: <http://www.sans.org/reading-room>

[7] The Method of Least Squares [on-line] [cit 2015.3.5]. Available online: <https://www.utdallas.edu/~herve/Abdi-LeastSquares06-pretty.pdf>

Shaping of 3D model of human head by the stereoscopic scanning

Ondrej KOVÁČ (4th year)
Supervisor: Ján MIHALÍK

Dept. of Electronics and Multimedia Communications, FEI TU of Košice, Slovak Republic

Ondrej.Kovac@tuke.sk, Jan.Mihalik@tuke.sk

***Abstract*— The article summarizes student´s work for the last year. In the last year main activity continued in analysis of the stereoscopic spatial coordinate estimation of the human head object. Based on the analysis, the parallel stereoscopic camera system was assembled. This system has imperfections that have to be removed via the software solution. However the main goal based on the stereoscopic estimation of geometric human head shape shaping of 3D human head model in regard to MPEG-4/SNHC standard.**

***Keywords*— shaping, stereoscopy, homography, anaglyph**

I. Parallel stereoscopic camera system

Fig. 1 shows the parallel stereoscopic camera system (PSCS) configuration. It is clear from the picture that optical axes of cameras are parallel and the planes of the images of left and right cameras are the same.

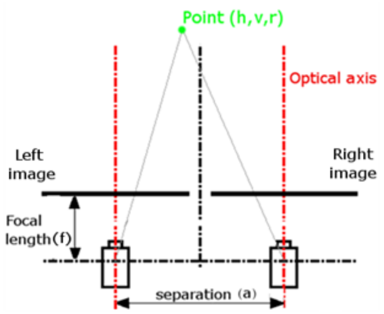


Fig. 1 The arrangement of cameras in ideal PSCS.

It is possible to estimate the spatial coordinates of arbitrary point of the 3D visual scene (3DVS) using thequations referred in [1]. Assembled PSCSwhich experiments were done with is shown on Fig. 2. Cameras of this system are separated by 41mm distance.

Fig. 2 Assembled parallel stereoscopic system.

Assembling of this system was accompanied by problems of the accuracy of the camera holder. I was forced to make it at home with my father due to insufficient tools in the department. I was not able to order the aluminum profile because no company was willing to deliver small amount. This is why it was necessary to drill into iron profile with the inaccurate drilling-machine. Therefore the holes for attachment of cameras are drilled with accuracy 0.5mm. Another problem was the cameras. I had to buy them by myself and so they are ordinary webcams. It is clear the assembled system has not ideal parameters and this had to be corrected by the software.

A. Correction and alignment of images

As mentioned above, in the real stereoscopic system it is hard to ensure that optical axes were truly parallel and cameras lie in one plane. These deficiencies are caused from the inaccuracy of the camera holder and inaccuracy of cameras themselves. Anaglyph of the calibration picture captured by the PSCS without correction is shown on Fig. 3 left. It is clear that this image cannot be used for the spatial coordinates estimation. This can be solved by using of homography we apply here rectification of captured images in regard to calibration picture [2]. This operation is for example important in measuring which requires a rectangular area of measuring and a parallel plane of this area with plane of the camera. This location of camera can be difficult and moreover camera system has distortion and curvature. Issues of the homography are described in [2], [3] . For computation of the homography we need at least 4 corresponding points at least in two images of the scanned plane. As first image is used original picture of the calibration image and the second is calibration image captured by the camera. Using the homography matrix it is possible to re-calculate coordinates of all points of stereoscopic images and to get rectified images. Anaglyph after rectification of stereoscopic images is shown on Fig. 3 right.

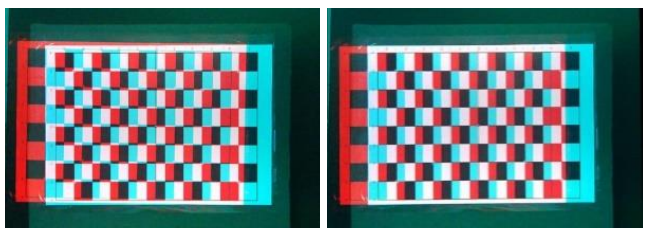


Fig. 3 Anaglyp aquired by PSCS before and after rectification.

Whereas those focal lengths of cameras were not known they need to be computed in process of calibration of the focal length

[2], [4], [5]. In spite of successful calibration and correction the estimation of distance or depth is not accurate. So correction of estimated coordinates is needed. With experimental was confirmed that the error was not linear. So there it is necessary do several measuring of estimated distance of the object and compare them with real distances. On the base of measured errors it is possible estimate their dependencies on the disparity and with this estimation of error the distance or depth can be corrected. [6].

B. Shaping of 3D human head model

Each human has a unique shape of body and of course, the head is not an exception. In MPEG-4 standard 84 facial definition parameters (FDP) are defined. [7]. They are arranged in to groups covering some parts of human head like eyes, nose, mouth and so on. With the feature points on the real head corresponding with FDP on the 3D model it is possible to shape this model in the way that the shape of 3D model corresponds to the shape of the human head. Experimental we determined that 56 features points are sufficient for shaping of the simplified model with 75 vertices if vertices on the edge of the 3D model except vertices of the chin are not shaped. Simplified model is shown on Fig. 4.

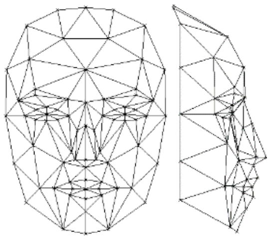


Fig. 4 Frong and side view of simplified base model.

In process of shaping it is also possible to use face symmetry and for simplification it is possible to reduce number of feature points approximate to 2/3. On Fig. 5 the images of human head with feature points from left and right stereoscopic camera are shown.

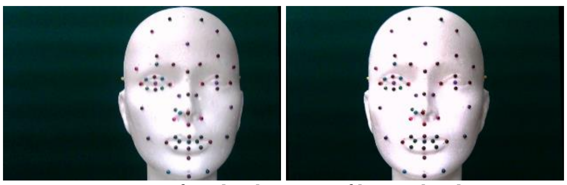


Fig. 5 Left and right image of human head.

It is necessary to say that for the estimation only central and left part of feature points are used. Points of right part are used only for check. On Fig. 6 a shaped 3D model is shown that overlay two orthogonal images of human head.

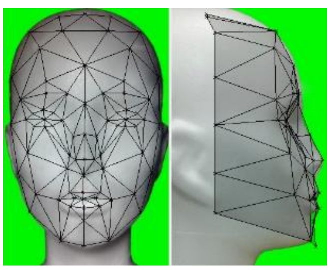


Fig. 6 Shaped 3D model overlaing two orthogonal images of human head.

Shaped 3D model was eventually textured by the texture obtained from front view on human head [4]. The texture can also be obtained directly, from stereoscopic images. Four views of textured 3D human head model are shown on Fig. 7.

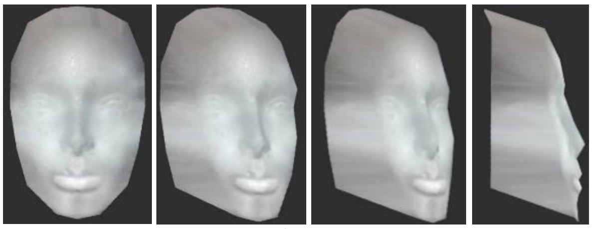


Fig. 7 Rotation of textured 3D model.

On Fig. 7 there are visible sharp distracting edges. These edges are caused by insufficient number of vertices for good surface representation. As mentioned above, number of vertices is reduced for the need of estimation. For suppression of this deficiency it is possible to use DNS approximation.[8], [9].

II. Conclusion

In the last year following results were especially achieved: Analysis and design of PSCS was performed. Based on this analysis PSCS consisting of two webcams was assembled. These webcams do not belong to professional class and they cause image distortion. Furthermore, manufacturer of these webcams does not provide information about focal lengths. The camera holder is not made accurately Therefore we propose simplified procedure of estimation of focal lengths. Alignment of cameras and other imperfections also partially distortion of optical systems of webcams are suppressed by the homographic rectification of stereoscopic images. For final accuracy increase of estimation spatial coordinates experimental obtained correction function was introduced. Estimation of spatial coordinates based on the stereoscopic scanning was applied in area of 3D processing of models. Based on estimation coordinates of feature points on human head placed in 3DVS the base 3D model of human head had been shaped and after that textured by the natural texture.

References

[1] O. Kováč and J. Mihalík, "Estimation of spatial coordinates of 3D objects by stereoscopic scanning," *Acta Electrotechnica et Informatica*, vol. 14, no. 3, pp. 43-48, 2014.

[2] O. Kováč and J. Mihalík, "Teoretická analýza geometrickej kalibrácie kamery," *Posterus*, vol. 7, no. 7, 2014.

[3] C. Loop and Z. Zhang, "Computing rectifying homographies for stereo vision," in *IEEE Computer Society Conference on. IEEE*, 1999.

[4] J. Mihalík and O. Kováč, Číslicové spracovanie textúr ľudskej hlavy, Košice: Technical university of Košice, 2014.

[5] O. Kováč and J. Mihalík, "Kalibrácia ohniskových vzdialeností kamery," in *Electrical Engineering and Informatics 5 : Proceedings of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice*, Košice, 2014.

[6] O. Kováč and J. Mihalík, "Tvarovanie 3R modelu ľudskej hlavy pomocou stereoskopického snímania," *Elektrorevue (Submitet by Editorial Board)*, 2015.

[7] J. Mihalík, "Štandardný videokodek MPEG-4," *Slaboproudý obzor,* vol. 60, no. 2, pp. 7-11, 2003.

[8] O. Kováč and J. Mihalík, "Modelovanie Ľudskej Hlavy," *Elektrorevue, (Submitted by Editorial Board)*, 2014.

[9] J. Mihalík, "Modeling of Human Head Surface by Using Triangular Splines," *Radioengineering*, vol. 19, no. 1, pp. 39-45, 2010.

Significance of Statistical Models in Facial Expression Recognition

[1]*Peter TAKÁČ*

Supervisor: [2]*Ján VAŠČÁK*

[1,2]Dept. of Cybernetics and Artificial Intelligence, FEI TU of Košice, Slovak Republic

[1]takac.p.x@gmail.com, [2]jan.vascak@tuke.sk

***Abstract*—Human faces play a central role in a human-human and also in human-computer interaction, hence they provide very specific information about the emotion state and attention. There are plenty of solutions making the Facial Expression Recognition possible. This paper introduces one of the most used method of Facial Feature Key-Points Extraction - the statistical models. We will show experiments about the accuracy and quality of the proposed method and draw conclusions from the performance.**

***Keywords*—statistical model, Active Shape Model, Active Appearance Model, emotions, Facial Expression Recognition**

I. INTRODUCTION

Interaction between humans and machines is forming a field of numerous research areas, which are important in information processing and transmission. Communicative events, that are present in this kind of interaction, must be similar to human-human communication to ensure natural perception of interaction from the human side. Therefore it is essential to consider multiple communication channels like speech, gestures and facial expression.

In this paper we are focusing on the problem of facial expression recognition, while the human face is capable of preserving miscellaneous identity and emotion information. We use the Ekman emotion model to determine the exact emotion expressed by the human face. Usage of this model gives us versatility in future progress, mainly because there are many distinct emotions, which can usually be created by the combination of the six basic emotional categories this model offers. Furthermore, Ekman's model is universal regarding to the emotions, which can be recognized in nearly all cultures. This gives us the opportunity to have a larger test and train set of natural human expressions, from which we can get better accuracy concerning classification performance. Our approach is to use statistical models for emotion recognition by which we can determine the wanted expression. [16]

II. FEATURE EXTRACTION METHODS

The origins of facial expression analysis go back into the 19th century, when Darwin originally proposed the concept of universal facial expressions in man and animals. Since the early 1970s, Ekman and Friesen (1975) have performed extensive studies of human facial expressions, providing evidence to support this universality theory. These 'universal facial expressions' are representing happiness, sadness, anger, fear, surprise, and disgust. To prove this, they provide results from studying facial expressions in different cultures, even primitive or isolated ones [12]. In one of the ground-breaking and most publicized works, Mase and Pentland (1990) used measurements of optical flow to recognize facial expressions.

Regarding feature-based techniques, Donato, Bartlett, Hager, Ekman, and Sejnowski (1999) tested different features for recognizing facial action units and inferring the facial expression in the frame. Oliver et al. [15] tracked the lower face to extract mouth shape information and fed them to an Hidden Markov model, recognizing again only universal expressions.

In present we can identify various ways of facial expression feature extraction, which can be divided into three categories: deformation feature extraction method, motion feature extraction method, and statistical feature extraction method. [11]

Deformation feature extraction method means extracting some facial deformation informations, such as: geometric deformation or texture changes, the former one mainly refers to the changed relative distance between feature points caused by the variety of expression, the latter one mainly refers to the textures' appearance or disappearance and changes in gradient caused by the changing expressions. [11, 18]

Motion feature extraction method is mainly used to extract some feature points or feature area's motion information from sequential expression images, such as: the movement distance and direction of feature points. The common methods include: feature point tracking, optical flow approaches, and model methods.[17]

Statistical feature extraction method describes the characteristics of expression images by statistics, such as: histogram or moment invariant. Choudhury et al. [8] processed special parts such as the eyes and eyebrows in the way of using consecutive frames subtraction, and then obtained the expression characterization. Shinoharal et al. [9] used 35 kinds of high-order local autocorrelation features as face features. Other researchers calculated the moment invariant of some regions as expression features. Because of it's invariance, statistical feature extraction method is more useful than other methods to the image's rotation, translation and size variation .At the same time it requires a long time for the large amount of computing, and it is easy to ignore specific information of local subtle features.

III. Facial Expression Recognition

As mentioned before, we use the Ekman's emotion model to recognize specific facial emotions. Ekman states that there is a connection between emotions, which are a psychological state, and facial expressions, which are certain behaviors responding to an event. In the case of recognition, we need to analyze the behavior of the face, which was influenced by an exact emotion. This change of behavior can be measured by the change of certain features. These features are considered to be multicultural and universal regarding to emotions. [16]

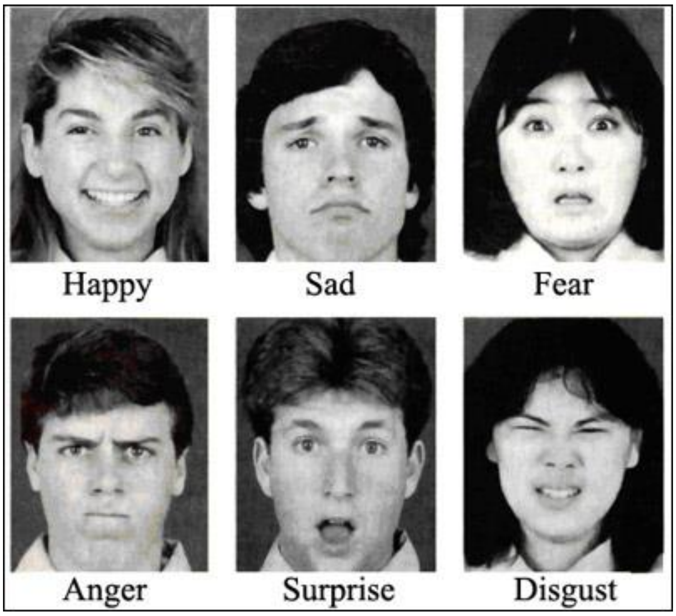


Fig. 1 Examples of the six basic Ekman's emotions.

Hence there are many measurable features on the human face, we can make a distinct model, which will be able to offer accurate simulation of emotions by feature analysis. In this paper we discuss the application of statistical models in face feature modeling and describe two most commonly used approaches: Active Shape Model and Active Appearance Model. While these approaches only deal with feature information extraction, for the problem of expression classification we must use a different solution. The final emotion recognition can be done by any classifier system, where we can input the rules of emotion differentiation described by Ekman in [16].

IV. Statistical models

One of the most important aspects of facial expression recognition is face processing namely the extraction of face feature points. Extraction of these feature points is crucial for correct model building, which should be invariant to gender, race, shape and to some extent also invariant to rotation. Statistical models deal with these problems by statistical analysis of various position of face feature points or facial landmarks. Position variations can be acquired from training sets of faces, which are labeled with emotions. This way we create a generalized model of landmarks, which should be capable to classify any new example to the given class. [1]

There are two approaches of building a statistical model. The first, Active Shape Models, manipulates a shape model to describe the location of structures in the target image, The second, Active Appearance Models, manipulates a model that is capable of synthesizing new images of the object of interest. [1]

A. *Active Shape Models*

Statistical models of shape are used to represent an object in the image. This object is described by a set of points, which are usually configured to be invariant to transformations, mainly translation, rotation and scaling. Main goal of active shape model (ASM) is to find the mean shape derived from the chosen annotated dataset, which represents a generalized model. To find a mean shape or base shape, we need to perform normalization operations (for example Procrustes Analysis) on the dataset and by applying Principal Component Analysis (PCA) on this normalized dataset. [2]

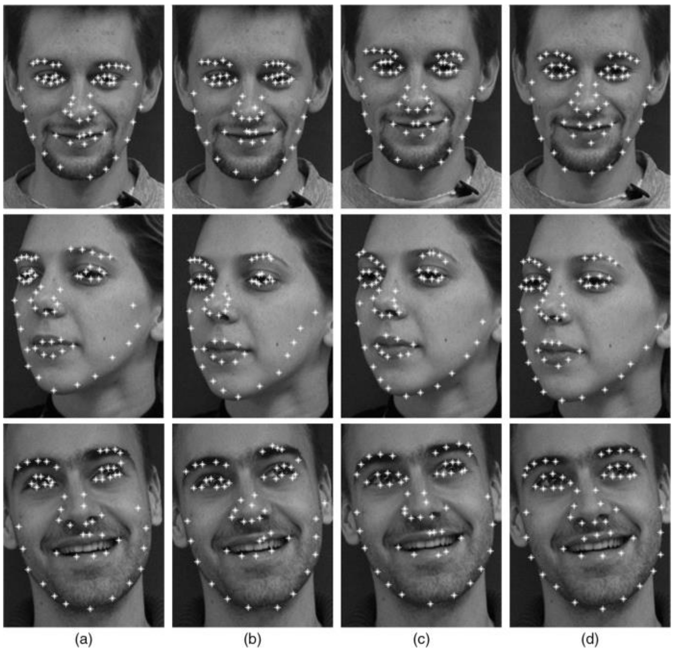


Fig. 2 Examples of the active shape fitting during various iterations (incrementing iterations from left to right) using different grayscale images of faces. The accuracy of fitting of the model is improving with each iteration.

Next step of this approach is the model fitting on a new sample. Since we can describe the position of landmarks in the image, *x*, by a function:

$$x = T_{X_t,Y_t,s,\theta}(\bar{x} + \Phi b)$$

where the function $T_{X_t,Y_t,s,\theta}$ performs rotation by θ, scaling by *s*, translation by (X_t, Y_t) on the current model, we can use traditional gradient based methods to estimate the nearest distribution of landmarks in the image. [1]

B. *Active Appearance Models*

Statistical model of appearance manipulate a model capable of synthesizing new images of the object of interest. The active appearance model (AAM) seeks to find the model parameters which generate a synthetic image as close as possible to the target image [4]. The concept of AAMs is closely related to the concepts of Active Blobs and Morphable Models, which are non-linear, generative and parametric models of certain visual phenomenon. We can also define the fitting of the AAM as solving a nonlinear optimization problem. To solve this problem, the AAM building consist of two parts: finding the mean shape from the dataset (same approach as building the active shape models) and generating the appearance variation. [2]

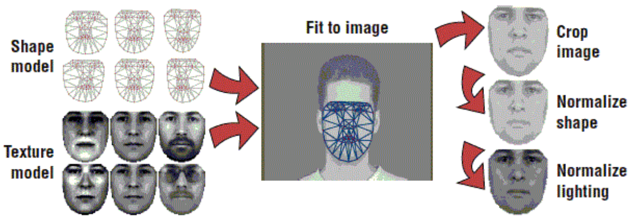


Fig. 3 This figure visualizes the process of fitting of active appearance model. With this approach we combine the shape and the appearance (texture) model into final active appearnace, which can be used for multiple purposes, e.g classification, tracking.

The difference between the shape building in the AAM approach and the ASM approach is in the form of the mean shape, while in the ASM approach we can use any shape of the model (for example contours, "snake" shaped models etc.), in the AAM approach we use meshes, which define the prescription of affine warp from vertices of triangles of the first shape to the vertices of triangles of second shape. The main goal of AAMs is to find such transformation of input image to fit to the generated appearance model, while minimizing the sum of square differences between the model and warped image. We define the error image in the coordinate frame of the AAM and denote it as:

$$E(x) = A_0(x) + \sum_{i=1}^{m} \lambda_i A_i(x) - I(W(x;p)).$$

where $A_0(x)$ denotes the base appearance, $\sum_{i=1}^{m} \lambda_i A_i(x)$ is the pixel mapping function, which maps each pixel of input image x to the corresponding pixel from warped image $I(W(x;p))$. This could lead us to use gradient based methods, but in this case these methods are inefficient because of slow computation caused by recomputation of Hessian direction in each iteration. More effective method is to use Ad-Hoc Fitting algorithms, which assumes, that between the image and additive increments to the shape is a constant linear relationship. Constant linear relationship means, that we don't need to recompute the model parameters in each iteration. [3]

V. Experiments

To test the capabilities of statistical models for face landmarking we experimented with active shape models with SIFT descriptors and MARS [19]. This method replaces the classic 1D gradient profiles of active shape model with a simplified form of SIFT descriptors [13] and uses Multivariate Adaptive Regression Splines (MARS) [14] for descriptor matching. By this modification it outperforms the classic approaches used in active shape models.

We tested this method using the Stasm landmarker [19] version 4.1., which incorporates the mentioned approach, on the Karolinska Directed Emotional Faces (KDEF) [7] and The Informatics and Mathematical Modeling (IMM) Face Database [6]. The KDEF face database contains a set of totally 4900 pictures of human facial expressions of emotion to test the ability of Stasm to annotate these faces with landmarks under different light, scale and rotation variations. The IMM face database contains 240 annotated images of 40 different human faces. These annotations were made manually and no other landmarking technique was used. We compare the accuracy of annotation using the Stasm landmarker and compare its annotations with the annotations made manually.

Fig. 4 Here we compare the manuall annotation and the Stasm landmarker accuracy of feature lanmarking. We can see the different position of landmarks and also the number of feature keypoints. The manuall landmarking scheme is much more simplier, since it uses only 58 points, while the Stasm landmarker uses 77. IMM faces (left), KDEF (right).

VI. Conclusion

We have demonstrated the fitting capabilities and landmark annotation of the Active shape model on frontal face images. The model has shown as a robust solution in the case of frontal face feature points extraction, which can be used as training data for a classificatory or any other approach. We also compared the two main approaches used in within the field of statistical models and showed the differences between them. The Active appearance model is more robust than the Active shape model, but since we mainly use only frontal faces with no rotation, there was no need to use the AAM. And while ASMs are much more universal regarding the possibilities of use, there is a slightly faster development than in case of AAM. In the future we will maybe see much better improvements of ASMs and optimized techniques for facial feature points extraction, which will generally help the field of expression recognition.

References

[1] Cootes, Timothy F., and Cristopher J. Taylor. "Statistical models of appearance for computer vision." (2004).

[2] Matthews, Iain, and Simon Baker. "Active appearance models revisited." *International Journal of Computer Vision* 60.2 (2004): 135-164.

[3] Edwards, Gareth J., Christopher J. Taylor, and Timothy F. Cootes. "Interpreting face images using active appearance models." *Automatic*

Face and Gesture Recognition, 1998. Proceedings. Third IEEE International Conference on. IEEE, 1998.

[4] Ioannou, Spiros V., et al. "Emotion recognition through facial expression analysis based on a neurofuzzy network." *Neural Networks* 18.4 (2005): 423-435.

[5] Wu, Ting, Siyao Fu, and Guosheng Yang. "Survey of the facial expression recognition research." *Advances in Brain Inspired Cognitive Systems*. Springer Berlin Heidelberg, 2012. 392-402.

[6] Nordstrøm, Michael M., et al. *The IMM face database-an annotated dataset of 240 face images*. Technical University of Denmark, DTU Informatics, Building 321, 2004.

[7] Calvo, Manuel G., and Daniel Lundqvist. "Facial expressions of emotion (KDEF): Identification under different display-duration conditions." *Behavior research methods* 40.1 (2008): 109-115.

[8] Bujarborua, Dhiren, et al. "Fluorescein angiographic features of asymptomatic eyes in central serous chorioretinopathy." *Retina* 25.4 (2005): 422-429.

[9] Wu, Ting, Siyao Fu, and Guosheng Yang. "Survey of the facial expression recognition research." *Advances in Brain Inspired Cognitive Systems*. Springer Berlin Heidelberg, 2012. 392-402.

[10] Ekman, Paul, and Wallace V. Friesen. *Unmasking the face: A guide to recognizing emotions from facial clues*. Ishk, 2003.

[11] Trier, Øivind Due, Anil K. Jain, and Torfinn Taxt. "Feature extraction methods for character recognition-a survey." *Pattern recognition* 29.4 (1996): 641-662.

[12] Ekman, Paul, ed. *Darwin and facial expression: A century of research in review*. Ishk, 2006.

[13] Lowe, David G. "Object recognition from local scale-invariant features." *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*. Vol. 2. Ieee, 1999.

[14] Friedman, Jerome H. "Multivariate adaptive regression splines." *The annals of statistics* (1991): 1-67.

[15] Brand, Matthew, Nuria Oliver, and Alex Pentland. "Coupled hidden Markov models for complex action recognition." *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*. IEEE, 1997.

[16] Ekman, Paul. "Facial expressions." *Handbook of cognition and emotion* 53 (1999): 226-232.

[17] Park, Dennis, et al. "Exploring weak stabilization for motion feature extraction." *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*. IEEE, 2013.

[18] Nikolaidis, Athanasios, and Ioannis Pitas. "Facial feature extraction and pose determination." *Pattern Recognition* 33.11 (2000): 1783-1791.

[19] Milborrow, Stephen, and Fred Nicolls. "Active shape models with SIFT descriptors and MARS." *VISAPP* 1.2 (2014): 5.

Social Learning for Personalization in Human-Robot Interaction

[1]*Gergely MAGYAR (2nd year)*
Supervisor: [2]Peter SINČÁK, Consultant: [3]Mária VIRČÍKOVÁ

[1,2,3]Dept. of Cybernetics and Artificial Intelligence, FEI TU of Košice, Slovak Republic

[1]gergely.magyar@tuke.sk, [2]peter.sincak@tuke.sk, [3]maria.vircikova@tuke.sk

***Abstract*—The paper gives an overview of our proposed system for personalization in human-robot interaction using social learning. We present a novel approach for adapting the robot's behavior based on interactive reinforcement learning and the technological tools used. It describes MIRACLE, a scenario for testing our approach and summarizes the work done in the last year.**

***Keywords*—social learning, human-robot interaction, cloud robotics, robotic middleware**

I. Introduction

When we want to have robots in households serving people, we have to resolve a critical issue that arises when thinking about such a scenario. Since we are unable to pre-program the behavior of a robot in every situation, users must teach it new skills. These skills can be social or non-social. For example a social skill can be an appropriate behavior in the presence of a specific person. Non-social skills relate to manipulations with various objects. Another issue is how these new skills can be acquired. Will users call a programmer to explicitly program the new behavior? Will they program these robots using the learning from demonstration technique? If so, will it be required to learn new ways of interacting with the robot or learn to control a particular device?

Another research issue is the short-term vs. long-term interaction. In the case of short-term interaction the robot doesn't need to adapt to the user's needs since the interaction occurs just once or a few times. On the other hand, long-term interaction requires an adaptation mechanism from the robot's side. The adaptation should dwell in changing the parameters of the robot's behavior in order to create an interaction that is natural and enjoyable for the human.

In our work we want to use artificial intelligence methods to create an adaptable robotic system which can be programmed through natural human interaction. In addition to this we want to replace the human operator in the experiments with autonomous behavior.

II. Social Learning in Human-Robot Interaction

Nowadays, HRI experiments are conducted using the Wizard of Oz (WoZ) technique, which means that the robots are not acting autonomously, but they are tele-operated by a human. One of our goals is to move from this approach towards an autonomous mode using social learning.

As it was described in [1], social learning in its original form has three steps. In our work we want to adapt this mechanism for personalization in HRI in the following form:

1. the robot will learn social behavior from the Wizard
2. the robot will apply the learned behavior in social human-robot interaction. It will then receive reinforcement from the human partner and based on this it will adapt its behavior
3. the robot will apply the adapted behavior in interaction with other people and will also ty to adapt them. This way it will create a personalized behavior for different people

In HRI, the main focus of learning from demonstration (LfD) is on teaching the robot new mechanical skills. In our work the role of the Wizard will be to teach the robot social behavior. The learning will be based on interactive reinforcement learning [2]. At the end of the process the robot will have a set of rules defining the appropriate behavior in various situations. Putting it more generally we can say that the robot's task is to approximate the Wizard's state-action policy.

After learning the appropriate social behavior, the robot will try it in a scenario called MIRACLE (Multimodal Interaction with a Robotic Affective Cloud-based Environment). The robot, in our case Aldebaran's NAO, will act as a guide for visitors of two laboratories, one at the Technical University of Kosice and one at the Kyushu Institute of Technology in Japan (Fig. 1). The core of the system will be a cloud service developed on top of the Windows Azure platform. The service will contain a database of knowledge gained during the interactions with the subjects and a set of AI bricks.

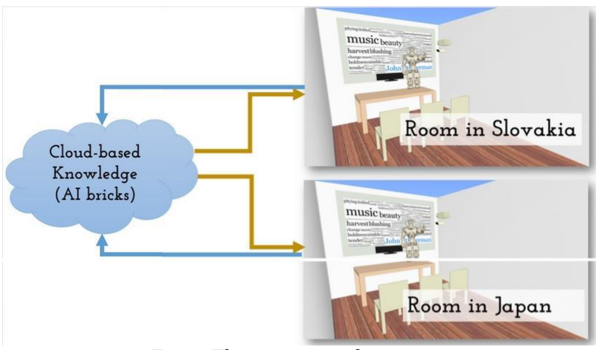


Fig. 1 The experimental set-up

During the presentation the system will sense the human partner's behavior through the Kinect sensor. Our main interest dwells in detecting the subject's non-verbal behavior, such as emotion and gesture. This information will be the input to the reinforcement learning algorithm. The output from the algorithm will be an adapted behavior of the robot. Under the term "adapted behavior" we mean the robot's behavior which parameters were changed during the interaction. In our case, the parameters can be the following: body gestures, emotional expression, voice, gaze.

The third step of the mechanism expects the existence of an adapted behavior created in step 2. The learning process will be the same as in step two with a difference that the robot will adapt the behavior from the previous step.

Our main goal is to replace the Wizard with the interactive reinforcement learning algorithm. To achieve this, only the first step will require the presence of the Wizard. In the other two phases the robot will run in autonomous mode.

The experiments will be evaluated qualitatively and quantitatively. In general it is hard to qualitatively evaluate human-robot interaction, since the participant's opinions are subjective and depend on various facts, e.g. whether they are from a robotic or non-robotic field, have they ever interacted with a robot, what the interaction was about, etc.

One option for such evaluation is the social Turing test. In our case it would ask the subjects whether the robot's verbal and nonverbal behavior during the interaction was generated by a human (Wizard) or by an autonomous agent. Additionally to this, the human judgments can be used as feedback on the interaction. In order to use it the learning algorithm has to be extended.

From the quantitative point of view we are focusing on the learning algorithm. The evaluation will be based on various technical attributes, such as speed, learning time, etc.

III. Cloud robotics and Robotic middleware

One of our technological goals is to combine the advantages of cloud computing [3] and robotic middleware [4]. In [5] it was discussed that robotics can benefit from the cloud computing infrastructure in many ways. In addition to this we want to further enhance the effectivity of creating cloud-based robotics systems using robotic middleware.

Every robotic middleware has a common feature which is the presence of a "master". This component is the core of the whole system, since it is responsible for the communication between modules of the system. In our work we want to port a robotic middleware into a cloud infrastructure. This solution has many possible advantages, such as:

1. master on the cloud
2. components in the cloud
3. robots as clients

In our work we plan to further examine ROS, RT-Middleware and YARP, since they are the most suitable for the human-robot interaction so for humanoid robotics. In addition to this, there are many existing components which can be used in various scenarios. From the cloud's side we will use Windows Azure from Microsoft. Since it offers many ways how to create their own virtual machines it is suitable for experimenting with the mentioned robotic middleware, since each of the has different requirements on hardware and software.

In addition to this, another reason why we want to use robotic middleware is the fact that designing an algorithm in such a system assures portability between different robotic platforms. This way we can test our social learning mechanism using various robots and compare the results.

IV. Current and future work

During the last year we proposed a system which uses a novel social learning mechanism based on interactive reinforcement learning for personalization in human-robot interaction [6]. The system from the technological point of view uses the advantages of cloud computing and robotic middleware.

We further examined several robotic middleware systems: ROS, RT-Middleware, OPRoS, Orocos and YARP. Based on the results we published a comparison study of the mentioned systems [7]. We also successfully completed a summer school of ROS at the FH Aachen University of Applied Sciences and done experiments with ROS in the MS Azure environment.

We developed the basic parts of the MIRACLE system and tested them in Slovakia and Japan. These are: recording data with the Kinect sensor, sending them to a cloud database and playing them on the robot NAO.

The next step will be the implementation of an interactive reinforcement learning algorithm as a ROS module. We also plan to create an online interface for teaching social behavior to the robot. This way we can crowdsource data from Wizards in order to create more diverse behaviors of the robot.

Acknowledgement

Research supported by the National Research and Development Project Grant 1/0667/12 "Incremental Learning Methods for Intelligent Systems" 2012-2015 and supported by the "Center of Competence of Knowledge Technologies for Product System Innovation in Industry and Service", with ITMS project number: 26220220155 for years 2012-2015.

References

[1] Gergely Magyar, *"Social Learning: A new paradigm for learning systems" in SCYR 2014: 14th Scientific Conference of Young Researchers*: proceedings from conference, 2014, pp.247-250

[2] Richard S. Sutton, Andrew G. Barto, *"Reinforcement Learning: An Introduction"*, The MIT Press, 1998, p. 322

[3] P. Mell, T. Grance, *"The NIST Definition of Cloud Computing Recommendations of the National Institute of Standards and Technology"*, Nist Spec. Publ., vol. 145, p. 7, 2011

[4] N. Ando et al., *"RT (Robot Technology) – Component and its Standardization – Towards Component Based Networked Robot Systems Development"*, SICE-ICASE International Joint Conference, pp. 2633-2638, 2006

[5] D. Lorencik, P. Sincak, *"Cloud robotics: Current trends and possible use as a service"* in SAMI 2013 – IEEE 11th International Symposium on Applied Machine Intelligence and Informatics, Proceedings, 2013, pp. 85-88

[6] G. Magyar, *"Social Learning for Personalization in Human-Robot Interaction"*, PhD thesis proposal, 2015

[7] G. Magyar, P. Sincak, Z. Krizsan, "Comparison Study of Robotic Middleware for Robotic Applications", in Advances in Intelligent Systems and Computing, Vol. 316 (2014), pp. 121- 128

Socially based mobility model

[1]Murat HASANI, [2]Ing. Michal Kovalčík
Supervisor: [3]Ľubomír DOBOŠ

[1,2]Department of Computers and Informatics, FEI TU, Letná 9, 042 00 Košice, Slovak Republic
[3]Department of Electronics and Multimedia Communications, FEI TU, Letná 9, 042 00 Košice, Slovak Republic

[1]murat.hasani@student.tuke.sk, [2]michal.kovalcik@tuke.sk

Abstract— **Verification of mobile ad hoc networking protocols depends almost entirely on the simulation. The results of verification are, therefore, very dependent on how realistic movement models that were used in the simulation are. Whereas in the public domain there is a very small number of available real traces, researchers have to focus into synthetic models for movement pattern generation. Currently, the simulation results are often based on randomly generated movement patterns and thus they may offer different results from those that can be obtained from real-life scenarios. Human movement is strongly affected by their needs to socialize. Fortunately, human patterns to associate are known in specific ways that can be mathematically modeled and measured. They have been studied in social sciences for years, so we are able to a certain extent mimic their mobility patterns.**

In this paper we propose a new socially based mobility model and compare it with random movement model that was mostly widely used in the past. The model allows hosts to be grouped together in a way that is based on social relationships among the individuals. This movement, which is influenced by social bonds and group movement, is then mapped to a topological space. Direction of movement and size of the groups may also change in time.

We have validated this proposed model with Random Walk Model by showing that the social mobility traces are better for evaluation and closer to the real-life movement.

Keywords— **Social Mobility Model, Mobility model, Social networks, Social behavior, RWM model.**

I. INTRODUCTION

In the recent years, we can detect the increasing use of the powerful mobile devices with the short-range wireless interfaces (Wi-Fi, Bluetooth, etc.). This resulted in reinforce of the idea of Mobile Ad-Hoc Networks studies. Mobile Ad-Hoc Networks (MANETs) is a class of wireless Ad-Hoc networks in which mobile nodes exhibit features of the free node mobility. Node mobility, in particular, can cause frequent and unpredictable topology changes. Connectivity between mobile nodes is created sporadically and highly depends on the movement of the users carrying those devices. This human movement is strongly influenced by the social relations [1]: Society is naturally divided into the smaller or bigger communities of people, according to their needs for cooperation or selection. We do not meet other people on a purely random basis. From sociology is known, that the idea of correlated interaction is that an individual of a given type is more likely to interact with another individual of the same or similar type (friends, colleagues, family) than with a randomly chosen member of the population.

As was mentioned, the human mobility is strongly dependent to the users' personal and social properties and behaviors as well as the environmental parameters. The mobile users that are attached to the specific locations or individuals, may have significant correlation between their locations constrains and movement patterns. For instance, there has been proven in [2] that mobile users tend to visit just a few places, where they sped most of their time and where they have strong social ties. Also in the most cases are users traveling over the short distances more likely than over the longer distances.

Considering the characteristics of the human mobility, we need to choose wisely an appropriate traces and models for the simulation. Since several models that have been proposed are very similar in nature, there need to be clearly understood every of the existing models and data sets. Generally we can obtain human mobility datasets thanks to the two ways: *realistic* (bounded environments such as colleges, campuses and industries, using Bluetooth or Wi-Fi technologies) or *simulation- based* models (main advantage is that they can be re-parameterized to be used for different scenarios).

Validation of mobile Ad-Hoc network protocols relies almost exclusively on the simulations, which, in turn, rely on realistic movement models for their credibility. Since there is a total absence of realistic traces in the public domain, synthetic models for the movement pattern generation must be used and the most widely used models are currently very simplistic, their focus being ease of implementation rather than soundness of foundation. The most widely used such models are based around random individual movement; the simplest, the *Random Walk Mobility Model* (equivalent to Brownian motion), is used to represent pure random movements of the entities of a system. Thus, the simulation results of the protocols are often based on the randomly generated movement patterns and, therefore, may differ considerably from those that can be obtained by deploying the system in real scenarios. Researchers, however, are in most cases focused mainly on the functioning of the networking protocol and take no emphasis on the environment behavior, in which they want their protocol validate. It was important to model the behavior of individuals moving in groups and between groups, as is likely in the typical Ad-Hoc networking deployment scenarios of disaster relief teams, clamps of animals, platoons of soldiers, etc. In order to capture this type of behavior, it was necessary to define relationships among the people carrying the devices. Existing group mobility models fail to capture this social dimension.

Thus, we propose a new mobility model that is founded on the social network theory, because this has empirically been shown to be useful as a means of describing human relationships. As was already mentioned, the most widely used models are those, which are based on the random movement of hosts. The simplest among all is the Random Walk Mobility model (RWM), which is used to represent pure random movements of the beings in the system [3]. Our approach is based on the simple consideration: mobile networks are social networks after all, since we assume that devices are carried by humans. Clustering is probable in the typical Ad-Hoc networking deployment scenarios (groups of vehicles, herds of animals, platoons of soldiers, etc.). In particular, the model allows collections of the hosts to be grouped together in the way that is based on the social behavior among the individuals. In order to capture this type of behavior, the movement of an individual is to a large extend adapted to movement of a group in which is located. Groups of hosts may become larger, smaller, or completely crumble. Designed model was called *Social Mobility Model*, abbreviated SMM and it is actually enhanced RWM model. As will be shown later in this paper, mobility models that are based purely on random mechanism show properties (such as mutual and close contacts percentages) that are way different from social based mobility model scenarios.

The paper has the following structure: In the Section II is summarized related work over the recent years. Section III describes Human Mobility Modeling and categories of human mobility. Section IV shows how we have decided to propose social mobility model. Section V illustrates the results of the evaluation of the social mobility model based on the comparison with the RWM model. Section VI concludes the paper, summarizing the original contribution of this work.

II. RELATED WORK

The human relationships and their nature have troubled mankind for centuries if not for thousands of years. Even through, we are still not able to fully understand and perceive the way people interact with each other. However, even if we are not able to accurately describe human relationships, we know they have some specific and very characteristic properties.

Concept of the Social Networks in the Mobility Models were for the first time introduced in 2009 by researchers Musolesi and Mascolo in the [4] where they categorized Human Mobility Models into the two types: *Real world traces* and *Synthetic mobility traces*, and they studied their advantages and disadvantages. Likewise, in the [5] is provided survey of the Real world and Simulation based traces as well as the Synthetic mobility models for multi-hop wireless networks. In the [6] are current Mobility Models categorized from an individual and also group perspective, with a brief analysis on their applications.

Later, in 2011 is the survey of simulation-based Mobility Models for OppNets [7] also there are discussed properties of movement extended with notions of the predictability and patterns. Then in the 2012, properties of some selected synthetic and trace-based mobility models are compared in [8], also the changing trends in the Human Mobility Modeling are summarized.

III. HUMAN MOBILITY MODELING

Despite the fact that many researches on the enumeration of the mobility have been performed, they've never been neither comprehensive nor detailed. Because of that the comprehensive classification of Human Mobility issues in the OppNets was never clearly introduced. Therefore, these issues were assigned in the [9] into three main groups (Fig.1): Human Mobility characteristics, Mobility Models, Mobility Prediction methods.

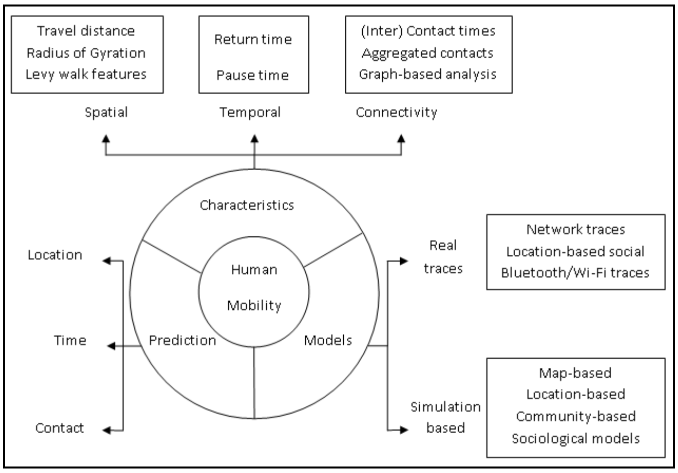


Fig.1 Three aspects of Human Mobility

In the following sections are closely described all three human mobility groups.

A. *Human mobility models*

Human mobility traces are collected from the Real-life Human Mobility or generated using the Simulation-based Methods, which make it possible to analyze and explore the patterns of trajectories of the mobile carriers. However, most of the Realistic movement traces are not scalable, flexible or accessible for public. The Synthetic Models were created for that purpose to capture the movement patterns of nodes in a realistic way.

Social properties play a big role in the proposal and designing of the efficient Delay Tolerant Network (DTN) protocols and therefore models should be able to capture social properties as accurately as possible. Good mobility models should be able to meet these three important requirements [10]:

- Be able to generate mobility patterns as realistic as possible
- Be mathematically tractable
- Be flexible enough to provide different mobility characteristics (qualitative and quantitative)

B. *Human Mobility characteristics*

Statistical characteristics of the Human Mobility are measured for purpose of discovery non-homogeneous behavior and movement patterns of the users' motion in the space and time. Also, they are used to predict users' future walks, destinations and contacts. In the OppNets could be those characteristics exploited to validate the synthetic mobility models [11] and study the effect of mobility in designing and performance of the routing protocols [12]. There exist three categories that explore human movement characteristics:

- **Temporal** - related to the time-varying features of the user mobility
- **Spatial** - refer to the trajectory patterns in the physical space
- **Connectivity** - concern the contact information of the users

C. Human mobility prediction

Tracking behavior of the users' future movement and quantifying their predictable regularity have been topics of considerable interest in recent years. Human mobility prediction capabilities can be used in various exciting mobile applications such as data sharing, epidemic modeling, traffic planning, and disaster response.

Predicting next locations, stay durations and future contacts of people is very important issue in human mobility models. There are many factors that are able to make influence on the user's future mobility patterns ranging from their social or personal factors to the environment parameters. Recent human mobility models demonstrate that human mobility can be predicted to a greater extent. It's not so far, when two groups of researchers [13], [14] found that human behavior is 93%, resp. 95% predictable.

The human mobility prediction methods can be divided in three groups. The first category of the proposals has strived to predict the next visiting locations of mobile users based on history of their trajectories (Location prediction). Other categories have strived to predict a user's arrival time to a location and his/her staying duration (Time prediction). In the third group is studied contact probability of mobile users, i.e. who will contact each other (Contact probability).

IV. DESIGN OF THE SOCIAL MOBILITY MODEL

In this section will be shown how the human mobility model based on the social network theories was proposed. The description of the mobility model is organized as follows:

- We describe the RWM model that we have decided to choose as a representative of the well-established models.
- Movement of the nodes in the both simulation models is introduced. Brownian motion with different speed and direction to achieve random movement was used in the simulation.
- We define social groups that were used as the social enhancement of the RWM model. Social Groups are defined as two or more nodes that interact with one another. Also will be described term Social Group Leader.
- In the next section we introduce and describe other properties of the human motion, which are contacts and their durations. We divided contacts into two main groups, namely Contact times (CTs) and Inter-contact times (ICTs). These properties will be measured and their results compared in our work.

For reasons of clarity, we make the assumption that only one device is associated with each individual. During the simulation process the number of nodes does not change, which means that no nodes are creating, or disappearing. However, the model can easily be extended. In the remainder of the paper, we note that the terms host, node and individual are equivalent and indicate a single moving entity in the mobility scenario.

A. Random Walk Mobility model

Einstein explained The Random Walk Mobility Model (RWM) mathematically in the 1926. It is a simple mobility model based on random directions and speeds. The RWM model was mainly developed to imitate the unpredictable changes. This model serves us as a basis for our SMM model.

In this mobility model, a mobile node moves from its current position to a new position by randomly choosing a direction and speed. Mobile node chooses any random direction to travel until the boundary of edge is found. The new speed and direction are both chosen from predefined ranges, [min-speed, max-speed] and [0, 2×π] resp. It is a memory-less mobility pattern because the current speed and direction of a mobile node is independent of its past speed and direction.

B. Node motion and Position update

In both models was used random walk for motion of nodes. A random walk, or Brownian motion, is a mathematical formalization of a path that consists of a succession of random steps. In the Fig.2 is shown an example of the simulation of the Brownian motion in the 2D environment.

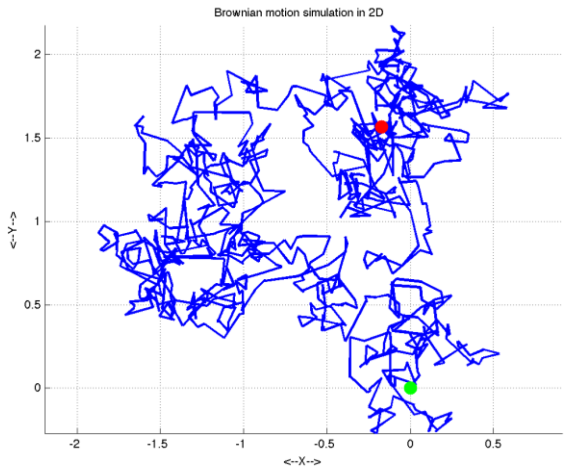


Fig.2 Example of Brownian motion

Position of the node that does not belong to the any social group is generated randomly, giving him random speed in the predefined range of 5− 10 km/h. The equation used for nodes new position is calculated as follows:

$$newX_{node_i} = current\ X_{node_i} + \overline{speed}_{node_i}\Delta t$$
$$newY_{node_i} = current\ Y_{node_i} + \overline{speed}_{node_i}\Delta t$$

Nodes belonging to the Social Group are moving inside the corresponding group in the group direction and speed. Their place in the group is randomly chosen. Every Social Group is moving with random speed that is in the predefined range and also direction. For the group movement, we have decided to choose a Social Group Leader that decides where the group should go. Node has options to remain in the Social Group and thus follow the leader of the group, or decide to leave the Social Group and move with its own random motion. We are using these simple equations for social group movement:

$$newX_{group_i} = current\ X_{group_i} + \overline{speed}_{leader\ of\ group_i}\Delta t$$
$$newY_{group_i} = current\ Y_{group_i} + \overline{speed}_{leader\ of\ group_i}\Delta t$$

C. Insertion of Social behavior

To the best of our knowledge, there are no publicly available human mobility datasets that also provide rich information about the personal interests of the network users. Due to this limitation, we have decided to enhance the social behavior in the RWM model that we were using till now. Mentioned RWM model has been over the years the most widely used model for testing MANET and OppNet protocols not only at our University, but throughout the world. The main disadvantage of this model is that it cannot imitate human behavior at all. However, it is still in use, because the most of researchers paid attention to the functionality of routing protocols and to the environment in which these protocols were tested they were not placing high demands. Problem occurred when routing protocols showed excellent results in simulation, but when they were subsequently implemented to use in practice, they could not even come close to simulation results. Grouping naturally arises at the social level, so creation of Social Groups was in our opinion a good direction in which we should go, if we want to design SMM model.

In the initial phase of the simulation, there is no social relationship between nodes, because nodes are randomly placed around the topology. Nodes start to socially behave immediately in the moment they join *Social Group*. We define Social Group as a set of nodes, which number is ≥2, that interact with each other. Under the term interaction we understand that nodes are in close, or mutual, contact with each other. There is no limitation for the number of groups in the simulation area, but there is a rule that holds:

$$no.of\ social\ groups \leq \frac{no.of\ nodes}{2}$$

In the main, individuals will move within the sphere of influence of the geographic group with which they are associated at any given point in time, but, occasionally, they will decide either to move between groups or to leave the group structure and to move completely independently. So, this means that each node that belongs to some Social Group owns sociability factor, which is set that on 75% node will remain in the social group and 25% that node decide to move in his own direction. Deciding whether the node remains in the social group always occurs at the end of the iteration time (we have set iteration time to 30 seconds). In the meantime, the node moves in the direction of its social group.

Under term *Social Group Leader* we understand that it is a node that is in charge of the some Social Group. When node is in the charge of the Social Group, it decides, in which direction and how fast Social Group has to move. Node has to have some properties to become leader of the Social Group. In our simulation, each node is labeled with his own ID number. For sake of simplicity we have decided that for the leader of the Social Group will be always chosen that node, which has the lowest ID number. In the equation below is rule that holds in every social group in our designed model.

$$ID\ number_{leader\ of\ group_i} < ID\ number_{members\ of\ group_i}$$

Finding of new leader occur always new node joins Social Group, or some node decides to leave the Social Group.

D. Durations of contacts

First of all, in OppNets the contact is defined as a state that occurs when the mobile devices are within mutual radio transmission range of each other. Understanding statistical properties of contacts is essential for the design of protocols and algorithms in irregular connected networks, because every contact represent an opportunity to forward data and bring them probabilistically closer to the destination, or a set of destinations. As was mentioned before, we take closer look at the *Contact times* (CTs), *Inter-contact times* (ICTs) and their durations.

1) Contact times (CTs)

Contact times (CTs), or contact durations, we can define as time intervals during which are devices in the mutual radio range. It can be the duration of a Bluetooth association, or staying under the coverage of common access point, it all depends on the considered scenario. The CTs are main factor, if we want to determine the transmission capacity between encounter devices in a mobile network. It holds that two devices can exchange more data during longer CTs. In our design we decided to divide CTs into two groups on the Close contact times (CCTs) and mutual contact times (MCTs).

We can define CCTs as time intervals, during which is couple of nodes in their close mutual radio range. We have used Euclidean distance for detection, whether or not are nodes in their close radio range. In mathematics, the Euclidean distance or Euclidean metric is the "ordinary" distance between two points in Euclidean space. With this distance, Euclidean space becomes a metric space. The associated norm is called the Euclidean norm. Older literature refers to the metric as Pythagorean Theorem. We can generalize this theorem to find the distance between two points $P(x_1,y_1)$ and $Q(x_2,y_2)$. We draw the triangle, as is shown in Fig.3, and see that the length of its legs is equal to the:

$$a = x_2 - x_1$$
$$b = y_2 - y_1$$

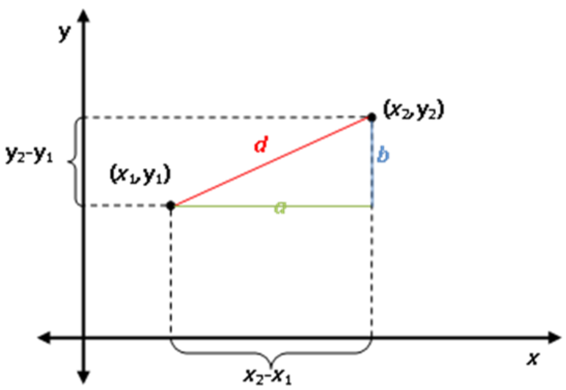


Fig.3 The distance between two points

Note that x_2-x_1 is the opposite of x_1-x_2, which means that x_2-x_1= -(x_1-x_2), so their squares are the same. It does not matter which point uses subscript "1" and which point uses the "2". Therefore the distance from (x_1,y_1) to (x_2,y_2) is:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Graphically, we can see in Fig.4, what we consider as close contact duration. Time t_0 represents initial state of the two nodes that are in this time in no contact. This state holds till time t_1, when nodes start to move towards each other. At the time t_2 we can say that nodes are in CCT duration, which lasts till time t_{n-1}, when nodes start to move from each other. At the time t_n are nodes again in no contact at all.

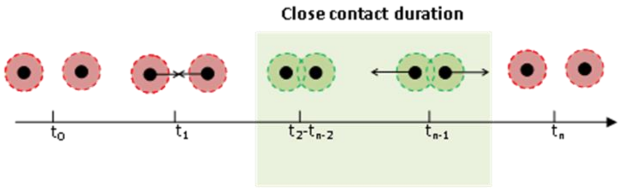


Fig.4 Example of the Close contact duration

From the perspective of close contact duration may be node situated in two states. Whether it is in close contact duration, or it is not.

We can define MCTs as time intervals, during which is number of devices in their close mutual radio range. For this purpose we needed an algorithm that is able to compute distances between all nodes and connect them into the groups, not just into the pairs, or couples of nodes. This problem is in mathematics also called the all-pairs shortest path problem. We have used All Path algorithm, which we obtained as a free source from the [15]. Algorithm gives us information about who is in contact with whom. It rapidly returns the shortest node-to-node distance along the edges of a graph, for all nodes in the graph.

The main difference between CCT and MCT is that we are looking after MCT after CCT is established. As we can see in the Fig.5, *node 1* and *node 2* are at the time t_0 in close contact, while *node 3* is not in any connection with somebody. This state holds till time t_1, when *node 3* starts to move towards *node 2*. At the time t_2 is possible to see two close contact durations between *node 1* and *node 2* and also between *node 2* and *node 3*. We can see that *node 2* is in two close contacts. We suppose that if *node 1* is in close contact with *node 2*, it is also in mutual contact with all his close contacts, in this case *node 3*. This case also applies vice versa, for *node 3*.

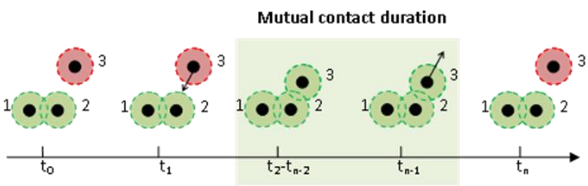


Fig.5 Example of Mutual contact duration

We can say that if node i is in close contact with node j, it is also with mutual contact with all close contacts of node j.

2) *Inter-contact times (ICTs)*

Inter-contact times (ICTs), is the amount of time elapsed between two successive contact periods for a given couple of devices. Shorter ICTs between couple of nodes means that nodes often encounter each other and thus could directly exchange data among each other. For longer ICTs, could be other nodes in the network selected as data carriers in order to deliver data.

In the Fig.6, it is possible to see an example of the ICT duration of the two nodes. Time t_0 represents initial state of CCT of two nodes. This contact holds till time t_1, when nodes start to move from each other. In this moment, at the time t_2 starts ICT duration, which lasts till time t_{n-1}, when nodes start to move towards each other.

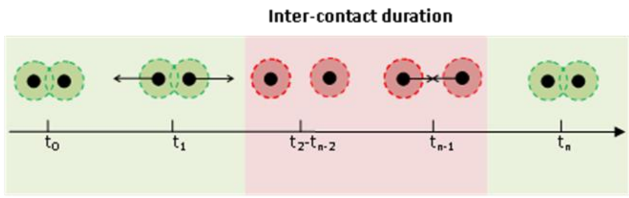


Fig.6 Example of the Inter-contact duration

V. IMPLEMENTATION OF THE MODEL AND SIMULATION RESULTS

We have implemented proposed model in MATLAB environment, which we use for the majority of our research in Mobile Ad-Hoc Networks.

We tested this mobility model using every time 10 runs to generate different scenarios and we compared the results with synthetic traces that were generated by the RWM model. New nodes cannot be created (connected), or disappeared (disconnected) during the simulation. We chose a multiple simulation scenarios, with a different transmission ranges, population density and topology sizes, in order to better see the differences in the results.

First of all we have set for each device equipped with a Wi-Fi, transmission range of 40, 80 and 120m. The size of the topology was $1000\times1000m^2$ with 100 nodes. The nodes were generated randomly throughout the simulation area. And the speeds of the nodes were randomly generated according to a uniform distribution in the range [5− 10] km/h. The duration of the simulation is set to 12 hours.

In the second case we were changing topology sizes and also population densities with the constant Wi-Fi transmission range of 40m.

All the parameters of the designed models are concluded in the Tab. 1.

Category	Parameter	Description
General	Simulation Area	$250\times250m^2$;$500\times500m^2$; $1000\times1000m^2$
	Initial Deployment	$100\times100m^2$; $200\times200m^2$; $400\times400m^2$
	No. of Nodes	60;80;100 nodes
Durations	of iteration	30 seconds
	of simulation	10 hours
Mobility model	Walking speed	5-10 km/h
Network	Transmission range	40m, 80m, 120m

Tab. 1 Simulation parameters

It is also important to note that there is no packet transaction behind contacts. So we do not measure packet deliveries or losses during contact. In our work we have decided to measure degree of connectivity between nodes based on their mutual and close contact durations.

A. *Statistical simulation results I*

1) *Transmission range of 40 meters*

In the Fig.7 are shown results of the CCTs of the both models. While RMM shows decreasing character of CCTs, the SMM model is able to maintain its connectivity during whole simulation.

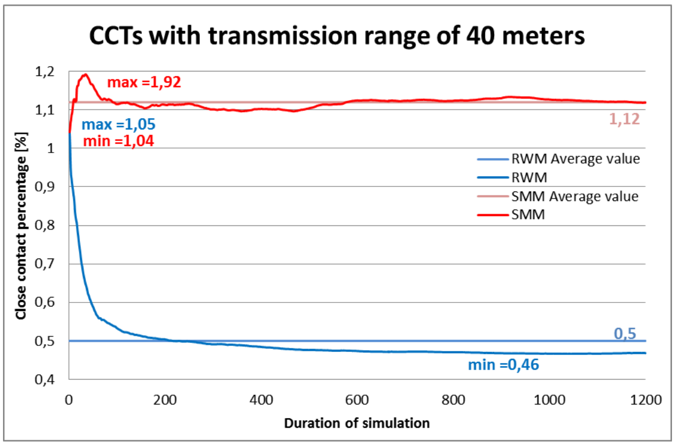


Fig.7 Simulation results of the CCTs with 40m transmission range

Measurement results of the MCTs are shown in the Fig.8. The MCTs had similar character like the CCTs in the previous graph. RWM model had decreasing character during whole simulation while the SMM model was able to increase its initial values.

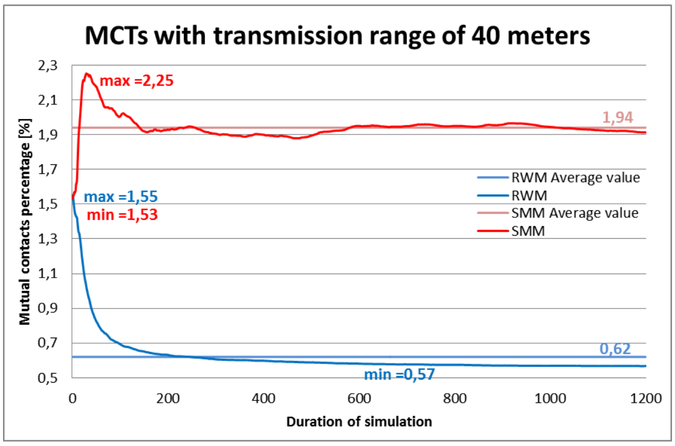


Fig.8 Simulation results of the MCTs with 40m transmission range

2) *Transmission range of 80 meters*

In the Fig.9 is we can see that the RWM model decreases its contact values from the simulation start till its end. The SMM model shows a little contact drops from the start of the simulation, but then it starts to increase CCTs.

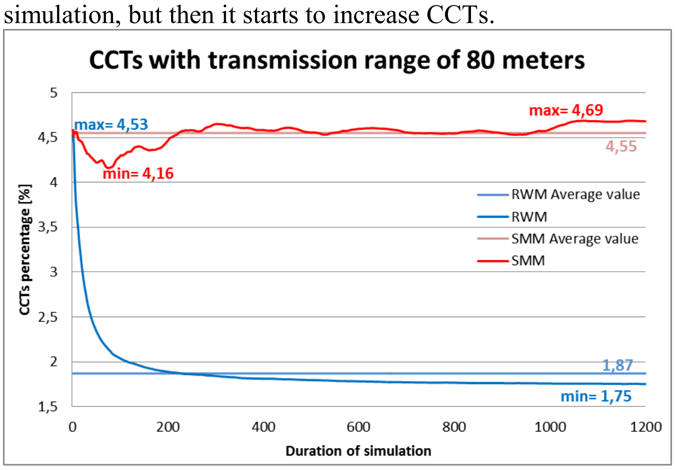


Fig.9 Simulation results of the CCTs with 80m transmission range

The graph with MCTs percentages is shown in the Fig.10. We can see contact drops from the both models, but SMM model is able to maintain higher MCTs percentage during whole simulation.

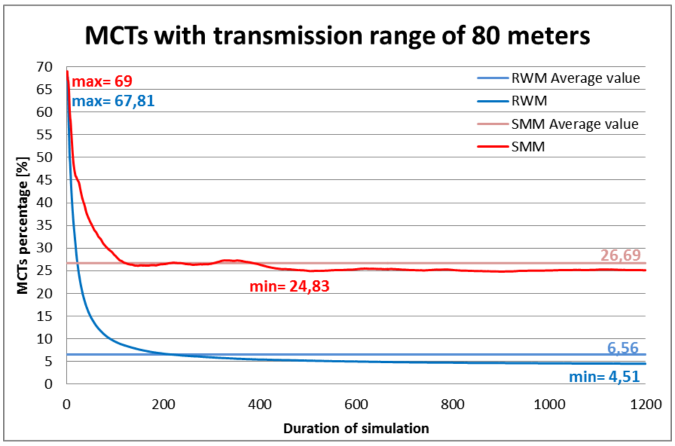


Fig.10 Simulation results of the MCTs with 80m transmission range

3) *Transmission range of 120 meters*

The Fig.11 shows the CCTs percentages of RWM and SMM model. In the case of the RWM model, the values have decreasing character, while SMM model has little drops during first iterations, but after them, CCTs starts to increase.

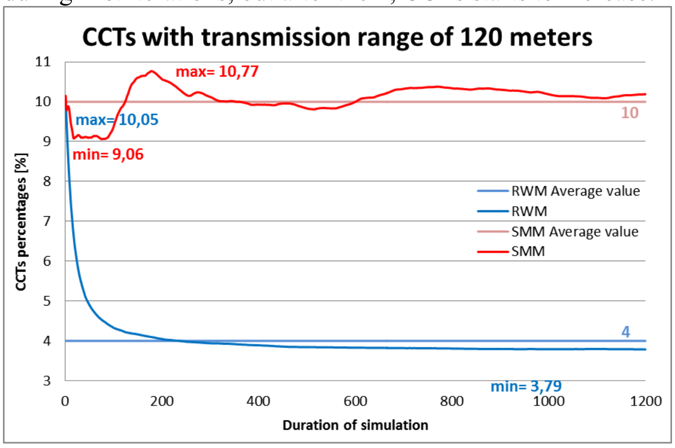


Fig.11 Simulation results of the CCTs with 120m transmission range

In the Fig.12 it is possible to see a graph with the data series of MCTs percentages of the SMM and RWM models. Nodes had the highest CTs in the beginning of the simulation. As the time goes, the nodes started to move over the whole topology and thus, connectivity started to decrease. We can see that the RWM model reaches lower MCTs values as the SMM model. The SMM model is able to maintain MCTs values higher because of the social behavior of the nodes.

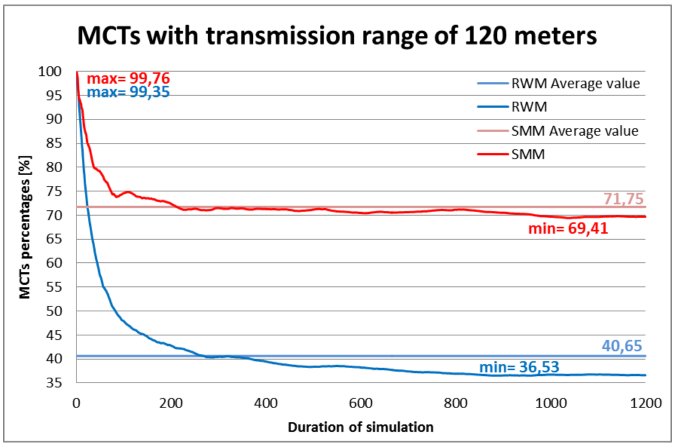


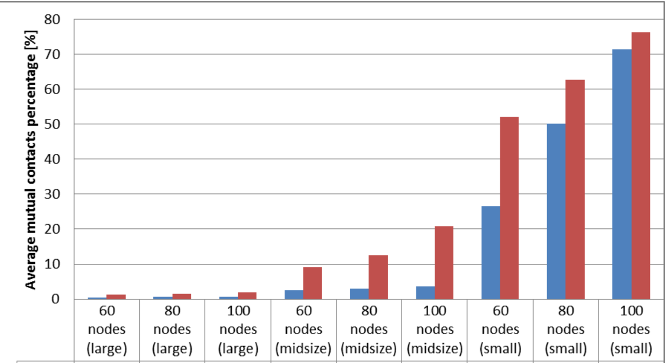
Fig.12 Simulation results of the MCTs with 120m transmission range

B. Statistical simulation results II

We have decided to take average values of close and mutual contacts that were obtained from measurements and compare their values to see how the mobility models managed to get on in the simulation scenarios.

We have to inform that the average values offer us a little skewed look at the values in which the results can vary, because they are largely influenced by the initial deployment of the nodes in the beginning of the simulation, which can in some cases have a big impact on the calculation of the results and thus increase or decrease these values.

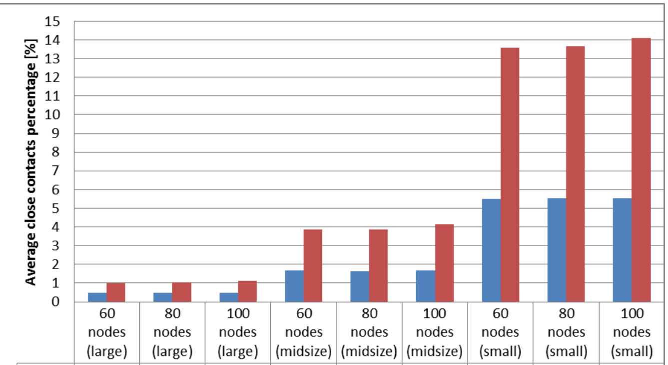
The Fig.13 shows performance of RWM and SMM models in terms of average MCTs. It is possible to see that SMM model outperformed RWM model in the all measured scenarios.



	60 nodes (large)	80 nodes (large)	100 nodes (large)	60 nodes (midsize)	80 nodes (midsize)	100 nodes (midsize)	60 nodes (small)	80 nodes (small)	100 nodes (small)
RWM	0,5378	0,5826	0,6188	2,505	2,9351	3,5821	26,6463	50,0739	71,3503
SMM	1,2967	1,5354	1,94	9,1372	12,4781	20,9071	52,0945	62,7658	76,2125

Fig.13 Results of average MCTs percentages

In the Fig.14 it is possible to see performance of given mobility models in the terms of average CCTs. We can see that here as well SMM model outperformed the RWM model and even the several fold.



	60 nodes (large)	80 nodes (large)	100 nodes (large)	60 nodes (midsize)	80 nodes (midsize)	100 nodes (midsize)	60 nodes (small)	80 nodes (small)	100 nodes (small)
RWM	0,4749	0,4887	0,4941	1,6572	1,6475	1,6634	5,5149	5,525	5,5258
SMM	0,9955	1,0437	1,1178	3,8795	3,8823	4,1396	13,5904	13,662	14,1099

Fig.14 Results of average CCTs percentages

After the evaluation of all the statistical results that we have decided to carry out, we came to the conclusion that designed SMM model suits better for future testing of MANET routing protocols because of higher nodes connectivity and more social behavior of the model.

C. Additional simulation results

Apart from the statistical results, we are also able to obtain "*non-statistical*" results that have redeemable values exactly for the type of the simulation that we have decided to investigate. We have created a simple simulation scenario as an example, for better understanding. We have set a bordered 150x150m simulation area that consist 10 nodes with speed in predefined range [5-10 km/h] and transmission range of 40m. Duration of simulation was set to one hour and nodes are set to have social behavior.

Results from this simulation that we can obtain in addition to those statistical results, are exact durations of MCTs and CCTs between various nodes. Symmetric matrix C_m represents output of the durations of the MCTs between the nodes and symmetric matrix C_c represents durations of CCTs between nodes.

$$C_m = \begin{bmatrix} 0 & 1701 & 1493 & 1190 & 1068 & 1909 & 1393 & 1000 & 1682 & 1838 \\ 1701 & 0 & 1731 & 1422 & 1288 & 1452 & 1514 & 1827 & 1795 & 1412 \\ 1493 & 1731 & 0 & 1038 & 1478 & 1437 & 1416 & 1495 & 1499 & 1454 \\ 1190 & 1422 & 1038 & 0 & 1078 & 1209 & 859 & 1577 & 1356 & 877 \\ 1068 & 1288 & 1478 & 1078 & 0 & 1481 & 1233 & 839 & 1282 & 971 \\ 1909 & 1452 & 1437 & 1209 & 1481 & 0 & 1264 & 1037 & 1676 & 1301 \\ 1393 & 1514 & 1416 & 859 & 1233 & 1264 & 0 & 1282 & 1513 & 1150 \\ 1000 & 1827 & 1495 & 1577 & 839 & 1037 & 1282 & 0 & 1096 & 1157 \\ 1682 & 1795 & 1499 & 1356 & 1282 & 1676 & 1513 & 1096 & 0 & 1329 \\ 1838 & 1412 & 1454 & 877 & 971 & 1301 & 1150 & 1157 & 1329 & 0 \end{bmatrix}$$

$$C_c = \begin{bmatrix} 0 & 901 & 740 & 861 & 608 & 1262 & 836 & 329 & 1350 & 1518 \\ 901 & 0 & 1090 & 804 & 868 & 1103 & 1164 & 1690 & 1467 & 1024 \\ 740 & 1090 & 0 & 687 & 1228 & 818 & 978 & 1091 & 762 & 1124 \\ 861 & 804 & 687 & 0 & 832 & 735 & 518 & 1209 & 866 & 632 \\ 608 & 868 & 1228 & 832 & 0 & 1130 & 878 & 534 & 702 & 483 \\ 1262 & 1103 & 818 & 735 & 1130 & 0 & 924 & 522 & 1171 & 972 \\ 836 & 1164 & 978 & 518 & 878 & 924 & 0 & 956 & 970 & 653 \\ 329 & 1690 & 1091 & 1209 & 534 & 522 & 956 & 0 & 548 & 700 \\ 1350 & 1467 & 762 & 866 & 702 & 1171 & 970 & 548 & 0 & 841 \\ 1518 & 1024 & 1124 & 632 & 483 & 972 & 653 & 700 & 841 & 0 \end{bmatrix}$$

As was mentioned, we set the duration of simulation to 1 hour, which is 3600 seconds. Numbers in the matrices represents duration of contacts in the seconds. Duration of CCT is always less or equal to the duration of MCT. In our simulations we also observe the ratio of how many percent of MCT duration between nodes makes CCT. This ratio can be seen in 10×10 matrix R.

$$R = \begin{bmatrix} 0 & 52{,}97 & 49{,}56 & 72{,}35 & 56{,}93 & 66{,}11 & 60{,}01 & 32{,}90 & 80{,}26 & 82{,}59 \\ 52{,}97 & 0 & 62{,}97 & 56{,}54 & 67{,}39 & 75{,}96 & 76{,}88 & 92{,}50 & 81{,}73 & 72{,}52 \\ 49{,}56 & 62{,}97 & 0 & 66{,}19 & 83{,}09 & 56{,}92 & 69{,}07 & 72{,}98 & 50{,}83 & 77{,}30 \\ 72{,}35 & 56{,}54 & 66{,}19 & 0 & 77{,}18 & 60{,}79 & 60{,}30 & 76{,}67 & 63{,}86 & 72{,}06 \\ 56{,}93 & 67{,}39 & 83{,}09 & 77{,}18 & 0 & 76{,}30 & 71{,}21 & 63{,}65 & 54{,}76 & 49{,}74 \\ 66{,}11 & 75{,}96 & 56{,}92 & 60{,}79 & 76{,}30 & 0 & 73{,}10 & 50{,}34 & 69{,}87 & 74{,}71 \\ 60{,}01 & 76{,}88 & 69{,}07 & 60{,}30 & 71{,}21 & 73{,}10 & 0 & 74{,}57 & 64{,}11 & 56{,}78 \\ 32{,}90 & 92{,}50 & 72{,}98 & 76{,}67 & 63{,}65 & 50{,}34 & 74{,}57 & 0 & 50{,}00 & 60{,}50 \\ 80{,}26 & 81{,}73 & 50{,}83 & 63{,}86 & 54{,}76 & 69{,}87 & 64{,}11 & 50{,}00 & 0 & 63{,}28 \\ 82{,}59 & 72{,}52 & 77{,}30 & 72{,}06 & 49{,}74 & 74{,}71 & 56{,}78 & 60{,}50 & 63{,}28 & 0 \end{bmatrix}$$

Besides that, we also count, how many times nodes connect and reconnect during the simulation. This information is also indicator of the social relations between the nodes. We can see results that were measured in this designed example of simulation scenario in the matrix P.

$$P = \begin{bmatrix} 0 & 9 & 9 & 10 & 11 & 12 & 8 & 5 & 9 & 16 \\ 9 & 0 & 11 & 7 & 9 & 13 & 11 & 11 & 13 & 13 \\ 9 & 11 & 0 & 15 & 17 & 13 & 14 & 10 & 7 & 13 \\ 10 & 7 & 15 & 0 & 14 & 15 & 13 & 13 & 10 & 9 \\ 11 & 9 & 17 & 14 & 0 & 15 & 13 & 11 & 10 & 8 \\ 12 & 13 & 13 & 15 & 15 & 0 & 14 & 18 & 11 & 11 \\ 8 & 11 & 14 & 13 & 13 & 14 & 0 & 14 & 13 & 10 \\ 5 & 11 & 10 & 13 & 11 & 18 & 14 & 0 & 8 & 14 \\ 9 & 13 & 7 & 10 & 10 & 11 & 13 & 8 & 0 & 12 \\ 16 & 13 & 13 & 9 & 8 & 11 & 10 & 14 & 12 & 0 \end{bmatrix}$$

Just as we have two types of contacts, we also have two types of inter-contacts, namely mutual ICT durations and close ICT durations. They are also measured in all simulation scenarios and in our case we can see them in the forms of the matrices; I_m for mutual ICT and I_c for close ICT durations.

$$I_m = \begin{bmatrix} 3600 & 1899 & 2107 & 2410 & 2532 & 1691 & 2207 & 2600 & 1918 & 1762 \\ 1899 & 3600 & 1869 & 2178 & 2312 & 2148 & 2086 & 1773 & 1805 & 2188 \\ 2107 & 1869 & 3600 & 2562 & 2122 & 2163 & 2184 & 2105 & 2101 & 2146 \\ 2410 & 2178 & 2562 & 3600 & 2522 & 2391 & 2741 & 2023 & 2244 & 2723 \\ 2532 & 2312 & 2122 & 2522 & 3600 & 2119 & 2367 & 2761 & 2318 & 2629 \\ 1691 & 2148 & 2163 & 2391 & 2119 & 3600 & 2336 & 2563 & 1924 & 2299 \\ 2207 & 2086 & 2184 & 2741 & 2367 & 2336 & 3600 & 2318 & 2087 & 2450 \\ 2600 & 1173 & 2105 & 2023 & 2761 & 2563 & 2318 & 3600 & 2504 & 2443 \\ 1918 & 1805 & 2101 & 2244 & 2318 & 1924 & 2087 & 2504 & 3600 & 2271 \\ 1762 & 2188 & 2146 & 2723 & 2629 & 2299 & 2450 & 2443 & 2271 & 3600 \end{bmatrix}$$

$$I_c = \begin{bmatrix} 3600 & 2699 & 2860 & 2739 & 2992 & 2338 & 2764 & 3271 & 2250 & 2082 \\ 2699 & 3600 & 2510 & 2796 & 2732 & 2497 & 2436 & 1910 & 2133 & 2576 \\ 2860 & 2510 & 3600 & 2913 & 2372 & 2782 & 2622 & 2509 & 2838 & 2476 \\ 2739 & 2796 & 2913 & 3600 & 2768 & 2865 & 3082 & 2391 & 2734 & 2968 \\ 2992 & 2732 & 2372 & 2768 & 3600 & 2470 & 2722 & 3066 & 2734 & 2968 \\ 2338 & 2497 & 2782 & 2865 & 2865 & 3600 & 2676 & 3078 & 2429 & 2628 \\ 2764 & 2436 & 2622 & 3082 & 3082 & 2676 & 3600 & 2644 & 2630 & 2947 \\ 3271 & 1910 & 2509 & 2391 & 2391 & 3078 & 2644 & 3600 & 3052 & 2900 \\ 2250 & 2133 & 2838 & 2734 & 2734 & 2429 & 2630 & 3052 & 3600 & 2759 \\ 2082 & 2576 & 2476 & 2968 & 2968 & 2628 & 2947 & 2900 & 2759 & 3600 \end{bmatrix}$$

Our results also offer the graphical view on the C_m, C_c, I_m and I_c matrices. Graphs show us, how many percent of the total time nodes spend in the given type of CT and ICT with each other. In the Fig.15 it is possible to see graphically, how many percent of the total time nodes spend with each other in MCTs and mutual ICTs.

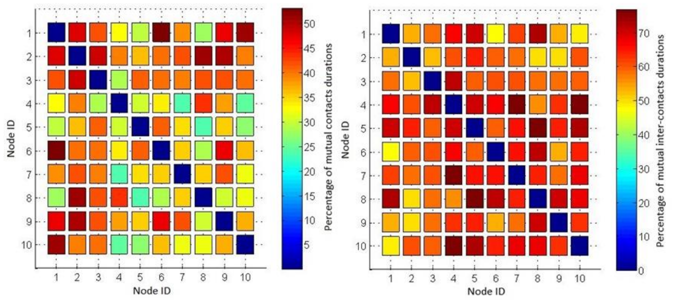


Fig.15 Graphical view of the MCT and mutual ICT durations

In the Fig.16 we can see the graphical representation of CCTs and close ICTs durations that were measured in this simulation scenario.

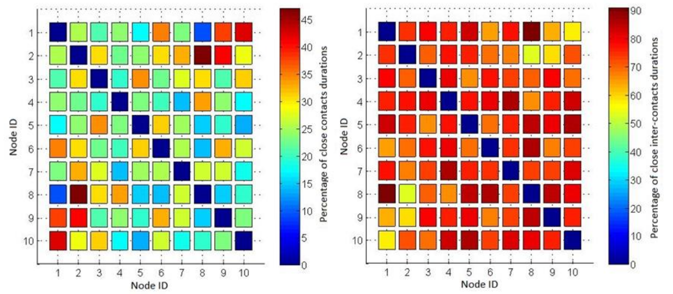


Fig.16 Graphical view of the CCTs and close ICTs durations

Because of the random initial deployment of the nodes and their random movement in the models, these results could not be measured statistically. However if we want to understand the behavior of nodes in some kind of specific scenario, as we did in our example, these results may help to understand how the nodes behaved and with who they were in the contact during the simulation.

In our case, we can say, for example that *node 1* and *node 2* were in MCT for 1701 seconds from total 3600 seconds. 52,97% of those MCTs was CCT. They were in contact 9 times and time, during which they weren't in the MCT, was 1899 seconds.

VI. CONCLUSION

We managed to create a mobility model based on social behavior which is, in this case, the formation of social groups in which the nodes decides whether they remain in the group and travels in it around the topology, or leave it and continue in its own movement. The Random Walk Mobility model (RWM) was created to serve as a basis for the designed model with social behavior. Designed model was called Social Mobility Model, abbreviated SMM and it is actually enhanced RWM model.

Statistical results that we obtained using MATLAB simulations show that the proposed SMM model outperformed RWM model in all sizes of topology and also in all deployments of the nodes. Nodes were moving around the topology in large social groups and thus they maintain greater percentage of contacts than in the RWM model. Besides statistical results we can observe in the simulations behavior and movement of the nodes more precisely, because in every simulation we are also collecting information about durations of the contacts between nodes, frequency of their meetings and also the duration of their inter-contacts. These results could not be statistically observed, since none of the simulations could be exactly repeated several times, because nodes will still be differently placed and they will always have random motion, when they will travel by themselves. Therefore, statistical results of these measurements would not have redeemable values.

We propose to continue in the research and improvements of this mobility model, since has the great potential and it is extremely necessary for the testing of MANET routing protocols in the most effective way, as possible. One of the improvements that could be done on this model is to adjust the placement of the nodes in order to remember their initial positions. Also nice idea is that once given the simulation carry out, the next simulations would then be based on the results of the first simulation. Social behavior of the nodes could then be affected by the history of their pervious contacts. Furthermore, create some obstacles in the topology, or even create some environment that mimics real world in order to make simulation more complicated for movement of the nodes. Also, in the real world is really rare symmetrical friendship. The fact that someone is our very close friend does not mean that we are also his. Therefore, it would be well to devote closer to the asymmetric relationships between individuals and also it would be interesting to see how the designed model would stand in compare to the real-trace models.

REFERENCES

[1] HUI, P.: *People are the network: experimental design and evaluation of social-based forwarding algorithms.* Technical report, UCAM-CL-TR-713, University of Cambridge, 2007 [Accessed 18.10.2014].

[2] PHITHAKKTNUKOON, S - SMOREDA, Z - OLIVIER, P.: *Socio-Geography of Human Mobility: A Study Using Longitudinal Mobile Phone Data.* PLoS ONE. 2012; 7(6):e39253 [Accessed 19.10.2014].

[3] EINSTEIN, A.: *Investigations on the Theory of the Brownian movement*. Dover Publications, 1956. [Accessed 19.10.2014].

[4] MUSOLESI, M. - MASCOLO, C.: *Mobility Models for Systems Evaluation*. A Survey. In: Garbinato B, Miranda H, Rodrigues L, editors. Middleware for Network Eccentric and Mobile Applications: Springer; 2009. p. 43-62 [Accessed 16.10.2014].

[5] ASCHENBRUCK, N. - MUNJAL, A. - CAMP, T.: *Trace-based mobility modeling for multi-hop wireless networks*. Computer Communications. 2011; 34(6):704-714 [Accessed 16.10.2014].

[6] ANDREA, R. - SOFIA, RC.: *A Survey on Mobility Models for Wireless Networks*. SITI, University Lusófona, 2011. [Accessed 16.10.2014].

[7] KARAMSHUK, D. - BOLDRINI, C. - CONTI M. – PASSARELLA, A.: *Human mobility models for opportunistic networks*. IEEE

Communications Magazine. 2011; 49(12):157-165 [Accessed 16.10.2014].

[8] MUNJAL, A. - CAMP, T. - ASCHENBRUCK, N.: *Changing Trends in Modeling Mobility*. Journal of Electrical and Computer Engineering. 2012:1-16. [Accessed 16.10.2014].

[9] PIROZMAND, P. - WU, G. - JEDARI, B. - Xia, F.: *Human mobility in opportunistic networks: Characteristics, models and prediction methods*. Volume 42, Pages 45–58, June 2014 [Accessed 16.10.2014].

[10] HSU, W.-J. - SPYROPULOS, T. - PSOUNIS, K. - HELMY, A.: *Modeling spatial and temporal dependencies of user mobility in wireless mobile networks*. IEEE/ACM Trans. Netw., vol. 17, no. 5, pp. 1564–1577, 2009 [Accessed 6.11.2014].

[11] NGUYEN, T. - SZYMANSKI, BK.: *Using Location-Based Social Networks to Validate Human Mobility and Relationships Models*. Proceeding of the 2012 IEEE/ACM International Conference Advances in Social Networks Analysis and Mining (ASONAM), 2012. p. 1215-1221. [Accessed 1.11.2014].

[12] ZE, L. - HAIYING, S.: SEDUM: *Exploiting Social Networks in Utility--Based Distributed Routing for DTNs*. IEEE Transactions on Computers. 2013; 62(1):83-97. [Accessed 1.11.2014].

[13] LU, X. - WETTER, E. - BHARTI, N. - TATEM, AJ. - BENGTSSON, L.: *Approaching the Limit of Predictability in Human Mobility*. Nature. 2013; 3. [Accessed 20.11.2014].

[14] SONG, C. - QU, Z. - BLUMM, N. - BARABÁSI, A-L.: *Limits of Predictability in Human Mobility*. Science. 2010b; 327(5968):1018-1021. [Accessed 2.1.2015].

[15] KLEDER, M.: "*All Pairs Shortest Path*" Graph Solver. 26 Dec 2007. Available at: <http://www.mathworks.com/matlabcentral/fileexchange/8808--all-pairs-shortest-path--graph-solver>. [Accessed 5.1.2015].

Source Code Projections and Projectional Code Editors

[1]*Ján JUHÁR (1st year),*
Supervisor: [2]Liberios VOKOROKOS

[1,2]Dept. of Computers and Informatics, FEI TU of Košice, Slovak Republic

[1]jan.juhar@tuke.sk, [2]liberios.vokorokos@tuke.sk

***Abstract*—The goal of this paper is to give an overview of current research in the field of projectional source code editing with regards to types of source code projections and existing projectional code editors. We present motives behind projectional source code editing and describe established types of source code projections based on reviewed research conducted in the recent years. We also overview some projectional editors that resulted from this research. Possible directions for future research are given at the end of the paper.**

***Keywords*—development environments, program comprehension, projectional editing, software concerns, source code editors**

I. Introduction

Contemporary Integrated Development Environments (IDEs) present main toolset of a programmer. They contain code editors and browsers, build automation tools, and debuggers that together aim for supporting all phases of the software life cycle and for higher productivity of developers.

Of all the tasks developers perform during software development and maintenance, program comprehension tends to be the most dominant one [1], [2]. Program comprehension is a process in which a programmer analyzes the source code and uses extracted knowledge to create a mental model of the system's behavior. In other words, it is a mapping of the solution domain to the problem domain. Although IDEs are equipped with tools that bring some automation to this cognitive process (e.g., inspectors, refactoring support), it still involves considerable amount of manual work. One of the problems that programmers must deal with during this process is the static structure of the source code.

Current approaches to the separation of concerns like procedures, object–oriented or aspect–oriented designs operate at the source code level [3]. When a programmer chooses some object-oriented design pattern or creates an aspect of a crosscutting concern, he or she is tying this decision to the base structure of the program. However, as argued in [4], structure of the source code that is considered as *good* by one programmer may be considered as *bad* one by the other. Even when considering only the author of the code, evolving experiences or system requirements may change their opinion on what is the best structure for dealing with a particular problem [5]. And new programmer must adapt to the chosen code design and to the mental model of the previous one. Thus, commonly used techniques of concern separation are insufficient to overcome the limitations that static code structure poses on the system decomposition [6].

As proposed in [7], an alternative approach to direct source code editing can be a manipulation of a base system definition through *projections* – a projectional editing. These projections should provide multiple, dynamically build views of the system according to the programmer's current needs, while maintaining the single underlying definition of the implemented system. In this paper we give an overview of existing types of projections as well as of source code editors that utilize such projections.

II. Types of Projections

On the basis of reviewed research in the field of projectional source-code editing we can identify the following two (though overlapping) types of projections:

- *IDE-based projections* that alter presentation form of the language to meet a specific demand, and
- *Concern-oriented projections* that create view of the source code based on the concerns it contains.

In this section we describe these types of projections.

A. IDE-based Projections

The larger and more general group of projections falls under the IDE-based type. It is based on the concept of multiple representations of the system with a single underlying base structure [7]. Typically, the base structure is the source code.

Modern IDE works with a language on three levels [4], namely (i) *notation* or *concrete syntax* that is used for serialization, (ii) *model* that represents the parsed abstract syntax tree (AST), and (iii) *view* that shows presentation syntax to the programmer. The main purpose for creating an AST by parsing the source code in IDE is to create a representation of the system that can be used for operations like code completion or refactoring [7]. Manual changes of the source code trigger re-parsing of the AST and any changes made directly to AST through various IDE operations are mapped back to the concrete syntax. The content of the editor can thus be considered to be a projection of the AST.

The AST can be utilized further to create well-known projections like tree of classes and methods in the current file, list of all usages of a given program element, and others. However, these projections are often not directly modifiable – they update only in reaction to changes to the source code. In the section III we present some of the existing tools that add projectional features to the IDE editors and thus create editable projections.

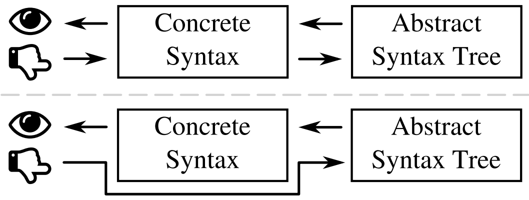


Fig. 1. Difference between parser-based (top) and AST-based editors [8]

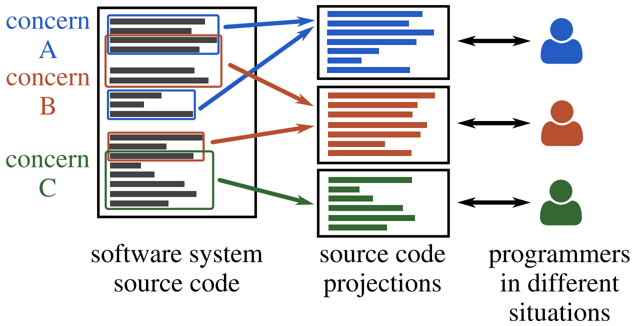


Fig. 2. The concept of the concern-oriented projections [4]

It should be noted that there is one highly discussed (e.g., in [8], [9], [10], [11]) group of editable projections. Their basic idea is that instead of using textual editor where programmer types in a sequence of characters that then needs to be parsed to build an AST, they allow programmer to directly edit the AST (see Fig. 1). In order to achieve this, a projection is used to create a user-editable representation of the AST.

An advantage of such approach is the elimination of the whole parsing phase. Basically, a programmer directly creates an AST by selecting required language concept and filling in the variable parts. Consequently, there is no place for semantic ambiguity that often poses a problem for design of a language that needs to be parsable [8], [12].

Another benefit lies in the advanced ability to compose multiple different languages and/or forms of program representation (e.g., text, tables, diagrams) into one view. This idea was promoted with *intentional programming* [9], [10] which advocated separation of concerns through generative programming and use of *Language Workbenches* – described in [13] as tools with support for language-oriented programming.

The nature of AST-based editor, however, poses a number of usability drawbacks to its users. These drawbacks are discussed to great detail in [8], where three general categories of issues are identified: (i) efficient textual code entering (how to support non-linear typing?), (ii) code selection and modification (how to introduce cross-tree parentheses in enforced tree structure?), and (iii) infrastructure integration (how to integrate with Version Control System tools that do not work well with serialized AST?) They evaluate mitigation techniques of these issues used by JetBrains MPS[1] – the state-of-the-art projectional IDE – to which we will get back in section III-C.

B. Concern-oriented Projections

Concern-oriented source code projections tackle the projection idea from a different perspective. They do not necessarily try to change presentation syntax of a language, but, as defined in [4], [5], are rather interested in providing multiple presentations of the source code of a particular system. Or, in other words, they give multiple dynamic structures to existing code base [5]. The ultimate goal of these projections is to show concerns (intentions) of a system that are scattered throughout the base structure in a more coherent way and consequently help the programmer to comprehend the source code. This is captured in Fig. 2.

Concern-oriented projections are typically, too, incorporated into IDEs, but have very specific function related to software concerns. In order for them to work, there must be some way of extracting concerns from the system for their subsequent use in projection building. There have been developed techniques for automated program analysis (e.g, reverse engineering) and for feature location. However, reverse engineering usually struggles with providing higher-level views as many concerns are lost during the transformation [4]. As for feature location, detailed survey of existing techniques and tools is presented in [14]. Authors of the survey state that it remains an open question to what extent these techniques make developers more productive.

Some of the tools for feature location presented in the survey have projectional capabilities, but available projections, yet again, do not allow direct modification of the underlying code. They are useful only either for investigation of scattered concerns or for storing the mappings between found concerns and source code (but, interestingly, never for both). To give an example, *Feature Exploration and Analysis Tool* [15] allows a programmer to navigate over extracted model of a Java program – so called *Concern Graph* – and to build up subset of the model according to required concern. Another tool is *ConcernMapper* [16]. This one is similar to tree view projection of classes and methods as it shows tree-like representation of the concerns. Both these tools are implemented as plugins for Eclipse IDE[2].

On the other side, in [4] authors promote explicit recording of concerns in the form of embedded metadata, e.g., source code annotations, structured comments or code conventions. Their projectional code editor that utilizes such metadata is presented in III-B.

III. Projectional Code Editors

In this section, we present an overview of some existing code editors that have projectional capabilities. We focus on their basic functionality and report results of the studies that were conducted with goal to evaluate their contribution.

A. Code Bubbles

Code bubbles (CB) is an user interface extension for the Eclipse IDE based on collections of editable code fragments – “bubbles”. It was introduced in [17] and further extended in [18]. Each bubble represents a fully editable view of a method or collection of class member variables that can be arranged with other bubbles to form working sets on a large, pannable 2D virtual space. This way it aims to overcome the file-based nature of contemporary IDEs where programmers spend considerable amount of time navigating the code. Furthermore, customized views should help developers to more quickly understand the code. Created working sets can be saved for later use or shared with other developers.

[1]http://jetbrains.com/mps

[2]http://eclipse.org

Fig. 3. Three methods opened in the *Code bubbles* interface

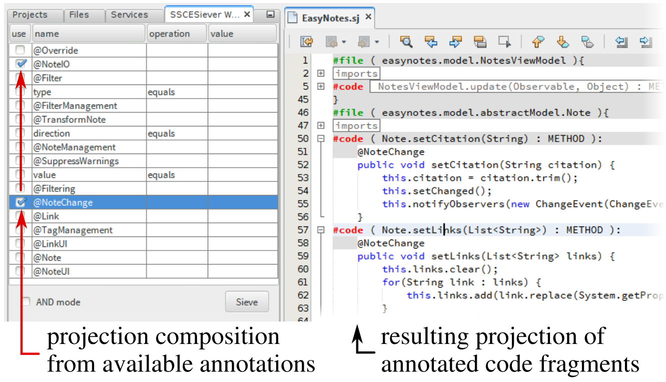


Fig. 4. Projection composition in the Sieve Source Code Editor

A fragment of the CB interface is shown in Fig. 3. It shows three methods opened in bubbles. The method on the left was opened first through a pop-up search box and the rest were opened through *Go to implementation* action. An arrow connects the line with a call to a method with the new bubble. The header of each bubble shows the name of the contained method. The views of related methods that can be explored through reference links and arranged by users fits our notion of concern-oriented code projections.

Qualitative evaluations of the proposed editor from [17] and [18] revealed a great amount of enthusiasm from the participating programmers. Participants positively reacted to the ability to see many function concurrently, to freely organizable working sets, integrated debugging capabilities or session management. However, they felt that plain files are still more useful in early stages of development. With the quantitative study in [17], authors focused on code understanding by comparing task completion time and success rate of developers working with CB to those using standard Eclipse editor (EE). They discovered that one task was completed significantly faster and with higher success rate by users of CB, while the other one has no significant difference between CB and EE users. Second quantitative study in [18] was focused on CB interface efficiency. Results showed that CB interface could show more functions simultaneously than EE. Yet, there were some cases when EE required less UI operations than CB.

One of the biggest limitation of the presented studies was their short time span, ranging from 45 to 90 minutes. Authors' suggestions for future work include extending CB interface to other areas of the IDE and more usage evaluations.

B. Sieve Source Code Editor

Sieve Source Code Editor (SSCE) is a concern-oriented source code projections tool implemented as an extension of the NetBeans IDE[3]. It was presented in [4]. It promotes explicit expression of concerns in a structured form directly in the source code as a form of metadata. Version of the SSCE tool presented in [4] used structured comments that started with `SsceIntent` keyword in the role of metadata capturing. Newer version provides an API for inclusion of multiple formats of concern capturing and comes with built-in support for annotation-based concern identification.

Construction of annotation-based projection with the SSCE tool is displayed in Fig. 4. The user interface consists of two main components: the table of available annotations and the code editor that displays constructed projections. In the table the user can select annotations that will be used to build the projection according to the concern(s) of interest. After confirming the selection with *Sieve* button, the resulting projection is displayed in the code editor. The content of a projection consists of program elements that were annotated with selected annotations, even if they occur in multiple files.

An experiment evaluating code comprehension with SSCE projections was presented in [4]. Experiment consisted of fixing a bug and extending functionality in unfamiliar code base (with pre-embedded concern annotations) by a group of 40 students. These tasks were designed to be a matter of program comprehension and not of algorithmic character. In general, results of the experiment showed positive feedback for projections, mainly with regard to navigation around the code. However, building of a program with concern-related metadata encoded to the comments was outside of experiment scope, mainly for restricted time frame. Therefore, the full cost of creating such projections is not reflected in the results.

C. JetBrains MPS

A prominent position among AST-based projectional editors is taken by JetBrains' Meta Programming System, or MPS for short. It represents a universal platform for designing domain-specific languages, their supporting tools and environments (thus the term *language workbench*), with focus on Language Oriented Programming that aims to "allow programmers to define languages as easily as they can write programs" [11]. Actual programs will be then written as a composition of custom DSLs, each suitable for different part of a program.

We have already described three levels at which IDEs work with language in section II-A. Similarly, MPS also works on three language levels that are described in [11], namely *structure*, *editor*, and *semantic* languages. Structure language defines abstract syntax of the language – supported concepts of the language and their possible arrangement. Editor defines concrete syntax along with means for its rendering and editing. And finally, semantic language defines how the language should be interpreted or transformed into executable form.

An example of custom DSL in MPS editor with view on AST behind the projection of selected code is shown in Fig. 5.

We already discussed issues that are typically associated with AST-based projectional editors in section II-A. According to [8], MPS tries to mitigate these issues with mixed success. For example, required manual disambiguation of entered code (in cases when multiple language constructs are matching the

[3]http://netbeans.org

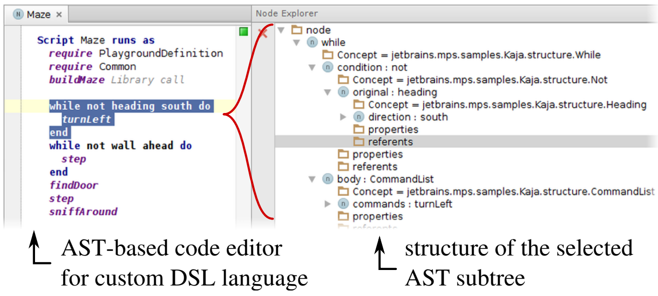


Fig. 5. Code editor and view of the AST in the JetBrains MPS

input) is resolved by code completion or context constrains. But only valid subtrees can be selected and integration with external text-based tools is complicated and rather impractical. However, results of the case study also presented in [8] shows that experienced programmers find MPS overall efficient for every-day work after the initial phase of getting used to it.

D. Registration-based Abstractions Editor

With the goal to help programmers better understand the code, an editor supporting *registration-based language abstractions* was introduced in [19]. Authors use the term *registration* to mean work needed to be done in order to understand part of a program. They argue that the common way in which innovations are introduced into the programming languages makes them tyrannical, mainly because programmers must understand them and all supporting tools must be updated.

The proposed *registration-based abstractions* (RBA) code editor is able to change presentation syntax of code fragments and thus add new language abstractions. As this is done only on the presentation level of the language, no changes in supporting tools are needed and programmers may choose which abstractions they want to use. From our perspective, RBA editor fits the concept of an IDE-based projection.

In a quantitative study presented in [20] authors observed understandability of selected RBAs by considering how often study participants needed to reveal source code behind them. They found out that majority of participants were able to understand the RBAs after seeing them less than five times. However, it is not clear how significant is such a result, as it basically states that participants were able to learn new language concepts. Another observation was the lack of significant difference between the rate at which participants revealed RBAs before and after a break between study tasks. Authors were unable to give conclusive reason behind this observation, given the overall short duration of the study.

IV. Conclusion and Future Research

Each code editor presented in the previous section has a different use of the projection idea. Helping with the program comprehension can be considered the most unifying aspect, with possible exception of MPS that is, after all, more focused on providing tailored notation for the system. But all of them provide valuable insight into the possibilities projectional editing can enable. However, the overview also brings forward some areas that require further research.

In our next work we want to expand on the idea of concern-oriented projections with connection to the development environments that could take advantage of such projections. As we have seen in the case of Code bubbles and SSCE, they work on the level of whole methods or list of member variables, but that do not necessarily match the level at which concerns are distributed throughout the code. Therefore we plan to explore how concerns detected in the system could be presented to the developers in the most faithful form that would provide multiple editable concern-based views on the source code of a system. We also want to take a look at ways in which concern-oriented projections could be used to create views composed of multiple languages. The presented editors do not deal with this case although it could help to convey concerns in many software systems that use such composition.

Acknowledgment

This work was supported by the project KEGA 008TUKE-4/2013 and by the Slovak Research and Development Agency under the contract No. APVV-0008-10.

References

[1] W. Maalej, R. Tiarks, T. Roehm, and R. Koschke, "On the Comprehension of Program Comprehension," *ACM Transactions on Software Engineering and Methodology*, vol. 23, no. 4, pp. 1–37, Aug. 2014.

[2] T. Kosar, M. Mernik, and J. Carver, "The impact of tools supported in integrated-development environments on program comprehension," in *ITI'11*, 2011, pp. 603–608.

[3] G. Kiczales and M. Mezini, "Separation of concerns with procedures, annotations, advice and pointcuts," in *ECOOP '05*, vol. 3586. Springer Berlin Heidelberg, Jul. 2005, pp. 195–213.

[4] J. Porubän and M. Nosáľ, "Leveraging Program Comprehension with Concern-oriented Source Code Projections," *3rd Symposium on Languages, Applications and Technologies. OASIcs*, vol. 38, pp. 35–50, 2014.

[5] M. Nosáľ, J. Porubän, and M. Nosáľ, "Concern-oriented source code projections," 2013, pp. 1541–1544.

[6] M. P. Robillard and G. C. Murphy, "Representing concerns in source code," *ACM Transactions on Software Engineering and Methodology*, vol. 16, no. 1, pp. 3–es, Feb. 2007.

[7] M. Fowler, "Projectional Editing," 2008. [Online]. Available: http://martinfowler.com/bliki/ProjectionalEditing.html

[8] M. Voelter, J. Siegmund, T. Berger, and B. Kolb, "Towards User-Friendly Projectional Editors," in *Software Language Engineering SE - 3*. Springer International Publishing, 2014, vol. 8706, pp. 41–61.

[9] K. Czarnecki and U. W. Eisenecker, *Generative programming: methods, tools, and applications*. ACM Press/Addison-Wesley Publishing Co., May 2000.

[10] C. Simonyi, M. Christerson, and S. Clifford, "Intentional software," *ACM SIGPLAN Notices*, vol. 41, no. 10, p. 451, Oct. 2006.

[11] S. Dmitriev, "Language Oriented Programming: The Next Programming Paradigm," *JetBrains*, november 2004. [Online]. Available: http://www.onboard.jetbrains.com/articles/04/10/lop/mps.pdf

[12] L. Diekmann and L. Tratt, "Eco: A Language Composition Editor," in *Software Language Engineering SE - 5*, 2014, vol. 8706, pp. 82–101.

[13] M. Fowler, "Language Workbenches: The Killer-App for Domain Specific Languages?" 2005. [Online]. Available: http://martinfowler.com/articles/languageWorkbench.html

[14] B. Dit, M. Revelle, M. Gethers, and D. Poshyvanyk, "Feature location in source code: a taxonomy and survey," *Journal of Software: Evolution and Process*, vol. 25, no. 1, pp. 53–95, Jan. 2013.

[15] M. P. Robillard and G. C. Murphy, "Concern graphs: finding and describing concerns using structural program dependencies," in *ICSE 2002*, 2002, pp. 406–416.

[16] M. P. Robillard and F. Weigand-Warr, "ConcernMapper: simple view-based separation of scattered concerns," in *eclipse '05*. New York, New York, USA: ACM Press, Oct. 2005, pp. 65–69.

[17] A. Bragdon, R. Zeleznik, S. P. Reiss, S. Karumuri, W. Cheung, J. Kaplan, C. Coleman, F. Adeputra, and J. J. LaViola, "Code bubbles: a working set-based interface for code understanding and maintenance," in *CHI '10*, 2010, pp. 2503–2512.

[18] A. Bragdon, S. P. Reiss, R. Zeleznik, S. Karumuri, W. Cheung, J. Kaplan, C. Coleman, F. Adeputra, and J. J. LaViola, "Code bubbles: rethinking the user interface paradigm of integrated development environments," in *ICSE '10*, vol. 1, 2010, pp. 455–464.

[19] S. Davis and G. Kiczales, "Registration-based language abstractions," *ACM SIGPLAN Notices*, vol. 45, no. 10, p. 754, Oct. 2010.

[20] J.-J. Nunez and G. Kiczales, "Understanding registration-based abstractions: A quantitative user study," in *ICPC '12*. IEEE, Jun. 2012, pp. 93–102.

Supervised Machine Learning with SVM Algorithm.

Milan KRENDZELÁK (2[st] year)
Supervisor: František JAKAB, PhD

Dept. of Computers and Informatics, FEI TU of Košice, Slovak Republic

krendzelak.m@gmail.com, frantisek.jakab@tuke.sk

Abstract—**Because of the ascent of use of virtual frameworks, ticket frameworks have become somewhat renowned. Addressing the issue tickets to proper individual or unit in the support group has discriminating significance with a specific end goal to give enhanced end client fulfillment while guaranteeing better allocation of support recourses. The task of ticketing to is still physically performed. Particularly everywhere associations, the manual task is not pertinent sufficiently. Numerous decision frameworks which give the client to pick the related classifications or unit inside characterized classes may appear like better, however the frameworks are not helpful in light of those clients, particularly new clients who have never utilized the framework previously, typically have no clue about the related classification or office. Additionally clients would prefer not to fill long ticket structures which are expected to distinguish the issue. In this framework, bag of word methodology, machine learning systems and different calculations which demonstrated execution in content handling are utilized. The suggested technique gives amazing client bolster and helps end-client fulfillment. It diminishes manual endeavors and human slips while guaranteeing high administration levels and enhanced end-client fulfillment.**

I. Introduction

A number of algorithms are used in online ticket support systems. These include the Decision Trees (DT), Generalized Linear Models (GLM), Naive Bayes, SVM, A-priori, k-Means, Non-negative Matrix Factorization, and others [4]. Of uttermost importance to us is the SVM algorithm. This mainly is a Support Vector Machine, which is a classification method in a Supervised Machine Learning approach. This algorithm used for text classification, regressions (binary and multi-class problem), anomalies detection (one class problem). Some other applications might be nested data problems e.g. transaction data or gene expression data analysis and pattern recognition. Distinct versions of SVM use different kernel functions to handle different types of data sets. Linear and Gaussian (nonlinear) kernels are supported, as well. SVM classification attempts to separate the target classes with the widest possible margin. SVM recession attempts to discover an endless function such that the maximum number of data points lie within an epsilon-wide tube round it.

II. Algorithm Overview

SVM algorithm concentrate just on points that are the most hard to distinguish one from the other, while other algorithms pay consideration on all points. The instinct behind the SVM methodology is that the classifier is great at the most difficult correlations, and afterward the classifier will stunningly better at the simple examinations. Perceptrons are built by taking one point at once and changing the isolating line likewise. When the majority of the focuses are divided, the perceptron calculation stops. However it could stop anyplace. There are a number of diverse isolating lines that separate the data. The perceptron's stopping criteria is basic: "separate the focuses and quit enhancing the line when you get 100% division". The perceptron is not expressly advised to discover the best differentiating line. Logistic relapse and linear discriminant models are assembled likewise to Perceptrons. The best isolating line amplifies the separation between the focuses nearest to one another. It's not important to take a glimpse at all of the focuses to do this. Truth be told, consolidating feedback from focuses that are far away can knock the line a little bit far [1]. SVM is built on the standard of structural risk minimization. In direct order, SVM makes a hyper plane that differentiates the information into two sets with the most extreme edge. A hyper plane with the greatest edge has the separations from the hyper plane to focuses when the two sides are equivalent. Scientifically, SVMs take in the sign capacity $f(x) = |wx + b|$, where $\mathbf{w}$ is a weighted vector in $\mathbf{R}^n$. SVM find the hyper plane $y = wx + b$ by differentiating the space Rn into two half-spaces with the greatest edge. Straight SVM can be summed up for non-direct issues. To do as such, the information is mapped into an alternate space **H** and we perform the straight SVM calculation over this new space. SVM has been effectively utilized on TC (text categorization) and they inferred preferable results over other Machine Learning procedures, for example, Naïve Bayes, Decision Trees, and KNN with reference to precision[2].

III. Utilizations

SVM has been utilized inside the Network frameworks to keep away from any interlopers into the system information who anticipate possibly to damage it. We complete different exercises like internet shopping, bill installment, film ticket reservation, air ticket reservation, structure filling for examinations etc. While completing exercises specified above we impart imperative resources so we require some security mechanisms for shielding our valuable resources from harm brought on by interlopers. For shielding and keeping our advantages from meddlesome assaults, Intrusion Detection System is actualized to identify the meddling conduct in the system [1]. Interruption discovery framework is grouped predominantly as host based interruption identification framework which breaks down movement of one single host in the system and system based interruption recognition framework which investigates exercises performed on the system. Method to recognize interruption in the system is sorted as abuse location and irregularity discovery. Abuse identification performs match of current exercises with the assault marks put away and if match happens, assault is recognized. In irregularity identification if any deviation is found is every day profile of client movement, framework is under assault and alert is activated. This is basically actualized utilizing the SVM calculation.

SVM is likewise utilized as a part of text categorization (TC) issues. TC, otherwise called text classification, is the undertaking of consequently sorting an arrangement of reports into classifications (or classes, or subjects) from a predefined set. Such undertaking is identified with IR and ML groups. Computerized text classification instruments are appealing since they free associations from the need of manual arrangement of documentations, which can be excessively costly, or essentially not achievable given the imperatives of the application or the quantity of reports included. TC includes numerous applications, for example, mechanized indexing of experimental articles as per predefined thesauri of specialized terms, recording licenses into patent registries, particular spread of data to data purchasers, computerized populace of various leveled lists of web assets, spam sifting, ID of record type, creation attribution, review coding and even robotized paper reviewing.

Joining example based and machine learning routines aides identify definitions for e-learning purposes which predominantly help make an enhanced e-learning stage. This is mostly done by recognizing the punctuation and afterward doing separating utilizing machine learning strategies. This helps learners to discover definitions in different fields effectively and quicker. Machine adapting additionally finds wide application in Intelligent Adaptive E-Learning Systems [2]. This basically goes for accomplishing learner's movement arrangement by essentially presenting new sorts of adjustment in e-learning frameworks, spanning the regular hole in the middle of substance and correspondence offices. Here, we approach this point utilizing movement mining and grouping.

Machine Learning is likewise connected in Information Mining as a method for improving e-Learning frameworks and courses of action. It upholds methodical adapting and useful realizing, which can adequately control clients through deliberate searching and request. With this capacity, it lives up to expectations a greater number of as an element inquiring about apparatus than static learning material. Then again, the SVM supports productive learning. In spite of the fact that the usefulness of a subject guide is predictable and deliberate, it is additionally achievable for undertaking based learning. From the valuable perspective, learners need assets from different hotspots with the end goal of autonomous examination. The instrument can suffice the investigation of different learning styles, inclinations of premiums, and expert capacities [3]. All the more significantly, this direction is not given by educators working in the classrooms, yet by a self-governing framework which is backed by an expert group with a wide cluster of assets. It transforms learning into data guided element. In this manner, this material helps clients to "find" new learning by exhibiting express and verifiable information so they find themselves able to see thoughts and ideas that are generally surprising. This procedure matches the essential standards of constructivist learning. Machine Learning can include a mixture of best calculations to draw course proposals frameworks. It additionally guarantees a methodology focused around feed forward neural systems to gather the learning styles of understudies consequently. The preference of this methodology is twofold. Initially, a programmed component for style distinction encourages the get-together of data about learning inclination, making it intangible to understudies. Second, the proposed calculation utilizes the late history of framework utilization so frameworks utilizing this methodology can perceive changes in learning styles or some of their measurements about whether [9].

IV. Conclusion

SVM approach has additionally been actualized in the saving money segment predominantly in assessment of credit[5]. The structure of SVM has numerous reckoning favorable circumstances, for example, exceptional direction at a finite sample and irrelevance between complexity of algorithm and the sample dimensions. A genuine MasterCard information experiment demonstrates that SVM technique has extraordinary appraisal capacity. Contrasted with the technique that is right now utilized by a noteworthy Chinese bank, the SVM system has an incredible potential prevalence in foreseeing accuracy.

The fundamental favorable circumstances of SVM can be summarized as:

1. SVM is utilized as a part of the circumstance of limited example information. It intends to get the ideal arrangement in view of the present data as opposed to the ideal worth when the quantity of test has a tendency to be interminable.

2. The calculation is at last changed into the advancement of quadratic project. Hypothetically, it will get a worldwide advancement esteem, which takes care of the unavoidable neighborhood improvement issue while utilizing neural systems.

3. The algorithm performs a nonlinear mapping from the first information space into some high measurement gimmick space, in which it builds a direct discriminant capacity to supplant the nonlinear capacities in the first information space. This extraordinary character guarantees that SVM has great speculation capacity.

All these different ways are supported by ticketing systems. It's thus evident that SVM is the best algorithm to be used in ticketing and other applications as compared to other approaches such as the Naïve Bayes. It offers greatest accuracy as compared to others.

References

[1] Atluri, V. (2008). *Data and applications security XXII 22nd Annual IFIP WG 11.3 Working Conference on Data and Applications Security, London, UK, July 13-16, 2008 : Proceedings. Berlin: Springer.*

[2] Marjanović, M., Kovačević, M., Bajat, B., & Voženílek, V. (2012). *Landslide susceptibility assessment using SVM machine learning algorithm. Engineering Geology, 225-234.*

[3] Mladenov, V. (2013). *Artificial neural networks and machine learning -- ICANN 2013: 23rd International Conference on Artificial Neural Networks Sofia, Bulgaria, September 10-13, 2013. Proceedings.*

[4] Wu, X. (2009). *The top ten algorithms in data mining. Boca Raton: CRC Press.*

[5] Yu, L. (2008). *Bio-inspired credit risk analysis: Computational intelligence with support vector machines. Berlin: Springer Verlag.*

Techno-economical aspects of spectrum sharing in the cognitive radio networks

[1]*Ján PASTIRČÁK (2^{nd} year),*
Supervisor: [2]*Dušan KOCUR*

[1,2]Dept. of Electronics and Multimedia Communications, FEI TU of Košice, Slovak Republic

[1]jan.pastircak@tuke.sk, [2]dusan.kocur@tuke.sk

***Abstract*—This paper presents a summary of the authors' recent research work in the area of joint design of spectrum sensing and sharing mechanisms for dynamic spectrum access in the cognitive radio networks. A brief overview of the state-of-the-art literature showing the importance and topicality of the selected topic is provided. The major authors' contributions are presented and supported by the list of his recent publications.**

***Keywords*—Spectrum sensing, spectrum trading, dynamic spectrum access.**

I. Introduction

The increased demand for spectrum usage in the recent years poses high requirements on the efficiency of spectrum utilization. Recent studies [1] show that licensed part of the spectrum is usually not fully utilized by the primary (licensed) users (PUs). One way of increasing the efficiency of spectrum utilization is to allow secondary (unlicensed) users (SUs) to access *spectrum holes* (i.e. unused parts of the licensed parts of the spectrum), while avoiding to cause interference to the PUs. This is the main idea of dynamic spectrum access (DSA) [2]. The key technology, which seems to be a promising candidate for an underlying platform of the next generation of wireless networks based on DSA is *cognitive radio*[1] (CR) [3]. One of the main challenges in DSA is to find an optimal trade-off between spectrum exploration (i.e. spectrum sensing) and spectrum exploitation (i.e. spectrum sharing). Moreover, joint design of the mechanisms for DSA, which take this trade-off into the account is necessary in order to enable the implementation of DSA in the practical applications.

This paper presents a short summary of the authors' research work in the area of spectrum management in cognitive radio networks. A brief overview of the state-of-the-art publications in this area is provided in II followed by the long-term research tasks to be solved within the authors' PhD. thesis in III. Selection of the most important accomplished tasks and major results from the previous year of the authors PhD. studies is presented in IV. Paper is concluded with the proposed tasks to be accomplished in the short- and mid-term horizon in V.

[1]Cognitive radio is defined in [3] as: „Cognitive radio is an intelligent wireless communication system that is aware of its ambient environment. A cognitive radio transmitter will learn from the environment and adapt its internal states to statistical variations in the existing RF stimuli by adjusting the transmission parameters (e.g., frequency band, modulation mode, and transmission power) in real-time and on-line manner."

II. Initial Status

The major restriction posed on the secondary spectrum access in DSA consists in the fact that the transmission of SUs does not cause any measurable interference to the PU [2]. Therefore, it is necessary for SUs to have timely and accurate information about the available spectrum holes. The most prevailing method for the identification of the spectrum holes is spectrum sensing, which can be used in the *shared use* model of DSA [1]. In spectrum sensing, SU analyzes its radio environment, creates a test statistics from the received signal and based on this statistics, decides between two hypotheses: *PU transmitting* or *PU idle*. The most discussed methods for signal detection in spectrum sensing are energy-based detection [4], cyclostationary detection [5] and most recently also subspace-based detection [6], [7], which with its *blind* (i.e. no a priori information about the PU signal are required) nature provides outstanding signal detection performance in various scenarios. Sensing results can be significantly improved by exploiting the spatial diversity of several SUs and making the final sensing decision based on the local sensing results. This is referred to as cooperative spectrum sensing [8].

Another main challenge in the area of DSA is the design of the spectrum trading mechanisms. Spectrum trading can be referred to as a process of leasing or selling unoccupied spectrum resources of the PU to the SU [9] and it can serve as a motivation for PUs to share unused parts of their spectrum with SUs.

Several spectrum trading mechanisms for the *dynamic exclusive use* model of DSA [1] have been proposed in the literature [10], [11]. In the *dynamics exclusive use* model of DSA, the SUs have information about the channel occupancy directly from the PUs, therefore spectrum sensing is not necessary and spectrum trading algorithms do not address the problem of *risk* caused by incorrect identification of spectrum holes. On the other hand, *shared model* of DSA is exposed to the performance degradation due to the imperfect spectrum sensing, resulting in the occurrence of *risk* in the spectrum auction process. Here, the term *risk* can be conceived as uncertainty in achievable channel capacity (or transmission rate) of the SUs due to the possible interference with PU(s). Thus, with the increased level of risk, the SU decision regarding the valuation of the channel is reduced. In spite of the high importance of this critical issue, by the authors best knowledge, there are only a few papers dealing with risk-aware spectrum trading and joint design of spectrum sensing and spectrum trading mechanisms [12], [13].

III. Proposed Long-term Research Goals

Motivated by the indubitable importance of the topic of joint design of spectrum sensing and trading mechanisms on one hand, and relative lack of works published in this area on the other hand, the author proposes the following long-term goals for his post-graduate research:

- Design of an agent-based model of spectrum sharing in the CR network working under the dynamic exclusive use and shared use model od DSA
- Implementation of the methods of auction-based spectrum trading on the designed agent-based model of CRN
- Design and implementation of the algorithms for cooperative spectrum sensing with the focus on maximal sensing performance and minimal computational costs.
- Analysis of the impact of imperfect spectrum sensing on the proposed mechanisms of spectrum sharing and spectrum trading.
- Design of methods for the elimination of risk connected with spectrum trading caused by imperfect spectrum sensing for the purpose of optimization of the spectrum trading process.

IV. Solved Tasks

In order to follow the proposed research goals, the author focused in his research during the foregone year mainly on two areas:

- Analysis of market saturation in the secondary spectrum sharing based on spectrum trading,
- Impact of spectrum sensing imperfections on the spectrum trading performance.

The research in these two areas resulted in two published paper [14], [15] and two papers, which are still under review [16], [17].

In [14], the author provided and extensive overview of the state-of-the-art in the area of spectrum trading in dynamic spectrum access.

In [16], a resemblance of the strategies of the primary users is analyzed in the secondary spectrum market with various load (number of secondary users). In the scenario under investigation, primary users compete between each other with the goal of selling as much spectrum resources to the secondary users as possible in order to maximize their revenue. It was shown that when the system is under/overutilized, the price competition and demand sensitivity as a major economic impact do not play a key role in the decision process of the SUs. Instead, the physical location of the SUs in the range of the PUs is of greater importance. Thus the resulting prices for the spectrum portions are uncorrelated in under/overutilized network. In contrary, if the network utilization is between 50-85%, the economic principles start to apply and strictly determine the market behaviour. Within that network utilization range, the price sensitivity of the SUs causes successful connections to the non-closest PUs, suppressing the effect of distance as the major physical factor considered in this paper. This research was made with the cooperation with the team from the Department of Finance at the Faculty of Economy, TUKE.

In [15] an agent-based model for spectrum trading in the shared use model of dynamic spectrum access is proposed. Spectrum trading is employed using the single-unit sealed bid first-price auction, which takes into the account risk due to the imperfect spectrum sensing. Bidding strategies of the bidder are controlled by the reinforcement learning algorithm. We consider cooperative energy-based spectrum sensing as a spectrum sensing mechanism. Two different decision fusion strategies, which provide different levels of risk are discussed. The results demonstrate that in risky environment, total revenue and total payoff of the auctioneer and bidder respectively is higher, than in the case of system with lower level of risk. On the other hand, normalized revenue and payoff per a single auction round is higher in the case with lower level of risk. Moreover, the results have shown that the optimum sensing time for maximizing revenue and payoff is different.

The spectrum sharing model proposed in [15] is extended in [17]. The main extension in this paper is application of multichannel auction and addition of *majority rule* decision fusion mechanism for cooperative spectrum sensing. The main result of this paper is that in risky environment, total revenue and total payoff of the auctioneer and bidder respectively is lower, than in the case of system with lower level of risk. However, this result is limited only to the scenario described in the paper and it would be too courageous to make any general conclusions from the results presented there.

V. Future Research

Author is currently undertaking research mobility at the Victoria University of Wellington, where he focuses in his research on the eigenvalue-based spectrum sensing. His goal for this 10 months long research mobility is to work on the analysis of the performance of eigenvalue-based detector for blind signal detection for spectrum sensing and setting of the optimal threshold based on the total probability of sensing error. The results of this research will be later used for the design of joint spectrum sensing and spectrum trading mechanism.

Acknowledgment

This work was supported by the Scientific Grant Agency of the Ministry of Education, science, research and sport of the Slovak Republic under the contract No. 1/0766/14.

References

[1] Q. Zhao and B. M. Sadler, “A survey of dynamic spectrum access,” *IEEE Signal Processing MAgazine*, vol. 24, no. 3, pp. 79–89, 2007.

[2] E. Hossain, D. Niyato, and Z. Han, *Dynamic Spectrum Access and Management in Cognitive Radio Networks*. Cambridge University Press, 2009.

[3] E. Hossain, D. Niyato, and D. I. Kim, “Evolution and future trends of research in cognitive radio: a contemporary survey,” *Wireless Communications and Mobile Computing*, 2013.

[4] Y.-C. Liang, Y. Zeng, E. Peh, and A. T. Hoang, “Sensing-throughput tradeoff for cognitive radio networks,” *Wireless Communications, IEEE Transactions on*, vol. 7, no. 4, pp. 1326–1337, April 2008.

[5] A. Dandawate and G. Giannakis, “Statistical tests for presence of cyclostationarity,” *Signal Processing, IEEE Transactions on*, vol. 42, no. 9, pp. 2355–2369, Sep 1994.

[6] N. Pillay and H. Xu, “Blind eigenvalue-based spectrum sensing for cognitive radio networks,” *Communications, IET*, vol. 6, no. 11, pp. 1388–1396, July 2012.

[7] Y. He, T. Ratnarajah, J. Xue, E. Yousif, and M. Sellathurai, “Optimal decision threshold for eigenvalue-based spectrum sensing techniques,” in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, May 2014, pp. 7734–7738.

[8] K. Letaief and W. Zhang, “Cooperative communications for cognitive radio networks,” *Proceedings of the IEEE*, vol. 97, no. 5, pp. 878–893, May 2009.

[9] D. Niyato and E. Hossain, “Spectrum trading in cognitive radio networks: A market-equilibrium-based approach,” *Wireless Communications, IEEE*, vol. 15, no. 6, pp. 71–80, 2008.
[10] S. Sengupta and M. Chatterjee, “Designing auction mechanisms for dynamic spectrum access,” *Mobile Networks and Applications*, vol. 13, no. 5, pp. 498–515, 2008.
[11] S. Gandhi, B. Chiranjeeb, L. Cao, H. Zheng, and S. Suri, “Towards real-time dynamic specturm auctions,” *Computer Networks*, vol. 52, no. 4, pp. 879–897, 2008.
[12] T. Wang, J. Wu, and C. Long, “Flexibly cooperative spectrum sensing and allocation in cognitive radio networks: An auction-based framework,” in *Computer Science and Network Technology (ICCSNT), 2011 International Conference on*, vol. 2, Dec 2011, pp. 1250–1254.
[13] M. Tehrani and M. Uysal, “Auction based spectrum trading for cognitive radio networks,” *Communications Letters, IEEE*, vol. 17, no. 6, pp. 1168–1171, June 2013.
[14] J. Pastirčák, J. Gazda, and D. Kocur, “A survey on the spectrum trading in dynamic spectrum access networks,” in *Proceedings of ELMAR-2014 : 56th International Symposium*, 2014, pp. 131–134.
[15] J. Pastirčák, L. Sendrei, S. Marchevský, and J. Gazda, “An agent-based model of the risk-based spectrum auction in the cognitive radio networks,” in *Proceedings of the 1st International Conference on 5G for Ubiquitous Connectivity*, 2014.
[16] J. Pastirčák, L. Friga, V. Kováč, J. Gazda, and V. Gazda, “An agent-based economy model of real-time secondary an agent-based economy market of real-time secondary market for the cognitive radio networks,” 2014, submitted in Springer Journal of Networks and System Management.
[17] L. Sendrei, J. Pastirčák, J. Gazda, and S. Marchevský, “Cooperative spectrum sensing schemes for cognitive radios using dynamic spectrum auctions,” 2015, submitted in 38th International Conference on Telecommunications and Signal Processing.

The usage of formal methods to describe the virtual worlds

[1]Zuzana DUDLÁKOVÁ (3rd year)
Supervisor: [2]Branislav SOBOTA

[1,2]Dept. of Computers and Informatics, FEI TU of Košice, Slovak Republic

[1]zuzana.dudlakova@tuke.sk, [2]branislav.sobota@tuke.sk

***Abstract*— Development of software or hardware system is multi-phase and long-term process where all requirements for the system must be taken into account. It is necessary to combine them with the functionality, safety and not least with the speed. Formal methods have been created to deal, respectively to facilitate the entire process. The paper describes the current state of the solved problems in the areas of formal methods and virtual reality. The main aim is to use formal methods to describe the virtual worlds.**

***Keywords*—formal methods, software development, virtual reality, virtual worlds.**

I. Introduction

Currently, most information systems are very extensive, especially in terms of their functionality. The probability of errors is much greater because of increasing complexity of systems. Moreover, some of these errors can cause catastrophic consequences such as loss of money, time, or even lives. The main goal of software engineering is possibility to develop such systems that operate reliably in spite of their great complexity. One way of achieving this goal is the use of formal methods, which are based on mathematical foundations and they provide techniques and tools for the specification and verification of these systems. Their usage does not guarantee correctness but they can significantly enhance the understanding of the system so that they detect discrepancies, ambiguities and incompleteness that might be overlooked.

Virtual reality has proved to be major new technology of last year with the arrival of high-resolution graphics, high-speed computers and interactive user devices. The main advantage of virtual reality application is the possibility of creating and simulating of new technological devices before its implementation. This may contribute to increased safety and ergonomics and reducing economic aspects of newly designed devices. The virtual model of the proposed system can detect hidden bugs whose elimination at later stages of development could cause considerable difficulties.

This paper surveys existing research areas. Section II mentions many areas where formal methods and virtual reality have found their application and it describes existing solutions built on the interaction of formal methods with virtual reality. Section III specifies the proposal of formal description of virtual worlds.

II. Existing Solutions

Formal methods and formal specification languages are based on a mathematical basis and provide a means to demonstrate that the specification of the system is feasible, complete, consistent and unambiguous. Even the most complex systems can be modelled using simple mathematical objects such as sets, relations and functions.

So far, researchers have proposed a number of formal methods for computer systems development. Examples include Petri nets [1], VDM [2], Z notation [3], B method [4], Event B [5] and Perfect language [6]. One of the first areas where formal methods have found their application is the design of hardware. Their usage, either in the design control or simulation, contributed to their commercial availability. Currently, the formal methods are used in various fields [7] such as aerospace systems[7], rail systems [8], nuclear power plants, medical systems, the storage and usage of weapons, the design of microprocessors [9], ATM cards, virtual reality, computer games …

Virtual reality is currently apply in almost all areas [10], whether it's industry, science, research, medicine [11], education, film, games, military, sports [12] and many others.

Systems: ICO (Interactive Cooperative Objects) [13], Geist3D [14] and APEX framework [15] use formalism - Petri nets, which is nowadays widespread. The main reason for its success is that it has a nice graphical representation and well-defined semantics allowing the formal analysis. Petri nets are among the most useful languages to model systems containing concurrent processes. These are the reasons for which the options for integrating Petri nets to development process and design of virtual worlds have been reviewed (Fig. 1).

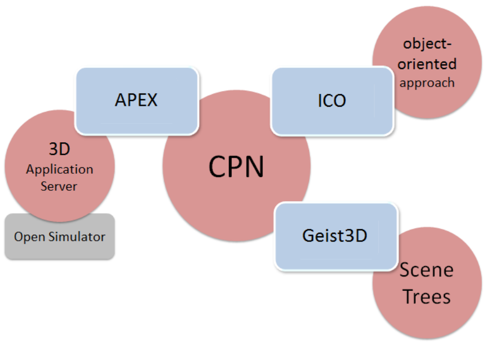


Fig. 1 Systems using formalism Petri nets

Mentioned systems use CPN (Coloured Petri Nets) to describe different aspects of modelling and controlling of virtual scenes [16]. ICO formalism using Petri nets in the context of object-oriented approach. Graphics engine Geist3D is built on CPN combined with tree structure for scene description - Scene Trees (Geometry-driven Petri nets). APEX framework combines 3D application server OpenSimulator with tool CPN Tools.

III. System Design

As it is necessary to formally describe the real world, so it is necessary to formally describe the virtual world. The formalisation of the system gives the option to user to operate with this system. A formal model allows examining even those things that may not be obvious at first glance for user. Because the procedure for creating a formal model of system can be shown only on some concrete example, the system for truck transport was chosen.

In the first phase of software development, the most accurate system requirements must be obtained. In the next stage, the most important one, there is necessary to process these requirements by correct and appropriate way. The formalisation gives the possibility to user to deal with this system before its creation and thus detect and find some hidden bugs, respectively streamline the system design. Another advantage may be an increase the level of its usability in practice, or prediction of possible situations or states that would normally be found on the basis of user experience acquired over several years.

A formal proposal of the system for truck transport is based on the distribution of its parts to dynamic (trucks) and static (roads, cities) parts. The static parts of the system are represented by a tree where vertices represent cities and edges represent roads. In the result, it is possible to detect an overload of some nodes (cities) or find those that provide new options to user through such an examination of provided data. Each building element (city, road, truck) has specific parameters. The functionality of the entire system depends on these parameters.

The visual output (Fig. 2) of the proposal (through simulation) helps to users to better imagine the functionality of the created system. The usage of virtual reality has very good potential to describe and illustrate the dynamic parts of the system to the user. In order for the resulting visualization was good, it is necessary to correctly describe the system (via formalisms).

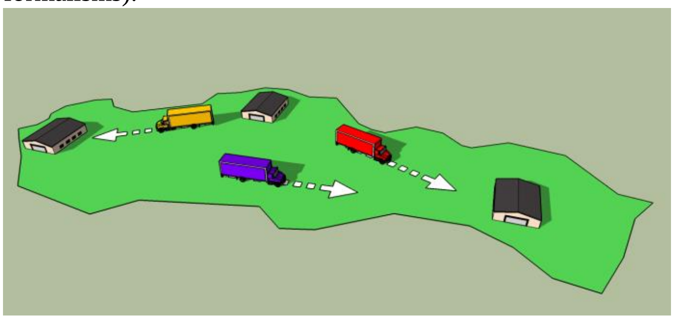


Fig. 2 Visualisation of system

IV. Future Work

Future work includes continuous development of the proposed solution, the improvement of formal description used in the design and experimental verification of this proposal on the real system.

Acknowledgment

This work was supported by the KEGA grant no. 083TUKE-4/2015 “Virtual-reality technologies in the process of handicapped persons education”.

References

[1] PETRI, C. A.: Introduction to General Net Theory. In: Proceedings of the Advanced Course on General Net Theory of Processes and Systems, Hamburg, 1979, Springer-Verlag, 1980, SS. 1-19.

[2] FITZGERALD, J., LARSEN, P. G.: Modelling Systems: Practical Tools and Techniques in Software Development. In: Cambridge University Press, 2009. ISBN: 0521626056

[3] SPIVEY, J. M.: The Z Notation: A Reference Manual. In: Prentice Hall International, Oxford, UK, 1992.

[4] ABRIAL, J. R.: The B book: Assigning programs to Meaning. In: Cambridge University Press, New York, 1996.

[5] ABRIAL, J. R.: Modeling in Event-B: System and Software Engineering. In: Cambridge University Press, 2010. ISBN: 9780521895569

[6] CROCKER, D.: Perfect Developer: A tool for Object-Oriented Formal Specification and Refinement. In: Proceedings of the Tools Exhibition Notes at Formal Methods Europe, 2003.

[7] DUDLÁKOVÁ, Z., VARGA, M.: Different approaches to utilisation of formal methods. In: Poster 2013 : 17th International Student Conference on Electrical Engineering : May 16, 2013, Prague. - Prague : Czech Technical University in Prague, 2013 P. 1-5. - ISBN 978-80-01-05242-6

[8] LECOMTE, T., BURDY, L., LEUSCHEL, M.: Formally Checking Large Data Sets in the Railways. In: Proceedings of DS-Event-B, Kyoto, Japan, November 2012.

[9] DAM, M., GUANCIALE, R., KHAKPOUR, N., NEMATI, H., SCHWARZ, O.: Formal verification of information flow security for a simple arm-based separation kernel. In: Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security, New York, USA, 2013, S.223-234. ISBN: 978-1-4503-2477-9

[10] SOBOTA, B., HROZEK, F., SZABÓ, C., VASKO, M., IVANČÁK, P., VARGA, M., DUDLÁKOVÁ, Z., PEŤKA, R., JANOVSKÝ, D., JACHO, L.: Concept of scalable virtual reality system in distributed computational environment. In: Computer Science and Technology Research Survey : volume 7. - Košice : TU, 2014 S. 198-206. - ISBN 978-80-553-1857-8

[11] AHN, W., DARGAR, S., HALIC, T., LEE, J. et al.: Development of a Virtual Reality Simulator for Natural Orifice Translumenal Endoscopic Surgery (NOTES) Cholecystectomy Procedure. In: Medicine Meets Virtual Reality 21, IOS Press, Manhattan Beach, California, February 2014.

[12] LAGOS, L., VASCHILLO, E., VASCHILLO, B., LEHRER, P., BATES, M., PANDINA, R.: Virtual Reality-Assisted Heart Rate Variability Biofeedback as a Strategy to Improve Golf Performance: A Case Study. In: The Association for Applied Psychophysiology and Biofeedback, Volume 39, Issue 1, 2011.

[13] NAVARRE, D., PALANQUE, P., BASTIDE, R., SCHYN, A., WINCKLER, M., NEDEL, L. P., FREITAS, C. M. D. S.: A Formal Description of Multimodal Interaction Techniques for Immersive Virtual Reality Aplications. In: Proceeding of the 2005 IFIP TC13 international conference on Human-Computer Interaction, Université Paul Sabatier, Toulouse, France, 2005, pp.170-183.

[14] WEBER-JAHNKE, J. H., STIER, J.: Virtual Prototyping of Automated Manufacturing Systems with Geometry-driven Petri Nets. In: Department of Computer Science, University of Victoria, Canada, 2009.

[15] SILVA, J. L., RIBEIRO, O. R., FERNANDES, J. M., CAMPOS, J. C., HARRISON, M. D.: The APEX Framework: Prototyping of Ubiquitous Environments Based on Petri Nets. In: Universidade do Minho, Braga, Portugal, 2010.

[16] DUDLÁKOVÁ, Z., KOREČKO, Š., SOBOTA, B.: Utilization of Petri Nets for Specification of Virtual Reality Systems. In: Informatics 2013, Proceedings of the Twelfth International Conference : Spišská Nová Ves, Slovakia, November 5-7, 2013. - Košice : TU, 2013 S. 104-109. - ISBN 978-80-8143-127-2

Towards Control and Modeling of Complex Systems Using Fuzzy Cognitive Maps

[1]*Michal PUHEIM (2nd year)*
Supervisor: [2]Ladislav MADARÁSZ
Consultant: [3]Ján VAŠČÁK

[1,2,3]Dept. of Cybernetics and Artificial Intelligence, FEI TU of Košice, Slovak Republic

[1]michal.puheim@tuke.sk, [2]ladislav.madarasz@tuke.sk, [3]jan.vascak@tuke.sk

***Abstract*—In this paper we present a current state of our research in the area of intelligent control and modeling of complex systems. We focus on recent progress in solving this research task with a proposal for a novel approach to system modeling using modified Fuzzy Cognitive Maps (FCM) which we refer to as the Three-Term Relation Neuro-Fuzzy Cognitive Maps (TTR NFCM). With regard to this novel method we describe the need for a new multi-purpose FCM library and introduce its proposal. In the last section of the paper we outline our plans for further work which will be based on application of given method to modeling and control of selected complex systems, in our case the small turbojet engine (STJE) model ISTC-21V and TJ-100.**

***Keywords*—fuzzy cognitive maps, modeling, control, complex systems,**

I. Introduction

The main objective of our research as stated by the title of my future PhD thesis is the “Utilization of means of computational intelligence in situational control of complex systems”. We have chosen the FCM as the selected intelligent method because of its universal properties which suitably combine the approximation capabilities of neural networks and readability of rule-based fuzzy systems. The STJE was chosen as the object of control because of its complex nonlinear internal structure and characteristics.

II. Initial State of Research

As a starting point for research in the recent year we used the FCM survey [1] and our own overview manuscript [2]. We based our further work on fundamental assumptions as stated in introductory article [3]. We used monograph [4] to acquire knowledge about state-of-the-art methods used for control, modeling and diagnostics of small turbojet engines.

III. Current State of Research

As stated in [3], the basic conception of FCM has several drawbacks which restrain its capability to model dynamic nonlinear relations within complex systems. Therefore our initial goal was to deal with these shortcomings.

A. *Three-Term Relation Neuro-Fuzzy Cognitive Maps*

In [5] we proposed a new TTR NFCM methodology, which enhances the conventional FCM by adding two new features. The first one solves the problem of relation dynamics by inclusion of trends in the concept update formula using the *Three-Term Relations* (TTR – see Fig. 1), which are inspired by control engineering methods, namely the PID controllers.

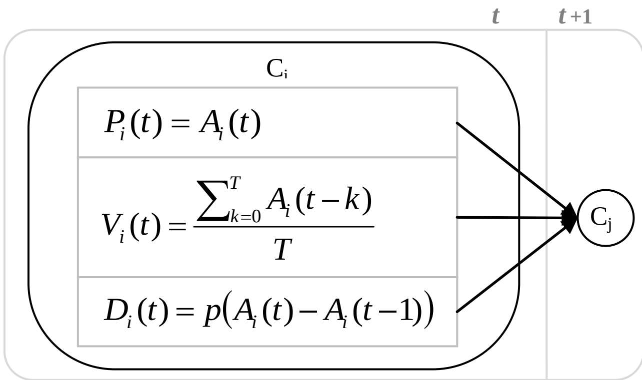


Fig. 1. Three-term relation between concepts C_i and C_j with proportional component P_i , moving average component V_i and derivative component D_i . T is the size of the time window used to calculate the moving average [5].

The second main feature is the replacement of simple linear weights between concepts by nonlinear feed-forward neural networks or multilayer perceptrons (MLPs). The applicability of the feed-forward MLPs instead of conventional weights comes from the fact, that every single FCM can be unwrapped into several simple feed-forward cognitive maps (see Fig. 2).

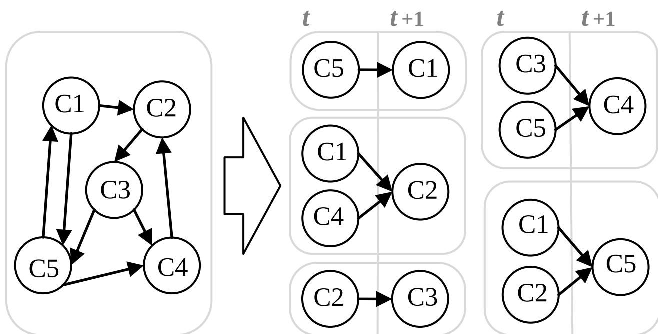


Fig. 2. Unwrapping of a simple FCM [5].

With an unwrapped FCM it is possible to replace all of the relations preceding a single concept with an MLP and transform the FCM into the hybrid *Neuro-Fuzzy Cognitive Map* (NFCM – see Fig. 3).

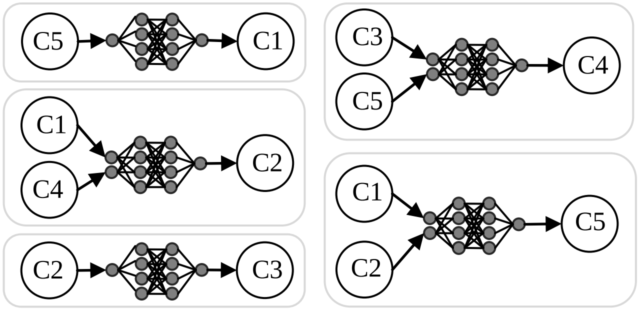


Fig. 3. Nonlinear relations between concepts using the MLPs [5].

B. Multi-Purpose Fuzzy Cognitive Maps Library

In order to implement the presented method and to tackle the problems and deficiencies of existing programs and tools used for FCM modeling, we proposed a new general multi-purpose library [6] (see Fig. 4).

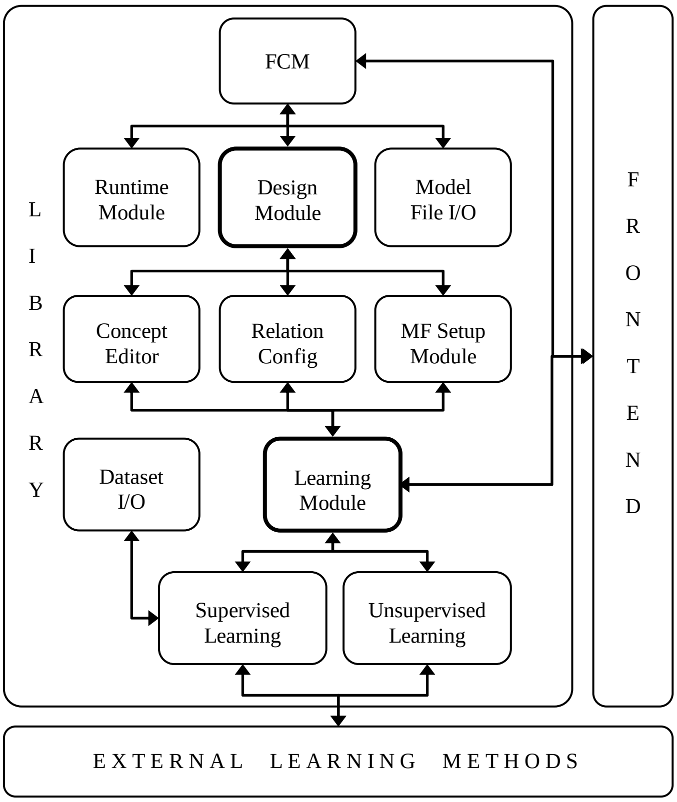


Fig. 4. A system diagram of the proposed FCM library[6].

We proposed the library with a goal of enabling fast and simple prototyping of system models. The library is also designed to supports various methods for relation expression (nonlinear, dynamic, neural, etc.) and to employ built in automatic learning mechanisms to adjust FCM parameters.

IV. Future Work

In accordance to the objectives of my future PhD thesis [7], our next goal is to implement the library and use it to create a model and an automated control system for STJE [8].

A. Implementation of Multi-Purpose FCM Library

After a careful consideration of several aspects related to the library implementation [6], we decided to develop it using the .NET framework, which is well integrated within modern versions of Windows OS. We aim to create a shared dynamic library (DLL), which can be incorporated into programs in various languages and environments. e.g. Python, Matlab, Simulink, WPF, etc.

B. Modeling and Control of STJE with use of FCM

The library will be used to create an experimental model of the ISTC-21V engine (see Fig. 5) and possibly also TJ-100 engine using the proposed TTR NFCM method.

Fig. 5. Small turbojet engine ISTC-21V [8].

After the method is successfully applied and evaluated for the purpose of system modeling, we will proceed with a design and implementation of an FCM controller for the ISTC-21V engine. Our goal is to simplify and outperform the existing control system (which consists of several different control algorithms) with a single unified approach.

Acknowledgment

This paper is supported by KEGA, Grant Agency of Ministry of Education of Slovak Republic, under *Grant No. 014TUKE-4/2015* – “Digitalization, virtualization and testing of a small turbojet engine and its elements using stands for modern applied lecturing.“

References

[1] E. I. Papageorgiou, J. L. Salmeron, “A Review of Fuzzy Cognitive Maps research during the last decade,” In: IEEE Transactions on Fuzzy Systems, vol. 21, no. 1, Feb. 2013. pp. 66-79. DOI: 10.1109/ TFUZZ.2012.2201727

[2] M. Puheim, J. Vaščák, L. Madarász, “Aplikácia fuzzy kognitívnych máp na riešenie problémov v oblasti leteckých turbokompresorových motorov.,” Unpublished.

[3] M. Puheim, “Introduction to Modeling of Complex Systems Using Fuzzy Cognitive Maps.” In: SCYR 2014: 14th Scientific Conference of Young Researchers: Proceedings. May 20th, 2014, Herľany, Slovakia. – Košice : TU, 2014. pp. 201-204. ISBN 978-80-553-1714-4.

[4] T. Lazar, L. Madarász, et al., “Inovatívne výstupy z transformovaného experimentálneho pracoviska s malým prúdovým motorom,” Monograph, Košice: Elfa, s.r.o., 2011, ISBN 978-80-8086-170-4.

[5] M. Puheim, J. Vaščák, L. Madarász, “Three-Term Relation Neuro-Fuzzy Cognitive Maps.” In: CINTI 2014: 15th IEEE International Symposium on Computational Intelligence and Informatics: Proceedings. November 19-21, 2014, Budapest. – Danvers: IEEE, 2014. pp. 477-482. ISBN 978-1-4799-5337-0.

[6] M. Puheim, L. Madarász, J. Vaščák, “A Proposal for Multi-Purpose Fuzzy Cognitive Maps Library for Complex System Modeling.” In: SAMI 2015: IEEE 13th International Symposium on Applied Machine Intelligence and Informatics: Proceedings. January 22-24, 2015, Herľany, Slovakia. – Danvers: IEEE, 2015, pp. 175-180. ISBN: 978-1-4799-8220-2.

[7] M. Puheim, “Využitie fuzzy kognitívnych máp v situačnom riadení a modelovaní zložitých systémov.” Písomná práca k dizertačnej skúške. Košice: TU, FEI, 2015. 77 pp.

[8] L. Főző, R. Andoga, L. Madarász, J. Kolesár, J. Judičák, „Description of an Intelligent Small Turbocompressor Engine with Variable Exhaust Nozzle,“ In: SAMI 2015: IEEE 13th International Symposium on Applied Machine Intelligence and Informatics: Proceedings. January 22-24, 2015, Herl’any, Slovakia. – Danvers: IEEE. 2015, pp. 157-160. ISBN 978-1-4799-8221-9.

Trajectory generator for mobile robot in plane

[1]Jakub ČERKALA (3^{rd} year),
Supervisor: [2]Anna JADLOVSKÁ

[1,2]Dept. of Cybernetics and Artificial Intelligence, FEI TU of Košice, Slovak Republic

[1]jakub.cerkala@tuke.sk, [2]anna.jadlovska@tuke.sk

***Abstract*—The research presented in this article focuses on author's work that was done last year according to dissertation theses.**

***Keywords*—mathematical modelling, robotics, mobile robot, trajectory.**

I. Introduction

The application potential of artificial intelligence methods as an extension of classical modelling and control approach for robotic systems is the main subject of dissertation theses. This paper's topics describe new results and work already done in fulfilling author's dissertation theses, the continuous research is presented in previous articles [1], [2]. The last year's research covers various partial tasks like the mathematical modelling of more accurate simulation mobile robot model with friction, a generation of various testing trajectories together with their properties analysis. A space is also devoted to dissertation main control structure with reference robot along with the description of 3D environment for robotic systems virtualization.

II. Previous analysis and achieved results in research field

Research field of robotics systems has been subject of author's previous research under goals of project VEGA No. 1/0286/11 Dynamic Hybrid Architectures ot the Multiagent Network Control Systems. During the robotic systems analysis, a key models of further interest were identified - the robotic arm and mobile robot with differential chassis as representatives of stationary and mobile robotics.

The robotic arm OWI 535 model and its control using USB interface was discussed in [1]. In terms of hardware and programming, the experience gained during work with rather different Laboratory Hydraulic model was applied.

Multiple kinematic and simple dynamic models of mobile robots have been implemented in Simulink environment as a library of robotic systems, which was the main focus of [2], the library is still under development with respect to Technicom project goals. Simulation models were verified in point and posture control structures.

Artificial intelligence methods in modelling and control can applied on a variety of robotic systems, the analysis conducted in dissertation prospectus have revealed potential application of neural models, especially in trajectory following control tasks, that is reflected in future dissertation theses.

Part of the final solution will be the OpenGL based environment programmed in C# language, designed for visualization of robotic systems, especially for mobile robots. First version of this simulator was already programmed, but it still require further attention to increase it's robustness.

The program outputs falling below dissertation theses, are mainly carried out in simulation programming language MATLAB/Simulink and in object programming language C#, previous publications also support the KEGA 021TUKE-4/2012 project.

III. Solved tasks and results

This article summarize the main results of various research tasks realized in this study year.

A. *Differential mobile robot with friction*

A significant part of last year's research [2] were devoted to modelling of mobile robot with differential chassis. Within a created simulation models library in Simulink environment, that contains the kinematic models with wheel radius r and distance between wheels b, based on

$$\begin{bmatrix} \dot{x}(t) \\ \dot{y}(t) \\ \dot{\varphi}(t) \\ \dot{\theta}_R(t) \\ \dot{\theta}_L(t) \end{bmatrix} = \begin{bmatrix} \frac{r}{2}\cos\varphi & \frac{r}{2}\cos\varphi \\ \frac{r}{2}\sin\varphi & \frac{r}{2}\sin\varphi \\ \frac{r}{2b} & -\frac{r}{2b} \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \omega_R(t) \\ \omega_L(t) \end{bmatrix}, \tag{1}$$

were implemented additional mobile robot parametric models with more dynamic properties.

Since these robots will be used as a data source for neural models training [3] as mentioned in previous article [2], to get a data with better accuracy, an implementation of generalised friction model [4] was subject of research.

Achieved results were presented at a conference with a personal participation supported by the project VEGA in article [5], which extended version were sent to Lambert Academic Publishing for further publication.

B. *Mobile robot reference trajectory generator*

Mobile robot reference trajectory is usually defined as a set of $[x_r, y_r]$ position coordinates, that are sampled in reasonable rate. An interval of these points for motion in plane can be approximated as an arc or line, the reference trajectories, commonly used for control algorithms verification are patterns like 8-shape [6], because they contain dynamically changing line/arc combinations.

A generator of various reference trajectories, that will be used in further control experiments were designed as a part of simulation models library with an options to define required pattern characteristics and scale, an example of pattern generators are depicted on Fig. 1.

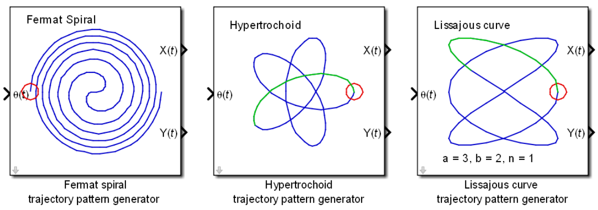


Fig. 1. Single trajectory generators - an example of created Simulink blocks.

The single trajectory generator summarize under one Simulink block multiple patterns:

- point / straight line / sine wave
- involute of circle / Fermat spiral,
- circle / ellipse / superellipse - rounded rectangle,
- 8-shape - lemniscate of Gerono / Bernoulli,
- epitrochoid / hypertochoid / Lissajous curve and more.

The parametric definition of each trajectory used to program generator ensure ability to modify the final shape by changing parameters. Dynamic mask of Simulink block enables to plot updated chosen trajectory after every change in parameters even before simulation.

These reference trajectory patterns will be used for testing of mobile robot classical and intelligent [3] control structures, however they can find an application with robotic arms or ball and plate model trajectory tasks.

C. Trajectory for mobile robot analysis

In terms of mobile robotics, the trajectory planning require a significant attention and importance. Mobile robots with dynamics have specific construction limits and therefore the required trajectory has to be analysed in advance - the robot must be able to realize defined trajectory.

Taking into account it's maximum linear and angular velocities together with sample rate, it is possible to calculate approximate linear velocity $v_r(t)$ for changing position coordinates $[x_r, y_r]$ as

$$v_r(t) = \pm\sqrt{\dot{x}_r^2(t) + \dot{y}_r^2(t)}, \tag{2}$$

where $\pm$ defines direction of robot's motion. The orientation angle φ_r for each trajectory point can be obtained as

$$\varphi_r(t) = \text{arctan2}(\dot{y}_r^2(t), \dot{x}_r^2(t)) + k\pi, \tag{3}$$

and $k = 0$ or $k = 1$ defines the selected rotation direction. The robot's approximate angular velocity $\omega_r(t)$ can be obtained from derivation of (3) as

$$\omega_r(t) = \frac{\dot{x}_r^2(t)\cdot\ddot{y}_r^2(t) - \ddot{x}_r^2(t)\cdot\dot{y}_r^2(t)}{\dot{x}_r^2(t) + \dot{y}_r^2(t)} = v_r(t)\cdot\kappa(t) \tag{4}$$

where $\kappa(t)$ denotes the path curvature. For trajectories generated from parametric definition applies that, with an exception of straight line and circle, the euclidean distances between two points are not equal - this is reflected in lower linear velocity $v_r(t)$ with greater angular velocity $\omega_r(t)$ for larger curvature $\kappa(t)$ and vice versa. These trajectory properties should be also taken into account during neural models training and experiments.

D. Trajectory tracking control for mobile robot

There are multiple approaches and control structures for mobile robot trajectory tracking. As mentioned in previous article [2], the control structure with a reference model, that includes feed-forward and feedback control was chosen as dissertation main control structure, depicted on Fig. 2.

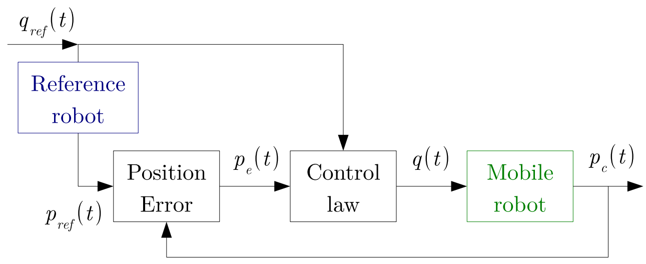


Fig. 2. Mobile robot with friction as simulink block.

The control law $q(t)$ is a combination of feed-forward part $q_{ref}(t)$, that can be obtained from defined reference trajectory using (2), (4) and feedback part, that depends on position error $p_e(t) = p_{ref}(t) - p_c(t)$. The simple kinematic model, mobile robot with dynamics and friction extended by internal PI control loop or trained neural model can be used as reference or controlled mobile robot.

IV. Proposal for next steps

In terms of mobile robot control, the priority is to include methods of artificial intelligence [3] into classical control structures, mainly to control structure with reference robot. Another important step is to modify reference trajectory to equidistant step, which will enable usage of additional control structures and can provide interesting results in comparison between classical and intelligent control approaches. At the same time, another task is to enhance OpenGL based environment for efficient visualisation of robotic systems.

Acknowledgment

This work has been supported by the Development Operational Program forproject: University Science Park Technicom for innovative applications with knowledge technology support, ITMS code 26220220182, co-financed by the ERDF.

References

[1] J. Čerkala, "Control of OWI 535 Robotic Arm Using C# and Possibilities of Future Applications," in *SCYR 2013 : 13^{th} Scientific Conference of Young Researchers*. Faculty of Electrical Engineering and Informatics, May 2013, pp. 23–25.
[2] ——, "Modeling and control of robotic systems," in *SCYR 2014: 14^{th} Scientific Conference of Young Researchers : proceedings from conference*. Faculty of Electrical Engineering and Informatics, May 2014, pp. 217–218.
[3] J. Velagic, N. Osmic, and B. Lacevic, "Neural network controller for mobile robot motion control," *International journal of intelligent systems and technologies*, vol. 3, no. 2, pp. 127–133, 2008.
[4] I. Virgala and M. Kelemen, "Experimental friction identification of a dc motor," *International Journal of Mechanics and Applications*, vol. 3, no. 1, pp. 26–30, 2013.
[5] J. Čerkala and A. Jadlovská, "Mobile robot dynamic with friction in simulink," in *Technical Computing Bratislava 2014 : 22^{nd} Annual Conference Proceedings : Bratislava, 4.11.2014. - Prague*. Institute of Chemical Technololgy, 2014, p. 10.
[6] G. Klancar, D. Matko, and S. Blazic, "Mobile robot control on a reference path," in *Intelligent Control, 2005. Proceedings of the 2005 IEEE International Symposium on, Mediterrean Conference on Control and Automation*. IEEE, 2005, pp. 1343–1348.

Understanding Malware Behaviour for its Future Analysis and Formalisation

[1]*Jana ŠŤASTNÁ (1st year),*
Supervisor: [2]*Martin TOMÁŠEK*

[1,2]Dept. of Computers and Informatics, FEI TU of Košice, Slovak Republic

[1]jana.stastna@tuke.sk, [2]martin.tomasek@tuke.sk

***Abstract*—At the beginning of malicious software research it is necessary to acquire deeper knowledge in this field. Perhaps unusually, this article does not deal with classification of malicious software on viruses, trojan horses, worms, etc. I leave that task to other researchers. The following text is centered around various types of malware behaviour and propagation of malware on infected systems. I provide a brief overview of research made by several authors, who analysed malware samples in their experiments. Based on a way the samples affected a host system, they were divided into several categories. This article also summarizes some of the problems associated with static and dynamic malware analysis and virtual analytic environments. It briefly describes techniques that are used for extraction of specific malware behaviour and categories of typical actions performed by malicious samples. Future plans for my work, based upon results in malware behaviour analysis, are outlined at the end of the article.**

***Keywords*—Dynamic analysis, malware behaviour, mutation, obfuscation, static analysis.**

I. Introduction

At this time, there is no uniform, globally accepted definition of malicious software, often referred to as malware. The informal ones differ. Computer security researchers often define malware according to the context of their current work or point of view on the topic.

One of well known informal definitions was formulated by Fred Cohen in his book [1], though he writes about a computer virus, not malware in general. According to it, "*a virus is a program that can 'infect' other programs by modifying them to include a, possibly evolved, version of itself*". As Peter Szor mentioned later in his book [2], Cohen's definition does not distinguish between copying a program for malicious purposes and copying a program as a desired side-effect. Szor slightly modified Cohen's definition and it mentions explicit replication and evolution of viruses, however, it still tells nothing about specific behaviour or manifestation on infected systems. In addition, the definition is not useful for other types of malware, which typically do not evolve or replicate on the host system.

Ding *et al.* mentions in his article that "*malware is a program that makes your system do something that an attacker wants it to do*" [3]. Again, this statement is vague and does not take into account other properties of malware.

It seems that informal definitions are not very useful for describing and deeper understanding of malicious software. Also Cohen's *formal* definition of the computer virus [1] does not describe all the malicious behaviour we should expect when trying to analyse and detect new malicious samples.

My goals are to create formal models of malware bahaviour and to use these models to improve malware analysis and early detection of unknown malicious samples, either new ones or evolved versions of those already in the wild. At the beginning of my work I need to learn as much as possible about malware and its characteristic behaviour. This article summarizes my research so far.

II. Frequent behaviour of malicious software

Malware writers implemented techniques for obfuscation of malicious source code and mutation of already existing samples, which defeat detection systems based on malware signatures. The problem is that signatures are based upon syntax of malicious samples and as Moser *et al.* mention, it is easier to create numerous malware variants in terms of the syntax rather than semantics [4].

Methods of malware concealment (e.g. encryption, dead-code insertion, register reassignment, subroutine reordering, instruction substitution, code transposition) [5] make also static analysis and detection very difficult, in some cases even impossible [4], [6].

Currently, a lot of malware researchers use dynamic analysis, which allows to observe behaviour of malware samples during their execution, for example Bailey *et al.* [7], Rieck *et al.* [8], Wagener *et al.* [9].

There are several techniques for dynamic malware analysis, which are described in a survey [10] by Egele *et al.*: *Function call monitoring* allows to record which specific functions (e.g. *Windows API* and *Windows Native API* functions, *system calls*) were called by analysed sample. *Function Parameter Analysis* is focused on current values of parameters of called functions, and their relationship with return values of functions called previously. *Information Flow Tracking* examines usage or manipulation of specified data, which are *tainted* with labels, during execution of analysed program. *Instruction trace* allows to examine details in behaviour of the sample at the level of machine instructions. The survey also mentions *Autostart Extensibility Points (ASEPs)*, described by Wang *et al.* in their work [11]. These points allow malware to be launched during boot process of the system or when specific program starts executing.

Every program can comprise numerous execution paths, also called execution traces. The disadvantage of dynamic analysis is that only one execution trace can be observed at a time. Concerning several traces, also static analysis with reverse engineering is problematic. However, Beaucamps *et al.* addressed this problem in their work [12] and proposed a method for static analysis of execution traces acquired from control-flow graphs. Also dynamic method was created by Moser *et al.* [13], which obtains control-flow points of a program according to inputs that the program requires. After that a system they designed is able to explore detected execution traces.

Although dynamic analysis is not absolutely reliable, it allows to detect and observe various malware behaviour. According to Wu *et al.*, operations that represent malware behaviour alter state (status) of the infected system [14]. Based on the type of alteration, they divided malware operations into 4 categories:

1) File actions,
2) Process actions,
3) Network actions,
4) Registry actions.

Unfortunately, they do not describe the environment in which malicious samples were investigated nor the methods used to carry out the analysis. However, it turned out that these categories of malware operations are appropriate. Bailey *et al.* used them in their work aimed at malware analysis and classification [7]. They correspond also to the work of Rieck *et al.* [8], Wagener *et al.* [9], and Bayer *et al.* [15] in which another possible category

5) Actions of graphical user interface

appeared.

A. File actions

Based on results of Bayer's *et al.* research [15], more than 70% of observed malicious samples affected filesystem during their execution. The actions that were performed related to

1) replication of malicious sample,
2) relocation of the program,
3) creation of non-executable files,
4) modification of various existing files.

In the first case, a new executable file is created, which is usually a plymorphic variant of the original sample. It could be interesting to search for such malware analytic techniques, which would compare newly created file with actually executing sample and in this manner discover similarities between them. This could indicate potential malware replication. Location of the new file may also be significant.

In the second case, according to Bayer *et al.* [15] the program or its components often relocate to Windows system directory.

In the third case, Bayer *et al.* mention [15] mainly temporary files (created by Internet Explorer), scripts, and `dll` libraries which are often located to system directory.

Modification of existing files occured in most cases indirectly, as a new record of an event in Windows system audit file. Bayer *et al.* claim that the data can be found in `system32\config\SysEvent.Evt`. However, I was not able to locate this file on my Windows 8 operating system. It is possible that my version of Windows places the audit file somewhere else.

Majority of observed malware samples tried to tamper with pre-installed programs such as Internet Explorer and Windows Media Player.

Deleting a file may also be included in the fourth category mentioned above. Malicious programs delete temporary files that they created during execution. They cover the tracks of their malicious actions. I expected it to be a very common phenomenon, but surprisingly, as Bayer *et al.* explain in their work [15], only few malware samples deleted a `.log` and `.evt` file[1]. It may be useful to learn more about how system handles these logs, if they are inspected or compared in some way. Perhaps these logs could help to get more detailed insight into behaviour of malware and its impact on the system. But first I need to learn what to look for, only then I will be able to filter desired data.

Bailey *et al.* noticed in their experiments [7] that newly created files have common names among numerous malware samples. They mention specifically `winhlp32.dat`.

B. Process actions

Concerning processes, an interesting observation was made by Bailey *et al.*, who note that many processes created by malware have the same name as common Windows utilities [7].

Rieck *et al.* mention infection of a running process by a computer virus and launching and terminating Windows services (e.g. disabling a malware protection) [8].

Wu *et al.* discovered that 67% of malware samples they observed, created sub-processes at the time of execution [14].

C. Network actions

Bailey *et al.* observed usage of mail ports (probably for spam), common IRC port 6667, HTTP ports for downloading of additional malicious content and command and control tunneling [7].

Results of Bayer's *et al.* work [15] show that almost half of the samples which perform network actions, query the DNS server for domain name resolving, yet 9.2% of them received no result. 19% of samples (mostly computer worms) performed scanning of specific ports, e.g. 139, 445, and 9988. Authors provide an overview of network activity, according to which the most used protocol was TCP (45.74% of samples), then UDP (27.34%), DNS queries follow with 24.53%, HTTP protocol was used by 20.75%, address scanning was used by 19.08%, ICMP traffic occured in 7.58% of cases, listening on ports appeared in 1.88% of cases, IRC protocol was used by 1.72%, SMTP by 0.89%, SSL by 0.23% and port scanning was used by 0.01%.

D. Registry actions

According to Bailey *et al.* some samples used randomly generated numbers and modified PRNG[2] registry keys [7].

Registry actions are better described in Bayer's *et al.* work [15]. About 62.7% of their samples produced new registry entry. These entries related to network adapter control, unique identifiers of COM objects, trusted certificates (specifically, registry entry with key `SystemCertificates\TrustedPublisher\Certificates` was created to install attacker's certificate). Registry

[1] `.log` files were deleted by 0.26% of observed samples and `.evt` files by 0.0018% of samples. Source [15] claims that the amount of unique samples was 901 294.

[2] Pseudo-random number generator

entry with key `Windows\CurrentVersion\Policies\System` caused that user was unable to launch task manager. Modification of existing registry entries was targeted at disabling Windows firewall, changing security settings (key `MSWindows\Security` was used), Windows services configurations (key `Services`) and setting autorun applications, which allowed malicious sample to be executed automatically in case of system restart (key `Currentversion\Run` was used).

E. Actions of graphical user interface

About a third of samples analysed by Bayer *et al.* [15] displayed a new window during the execution. 2.2% of them was caused by failure of the program. 4.47% of displayed windows did not have any title nor text content. Majority of created windows contained simple error messages, which should distract user's attention or decrease suspicion of malware attack.

Bailey *et al.* observed in their experiment [7] that sometimes malware samples even use existing programs, e.g. `iexplore.exe`, for launching popup windows with advertisements.

III. Malware behaviour extraction

According to Lu *et al.* the most modern way to represent malware behaviour is by creating a behaviour graph, which shows how information flow between system calls that the program executes [16]. However, Moser *et al.* showed, that it is possible to inhibit (probably static) analysis of control-flow and data-flow in programs with obfuscation techniques which hide locations of data and data usage [2].

It seems that extraction of malicious behavior from system calls is frequently used technique. Wagener *et al.* described in their work [9] a method for creating models of malware behaviour. Before they started analysing a malware sample, they recorded the initial state of the virtual environment in which the sample was going to execute. After the experiment, they again recorded state of the system. By comparing the initial and final state of the system, they obtained a first overview of how the investigated sample affected the system. In addition to that, their analytic system generated reports about execution of the sample. At the end, they were able to extract executed system calls from the collected data. Regarding data representation, Wagener *et al.* assigned number to each type of system calls, e.g. to `LoadLibraryA` was assigned number 1. According to the sequence of system calls from generated report a sequence of numbers (codes) is created. These sequences are then compared in form of a matrix and system looks for similar and different segments.

Bailey *et al.* question adequacy of malware behaviour extraction based on system calls. In their article [7] they claim, that system calls are too low-level to provide meaningful information, and that their level of abstraction is not suitable for comprehensive description of malware. They describe malicious behaviour by analysing changes of the state of a testing system. It is hard to determine whether their statement is correct or incorrect. A study examining precision of system calls-based detection [17] may shed more light on this topic.

IV. Present challenges

For security reasons, dynamic analysis of malware is often performed in a virtual environment or a sandbox. A problem may arise when some kinds of malware are able to detect this uncommon environment. In case this condition occurs, malware samples start to execute different execution path so that no suspicious action is performed. Consequently, the analysed program looks harmless and draws no attention.

Bayer *et al.* address also this problem in their research [15]. They divide anti-sandbox techniques, according to method of detection of this environment, into:

1) detection technique on a level of symbolic instructions,
2) detection on a level of Widows API calls.

The first method verifies if CPU instructions correspond to real CPU or to emulated CPU. It is possible to lower efficiency of this method by using customized environment which resembles real environment at the level of instructions. Bayer *et al.* did not study this technique deeper, particularly because of performance limitations and time-consumption of monitoring of each instruction.

With the second technique malware inspects the execution environment by means of certain API function calls. The article [15] by Bayer *et al.* discusses comparison of return values obtained from API functions, e.g. from `GetComputerName` and `GetCommandLine`. The return value of `GetComputerName` is hard-coded in some of emulated environments, so if malware writers recognize this value, it is not difficult for them to add corresponding detection mechanism into malicious programs. Some emulation systems set the same name for every program executed with them. In analogy to the previous case, it is possible to discover this property with comparing the return value from function `GetCommandLine`.

Another problem that may occur with dynamic analysis in virtual environment is that analysing a massive amount of samples is time consuming. Usually, every sample can get only limited amount of execution time, therefore certain malicious events which trigger later may not be revealed. Crandall *et al.* address this problem in their work [18]. They proposed an automated analytic technique which tries to determine time-triggered actions. It may be promising, however, they tested the system only on 6 samples of computer worms, which I consider too few to draw some conclusion about the system's efficiency.

To learn more about these issues I plan to study work of Raffetseder *et al.* [19], in which they compare detectability of system emulators and casual virtual machines and also analyse techniques for detection of such environments. Ferrie *et al.* also discuss detection techniques, but in addition propose defensive methods against them [20]. Ormandy evaluates in his work [21] practical security of virtual machine implementations used for malware analysis. It could be beneficial to take his findings into consideration in future preparations of malware analytic environment. Dinaburg *et al.* designed in their work [22] a transparent malware analyser that tries to solve problems mentioned above.

To prevent malware detectors from discovering malicious samples, these are changed - mutated, e.g. with methods of code obfuscation [5], compression and encryption. According to results of mutation we can distinguish between two types:

- *Mutations preserving semantics of a program.* Program's instructions are adjusted with obfuscation techniques, it will exhibit the same malicious behaviour, however, it will be different syntactically. Its detection with malware signatures becomes impossible and it becomes harder to analyse statically. As Konstantinou also stated in his work [23], this leads to malware metamorphism.

- *Mutations modifying semantics of a program*. As a consequence, following generation of the malicious program exhibits somewhat different behaviour. Konstantinou mentions this as an advanced malware evolution. An example is infection of a virus `W32/HLLW.QAZ` by another virus `W32/FunLove`. The resulting `QAZ` virus carried new behaviour taken from `FunLove`. This possibility represents a serious threat.

V. Perspective for the Future

To perform analysis of malicious software, a system must be infected with malicious sample in order to observe its behaviour. Firstly, it is necessary to prepare a suitable environment - an analytic laboratory customised for this purpose. I consider preparation of fully controlled virtual system within one machine a suitable step to begin with. I will examine features of virtual systems that are available and choose the most suitable one. As previous chapter mentions, several types of malicious software are able to detect if they are executed in a virtual environment, therefore I need to study work of the researchers who are devoted to the problem.

Virtual environment simplifies management of several analytic systems and enables fast recovery of the system in case of problems. Another advantage of virtual system is that it allows to record its current state and control the state before infection with malicious sample and after it. Network interaction of malware will take place within the infrastructure of virtual machines and it will be possible to control it. However, besides the problem with virtual environment and sandbox detection, another drawback is that virtualisation places high requirements on RAM and hard drive.

Concerning experiments, I plan to employ techniques of both static and dynamic analysis. I believe that by using both approaches the results will be more detailed and accurate, in comparison with using only one approach. I plan to test various tools, e.g. Process Monitor[3], Process Hacker[4], Wireshark[5], Regshot[6], OllyDbg[7], to examine behaviour of already known malicious samples. The next step will involve extraction of specific actions, performed by samples, which could be considered as undesired and malicious and therefore could be used as *malware markers*. I expect that this task will be problematic due to the ambiguous nature of some actions performed by programs in the context of malware research. Some analytic tools allow to export results of analysis into external file, so this may be utilized by a custom program which will search the file for potential *markers*. The plan is to create a set (or various sets) of such *markers* and use them as a basis for formulation of malware models.

References

[1] F. B. Cohen, *A Short Course on Computer Viruses*, 2nd ed. New York, NY, USA: John Wiley & Sons, Inc., 1994.

[2] P. Szor, *The Art of Computer Virus Research and Defense*. Addison-Wesley Professional, 2005.

[3] Y. Ding, X. Yuan, K. Tang, X. Xiao, and Y. Zhang, “A fast malware detection algorithm based on objective-oriented association mining,” *Computers & Security*, vol. 39, Part B, no. 0, pp. 315 – 324, 2013.

[4] A. Moser, C. Kruegel, and E. Kirda, “Limits of static analysis for malware detection,” in *Computer Security Applications Conference, 2007. ACSAC 2007. Twenty-Third Annual*, Dec 2007, pp. 421–430.

[5] I. You and K. Yim, “Malware obfuscation techniques: A brief survey,” in *2010 International Conference on Broadband, Wireless Computing, Communication and Applications (BWCCA)*, Nov 2010, pp. 297–300.

[6] W. Landi, “Undecidability of static analysis,” *ACM Lett. Program. Lang. Syst.*, vol. 1, no. 4, pp. 323–337, Dec. 1992.

[7] M. Bailey, J. Oberheide, J. Andersen, Z. M. Mao, F. Jahanian, and J. Nazario, “Automated classification and analysis of internet malware,” in *Proceedings of the 10th International Conference on Recent Advances in Intrusion Detection*, ser. RAID’07. Berlin, Heidelberg: Springer-Verlag, 2007, pp. 178–197.

[8] K. Rieck, T. Holz, C. Willems, P. Dssel, and P. Laskov, “Learning and classification of malware behavior,” in *Detection of Intrusions and Malware, and Vulnerability Assessment*, ser. Lecture Notes in Computer Science, D. Zamboni, Ed. Springer Berlin Heidelberg, 2008, vol. 5137, pp. 108–125.

[9] G. Wagener, R. State, and A. Dulaunoy, “Malware behaviour analysis,” *Journal in Computer Virology*, vol. 4, no. 4, pp. 279–287, 2008.

[10] M. Egele, T. Scholte, E. Kirda, and C. Kruegel, “A survey on automated dynamic malware-analysis techniques and tools,” *ACM Comput. Surv.*, vol. 44, no. 2, pp. 6:1–6:42, Mar. 2008.

[11] Y.-M. Wang, R. Roussev, C. Verbowski, A. Johnson, M.-W. Wu, Y. Huang, and S.-Y. Kuo, “Gatekeeper: Monitoring auto-start extensibility points (aseps) for spyware management,” in *Proceedings of the 18th USENIX Conference on System Administration*, ser. LISA ’04. Berkeley, CA, USA: USENIX Association, 2004, pp. 33–46.

[12] P. Beaucamps, I. Gnaedig, and J.-Y. Marion, “Abstraction-based malware analysis using rewriting and model checking,” in *Computer Security ESORICS 2012*, ser. Lecture Notes in Computer Science, S. Foresti, M. Yung, and F. Martinelli, Eds. Springer Berlin Heidelberg, 2012, vol. 7459, pp. 806–823.

[13] A. Moser, C. Kruegel, and E. Kirda, “Exploring multiple execution paths for malware analysis,” in *Security and Privacy, 2007. SP ’07. IEEE Symposium on*, May 2007, pp. 231–245.

[14] L. Wu, R. Ping, L. Ke, and D. Hai-xin, “Behavior-based malware analysis and detection,” in *Complexity and Data Mining (IWCDM), 2011 First International Workshop on*. IEEE, Sept 2011, pp. 39–42.

[15] U. Bayer, I. Habibi, D. Balzarotti, E. Kirda, and C. Kruegel, “A view on current malware behaviors,” in *Proceedings of the 2Nd USENIX Conference on Large-scale Exploits and Emergent Threats: Botnets, Spyware, Worms, and More*, ser. LEET’09. Berkeley, CA, USA: USENIX Association, 2009, pp. 8–8.

[16] H. Lu, X. Wang, B. Zhao, F. Wang, and J. Su, “Endmal: An anti-obfuscation and collaborative malware detection system using syscall sequences,” *Mathematical and Computer Modelling*, vol. 58, no. 56, pp. 1140 – 1154, 2013.

[17] D. Canali, A. Lanzi, D. Balzarotti, C. Kruegel, M. Christodorescu, and E. Kirda, “A quantitative study of accuracy in system call-based malware detection,” in *Proceedings of the 2012 International Symposium on Software Testing and Analysis*, ser. ISSTA 2012. New York, NY, USA: ACM, 2012, pp. 122–132.

[18] J. R. Crandall, G. Wassermann, D. A. S. de Oliveira, Z. Su, S. F. Wu, and F. T. Chong, “Temporal search: Detecting hidden malware timebombs with virtual machines,” *SIGARCH Comput. Archit. News*, vol. 34, no. 5, pp. 25–36, Oct. 2006.

[19] T. Raffetseder, C. Kruegel, and E. Kirda, “Detecting system emulators,” in *Proceedings of the 10th International Conference on Information Security*, ser. ISC’07. Berlin, Heidelberg: Springer-Verlag, 2007, pp. 1–18.

[20] P. Ferrie, “Attacks on more virtual machine emulators,” 2007. [Accessed 2015-03-06]. [Online]. Available: http://pferrie.tripod.com/papers/attacks2.pdf

[21] T. Ormandy, “An empirical study into the security exposure to hosts of hostile virtualized environments,” 2007. [Accessed 2015-02-19]. [Online]. Available: http://taviso.decsystem.org/virtsec.pdf

[22] A. Dinaburg, P. Royal, M. Sharif, and W. Lee, “Ether: Malware analysis via hardware virtualization extensions,” in *Proceedings of the 15th ACM Conference on Computer and Communications Security*, ser. CCS ’08. New York, NY, USA: ACM, 2008, pp. 51–62.

[23] E. Konstantinou, “Metamorphic virus: Analysis and detection,” 2008. [Accessed: 2015-03-13]. [Online]. Available: https://www.ma.rhul.ac.uk/static/techrep/2008/RHUL-MA-2008-02.pdf

[3] Available: https://technet.microsoft.com/en-us/sysinternals/bb896645.aspx
[4] Available: http://processhacker.sourceforge.net/
[5] Available: https://www.wireshark.org/
[6] Available: http://sourceforge.net/projects/regshot/
[7] Available: http://www.ollydbg.de/

Universal Image Steganalysis

[1]*Martin BRODA (3rd year)*
Supervisor: [2]Dušan LEVICKÝ

[1,2]Dept. of Electronics and Multimedia Communications, FEI TU of Košice, Slovak Republic

[1]martin.broda@tuke.sk, [2]dusan.levicky@tuke.sk

***Abstract* — This paper is summarisational article, which is overview of my work for the last year of post gradual study. The aim of this paper is summarization of results that were reached in the design of steganalytic tool and comparison of specific classifiers in training and testing process. The last year of my work was dedicated to comparison of classifiers and their characteristics. Specifically, Bayes classifier and SVM (Support Vector Machine) classifier were verified. The next part of my work was focused on improvement of detection accuracy for proposed steganalytic tool. It was implemented by increasing the number of extracted parameters, from original 274 to 285 statistical features. Created models were tested for the detection accuracy, computational and time complexity. Popular steganographic methods (e.g. F5, Outguess and other) were used for testing of proposed steganalytic tool.**

***Keywords* — image, JPEG, secret message, steganalysis, steganography.**

I. INTRODUCTION

The main steganography objective is to ensure a secret transmission of data in background of non-privacy communication without additional security elements. Steganographic method can be used to illegal activities as terrorism or applied in environment where communication channel is monitored and user needs to transmit a secret message.

On the other hand, steganalysis is focused on detection and revealing of such communication mostly by employing a statistical analysis. Generally, the steganographic system is considered as broken once the steganalytic tool can decide whether testing medium contains a secret message or the original medium is intact while success probability is higher than random guessing.

II. THE INITIAL STATUS IN THE SOLVING OF THE RESEARCH TASK

Previous years, my research work was focused on the development of steganalytic tool that was able to detect secret message in static JPEG images. The principle of proposed method is based on the extraction of statistical features from DCT transform domain of static images. Image database is created from the group of cover and stego images and database diversity (images from different types of camera, various image resolutions and etc.) should be respected. The group of extracted statistical features has assigned identifier depending on whether parameters were extracted from a stego or cover image. The output of this process is subsequently used for the training of model using specific type of classifier. Finally, the created model for the specific steganographic tool (e.g. F5 [1] and Outguess [2]) is utilized in the detection of secret message in still image. Extracted statistical features are compared by classifier with trained model and the output of this comparing is result, whereby tested features can be assigned to cover or stego image. Block diagram of proposed steganalytic tool is illustrated in Fig. 1.

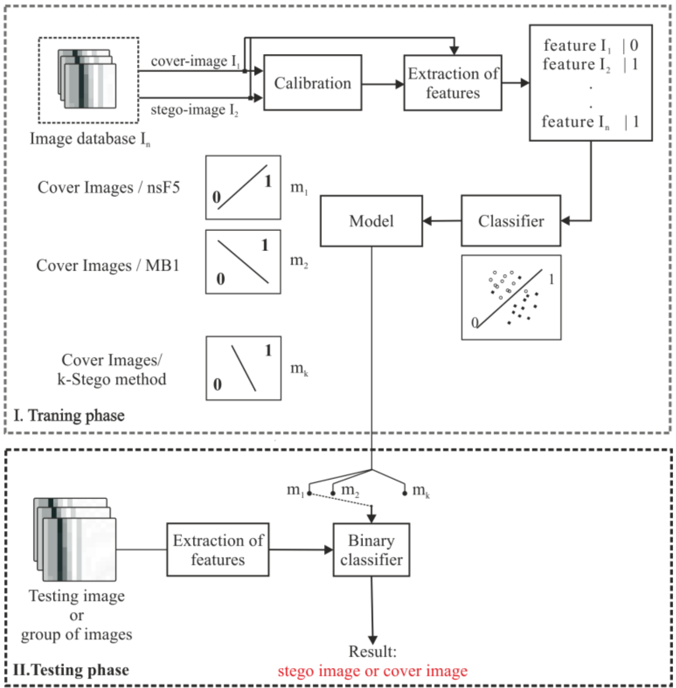


Fig. 1 Block diagram of proposed steganalytic tool for JPEG images

Previous years of study, research activity was focused on design of steganalytic tool and creating of required image database. The last year, my work was oriented on increasing of the detection accuracy. It was implemented to addition of statistical features and verification of different classifier types that can be used in this area. These steps are described in the next chapter in detail.

III. OBTAINED RESULTS

The last year of my research activity was focused on modification of steganalytic algorithm. It was implemented by increasing the number of extracted statistical parameters from 274 (specific features are described in [3], e.g. features from global, local histogram and others) to 285 parameters. Specifically, inter-block correlation parameters of DCT coefficients were added in this case (1).

$$D_{i,j} = d_{i,j} - d_{i,j+1} \tag{1}$$

where $D_{i,j}$ is matrix that is calculated by the difference between all adjacent blocks of DCT coefficients using horizontal sampling. Consequently, histogram (2) is defined from this matrix in interval <-5,5>, where is situated its maximum.

$$D = D_L, \ldots, D_R \tag{2}$$

where $L = min_{i,j,k}\, d_{i,j}$ a $R = max_{i,j,k}\, d_{i,j}$.

By this approach we can get next 11 statistical parameters, subsequently the length of statistical vector will be 285 features.

Comparison of detection accuracy for models with different length of statistical vector for steganographic tool Outguess is shown in Tab. 1. There is illustrated, that the addition of statistical parameters caused an increase of accuracy and the highest value is achieved currently in the new extraction scheme of 285 statistical features. It was published in paper on the international conference ELMAR in Croatia [3].

TABLE I. TRUE POSITIVE RATE AND ACCURACY OF DETECTION FOR BINARY CLASSIFICATION OF SPECIFIC STEGANOGRAPHIC METHODS

Algorithm	bpnz	FBS(66)		FBS(274)		FBS(285)	
		TPR[%]	ACR[%]	TPR[%]	ACR[%]	TPR[%]	ACR[%]
Outguess	0,05	70	71,2	53,6	74,6	56,9	76,1
	0,1	89,6	81	89,6	92,6	90,1	94,1
	0,15	96,4	84,4	95,6	95,6	95,1	95,9
	0,2	98,4	85,4	98,4	97	98,5	97,9
	0,25	97,9	85,2	98,8	97,2	98,9	98,1
	0,4	98,2	85,3	100	97,8	100	98,1

The significant part of research activity was focused on training of specific model that is comprised from cover and stego images obtained by MHF-DZ (Modulo Histogram Fitting with Dead Zone) steganographic tool. This method was proposed at department of electronics and multimedia communication. The model trained by MHF-DZ also achieved high detection accuracy for other steganographic tools (detected also steganographic methods F5, Outguess, JPHS and MB). Whereby we obtained simple universal steganalytic model trained using by one steganographic tool. These results were also published in the Radioengineering journal [5] covered by SCOPUS (IF 0.796).

Significant part of proposed image steganalytic method is represented by classifier. Therefore, the last year of PhD. study was also dedicated to this area. Comparison of classifier SVM (Support Vector Machine) a Bayes classifier was realized in the view of the influence on the detection accuracy, but also in the view of computational complexity for training model or testing. The comparison is showed in [6], where SVM classifier achieves better results of detection accuracy, but time and computational complexity is higher. For example, Bayes classifier was able to perform training of model Cover - MB2 with the 2000 images in less than 30 seconds. On other hand, SVM classifier achieved training time: 10 minutes in the same case. Specific test was executed using Intel Core i5 processor with the clock frequency 2,5 GHz. Comparison of the detection accuracy for different steganographic tools is shown in Fig. 2.

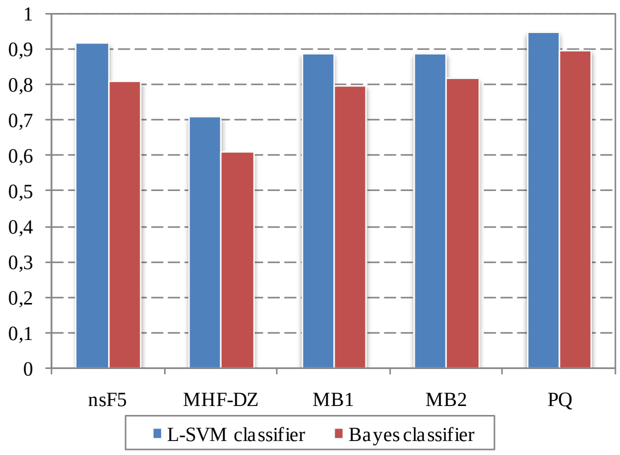


Fig.2 Comparison of accuracy for specific steganalytic model using L-SVM or Bayes classifier

It follows, that Bayes classifier has advantage in the area, where is necessary to achieve fast training model despite lower detection accuracy of secret message.

The last year, I also participated on Erasmus intensive program SUSCOMTEC (Intercultural Knowledge Transfer in Engineering for a Sustainable Global ICT Community) in Sofia, Bulgaria.

IV. FURTHER RESEARCH

Based on the analysis of previous results can be designed further research in the area of image steganalysis and steganography. This year, my work will be focused on design of steganalytic tool in DWT domain. In the next step, reducing computational complexity of popular and novel image steganalytic methods will be investigated.. Finally, my research will be focused on testing of detection accuracy of proposed steganalytic tools and comparing of obtained results with the existing systems of image steganalysis.

ACKNOWLEDGMENT

I would like to express my deep gratitude to prof. Ing. Dušan Levický, CSc., my research supervisors, for his patient guidance, enthusiastic encouragement and useful critiques of this research work.

REFERENCES

[1] A. Westfeld, "High capacity despite better steganalysis (F5 – a steganographic algorithm)," In I. S. Moskowitz, Information Hiding. New York: Springer-Verlag, pp. 289-302 (2001).

[2] N. Provos, Defending Against Statistical Steganalysis. Proc. 10th USENIX Security Symposium. Washington, DC, 2001.

[3] G. Bugár – V. Bánoci – M. Broda – D. Levický, A Novel Approach for Image Steganalysis, In: Proceedings of ELMAR-2014 : 56th International Symposium, Zadar, Croatia: Faculty of Electrical Engineering and Computing, University of Zagreb, 2014 P. 171-174. - ISBN 978-953-184-199-3 - ISSN 1334-2630.

[4] M. Broda – D. Levický – G. Bugár – V. Bánoci, "Universal Image Steganalytic Method Based on Binary Similarity Measures", In: Radioelektronika 2014. April 15-16, 2014, Bratislava, Slovak Republic. - Bratislava : STU, 2014 S. 1-4. - ISBN 978-1-4799-3713-4.

[5] V. Bánoci – M. Broda – G. Bugár – D. Levický, "Universal Image Steganalytic Method", In: Radioengineering, December 2014, Volume 23, Number 4, pp. 1213-1220. ISSN 1210-2512.

[6] V. Hajduk – M. Broda – D. Levický, "Bagging efficiency of machine learning in steganalysis", In: Electrical Engineering and Informatics 5 : Proceedings of the Faculty of Electrical Engineering and Informatics of the Technical University of Košice. - Košice : TU, 2014 S. 864-868. - ISBN 978-80-553-1704-5.

Utilizing of Disaggregation for Non-standard Situation Evaluation

[1]*Michal KOVALČÍK (2nd year)*
Supervisor: [2]František JAKAB
Consultant: [3]Peter FECIĽAK

[1,2,3]Dept. of Computers and Informatics, FEI TU of Košice, Slovak Republic

[1]michal.kovalcik@tuke.sk, [2]frantisek.jakab@tuke.sk, [3]peter.fecilak@tuke.sk

***Abstract*— Presented paper deals with the disaggregation of household energy consumption profile. Disaggregation in this context represents segmentation of total energy consumption profile to specific profiles of household appliances. To achieve this, NIALM algorithm and Montreal methods are utilized. Experimental disaggregation procedure of selected appliances are further presented. The output of the algorithm is in a form of an activity map of household appliances. In the conclusion, the purpose of acquired data is described, with the outlook to future research on the evaluation of non-standard situations regarding household appliances.**

***Keywords*—NIALM, smart metering, reactive power, active power.**

I. Introduction

The objective of the experiment described in this paper is to present the possibility of detecting the abnormal situations based on the energy consumption profile. Real-time reading consumption is enabled by sensors, once appliance is located beyond such sensor, detection of appliance's activity is possible. Example of such sensor is smart meter, device used to measure electric power. Disaggregation methods applied on signal recorded by smart meter allow to determine the appliance's activity. Another type of sensor is smart meter for flow of water or gas consumption measurements.

Beginnings of power disaggregation date back to 1980s. One of the first disaggregation methods is Nonintrusive appliance load monitoring (NIALM).

The output of the disaggregation, based on the overall power profile, determines what appliances are located in specific household. Another parameters that may be determined are time when specific appliance was turned on and off, and the duration of its active operation. Following this data, the map of standard behaviour is to be created for individual appliance. Self-organizing maps (SOM) may be used to evaluate the probability of presence of the appliance in the household.

II. Disaggregation of Selected Appliances from Overall Power Profile

The purpose of the experiment described in this chapter is the disaggregation of the power profile of selected appliance. Calibrated smart meter, having option of data export at minimal frequency of 1 Hz, is to be used for the exact disaggregation. This experiment is performed by 3 phase smart meter which allows calculation of apparent, active and reactive power. Data having the component of active and reactive power exported from smart meter are evaluated with more accuracy when compared to outputs of smart meter, which provides only information about apparent power. Relationship (1) between apparent, reactive and active power is as follows:

$$S = \sqrt{P^2 + Q^2} \qquad (1)$$

From this relationship can be concluded that various appliances can have the same apparent power. Identification of particular appliance may be still possible once we consider that each appliance has specific component of reactive power and specific component of active power. Example is comparison between electric kettle and electric motor of milling cutters – both having power rating of 2000 W, where by this power we mean power factor. However both of these appliances differ strongly in the ratio of reactive and active component. In kettle the active component of power prevails, it is due to its ohm nature, while only the parasitic capacitance or inductance is present. On the contrary the electric motor prevails while having inductive character and containing only parasitic capacitance or resistance. At first it might seem that both of these appliances are the same when it comes to the apparent power, however each of them is unique once the energy is decomposed to reactive and active power. Experimental procedure is to be based on just described fact.

Kettle and microwave oven will be used for the evaluation of experimental procedure, whole procedure may be described in three steps:

First step consists of generating the appliance's signature. This phase is also referred to as learning phase. Process itself utilizes calibrated smart meter, which allows export of reactive and active component. Data collection starts with the appliance being in the off state and ends when it is turned off again. Exported data may be displayed in a form of graph. The resulting graph is depicted in Fig. 1: the signature of apparent, active and reactive component of kettle (left) and microwave oven (right) is shown.

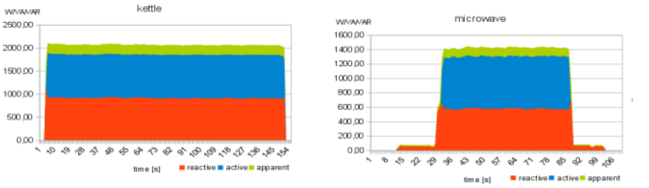


Fig. 1. Signature of kettle (left), microwave (right). Green color represents apparent, blue active and red reactive power.

Second step consists of gathering the data from all the household. Smart meter measures all three power components (in second intervals) and exports them. Fig. 2 illustrates the example of such data from two phases – phase no.3 on left and phase no.1 on right. From this graph can be concluded that in last seconds the activity of appliances appeared. Such activity was initiated by turning on the appliance until the end (turn off) of its operation, i.e. same state as before its operational state.

Fig. 2. Overall power collected by smart meter on third phase (left) and first phase (right) from household (interval of 15 minutes)

Third step: Algorithm NIALM and method Montreal is applied, see [1] for its detailed description. In [2], [3] its fundamental principle as creation of steady state levels of powers with certain tolerance is described. Following, the edge detection of two consecutive steady state levels is carried out. These edges are paired based on the difference values of consecutive steady state levels, i.e. positive values are paired with negative. Match is sought based on the comparison of these pairs with signature database (for all power components), further start and end of the appliance's activity is determined. The output of this algorithm has to following structure:

- appliance ID,
- additional information about appliance,
- power as collected by smart meter after disaggregation,
- time information of appliance's operation (start),
- time information of appliance's operation (end),
- duration of appliance's operation in seconds.

Output of the algorithm (example):

Paering >> 2095,8 <==> 1995,6 DATE 2015-03-23 17:53:47
Paering >> 2095,8 <==>-1950,8 DATE 2015-03-23 17:57:45
FINAL> kettle==> 1973,2 DATE FROM 2015-03-23 17:53:47 DATE TO 2015-03-23 17:57:45 DIFF 238sec.

This output may be represented in the form of graph depicted in Fig. 3 – red colour marks disaggregated appliance from total household energy profile (marked with yellow colour).

NIALM algorithm can be applied to any time sequence with the accuracy of 80-90 percent. Disaggregation in the real time may be achieved by NIALM algorithm and Montreal method providing that signature database of the whole household is available. Contribution of such process is the ability to calculate the consumption for specific appliances, monitoring of their operation with related user notification, further statistical processing and other options are to be available.

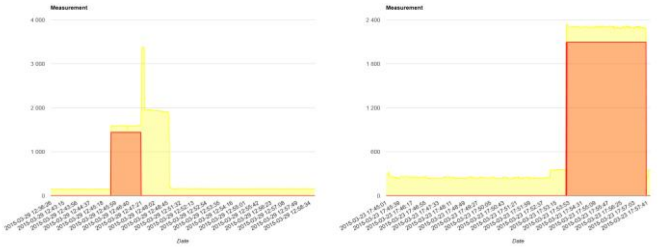


Fig. 3. Result after applying NIAM algorithm. Disaggregated power of kettle (left) and microwave oven (right)

III. Conclusion

The goal of this experimental procedure was to determine the accuracy of the algorithm presented in [4], [5] and its utilization for further research. During the testing, the algorithm was able to determine the number of appliances in operation, both for kettle and microwave oven.

The database of appliances operation in household is created based on the data acquired from the disaggregation algorithm. Further step is to apply self-organizing maps (SOM) mechanism and create the map of household standard behaviour. In this way is to possible to deploy the learning process, which is to be applied to a system for detection of non-standard behaviour, further described in [1]. Next we will be able to detect non-standard behaviour in relation to disaggregation of appliances. This system is to allow non-standard behaviour detection, which can be a result of abnormal behaviour of appliances, e.g. its high power consumption, excessive use, inefficiency or simply it may detect appliance that should be turned off. As for the near-future goals, we focus on the enhancement of the model for non-standard behaviour detection in household appliances.

Acknowledgment

Paper is the result of the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDFF. We support research activities in Slovakia/This project is being co-financed by the European Union.

References

[1] Kovalčík, M.: Metódy a prostriedky vyhodnocovania neštandardných situácií v inteligentnej domácnosti: Písomná práca k dizertačnej skúške: Košice: TUKE, 2015
[2] MARCEAU, Medgar Louis; ZMEUREANU, R. Nonintrusive load disaggregation computer program to estimate the energy consumption of major end uses in residential buildings. Energy Conversion and Management, 2000, 41.13: s. 1389-1403.
[3] G.W. Hart, Nonintrusive Appliance Load Monitoring, Proceedings of the IEEE 80 (1992) 1870-1891.
[4] S. Hayes and J. Cone. Reduction of residental consuption of electricity through simple monthly feedback. Applied behavior analysis.
[5] ZEIFMAN, Michael; ROTH, Kurt. Nonintrusive appliance load monitoring: Review and outlook. Consumer Electronics, IEEE Transactions on, 2011, 57.1: s. 76-84.

Author's index

15th Scientific Conference of Young Researchers
of Faculty of Electrical Engineering and Informatics
Technical University of Košice

Proceedings from Conference

Published: Faculty of Electrical Engineering and Informatics
Technical University of Košice
I. Edition, 334 pages
Number of CD Proceedings: 50 pieces

Editors: Prof. Ing. Alena Pietriková, CSc.
Ing. Milan Nosáľ
Ing. Jana Šťastná
Ing. Emília Pietriková, PhD.

ISBN 978-80-553-2130-1