ISSN 1310-8271

JOURNAL

OF THE TECHNICAL UNIVERSITY - SOFIA
PLOVDIV BRANCH, BULGARIA

FUNDAMENTAL SCIENCES AND APPLICATIONS

Volume 23, 2017

Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"

EDITORIAL BOARD

© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 23, 2017

CONTENTS

© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 23, 2017

FUZZY MODELING AND SIMULATION OF GAS TURBINE USING FUZZY CLUSTERING ALGORITHM

ABDELHAFID BENYOUNES [(1)], AHMED HAFAIFA [(1*)], ABDELLAH KOUZOU [(1)] , MOULOUD GUEMANA [(2)]

Abstract: *Gas turbines are one of the major parts of modern industry. They have played very important role in aeronautical industry, power generation and main mechanical drivers for large pumps and compressors. This study addressed the modeling and the simulation of the Industrial Gas Turbine GE ms5001P, located in the electrcal production station of M'sila in Algeria. The used method for modeling of this gas turbine is based on fuzzy inference system with the use of Gustafson-Kessel clustering (GK) algorithm.*

Key words: *Gas turbine modeling; fuzzy modeling; fuzzyinference system; GK clustering algorithm*

1. INTRODUCTION

Currently, the development of mathematical models for the representation and approximation of complex nonlinear systems is an essential subject in several engineering disciplines. The need for a strong understanding of physical phenomena in industrial systems is a great restriction at the practical level when dealing with complex non-linear systems. Indeed, the equation of the laws governing such systems generally leads to a too complex model of knowledge and its implementation is delicate. In this case, the use of modeling techniques developed from the input / output measures collected on the system is required.

For this reason, this work relies on fuzzy logic and its tools to present a complete and integrated approach to solve all the problems encountered in the use of classical modeling and control methods. The problem of exploiting fuzzy models based on Takagi-Sugeno approach, using input / output data collected on the system being tested, for a gas turbine system application.

Indeed, Takagi-Sugeno fuzzy modeling is a universal approximation of real systems, which has shown these efficiencies in several applications in the literature [17, 24, 28, 32, 38, 81, 86]. In 1985 Takagi and Sugeno proposed the use of properties of fuzzy sets and the use of rules in a fuzzy model (TS fuzzy model). Since then, the fuzzy model of Takagi and Sugeno TS has proved its effectiveness in the study of many other nonlinear dynamic systems. Conversely, conventional approaches use a single model to describe the overall behavior of a nonlinear system, whereas fuzzy Takagi and Sugeno TS models essentially use a multi-model approach, in which the simple sub-models are linear and are combined between them, for the purpose of describing the overall non-linear system behavior.

This strong property of fuzzy models of type Takagi and Sugeno can be applied in several dynamic systems modeling applications, that can be described by differential equations. In this paper, fuzzy identification and modeling from experimental data will be presented to approximate nonlinear systems. In this framework, this work is intended for the design of an original method for the identification and modeling of industrial systems, with a view to its application to a gas turbine system. This model was trained by use of real operational data of a GE MS5001P gas turbine used for electrical energy production.

2. PROCESS DESCRIPTION

The GE MS 5001P gas turbine engine is made of three main sections [1]: the axial compressor, the

 ISSN 1310 - 8271

combustion chamber and the turbine. There is a variable inlet guide vane (IGV) in the inlet of the axial compressor and a variable nozzle guide vane (NGV) in the turbine section [3]. In other part, two valves are used as the main controlling devices of the gas turbine located on the inlet fuel line of the combustion chamber, the first one is for the control of the inlet fuel pressure which called the stop ratio valve (SRV) and the second, the most important one, is used for governing the speed of the load shaft this one called the gas control valve (GCV). The figure (figure.1) represents the schematic of the gas turbine system.

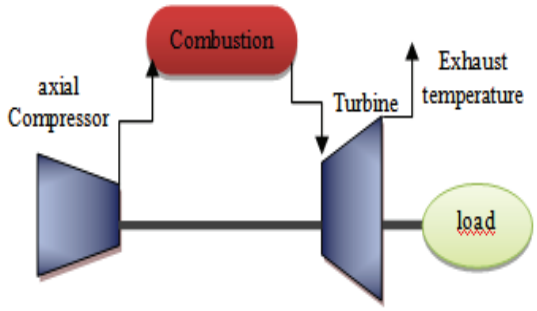


Fig. 1. *Signal shaft gas turbine GE MS 5001P*

The gas turbine under investigation in this work is the GE MS 5001P industrial gas turbine specifically, this turbine is designed for mechanical drive applications with a wide operating speed range to meet operating conditions of the most common driven equipment In our case it is used as an turbo generator in the M'sila electrical production station located in M'sila –Algeria (figure 2). The design point specifications of this gas turbine are presented in Table I. The detailed information can be found in the product manuals.

Fig. 2. *M'sila electrical production station*

Table 1. *Examined gas turbine specifications parameters*

Quantity	Value
Compressor stage	16
Firing temperature	1.730 (F°)
Exhaust temperature	898 (F°)
Air flow	928.5 (Lb/hr)
Output	24,700 (kw)
Heat rate	12,950 (kJ/kW-h)

A. Nonlinear System Modeling

The fuzzy modeling from experimental data are effective tools to approximate a non-linear system. Among the models are widely used in the modeling techniques we find those of Takagi-Sugeno (TS) [12, 15].

A TS model uses the idea of linearization of the fuzzy regions in the state space. Based on these regions fuzzy, a non-linear system can be decomposed into a multi structure models consisting of several linear models which are not necessarily independent [2]. The fuzzy sets premises partition the input space in a number of fuzzy regions, while the functions consequences describe the behavior of the system in these regions. The fuzzy model TS is generally built in two steps:

Step 1: Determine the membership functions (MFs) antecedent of the rules;

Step 2: Estimate the parameters of the functions consequences.

One of the techniques used to achieve the first step is the fuzzy clustering, in this paper apply the Gustafson-Kessel (GK) algorithm.

B. Gustafson-Kessel algorithm

The GK algorithm constitutes a reference among the different among different methods of fuzzy classification based on minimizing the objective function of the form:

$$J_{FCM}(Z;U;V)=\sum_{i=1}^{c}\sum_{K=1}^{N}(\mu_{ik})^{m}D_{ikA}^{2} \qquad (1)$$

Where Z is the data set, $U=[\mu_{ik}]$ is the matrix of fuzzy partition (size $C\times N$) and $V=[v_1,v_2,...,v_c]$

is the vector of the center of classes to be determined with $v_i \in \Re^n$ the center of the i^{th} class $1 \prec i \prec c$, $m \in [1,+\infty]$ is a factor that denotes the degree of fuzziness of the partition.

The standard quadratic distance in space in question, which defines the distance measure between the Z_k observation and the v_i center within the meaning of the metric induced by A.

Gustafson and Kessel in employing a standard of adaptive distance in the purpose of detecting of classes of different geometrical shapes in a set of data. In this case, each class has its own matrix of standard, which leads to:

$$D_{ikA}^2 = \|z_k - v_i\|_A^2 = (z_k - v_i)^T A_i (z_k - v_i) \quad 1 \le i \le c, 1 \le k \le N \tag{2}$$

It is assumed that the matrix A_i tested the hypothesis:

$$|A_i| = \rho_i \ , \ \rho_i > 0 \tag{3}$$

Where ρ_i is fixed for each class.

In this case, the optimization of (1) gives us the following expression for A_i :

$$|A_i| = [\rho_i \det(F_i)]^{\frac{1}{n}} F_i^{-1} \tag{4}$$

Where F_i is the covariance matrix blurred the i-th class given by:

$$F_i = \frac{\sum_{K=1}^{N} (\mu_{ik})^m (z_k - v_i)(z_k - v_i)^T}{\sum_{K=1}^{N} (\mu_{ik})^m} \tag{5}$$

In the equation (1), the measurement of non-similarity is expressed by of the sum of squares of distances between each data vector and the center of the corresponding class. The effect of this distance is weighted by the degree of activation μ_{ik}^m corresponding to the Z_k data vector, the value of the cost function $J_{FCM}(Z;U;V)$ can be seen as a measure of the total variance of Z_k with respect to the v_i centers. The minimization of the objective function (1) is given as flows:

$$\mu_{ik} = \frac{1}{\sum_j^c 1 \left(\frac{D_{ikA}}{D_{jkA}} \right)^{\frac{2}{m-1}}} \qquad 1 \le i \le c, 1 \le k \le N$$

$$v_i = \frac{\sum_{K=1}^{N} (\mu_{ik})^m z_k}{\sum_{K=1}^{N} (\mu_{ik})^m} \tag{6}$$

This leads to the GK algorithm given in three steps:

Step 1: Calculate the cluster centers

$$v_i^l = \frac{\sum_{K=1}^{N} (\mu_{ik}^{(l-1)})^m z_k}{\sum_{K=1}^{N} (\mu_{ik}^{(l-1)})^m} \qquad 1 \le i \le c$$

Step 3: Calculate the covariance matrix

$$F_i = \frac{\sum_{K=1}^{N} (\mu_{ik}^{(l-1)})^m (z_k - v_i^{(l)})(z_k - v_i^{(l)})^T}{\sum_{K=1}^{N} (\mu_{ik})^m}$$

Step 3: Calculate the distances

$$D_{ikA}^2 = (z_k - v_i^l)^T A (z_k - v_i^l) \quad 1 \le i \le c, 1 \le k \le N$$

Step 3: Update the partition matrix

If $D_{ikA}{}^2 > 0$ *for* $1 \le i \le c, 1 \le k \le N$

$$\mu_{ik}^{(1)} = \frac{1}{\sum_j^c 1 \left(\frac{D_{ikA}}{D_{jkA}} \right)^{\frac{2}{m-1}}}$$

Otherwise $\mu_{ik}^{(1)} = 0$ if $D_{ikA} > 0$ and $\mu_{ik}^{(1)} \in [0,1]$ with $\sum_i^c \mu_{ik}^{(1)} = 1$ too $\left\| U^{(l)} - U^{(l-1)} \right\| < \varepsilon$.

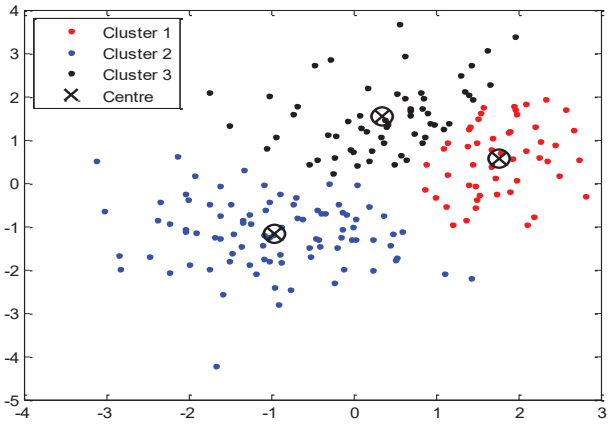


Fig. 3. *Gustafson Kessel clustering algorithm*

C. Construction of Takagi-Sugeno models

Consider a system described by the equation (7) :

$$y_k = f_{NL}(x_k) \qquad (7)$$

Our objective is to approximate the nonlinear function f_{NL} of equation (7) by Takagi-Sugeno model (TS):

R_i : if x_1 is A_{i1} and x_2 is A_{i2} and … x_p is A_{ip}

Then $y_i = a_i x + d_i \quad i = 1,...,r$

$a_i^T = [a_1, a_1, ..a_{in}]$ With R_i represents the $i^{iéme}$ rule

$x = [x_1, x_1, ..x_n]$ Vector of observations

A_{i1}, A_{i2},…, A_{in} : Represents the fuzzy sets,

y_i : represent the output of $i^{iéme}$ rule $a_i \in R^p$ is the vector of parameters and d_i is scalar .

The determination of f_{NL} is done in two steps:

Step1. It starts first of all by apply the clustering algorithm Gk, in order to calculate the matrix of fuzzy partition U.

Step 2. We then estimate the parameters *a* and *b* . In effect, the method of defuzzification used in the model of Takagi-Segeno, is linear in relation to the parameters. However, these parameters can be estimated by using the techniques of least-squares.

3. GAS TURBINE SYSTEM MODELING

In this study, the identification of gas turbine system is carried out in closed loop, or the control type is isochronous, the use of an open-loop procedure for identification is not preferred. The outputs inputs data have to be correlated together and generated in normal operating conditions during 5 hours and half, the inputs parameters are axial compressor discharge pressure and temperature (PCD)(TCD), Gas control valve opening (GCV), and the outputs are the speed of the turbine and Exhaust temperature (ET) Figure(4).

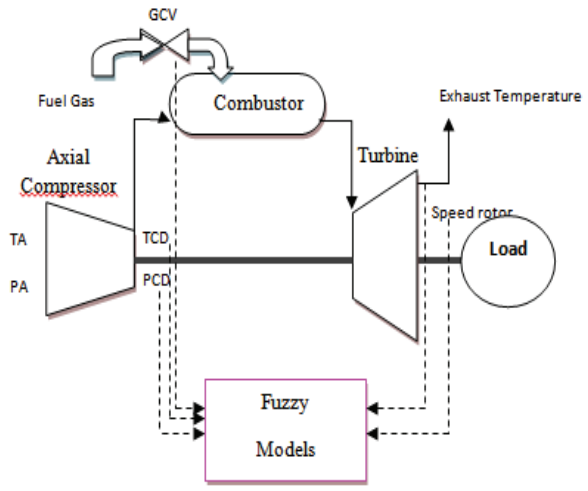


Fig. 4. Examined gas turbine inputs and outputs used in modeling

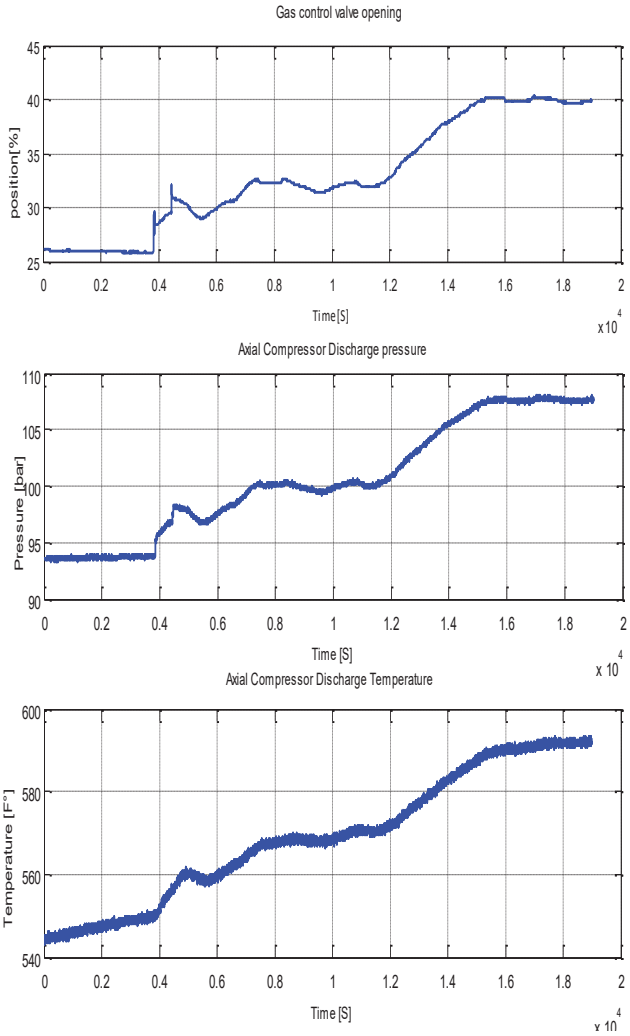


Fig. 5. Validation data of the examined gas turbine inputs

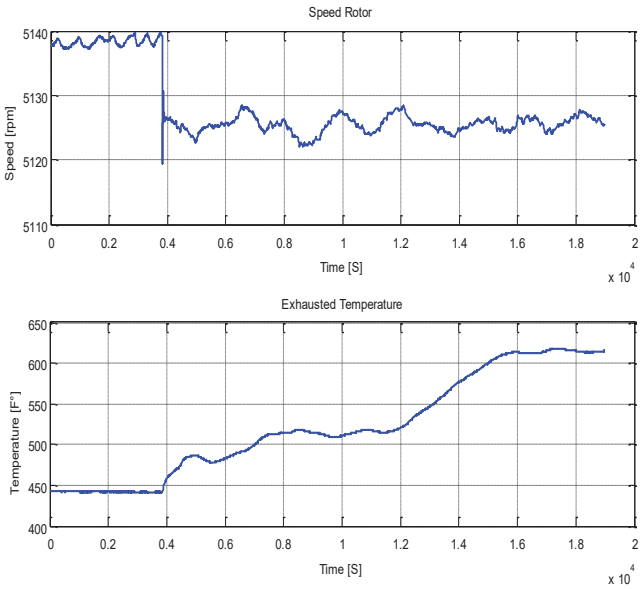


Fig. 6. *Validation data of the examined gas turbine outputs*

The obtained result of this simulation is shown in Figures 7 and 8 representing respectively the two main outputs of the fuzzy models ; speed turbine and exhausted temperature. Figure 9 and 10 shown the membership functions used in fuzzy models.

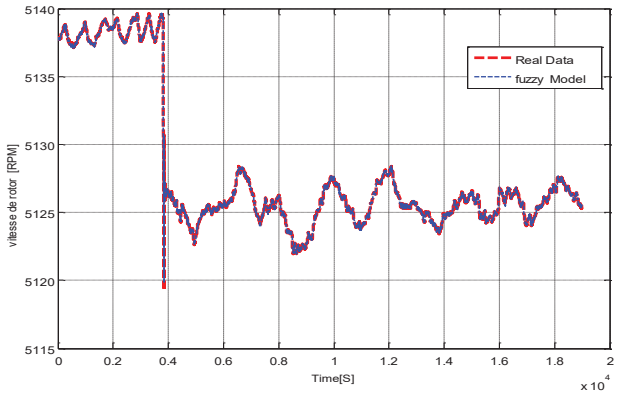


Fig. 7. *Variation of obtenied model of speed rotor*

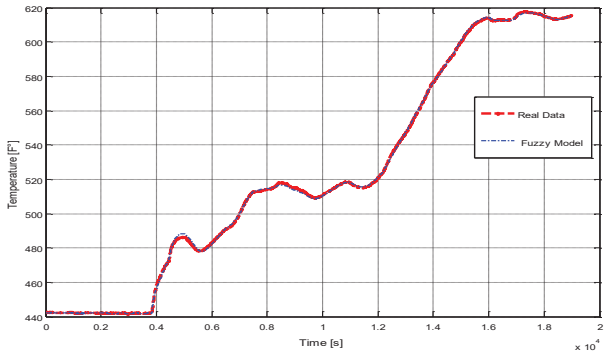


Fig. 8. *Variation of obtenied model of the exhaust temperature*

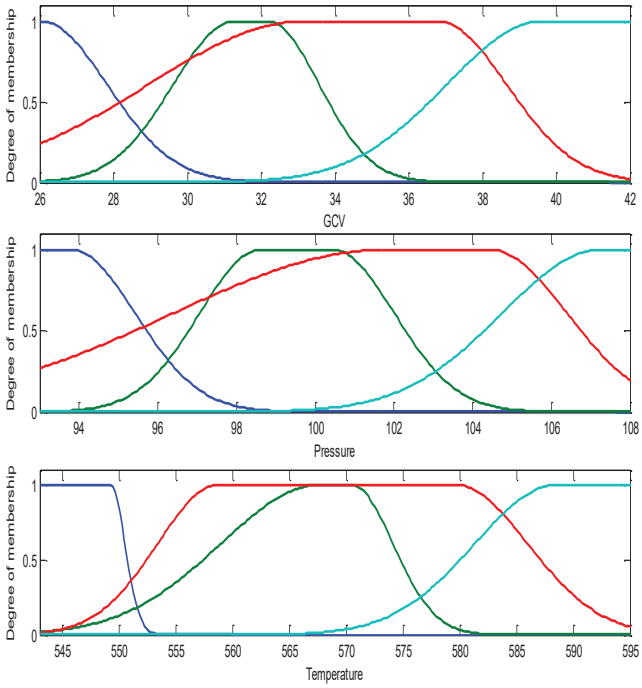


Fig. 9. *Membership functions used in fuzzy models inputs*

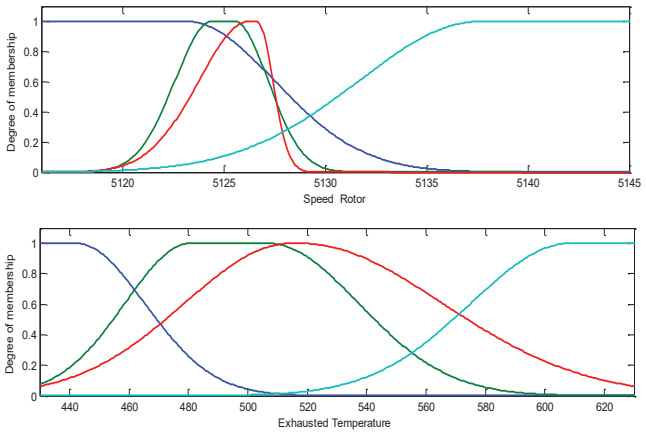


Fig. 10. *Membership functions used in fuzzy models outputs*

The parameter used for the numeric validation is the root-mean-square error (RMSE). and is shown in Table II.

D. Root mean square error (RMSE)

The RMSE is a frequently used measure of the difference between values predicted by a model and the values observed from the environment that is being modeled. These individual differences are also called residuals.The RMSE of a model prediction with respect to the estimated variable X model is defined as the square root of the mean squared error, given by:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (X_{obs,i} - X_{mo\,del,i})^2}{n}} \quad (8)$$

Table 2. *Table I: RMSE comparison between the two model*

	Gas turbine parameter	
Model type	LP shaft speed	ET system
Fuzzy model	0,0076	0,011846

4. CONCLUSION

This work has addressed one of the major problems when looking for a reliable mathematical representation; the proposed Fuzzy model provides a good improvement in performance during its operation in the examined gas turbine. The use of a GK fuzzy clustering algorithm has the important advantage to allow the automatic generation of membership functions of fuzzy regions from studied data. This work, confirm that the development of digital approach is obviously More flexible to implement. The obtained results from data classification with the associated models construction offer advantageous performance in modeling of the examined gas turbine system. This approach can provide reliable models for controlling the gas turbine and system Fault diagnosis.

REFERENCES

1. Abdelhafid Benyounes, Ahmed Hafaifa and Mouloud Guemana, Fuzzy logic addresses turbine vibration on Algerian gas line. Oil & Gas Journal, 2016, vol. 114, no. 1, pp. 22-28.
2. Abdelhafid Benyounes, Ahmed Hafaifa and Mouloud Guemana, Gas turbine modelling based on fuzzy clustering algorithm using experimental data. Journal of Applied Artificial Intelligence, Taylor & Francis, 2016, vol. 30, no. 1, pp. 29-51.
3. Abdelhafid Benyounes, Ahmed Hafaifa, Abdellah Kouzou1 and Mouloud Guemana, Gas turbine modeling using adaptive fuzzy neural network approach based on measured data classification. Mathematics-in-Industry, Case Studies, vol.7, no.4, Springer 2016, DOI 10.1186/s40929-016-0006-3.
4. Ahmed Boubenia, Ahmed Hafaifa, Abdellah Kouzou, Kamal Mohammedi and Mohamed Becherif, Carbone dioxide capture and utilization in gas turbine plants via the integration of power to gas. Petroleum Journal, March 2017,
5. Ahmed Hafaifa, Guemana Mouloud, and Belhadef Rachid, Fuzzy Modeling and Control of Centrifugal Compressor Used in Gas Pipelines Systems. Multiphysics Modelling and Simulation for Systems Design and Monitoring, Book Chapter, Applied Condition Monitoring vol. 2, 2015, pp.379-389.
6. Babuška R., Verbruggen H.B., Identification of composite linear models via fuzzy clustering. Proceedings of the European Control Conference, Rome, Italy, 1995, pp. 1207–1212.
7. Babuška Robert, Fuzzy modeling for control. International Series in Intelligent Technologies, Volume 12, Kluwer Academic, Publishers Norwell, MA, USA, ISBN: 0792381548, 9780792381549, 1998.
8. Balazs Feil, Janos Abonyi, Ferenc Szeifert, Model order selection of nonlinear input–output models: A clustering based approach. Journal of Process Control, 2004, vol. 14, no. 6, pp. 593-602.
9. Berchtold M., Riedel T., Decker C. and Van Laerhoven K., Gath-Geva specification and genetic generalization of Takagi-Sugeno-Kang fuzzy models. IEEE International Conference on Systems, Man and Cybernetics. 2008, vol. SMC 2008, pp. 595 – 600.
10. Bezdek J.C. and Dunn, J.C., "Optimal Fuzzy Partitions: A Heuristic for Estimating the Parameters in a Mixture of Normal Distributions". IEEE Transactions on Computers, 1975, vol. C-24, no. 8, pp. 835 – 838.
11. Bezdek J.C. and Dunn, J.C., Optimal fuzzy partitions: A heuristic for estimating the parameters in a mixture of normal distributions. IEEE Transactions on Computers, 1975, vol. C-24, no. 8, pp. 835 – 838.
12. Buckley J.J., Universal fuzzy controllers. Automatica, 1992, vol. 28, no. 6, pp. 1245–1248.
13. Enrique H. Ruspini, A new approach to clustering. Information and Control, 1969, vol. 15, no. 1, pp. 22–32.
14. Ford C.L., Carrotte J.F., Walker A.D., The application of porous media to simulate the upstream effects of gas turbine injector swirl vanes. Computers & Fluids, 2013, vol. 77, pp. 143-151.
15. Günyaz Ablay, A modeling and control approach to advanced nuclear power plants with gas turbines. Energy Conversion and Management, 2013, vol. 76, pp. 899-909.

16. Gustafson D.E. and Kessel W.C., Fuzzy clustering with a fuzzy covariance matrix. Proceedings of IEEE Conference on Decision and Control including the 17th Symposium on Adaptive Processes, 1978, vol. 17, Part 1, pp. 761 – 766.
17. Kulikov G.G. and Thompson H.A., Dynamic modeling of gas turbines. Book, 1st edition 2004, London, Springer.
18. Mohamed Ben Rahmoune, Ahmed Hafaifa, Mouloud Guemana, Fault Diagnosis in Gas Turbine Based on Neural Networks: Vibrations Speed Application. Book Chapter in Advances in Acoustics and Vibration, Volume 5 of the series Applied Condition Monitoring pp 1-11, on line. 02 September 2016, ISBN: 978-3-319-41458-4, 2017.
19. Nadji Hadroug, Ahmed Hafaifa, Abdellah Kouzou and Ahmed Chaibet, Faults detection in gas turbine using hybrid adaptive network based fuzzy inference systems to controlling there dynamic behavior. DIAGNOSTYKA the Journal of Polish Society of Technical Diagnostics (PSTD), 2016, vol. 17, no. 4, pp. 3-17.
20. Nadji Hadroug, Ahmed Hafaifa, Kouzou Abdellah and Ahmed Chaibet, Dynamic model linearization of two shafts gas turbine via their input / output data around the equilibrium points. Energy Elsevier Journal, 2016, DOI: http://dx.doi.org/10.1016/j.energy.2016.11.099
21. Rowen W.I., Simplified mathematical representations of heavy duty gas turbines. ASME J. Eng. Power, 1983, vol. 105, pp. 865–869.
22. Ruano, A. E., Fleming, P. J., Teixeira, C., Rodrı Guez-V azquez, K., and Fonseca, C. M., Nonlinear identification of aircraft gas turbine dynamics. Neurocomputing, 2003, vol. 55, no. 3–4, pp. 551–579.
23. Setnes M., Babuška R., Verbruggen H.B.,Complexity reduction in fuzzy modeling. Mathematics and Computers in Simulation, 1998, vol. 46, no. 5–6, pp. 507-516.
24. Takagi T. and Sugeno M., Fuzzy identification of systems and its applications to modelling and control. IEEE Transactions on Systems Man and Cybernetics, 1985, vol. SMC-15, no. 1, pp. 116 – 132.

Authors' contacts

Abdelhafid Benyounes is with the Applied Automation and Industrial Diagnostics Laboratory, Djelfa University, Algeria (Email: ab_benyounes@univ-djelfa.dz).

Abdellah Kouzou is with the Applied Automation and Industrial Diagnostics Laboratory, Djelfa University, Algeria (e-mail: kouzouabdellah@ieee.org).

Mouloud Guemana is with the Science and Technology Faculty, Médéa University, Algeria. (e-mail: guemana.mouloud@univ-medea.dz)

Ahmed Hafaifa is with the Applied Automation and Industrial Diagnostics Laboratory, Djelfa University, Algeria. (e-mail: hafaifa.ahmed.dz@ieee.org).

© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 23, 2017

POWER PLANT SYSTEM IDENTIFICATION BASED ON EXTENDED LEAST SQUARE KRONECKER OPERATOR

BACHIR NAIL[1], ABDELLAH KOUZOU[1], AHMED HAFAIFA[1], AHMED CHAIBET[2]

Abstract: *This paper focuses on finding an approximate Multivariable Output-Error (MIMO OE) model of energy process (power plant) based on experimental information acquired on site. In order to obtain the parametric model, the Extended Least Square with the help of the Kronecker Operator (ELSK) and the Left Matrix Fraction Description (LMFD) theory were used. four validations criteria are used in order to select the best model order, and taking into account that the model must be left coprime (no (pole/zero) cancellation). With the objective to prove the reliability of this model and the estimator, a comparative study with recent estimation algorithm was conducted, such as: Multivariable Output Error State Space (MOESP) and with artificial intelligent model, Adaptive Neuro-fuzzy Inference System (ANFIS).*

Key words: *Multivariable model, Output-error, ELSK, LMFD, Validations criteria, MOESP, ANFIS, energy process, power plant.*

1. Introduction

The implementation of a strategy of control and regulation of an industrial engineering systems requires the use of reliable models of it which can more exploitable in control. The main aime in this work is the modelling of the dynamics behavior of a energy process, based on system identification theory using experimental data. Due to the increasing complexity of this equipment and their severe operating constraints increase with their use of added values on these supervisory strategies.

In this framework, the examined system is a Turbo-alternator (gas turbine+alternator) power plant installed in unit of production of electricity in M'sila Algeria, this plant is an internal combustion engine that uses the gaseous energy of the air, converting the chemical energy of the fuel into mechanical energy, it is designed to extract, as far as possible, the energy of the fuel. Indeed, gas turbines are also called combustion turbines, are used in a wide range of applications, including power generation, natural gas transmission, as well as various process applications..etc. However, a gas turbine is an internal combustion engine, which operates with rotary motion and reciprocation. These gas turbines are essentially composed of three main elements: the axial compressor, the combustion chamber and power turbine , and alternator drived by this gas turbine (GE MS 5001P).

MANY results have been published about the modelling of gas turbine and energetic plant using largely the artificial intelligence theory (ANFIS, Fuzzy Logic and Neural Network models), for the following grounds easy to apply and unconditioned constraints in the quality and the characterisation of the data (the convergence, the smoothing and the number of samples is not required), but there are disadvantages on these models, the difficulty of controlling, and the problem of instability, also the limitations of applying the theory of advanced control for the reason of the lack of the mathematical model, majority of the researches they have been done in recent years in this area : Gas turbine modeling based on fuzzy clustering algorithm [1], Fuzzy Modeling of Centrifugal Compressor [2], Gas turbine modelling using fuzzy logic and artificial neural networks[3], and others.

Therefore, the parametric system identification theory is used to solve the problem of approximate mathematical model which has the same characterization of the empirical model, this paper dealing with the implement the theory of left matrix fraction description (MFD) to identify the behavior of the turbo-alternator plant multivariable output-error model. In the literature many paper have been done in field of the (MFD) system identification in the left and the right among them : Identification of Turbo-compressor using (LMFD)

 ISSN 1310 - 8271

[4], instrumental variable identification and extending the SRIV algorithm methods for LMFD models[5],[6] , ARMAX models identification using MFD [7], and MIMO least squares using MFD [8], and others [9],[10].

In our study, the MIMO OE model is considered to represent the dynamic behaviour of energy process (Turbo-alternator) by the extended least square kronecker operator estimator based on LMFD theory, using inputs/outputs real data, the inputs signals, GCV: gas control valve, PCD: Pressure compressor discharge and TCD: Temperature compressor discharge, and the outputs signals, ET: Exhaust Temperature, RS: Rotor Speed, and PG: Power Generation, the experimental data used in this identification acquired on site of length=10000 samples in time of 167 minutes, the selection of the final model needed to the validations criteria decision, we have using four criteria (AICs, FPEs, RMSE, and VAF) [11], in order to test the ratability of the obtained model and implemented estimator a comparative study has been done with intelligent artificial model ANFIS and MOESP algorithm introduced by Katayama [12].

The paper is organized as follow: first an introduction discussed the historical and the recent development of system identification, then it is followed by the main procedure of ELSK with application on energetic plant and a discussion of the obtained results. Finally comments, perspectives and a conclusion will finish the paper.

2. System identification

A MIMO OE (output-error) model given as:

$$A(q^{-1})y[k] = B(q^{-1})u[k] + e[k] \quad (1)$$

Can be written in LMFD form as:

$$y[k] = A(q^{-1})^{-1}B(q^{-1})u[k] + A(q^{-1})^{-1}e[k] \quad (2)$$

Where $u[k] \in R^m$ and $y[k] \in R^p$ are inputs and outputs vectors of the system respectively, while $e[k] \in R^p$ is a white-noise signal and the polynomial matrices $A(q^{-1})$, $B(q^{-1})$ have the following structure:

$$A(q^{-1}) = I_p + A_1 q^{-1} + ... + A_{na} q^{-na} \quad (3)$$

$$B(q^{-1}) = B_1 q^{-1} + ... + B_{nb} q^{-nb} \quad (4)$$

The objective is to identify the matrix coefficients $A_i \in R^{p\times p}$ and $B_i \in R^{p\times m}$ of the matrix polynomials $A(q^{-1})$ and $B(q^{-1})$.

Expanding equation (2) yields

$$\begin{aligned} y[k] = &-A_1[k-1] ... - A_{na} y[k-na] + \\ &B_1 u[K-1] ... + B_{nb} u[k-nb] + e[k] \end{aligned} \quad (5)$$

Then using the Kronecker operator we get

$$\begin{aligned} \left[I_P \otimes y[k]^T\right] col(I_P) = &-\left[I_P \otimes y[k-1]^T\right] col\left(A_1^T\right) \\ &... - \left[I_P \otimes y[k-na]^T\right] col\left(A_{na}^T\right) \\ &+ \left[I_P \otimes u[k-1]^T\right] col\left(B_1^T\right) \\ &... + \left[I_P \otimes u[k-nb]^T\right] col\left(B_{nb}^T\right) \\ &+ e[k] \end{aligned} \quad (6)$$

Here, the col{.} operator is one which forms a vector from a matrix by stacking its columns on top of one another.

Moreover, equation (6) can be written as

$$\begin{aligned} e[k] = &\left(\left[I_P \otimes y[k]^T\right] col(I_P) + \left[I_P \otimes y[k-1]^T\right]\right. \\ &... + \left[I_P \otimes y[k-na]^T\right] - \left[I_P \otimes u[k-1]^T\right] \\ &\left. ... - \left[I_P \otimes u[k-nb]^T\right]\right) \begin{pmatrix} col\left(A_1^T\right) \\ \vdots \\ col\left(A_{na}^T\right) \\ col\left(B_1^T\right) \\ \vdots \\ col\left(B_{nb}^T\right) \end{pmatrix} \end{aligned} \quad (7)$$

Or simply,

$$e[k] = \left[I_P \otimes y[k]^T\right] col(I_P) - \phi_f^T[k]\theta \quad (8)$$

Let's define the new signals

$$y_{ff}[k] = \left[I_P \otimes y[k]^T\right] col(I_P) \quad (9)$$

$$y_f[k] = \left[I_P \otimes y[k]^T\right] \quad (10)$$

$$u_f[k] = \left[I_P \otimes u[k]^T\right] \quad (11)$$

$$\hat{\theta} = \left[\Phi_f^T \Phi_f\right]^{-1} \Phi_f^T Y_f \quad (12)$$

These signals $y_{ff}[k]$, $y_f[k]$ and $u_f[k]$ belong to $R^{p\times 1}$, $R^{p\times pp}$ and $R^{p\times pm}$ respectively

$$Y_f = \begin{pmatrix} y_{ff}[1 + p\times na] \\ : \\ y_{ff}[p\times N] \end{pmatrix} \quad (13)$$

Then θ can be estimated using least squares

$$\Phi_f = \left[\Phi_{yf} \quad \Phi_{uf}\right] \quad (14)$$

Where

$$\Phi_{uf}=\begin{pmatrix}\begin{bmatrix}u_f\left[1+p\times(na-1),:\right]\\ :\\ u_f\left[p\times(N-1),:\right]\end{bmatrix} & \cdots & \begin{bmatrix}u_f\left[1+p\times(na-nb),:\right]\\ :\\ u_f\left[p\times(N-nb),:\right]\end{bmatrix}\end{pmatrix} \quad (15)$$

Φ_{uf} and Φ_{yf} are constructed as follows :

$$\Phi_{yf}=\begin{pmatrix}\begin{bmatrix}-y_f\left[1+p\times(na-1),:\right]\\ :\\ y_f\left[p\times(N-1),:\right]\end{bmatrix} & \cdots & \begin{bmatrix}-y_f\left[1,:\right]\\ :\\ -y_f\left[p\times(N-na),:\right]\end{bmatrix}\end{pmatrix} \quad (16)$$

N is the number of I/O data.

3. MOESP and ANFIS Approaches

In this paper, we have conducted a comparison between our algorithm EKLS and MOESP and ANFIS approaches, and due to the lack of ability to display all the theories related to them, cite only the sources for those wishing to learn more, for MOESP algorithm introduced by Katayama [12], ANFIS, in matlab there is a interface dedicated to identify and modelling any inputs/outputs data with ANFIS, but it is limited to give a model with single output therefore we are obliged to obtain the numbers of outputs of sub-models in our case is three [3].

4. Application

The system under examination is power plant used in the generation of electricity in unit of Sonelgaz M'sila Algeria. It consists of two main parts, first alternator which generate the electricity and trained by the the second part is the gas turbine (GE MS5001P), Figure 1 present real show of gas turbine, and the figure 2 show the schematic block diagram of Turbo-alternator with inputs/piputs position in the Turbo-alternator model, the length of the measured inputs/outputs real data used in this identification is 10^4 samples taken during 10^4 seconds.

Fig. 1. *GE MS 5001P Gas Turbine*

4.1. Validation of the Model

The model that have been select must have order n of (3,6,9...), because each l , $l=\frac{n}{p}$ matrices blocks , the blocks have a dimension of 3×3, because we have m=3 inputs and p=3 outputs . The table1 shown below clarify the change of the orders n until we stop at the best one.

The system write in linear discrete-time state space Block Observable canonical Form as follow :

Table 1. *Turbo-alternator power plant model orders n and blocks l with validations criteria values*

ELSK						MOESP				ANFIS			
l	*n*	AIC	FPE	RMSE	VAF	AIC	FPE	RMSE	VAF	AIC	FPE	RMSE	VAF
1	3	2.5483	12.7855	9.65e+27	-5.59e+51	2.5769	13.156	9.44e+28	6.89e+52				
2	6	0.0453	1.0463	1.54e+5	-1.20e+6	0.0260	1.0263	2.44e+5	2.50e+6	-2.43	0.08	186.25	98.16
3	9	0.2630	1.3008	358.1150	92.4540	0.3732	1.4524	388.1250	90.3245				
4	**12**	**0.1152**	**1.1221**	**239.6157**	**96.6654**	0.1451	1.1561	248.5187	94.9854				
5	15	0.1997	1.2212	245.3060	95.4452	0.4326	1.5412	254.4860	93.0548				

$$\begin{cases}x(k+1)=Ax(k)+Bu(k)\\ y(k)=Cx(k)+Du(k)\end{cases} \quad (16)$$

Where

$$A=\begin{pmatrix}O & O & O & -A_4\\ I & O & O & -A_3\\ O & I & O & -A_2\\ O & O & I & -A_1\end{pmatrix},\ B=\begin{pmatrix}B_4\\ B_3\\ B_2\\ B_1\end{pmatrix},\ C^T=\begin{pmatrix}O\\ O\\ O\\ I\end{pmatrix}$$

$$D=\begin{pmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}, I=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}, O=\begin{pmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}$$

From the table 1, the best model having both minimums AICs/FPEs and RMSE and maximum

VAF with no Pole/Zero cancellation is the one having order n=12 with l=4 blocks.

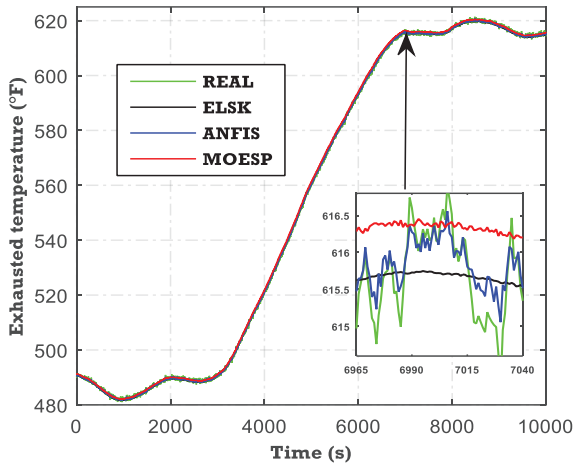


Fig. 2. *Exhausted Temperature signal*

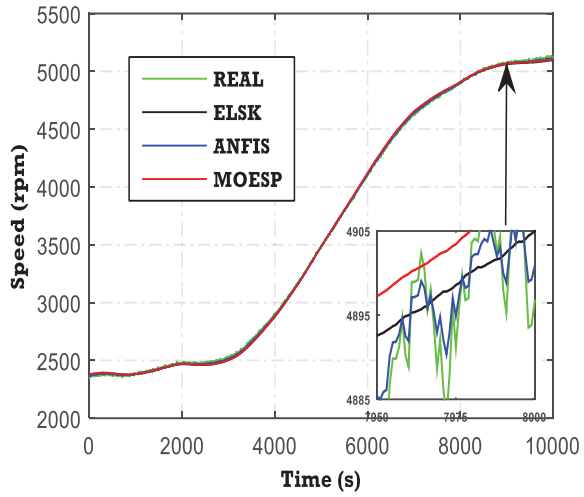


Fig. 3. *Rotor Speed signal*

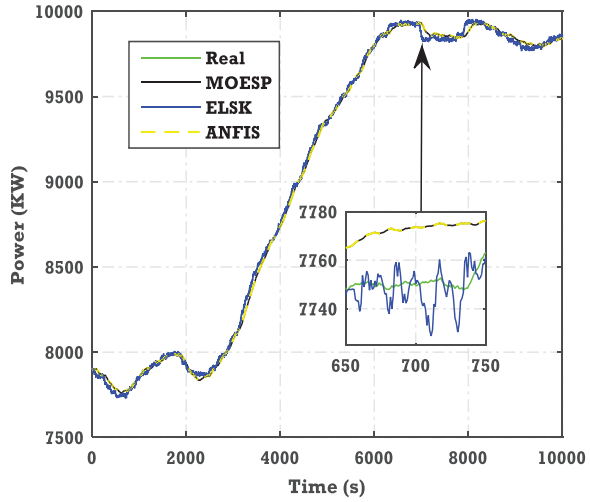


Fig. 4. *Power Generation signal*

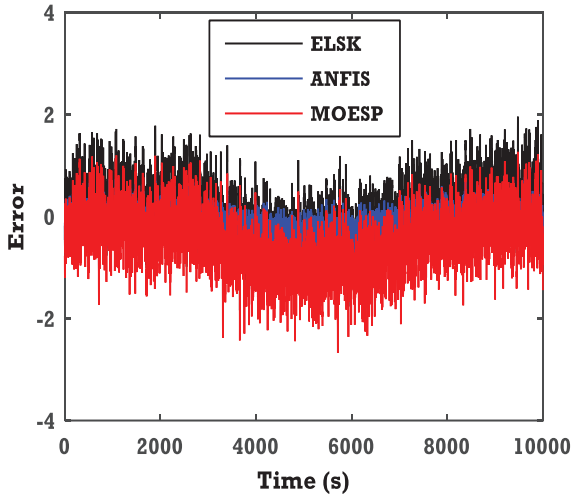


Fig. 5. *Error validation of Exhausted Temperature* signal

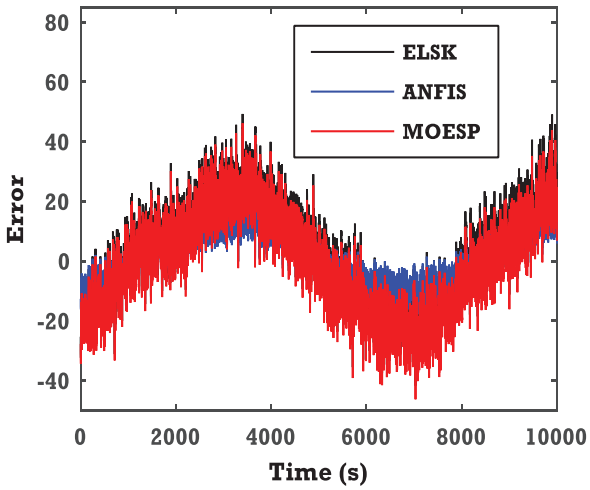


Fig. 6. *Error validation of Rotor Speed signal*

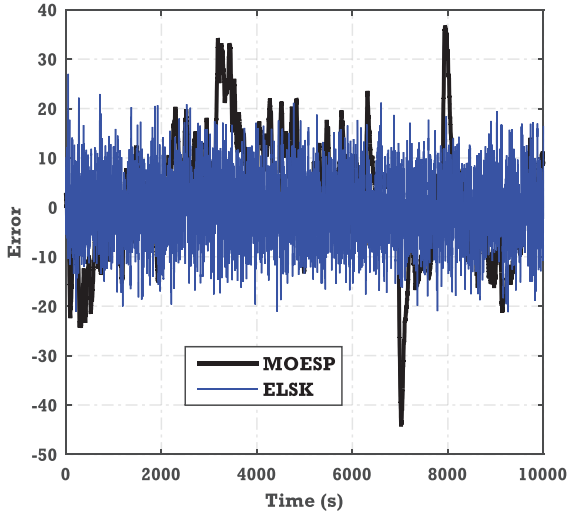


Fig. 7. *Error validation of Power Generation*

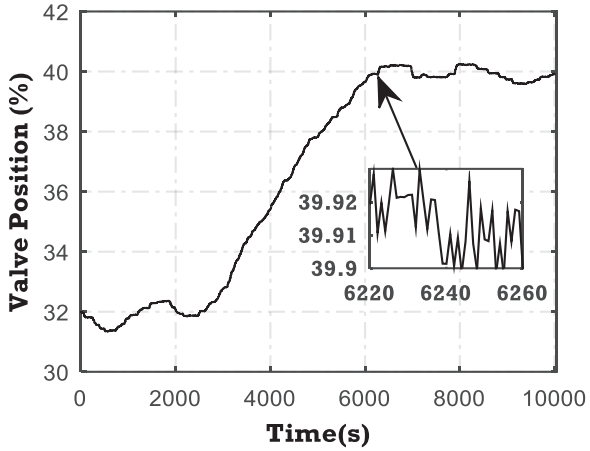


Fig. 8. *Gas Control Valve signal*

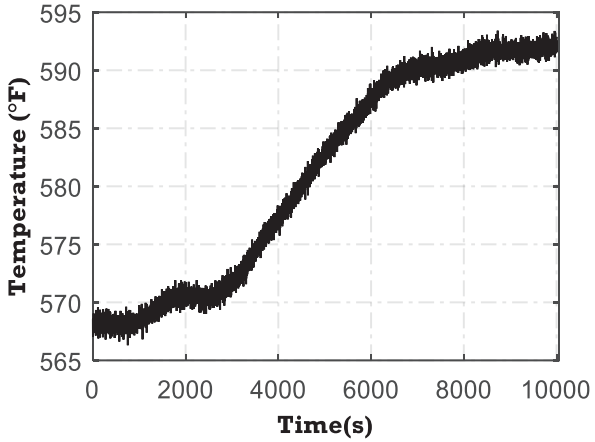


Fig. 9. *Axial Compressor Discharge Temperature signal*

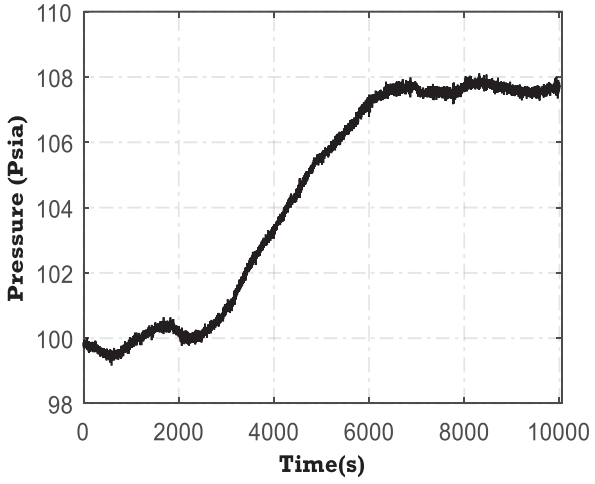


Fig. 10. *Axial Compressor Discharge Pressure*

4.1 Discussion of the obtained results

From the obtained results shown in the previous figures and due to the values numbers of the validations criteria in Table 1, can we say the obtained model is written in the block observable canonical form of index l=4 blocks, on the other hand by compared ELSK model with MOESP and ANFIS models, it is clear that obtained model is better than the MOESP model, and more suitable with ANFIS model, and through the figures of the outputs signals, the figures 2-4, the real signals of the exhausted temperature, rotor speed and power generation are identical with the estimated signals in all the time (0:1000) seconds and no statistical error and peak , the figures 5-7 demonstrated the convergence of the this identification (unbiased estimator), where the error converge asymptotically, the outputs signals, the figures 8-10, gas control valve, axial compressor discharge pressure, and axial compressor discharge temperature are riches with frequency which covered all the behavior of the Turbo-alternator.

Conclusion

To the best of our knowledge, this is a new contribution combined the (LMFD) theory in identification of class of MIMO OE model with the help of EKLS estimator and the first application on the energetic domain Turbo-alternator investigation, using inputs/outputs real data, based on the richest of the acquired informations with frequency about the black box (Turbo-alternator) plant, and the efficient of the kronicker product in the extended least square estimator and the flexibility of the LMFD theory all of this advantages allowed us to obtain an accurate identification and a valid and reliable model with no factors common (coprime) , this is according to the results of the validations criteria (AIC, FPE, RMSE and VAF) and the comparative study with recent MOESP algorithm which generate subspace model and with artificial intelligence model ANFIS, and recognized that they give high-fidelity models, and can say that EKLS is unbiased estimator, despite this is prejudge, due to this identification is off-line, but in our case in this work for more precision. As further works, can make this EKLS estimator in on-line system identification (recursive loop) and done a comparative study, also can applied many advanced control theory on this obtained dynamical model.

REFERENCES

1. Smith, S.E. (2004). *Name of book in italics*, page or chapter numbers if relevant. Publisher, Place of publication.
2. Brown, F., Harris, M.G., and Other, A.N. (1998). Name of paper. In Name(s) of editor(s) (ed.), *Name of book in italics*, page numbers. Publisher, Place of publication.
3. Smith, S.E. and Jones, L.Q. (2008). Name of paper. *Name of journal in italics*, volume (number), page numbers.
4. Abdelhafid Benyounes, Ahmed Hafaifa & Mouloud Guemana, Gas turbine modelling

based on fuzzy clustering algorithm using experimental data. *Journal of Applied Artificial Intelligence, Taylor & Francis*, 2016, vol. no., pp.

5. Ahmed Hafaifa, Guemana Mouloud, and Belhadef Rachid, Fuzzy Modeling and Control of Centrifugal Compressor Used in Gas Pipelines Systems. Multiphysics Modelling and Simulation for Systems Design and Monitoring, *Book Chapter, Applied Condition Monitoring*, 2015, vol. 2, pp.379-389.
6. Abdelhafid Benyounes, Ahmed Hafaifa and Mouloud Guemana, Gas turbine parameters modelling based on fuzzy logic and artificial neural networks: Solar TITAN 130N investigation. *5eme Symposium de lAssociation Algrienne de lIndustrie du Gaz AIG Sonatrach*, 16-17 Fvrier 2016, Alger.
7. Bachir Nail, Abdellah Kouzou, Ahmed Hafaifa, and Choayb Djeddi, Parametric Identification of Multivariable Industrial System Using Left Matrix Fraction Description. *J. Automation & Systems Engineering* , 10-4 (2016): 221-230.
8. Akroum, M., Hariche K., An optimal instrumental variable identification method for LMFD models, *Studies in informatics and control*, Vol 17, No.4, December 2008.
9. Akroum, M., and Hariche K., Extending the SRIV algorithm to LMFD models, *Journal of electrical engineering and technology*, Korean Instituteof electrical engineers, Vol. 4, No.1, march 2009.
10. Akroum, M., and Hariche K., A matrix fraction description-based identification algorithm for MIMO ARMAX models, submitted to *Asian journal of control*, 2008
11. R.A. de Callafon, D. de Roover, P.M.J. Van den Hof Multivariable least squares frequency domain identification using polynomial matrix fraction descriptions. Decision and Control, 1996., *Proceedings of the 36th IEEE Conference on Decision and Control*,13-13 Dec. 1996.
12. Nehorai, A., and Morf, M., Recursive identification algorithms for Right Matrix Fraction Description models . *IEEE Transactions on Automatic Control*, 29, 1103-1106, 1984.
13. R. Johansson, Multivariable system identificationvia continued-fraction approximation, *IEEE Transactions on Automatic Control*, 40(3), 507-512, 1995.
14. Ljung, L.,*System* Identification : Theory for the user (Prentice Hall),1999.
15. Katayama, T., Subspace methods for System Identification, *Springer-Verlag*, 2005.
16. Bachir Nail, Abdellah Kouzou, Ahmed Hafaifa, Belkacem Bekhiti, Parametric output feedback stabilization in MIMO systems:Application to gas turbine power plant, 8th International Conference on Modelling, Identification and Control *(ICMIC),* 971-976, November 2016.
17. L. S. Shieh and Y. T. Tsay, Transformation of a class of multivariable control systems to Block companion forms, *IEEE Truns. Autonwt. Control* 27: 199-203 (1982).
18. L. S. Shieh and Y. T. Tsay, Transformation of solvent and spectral factors of matrix polynomial, and their applications,Internat. *J. Control* 34:813-823 (1981).
19. I. Gohberg, P. Lancaster, L. Rodman, *Matrix Polynomials*,Academic Press, 1982.

Bachir NAIL received his license and master degrees in Electrical engineering, Automatic control from Ziane Achour University Djelfa, Algeria in 2013 and 2015 respectively. He is currently a Ph.D. student at the faculty of science and technology Djelfa, Algeria. His current research interests include Fault Tolerance control and Detection, MIMO linear and nonlinear control system, estimation theory and system identification.

[1]The Applied Automation and Industrial Diagnostics Laboratory LAADI, University of DJELFA, Algeria.
[2]Aeronautical Aerospace Automotive Railway Engineering school, ESTACA Paris, France.
b.nail@univ-djelfa.dz
kouzouabdellah@ieee.org
hafaifa@hotmail.com
Ahmed.CHAIBET@estaca.fr

© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 23, 2017

BILATERAL FILTER INTEGRATION INTO THE OPEN SOURCE SOFTWARE GELJ

ANGEL DANEV, ATANASKA BOSAKOVA-ARDENSKA, ILIYAN DOBREV, TODOR STAMENOV

Abstract: *This paper presents the application of the computer program GelJ with additional bilateral filter implementation in the context of DNA analyses of lactic acid bacteria recombinants. The results show that GelJ could be applied for lactic acid bacteria recombinants study.*

Key words: *open source software, GelJ, DNA analyses, bilateral filter, lactic acid bacteria recombinants.*

1. Introduction

The open source software began in the late 1970s and early 1980s when the shearing of source code from a computer program and the proprietary software began to really come into conflict [1].

The term "open source software" means computer software that is made available for people to study, modify and distribute the source code to anyone and for any purpose. The copyright holder provides these rights and there is no problems for software developers to inspect, change and enhance the source code. The difference between open source software and other types of software is that the authors of open source software make its source code available to others who would like to view that code, copy it, learn from it or share it. On the other hand some software has source code that only the person or organization who created it can modify or share. Only the original authors can legally copy, inspect and alter that software.

There are different applications of open source software. In this paper is presented a computer open source program that is used for specific analyses of *Lactobacillus* recombinants in gel electrophoresis.

GelJ version *2.0* is developed by Jónathan Heras (joheras@gmail.com), César Domínguez, Eloy Mata, César Larrea, and Vico Pascual at the Department of Mathematics and Computer Science of University of La Rioja (Spain) and Mysic [2]. The program is an open source software for analyzing DNA fingerprint gel images. DNA fingerprinting is a technique for comparing DNA patterns that has applications in a wide variety of contexts [3]. There are several freely-available and commercial tools that can be used for analyzing DNA fingerprint images, but the commercial software is expensive and difficult to use in some cases. On the other hand free tools support the basic functionality for DNA analysis. So, the main advantage of GelJ is that it is free and open source software. This program is user-friendly, platform-independent and feather-weight. GelJ provides a lot of outstanding features such as digital image processing including different types of filters, mechanisms for accurate lane- and band-detection, a number of band- and curve-based similarity methods, different techniques for generating dendrograms, comparison of banding patterns from different experiments, and database support.

2. Bilateral filter

The process of filtering in terms of digital image processing is fundamental operation in the science of computer vision and image processing. The term "filtering" means that the value of the filtered image at the given location (pixel of the image) is a function of the values of the input image in a small neighborhood of the same pixel or entire image.

Bilateral filter is firstly presented by Tomasi and Manduchi in 1998 [4]. The concept of the bilateral filter has been also presented in [5] as the SUSAN filter and in [6] as the neighborhood filter.

In brief, bilateral filtering smooths images while preserving edges, by means of a non-linear combination of nearby image values. Each pixel is replaced by an average of its neighbors.

This type of filter is an effective image denoising technique and it can be used for blurring an image while respective strong edges too [4, 7]. The bilateral filter is also defined as a weighted average of nearby pixels in a manner very similar to Gaussian convolution. The difference is that the

 ISSN 1310 - 8271

bilateral filter takes into consideration the difference in value with the neighbors to preserve edges while smoothing. The bilateral filter is defined by the following formula:

$$\boldsymbol{BF[I]_p = \frac{1}{W_p}\sum_{q\epsilon S} G_{\sigma_s}(\|p-q\|)\, G_{\sigma_r}(|I_p - I_q|) I_q} \quad (1)$$

where $\boldsymbol{I_p}$ is notation for the gray-scale image value at pixel position $\boldsymbol{p}$; the weight of pixel $\boldsymbol{q}$ is defined by the Gaussian $G_\sigma(\|p-q\|)$ where $\boldsymbol{\sigma}$ is a parameter defining the neighborhood size; $\boldsymbol{s}$ is the set of all possible image locations; $\boldsymbol{r}$ is the set of all possible pixel values. Normalization factor W_p ensures pixel weights sum to 1.0 [7]:

$$\boldsymbol{W_p = \sum_{q\epsilon S} G_{\sigma_s}(\|p-q\|)\, G_{\sigma_r}(|I_p - I_q|)} \quad (2)$$

where the parameters σ_s and σ_r specify the level of filtering for the image $\boldsymbol{I}$ [7].

Median and bilateral filter are applied to an image in figure 1 (A) to yield the image in figure 1 (C).

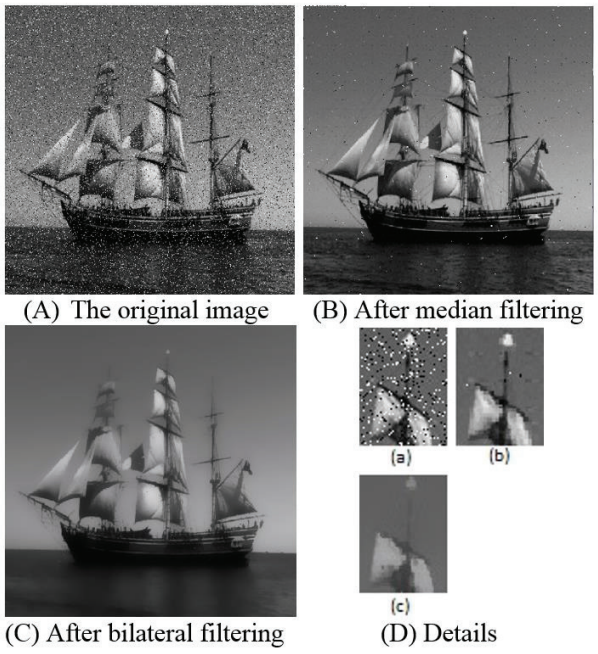


***Fig. 1**. A picture before (A) and after (B) median and (C) bilateral filtering.*

The original image has salt-and-pepper noise. It can be seen that the most of the fine texture has been filtered away, and yet all contours are as crisp as in the original image. Figure 1 (d) shows a detail of figures 1 (a), 1 (b), and 1 (c).

The bilateral filter can be applied to either grayscale images and to colored images.

There are variety of applications that the bilateral filter can be used. According to the specific needs this type of filter is used for image denoising, texture and illumination separation, data fusion, 3D fairing and e.t.c. Denoising of an image is the primary goal of the bilateral filter and it is used in applications for medical image processing, tracking, movie restoration and so on [7].

3. GelJ structure and bilateral filter integration

GelJ is a software product for analyzing DNA fingerprint gel images. It is developed using the objected oriented language Java. GelJ is platform independent program because Java is platform independent language. The product provides different features for image processing using the ImageJ library [8]. On the other hand a library called Weka [9] provides variety of machine learning algorithms for data mining tasks [9, 10]. GelJ contains a special embedded database that is integrated using a library called JavaDB. GelJ has three common concepts – experiment, comparison and research. The process of making an experiment refers to analyzing the DNA fingerprints gel image provided from the biological experiment. Than the process of comparison evaluates the similarities between the samples from one or more experiments. At the end researching process collect all data from the experiments and comparisons. The main window of the GelJ program is shown in figure 2.

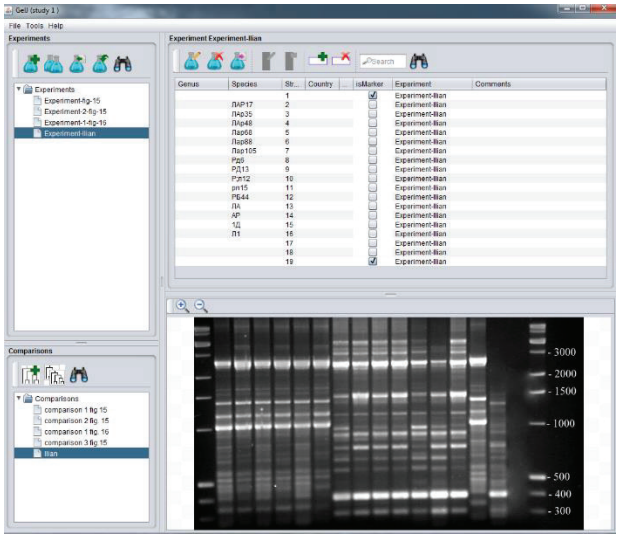


***Fig. 2**. GelJ main window*

The process of making experiments consists of several steps that the user have to go through – preprocessing of the DNA fingerprint gel image, lane detection, normalization and bands detection.

Figure 3 illustrates all steps needed for gel image processing in details. On the first step the user can crop the image to cut out a specific important region. On the other hand the user can adjust brightness and contrast of the image; invert the colors and so on. The second step is lanes detection. The user have two options- automatically or manual

detection. He can choose to detect the lanes automatically but after that manually editing is required. It is important to choose and mark the reference lanes. Bands detection is the third step of every analysis. First the user can automatically detect all of the bands in the entire image. For greater precision manually picking and editing the bands is required. The program provides features for adding and removing bands. The step of normalization is important for creating accurate analysis. There is a feature for adding or loading reference marker.

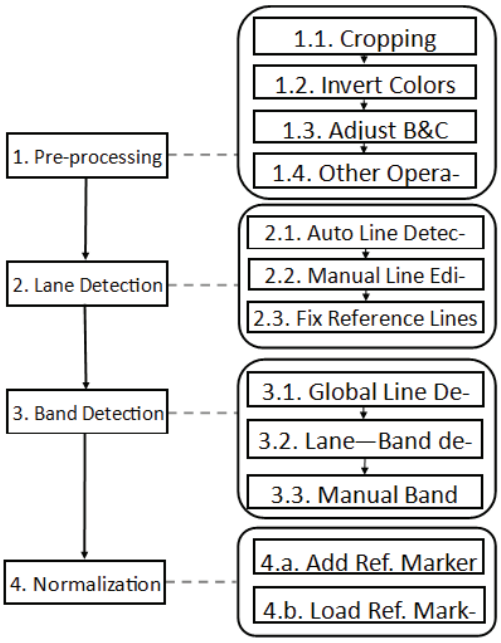


***Fig. 3**. Steps needed for agarose gel image processing*

Figure 4 illustrates the user interface for selecting the bilateral filter and all of the necessary parameters.

Below are listed all of the available parameters that the user can change in order to use the bilateral filter.

Sigma Range – as the Sigma Range parameter increases the bilateral filter becomes closer to Gaussian blur.

Sigma X, Y, Z – As the spatial parameter of bilateral filter (σ_d) increases, the filter smooth larger features. In this parameter x – is the amount of pixel neighborhood in x – axis of the coordinate system; y – is the amount of pixel neighborhood in y – axis of the coordinate system; z – It is a coordinate value for smoothing 3D meshes while preserving their most prominent features [7].

Method Channels – The possible values are: 0 – Best channels; 1 – All channels; 2 – By Truncation; 3 – By Tolerance;

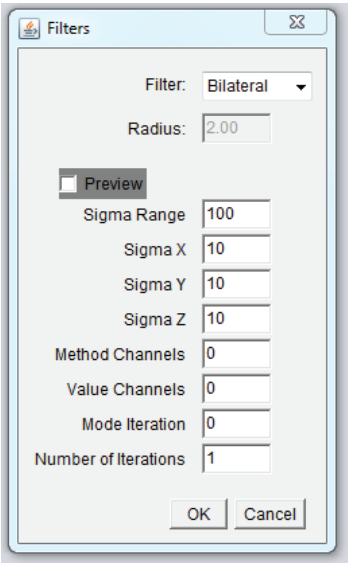


***Fig. 4**. Dropdown menu for filter type selection*

Value Channels – This parameter specify the number of channels. Every image has color channels for example some colored image has 3 channels for Red, Green and Blue (RGB) components.

Mode Iteration – The bilateral filter can be iterated. The value of this field specify the iteration schema: 0 –None; 1 –Range Iteration; 2 – Space Iteration; 3 –Overall Iteration;

Number of iterations – in this field the number of iterations can be specified. Larger number of iterations leads to achieve cartoon-like renditions of images.

For integration of bilateral filter in GelJ as an additional feature *BilateralFilter_.jar* file is downloaded from:

bigwww.epfl.ch/algorithms/bilateral-filter/ [11, 12].

After that the file is included in a directory called *lib* where the external libraries are situated. In package *ij.plugin* a class named *Bilateral_Filter_Instant.java* was created. All of the necessary variables are initialized in it. They are passed to the method named *process* where the source code for implementation of bilateral filter is placed. All of the classes used in the bilateral filter library are described below.

Allocation – memory allocation for storage all intermediate result from bilateral filter.

Data – contains the input data needed for image processing.

Filter – contains the logic for image processing.

Gaussian – contains the program logic needed for the Gaussian case.

Initialization – a class responsible for initialization process of tables for *sin* and *cos*.

All of the methods that are part of the *Filter* class are listed below.

approximate() – returns three indicators on the quality of the estimate of the Gaussian function, discretized on bin items;

computeOrder() – establishing the execution order of the filter;

execute() – processing the image

selectChanels() – determines the choice of channels;

setIterativeScheme ()– determines the iterative scheme;

setLog() – specifies whether to generate a log file after treatment;

setMultithread() – determines whether the image will be processed multithreaded;

The Gaussian class implements the interface *java.lang.Runnable*. It is used following methods: *isRunning()* - to check whether it is running image processing; *run()* – for running.

4. Analyses of lactic acid bacteria recombinants

The method of agarose gel electrophoresis allows the creation of images containing valuable information for DNA analysis. Today's modern computer technologies provides variety of open source software products for creating DNA analyses easier in comparison to manual analyses.

In this research, one modern computer program for DNA analyses was used to analyze lactic acid bacteria recombinants. Randomly Amplified Polymorphic DNA (RAPD) analysis was conducted. It was used an image obtained from biological experiment. Two main primers are selected – OPP-7 and j-8 [13, 14, 15] that can work with the genomes of the couples parental strains - L. rhamnosus 1D and L. acidophilus 2; L. rhamnosus 1D and L. acidophilus Ar; L. acidophilus 2 and B. bifidum L1. The image is shown in figure 6.

After a few steps of the image pre-processing the user can continue his DNA analysis with either automatically or manually lane detection. The process of automatically lane detection often requires manually editing the height, width and position of the lanes. As a next step in the DNA analysis using GelJ takes place the procedure of bands detection. Again it can be done either manually or automatically. In the most cases the user have to do this manually for greater accuracy.

All of the bands automatically detected and after that manually processed using the software product GelJ are shown in figure 6. A band marker should be placed on every important bright peak. The groups of red colored markers determines which are the reference lanes.

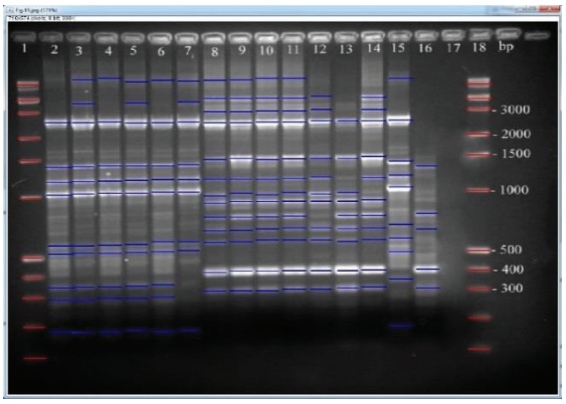


Fig. 6. *Automatically detected bands and manually processed*

As an important step in procedure of DNA analyses of lactic acid bacteria recombinants is creating a special dendrogram which is designed to illustrate the grade of similarity between couples parental strains and the recombinant clones. Such a dendrogram is shown below in figure 7. It is generated by the GelJ software using the Similarity method called Dice [16]; UPGMA linkage and the output is selected to be only dendrogram. The Dice method is also known as Dice's coefficient. It is used for comparing the similarity of two samples. The output can be dendrogram, dendrogram + images, dendrogram + images + bands. The window for creating a dendrogram is shown in figure 8.

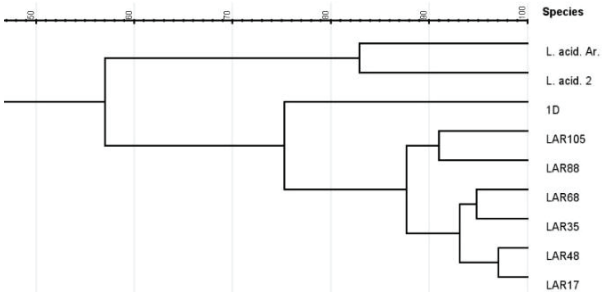


Fig. 7. *A dendrogram illustrating the grade of similarity between couples parental strains and the recombinant clones*

Creating a dendrogram follows the following iterative process: on each step the nearest two groups are combining one another in a group of higher level. Every single generated dendrogram can additionally contains information about the lanes, the original image or the image with detected bands.

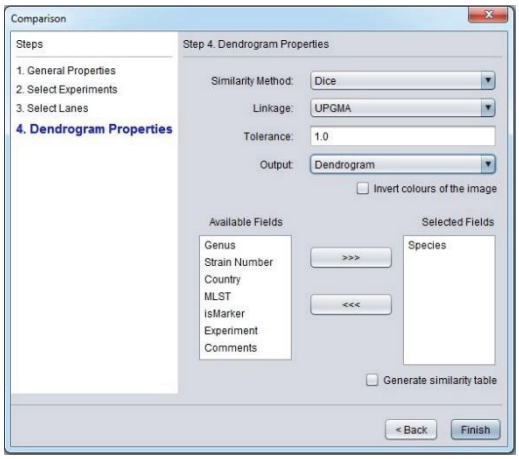


Fig. 8. *Comparison window for generating a dendrogram*

5. Conclusion

The computer program GelJ as an open source software for creating DNA analyses could be applied for lactic acid bacteria recombinants study. Using GelJ by lactic acid bacteria recombinants analysis sufficiently accurate results were received in comparison with commercial software products [13, 14, 15]. GelJ provides a lot of features thanks to which the users can easily create their analyses. In addition the open source conception of the product allows computer programmers to develop better tools and features.

REFERENCE

1. Wong, K., Sayo, P. (2004). Free/Open Source Software. A General Introduction. *United Nations Development Programme's Asia-Pacific Development Information Programme (UNDP-APDIP) Kuala Lumpur, Malaysia*, ISBN: 983-3094-00-7
2. Heras, J., Domínguez, C., Mata, E., Larrea, C., and Pascual, V. (2016). GelJ manual v.2.0, *Department of Mathematics and Computer Science of University of La Rioja (Spain) and Mysic*.
3. Heras, J., Domínguez C., Mata, E., Pascual, V., Lozano, C., Torres, C. (2015). GelJ – a tool for analyzing DNA fingerprint gel images, *BMC Bioinformatics*
4. Zhang, M. (2006), Bilateral filter in image processing, *B.S. Department of Electrical and Computer Engineering*, Beijing University of Posts and Telecommunications.
5. Smith, S. M. and Brady, J. M. (1997), Susan - a new approach to low level image processing, *Int. Journal of Computer Vision*, volume (23), pp.45–78.
6. Yaroslavsky, L. (1985). Digital Picture Processing - An Introduction, *Springer Verlag*.
7. Paris, S., Kornprobst, P., Tumblin, J. and Durand, F. (2008). Bilateral Filtering: Theory and Applications, Foundations and Trends in Computer Graphics and Vision, Vol. 4, No. 1, pp.1–73.
8. Schneider, CA, Rasband, WS, Eliceiri, KW (2012). NIH Image to ImageJ. Nat Methods; 9:671–5.
9. Markov, Z., Russell, I. (2006). An Introduction to the WEKA Data Mining System, pp. 367–368
10. Amara,l K., Netbeans IDE Tutorial for using the Weka API, University of Massachusetts Boston
11. Sage, D., Chaudhury, K.N. (2011). Constant-time bilateral filter, Biomedical Imaging Group (BIG) Ecole Polytechnique Federale de Lausanne (EPFL) Lausanne, Switzerland.
12. Chaudhury, K.N., Sage, D., Unser, M. (2010). Fast O(1) bilateral filtering using trigonometric range kernels, IEEE Transactions on Image Processing, submitted Dec.
13. Dobrev, I.N. (2011). Application of lactobacilli and bifidobacteria in the manufacture of healthy foods from goat milk, PhD thesis, University of food technology, Plovdiv.
14. Denkova, R., Dobrev, I., Denkova, Z., Yanakieva, V., Urshev, Z., Yordanova, M., Ilieva, S. (2012). Genetic, biochemical and physiological characteristics of the recombinant strain Lactobacillus RL15, obtained by intergeneric hybridization between Lactobacillus acidophilus 2 and Bifidobacterium bifidum L1. Food and Environmental Safety – Journal of Faculty of Food Engineering, Stefan cel Mare University – Suceava, Volume XI, Issue I – 2012, pp.5-14.
15. Goranov, B., Denkova, R., Teneva, D., Denkova, Z., Popova, P. (2015). Molecular-genetic identification of Lactobacillus

strains, isolated from homemade yoghurt. Ukrainian Food Journal 4 (1): pp.67-76

16. Dice, Lee R. (1945), Measures of the Amount of Ecologic Association Between Species, University of Michigon, Ecology, 26 (3): pp.297–302

Контакти:

Ангел Милчев Данев
УХТ – гр. Пловдив
бул. „Марица“ 26
тел. 032 603 860
angel_danev_bg@abv.bg

Атанаска Босакова-Арденска
УХТ – гр. Пловдив
бул. „Марица“ 26
тел. 032 603 860
a_bosakova@uft-plovdiv.bg

Илиян Добрев
УХТ – гр. Пловдив
бул. „Марица“ 26
тел. 032 603 696
indobrev@abv.bg

Тодор Стаменов
УХТ – гр. Пловдив
бул. „Марица“ 26
тел. 032 603 860
todor.stamenov93@gmail.com

© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 23, 2017

AUTOMATED SYSTEM FOR CONTROL AND MANAGEMENT OF A MACHINE FOR AUTOMATIC ASSEMBLY OF CURTAIN BRACKETS AND HOOKS

PENKO MITEV, GEORGI SHTEREV

Abstract: *The automated system for control and management is intended to assemble curtain brackets and hooks by a predefined movement sequence, registering each of the details "roll" or "bracket" with their subsequent assembly. In the work is pointed out the way of choosing of a technological scheme for control and management of the executive devices for a real automated system built and integrated in the production. The invention is implemented by "KMS Engineering" LTD - Plovdiv in "UYUT" LTD - Sankt Petersburg.*

Key Words: *fiber optic converters, logic controller, pneumatic distributors, ejector nozzle*.

1. Introduction

The target of automation is an automatic assembly machine which joins together two plastic parts - "hook" and roller". This features the following processes:

- Initial check of all default positions of actuators
- Presence of all start conditions related to the automatic mode;
- Start of automatic mode;
- Stop of automatic mode;
- Control of all mechanisms in manual mode;
- Control of each of the two vibratory bowl feeders depending on the signals from the sensors for minimal and maximal parts quantity;
- Control of the signal lamp tower with red and green lights;
- Control of the sequence of actions, which take part in the assembly process;
- Tracking of statistical parameters - "productivity", "number of errors for a period of time", etc.

The main tasks are related to the design and assembling of the described system, based on standard components, writing of control algorithm for the technological process and creation of a simulative laboratory unit demonstrating the various process activities.[1, 2, 3, 4]. The machine control system if based on a programmable logic controller which accepts signals from fiber optics sensors for detection of parts presence in assembly position and controls four pneumatic valves through relay outputs.

2. Networks of possible variants

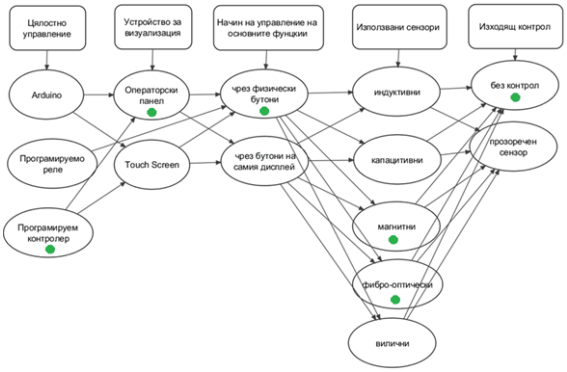


Fig. 1. Network of the possible variants for project realization

Fig. 1 shows a network of the possible variants for project realization based on the main functions which are required from the control system. An optimal solution is to be found among these variants. Some devices are

 ISSN 1310 - 8271

able to do more than one operation or two fulfill two functions simultaneously. The process of decision taking is connected to a complex analysis of many factors and for each solution there are various advantages and disadvantages. Green dots mark the taken decisions for this specific task.

3. Cyclogram of the machine process

The assembly process is cycle-based where each cycle takes 1,5 seconds. Fig.2 demonstrates the time for the working stroke of every of the four pneumatic cylinders(lifting, assembly, opposing cylinders as well as a blowout nozzle). In the event of anomaly in the working cycle or in case of other emergency situation, the control system activates light and sound signals to attract attention of the site personnel.

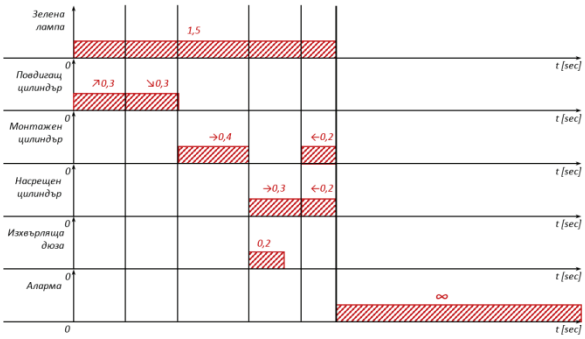


Fig. 2. Cyclogram of the machine process

4. Working algorithm

It is based on several modules - preparatory, subroutines calling module, module for delaying fiber-optic sensors' signals, control module for vibratory bowl feeders and alarms handling module. All modules are written in LADDER language and each module does not surpass 30 program steps.

5. Experimental analysis

5.1. Experimental analysis on the productivity of the vibratory bowl feeder for hooks

The vibratory bowl feeder for hooks is a critical part of the whole system. Depending on the number of arriving and properly oriented parts, the machine will either have minimal presence of parts for normal work cycles, or will need to wait until this quantity is attained. The following statistical information was acquired based on 10 trials.
(table 1):

Table 1

TESTS FOR HOOKS FEEDER PRODUCTIVITY

Parts quantity: 76

Duration for the above quantity:

Minute	Seconds	Total: (seconds)	Parts / sec	Parts / min	Parts / hour
1	30	90	0,84	50,67	3040,00
1	40	100	0,76	45,60	2736,00
1	25	85	0,89	53,65	3218,82
1	15	75	1,01	60,80	3648,00
1	20	80	0,95	57,00	3420,00
1	32	92	0,83	49,57	2973,91
1	45	105	0,72	43,43	2605,71
1	15	75	1,01	60,80	3648,00
1	20	80	0,95	57,00	3420,00
	AVERAGE:	87	0,89	53	**3190**

Based on the results, the average feeder productivity is 3190 parts / hour. However, the assembly mechanism is able to work faster.

In case the bowl feeder could not feed enough parts, its own feeder rate is the maximum possible productivity value for the machine.

5.2. Experimental analysis on the productivity of the vibratory bowl feeder for rollers

- after 10 trials the results show that the rollers feeder has higher feed rate than the hooks feeder. This is easily explained with the smaller roller dimensions and the easier orientation in general.

The results are shown in Table 2:

Table 2

TESTS FOR ROLLERS FEEDER PRODUCTIVITY

Parts quantity: 40

Duration for the above quantity:

Minute	Seconds	Total: (seconds)	Parts / sec	Parts / min	Parts / hour
0	38	38	1,05	63,16	3789,47
0	43	43	0,93	55,81	3348,84
0	45	45	0,89	53,33	3200,00
0	45	45	0,89	53,33	3200,00
0	40	40	1,00	60,00	3600,00
0	42	42	0,95	57,14	3428,57
0	43	43	0,93	55,81	3348,84
0	45	45	0,89	53,33	3200,00
0	45	45	0,89	53,33	3200,00
	AVERAGE:	43	0,94	56	**3368**

5.3. Experimental analysis on the PLC scan cycle

- a "scan cycle" by PLCs is the time needed to execute the following sequence of operations:

- check of states on every input - logic "0" or logic "1"
- program execution
- update of output signals based on the changes in the current PLC cycle.

The PLCs from the XINJE XC3 series have a scan cycle from 1 ms to 10 ms, depending on the program complexity and the number of additional expansion modules, connected to the base PLC. During the phase of PLC program creation, the following statistical information was acquired:
(table 3):

Table 3

Number of steps inside the PLC program	**PLC scan cycle (ms)**
100	0
200	1
300	1
400	1
500	2
600	2
700	2
800	3
900	4

The following additional research on the scan cycle has been conducted independently from the system with a resulting scan cycle of 20 ms:

- Base PLC XC3-60PRT-C with 36 inputs and 24 outputs
- 7 additional expansion modules (a total of 140 inputs / 100 outputs)
- 5000 program steps inside the program

6. Conslusion:

1. The experience gathered from the project for automatic assembly of hooks and rollers could be applied in many other projects having similar parts. During the development of a specific manufacturing process, a detailed research activity related to industrial sensors used in industrial environment, automation components and innovations in IT and communication technologies is to be done. Additionally, analysis on the application of programmable logic controllers for industrial automation projects and other areas, based on the end user's inquiry are to be researched;
2. The design of the electrical and pneumatic diagrams is connected with the necessary number of input/output signals and parameters of the assembly process;
3. Economical parameters are to be taken into account when making decisions on the choice of automation components.

7. REFERENCES

1. Гановски В., Нешков Т., Бояджиев И., Ликов Ц., Механизация и автоматизация на монтажнитепроцеси в машиностроението“, изд. „Техника“, София, 1986 год.
2. KrauseWerner, Geratekonstrution, VEB VerlagTechnik, Berlin, 1986.
3. Лебедовский М.С., А.И. Федотов, Автоматизация сборочныхработ, Л., Лениздат, 1970.
4. Замятин В.К., Технология и автоматиза-циясборки, Москва, Машиностроение, 1993.

8. About the authors

Penko Valkov Mitev: engineer, Msc, "Machine building" (2017), TU-Sofia, branch Plovdiv;

Georgi Panayotov Shterev: electrical engineer, Msc „Electrical measurement equipment” (1978), Ph.D (2007), Assoc. Prof. (2009), TU-Sofia, branch Plovdiv

© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 23, 2017

MULTI-FUNCTION AND RECONFIGURABLE ROBUST ANALOG I/O MODULE DESIGN METHODOLOGY FOR INDUSTRIAL CONTROLLER SYSTEMS

CAGRI CANDAN, EMRE CAN DIKMEN, T. CETIN AKINCI, SERHAT SEKER

Abstract: *In this paper, applied design methodology of robust analog I/O (input and output) module of industrial controller is explained. Analog I/O modules are one of the essential part of the controller. Main aspects of the design methodology are reducing microprocessor processing complexity, selecting proper ICs (Integrated Circuits), analog signal correlation, transducer reading performance improvements, ADC (analog to digital converter) and DAC (digital to analog converter) implementation techniques. Additionally, test results and serial communication protocols also mentioned.*

Key words: *Analog module design, transducer performance improvements, noise filtering, IC selection, signal converter, signal generator.*

1. Introduction

Industrial Control Systems (ICS) is a general term covering various types of control systems and associated instruments used in industrial production, process control, machinery and critical infrastructures [1]. High capabilities and multiple functions of industrial controllers have caused widely usage and therefore have become a crucial part of the industry. As systems became more and more complex, demands of more advanced controller has gained momentum [2].

This system provides more flexible and effective measurement solutions to industry this utilizes the efficiency of overall process. System has efficient processing power and connectivity capabilities on the industrial computers thus systems. In this decade the industry which has depended on electronic design has becoming more advanced because of unique designs products. Because of these the requirement of conversion systems increased in terms of flexibility, accuracy and reliability. The previous version of the conversion systems has seen that because of their design issues the data accuracy and capacity of system is quite low.

To obtain better performance over the industrial systems several methods has been researched. In early 2000's technology a central process unit has to be connected to the all of the sensors and devices via reading or transmitting voltage [3]. The experiments showed that regarding on the process capacity, time consumption might increase exponentially. In order to decrease complexity, a modular approach has been used. To solve problem extra microprocessors has been used. Implementing microprocessor increased the speed of device furthermore it implements several benefits on device as well such as well-known communication tools (RS232-485), local algorithms which can embedded on its registers or OLED screens for notify users the status of process and device [4].

In the following level several components and their purpose has been explained. Overall system requirements process levels of methods explained.

2. Design Methodology

From the analysis of literature review, most of conversion systems has not capable of fulfilling the system requirements of industrial processes. Most of the systems are not supply current technologies demand. Therefore, our design offers low cost, low power consumption, portability, high speed and high accuracy in terms of both academic perspective and industrial demands. Our design is a compact size device. It uses 32 bit microprocessor which has the speed of up to 48 MHz Beside of this 256 Kb EEPROM. 16 Mb Internal RAM. It includes 4 channel 24 Bits A/D converters for connecting several sensors. For dominating the sensors and actuators it includes 16 bit D/A converter with both

 ISSN 1310 - 8271

current and voltage options. Device has capable of supplying all industry demands such as variety of voltages (+-10V, +-5V) and for current (0-20, 4-20, 4-24 etc.).

2.1 Central Process Unit (CPU) Selection

The demand of the industrial projects has high microprocessors speed. Due to the natural environment of microprocessors this issue solved with high clock cycle, long resolution bit and smart algorithms which can able to solve big cluster of data packets and communication protocols simultaneously [5]. Almost all of the industrial processes count as nonstop processes generally the systems work on 24 hours and 7 days because of this the selected microcontroller of the systems should consume less energy from that perspective energy consumption is equal enough with processing speed. In our design we used one well known microcontrollers company's products.

To accomplish complex industrial projects a smart communication and sensor analyze algorithm required. The easiest way to solve this problem is using a fast, powerful microprocessor. In this design ARM Core based microprocessor has been used. ARM based microprocessors have supreme flexibility peripherals and cortex layers which increase our systems compatibility. The required microprocessor has to have high frequency which is up to 40 MHz or higher in order to achieve acceptable computation speed. To balance power consumption of the design, a low power consumed microprocessor required. In order to achieve process in required power level and speed the well-known company the ST is selected. Consider to GPIO requirements of the system 48 pin STM32F0xxx microprocessors series is selected.

2.2 Analog to Digital Converter (ADC) Selection and Implementation

Due to the accomplished a successful industrial project sensor communication required. Industrial devices, peripherals have different level of scale however each of which most commonly uses well known standards such as 4-20 mA or 0-10 V output. This design has capability of both 4-20 mA standards and 0-10 V standards. Design has internal jumper sets each of which for making decision of whether current or voltage uses. Receiving analog input requires a high precession and speed. To solve this, issue a four channel -24 bit delta sigma analog to digital converter has been used. The converter can able to work with 50Hz or 60 Hz analog signals. The noise resistance of integrated circuits has up to 600nB RMS [6].

2.2.1. Isolation

In this decade the design of the Analog input/output modules has been changed. All of the analog/output modules separated with the host microprocessors via some separators and controllers. There were been several methods for separating the modules from the processor. In this design we used galvanic isolation because of several advantages over the other well-known separation methods [7]. The Digital to analog converter chip has specialized for industrial projects. Selected chip especially capable of providing to 20 mA, 0 to 20mA or 0 to 24 mA current outputs or 0-5 V, 0-10 V, ±5 V or ±10 voltage outputs with a10% false supply accuracy range (0-5.5 V, 0- 11 V, ±5.5 V, or ±11 V). Because of these properties in this design our team select TI-DAC8760 model IC chip.

2.2.2. Voltage Output

The DAC chip has property of configurability. When chip is configured for voltage output, max resistivity is 1 K Ohm at 10 mA. The voltage level of output the modules output can extend up from +15V to -15V. The voltage source of the chip is supplied by power rail; therefore, the rail must be at least +- 10 V supply range in order to supply the IC chips.

2.3 Digital to Analog Converter (DAC)

As known as to accomplish an industrial project the designer of systems needs to control several sensors and peripherals. Besides of shut of/on devices any peripherals or sensors requires a wide range of input streams such as 0-10 V or 4-20 mA or 0-20mA or so. To accomplish these requirements a high precision digital to analog converter (DAC) has required. The DAC has to be work with either 0-10 V or 4-20mA or both. To accomplish this standards 16 bit dual (both V and mA output) integrated chip has been selected [8].

The DAC chip has 16 bit resolution. This range has been chosen because of the requirement of high level integration. The IC chip has max 0.1 full scale range (FSR). This range includes offset error, gain error, calibration error, scale error etc. This FSR scale range is acceptable for all of the voltage and current supply range. The maximum nonlinearity is +-1 V or +-2 mA regarding to our experiments in real-time environment. The IC chip offers a programmable interface for determining the voltage and current ranges, with using calibration register user able to program the gain error and zero error and other calibration parameters. The slew rate of output is also programmable. All of the communication between The IC chip and host processor have achieved through SPI interface. SPI is a well-known fast communication protocol which

is increased the speed of the overall design. For each DAC module one microprocessor GPIO has been assigned and with full-duplex SPI communication protocol all of devices programmed separately. The calibration parameters and design information have stored in an EEPROM. The storage unit also uses the SPI buss therefore the communication speed is sufficient for obtaining all of the information from converter peripherals [9].

. In this system there has 8 blocks which represent the most important functions. The sequence of device starts with the input signal from sensors. The input determined from user side at early system execution. The user could either control the analog output by DAC or receive information from analog input by ADC peripheral. All of the communication ways handled through serial communication line [10]. Both of the information of analog input and output peripherals values stored at micro controller. After each sequence the memory part which contains the information of devices peripheral is stored on the EEPROM for providing secure information and safety to user in case of power failure. The CPU handles all of these events with using high clock frequency which is up to 42MHz. Microcontroller gather input data form ADC and generate output data which has the same parameter with input data to DAC. Thanks to that the input data remaining scaled and secured without losing information. After these processes all of the information sends to viewing module display for showing the final results.

3. Device Working Principle

Designed to comply with IEC61000-4 standards

•Selectable Voltage Output ranges from± 10V, 0 -10V, ± 5V, 0 -5V with output

•Selectable Current Outputs ranges from: 0 - 20 mA, 4 -20 mA, 0 -24mA

•Output filtering, protection up to 15kV for Electrostatic Discharge protection

•SPI interface applied with isolation which has communication speed up to 20MHz.

•High temperature protection, short circuit protection.

4. Test Results

The chip has capability of working the view of ADC on integrated circuit on 2.7 to 5 V periods. The device has been tested interns of current and voltage level and the temperature over the voltage and/or current consumption.

All of the accuracy and speed test have made with 100 sample rate with different voltage/current levels. 7 tests have been made with this device 3 of them for voltage and current test one of them for temperature test. The first test has been made 100 ms period of time. In between two following 100 ms time period the microcontroller adjusted to send a new voltage level starting from -10 V to +10 V and the devices output level measured by voltage source. As seen in Figure 1 the integrated chip cannot reach the speed level of microprocessor continuously therefore the level of voltage cannot increase sequentially. We observed that the voltage level increased like a saw-tooth wave.

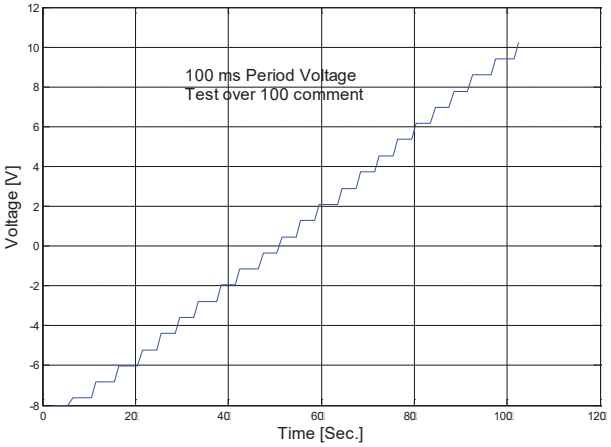


Fig. 1. *Accuracy test +-10V level 100 ms Period for 100 samples*

Second test has been made with 250 ms time interval. To measure accuracy of the device the time interval increased by 150 ms. The results get better however instead of overall linear graph we have been seen that ripples between sequential voltages. As shown in figure two the voltage become more stable than previous test however we can clearly assume that the time period is not long enough. Third test has been made with 500 time interval. To get optimum result 500 ms have been selected. The results are get optimum as expected. As seen in figure 3 the voltage and the time period gets completely linear.

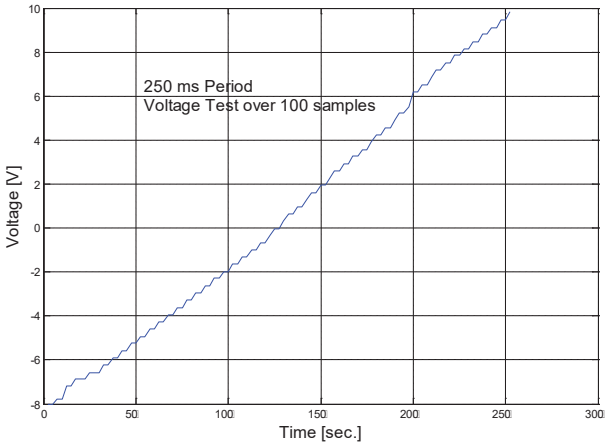


***Fig. 2**. Accuracy test +-10V level 250 ms Period for 100 samples*

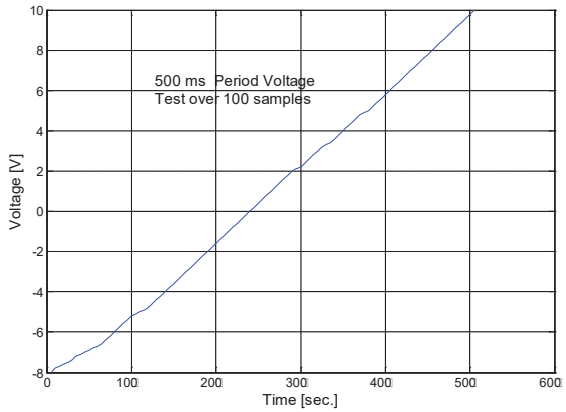


***Fig. 3.** Accuracy test +-10V level 500 ms Period for 100 samples*

After linearization of voltage test the second phase began. In second phase the level of current accuracy has been tested in the same condition with voltage tests. The first test has been made 100 ms period of time. In between two following 100 ms time period the microcontroller adjusted to send a new current level starting from 4 mA to 20mA and the devices output level measured by current source.

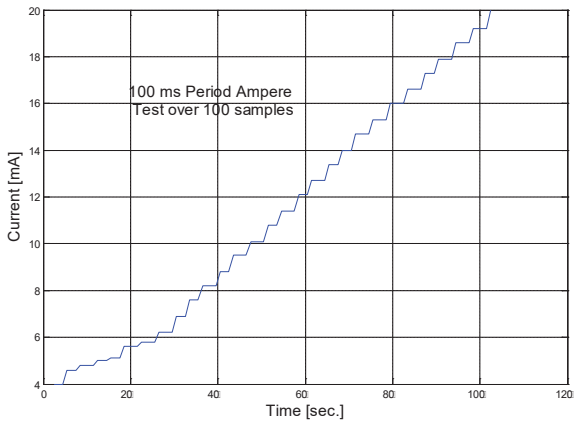


***Fig. 4.** Accuracy test 4-20 mA level 100 ms Period for 100 samples*

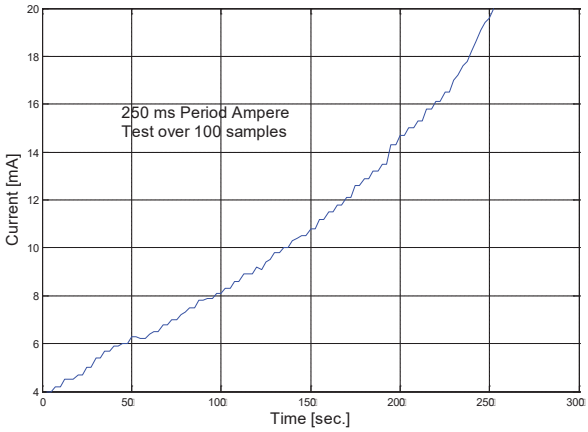


***Fig. 5.** Accuracy test 4-20 mA level 250 ms Period for 100 samples*

As seen in figure 4 we obtained a nonlinear graph in our first 30 trial we observed almost correct output however when the current level increases the chip cannot reach the level of microcontroller and some additional delays occurred. Second test has been made with 250 ms time interval. As it seen with the voltage test the result becomes sharper however as seen in figure 5 additional ripples are still disrupt the overall current signals.

Third test has been made with 500 time interval. As seen in previous test set we observed that for the optimum result the time interval must be higher than 250 in order to compere between two test sets the level adjusted to 500 ms time interval. As seen in figure 6 after adjusting the right time interval the current level become completely linear.

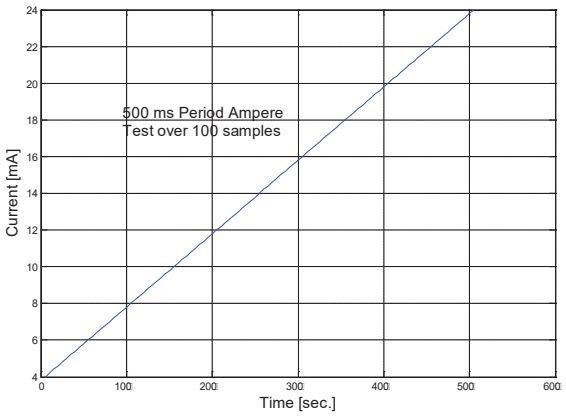


***Fig. 6.** Accuracy test 4-20 mA level 50 ms Period for 100 samples*

After finishing the accuracy test a temperature test has been started. The device applied with various level of current and voltage output and the temperature of chip measured externally. As seen in the figure 7 the level of temperature increase regarding to level of output up to a point. We observed that the device becomes stable after level of the current or voltage has been

produced by the chips. The stabilization may cost because of the space between the chips on PCB as shown figure 8. The vain spaces could absorb the temperature of device which can reach over a certain range

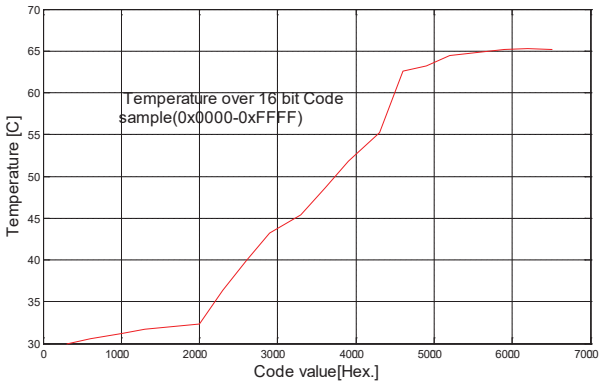


Fig. 7. *Temperature test 20 samples with the range of 16 bit data from 0 to 65535*

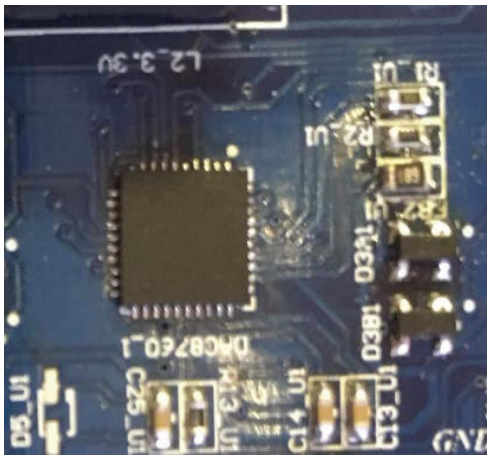


Fig. 8. *the DAC on PCB frontal seen*

5. Conclusion

In this paper an industrial sensor peripheral reader controller has been explained. The device currently in is working on with several different applications in throughout several different places and facilities. The combination between DAC&ADC enhanced the device usage thus several industrial processes has speeded up and get some extra qualifications such as combinations of analog input outputs, smart control over the process thanks to microprocessor etc. Improvement of smart devices will increase the process speed and reduce the errors.

The IC chip has capability of producing vide range of voltage and currents. The IC can also be programmed by designer in order to change the level of voltage and current. The general design of device can be seen in Figure 9.

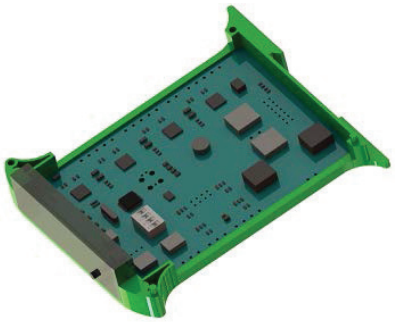


Fig. 9. *Module simulation overview*

REFERENCES

1. Hosny, A.A., Mohammed, H.A. (2015). On the Adoption of Multi-Agent Systems for the Development of Industrial Control Networks: A Case Study, ICAS 2015, Rome, Italy.
2. Balasubramanian, S., Brennan, R.W., Douglas, H., Norrie, H. (2001) An architecture for metamorphic control of holonic manufacturing systems, *Computers in Industry*, vol.46, no.1. pp.13-31.
3. Al-Zobi, Q., Al-Tawil, I., Gharaibeh, K., Al-Kofahi, I.S. (2008). Design of Power-Line Communication System (PLC) Using a PIC Microcontroller, *Journal of Active and Passive Electronic Devices*, vol. 3, pp. 331–340.
4. Almeida, E.E., Luntz, J.E., Tilbury, D.M. (2007). Event-Condition-Action Systems for Reconfigurable Logic Control, *IEEE Transactions on Automation Science and Engineering*, vol.4, no. 2. Pp.167-180.
5. Myers, G.J., Yu, A.Y.C. House, D.L. (1987). Microprocessor Technology Trends" Microelectronics Reliability, Proceedings of the IEEE, vol.74, no.12, pp.1605-1622.
6. Walden, R.H. (1999). Analog-to-digital converter survey and analysis, *IEEE Journal on Selected Areas in Communication*, vol.17, no.4, pp.539-550.
7. Inoue, S., Akagi, H. (2007). A Bidirectional Isolated DC–DC Converter as a Core Circuit of the Next-Generation Medium-Voltage Power Conversion System, *IEEE Transactions on Power Electronics*, vol.22, no.2, pp.535-542.
8. Ching-Wen Hsue, Chih-Jen Lin (1993). Built-in Current Sensor for IDDQ Test in CMOS, IEEE International Test Conference, pp.635-641.
9. Leens F. (2009). An Introduction to I2C and SPI Protocols, *IEEE Instrumentation & Measurement Magazine*, pp. 8-13.
10. Machacek, J., Drapela, J. (2008). Control of Serial Port (RS-232) Communications in LabVIEV, Modern Technique and

Technologies MTT 2008 Conference, pp.36-40.

Authors' contacts

Cagri Candan,
Istanbul Technical University, Department of Electrical Engineering, Istanbul-Turkey.
E-mail: cagricandan91@gmail.com

Emre Can Dikmen,
Istanbul Technical University, Department of Electrical Engineering, Istanbul-Turkey.
E-mail: emrecan@uestco.com

Tahir Cetin Akinci,
Istanbul Technical University, Department of Electrical Engineering, Istanbul-Turkey.
E-mail: akincitc@itu.edu.tr

Serhat Seker,
Istanbul Technical University, Department of Electrical Engineering, Istanbul-Turkey.
E-mail: sekers@itu.edu.tr

© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 23, 2017

INDUSTRIAL NETWORK DESIGN USING LOW ENERGY PROTOCOLS

STOITCHO PENKOV, ALBENA TANEVA, MICHAIL PETROV, VASIL KALKOV

Abstract: *In this work a network development for industrial application is presented. Solutions related to the network protocols and standards for industrial applications are summarized. The main goal is focused on net operation by combining low energy protocol and transmission method. In the developed network project for industrial data exchange was use MQTT protocol and LoRaWAN. An application for security purpose is obtained. The real test and verification with sending and receiving data between the connected nodes are made. The advantages of the developed network with MQTT and LoRa are given.*

Key words: *Industrial Network, MQTT protocol, LoRa*

1. Introduction

Using network in security applications is usual, complex and sometimes dangerous tasks. There are many different ways to solve, depending of cases. An example a security in places without electricity in order to be organized, is not easy task. How to secure buildings or value staff against tiffs and there no electricity? These tasks over a long period of time are performed. It is case of highly dynamic phenomena, such as change of security status because of fire alarm or intruder, means that the results do not reflect the real value. A solution of such problem is to use low cost nodes equipped with relevant networked sensors for data collection. Several nodes can be organized and formed a grid which will bring more complexity, and we will gather all needed information. If the goal is to secure object, or to gather info under the sky, it could be use nodes with GPS to know exact location, but considering power plan that is not energy efficient. Where is needed can be used mobiles and flying robots for deploying these nodes, and at moment when nodes are deployed, it can be marked GPS position. Thus are covered large areas and different surfaces. One of the main factors for such system development is the implementation cost. In order to achieve it is appropriate to use small, low-power nodes, which are controlled and supervised by a main network grid organized by LoRaWAN.

The paper presents a WSN extension to provide communication between the nodes and the gateway over LoRaWAN, instead of IEEE 802.11 (WiFi), instead of 3,4G.

This paper is focused on configuration and development of the networked sensors for monitoring and control the environment and security system. Many articles show variety of protocols and communication networks [2, 3, 4, 5]. Generally efforts are focused on problems with network possibility and delays, packet dropouts, address channel limitations related to the packet-rates, but there is no universal solution. In the paper problems definition, development suggestions and maintenance for networked sensors are made. This work summarizes solutions related to the network protocols and standards for predefined application. Proposition of node based sensor network with specified low energy protocols are also presented.

2. Layer protocols

Regarding to the OSI model is important to discuss the way of data transmission through layers. In this point of view the transmission method will be summarized.

2.1 LoRa™ is a proprietary spread spectrum modulation scheme that is derivative of Chirp Spread Spectrum modulation (CSS). It trades data rate for sensitivity within a fixed channel bandwidth. It implements a variable data rate utilizing orthogonal spreading factors, which allows the system designer to trade data rate for range or power. Furthermore the network performance is optimized in a constant bandwidth. LoRa™ is a PHY layer implementation and is agnostic with to higher-layer implementations. This allows LoRa™ to coexist and interoperate with existing network architectures. This application note explains some of the basic concepts of LoRa™. Modulation and

 ISSN 1310 - 8271

the advantages of the scheme can provide when deploying both fixed and mobile low-power real-world communications networks [6].

`

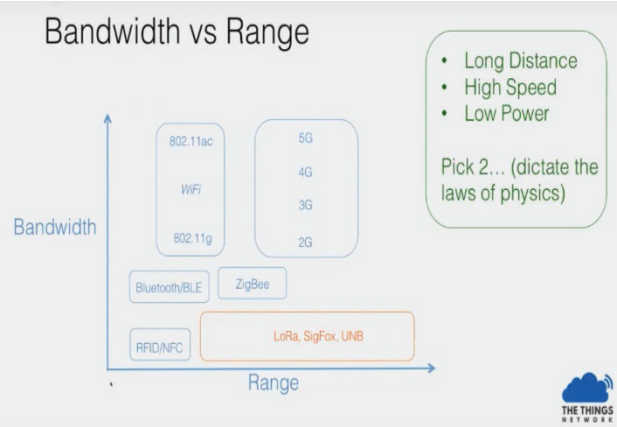


Fig. 1. *Representation of the relationship between Range and bandwidth of the signal*

Fig.1 presents relations between signal range for successful connection establishment and frequency bandwidth. In case of frequency decreasing regarding to the modulation the signal path increases with LoRa.

2.2 LoRaWAN™ /layer2/ is a Low Power Wide Area Network (LPWAN) specification intended for wireless battery operated Things in a regional, national or global network LoRaWAN would correspond to the Media access control (MAC) layer. LoRaWAN targets key requirements of Internet of Things such as secure bi-directional communication, mobility and localization services. It specification provides seamless interoperability among smart Things without the need of complex local installations and gives back the freedom to the user, developer, businesses enabling the roll out of Internet of Things.

The choice of transport protocol, when internet connectivity is need basically, is reduced to two options TCP and UDP. The Protocols allow multiple devices to communicate effectively using the Internet. However, the determination way for different data types, how are divided and stored in frames, are required. In case of a system design for collecting data in order to reduce the workload, related to the data exchange organization, it is possible to use application-layer protocols. There is variety of options. One is to use industrial automation protocols. In problem definition arises high cost of implementation.

By analyzing network protocols, extra attention should be pay to the model they use for data exchange. Many of the technologies used in modern computer systems use a data exchange model referred to as Request-Response. However, when you try to use such a model for data exchange in the sensor network you can encounter some difficulties. A possible solution is using the Publish-Subscribe method. In this method, the data publishing modules send it to a server called the broker, which then sends the data to clients subscribed to certain information. Using these methods of data exchange allows the clients to receive not all the information sent by the node, but only the data that interest. There is also no need to constantly calling the modules that generate information about the data.

2.3 MQTT protocol

Among different options described one of the most appropriate is MQTT protocol (Message Queue Telemetry Transport), details in [5]. It was designed in 1999 for transferring data from telemetry devices. The main goal of the designers was to create an efficient protocol to transfer data from devices with limited hardware resources, which is equipped with a low-performance microprocessors and a small amount of memory. Also expected to work in networks with severely limited bandwidth for data transmission. The protocol uses a publish-subscribe method and transmits the data over TCP/IP or UDP. In its implementation requires a special computer called a messages broker. The task of the broker is to collect messages and sending them to devices interested in specific information. Fig. 2 shows the organization diagram for data exchanging between Publishers and Subscribers by MQTT Broker.

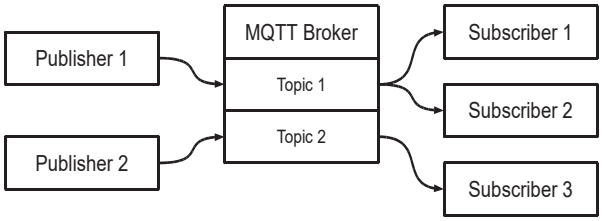


Fig. 2. *Organization of data exchange in MQTT.*

MQTT protocol messages are assigned to names that are topics. In context of the client and the broker, there is no need to configure the topic. The client sends a message to a specific topic. If there is a particular topic the broker will update its data, in the absence a new topic will be created automatically, to which will be assigned the information transmitted in the message. Topics may be organized in a hierarchical manner using the separator in the form of a forward slash (/). This allows us to organize data in a broker in a manner

similar to the file system. Example topic for networked grid nodes may have the following form:

Network22/NODE11/sensor33/DATA

An important feature of the MQTT protocol is the ability to manage the quality of service by implementing QoS (Quality of Service). It allows you to manage the way to deliver a message and confirmation of its receipt.

3. System design and configuration

The developed solution for industrial network combines LoRa and MQTT protocol.

An important step in designing communication layer using MQTT protocol is to determine the structure of topics and related messages. In current system only test messages were send, which allows reading data from nodes sensors.

Therefore is achieved LPWAN based LoRaWAN, presented on fig.3. It is evident there is variety of tasks: many sensor types, devices through Gateway to the many user defined applications.

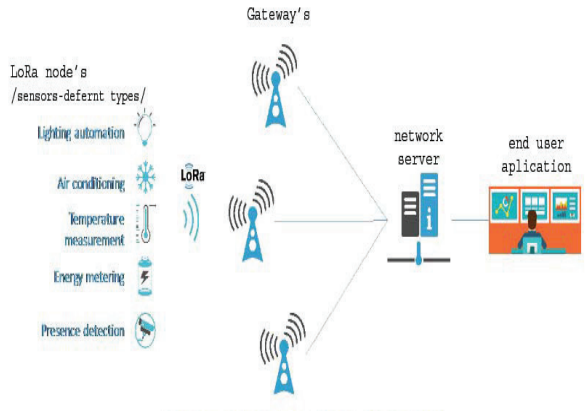


Fig. 3. *Overview of the developed network with industrial applications*

The network can consist of thousands nodes. In case of new LoRaWAL project is need a Gateway in case with no LoRa coverage. The proposed solution includes Gateway. This can be viewed as an advantage of the application, fig.4. LoRaWAN uses licence-free spectrum, usually ISM (Industrial, Scientific, Medical) bands to communicate over the air. In Europe, ETSI regulates the ISM band access on the 868MHz and 433MHz bands.

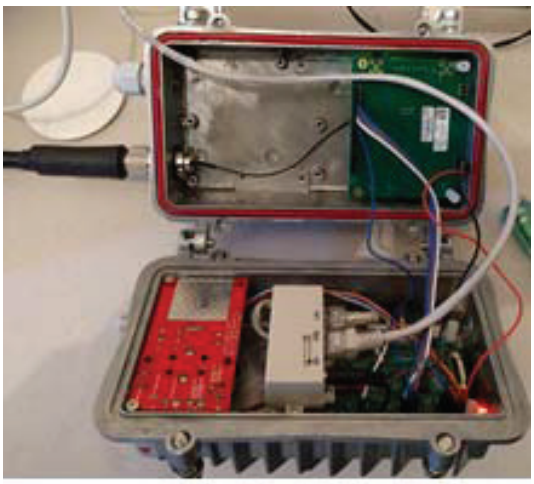


Fig. 4. *Designed and Assembled Gateway included in LPWAN*

The usage of these bands is submitted to limitations: The output power (EIRP) of the transmitter shall not exceed 14dBm or 25mW, and the duty cycle imposed in Europe by ETSI is limited to 1% (for devices) or 10% (for gateways) depending on the used sub-band. In this project the Gateway is placed on the laboratory roof. As an option in a startup it was assumed. The achieved signal coverage was good enough. In future work it will be improved. On the next fig. 5 are shown the used and connected LoRa nodes based on Microchip ®. In this way is developed and obtained network based on LoRAWAN.

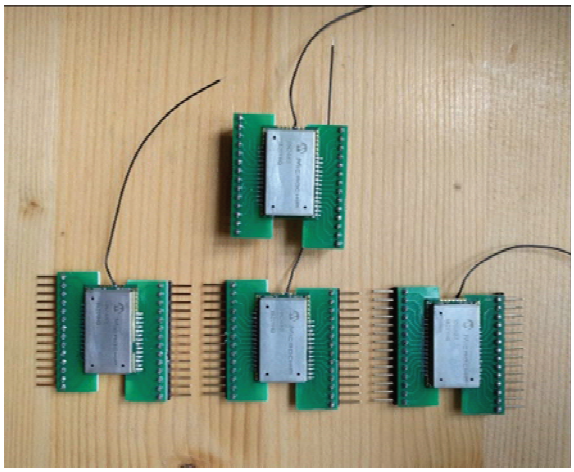


Fig. 5. *Numbers of the used nodes based on Microchip in the developed network.*

The architecture including three different classes (A,B,C), of communication profiles are available in LoRa networks between devices and applications, are presented on fig.6.

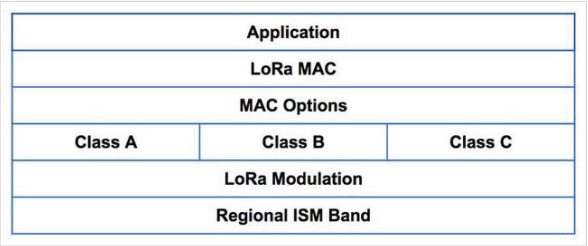


Fig. 6. *Typical system architecture of a LoRaWAN node*

Class name	Intended usage
A (« all »)	**Battery powered sensors**, or actuators with no latency constraint Most energy efficient communication class. Must be supported by all devices
B (« beacon »)	**Battery powered actuators** Energy efficient communication class for latency controlled downlink. Based on slotted communication synchronized with a network beacon.
C (« continuous »)	**Mains powered actuators** Devices which can afford to listen continuously. No latency for downlink communication.

Fig. 7. *Description of A,B and C profiles*

Each class serves different application needs and has optimised requirements for specific purposes. The key difference between A,B and C profiles is the trade-off made between latency and power consumption, fig.7. It is a difficult task to gather information from many and different points. In a stage of problem definition the task is related to a network design and configuration of nodes equipped with sensors. The proposed solution is based on a grid of networked nodes. If there is a lot of equipment, it is important to find a low-cost solution with low energy consumption. In order to reduce costs it is necessary to minimize the tasks of networked nodes. Therefore cable connections are not relevant or suitable. In such case of many distributed nodes cooperation, the option is to use radio communication. The solution performs the data transmission from the sensors to the gateway /one or many/. This allows using of nodes with low-cost microcontrollers. The power consumption. Easy operation and reconfiguration of the system are also given and discussed in [6].

Nowadays there are many radio communication standards, but not all are well suited to the task. The ability and widely used popular, not expensive modules with easy connection to the Internet is very important.

First one solution that comes to mind is the standard IEEE 802.11 (WiFi). It allows connecting multiple devices to the network, and easy integration with the Internet. However, it has the disadvantage of relatively high demand for energy, which is particularly important in the case of using small nodes to collect data. Among the standards for devices with low power the most interesting solution is using **LoRaWAN** network

IEEE standards are for the physical layer of the OSI model. They allow transferring data. But do not provide a convenient way to control communications and meeting the requirements of QoS (Quality of Service). Accordingly, the next step is the selection of a transport layer protocol.

Level2 Sections

For second level of headings use „SectionL2” mark-up style.

Level3 Sections

For second level of headings use „SectionL3” mark-up style.

4. Application in security area

There are many places around cities (levers, fields, car morgues, etc.) without electricity and 220V power supply. From security point of view this is a big problem and complex task. To establish and maintain the communication with such plant. The options are very less. Generally is used security system equipped with transmitters and detectors using 220V power supply. They are not applicable in such case.

It is possible to use GSM operators. We are facing with need to use special equipment with batteries. These batteries should therefore be charged. Even in this way the operability depends of the power supply. The lifecycle is not so long. This weakness is other come with the developed solution combining LoRaWAN and MQTT. In this way can be achieved promising security without 220V power supply.

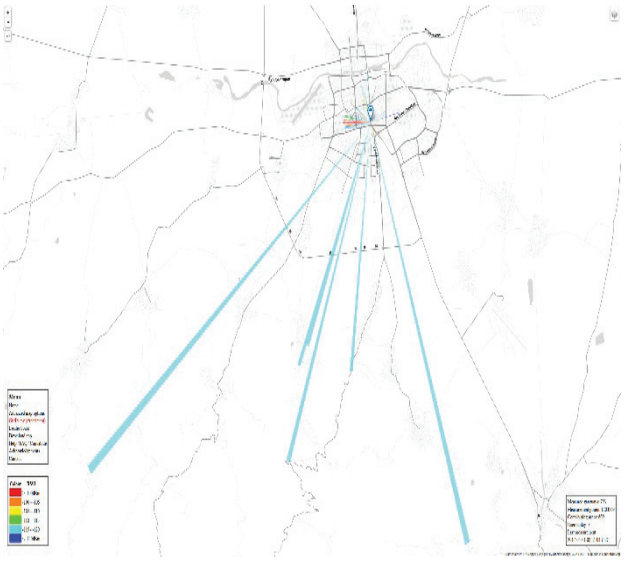


Fig. 8. *The coverage measurement with TTN mapper*

.LoRa uses low energy hence the workability is expected to be 5-10 years. To one node could be connected more than one detector and passive detectors as well.

The radio coverage of developed network, fig.8 can be viewed as an industrial solution in case with IoT implementation. The signal covers the laboratory location and connected nodes. To investigate and conduct the preliminary experiments is used software named TTN /The Things Network/ mapper, fig.9. The signal coverage and connected node was monitored. The results demonstrate the good enough signals. It shows the achieved passing of the signal at a large distance with low power.

TX- is the emitted power by the transmitter. The signal is gained by the Gateway. The signal above the dashed red line confirms the established and "good" connection.

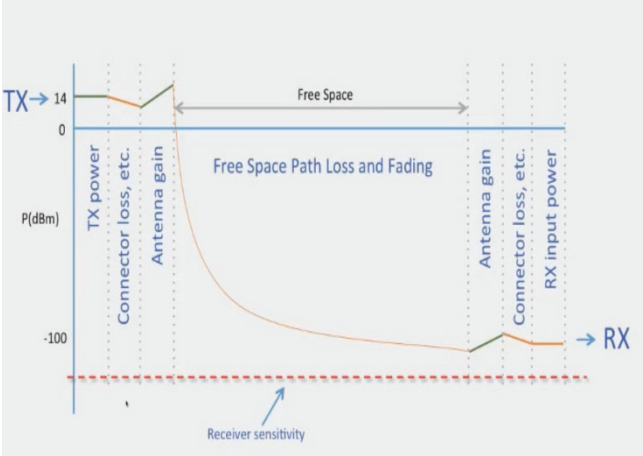


Fig. 9. *The TX and RX signals*

5. Conclusion

This work is devoted on developing, configuration and investigation of the network for security purpose in the case of missing 220V supply. A solution is found out by combination between the LoRaWAN network and MQTT protocol. All developed nodes covered IP 67 standard.

Some advantages of the system are: easy maintenance; low cost; low energy consumption; reliability and security; easy connection to the Internet, long lifecycle. The system nodes (the grid) can be enlarged with up to more than 1000 nodes connected with one gate.

In the particular test application for the industrial network development it is obtained promising results. There are eight connected nodes in this case study. Furthermore it will be possible and very easy to extend the node numbers if necessary. The idea could be associated with networked sensors for robots to perform various tasks. To industrial environment for monitoring process values like temperature, air conditioning systems and etc. Each "node" could present sensor one or many, or robot sensors, connected the main node through the Gatrway with LoRa. It is possible to use robots to spread sensors in case of hostile environment.

The paper shows an example of the use of open network protocols in networked mobile robot sensor system. MQTT protocol allowed for the use of standardized methods for data exchange in the sensor network. Also greatly simplified the integration of new nodes, the use of nodes information from other systems and integrating with the Internet.

The use of open protocols simplifies software development work, especially when it consists of a large number of independent nodes. Also simplifies the maintenance process, since it is possible to read information about the industrial values and the status of individual sensors without having to use special tools. Future work will be focused on industrial network implementation.

REFERENCES

1. Souza R., L. Agostinho, F.Teixeira, D.Rodrigues, L. Olivi, E. Guimaraes, E.Cardozo. *Control of Mobile Robots Through Wireless Sensor Networks.* XXIX Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos, pp.805-818.
2. GreenPeak Technologies (2015), *Wireless Communication Standards for the Internet of Things,* White paper
3. Silicon Laboratories, Inc., *The Evolution of Wireless Sensor Networks*, 2015.
4. ITU-T, *Ubiquitous Sensor Networks (USN)*, ITU-T Technology Watch Briefing Report Series, No. 4, February 2008.
5. MQTT V3.1 Protocol Specification. International Business Machines Corporation, Eurotech, IBM, 2010
6. Beser, Nurettin Burcak. "Operating cable modems in a low power mode." U.S. Patent No. 7,389,528. 17 June 2008. – тва не го знам какво е
7. SEMTECH, AN1200.22 LoRa™ Modulation Basics
8. LoRa Device Developer Guide, Orange Connected Objects & Partnerships

ACKNOWLEGEMENTS

The authors would like to acknowledge Ministry of Education and Science of Bulgaria, Research Fund Project No: DFNI-I 02/6/2014.

Department of Control Systems
Technical University–Sofia, Branch Plovdiv
25 Tsanko Diustabanov St.
4000 Plovdiv

Phone: +359 032 659 585
E-mail: stoitcho@abv.bg
E-mail: altaneva@tu-plovdiv.bg
E-mail: mpetrov@tu-plovdiv.bg
E-mail: vassilkalkov@gmail.com

© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 23, 2017

DIPOLE ANTENNA OVER EBG STRUCTURE FOR UHF RFID APPLICATIONS

ANGEL SLAVOV, PETER Z. PETKOV, STOYAN ILIEV

Abstract—*This paper examines the possibility of designing a low profile antenna with high gain for RFID identification in the frequency range 865 to 868MHz. The advantages of the suggested structure are high gain and low profile compared to traditional antennas.*

Key Words: *EBG, antenna, dipole, gain, RFID*

DIPOLE ANTENNA OVER EBG STRUCTURE FOR UHF RFID APPLICATIONS

ANGEL SLAVOV, PETER Z. PETKOV, STOYAN ILIEV

Abstract—*This paper examines the possibility of designing a low profile antenna with high gain for RFID identification in the frequency range 865 to 868MHz. The advantages of the suggested structure are high gain and low profile compared to traditional antennas.*

Key Words: *EBG, antenna, dipole, gain, RFID*

1. Introduction

Modern antenna technologies have evolved extremely and rapidly last years. One of the driving reason for that evolution is so-called metamaterials. In this paper we will focus on a specific part of metamaterials namely Electromag- netic Band Gap structures (EBG). These metamatirials exhibit electromagnetic features which may not exist in nature. These materials have simultaneously < 0 and $\mu < 0$. [1] This type of structure is applicable to a wide range of applications in antenna and propagation fields. For instance it can be used as a spiral and curl antenna to achieve low profile design. [2] One of the most interesting properties of metamaterials is the reflection phase. It is defined as the phase of the reflected electric field at the reflecting surface compared to the incident one. [3] It is known that a perfect electric conductor (PEC) has 180 ◦ reflection phase for a normal incident plane wave. A perfect magnetic conductor (PMC) has the reflection phase 0 ◦ but does not exist in nature. [1], [3]

However, EBG structures are more than the PMC surface. The reflection phase of the EBG surface varies from 180 ◦ to−180 ◦ versus frequency, not only 180 ◦ for the PEC surface or 0 ◦ for the PMC surface. This reflection phase feature makes EBG unique. For example one possible application of that surface it is usage as a ground plane of wire or patch antennas for low profile design which is desirable in many wireless communications systems. [3]

A basic question here is: How effective is the EBG struc- tures for low profile antenna applications? Mushroom-like is known to have effective bandgap for a surface-wave propa- gation. It can be used to improve antenna radiation patterns. When the incident wave is a plane wave (kx 2 + ky 2 ≤ k0 2 , kz has a real value), the reflection phase of the EBG structures varies with frequency. At certain frequency the reflection phase is zero degree, which resemble the perfect magnetic conductor that does not exist in nature. [1]

To understand how the EBG improves antenna parameters, a vector dipole antenna is used. The antenna consists of two orthogonal dipoles. The advantage of the suggested construction is that the antenna can realize circular polarization which is

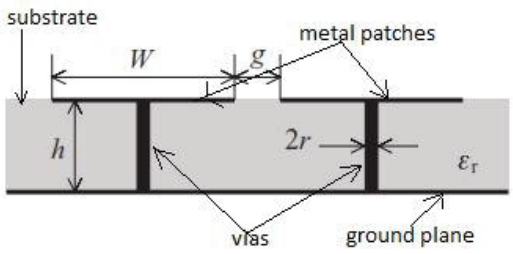
 ISSN 1310 - 8271

important for the RFID communications if both dipoles are used. Here only one dipole is used as radiator, for the sake of construction simplification.

The paper focuses on mushroom-like EBG structures in- vented by Sievenpiper et al. [4]. The structure studied here has a feature of compactness and can be integrated into printed circuit boards, which is very critical in handheld devices. To prove that it is necessary to be used the vector dipole antenna, the conventional ground plane and the EBG structure. Consequently three comparisons on antenna gain are presented: the vector dipole antenna, the dipole antenna with metal reflector and the dipole antenna with EBG structure.

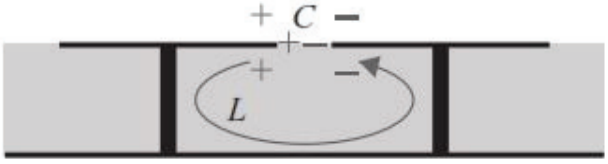
2. The EBG Structure

The mushroom-like EBG structure displayed here consisted of four elements: metal patches, a substrate, a ground plane and vias, which connected metal patches and the ground plane. This is illustrated in Fig. 1.



Фиг. 1. *Schematic diagram of the EBG structure*

The parameters in Fig. 1 are: w - width of the patch; g - patch distance; h - substrate height; r - radius via; r = 4, 3. The functionality of the EBG structure can be explained by a LC equivalent circuit model [1], shown in Fig. 2.



Фиг. 2. *Equivalent scheme of the EBG structure*

The inductance in that model comes from the current flowing along adjacent patches and connecting vias. The ca- pacitance result comes from the fringing electric field between adjacent metallic patches. The equivalent circuit model is able to predict the reflection phase as well as some surface wave properties. The capacitance and inductance in the equivalent circuit can be approximated by following formulas [5].

$$L = \mu_0 \mu_r h \quad (1)$$

$$C = \frac{\omega \varepsilon_0 (1 + \varepsilon_r)}{\pi} \cosh^{-1}\left(\frac{\omega + g}{g}\right) \quad (2)$$

The fractional bandwidth (BW) ±90 ◦ of the reflection phase is received by the following formulas [5].

$$\omega_0 = \frac{1}{\sqrt{LC}} \quad (1)$$

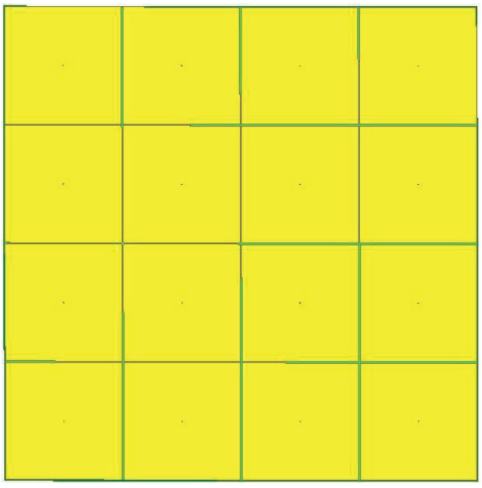
$$BW = \frac{\Delta\omega}{\omega_0} = \frac{1}{Z_0}\sqrt{\frac{L}{C}} \quad (2)$$

3. Results

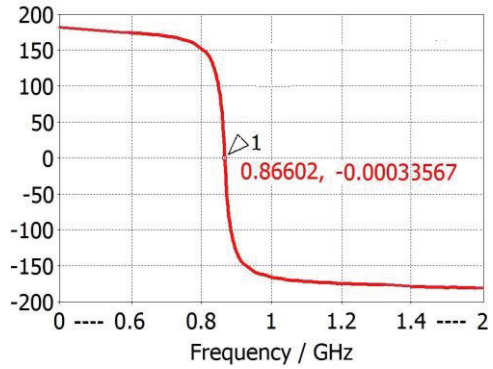
3.1. EBG Structure

For the RFID frequency range 865 MHz to 868 MHz the EBG structure is made with finite dimensions 300×300mm. The substrate material is FR-4 with thickness 2, 4mm. Accord- ing to formulas 1, 2 and 3 patch width is approximately w = 70 mm. Consequently the structure consists of 16 patches - 4×4. To simulate radiation structure described above commercial software CST Microwave Studio is used. The final dimensions after fine tuning are: patch width - 74, 24 mm; gap width - 0, 76 mm and via radius - 0, 25 mm. The EBG surface is shown in Fig. 3.

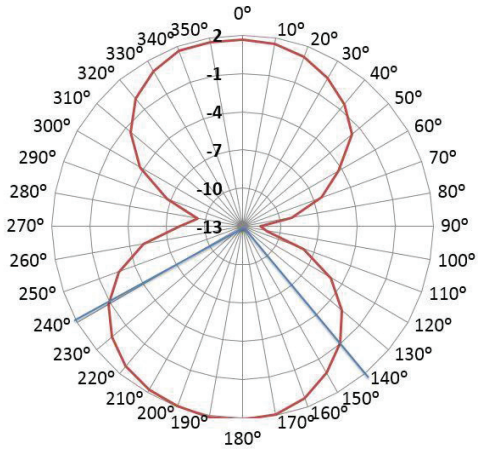
The reflection phase of the structure is determined by CST and presented in Fig. 4. As one can see from Fig. 4 the surface has reflection phase 0 ◦ at 866 MHz which is in desire bandwidth between 865 -868 MHz. In ±90 ◦ the construction has 30 MHz bandwidth from 850 to 880 MHz.



Фиг. 3. *The finite EBG structure*



Фиг. 4. *The Reflection phase of the EBG structure*



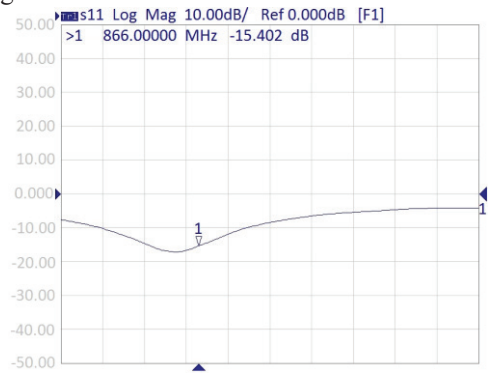
Фиг. 5. *The Radiation pattern of the dipole antenna (no reflector is employed)*

3.2. The Dipole antenna

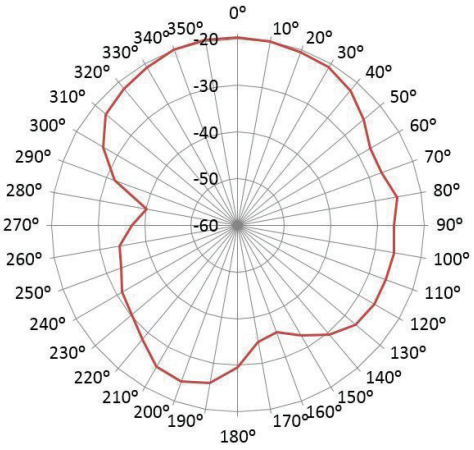
In this subsection the measurement results of the dipole antenna are presented. The antenna dimensions are: width – 104 mm, length – 104 mm, height – 2, 4 mm and substrate – FR-4. For feeding of the antenna a coaxial cable (50Ω) and balun mounted on a small PCB circuit in the middle of the antenna are used. The radiation pattern of proposed dipole is displayed in Fig. 5, resulting in gain of 2 dBi. The input return loss is displayed in Fig. 6.

3.3. The Dipole antenna with metal reflector

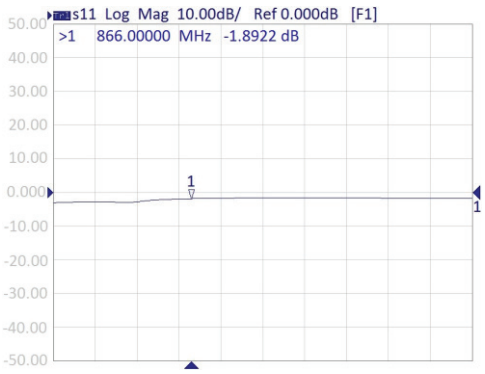
In this subsection the results from the dipole antenna over a plain sheet reflector are presented. The distance between the antenna and metal reflector is 1 mm. Later will be shown why a gap is chosen exactly 1 mm. For feeding of the antenna is selected again balun and coaxial cable (50Ω). There is a small hole on the metal reflector used to pass thru the coax feed cable. The dipole is fixed to the reflector with double sided adhesive tape with exact thickness of 1 mm. The measured radiation pattern of the dipole reflector struc- ture is displayed in Fig.7.



Фиг. 6. *Return loss of the dipole antenna (no reflector)*



Фиг. 7. *The Radiation pattern of the dipole antenna over a metal reflector*



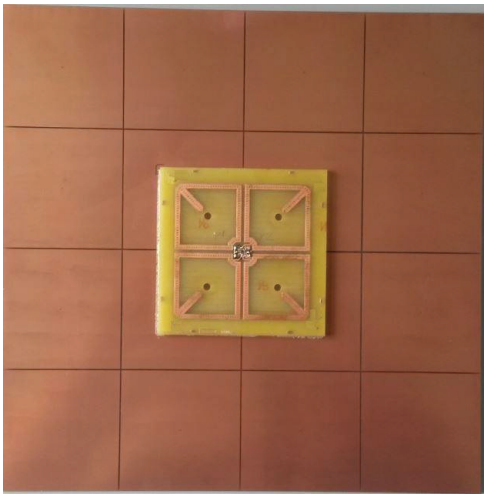
Фиг. 8. *The Return loss of the dipole antenna over ground plain sheet metal reflector*

3.4. D. Dipole antenna with EBG reflector

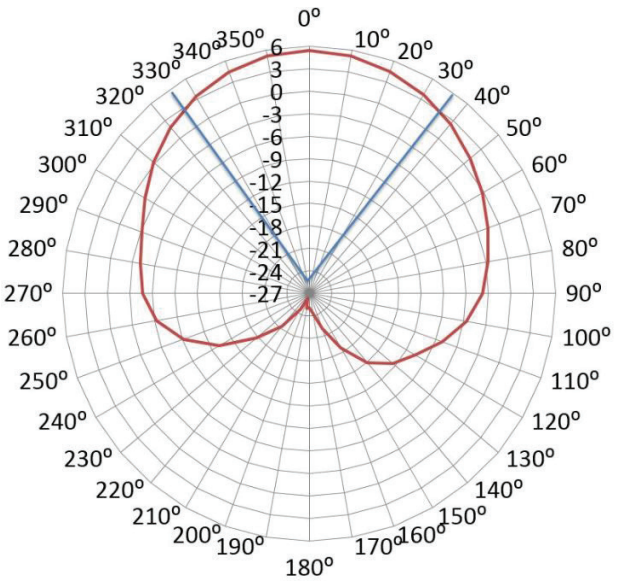
A driving reason for researching the low-profile (closely spaced reflector) antenna is the necessity of compact systems for handheld applications. To achieve high gain and low profile at the same time, a parametric analysis is performed. The results are shown in table I.

From table I it can be concluded that a reasonable com- promise between high gain and low profile antenna can be achieved at 1 mm reflector distance from the radiating element. When the distance becomes over 15 mm the efficiency of the EBG surface significantly decreases.

The real construction is shown in Fig. 9. The final dimen- sions are: width – 300 mm, length – 300 mm and height – 5, 8 mm. For feeding of the antenna here also are used coaxial cable (50Ω) and balun which is mounted on a small PCB circuit in the middle of the antenna. Also on the EBG reflector there is a small pass-thru hole. The dipole and the EBG reflector are connected with double sided adhesive tape which has 1 mm thickness such as on the previous model with plain reflector.



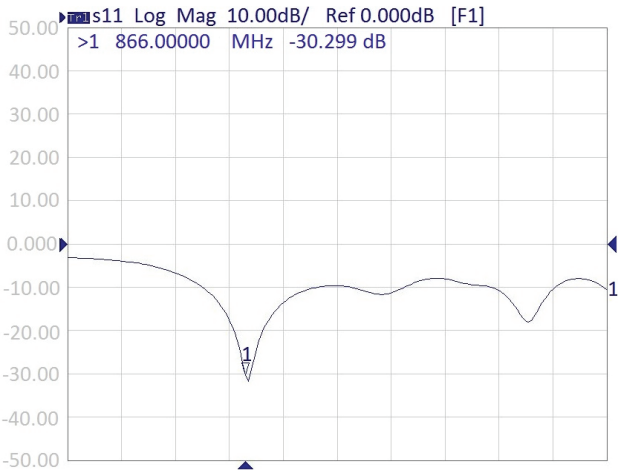
Фиг. 9. *The Real model of the antenna*



Фиг. 10. *The Measured radiation pattern of the dipole antenna over EBG structure*

Figure 10 shows the radiation pattern of the antenna with EBG structure. Also figure 10 shows gain of the antenna is G = 5, 7 dBi. The width of the main lobe is 70 ◦ which is a typical angle for RFID-application antennas. The return loss of the antenna is displayed in Fig. 11.

According to Fig. 11 the antenna has good impedance matching with reasonable bandwidth which makes antenna insensitive to manufacturing tolerances.



Фиг. 11. *The Impedance matching of the dipole antenna over EBG structure*

4. Conclusion

This paper introduced possible design of dipole antenna with EBG surface. EBG surfaces are used to increase antenna gain and decrease the antenna thickness compared with tra- ditional air gap antennas. All this features give the antenna possibility for mass production. It is intended mainly for RFID communications.

REFERENCES

1. F Yang and Y Rahmat-Sami, *Electromagnetic Band Gap Structures in Antenna Engineering*, Cambrigde University Press, 2009.
2. T.H. Liu, W.X. Zhang, M. Zhang and K.F.Tsang, ”Low profile spiral antenna with PBG substrate*”, Electron Lett*., vol.36, no.9, 2000, pp. 779 − 780
3. F Yang and Y Rahmat-Sami, ”Reflection Phase Characterizations of the EBG Ground Plane for Low Profile Wire Antenna Applications”, *IEEE Transactions on Antennas and Propagation*, Vol. 51. 10, October 2003, pp. 2691 - 2703
4. D Sievenpiper, L. Zhang, R. F. J Broas, G. Alexopolus and E. Yablonovich ”High-impedance electromagnetic surface with a forbidden frequency band”, *IEEE Trans. Microwave Theory Tech*., vol. 47., November 1999, pp. 2059 - 2074
5. T.T. Nguyen, D. Kim, S. Kim and J. Jang, ”Design of a wide- band mushroom-like electromagnetic bandgap structure with magneto- dielectric substrate”, *ICITA*, 2000.

Contacts
Angel Slavov
Technical University of Sofia
Sofia, Bulgaria
email: angel.slavov.13@gmail.com

Peter Z Petkov
Technical University of Sofia
Sofia, Bulgaria
email: pjpetkov@tu-sofia.bg

Stoyan Iliev
Kathrein RFID
Stephanskirchen, Germany
email: s.iliev@kathrein-rfid.de

© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 23, 2017

A REVIEW ON RECENT ANTENNA DESIGNING TECHNIQUES FOR ELECTROMAGNETIC COMPATIBILITY(EMC) TEST

SARANG M. PATIL, PETER Z. PETKOV, BONCHO G. BONEV

Abstract: *All devices from electronics category must meet EMC requirement. In this article presents a review summary of various antenna structure designed for EM signal radiated emission and susceptibility measurements currently used for evaluating the electromagnetic compatibility/interference (EMC/EMI) characteristics of electronics systems and devices. For accurate EMC/EMI test require to capture of unknown electromagnetic radiation from an equipment using suitably calibrated EMI sensor. The response of sensors is analyzed in terms of the antenna factor-which is the ratio of the incident electric field on the antenna surface to the detected voltage at load. At the end comparison is done of all antennas specifications and new idea is proposed for improvement in EMC measurement*

Key words: *Antenna Factor, Electromagnetic compatibility, Electromagnetic Interference sensor, Immunity.*

1. Introduction

All the electronics devices and appliances are important part of our day today life, like kitchen tools to satellite, because of the more demands for wireless devices, the electromagnetic territory has becomes polluted. Electromagnetic interference (EMI) both inter- and intra- device is the well-known "pollutant".

Electromagnetic compatibility is the member of electrical science which deals with unintentional propagation, reception and generation of electromagnetic energy with reference to the unwanted reaction that such energy may induce. Electromagnetic compatibility testing of electronics equipment Necessitate to measurement of field strength radiated form devices [1].Radiated emission from equipment are measured and calibrate the field level for immunity test. Ordinarily, antennas with large bandwidth are used for emission measurement and diode based field sensors are widely used for calibrating field level for immunity test [2].The definition of EMC/EMI as per IEEE dictionary as follows [3].

Efficient EMC measurement, observes and determines the behaviour of electrical apparatus from the prospect of EM radiation with electronic circuits [4].In communication system various electronics equipments are used, at side of transmitter are designed to radiate a specific radio-frequency, power at selected frequencies to antenna, antenna radiated signal intentionally but along with that another uninitiated signal also transmitted, it is form of EMI. Another non-radiators devices is also present in surrounding like computers, ideally computer performs task and during the computation time its generate some signals internally because of electronic circuits within the system will be contained and not be radiated, however some signals are radiated as EMI because of internal signals are not contained totally [5].

Electromagnetic pre-compliance measurement applications antenna is not used widely used because of large in size and are sensitive to nearby reflections and interact with surrounding metal objects .If EMC testing is carried out during the design phase of product it provides following advantages,

- Passing rate of final compliance test is increased.
- Retest count is decreased at EMC test laboratory.
- Eliminates surprises late in the design.
- Conform that EMC considerations are part of the original.

As per the data received for EMC test laboratory, 50% submitted products are fail in compliance testing at first attempt. To reduce this statistic and improve the percentage of passing form 50% to 90% need to use EMC measurement antenna /Probe at designers/Manufacturers end [6].

A variety of antennas have been designed to help the EMC engineer test quickly over the range from 30MHz to 1GHz, and they can have quite interesting shapes [7]. A very significant concept about all antennas is that they required calibrating, and their calibration factors must be taken into consideration in any measurement of

 ISSN 1310 - 8271

radiated emissions [8]. Test field calibration is described and analyzed in [9].

2. Common Antennas Used for EMC Testing

Now a day various antenna types and geometry are present, some of them are commonly used for EMC test, they are as below:

2.1. Tunable Dipole Antenna

Electrical field measurement purpose tunable antenna is widely used, this antenna support for the range of frequency over 25-1000 MHz Tunable dipole having two types- the resonant dipoles and half-wavelength dipoles.

2.2. Biconical Antenna

This antenna is called broadband dipoles that consist of two conical conductors having a common axis and vertex. The receiver is connected at vertex. Biconical antenna operates in the range of 30-300 MHz, this antenna is best for vertical polarization measurement because of smaller in size.

2.3. Log-Periodic Antenna

This antenna is called log-periodic because of structural geometry that its impedance and radiation characteristics repeat periodically as logarithm of frequency. It's operating range of frequency from 200-1000MHz

2.4. Bilog Antenna

This antenna is called log-periodic because of structural geometry that its impedance and radiation characteristics repeat periodically as logarithm of frequency. It's operating range of frequency from 200-1000MHz

2.5. Loop Antenna

The shape of this antenna is like coil and is highly sensitive for magnetic field but shielded against electric field. Measurement of electromagnetic field takes place with electrically small loop for the range of frequency approximately 20Hz-30MHz.

2.6. Horn Antenna

This antenna is specially used for high range frequency measurement of electromagnetic signals. This measure frequency above the 1GHz. Range of frequency is 1-40MHz.

3. literature survey

A number of researchers work on antenna designing for electromagnetic compatibility test. In this session consider some recants important and efficient antenna designs with specifications.

N.Narang et el.[10]. Present A Coplanar Microstrip Antenna as a Dosimetric E-field Probe for GSM /UMTS Applications. Co-planar Microstrip antenna used as E-field sensor. This probe is specifically used for Specific Absorption Rate (SAR) measurement in mobile communication. In this probe used antenna is dual band Microstrip, its support for the frequency band i.e. 915MHz and 2.10GHz in Global System for Mobile communication frequency. Over all dimensions of antenna is 30mm*100mm in size. The Microstrip design is made on a low dielectric substrate, FR4 epoxy of er = 4.4 and thickness t = 1.6 mm. structure of antenna is see in figure 1.

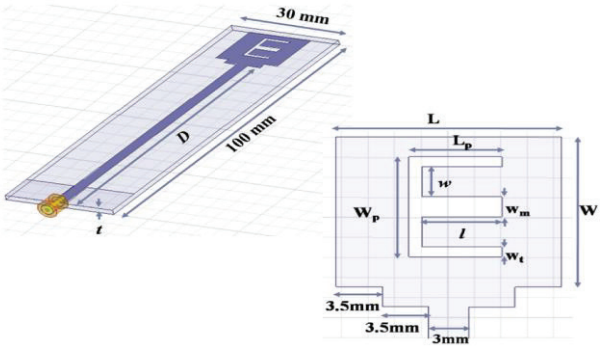


Fig. 1. *Proposed E-field probe [10].*

Mohammed Siddeqet el. [11], A New Printed Monopole Antenna as EMI Sensor is proposed, structure of proposed antenna is shown in figure 2, Tthis antenna operates in triple band resonating frequency i.e. 2.6, 3.5, 5.7 GHz. Proposed antenna overall size of 33.06mm*25mm*1.6, etched on FR-4 epoxy substrate, testing result observed like the percentage bandwidth of 6.9% and 12.2% is observed in the structure at 2.6GHz and 5.7GHz.

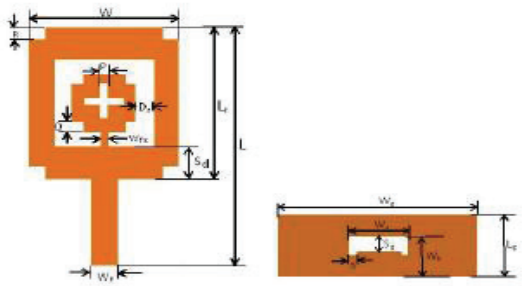


Fig. 2. *Proposed E-field probe [10].*

Zdenek Kubík et el [12], see in fig.3, has developed Optimization of electrical properties of parallel plate antenna for EMC testing, this parallel plate antenna is also called as Strip line antenna or TEM cell antenna, for the measurement device under test is kept within the uniform EM field region.

Fig. 3. *Proposed parallel plate antenna [12]*

Parallel plate antenna designed specifications are: W = 0.3m, H = 0.4 m, and thickness of plate's t = 0.0012m, this design support for the range of frequency from 9KHz to 30MHz.

Amit Kumar Srivastava et el. [13], , first proposed design of patch antenna is simulated on HFSS. After the simulation simulated module is fabricated, with substrate 'FR4' having electric permittivity, εr=4.3 with the height, h= 1.588mm is used. This antenna operates the 2.45GHz to 3.00GHz.so circular patch is act as good EMI sensor is concluded. Observe dimensions of proposed antenna in figure 4.

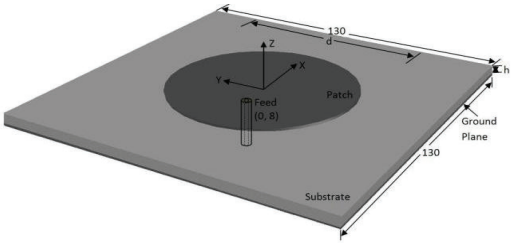


Fig. 4. *Proposed circular patch antenna [13]*

S. Capdevila, J. Romeu et el. [14], design of an Modulated Scatterer Technique (MST) probe antenna to be used in sensing applications, antenna designed with substrate RO4003C with a thickness of 1.52mm, has been scaled down to accommodate the new frequency range (from 868MHz to 2.45GHz). Figure 5 indicate proposed structure of MST antenna with dimensions in mm.

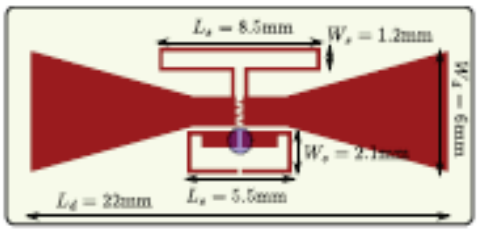


Fig. 5. *Proposed MST Probe antenna [14]*

Surbhi Mittal et el. [15], approach to electrically small loop electrically short field probe to obtain a flat frequency response using the combination with active oscilloscope probe. Proposed design supports to operate up to 5GHz.loop dimensions are 3*3mil and it is very small. Measurement results agree very well with the expectations from 0. 1 GHz to 3 GHz. Figure 6 is small loop field probe.

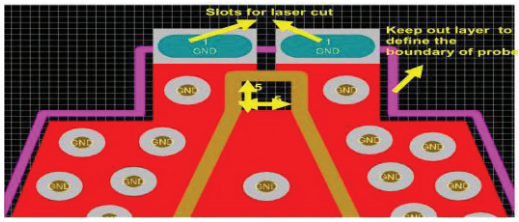


Fig. 6. *Proposed Design of Flex Circuit Probe [15]*

M. F. P. Tartaglia1 et el. [16], discussed the Metamaterial-based Probe for EMC Measurements, in this antenna author try to improve the gain, bandwidth and resonance with the supports of metamaterials. In this work planer probe used with square format along with 5mm of track width with 35 mm of thickness on the o ROGERS R3003 substrate with 1.5mm of thickness with a full ground plane shown in figure 7. EMC measurement can be performed with low cost planer probe is concluded.

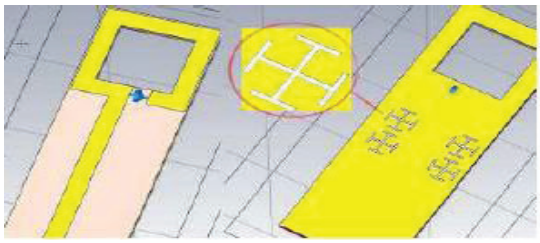


Fig. 7. *Proposed Design of Flex Circuit Probe [15]*

H. Lv et el.[17] , see Fig. 8 is the Indoor Electromagnetic Testing Antenna Design with Better Standing Wave Ratio and Gain is proposed, this covers frequency range between 0.14GHz and 6.3GHz so relatively bandwidth is 191.3% proposed. The antenna is constructed with dielectric constant of 2.55 and 0. 8 mm thickness of the dielectric substrate of polytetrafluoroethylene.

Fig. 8. *Prototype antenna module. [17]*

Xingyu Zhang et el. [18],this article present highly sensitive integrated photonic electromagnetic field sensor based on a silicon-organic hybrid modulator driven by a bowtie antenna shown in

figure 9, fabrication and characterization of this compact antenna is analyze. Device capable to measure electromagnetic power density of 8.4mW/m2, corresponding to a minimum detectable electric field of 2.5V/m. at 8.4GHZ.

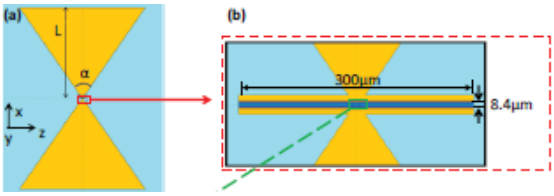


Fig. 9. *(a) Schematic top view (b) Magnified image of the feed gap region in (a). [18]*

F. H. Wee et el.[19] ,electromagnetic wave detection based on multiband antenna design proposed, antenna design is shown in figure 10 to solve the hazard electromagnetic wave detection problem a multi-band Microstrip is designed. Antenna is simulated and analyzed the results like, efficiency, VSWR, return loss, radiation pattern etc. Antenna radiates for multiple frequency like ,2.3, 4.5, 4.68, and 5.2GHz.overall dimensions of antenna is 80mm*70mm with patch 30mm*60.32mm.For mentions dimensions antenna return loss below -10dB and input impedance equal to 50Ω .

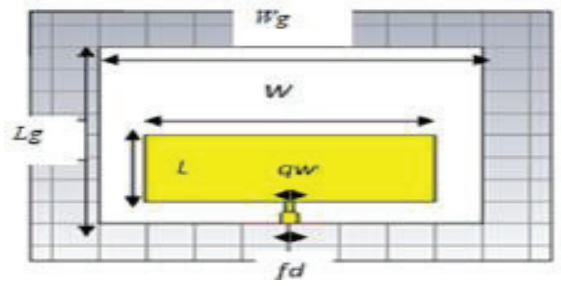


Fig. 10. *Proposed Design of Detector antenna [19]*

Nisha Gupta et el. [20] has Design of Printed Log Periodic EMI Sensor, this printed log-periodic dipoles called array with linear polarization with desired frequency range. Antenna is printed on FR-4 substrate with dielectric constant 4.4 and thickness 1.6mm, printed pattern separated by dielectric and named as element1 to 5. This 5 element dipole array supports the frequencies of 720 MHz, 1.344 GHz. and 2.64 GHz. Proposed dipole array is shown in figure 11.

Fenghan Lin et el.[21], invented 0.7–20-GHz Dual-Polarized Bilateral Tapered Slot Antenna for EMC Measurements. DBTSA in constructed with bilateral tapered slot antenna, this maintain wide operating frequency range from 700MHz to 20GHz. Figure 12 indicate Antenna structure,

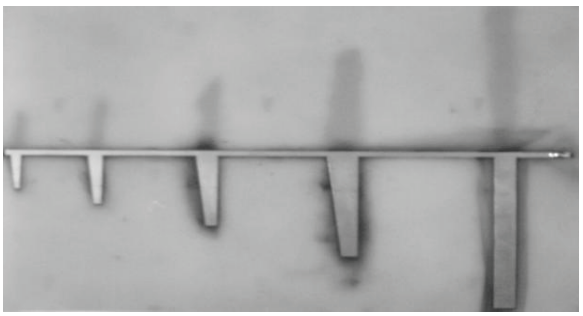


Fig. 11. *Proposed antenna Top view [20]*

The constructed of two-layer printed circuit board, substrate material is Rogers RT6002, relative dielectric constant *εr* of 2.94 and a thickness *t* of 0.762 mm (overall 1.524 mm), After result analysis of DBTSA, concluded that it is a good candidate for a dual-polarized sensor antenna for applications in wireless EMC measurement.

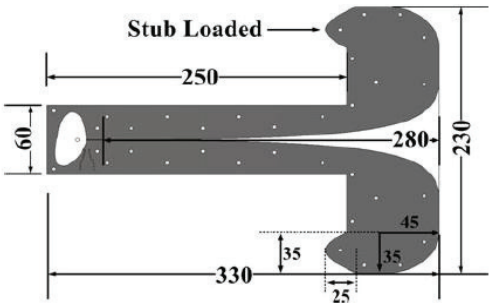


Fig. 12. *Proposed BTSA antenna [21]*

4. Important parameters for EMC/EMI Antenna Selection

The most common fundamental parameters of technical tools, the EMC antenna is used. The review of the basic parameters and related antenna parameters are discussed below:

4.1. Antenna Factor

In Electromagnetic, AF is the ratio of the electric field strength to the voltage V induced across the terminal of an antenna. This is linear definition expressed in volt per meter (V/m)

4.2. Frequency Response

Frequency response is most crucial parameter in probe characteristics, is defined as the range the probe will respond to. Probe should provide flat frequency response over all frequency range

4.3. Sensitivity

Sensitivity is defined as the how small and RF signal a probe can respond to accurately. Sensitivity is most important when small RF field should be measure. It is measured in V/m.

Table 1. *Comparative analysis of 3 -vector Probe*

Ref. Paper No.	Frequency Range.		Simulator Used	Over all antenna size.	Sensitivity	Antenna Factor (AF)dB/m	Applications
	From	To					
Ref.10	915MHz	2.10GHz	HFSS 13.0	30mm*100mm	1mV/m	50 dB/m	SAR Measurement
Ref.11	2.5 to 2.68 ,3.1 to 4.1	5.3 to 6GHz	HFSS 13.0	33.06mm × 25mm ×1. 6mmm	-	-	EMI sensor
Ref.12	9KHz	30MHz	-	W=0.3m, H=0.4m	-	-	EMC,EMI Testing of GSM cell phones, radio and TV Transmitters
Ref.13	2.45GHz	3GHz	CST Microwave Studio	130*130mm	-1.8930 to 2.4728V/m	41.7375 to 19.6dB/m	EMI sensor
Ref.14	868MHz	2.45GHz	-	22*6mm	-	-	EMC measurement
Ref.15	0.1GHz	3GHz	-	3x3 mil	3.3mV	-	EMC measurement
Ref.16	-	-	-	6cm*5cm	-	-	EMC measurement
Ref.17	0.14GHz	6.3GHz	HFSS	220*111mm	-	-	electromagnetic detection
Ref.18	8.4GHz	10GHz		12*6mm	2.5V/m.	-	electromagnetic field sensor
Ref.19	2.3GHz	5.2GHz	CST microwave studio 2014.	80*70mm	-	-	electromagnetic Wave detection
Ref.20	625MHz	2.6GHz	IE3D	Array Length=206mm		13.08-24.24dB/m	EMI Sensor
Ref.21	700MHz	20GHz	CST microwave studio	320*230mm	-	-	wireless EMC measurements

4.4. Dynamic Range

Dynamic range is the total range of RF field coverage a probe will respond to. The greater the Dynamic range the better a probe is suited to address test applications that span the gamut from low to high field strengths

5. Proposed Module

In this proposed solution for EMC/EMI measurement system, antenna constructed for three axial field measurements, designed antenna has wide bandwidth, Good antenna Factor, sensitivity should be measured in V/m ,small in size along with multiple polarization capability. Features of proposed module shown in Fig.21.

1. Conclusion

In this survey paper describes and reviews the various antenna designing techniques/ methods with different antenna geometries, Nevertheless, useful solution are still less and suffer from different problems like, Improvement in antenna factor, complexity of structure, gain, sensitivity etc, whatever antenna's currently used as EMI sensor are only detect one or two axial components, but accurate measurement of electromagnetic compatibility the three vector antenna must.

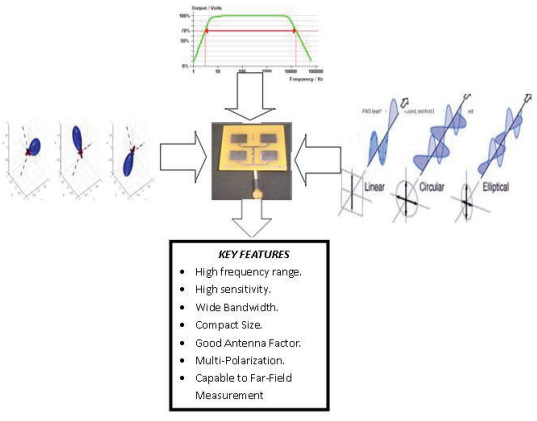


Fig. 13. *Proposed Antenna Features*

Hence, we have proposed three axial antenna geometry for calculating the average 3-axisal vector field for EMC measurement.

REFERENCES

1. Suk-Hwan Choi and Dong-Seok Kim, "*Three Axis Isotropic Field Strength Measuring Antenna*,"the journal of korean institute of electromagnetic engineering and science. 2014 sep.; 25(9), 879-885
2. Benjamin Loader and Martin Alexander, "*development of optical electric field sensor for EMC measurement*,"www.ieice.org/ proceedings/EMC14.
3. IEEE standard dictionary of electrical and electronics terms,ANS/IEEE Std 100-1977, 2nd ed.,Jay, ed, New yorl:wiley 1977.
4. J. Leutcher, V. Stekly, E. Blasch, "*Invetsigation of Avionics Power Switch Loading versus Aircraft Electromagnetic Compatibility*," IEEE AESS Magazine, Vol. 30, Issue 9, pp. 24-34, Sept. 2015
5. Mark T.MA,and Motohisa Kanda," *A review of electromagnetic compatibility/interference measurement methodologies*," procedding of the IEEE,Volume:73,Issue:3, page 388-411,yeat-1985.
6. Mark T.MA, Motohisa Kanda and Myron L.Crawford, "*A review of Electromagnetic compatibility/Interference Measurement Methodologies*," Proceeding of the IEEE, Vol 73, No3, March1985.
7. Tim Williams, "*What to look for in an EMC antenna*," Compliance Engineering European.
8. M J Alexander, "*Calibration and use of EMC antennas*, "Measurement good practice guide No. 4 from the National Physical Laboratory
9. E. L. Bronaugh, J. D. M. Osburn, " *Whole-vehicle radiated EMI immunity tests in automotive EMC: establishing and calibrating the test field*," Electromagnetic Compatibility, 1992., Eighth International Conference on, 21-24 Sep 1992, Edinburgh, pp. 39-42, 1992.
10. Narang, N., Dubey, S.K., Negi, P.S. et al., "*A Coplanar Microstrip ntenna as a Dosimetric E-field Probe for GSM Frequencies*,"MAPAN journal of metrology society of india,pp 1-5,jan 2017.
11. Mohammed Siddeq and Dr M. Satyanarayana, " *A new printed monopole antenna as EMI sensor*," International Journal of Engineering Research & Technology (IJERT), ISSN: 2278-0181, Vol. 2 Issue 10, October – 2013.
12. Zdeněk Kubík , Denys Nikolayev et el, "*Optimization of electrical properties of parallel plate antenna for EMC testing*," Journal of omputational and Applied Mathematics 270, May 2014.
13. Amit Kumar Srivastava," *Analysis of Circular Patch Antenna as an Electromagnetic Interference Sensor*," Conference on Electrical, Electronics and Computer Science, March 2012 IEEE Students.
14. S. Capdevila, "*Design of a Small MST Probe for EM-Field Measurements and Sensing Applications*," Antennas and Propagation Society International Symposium (APSURSI), 2012 IEEE, 8-14 July 2012.
15. Surbhi Mittal, "*Active Probes for Creating H-Field Probes for Flat Frequency Response*," Electromagnetic Compatibility, 2009. EMC 2009. IEEE International Symposium , 17-21 Aug. 2009.
16. M. F. P. Tartaglia1 et el, "*A Metamaterial-based Probe for EMC Measurements*," PIERS Proceedings, Guangzhou, China, August 25{28, 2014.
17. H. Lv, X. S. Xia, Y. L. Yu, Z. X. Hua, "*An Indoor Electromagnetic Testing Antenna Design with Better Standing Wave Ratio and Gain*," International Conference on Artificial Intelligence and Industrial Engineering (AIIE 2015), Published by Atlantis Press,May 2015.
18. Xingyu Zhang, Amir Hosseini,, "*Integrated Photonic Electromagnetic Field Sensor Based on Broadband Bowtie Antenna Coupled Silicon Organic Hybrid Modulator*," Journal of Lightwave Technology, Volume: 32, Issue: 20, oct.15, 15 2014.
19. F. H.Wee, F. Maleke et el., "*electromagnetic wave detection based on multiband antenna design*, ARPN Journal of Engineering and Applied Sciences, ISSN 1819-6608, VOL. 11, NO. 8, APRIL 2016.
20. Nisha Gupta and Md. Anjarul Haque, "*Design of Printed Log Periodic EMI Sensor*," international journal of microwave and optical technology,vol.4 no.4 july 2009.
21. Fenghan Lin et el ," *0.7–20-GHz Dual-Polarized Bilateral Tapered Slot Antenna for EMC Measurements*," IEEE transactions on electromagnetic compatibility, Page -1271 - 1275, Volume: 56 Issue: 6,Dec 2014.

Authors' contacts
Organization: Faculty of Telecommunication,
Technical University of Sofia,
Address: Post Code 1000, boulevard Kliment hridski 8,Sofia, Bulgaria
E-mail: Sarang.p86@gmail.com,
pjpetkov@tu-sofia.bg,
bbonev@tu-sofia.bg.

© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 23, 2017

DEFINING THE COEFFICIENT OF ADDITIONAL LOSSES BETWEEN BUSBARS EXPERIMENTALLY AND NUMERICALLY

IVAN HADZHIEV, DIAN MALAMOV, VASIL SPASOV

Abstract: *Experimental tests and numerical study of the influence of both skin and proximity effects on parallel current-carrying busbars have been carried out and the results have been described in this paper. The experiments have been conducted by means of an experimental installation, specially developed for the purpose. The numerical study has been performed with the help of computer models, developed in the software product Comsol. A comparison has been made between the obtained experimental and numerical data of the coefficient of additional losses between current-carrying busbars.*

Key words: *coefficient of additional losses, current-carrying busbars, FEM*

1. Introduction

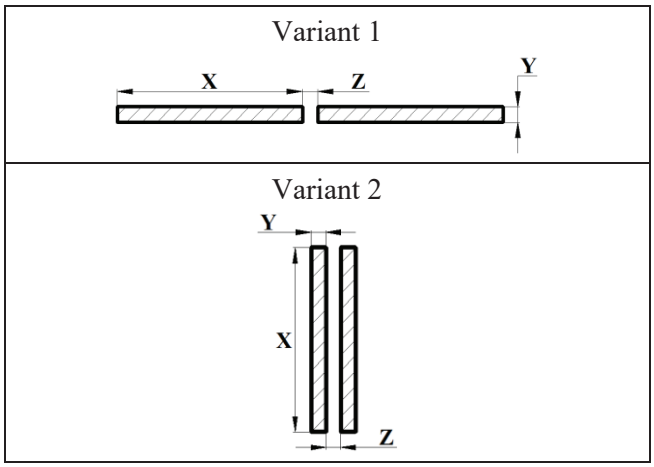
It is known that when alternative current flows through current-carrying elements, additional losses occur due to skin and proximity effects [1]. These losses depend in a complex way on the dimensions of the current-carrying elements, on the frequency of the current, on the properties of the material, etc. It makes their direct calculation quite difficult. The skin effect is described by differential equations [2], [3], [4]. Analytical solution is only possible for simpler cases, such as the ones, presented in [5]. Papers [3] and [4] consider the skin effect in a massive busbar and give an approximate analytical solution. An electric device has been shown in [2] for measuring the flux density, while in [6] the skin effect has been discussed and the magnetic field in the conductors has been calculated by the finite element method. The numerical results have been presented in a graphical format and compared with analytical ones. Paper [7] presents a study of the skin and proximity effects in case of using U-shaped busbars in terms of their dimensions and location, as well as vs. the frequency.

This paper describes both experimental tests and a numerical study of the coefficient of additional losses between current-carrying busbars. A comparison between the obtained experimental and numerical results has been made.

2. Busbar location variants

Table 1 presents the variants of location of the busbars, for which the coefficient of additional losses has been studied.

Table 1. *Types of location of the investigated busbars*



3. Studying the coefficient of additional losses between current-carrying busbars experimentally

A special experimental installation has been developed for the purpose of studying the coefficient of additional losses experimentally, as shown in Fig. 1.

 ISSN 1310 - 8271

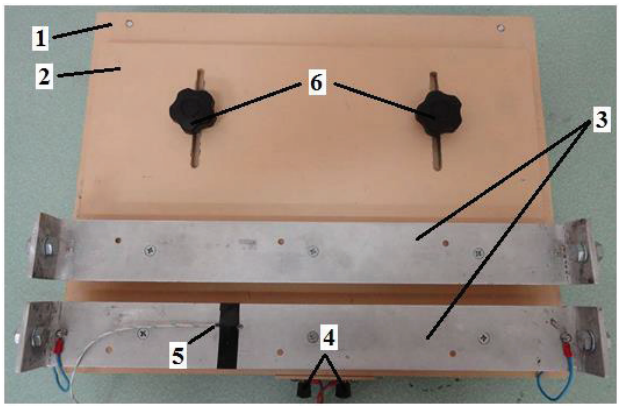


Fig. 1. *Experimental installation for studying the coefficient of additional losses:* ***1*** *- stationary part;* ***2*** *- movable part;* ***3*** *- aluminium busbars with dimensions [60x6]mm;* ***4*** *- terminals for measuring the voltage drop;* ***5*** *- K type thermocouple;* ***6*** *– fixer.*

The electrical circuit of the experimental installation is shown in Fig. 2.

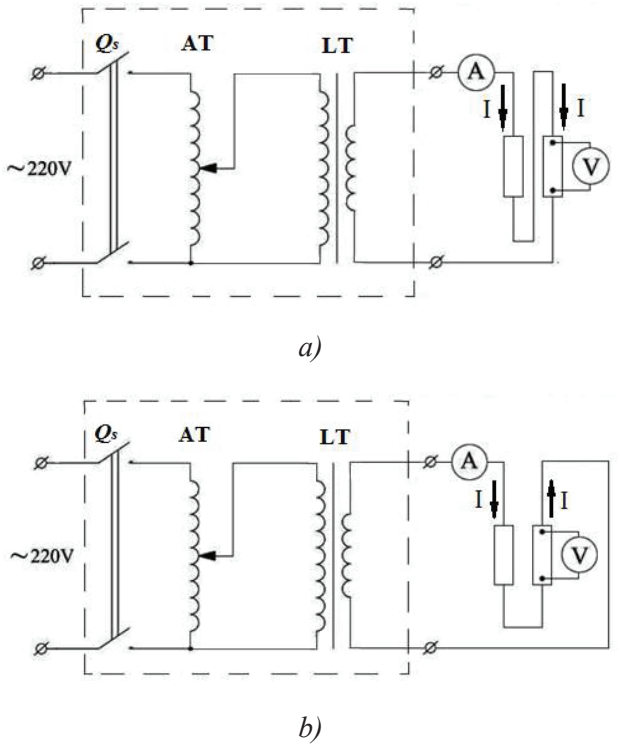


Fig. 2. *Electric circuit of the experimental installation for studying of the coefficient of additional losses at: a) – currents in the same direction; b) – currents in opposite directions; Q_s – circuit breaker; AT – autotransformer; LT – load transformer; A – ammeter; V – voltmeter.*

The coefficient of additional losses is defined according to the formula:

$$k = \frac{P_{\sim}}{P_{=}} = \frac{I^2.R_{\sim}}{I^2.R_{=}} = \frac{R_{\sim}}{R_{=}}, \qquad (1)$$

where: $P_{\sim}$ are the losses per unit of length at alternating current with r.m.s. value I; $P_{=}$ are the losses per unit of length at direct current I of the same magnitude; $R_{\sim}$ and $R_{=}$ are the resistances of the busbars at alternating and direct current respectively.

The resistances of the busbars $R_{\sim}$ and $R_{=}$ are defined according to the following formulae:

$$R_{\sim} = \frac{\Delta U}{I_{\sim}}; \qquad (2)$$

$$R_{=} = \frac{\Delta U}{I_{=}}. \qquad (3)$$

Here ΔU is the voltage drop of the busbars.

3.1. Experimental results for the coefficient of additional losses

The coefficient of additional losses has been defined by means of the experimental installation described above. Fig. 3 and Fig. 4 illustrate the dependence of the coefficient of additional losses on the distance between the busbars for variant 1 in case of currents in the same direction and in opposite directions.

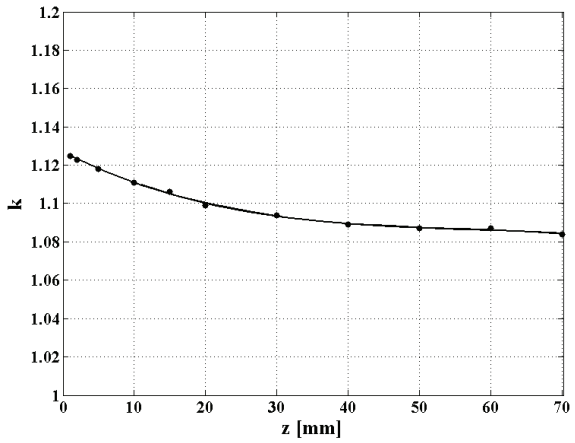


Fig. 3. *Dependence of the coefficient of additional losses on the distance between the busbars for variant 1 and currents in the same direction.*

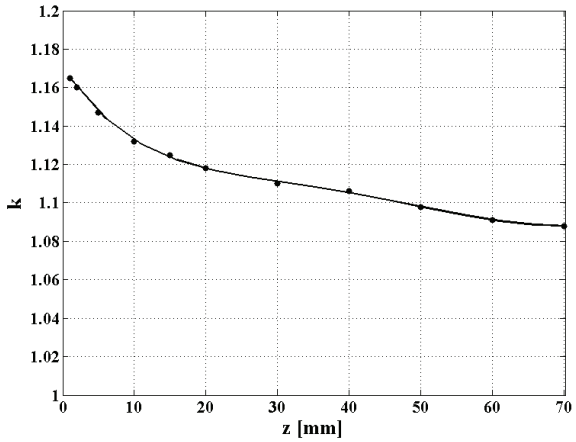


Fig. 4. *Dependence of the coefficient of additional losses on the distance between the busbars for variant 1 and currents in opposite directions.*

Fig. 5 and Fig. 6 show the dependence of the coefficient of additional losses on the distance between the busbars for variant 2 in case of currents in the same direction and opposite directions.

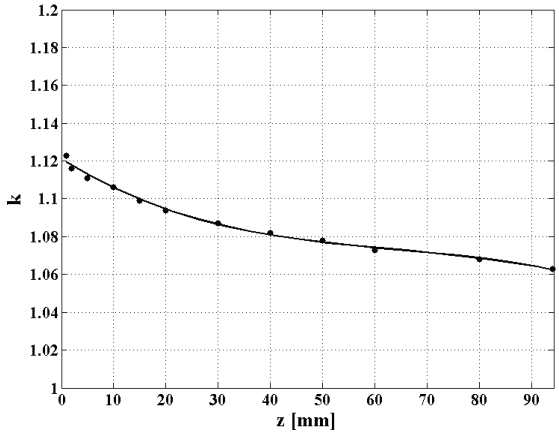


Fig. 5. *Dependence of the coefficient of additional losses on the distance between the busbars for variant 2 and currents in the same direction.*

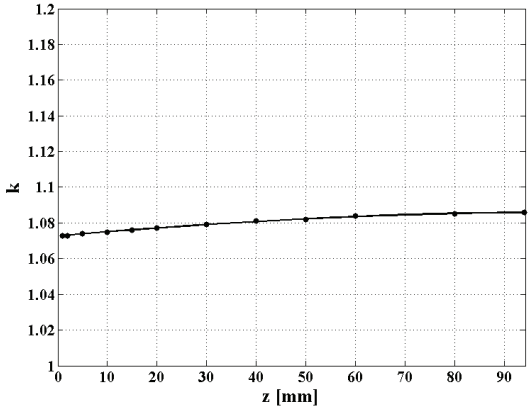


Fig. 6. *Dependence of the coefficient of additional losses on the distance between the busbars for variant 2 and currents in opposite directions.*

4. Numerical study of the coefficient of additional losses between current-carrying busbars

4.1. Mathematical model

The mathematical model of a quasi-stationary electromagnetic field is described by the Helmholtz equation in a complex type [8]:

$$\nabla^2 \dot{\mathbf{A}} - j\omega\mu\sigma \dot{\mathbf{A}} = -\mu \dot{\mathbf{J}}_{\mathbf{e}}, \tag{4}$$

where: $\mathbf{A}$ is the magnetic vector potential; ω is the angular frequency; μ is the magnetic permeability; σ is the specific electrical conductivity; $\mathbf{J_e}$ is the current density of external sources. The second member of the equation reflects the eddy currents.

The solution of the electromagnetic problem is given at a predetermined boundary condition along the border of the buffer zone of the type:

$$\mathbf{n}\mathrm{x}\mathbf{A} = 0. \tag{5}$$

The losses per unit of length at alternating current are obtained from the electromagnetic field problem solution according to the formula:

$$P_{\sim} = \iint_S \mathbf{J}.\mathbf{E}dS. \tag{6}$$

The losses per unit of length at direct current are calculated by the formula:

$$P_{=} = \frac{I^2}{\sigma S}, \tag{7}$$

where I is the current through the busbar; S is the cross-section of the busbar.

The coefficient of additional losses is defined by the formula:

$$k = \frac{P_{\sim}}{P_{=}}. \tag{8}$$

4.2. Analysis by the finite element method

Numerical models for defining the characteristics of the electromagnetic field for the variants presented in Table 1 at frequency of 50 Hz have been created by the software product COMSOL [9].

Fig. 7 shows the mesh of finite elements.

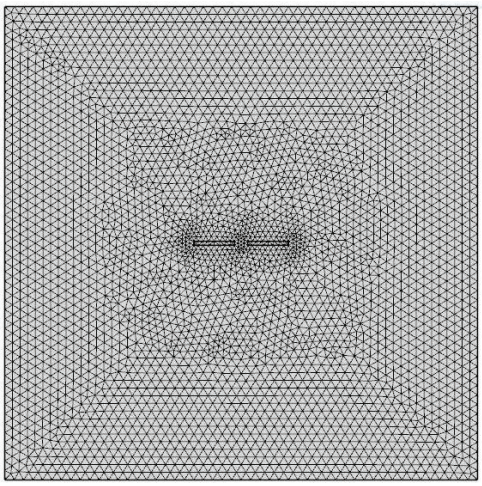


Fig. 7. *Mesh of finite elements.*

Fig. 8, Fig. 9, Fig. 10 and Fig. 11 illustrate the distribution of the magnetic flux density for variant 1 and variant 2 at dimensions x = 60mm, y = 6mm and z = 20mm.

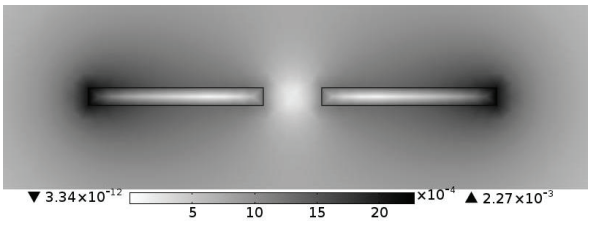


Fig. 8. *Distribution of the magnetic flux density B[T] for variant 1 and currents in the same direction.*

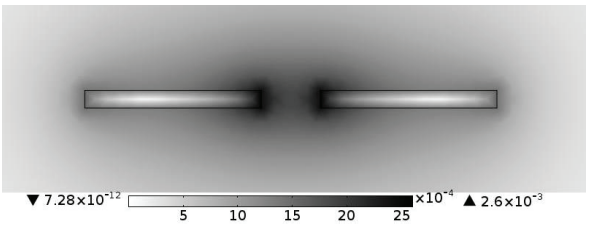


Fig. 9. *Distribution of the magnetic flux density B[T] for variant 1 and currents in opposite directions.*

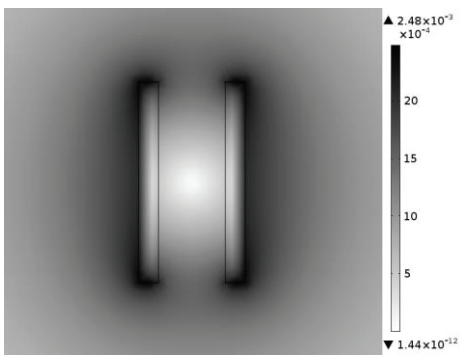


Fig. 10. *Distribution of the magnetic flux density B[T] for variant 2 and currents in the same direction.*

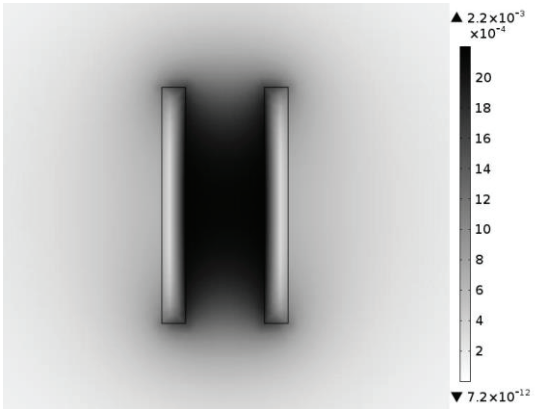


Fig. 11. *Distribution of the magnetic flux density B[T] for variant 2 and currents in opposite directions.*

The obtained distribution of the current density in the busbars for variant 1 and variant 2 at dimensions x = 60mm, y = 6mm and z = 20mm is shown in Fig. 12 and Fig. 13.

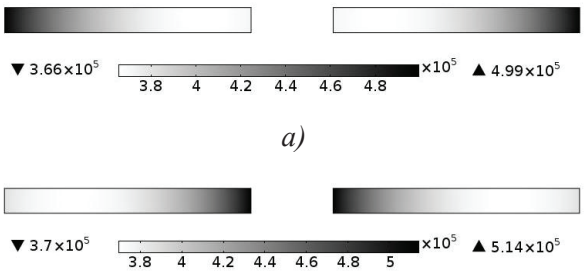


Fig. 12. *Distribution of current density [A/m^2] for variant 1 at: a) - currents in the same direction; b) - currents in opposite directions.*

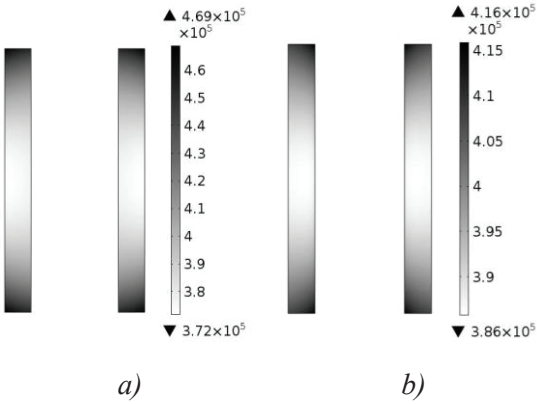


Fig. 13. *Distribution of current density [A/m^2] for variant 2 at: a) - currents in the same direction; b) - currents in opposite directions.*

The obtained distribution of the specific losses in the busbars for the different variants at dimensions x = 60mm, y = 6mm and z = 20mm is shown in Fig. 14 and Fig. 15.

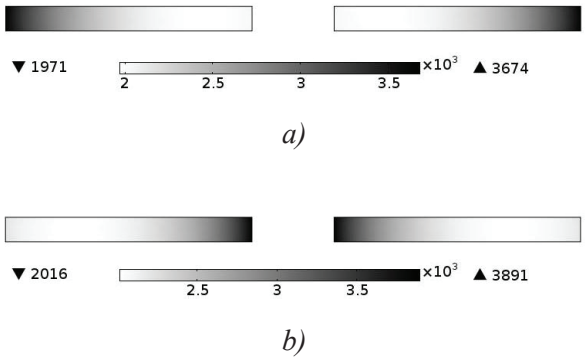


Fig. 14. *Distribution of the specific losses [W/m³] for variant 1 at: a) - currents in the same direction; b) - currents in opposite directions.*

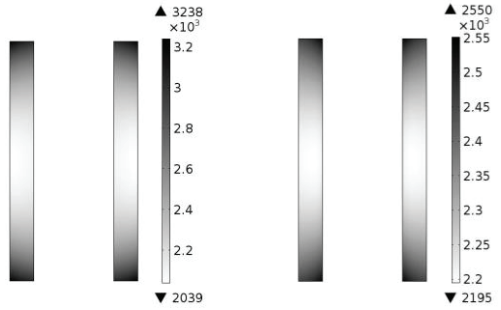


Fig. 15. *Distribution of the specific losses [W/m³] for variant 2 at: a) - currents in the same direction; b) - currents in opposite directions.*

5. Comparison between the obtained numerical and experimental results

Fig. 16, Fig. 17, Fig. 18, and Fig. 19 show the graphic dependencies of the coefficient of additional losses on the distance between the busbars obtained from both the experimental tests (curve 1) and the numerical study (curve 2).

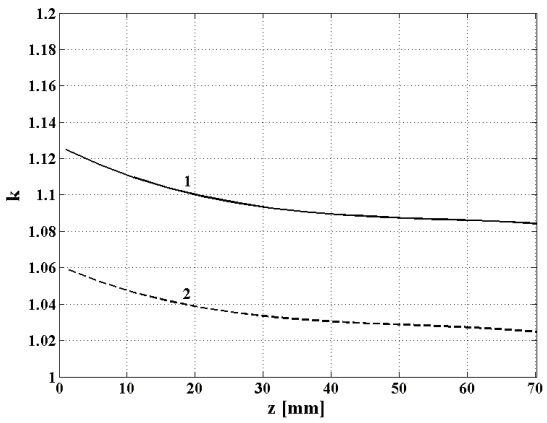


Fig. 16. *Dependence of the coefficient of additional losses on the distance between the busbars for variant 1 and currents in the same direction.*

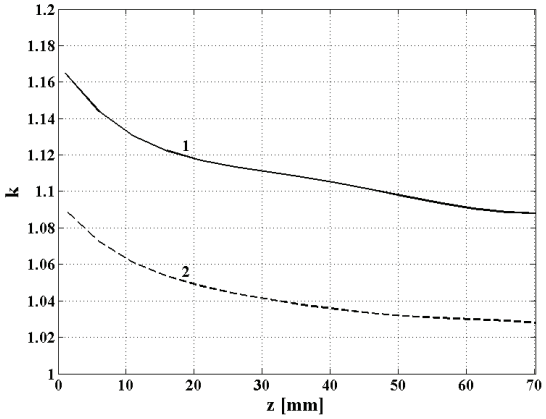


Fig. 17. *Dependence of the coefficient of additional losses on the distance between the busbars for variant 1 and currents in opposite directions.*

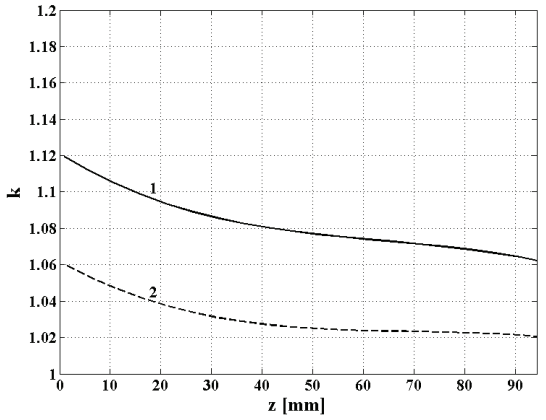


Fig. 18. *Dependence of the coefficient of additional losses on the distance between the busbars for variant 2 and currents in the same direction.*

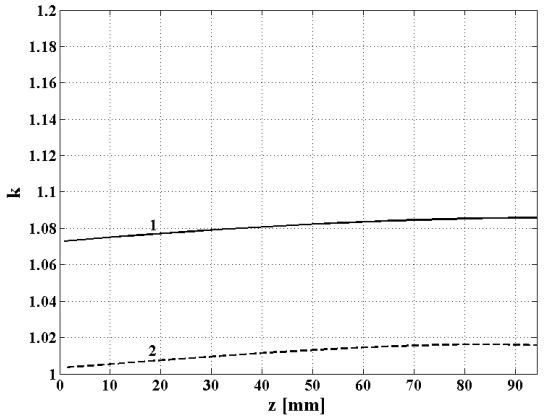


Fig. 19. *Dependence of the coefficient of additional losses on the distance between the busbars for variant 2 and currents in opposite directions.*

6. Conclusion

Based on the results from the numerical study and the experimental tests carried out, the following conclusions can be made:

- the pattern of change of the coefficient of additional losses considerably depends on the distance between the busbars;

- the coefficient of additional losses for variant 1 decreases with the increase in the distance between the busbars, regardless of the direction of the current;

- the coefficient of additional losses for variant 2 decreases with the increase in the distance between the busbars at currents in the same direction, while at currents in different directions the situation is just the opposite;

- the coefficient of additional losses has the highest value when the busbars are positioned as in variant 1 and in case of currents in opposite directions.

The difference between the obtained numerical and experimental results is 7%. It allows for studying the coefficient of additional losses numerically and obtaining arrays of data in order to be able to define this coefficient directly.

REFERENCES

1. Alexandrov, A. (2002). *Electrical apparatuses - part I (Electromagnetic and thermal problems),* Heron press, Sofia (in Bulgarian).
2. Kosek, M., Truhlar, M., Richter, A. (2011). Skin-effect in massive conductors at technical frequencies, *Przeglad Elektrotechniczny (Electrical Review),* ISSN 0033-2097, R. 87 NR 5.
3. Kosek, M., Truhlar, M., Richter, A. (2012). Skin-effect in conductor of rectangular cross-section – approximate solution, *Przeglad Elektrotechniczny (Electrical Review),* ISSN 0033-2097, R. 88 NR 7a.
4. Gerling, D., Approximate analytical calculation of the skin effect in rectangular wires, *University of Federal Defense Munich, 85579 Neubiberg,* Germany.
5. Hanka, L. (1975). *Theory of electromagnetic field*, SNTL, Prague. (in Czech).
6. Greconici, M., Madescu, G., Mot, M. (2010). Skin effect analysis in a free space conductor, *Elec. Energ*. vol. 23, No. 2, pp. 207-215, Serbia.
7. Lažetic, K., Prša, M., Đuric, N. (2009). Proximity effect against skin effect in two coupled U–shaped busbars, *Elec. Energ*. vol. 22, no. 3, pp. 293-303, Serbia.
8. Yatchev I., Marinova, I. (2011). *Numerical methods and modeling of circuits and fields*, *Part one*, Sofia. (in Bulgarian).
9. *COMSOL Version 4.2 User'guide*. (2011).

Assistant Prof. Ivan Hadzhiev, Ph.D.
Department of Electrical Engineering
E-mail: hadzhiev_tu@abv.bg

Assoc. Prof. Dian Malamov, Ph.D.
Department of Electrical Engineering
E-mail: deanmalamov@abv.bg

Assoc. Prof. Vasil Spasov, Ph.D.
Department of Electrical Engineering
E-mail: vasilspasov@yahoo.com

Technical University - Sofia,
Branch Plovdiv
25 Tsanko Dyustabanov Str.
4000 Plovdiv, Bulgaria
Telephone number: +359 32 659686

© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 23, 2017

HIGH SENSITIVE LONG DISTANCE SCANNING FIBER-OPTICS LASER SENSOR SYSTEM

MARGARITA DENEVA, MARIN NENCHEV

Abstract: *The principle of the proposal is to use laser with long (25 m and more) resonator, composed by series of short free space parts, coupled by long optical fibers. We combine the laser with such resonator construction with spectral characteristics of the generation, which are appropriate for intra-cavity laser spectroscopy registration of expected atoms and molecules (including atmospheric non-desired pollutions – SO_2, NO_2, Na etc.). The registration on the base of the intra-cavity laser spectroscopy (ILS) method makes the laser developed extremely sensitive for atoms or molecules presence in the investigated spaces of the resonator free-space region, in principle down to single atoms. The laser spectral control includes also original solution for producing suitable controlled multi-bands spectrum for simultaneous monitoring in different spectral bands. The laser system proposed is very convenient for controlling the air purity in number of rooms in given laboratory, enterprises, where the presences of the non-desired pollutants are expected. We treat theoretically the system action and present the experimental laboratory test on the example of registration of Sodium atoms presence in the air.*

Key words: *intra-cavity spectroscopy, special laser system, long resonator, multi registration parts, theoretical treatment, experiment with Sodium atoms*

1. Introduction

Specialized laser systems are established at present as important tools for distant monitoring of air pollutions, in ecology, in industrial enterprises producing different type chemical materials, in military production, etc. [e.g. 1-4]. Here we propose high-sensitive long-distance scanning fiber optics laser sensor system for such ecological control, especially for monitoring of close working places (number of separated rooms). The principle of the proposal is to use laser with long (50 m and more) resonator, composed by series of short free space parts in each monitored place (room), coupled by long optical fibers and combined by the laser with such resonator construction with spectral characteristics of the generation that are appropriate for intra-cavity laser spectroscopy registration of expected atoms and molecules (including atmospheric non-desired pollutions – SO_2, NO_2, Na, etc.). The registration on the base of the intra-cavity laser spectroscopy method makes the laser developed extremely sensitive for atoms or molecules presence in investigated free-space resonator region, in principle down to single atoms. Except the cheapest engineering realization, the important advantage is the completely elimination of possibility for spark formation that is of essential importance for the military production specialized enterprises. We present also the experimental laboratory test of developed such system by registration of Sodium atoms presence in the air in different places.

2. Details of the principle of the proposed system

2.1. General description

The laser source, which is suitable for intra-cavity spectroscopy, is with homogeneously broadened amplification and with controlled wide-band generation. As an additional advantage in our case, the spectrum is tunable in the laser gain. The general condition for application of the laser source in intra-cavity laser spectroscopy is that the spectral band of the emission must be smooth, structure-less. The principle of the extremely high sensitive intra-cavity registration consists in this that the spectral absorption lines of the investigated atoms or molecules presented in the laser cavity make holes in the smooth laser spectrum.

The drastically increasing of the sensitivity in comparison with this one in the case of the extra-cavity registration is result of strong competition between the generation at different lines in the wide

 ISSN 1310 - 8271

spectrum in the homogeneous broadened active laser medium [5,6]. The introduction of smallest losses for one line leads to strong decreasing or suppressing of the lasing at this line. The additional important factor to increase the competition and respectively - the sensibility is to use temporally long generation – long pump pulses (flashlamp-pumping with long exciting pulses or cw diode pumping) or cw operation. The condition to be avoided the typical structure in the wide-band laser spectrum is that in the resonator are precluded any parallel reflective surfaces (including low reflection), forming the resonance structure of Fabry-Perot type. As very convenient laser for intra-cavity laser spectroscopy, which is applied here, following our earlier work [7], is a waveguide type laser that naturally enables to exclude the parallel surfaces into the cavity and to obtain smooth spectrum.

The schematic of the system proposed, taking into account the given general description, is presented in Fig.1.

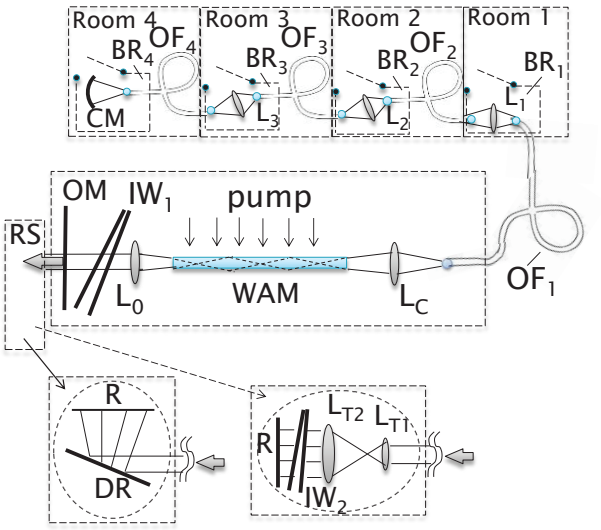


Fig. 1. *The schematic of the proposed system*

In the proposed system, the laser is composed by a capillary-tubular cuvette filled with liquid active medium (WAM) or other type waveguide realization active medium; pump source, and the special resonator. In one side of the WAM the resonator reflector is formed by output coupling using wedged mirror OM (R= 0.96), spectral selective and tuned Interference Wedge IW [8] and focusing lens L_o. The last transforms the emitted light from the output of the waveguide AM. In the other side, the resonator reflector is formed by series, numbered with *i*, (*i* = 1,2,3,4), of boxes for registration BR_i connected via long optical fibers (OF_i) that are coupled each other with the corresponding lens (L_i *; all lenses in the system are AR coated*) and at the end of fiber OF_4 with retro-reflected output by the end-concave mirror CM. The each box is disposed in the corresponding room – $Room_i$. The registration subsystem RS consists of spectrum analyzer, which in one variant is based on diffraction grating DR in near grazing incident angle, where the resolution is highest, and photograph-plate R with corresponding construction arrangement (first inset in Fig.1, details are not shown) or realized on the base of CCD matrix. In equal manner, the spectrum can be registered after expanding telescope (L_{T1}-L_{T2}) and using convenient Interference Wedge IW_2 [8, 9] second inset in Fig.1. Below, as a point of the presentation, we give general detailed description of the characteristics of the system, which leads to the reality of its expected advantageous practical functioning.

The laser is the waveguide flashlamp pumped dye laser (in other realization can also be suitable spectrally semiconductor laser and flashlamp pumped Ti:Sapphire laser). The dye active solution – 3.10^{-4} mol/l Rh6G in ethanol, flows through the cell that is cylindrical glass capillary-tube with internal diameter of 1 mm, external diameter of 10 mm and length of 100 mm. The pumping is by flashlamp in elliptical type reflector thus assuring the exciting pump energy of 10 J in 15 μs pulse (rise front 2 μs, plato 4 μs and fall front 7 μs - at half maximum). The focal length of the lens L_o is 10 cm and for the lens L_c it is 5 cm, respectively; the radius of end-concave mirror CM is 5 cm and its reflectivity is 99 %. The IW_1 and IW_2 are composed by dielectric layers and both are with thickness of 3 μm, wedge angle $5x10^{-5}$ rad and reflectivity of the composed mirrors for IW_1 – 0.8 and for IW_2 - 0.3. The mechanical construction of each box (20x20x25 cm) contains a door that can assure its prevention of the registration, i.e. the corresponding room to be out of control. The aim of such isolation is to assure possibility to monitor each room separately. The total length of the high quality connected optical fiber is with β = 7 dB/km (multimode, 125/250 μm) and the total length of the fibers is 15 m. Accepting that the loss of any connection (by lenses L_i) is (0.2-0.3) dB, the ratio (returned power/incident power) for total fiber system can be evaluated to be of order of (0.55 – 0.6), or the fiber part can be accepted as equivalent mirror with reflection of ~0.6.

2.2. Theoretical description of the system action.

A discussion of the strongly increased sensitivity obtained by applying intra-cavity spectroscopy technique will be presented.

In the consideration we will use the values of laser system discussed in the previous point. The

differential rate equations system [5], adapted for description of the case under investigation described Rh6G dye solution and pumping with the system parameters described above, is:

$$\begin{cases} \dfrac{dN_2}{dt} = R_p(t) - \left(\sum_i B_i \cdot q_i \right) \cdot N_2 - \dfrac{N_2}{\tau} \\ \dfrac{dq_i}{dt} = Va \cdot B_i \cdot q_i \cdot N_2 - \dfrac{q_i}{\tau_{c_i}} \end{cases}$$

with $P_{out_i}(t) = (\gamma_1 \cdot c/2L') \cdot h\nu \cdot q_i(t)$

Here q_i are the generated photon number for the corresponding wavelength in the considered laser spectrum; P_{out_i} is the corresponding output power, which integration in the time (from 0 to the length pulse) gives the output energy. In the systems, with N_2 is noted the population of the upper laser level per unit volume in the WAM. The term $B_i = \left(\sigma_e^i \cdot l \cdot c\right) / \left(V_a \cdot L'\right)$ [s^{-1}], where σ_e^i is the emission cross-sections for the given wavelength in the laser spectrum; V_a is the working volume (0.05 cm^3); c = 3x10^{10} cm/s is the light velocity; $L' = L + (n-1) \cdot l$ - the optical length of resonator, where l=*10 cm* is the length of the active medium, n is the refractive index and L is 25 m. The time term τ of 3 ns is the lifetime of the upper laser level for the Rh6G WAM. The dumping time of a photon in the resonator is $\tau_{c_i} = L' / (c \cdot \gamma_i)$, where γ_i [5] describes the loss into the resonator for the wavelength in the considered laser gain spectrum (depending on the reflectivity of the combined end resonator mirror OM, IW_1 and Lo). The system was solved numerically by Runge-Kutta-4 method. From the solution we obtain q_i(t) and the respective output power for each wavelength in the spectrum; γ_1 characterises the output of the laser resonator. The calculations are prepared for the spectral range (588 – 591) nm, in which the Sodium D-lines are also comprised.

In the example considered we will compare the extra-cavity and intra-cavity registration of presence of Sodium vapors outside and inside of the described special resonator. We will study the behavior of the formed by the Sodium atoms absorption holes in the spectrum at the Sodium D-lines (doublet D_2 =589.0 nm and D_1=589.6 nm). Let's we accept that the same volume of Sodium atoms with the same concentration and length – path of the laser beam through the volumes, is disposed one time outside the resonator and after this - inside the resonator. Let's the Sodium atom concentration provides absorption losses for single pass at D_2 line of 1% (respectively – 0.5 % for D_1 line). Using the given above differential equation system, we obtain the spectra of laser light for the two cases.

For the extra-resonator case (simple passage) we obtain practically non-observable (or specially detected holes due to of 1 and 0.5 % absorptions at the lines). The calculated output laser spectrum for this case is shown in Fig.2, in the inset the formed hole by the absorption at D_2 line in expanded scale is given.

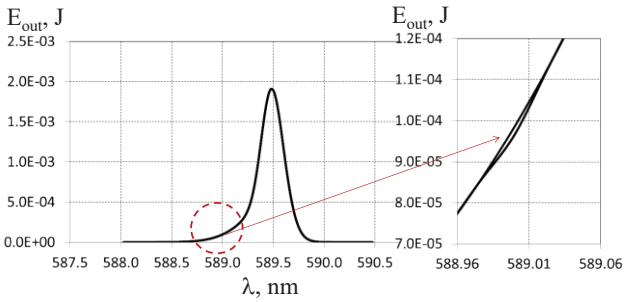


Fig.2. *The calculated output laser spectrum for the case of registration by single passage through the Sodium D-lines. In the inset is shown the formed hole by the absorption at D_2 - line (the figure is expanded).*

The calculations for the intra-cavity case of emission spectrum with the same volume of Sodium atoms, gives the picture, shown in Fig.3. The drastically increasing of the holes at the D-absorption lines can be seen. This confirms our expectations for the possibility for high sensitive registration of desired atoms presence by application of the proposed laser system and for monitoring of such presence in few rooms.

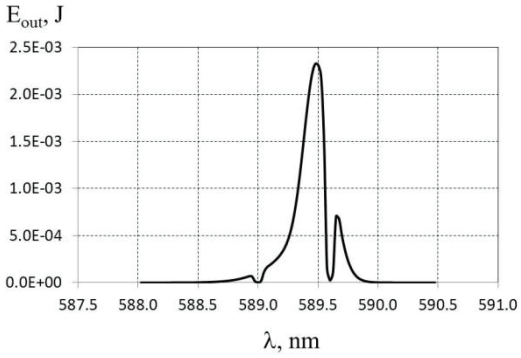


Fig.3. *The calculated output laser spectrum for the case of intra-cavity registration by single passage through the Sodium D-lines.*

Let us evaluate the concentration of the Sodium atoms that leads to 1% decreasing of the transmission – i.e. the case considered above. Following Ref. [9] we can write:

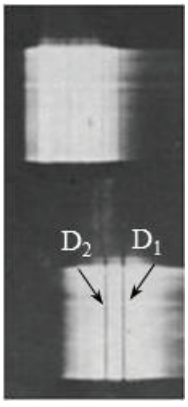
$$I/I_o = \exp(-\sigma \cdot N \cdot l_a)$$

where I_0 is the incident intensity, I – the intensity of the transmitted beam, both I_0 and I for wavelength of the absorption line, $\sigma = 6.10^{-12}$ cm^2 is the absorption cross-section for the D_2 line and l_a=5 cm is the path of the beam trough the volume with Sodium atoms. For these conditions, after a simple arithmetic, we obtain that N is of order of $(10^7 \div 10^8)$ cm^{-3}. Thus our system assures reliable registration of presence of Sodium atoms with such concentration.

3. Experimental test

In the given in the work practical test-realization (example) that clarifies the system action, we use only two structures - BR_1 and BR_4, excluding BR_2 and BR_3. Both optical fibers OF_1 and OF_4 were with length of 1 m each and damping factor β = 300 dB/km and four connections with the lenses (equivalent mirror with reflection ~ 60 %). The other used characteristics of the elements and of the construction are the given above. The volume of Sodium atoms is formed by burning of three standard matches with the flame outside the cavity near the output mirror OM (extra-cavity spectroscopy) or inside the cavity - in the space between the cell output and lens Lc.

The typical real spectra of the emission of laser system with spectrum of generation tuned at Sodium D-lines are given in Fig.4. The top spectrogram (Fig.4a) is for the case when the Sodium atoms are formed outside the cavity and the laser light passes through the space with the Sodium atoms (absorption outside the cavity). The bottom spectrogram (Fig.4b) is for the case of presence of the Sodium atoms inside the complex laser cavity.



(a)

(b)

Fig.4. *Typical real spectrums of the laser system emission with spectrum of generation tuned at Sodium D-lines. The top spectrogram is for the absorption outside the cavity and the bottom - for the Sodium atoms inside the complex laser cavity. The strong increasing of the sensitivity is evident.*

In Fig.5 are shown the corresponding trace-curves for the two spectrograms – left for the given in Fig.4(a) and right – for the spectrogram in Fig.4(b).

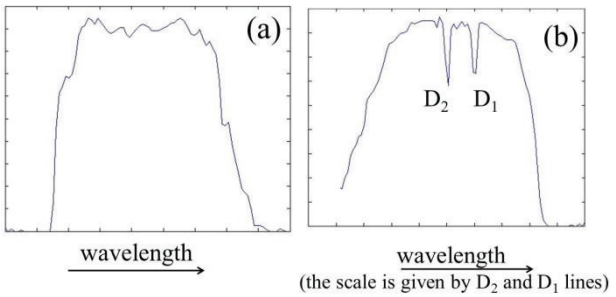


Fig.5. *The corresponding trace-curves for the two spectrograms in Fig.4 – left for case (a) – Sodium atoms are outside the cavity (practically the absorption D-lines are not observed) and right for the case (b)– the atoms are inside the cavity (real observation of the D-lines). The scale is given by the holes of D_2 and D_1 lines marked on the figure.*

Following the theoretical considerations, given above, and the obtained spectrograms, we can conclude that practical laboratory model of the system proposed is able for real registration of Sodium atoms with concentration ~ 10^7 cm^{-3}. This concentration can be accepted to correspond to the obtained one in the matches flame.

4. Discussions

Following the results from the theoretical analysis and from the carried out experiment we make in evidence the feasibility of the proposed system and its advantages for specific ecologic monitoring and survey for dangerous pollutions, accompanied the industrial production. The sensibility of such intra-cavity spectroscopy based technique is more than of order of magnitude higher than the standard extra-cavity laser technique. Very suitable laser sources are these one that avoid the flat parallel reflected (as well as partially) surfaces, what in high degree is natural property of the waveguide laser sources. With application of modern low-losses optical fibers and suitably treated elements (without parallel surfaces, there is no parallel passage of the light rays) the proposed system can be organized as very suitable technical instrumentation for extremely sensible detection of desired pollutant atoms. Note that in particular case of many rooms disposed at both sides of building corridor, the optical fiber part of the resonator can be adapted to form a ring type laser resonator. This will permit to use more effectively fiber parts of the resonator.

5. Conclusion

In the work is proposed the principle of the special laser system with an optical fiber based long (20 m and more) resonator, composed by series of short free space parts, coupled by long optical fibers. We combine the laser with such resonator construction with spectral characteristics of the generation that are appropriate for intra-cavity laser spectroscopy registration of expected atoms and molecules. As it is shown, the registration on the base of the intra-cavity laser spectroscopy (ILS) method makes the developed laser extremely sensitive for atoms or molecule presence in investigated spaces of the resonator free-space region. The laser spectral control includes also the original solution with Interference Wedge for producing suitable controlled multi-bands tunable spectrum (for monitoring in different spectral bands). The theoretical treatment of the system action and the experimental laboratory test for registration of very low concentration presence of Sodium atoms in the air confirm the feasibility of the system and its expected advantages. The laser system proposed is very convenient for controlling the air purity in number rooms of given laboratory, enterprises, where the presence of the non-desired pollutants is expected.

Acknowledgement: The work is supported by DN 08/13 (2016) Contract with NSF-BG.

REFERENCES

1. W. Demtröder *Laser spectroscopy: basic concept and instrumentation*, (2003) 3rd ed. Springer, Germany and the literature therein.
2. P. Sudhakar, P. Kalavathi, D. Ramakrishna Rao, M. Satyanarayana (2014). Design of Laser Based Monitoring Systems for Compliance Management of Odorous and Hazardous Air Pollutants in Selected Chemical Industrial Estates at Hyderabad, India. *Remote Sensing and Spatial Information Sciences*, volume XL-8, *and the literature therein* .
3. E. D. Hinkley, P. L. Kelley. Detection of Air Pollutants with Tunable Diode Lasers**,** *Science* 19 Feb 1971: Vol. 171, Issue 3972, pp. 635-639, DOI: 10.1126/science.171.3972.635
4. Бураков В.С., Исаевич А. В., Мисаков, П.Я., Ненчев М.Н., Патриков Т.В., Пашов А.Е., Пешев З.И., Разширение аналитических возможностей внутрирезонаторных лазерных спектрометров, *Приб. и техн. Эксперимента* - *ПТЭ*, No 5, 1994, 150-156, (RUSSIA, trad. and parallel ed. in USA – *Prib. Tech. Eksperimena*, (1994) Improvement analytical …, Burakov V. S. et all); M.Deneva, E.Krasteva. P.Bakardjiev, M.Nenchev (2001). Tunable laser devices with inverted spectrum for application in intracavity spectroscopy and spectral selective excitation of atoms, *Proc. of LTL Plovdiv'2001, International Symposium* (2001) 153-156.
5. O.Svelto. *Principles of lasers,* 5th ed. Springer Science-Business Media, 2008.
6. M.Deneva, P.Uzunova, M.Nenchev, Tunable subnanosecond laser pulse generation using an active mirror concept", *Opt. and Quant. Electronics*, 39 (2007) 193-212, **USA**
7. M.Nenchev, V.Stefanov, A. Gizbrekht, Study of Self frequency locking in dye laser at the Sodium D-lines. 3rd Confer. on Luminescence, Szeged, Hungary,1979, Confer. Digest,. Vol.1, 135-139 (in Russian); M.Nenchev, V.Stefanov, A.Gizbrekht, A.Bokhanov, Self frequency locking in lasers at lines of intracavity gases, Bulg. J. Phys. 7, (1981) 521-524
8. E.Stoykova, M. Nenchev, Gaussian Beam Interaction with Air-gap Fizeau Interferential wedge, *J. Opt. Soc. America A*, 27 (1), 58-68 (2010)
9. Y.H.Meyer, M.N.Nenchev, On intracavity absorption and self-frequency locking in pulsed dye laser. *Opt. Communs.*, vol.4, No5, (1982) 292-294 and the literature therein.

Contacts:
M. Deneva, M. Nenchev
Technical University-Sofia, Branch Plovdiv
Address:
Dept. Optoelectronics & Lasers
and Scientific Laboratory QOEL
25 Tsanko Diustabanov St,
4000 Plovdiv
Bulgaria
E-mail:
mar.deneva@abv.bg

© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 23, 2017

THERMO-SENSITIVE PAPER APPLICATION AS METHOD FOR LASER BEAM SPOT STUDY – CASE OF Q-SWITCHED LASER PULSES

VALKO KAZAKOV

Abstract: *Recently, we have developed, as a new technique for laser spot profile investigation, a combined spot registration on appropriately chosen thermo-paper with convenient computer treatment [3]. Such technique presents some essential advantages: spectral insensitivity in very large range (from UV to IR); the registration is not affected by electromagnetic noise; as addition, it is extremely cheap and accessible. Here, on the base of our previous experience, obtained mainly with free lasing Nd:YAG laser pulses, we present results of application for Q-switched pulses of this laser. Some specific behavior of the registration under condition of more than thousand times shorter pulse influence than in free lasing are shown and the condition – power density and energy density limits of correct registration are carried out. The technique under carefully laser illumination control can give acceptable results also in the considered case.*

Key words: *laser spot study, thermo-sensitive paper, Q-switching.*

1. Introduction. General aim of the work.

The aim of the work is extension of our previous results, related with development of the practical application of **T**hermo-**S**ensitive **P**aper **R**egistration **T**echnique (TeSPeRT) for laser spot imaging. This old technique [1, 2] is used in the literature and in the laboratory practice only for visual illustration (including also successfully - for the interesting cases) of the laser spot. The base of the technique, as a rule, is the whitening in the laser beam incident area on the black thermo-sensitive paper. In previously reported applications [1, 2] are used accidentally taken sensitive materials and without conclusion about energetic characteristics of the spot. As we have shown in Ref. [3], in condition of combining the convenient chosen thermo-sensitive paper for registration and convenient computer treatment, and with correctly defined illumination, this technique can be successfully used for laser beam energetic parameter study. In the noted work we demonstrate the correct determination of the energy density distribution in the beam cross-section, as well as evaluation of the beam energy. Our previous investigation is by the use of the free lasing Nd:YAG laser light (pulse length 0.1 – 3 ms; power ~ 10 - 40 kW). The obtained results confirm the usefulness and the expected potential of the technique developed by us, especially for the condition of typical free lasing light emission. Taken into account the essential difference between the characteristics of the free lasing and Q-switched laser pulses, especially temporal (ms against ns) and power (kW against hundreds kW for comparable energies), the application of the developed by us technique for the Q-switching case needs particular treatment of the Q-pulses registration, which is the aim of the presented work. Here, we present the results of noted above study and namely concerning the energetic aspects and conditions for correct applications of the TeSPeRT for Q-switched laser pulses treatment. Note that such technique presents some essential advantages in comparison with electronic based techniques of this type: spectral insensitivity in very large range (from UV to IR); the registration is not affected by electromagnetic noise; and as addition, it is extremely cheap and accessible. The beam-profile knowledge is of essential importance in the laser manufacturing and in the scientific and practical application of the lasers [1 - 5].

2. Experimental investigations and results.

We start the proposed study using the results, given in our previous work [3], about the free laser pulse light treatment applying TeSPeRT. In the work [3] we have presented and discussed the principle and the convenient condition to obtain correct energy density distribution in the laser beam cross section. The given there results are for pulse

 ISSN 1310 - 8271

energies in the range of 0.3 J to 10 J with pulse duration in the range 100 - 3000 μs – i.e. the typical free lasing operation of used in the tests Nd:YAG lasers (at 1.06 μm and 1.36 μm generation). Taken into account the well-known essential difference between the characteristics of free lasing and Q-switched laser pulses, especially temporal and pulse power characteristics [1, 2], we provide systematic study for the case of the Q-pulses registration by the TeSPeRT.

In the study of free lasing case in Ref. [3], we have shown that a very suitable material for the beam spot marking is the tracing paper, blacked at one side through the copy machine and illuminated at the non-blacked side by the laser beam. Here, following our works [3] and complementary investigations, we have used for beam spot registration the same type of tracing paper (A4 92 gr/m^2 from Sihl Digital Imaging Company), blacked one (1X), two (2X) or three (3X) times through the standard xerography type action machine - Konica Minolta Dialfa Di 5510. The blacking-out was made by copying the tracing paper at normal blacked position of the machine option. We do not observe noticeable difference comparing few tracing papers from different producers and also the blacking by different copiers of the noted before type (Sharp-MX-3500, Toshiba 2500c) for normal blacking operation. The transmission of pure tracing paper is measured to be ≈ 50 % when the paper is placed closely to the entrance of the light power meter (Thorlabs SN6050506). In this manner the measurement with flat and large entrance of 2 cm diameter accept all the diffused light. We used both He-Ne yellow (0.595 μm) laser beam with incident diameter on the paper of ~ 2 mm and power of 3.6 mW, and focalized light from standard electric bulb with tungsten filament. There is no notable difference of the measured transmissivity. The corresponding measurements of the transitivity after blacking are: for 1X - (10 ± 2) %; for 2X - (1 ± 0.1) % and for 3X ≈ 0.1 %. The transmission through marked spot by Q-switched pulse is measured to be ≈ 25 %, which is ~ 2 times more than the transmission through spot illuminated by free lasing pulse with the same energy – given in Ref. [3]. There are quite differences in 2D and 3D graphics, when the Q-switched marked spot is scanned with white cover and with black cover. The graphics of spots, scanned with black cover are with unusual form and lower height than is expected to be. Hence, in case of Q-switched pulses there is not only whitening of the blacked paper, but also cleaning at some points of the spot where is a higher power density and there the blacked paper is transparent like a pure tracing paper. The computer treatment in this work is made after scanning the spots by scanner with white cover. Thus, the cleaned/ transparent points of the spots are scanned like white points, which gives more accuracy, because the base of the method is whitening proportional to the energy/ power density. In case of free lasing, there is no notable difference between scanning with white or black scanner cover, because there is only whitening without cleaning of the blacked layer.

As a laser beam source for the investigation we used passively Q-switched Nd:YAG laser. The study was performed at laser parameters: generally for the energy between 0.1 J to 1.5 J and pulse length 1 μs ± 0.2 μs; repetition rate – single pulse (to 1 Hz); emitting at line 1.06 μm. In most of the experiments the energy of the beam, illuminating the studied material, was equal to the laser output energy, except the case where fine plate filters were used. The energy was measured with FIELDMAX energy meter, Coherent (USA), the pulse duration - with two-channel 200 MHz storage oscilloscope RIGOL DS1202CA (USA) and light detectors with resolution better than 5-ns - R-108 (BG).

At the left side of Fig.1 are shown the typical marked spot by Q-switched pulse (Fig.1a) and the graphs, obtained by 3D (Fig.1b) and 2D (Fig.1c, across the spot diameter of ~ 0.7 cm) computer treatment of its whitening. The Q-pulse duration is ~ 1 μs and the pulse energy is 0.88 J on the 1X type of paper. For comparison, in the same picture at the right side are shown - the spot, made by the free lasing laser light (Fig.1d) and the 3D (Fig.1e) and 2D (Fig.1f) computer treatment of the white spot. The free laser pulse duration is ~ 300 μs, pulse energy ~ 1.17 J and spot diameter is ~ 0.7 cm. The computer curves are normalized to their maximal values.

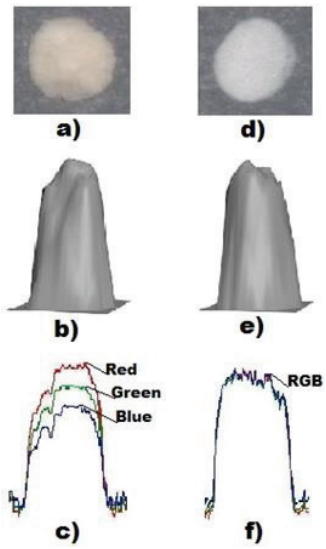


Fig.1. *Marked spots by Q-switched pulse (a) and free lasing (d), and their corresponding 3D (b, c) and 2D (e, f) images for 1X type of paper.*

The general difference between the Q-switched and free lasing spots, as it can be seen also in Fig.1, is that the Q-spots are colored in red. This leads to different colors presentation of the 2D image of the spot – red, green and blue curves. This is quite different from the 2D image of the spot for the free lasing, where the three colors practically coincide completely. We attribute this difference with formation of products of burning and/ or photochemical reaction at the black layer due to the high power of the illuminating light in the Q-pulses. The nature of the coloring needs future chemical investigation. Thus, the first step of the investigation is to see: i) at which degree the registration (near constant R, see below) is correct when we use for beam chosen color and ii) which color corresponds better with the real energy density distribution in the spot. The curves of the three colors that describe the 2D description of the graph are shown in the figure with notation – Red, Green, Blue and their composition as RGB.

We accept, as approach to computer treatment of the spot, the methodology argumented, developed and experimentally tested in details in Ref. [3] (our previous work for the free lasing case). Thus, following the discussion of this work, for criterion of correctness for the obtained energy density distribution $W_E(x, y)$ J/cm^2 in the spots, we will use the ratio $R = E/V$ of the pulse energy E and the calculated volume V under the computer obtained envelop graph $W_H(x, y) = W_E(x, y)/K_1$ of the whitening. The constant value of $K_1 = R = E/V$ can serve as a criterion for correctness of energy density distribution, obtained by the experimental measurement of E and computer treatment - the arbitrary graphs of whitened and computed volumes under the 3D graphs.

Following the formulated above problems, we treated series of stamps with spots on the described blacked tracing papers, illuminated by variety of Q-switched pulses (Nd:YAG, 1.06 μm). The passive Q-switched operation of the laser was with controlled variable energy output between 0.1 J and 1.5 J and pulse length of ~ 1 μs. For the spots with diameter of 0.7 cm, the energy (and the power) density vary respectively between ~ 0.5 J/cm^2 (0.5 MW/cm^2) and 3 J/cm^2 (3 MW/cm^2).

Preliminary question is the dependence of the pulse energy (and power) density and especially limits of notable marking and saturation. We found that for the three investigated type of papers - 1X, 2X, 3X the noted limits are relatively close to each other - minimum energy density for reliable registration was ~ 0.5 J/cm^2 (power density ~ 0.5 MW/cm^2) and the saturation was observed for ~ 3 J/cm^2 (~ 3 MW/cm^2).

As example of the realized investigations, a collection of Q-switched laser light spots marked with different pulse energies on the paper 1X, noted as Q1, Q2, Q3, Q4 and their computer treatment are given in Fig.2. For comparison are given also the two spots F1 and F2 of free lasing pulses.

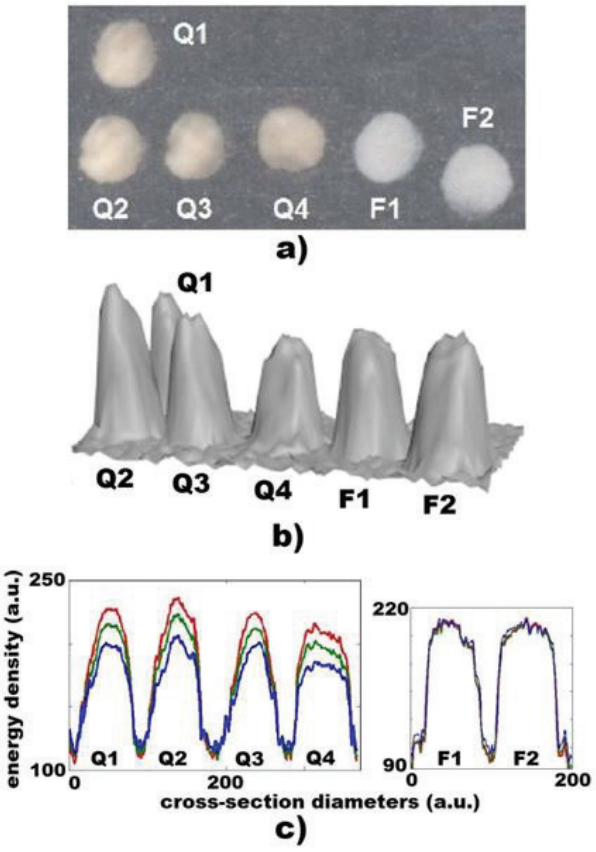


Fig.2. *Marked spots by Q-switched pulses Q1 - Q4 and free lasing light formed spots F1, F2 on 1X type of paper (a); their 3D (b) and 2D (c) graphics.*

The correctness of TeSPeRT applications for the free lasing case is shown in Ref. [3]. Thus, we treat here the Q-switched pulse spots. The pulse length is ~ 1 μs and the pulse energies are given in Table 1. In the same table, the calculated volumes under the 3D graphs and the ratio R_i for the RGB registration are also given. With RGB are noted the averaged curves, giving the average value for each point from the sum of the corresponding points for the three curves. The computer program automatically gives the averaged volume of the volumes for the three colors scanning. We have also study separately the 3D images for each color and the results are shown in the Table 2 and Table 3.

Processing of each paper type will be detailing described on the example of 1X, for which the best results are obtained. After computer treatment of the four (Q1 – Q4) spots shown in Fig.2 and with measured light energy that forms each of them, we obtained the average value $R_{av}=(\sum V_i/E_i)/N=\sum R_i/N$, where sum is for i = 1 - 4,

V_i and E_i are the volume and the energy of the spots numbered as ***i*** and $N = 4$; $R_{av} = 1.29$. We defined the deviation (or error) ΔR_i for each spot as $\Delta R_i = R_i - R_{av}$ and the relative average deviation for the considered group spots of 1X sheet was $\Delta R_{av} \approx 0.12 = 12$ %.

Table 1 and Table 2 are illustrations of the analisys - measured energies and calculated characteristics of the computer treatment for the averaged three colors (RGB) and separately – for each color – red, green, blue. The sheet is 1X, which gives the best results.

In the same manner, we have treated the noted other cases. The results are given in Table 3 - comparison of the results of treatment for the two types blacked tracing paper - 2X and 3X.

Table 1. *Characteristics for 1X type of paper – energy and results for averaged three colors (RGB)*

№	E, [J]	V(RGB), [a.u]	R_i, [a.u.]	R_{av}, [a.u.]	$\lvert\Delta R\rvert$, [a.u.]	ΔR_{av}, [a.u.]	$\frac{\Delta R_{av}}{R_{av}}$, %
Q1	0,95	1,09	1,14	1,29	0,14	0,12	9
Q2	0,82	1,21	1,48		0,19		
Q3	0,76	1,01	1,32		0,04		
Q4	0,88	1,06	1,20		0,08		

Table 2.a. *Characteristics for 1X type of paper – volume and R_i for R, G, B colors*

№	V(Red), [a.u.]	R_i, [a.u.]	V(Green), [a.u.]	R_i, [a.u.]	V(Blue), [a.u.]	R_i, [a.u.]
Q1	1,19	1,24	1,06	1,19	0,93	0,98
Q2	1,34	1,63	1,19	1,45	1,02	1,25
Q3	1,01	1,44	0,99	1,30	0,87	1,15
Q4	1,17	1,33	1,03	1,17	0,9	1,02

Table 2.b. *Characteristics for 1X type of paper - R_{av}, ΔR_{av}, $\Delta R_{av}/R_{av}$ for R, G, B colors*

Color	R_{av}, [a.u.]	ΔR_{av}, [a.u.]	$\Delta R_{av}/ R_{av}$, %
Red	1,41	0,13	8,8
Green	1,26	0,12	9,1
Blue	1,26	0,10	8,8

Table 3. *Characteristics for 2X and 3X type of paper*

Sheet	R_{av}, [a.u.]	ΔR_{av}, [a.u.]	$\Delta R_{av}/ R_{av}$, %
2X (RGB)	1,54	0,17	11,4
2X (Red)	1,68	0,19	11,2
2X (Green)	1,51	0,17	11,3
2X (Blue)	1,36	0,17	12,2
3X (RGB)	1,79	0,44	24,8
3X (Red)	1,89	0,45	23,8
3X (Green)	1,77	0,44	25
3X (Blue)	1,55	0,4	25,6

3. Discussion.

The syntheses of the typical most important results of the investigation are summarized in the examples in Table 1, Table 2 and Table 3. Following the results, we can conclude that the application of the technique developed gives acceptable accuracy of the beam profile registration. For the defined parameters of the Q-switched illuminating laser, described in the work, as most suitable blacked paper can be considered the 1X paper type. The average error for this case is ~ 12 %, which is acceptable information for the desired energy density distribution. Interesting fact is that the treatment taking the three colors gives not very different results - less than 1 %.

4. Conclusion.

In the work we present results of developed by us simple, cheap and accessible technique (TeSPeRT) for Q-switched laser beam profile study. We have shown this study using appropriate, well defined and reproducible as a production sensible materials – xerox blacked tracing papers for laser spot registration in combination with suitable computer treatment. This technique can be successfully employed as beam-profile registration method with well acceptable accuracy. We present the conditions (our experiments), when the developed technique shows such applicability. Note that, such technique represents some general essential advantages in comparison with electronic based beam-profile treatment techniques: the registration is not affected by electromagnetic noise and has spectral insensitivity in very large range (from UV to IR).

Acknowledgement:

The author thanks for the support for the present publication to the Contract DN8/13 FSI–BG, assoc. prof. M. Deneva and prof. M. Nenchev.

REFERENCES

1. Ifflander, R. (2001). *Solid-State Lasers for Materials Processing*. Springer-Verlag Berlin. Heidelberg, New York, series in Optical Sciences. ISSN0342-4111, ISBN 3-540-66980-9 (2001).
2. Carlos B. Roundy, Spiricon, Inc., http://aries.ucsd.edu/LMI/TUTORIALS/profile-tutorial.pdf, Beam Profile Measurements and the literature therein.
3. V.P. Kazakov, M.A. Deneva, M.N. Nenchev, N.Y. Kaymakanova. Xerox treated tracing paper as suitable and accessible material for development of new laser beam–profiler technique**.** Journal of the Chemical Institutes of

the Bulgarian Academy of Sciences and of the Union of Chemists in Bulgaria 2016 Volume 48 / Special issue G, pp 85-90.;
M. Deneva, M. Nenchev, V. Kazakov, N. Kaymakanova. Development of the thermo-sensitive paper application as method for laser beam spot study. Journal of the Technical University – Sofia, Plovdiv branch, Bulgaria "Fundamental Sciences and Applications", Volume 22, 2016, pp 92-95.

4. P. Uzunova, S. Rabadgiiska, M. Deneva, M. Nenchev et al**.** Proc. SPIE, Vol.8770-54, (2013). Study the penetration of IR laser radiation in human teeth: determination of the absorbed and scattered parts.
5. Deneva M., Nenchev M., Wintner E., Topcu S. (2015). Coaxial-geometry tunable dual-wavelength flashlamp-pumped Nd:YAG laser. *Opt. Quant. Electronics*, DOI 10.1007/s11082-015-0205-5 online publ.

Contacts:
V. Kazakov
"QOE" Scientific Laboratory, R&D Dept., Technical University of Sofia and OELE Dept., TU-Branch Plovdiv, BG
v_p_k@mail.bg

© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 23, 2017

IMAGE PROCESSING ON ANDROID DEVICES WITH OPENCV

CEM TASKIN, ALTUG YIGIT, DENIZ TASKIN, AYSE CELIK TASKIN

Abstract: *Nowadays, computer vision is a popular approach in computer science. Many applications running on personal computers use computer vision libraries in order to have object detection capability for many years. Today mobile devices have powerful hardware like quad core cpus, high density cameras and high speed internet connection. With the hardware achievements in mobile devices, programmers developed mobile versions of computer vision libraries. In this paper, we are testing an Android application using computer vision library for distinguishing colors. This sample application works on real time camera and distinguish a color from others in the scene that was selected by user.*

Key words: *image processing, mobile device, android, color, filtering, opencv*

1. Introduction

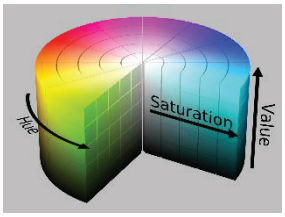
Object detection is one of the most challenging problems in computer vision. It is difficult to detect due to the significant amount of variation between images belonging to the same object category. Other factors, such as changes in viewpoint and scale, illumination, partial occlusions and multiple instances further complicate the problem of object detection. Most state-of-the-art approaches to object detection rely on intensity-based features that ignore color information in the image. Color has been shown to yield excellent results in combination with shape features for image classification. [1]

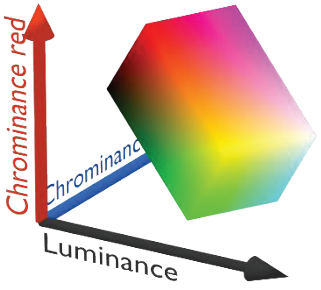
Usually computer vision libraries use captured images as input. But in real time applications, the input is always a collection of continuous frames (pictures). It is a main problem processing these frames in short times. And also it is very important to give the result of process as soon as possible. Computer vision techniques like 3D segmentation and modelling is not suitable for real time processing. In this study, color filtering method is used for distinguishing colors and tested on an Android mobile device.

2. Color Filtering

In color filtering method, determining the correct threshold value is the most important fact on accuracy. In order to get consistent results, appropriate color space must be selected before processing images.

Restricting the color information merely to RGB components is a simple abstraction that dismisses the information available within the color object. A color is fully defined by its complete wavelength response, whereas the RGB color space represents only three wavelengths. The perceived color also depends on the illumination condition, viewing angle, and sensor type. As a consequence, efficient color image processing requires an adequate color representation. Different color spaces can be used to represent various color components, with different degrees of interdependency among them. Of the four classic color spaces, the hue/saturation/value (HSV), YCbCr , L*a*b*, and RGB shown as Figure 1. [6]





 ISSN 1310 - 8271

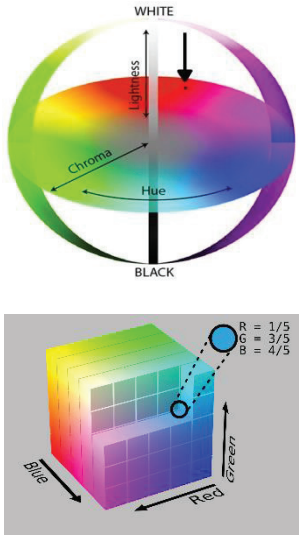


Fig. 1. *HSV, YCbCr , L*a*b*, and RGB Color Space*

The HSV (Hue-Saturation-Value) theory is the most common representation of points in an RGB (Red-Green-Blue) color technical model. Computer graphics pioneers developed the HSV model in the 1970s for computer graphics applications (A. R. Smith in 1978, also in the same issue, A. Joblove and H. Greenberg). A HSV theory is used today in color pickers, in image editing software, and less commonly in image analysis and computer vision. [7] The HSV color space is more related to human color perception. [8]

HSV color space separating luminance component with chrominance component. The luminance component lies in V (intensity value) and the chrominance component lies in H (Hue) and S (Saturation). In HSV model, Hue defines the kind of color like red, yellow, green, blue, or the combination between those colors. Hue value is between 0-360. Saturation gives a measure of the degree by which a pure color is diluted by white light, and value show the intensity of light-dark color.[9]

Color is a strong tool for image segmentation. Color-based segmentation is better than edge-based and luminance histogramming techniques because color is computational inexpensive, and it can give more information than a luminance-only image or an edge-segmented image which need more computational resources, so they make real time application systems hard to realize.[5]

3. OpenCV and Android

OpenCV (Open Source Computer Vision) is a library of programming functions mainly aimed at real-time computer vision. Originally developed by Intel's research center in Nizhny Novgorod (Russia), it was later supported by Willow Garage and is now maintained by Itseez. The library is cross-platform and free for use under the open-source BSD license. The image processing algorithms are implemented using OpenCV. OpenCV was designed for high computational efficiency with a strong focus on real time applications. OpenCV was written with optimized C. [3,4] OpenCV library is a common library used in image processing applications.

Table 1. *Worldwide Smartphone Sales by Operating System*

Worldwide Smartphone Sales to End Users by Operating System in 4Q15 (Thousands of Units)

Operating System	4Q15 Units	4Q15 Market Share (%)	4Q14 Units	4Q14 Market Share (%)
Android	325,394.4	80.7	279,057.5	76.0
iOS	71,525.9	17.7	74,831.7	20.4
Windows	4,395.0	1.1	10,424.5	2.8
Blackberry	906.9	0.2	1,733.9	0.5
Others	887.3	0.2	1,286.9	0.4
Total	**403,109.4**	**100.0**	**367,334.4**	**100.0**

Source: Gartner (February 2016)

It is possible to use this library on Android Studio by adding OpenCV Android Software. Development Kit (SDK). This SDK can be downloaded from the http://opencv.org/downloads.html.

As seen on Table 1, Android OS is the common operating system used by mobile devices. Android as a development platform use the Java Standard Edition with additional Android classes. Applications are developed in the Java language using Android Software Development Kit (SDK). For a low level interaction with hardware, a modified Linux kernel is used [2].

4. OpenCV on Andoroid OS

In our study, we are testing an Android application using OpenCV for real time color based image segmentation. The flowchart of the application is shown on Figure 2. OpenCV's onCameraFrame method captures frames in RGBA color space. These frames are converted into 8-bit 4-channel pixel matrix. After that step, RGBA to HSV conversion are applied.

5. Application Flow Chart

In the RGBA color space R (Red), G (Green), B (Blue), A (alpha) channels are available. For each channel, color values are stored in 8 bits and has a range between 0 and 255. Alpha channel is ignored on conversion from RGBA to HSV. In the first step of converting we divide each color value with 255 and we obtain a conversion value between 0 and 1.

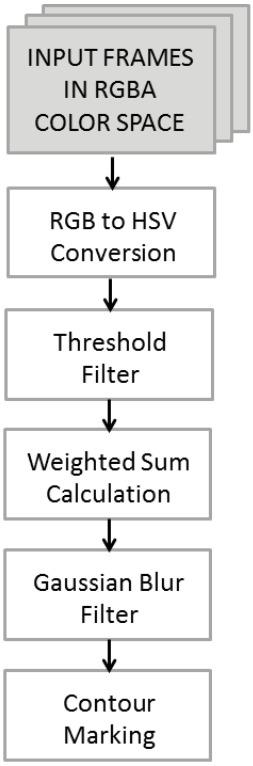


Fig. 2. *Flow Chart of Our Application*

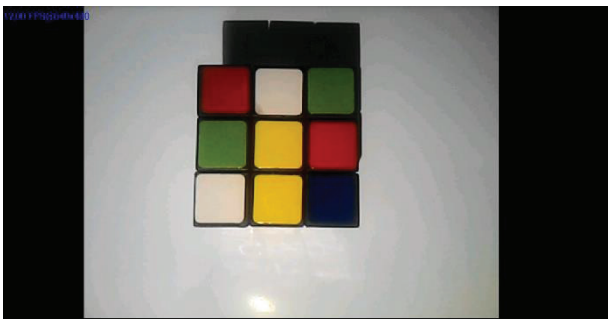
In the next step, we are using following formulas to calculate HSV values [11].

$$V \leftarrow max(R, G, B)$$

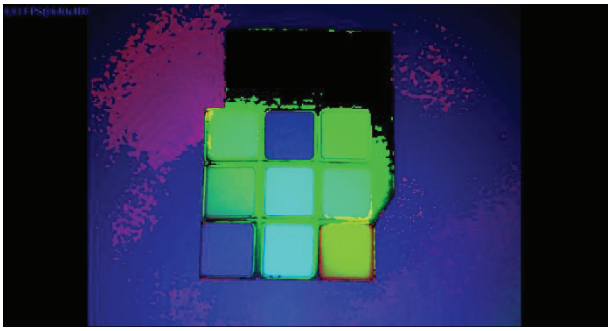
$$S \leftarrow \begin{cases} V - \dfrac{\min(R,G,B)}{V} & If\ V \neq 0 \\ 0 & otherwise \end{cases} \quad (1)$$

$$H \leftarrow \begin{cases} \dfrac{60(G-B)}{V - \min(RGB)} & If\ V = R \\ 120 + \dfrac{60(B-R)}{V - \min(RGB)} & If\ V = G \\ 240 + \dfrac{60(R-G)}{V - \min(RGB)} & If\ V = B \end{cases}$$

This conversion can be performed in OpenCV with function cvtColor. In Figure 3 a conversion from RGB to HSV is shown



Original image



Converted Image

Fig. 3. *RGB to HSV Conversion*

5.1. Threshold

After RGB to HSV conversion, we select a color for filtering. In our test, we select Red color. Before the threshold filtering, two color range values are selected (The low and high color tone). For threshold filter we use OpenCV's inRange function. The formula used by this function is shown in Formula 2.

$$Output \leftarrow \begin{cases} 1 & If\ Input \in Range \\ 0 & If\ Input \notin Range \end{cases} \quad (2)$$

In Figure 4, the output of this function applied to sample image is shown.

Fig. 4. *Output After Threshold Filter*

5.2. Weighted Sum of Two Arrays

In order to increase performance of color filtering, weighted sum of two threshold output matrix is calculated. In this calculation operation, two matrix has equal weight (50% and 50%). We use Formula 3 for this operation. In OpenCV, for calculating the weighted sum of two threshold output matrix,addWeighted function is used.

(3)

$$dst(I) = saturate(src1(I) * alpha + src2(I) * beta + gamma)$$

Parameters:

src1 – first input array.
alpha – weight of the first array elements
src2 – second input array
beta – weight of the second array elements
dst – output array
gamma – scalar added to each sum.[11]

5.3. Gaussian Blur

We use Gaussian Blur to reduce noise of the weighted sum value obtained. When the Gaussian blur applied, pixel values apart from the others become closer. We use Formula 4 for this filter. In OpenCV, for applying Gaussian Blur filter GaussianBlur function is used.

(4)

$$G_0(x,y) = Ae^{\frac{-(x-\mu_x)^2}{2\sigma_x^2} + \frac{-(y-\mu_y)^2}{2\sigma_y^2}}$$

In Figure 5, the output of Gaussian Blur filter applied to sample image is shown.

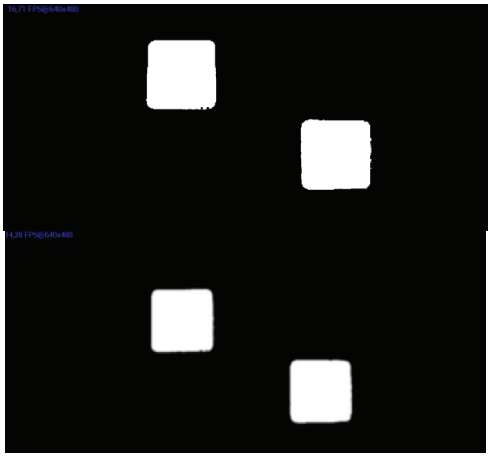


Fig. 5. *Output After Gaussian Blur Filter*

5.4. Contour Marking

After applying Gaussian Blur filter we obtain a smooth image. In this output image, we find the positions of white pixels and the relationship of each other.

Contours can be explained simply as a curve joining all the continuous points (along the boundary), having same color or intensity. The contours are a useful tool for shape analysis and object detection and recognition. In OpenCV that can be applied by calling the findContours function. This function has three arguments. First one is source image. The second is contour retrieval mode. Third parameter is contour approximation method. And it outputs the contours and hierarchy. [13] Contour retrieval mode selected for this Android application is CV RETR TREE. In this mode all of the contours output vector are obtained with hierarchical relationships. [12] The approximation method selected for this Android application is CV CHAIN APPROX SIMPLE.

By using CV CHAIN APPROX approximation method all redundant points are removed and that saves memory which is very important for mobile devices.

After this step, the pixel locations of the rectangle are obtained. In OpenCV rectangle function is used for drawing a rectangle. In Figure 6, the result of color filtering is shown. In our test, the application follows the color and draws contours in real time video.

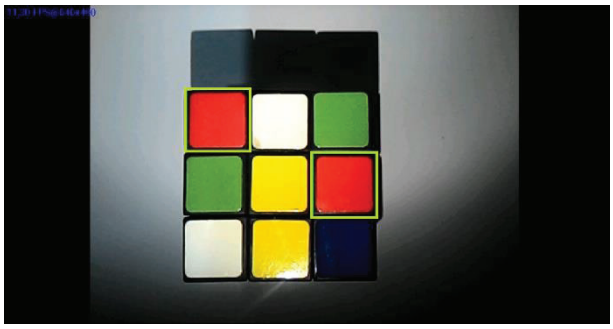
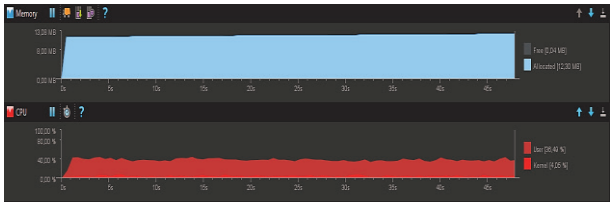


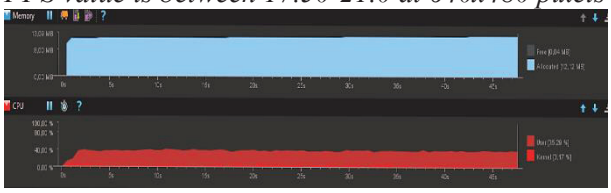
Fig. 6. *Result Of Color Filtering*

6. Performance Testing

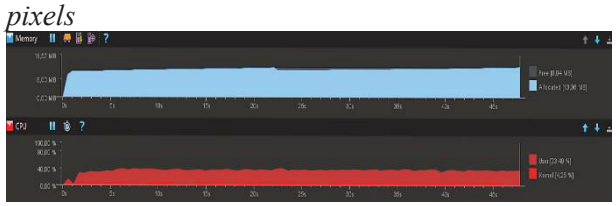
Application performance was tested with an Intel® Atom™ Z2560 1.6 GHz CPU equipped mobile device on Android Studio. The results of CPU usage, Memory usage and FPS(Frame Per Second) are shown in Figure 7.



Threshold Filter performance
FPS value is between 17.50-21.0 at 640x480 pixels



Gaussian Blur filter performance
FPS value is between 15 and 17.50 at 640x480 pixels



Contour Marking
FPS value is between 9.0 and 16.50 at 640x480 pixels

Fig. 7. *Test Results*

7. Conclusion

As a result of our mobile application tests, we can say that mobile devices has capable of color filtering in real time. Our future work is to test real time object detection methods that use deep learning techniques.

REFERENCES

1. Fahad Shahbaz Khan, Rao Muhammad Anwer, Joost van de Weijer, Andrew D. Bagdanov, Maria Vanrell, Antonio M. Lopez. Color attributes for object detection. In CVPR, 2012.
2. Damir Demirovi´c, Amira Šerifovi´c-Trbali´c, Naser Prlja˘ca, Philippe C. Cattin. Evaluation of image processing algorithms on ARM powered mobile devices. MIPRO, 2014.
3. https://en.wikipedia.org/wiki/OpenCV#cite_note-2
4. Deepthi.R.S, S.Sankaraiah. Implementation of mobile platform using qt and open cv for image processing applications. IEEE Conference, 2011.
5. Houssem Lahiani, Mohamed Elleuch, Monji Kherallah. Real time hand gesture recognition system for android devices. ISDA, 2015.
6. Hossein Nejati, Victor Pomponiu,Thanh-Toan Do, Yiren Zhou,Sahar Iravani, and Ngai-Man Cheung. Smartphone and mobile image processing for assisted living. IEEE Signal Processing Magazine, 2016.
7. Ján Haluska, On fields inspired with the polar HSV -- RGB theory of colour. In arXiv, 2015.
8. Alberto Albiol, Luis Torres, Edward J. Delp. Optimum color spaces for skin detection. IEEE Image Processing Conference, 2001.
9. Priyanto Hidayatullah, Miftahuddin Zuhdi. Color-texture based object tracking using HSV color space and local binary pattern. In International Journal on Electrical Engineering and Informatics, 2015.
10. http://docs.opencv.org/2.4/modules/imgproc/doc/miscellaneous_transformations.html
11. http://docs.opencv.org/2.4/modules/core/doc/operations_on_arrays.html
12. http://docs.opencv.org/2.4/modules/imgproc/doc/structural_analysis_and_shape_descriptors.html
13. http://docs.opencv.org/trunk/d4/d73/tutorial_py_contours_begin.html

© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 23, 2017

ALTERA AND XILINX TOOLS IN EDUCATION

ATANAS KOSTADINOV

Abstract: *In this paper, Altera and Xilinx tools used in Technical University - Sofia, Plovdiv branch are presented. Initially, Altera (now the Intel Programmable Solutions Group) tools are applied in Reconfigurable Logic and VLSI (Very-large-scale integration) design courses. Recently, it has been received donation from Xilinx university program (XUP) in the form of FPGA (Field-programmable gate array) design kits and Vivado Design Suite: System Edition. Now tools from both reconfigurable logic market leaders are used as themes in bachelor's and master's degree syllabuses.*

Key words: *Altera and Xilinx tools, Xilinx university program, Vivado Design Suite: System Edition, Nexys4 DDR*

1. Introduction

The reconfigurable integrated circuits especially CPLD (Complex programmable logic device) and FPGA (Field-programmable gate array) are used in higher education very often [1, 2, 3, 4, 5, 7, 8, 9, 10, 11] in different courses. The subjects can vary from Digital logic design [1, 2, 10, 11], Embedded Systems design [7, 8, 9] to Reconfigurable Logic and VLSI (Very-large-scale integration) design [4, 5]. In the lectures are presented some basic issues regarding reconfigurable integrated circuits, their internal organization and the way of their reconfiguring. In laboratory exercises are used CPLD and mainly FPGA boards and students implement some digital and microprocessor circuits. This practically oriented higher education helps students to make required steps from using fixed logic to applying reconfigurable one and corresponding hardware description languages as Verilog and VHDL (Very high speed integrated circuits description language).

Technical University - Sofia, Plovdiv branch has about fifteen years tradition of using contemporary reconfigurable integrated circuits in the educational process [2, 3, 4, 5]. During these years some improvements have been done in course syllabuses as well as in software and hardware tools. The last change was received donation from the Xilinx University Program (XUP) [12] in the form of five FPGA design kits of type Nexys4 DDR and twenty licenses for Vivado Design Suite: System Edition. Together with previously used Altera tools now design software and FPGA boards from both major reconfigurable logic suppliers are used in bachelor's and master's degree syllabuses. In the next sections are presented some details about reconfigurable integrated circuits and tools used from the beginning up to now.

2. Altera (now the Intel Programmable Solutions Group) tools

Initially, the current lecturer of Reconfigurable Logic and VLSI design courses has graduated from ALaRI (Advanced learning and research institute) [13] in Lugano, Switzerland. During education process, there were two subjects related to reconfigurable integrated circuits produced by Altera and VHDL. After finishing master degree course and returning to Bulgaria, I have been involved in teaching of similar material to students from Technical University - Sofia, Plovdiv branch. Together with it, the education in reconfigurable circuits has been started in the former John Atanasoff Technical College. In order to get licensed version of the design software and to obtain boards for laboratory exercises, it has been applied for a donation through Altera University Program [14].

2.1. First CPLD board

Fig. 1. *UP2 board*

 ISSN 1310 - 8271

The application was successful. The first donation from Altera Corporation has been in form of licensed software (Quartus II and ModelSim) and one CPLD board (UP2). The received UP2 (University Program 2) is presented in Fig. 1.

It has been realized a license server in 2202 computer hall and the students have been used the licensed versions of Quartus II and ModelSim design software. The obtained single UP2 board was a good starting point in the lab equipment process. It has been used as a demonstration of loading a simple design into one of two CPLDs located on the board. UP2 connection to personal computer (PC) was parallel interface using printer port (LPT - Line print terminal 1) of the PC. The length of this cable was not enough UP2 to be put on the laboratory desk. It has been decided to be found another way in order to be increased the number of the FPGA boards that are needed for the laboratory exercises.

2.2. Support from the University

With the financial support of the Technical University - Sofia have been obtained two Cyclone II FPGA starter development boards. One of them is presented in Fig. 2.

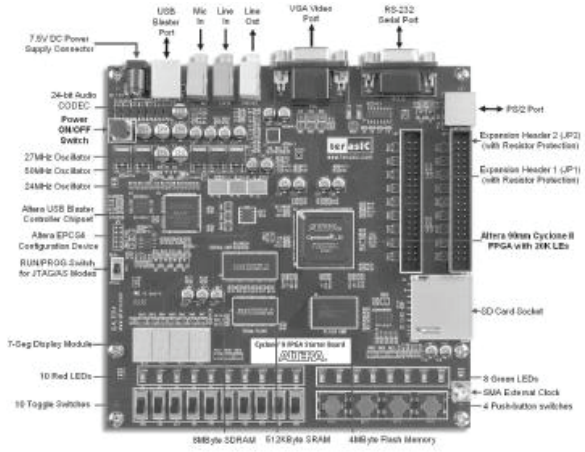


Fig. 2. *Cyclone II FPGA starter development board*

These boards are used in the laboratory exercises of the courses connected to the reconfigurable integrated circuits.

2.3. Donation from the firm partner

With help of the firm partner (ASIC Depot) of Computer Systems and Technologies Department has been added additionally two Cyclone II FPGA starter development boards. In this way the number of existing FPGA boards are very close to the sufficient amount required for the laboratory exercises.

Unfortunately, we had no laboratory equipment and design software by another major FPGA producer - Xilinx. We have decided to apply for Xilinx University Program [12]. The results of this application was again successful. Details about new donation are presented on the next lines.

3. Xilinx tools

We have applied for the five FPGA boards of type Nexys4 DDR. The application was again successful, as it mentioned before, and we have obtained these boards. The view of one of them is presented in Fig. 3.

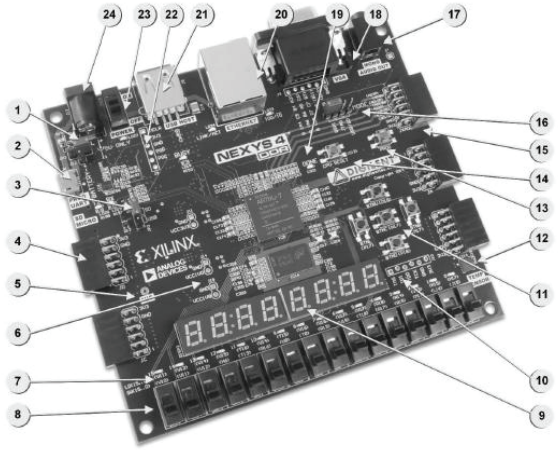


Fig. 3. *Nexys4 DDR board*

Additionally, we have received twenty licenses for the Vivado Design Suite: System Edition. Now we have FPGA boards and corresponding design software from both global players of reconfigurable integrated circuit market. These tools we use in the educational process in Technical University, Plovdiv branch as well in the research activities, too.

4. Altera and Xilinx tools used in the teaching and research activities

4.1. Educational process using both tools

At the beginning, we have started using Altera (now the Intel Programmable Solutions Group) tools. The corresponding syllabuses were presented in the paper [5]. As example, in the lectures are introduced IPs (Intellectual Properties) modules. In the Fig. 4 is presented a block diagram of the PLL (Phase - locked loop) circuit. Then corresponding IP module called ALTPLL Megafunction is presented [15].

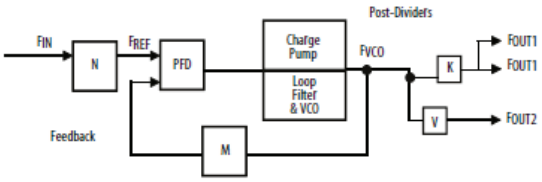


Fig. 4. *PLL block diagram*

At the end of Reconfigurable logic course a small design task is given to the students. They have to describe in VHDL a simple test system. This system uses a 4-bit binary counter and a LFSR (Linear feedback shift register) in order to be generated input stimuli to device under test. One solution of this task is presented in Fig. 5 [5].

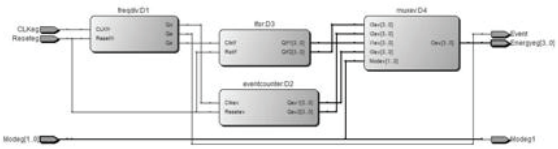


Fig. 5. *Synthesized test circuit*

Now, in the lectures' material an information about Nexys4 DDR and Vivado Design Suite: System Edition has been added. In the Fig. 6 an electronic circuit of Nexys4 DDR basic I/O (connected Input / Output devices) is presented.

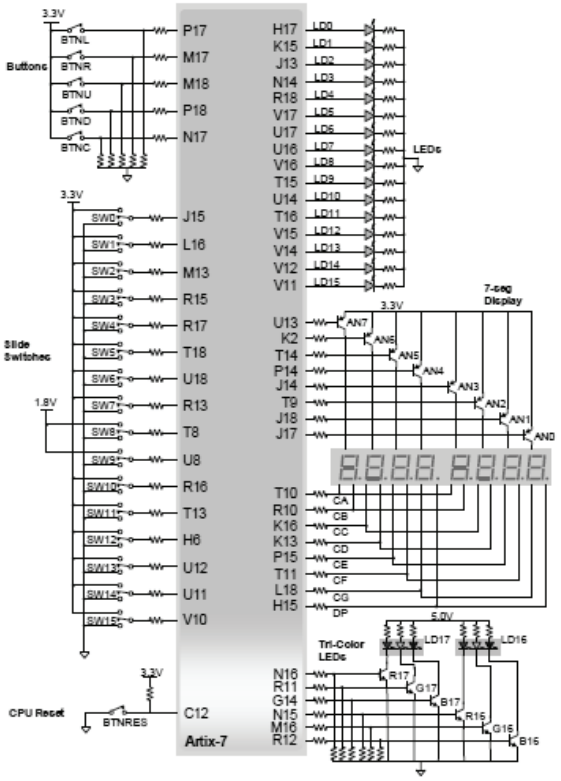


Fig. 6. *Nexys4 DDR basic I/O electronic circuit*

In the Fig. 7 a screenshot of a simple VHDL design (multiplexer) realized using Vivado Design Suite: System Edition is presented . This a part of the lectures in Reconfigurable Logic and VLSI (Very-large-scale integration) design courses.

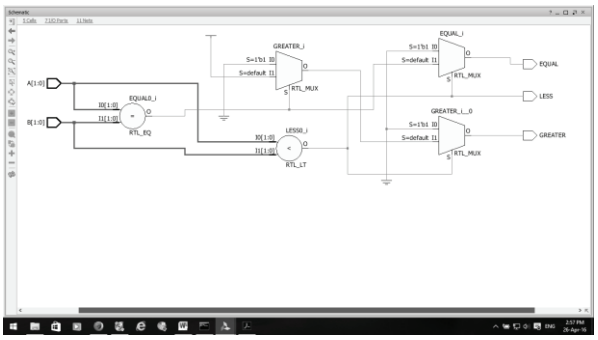


Fig. 7. *Vivado Design Suite: System Edition project*

4.2. Research achievements using both tools

It has been published many research papers in which FPGA boards and tools (mainly from Altera Corporation) have been used [2, 3, 4, 5]. Some research results have been included in the book section [6].

As example, a predicate logic processor (PLP) has been designed and tested during ERCIM (European research consortium for informatics and mathematics) postdoctoral fellowship in NTNU (Norwegian University for Science and Technology) under guidance of Prof. DSc. Guennadi Kouzaev.

A simplified block diagram of the PLP is presented in Fig. 8.

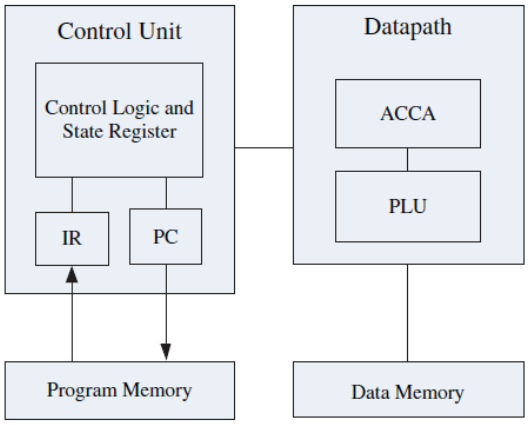


Fig. 8. *PLP block diagram*

Additionally, a patent application has been done [16]. Now, in the future research activities will be added some results using Nexys4 DDR FPGA board and Vivado Design Suite: System Edition. We have already started with a bachelor thesis connected to design of a master device controlling I2C (Inter-Integrated Circuit) bus. After successful defense and with an additional work, the material is able to be presented on a scientific conference.

5. Conclusion

In the paper, the organization of the educational process and research activities connected to the reconfigurable integrated circuits have been presented.

It has been given information regarding using tools from both major suppliers of CPLDs and FPGAs.

It has been mentioned that the XUP donation is already used in the teaching activities. The current lecturer plans that this donation will be used also in the future research activities.

6. Acknowledgements

I gratefully acknowledge the funding received towards my two postdoctoral fellowships from the European Research Consortium for Informatics and Mathematics (ERCIM) as well as from FP7 (Framework Program 7) of the EU (European Union).

I would like to thanks to Prof. DSc. Guennadi Kouzaev for his excellent supervisory role.

I would like to thanks to Prof. PhD Grisha Spasov for his support during the last application and received donation from Xilinx University Program (XUP).

REFERENCES

1. Kiray, V. and Jambulov, A. (2009). FPGA based digital electronics education and a simulator core design for a/d communication. *Academic Journals*, vol. 4, no. 11, pp. 1292 - 1301.
2. Kostadinov, A. (2009). Using of a hardware description language in the digital design subject. *Journal of the Technical University – Sofia, Plovdiv branch "Fundamental Sciences and Applications"*, vol. 14, no. 1, pp. 421-426.
3. Kostadinov, A. (2009). Using of embedded logic analyzer in laboratory exercises of testing and diagnostics subject. *Journal of the Technical University – Sofia, Plovdiv branch, "Fundamental Sciences and Applications"*, vol. 14, no 1, pp. 427-432.
4. Kostadinov, A. (2009). Using HDL translators in VLSI design laboratory exercises. *Annual Journal of Electronics*, vol. 3, no. 2, pp. 56-59.
5. Kostadinov, A. (2013). Field-Programmable-Logic device education, *Annual Journal of Electronics*, vol. 7, pp. 46-49.
6. Kouzaev, G. (written together with Kostadinov A.). (2013) *Lecture notes in electrical engineering 169, Applications of advanced electromagnetic, Components and systems, Predicate logic, gates and processor for topologically modulated signals (TMS)*, pp. 451 – 460, Springer, Germany.
7. Kumar, A., Fernando, S. and Panicker, R. (2013). Project-Based Learning in Embedded Systems Education Using an FPGA Platform. *IEEE Transaction on Education,* vol. 56, no. 4, pp. 407 - 415.
8. Mahmoodi, H. at al. (2012). Hands-on Teaching of Embedded Systems Design Using FPGA-Based tPad Development Kit. Interdisciplinary Engineering Design Education Conference, Santa Clara, USA, pp. 1 - 6.
9. Nakano, K. and Ito, Y. (2008). Processor, Assembler, and Compiler Design Education Using an FPGA. 14th IEEE International Conference on Parallel and Distributed Systems, Melbourne, Victoria, Australia, pp. 723 - 728.
10. Petrescu, I., Păvăloiu, I. and Drăgoi, G. (2014). Digital Logic Introduction Using FPGAs. The 6th International Conference Edu World, Pitesti , Romania, pp. 1507 - 1513.
11. Zeyad, A., Sulaiman, N. and Hamidon M. (2009). FPGA-based Implementation of Digital Logic Design using Altera DE2 Board. *International Journal of Computer Science and Network Security,* vol. 9, no. 8, pp. 186 - 194.
12. https://www.xilinx.com/support/university.html
13. http://www.alari.ch/
14. https://www.altera.com/support/training/university/overview.html
15. https://www.altera.com/en_US/pdfs/literature/ug/ug_altpll.pdf
16. https://www.ipo.gov.uk/p-ipsum/Case/PublicationNumber/GB2508162

Authors' contacts
Organization: Technical University - Sofia, Plovdiv branch
Address: 25 Tsanko Diustabanov Str.
Phone: +359 32 659 726
E-mail: kostadat@tu-plovdiv.bg

© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 23, 2017

IMPLEMENTATION OF SOFTWARE ARCHITECTURE FOR COMPLEX ANALYSIS OF SOLAR CORONA IMAGES

DIMITAR GARNEVSKI

Abstract: *A study of the processes in the nearest star - the Sun is important because of its influence on space weather and life of the planet as a whole. As in any scientific fields research in the solar corona require specialized software for processing the received data. This publication will investigate aspects of the architecture and the way of implementation of software for processing and analyzing images of the solar corona and in particular data retrieval for CME (Coronal Mass Ejection).*

Key words: *image processing, filters, parallel computing, solar corona, OpenCL, file formats, meta-data*

1. Introduction

In process of analysis of Solar corona and especially research of events and processes in magnetic field scientists uses special software which meets their requirements. Common tasks can be divided in few groups like initial discover of the event (CME - Coronal Mass Ejection), estimation of parameters of the event, additional classification of the CME, tracing of changes in the CME in time. Software must also present information to the scientists in proper format for additional processing.

2. Requirements

We can be define following requirements which must be fulfilled by the implemented software

- defining model of processed data (solar corona images and meta-data)
- defining states of images quality estimation algorithms
- implementation of adaptive model of input images pre-processing depending on images parameters
- implementation of pre-processing of the images like noise filtering etc.
- implementation of common processing algorithms stages
- visualization of the result and output data formats

3. Stages of image processing

Relying on initial requirements we can define following stages of solar corona image processing in developed software. The diagram at Fig.2 presents different stages of processing of image and flow of data between steps.

3.1. Pre-Processing

One of the additional steps which can be implemented and can improve results of image processing is quality estimation of input images and their initial pre-processing. Although currently significant amount of solar corona images which captured from Earth surface (capturing Solar corona with coronagraph with CCD or CMOS technology camera) has parameters which in significant amount cases meet requirements of used algorithms for image analysis and processing. When we use images received from cosmic probes cannot always guarantee good quality of input images. If we leave aside the parameters such as number of bits per pixel and levels of compression of images, most modern tools are able to provide color or black and white images with dimensions at least 1024 x 1024 pixels, it provides a starting point for further processing.

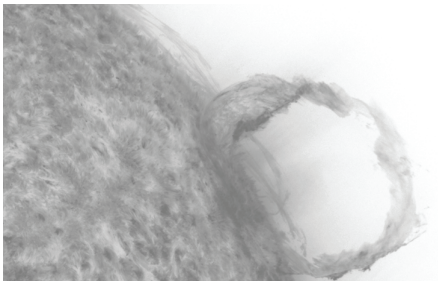


Fig. 1. *Prominence sample image*

 ISSN 1310 - 8271

Initial image quality estimation include estimation of noise level and level of shining from camera lenses. Noise level estimation is performed with stochastic approach.

Undesirable shining at the edges of the image can be removed by estimation of these pixel values and following normalization against average values available in entire image.

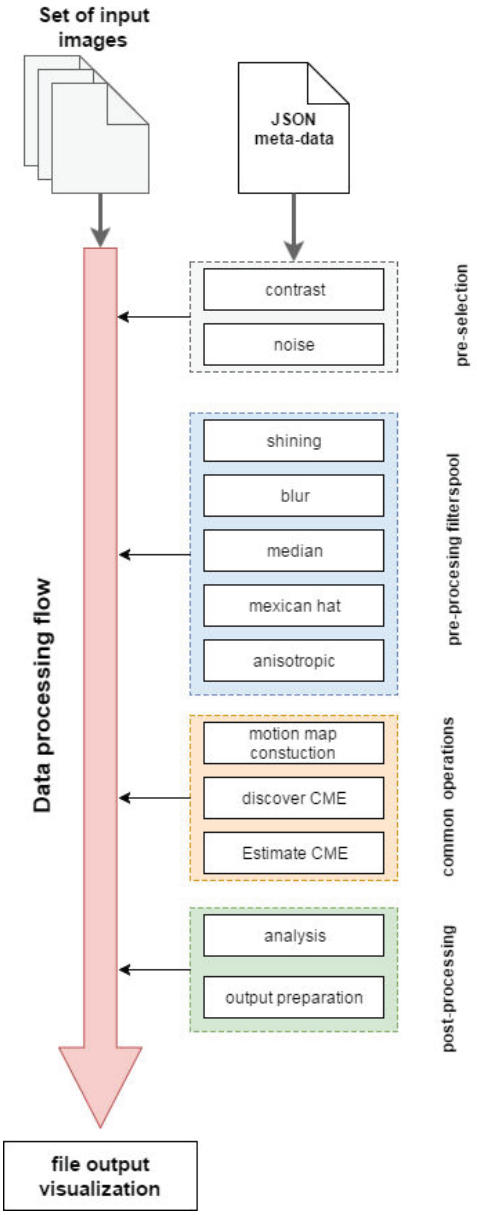


Fig. 2. *Data processing flow*

Noise removing is performed via filtering input images with one or combination of multiple filters like Mexican hat, median, anisotropic filter, etc. Common task of this operation is removing of pixel values which are out-of desired range without damaging structure of captured objects (lines, edges). In this situation anisotropic filtering has some benefits against other filters like blur because anisotropic filtering blurs image without damaging edges between objects in image [1][2][6][8].

3.2. Meta-data

For the purposes of the processing and in particular tracking changes over time are necessary metadata to images (date and time of capture, necessary adjustments, etc.). The date and time can both be obtained from the file or from meta-data (in cases when the data file does not correspond to actual shooting time). In implemented application when as input the data files FITS time are obtained from the file. In cases when using JPEG and processed series of images to store metadata using text file containing JSON (JavaScript Object Notation) with the necessary meta-data (in the basic version it is a list of files for processing and shooting time).

Format used contains data preprocessing:

- image contrast requirements - min and max values
- noise level
- possibility for preprocessing: pre-filter: shining removing, blur, median, Mexican hat, anisotropic (2 types with parameters)
- motion map stage
- output dir
- output file format

Fragment of a JSON file metadata is presented below:

```
{
  "inputDir": "C\\TestFilesInput",
  "inputFiles": [
    {
      "file": "test1.jpg",
      "timestamp": "2017-02-02 20:23:25",
      "preprocess": "true"
    },
    {
      "file": "test2.jpg",
      "timestamp": "2017-02-02 21:24:25",
      "preprocess": "true"
    }
  ],
  "preProcess": [
    {
      "filter": "glow",
      "mask": 3
    },
    {
      "filter": "median",
      "mask": 3
    },
    {
      "filter": "mexican",
      "mask": 3
    }
  ],
```

```
"process": [
  {
    "step": "motionmap"
  },
  {
    "step": "discover"
  },
  {
    "step": "cme_params"
  }
],
"postProcess": [
  {
    "step": "image",
    "format": "jpg"
  },
  {
    "step": "video",
    "format": "avi"
  },
  {
    "step": "cme_params",
    "format": "csv"
  }
],
"outputDir": "C\\TestFilesOutput"
}
```


Fig. 3. *JSON file with metadata*

3.3. Image analysis

Image processing core of the application include algorithms like generation of motion map, discovering of prominence at the boundary of occulting disk, estimation of prominence parameters like height, direction and their evaluation in time and estimation of connection between changes in Sun brightness and probability for solar prominence.

Generation of motion map is performed via algorithm described in [3][4][5]. In application is used its OpenCL implementation with platform independent optimizations.

Once built the map with the movement of the particles may be estimated height of the prominence and its direction.

Discovery the borders of occulting disk (artificial moon). In assessing the images and special about calculating the data for the height of the protuberance having matching the size of the solar disk image and the actual size of the Sun. In this case the work is performed with parts of images that include approximately 1/2 to 2/3 of the entire solar disk.

In this case, be carried out two operations, alignment of the images with respect to each other and detecting the boundary of the artificial moon. Artificial moon itself is depicted as a dark area defining the border is done by detecting the two endpoints that are located in border areas of the image and the central point, which is positioned between them.

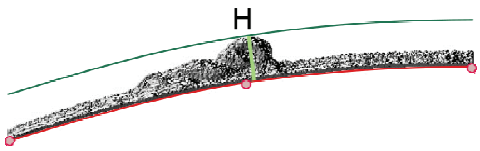


Fig. 4. *Generated motion map with prominence height estimation markers*

With these data points are build arc that defines the boundary of occulting disk and allows to calculate the correlation in size.

Estimation of connections between changes in Sun brightness and solar prominence is performed via processing of number of images in selected interval and discovering prominence in this interval.

For the implementation of the software used language C++, OpenCL environment for parallel processing, OpenCV library for image processing and number of additional libraries for FITS and JSON processing. For the implementation of key filtering functionality used OpenCL so calculations can be performed on CPU and GPGPU. Using of OpenCL provides needed flexibility in selection of different computational platforms. Also each filter can be defined as one separate kernel function in OpenCL which provides standardized approach in filter execution and opportunity for filter replacement.

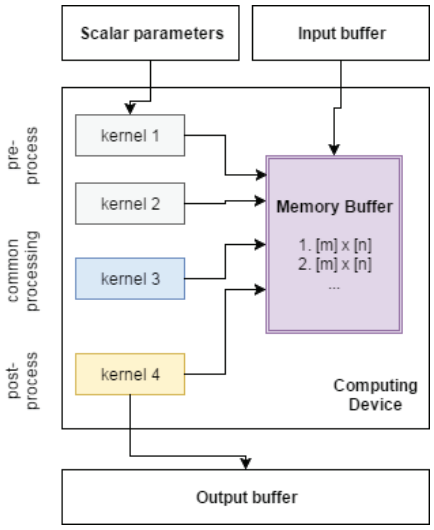


Fig. 5. *Kernel processing model*

Additionally be can perform changes in filter algorithm (kernel function implemented in OpenCL language) during development process without performing changes in entire application and we don't recompile it.

3.4. Post-processing

Post-processing of results include generation of out data in human readable format or format which can be used in another application. Developed application creates output in common images formats like JPEG or FITS (Flexible Image Transport System) [9] and also generates video based on frame sequences. For numerical data we use CSV (Comma-Separated Values) [11] format which is human readable and can be imported in most of data processing software like Microsoft Excel, it's open-source alternatives or other applications. Basic example of generated CSV is shown at image.

Table 1. *Sample CSV data format generated by the application*

1	2017-02-02 20:23:25	2017-02-02 21:24:25	128	46720
2	2017-02-02 20:24:25	2017-02-02 21:27:25	134	48910

CSV data format generated by the application is shown in Table 1. Individual columns are separated by the comma or tab-space depending on preferences. Shown in Table 1 file contains information amount measured high of discovered coronal mass ejection (CME). First column represents sequence number, 2nd and 3rd start and end time of measured interval (filled against input meta-data or timestamp of the input images), 4rt columns represent height of the CME in pixels and the last columns shows height in km, which was calculated depending of scale factor (filled in meta-data) and image resolution.

3.5. Result visualization

As we mentioned in 3.4 application generates images depending on desired result. So they can be viewed by external image viewers. Soft has embedded viewer which uses OpenCV [7] image processing library for additional image processing before visualization. This type of additional processing includes conversion between color spaces, basic filtering or adding of geometrical objects (lines, arc, markers, etc.) to the image.

For in example image at the Fig.4 are visible object (arcs and markers) which has been added with usage of OpenCV image processing library.

Generation of video files (.avi format) is also performed via OpenCV library.

4. Conclustions

Adaptation of the application to work with web-based storage systems and organization of images. This will include implementation of API which can be used by external systems.

Implementation of additional pre-processing and common filtering algorithms.

Integration of distributed computations like MPI (Message Passing Interface)

REFERENCES

1. Al Bovik. (2004). *Handbook of image and video processing*, Academic Press. Canada.
2. Rafael C. Gonzalez Richard E. Woods (2002) *Digital Image Processing*, 2nd Edition, Prentice Hall
3. Petya Pavlova, Elena Duncheva, Kostadinka Koleva (2010). Method for tracking and mapping a motion based on images of the solar corona. *Proceedings of the VII Bulgarian-Serbian Astronomical Conference (VII BSAC)*, Chepelare, Bulgaria.
4. Petya Pavlova, Kristina Staneva (2005). Dependence of the Colour in Computer Models on the Colour-reproducing Signals Discretization. *E+E*.
5. Petya Pavlova, Dimitar Garnevski, Kostadinka Koleva (2016). Optimization of a motion tracking and mapping method based on images of the solar corona. *Bulgarian Astronomical Journal, Volume 24*
6. S. G. Hoggar, *Mathematics of digital images. Creation, Compression, Restoration, Recognition*, Cambridge
7. http://opencv.org/, OpenCV (Open Computer Vision) Reference
8. Frery A. C, Perciano T. (2013*). Introduction to Image Processing Using R*, Chapter 2, Springer
9. https://fits.gsfc.nasa.gov/, The FITS Support Office, NASA/GSFC
10. http://www.json.org/, ECMA-404 The JSON Data Interchange Standard.
11. https://tools.ietf.org/html/rfc4180, Common Format and MIME Type for Comma-Separated Values (CSV) Files
12. Markus J. Aschwanden, A Code for Automated Tracing of Coronal Loops Approaching Visual Perception, *Springer Science+Business Media B.V. 2010*

Dimitar Garnevski
Department of Electrical Engineering
Technical University–Sofia, Branch Plovdiv
25 Tsanko Diustabanov St.
4000 Plovdiv
E-mail: garnevsky_dm@abv.bg

© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 23, 2017

ADVANCED ALGORITHM FOR CANNY EDGE DETECTION GRADIENT MAGNITUDE COMPUTATION TO BE IMPLEMENTED ON FPGA

DIMITRE KROMICHEV

Abstract: *The paper presents a Pythagorean theorem computing algorithm aimed at simultaneously satisfying both the requirement for mathematical exactness and the demand for maximum speed of gradient magnitude computations in the FPGA orientated Canny edge detection realization. Proposed are several criteria to be satisfied for a computational algorithm to be labelled as advanced. The focus is on advantageously utilizing the FPGA integer arithmetic capabilities for providing a hundred per cent accuracy of square root calculations at optimal speed.*

Key words: *Canny edge detection, gradient magnitude, FPGA, accuracy, speed, algorithm, square root, interval*

1. Introduction

Canny edge detection is a low level digital image processing technology. It was intended to be used in software. Field Programmable Gate Array (FPGA), as a representative of the widely applicable programmable logic, has been increasing its market share over the last two decades. The two capital requirements for FPGA-based hardware implementation of Canny edge detection are accuracy and speed, the former being a function of the mathematical exactness of computational algorithms within each of the modules framed by the maximun utilization of FPGA capabilities. Speed depends upon several factors, among them of tremendous importance being the algorithms calculations are realized through. FPGA functionalities and characteristics should be taken into account by the computational algorithms in terms of utilizing those favourable to fast executions whereas avoiding operations and methods serving as bottlenecks. Speed is worthy of being set as an accomplishable goal only on the peremptory platform defined as total accuracy of calculations and veracity of detected contours. Thus, in FPGA-based Canny every single algorithm requires a multifaceted approach targeting the results' being reliable as well as feasible.

The implementation of Pythagoras aimed at computing the gradient magnitude in the FPGA orientated Canny which is described in the literature [7][8][9][10] relies on the approximation

$$|G_M| = |G_X| + |G_Y| \qquad (1)$$

where Gx and Gy are the x- and y-gradients. Although fast, this technique is very inaccurate. So far, in the literature there has been no algorithm addressing both the requirement for mathematical exactness and the demand for speed.

The objective of this paper is to present an advanced gradient magnitude algorithm aimed at speeding up the FPGA-based Canny on the basis of providing a hundred per cent mathematical accuracy of results. The task is to describe in detail the sequence of steps, to thoroughly analyze the computational reliabilty, to expose the characteristics, and to point out the applicability of the proposed algorithm. taking into account the FPGA functionalities. Essential criteria to be satisfied for an algorithm to be labelled as advanced are set forth. The targeted hardware is Altera FPGAs. Relevant to the analyses and conclusions arrived at in this paper are only gray-scale images.

2. Criteria for advanced algorithm

Canny edge detection includes these consecutive modules: Gaussian smoothing, computing the orthogonal gradients, computing

 ISSN 1310 - 8271

gradient magnitude and direction, non-maximum suppression, hysteresis thresholding. Labelling an algorithm should reflect its essential computational characteristics, and rely upon definite criteria, the latter in the FPGA-based Canny being as follows.

1) Satisfying the peremptory demand for a hundred per cent mathematical accuracy of results. This is the solid guarantee for veracity and reliability of detected contours. In this respect, gradient magnitde is of tremendous importance for the overall plausibility of Canny computations in terms of procuring the values to be compared in the non-maximum suppression module, thus directly and immediately impacting one of the few most significant and applicabale qualities of contour – its precise localization.

2) Serving as a tangible tool for optimally enhancing the speed parameter of Canny computations on FPGA. This criterion has a quantitatvely measured characteristic presented by the smallest possible amount of clock cycles requred for executing the surveyed algorithm. Two particular aspects are to be taken into account here. Speed is a function of the appropriate and expedient organization of calculations in the algorithm being analyzed. On the other hand, fast and computationally sofisticated as it can seem, any algorithm should comply with the integer arithmetic peculiarities and benefit from the advantageous functionalities of the FPGA hardware implementation. Thus, the more mathematically complicated the algorithm (such as Pythagoras), the heavier the speed dependence on the programmable logic's computation accelerating capabilities.

3) Boosting the efficiency of pipelining, the latter being indispensable to the FPGA–based Canny edge detection implementation focused on speed. This criterion proves to be of particular importance in case of algorithms being executed in parallel within the main flow of computations on FPGA, gradient magnitude and gradient direction algorithms being indispensable here. Taking into account that calculating the local maximum requires a square neighbourhood of pixels, the calculated gradient magnitude values have to be ready for use at a rate that allows for a comparison of the central pixel with the adjacent pixels along the axis of gradient direction without any delays.

4) Minimizing the impact of input data width. Confining the input data to 8 bits is of significant importance in the Canny modules as a tool for diminishing delays. On Altera FPGAs, the fastest integer arithmetic execution in terms of clock cycles is guaranteed for 8-bit values [1][2][3][4][5]. In respect to this criterion, an advanced algorithm should optimally commence computations with integer arithmetic which takes exactly one clock cycle to execute, whatever the width of input values. Fulfillment of that requirement is well manifested for the Pythagoras, the FPGA implementation of the algorithm peremptorily starting with multiplication.

3. The proposed algorithm

Mathematically, gradient magnitude G_M is computed using

$$G_M = \sqrt{G_x^2 + G_y^2}, \quad (2)$$

where

Gx and *Gy* are the *x*- and *y*-gradients.

Altera provides a specialized integer square root function – ALTSQRT [5]. Although the 8-bit gray scale image values are definitely advantageous to the speed parameter, the maximum clock rate the function can be executed at [5] is not campatible with the accomplishment of the speed optimization goal – calculating the correct result requires nine clock cycles. Thus, an optimized approach is needed.

3.1. Input data

Gx and *Gy* are computed in the Canny orthogonal gradients module by employing two approaches:

1) Exact mathematics. The equations used to accurately calculate the *x* gradient and the *y* gradient are:

$$G_{\dot{x}} = \frac{(N_3 + 2N_4 + N_5) - (N_1 + 2N_8 + N_7)}{4} \quad (3)$$

$$G_y = \frac{(N_7 + 2N_6 + N_5) - (N_1 + 2N_2 + N_3)}{4}, \quad (4)$$

where

N1..N8 are the neighboring pixels in a 3x3 neighborhood.

The values calculated here are within the intetval [-255,255].

2) Approximation.

$$Gx = [(N_3 + 2N_4 + N_5) - (N_1 + 2N_8 + N_7)] \quad (5)$$
$$Gy = [(N_7 + 2N_6 + N_5) - (N_1 + 2N_2 + N_3)], \quad (6)$$

where

N1..N8 are the neighboring pixels in a 3x3 neighborhood.

With division by the power of 2 being dropped out, the values calculated here are within the intetval [-1020,1020]. Consequently, proper scaling is required for all the results to fit within the maximum gray scale image pixel range.

3. 2. Advanced gradient magnitude computation

On FPGA, computing Pythagoras starts with multiplying $|Gx|$ by $|Gx|$, and $|Gy|$ by $|Gy|$. The largesr positive pixel value being 255, the results of both multiplications are within the interval [0, 65025].

In the next step, the multiplication results are added, and consequently the numbers to be square rooted are within [0, 130050].

Thus, all the possible values at the input of the square root function are 130051, And all the possible values at its output are 2^8, taking into account that only numbers in the interval [0, 255] are relevant for pixels in the gray scale matrix. Consequently, the exact mathematical results of the square root function can be determined by 256 closed intervals, each of them representing a single integer from the interval [0,255]. These intervals contain all the 130051 values calculated by adding the squared gradients.

Defining the smallest and the largest in each interval is based on the fact that the difference between squares of two consecutive integers is equal to the smaller integer multiplied by 2, and then 1 is added to the result. Thus, the smallest of all the consecutive values contained in an interval is determined through:

1) squaring the gray scale image pixel value represented by this particular interval;

2) the gray scale image pixel value represented by this particular interval is decremented, and the result is subtracted from the result computed in 1) .

The largest of all the consecutive values contained in an interval is determined through:

1) squaring the gray scale image pixel value represented by this particular interval;

2) the gray scale image pixel value represented by this particular interval is subtracted from the result computed in 1) . The only exception is the value in the interval representing 255 – here the largest possible number is 130050.

The comparisons of all values for computing the gradient magnitude are executed simultaneously on FPGA, utilizing one of the most significant capabilities of hardware implementation – parallel computations. Thus the proposed algorithm satisfies the criteria for being labelled as advanced. It guarantees a hundred per cent mathematical accuracy of gradient magnitude computation. It also utilizes the fastest integer arithmetic on Altera FPGAs. The presented algorithm is not susceptible to any input data width driven delays. Being that fast, it is capable of increasing the efficiency of pipelining in the FPGA-based Canny.

4. Conclusions

The paper focuses on presenting and analyzing a new gradient magnitude algorithm addressing the goal of speed enhanced and mathematically exact Canny edge detection computations intended to be implememted on FPGA. Essential criteria to be satisfied for an algorithm to be labelled as advanced are defined. A thorough description of the algorithm's peculiarities is set forth. The algorithm is scrutinized in terms of its applicability, computational reliability and speed characteristics. The technological capabilities of efficiently avoiding speed eroding computational approaches and satisfying the peremptory demand for mathematical accuracy and plausibility of results pinpoints the feasibility of the proposed algorithm as a tangible tool for enhancing Canny's performance on FPGA.

REFERENCES

1. Altera Corporation. Cyclone IV Device Handbook. 2014, Volume 1
2. Altera Corporation. Cyclone V Device Handbook. 2014, Volume 1
3. Altera Corporation. Stratix III Device Handbook. 2014, Volume 1
4. Altera Corporation. Stratix IV Device Handbook. 2014, Volume 1
5. Altera Corporation. Integer Arithmetic IP Cores User Guide, 2014
1. Altera Corporation. Cyclone II Device Handbook, Volume 1, 2014
7. Chandrashekar N.S., K. R. Nataraj. Design and Implementation of a Modified Canny Edge Detector based on FPGA, *International Journal of Advanced Electrical and Electronics Engineering*, (IJAEEE). 2013, Vol.2, (1), pp. 17-21
8. Chandrashekar N.S., K. R. Nataraj. NMS and Thresholding Architecture used for FPGA based Canny Edge Detector for Area Optimization, *Proceeding of International*

Conference on Control, Communication and Power Engineering. 2013, pp. 80-84
9. Divya. D., P. S.. Sushma. FPGA Implementation of a Distributed Canny Edge Detector, *International Journal of Advanced Computational Engineering and Networking.* 2013, Vol. 1, (5), pp. 46-51
10. Shamlee V., Jeyamani.. A Split Canny Edge Detection: Algorithm and its FPGA Implementation. *International Journal of Science and Research* (IJSR). 2014, Vol. 3, pp. 1198 -1205

Plovdiv University Paisii Hilendarski
Plovdiv 4000
24 Tzar Asen Street
e-mail: dkromichev@yahoo.com

© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 23, 2017

APPROXIMATION IN THE FPGA BASED CANNY EDGE DETECTION COMPUTATIONS

DIMITRE KROMICHEV

Abstract: *The paper deals with the technology of approximation and its impact on the computational results in the FPGA based Canny edge detection modules. Presented are four general types of approximation and their essential functions and characteristics are exposed. The impact of approximation on the mathematical exactness of calculations in the Canny algorithm is scrutinized with respect to the peculiarities of hardware implementation and the targeted veracity of the mapped contours. Analyzed is the relevance of approximation in terms of speed vs. accuracy.*

Key words: *Canny, FPGA, approximation, type, algorithm, accuracy, speed*

1. Introduction

For decades now, Canny edge detection has gained a well deserved reputation of being an efficient and reliable low level digital image processing technology. Inasmuch as this precise yet complicated operator is orientated towards being used in software, its hardware implementation on the Field Programmable Gate Array (FPGA) is particularly difficult and requires that the computations be executed by strictly taking into account the capabilities and functionalities of that widely used representative of programmable logic. The goal to enhance performance while sparing the area of FPGA generaly involves certain precision decreasing approaches in the realization of the basic computational algorithms Canny relies upon. Thus, addressing the quality and plausibility of detected edges, the technique of purposely replacing exact mathematical values with numbers that are close demands an in depth analysis. In the FPGA based Canny, approximation has several aspects of utilization, each of them impacting the calculations in its own way and worthy of being duly assessed.

Described in the literature are different uses of approximation related to Canny edge detection on FPGA. In [10], the approximation is focused on computing the gradient magnitude in terms of replacing the square root function with the sum of positive values of x- and y-gradients, and is aimed at enhancing the speed performance. In [8], split Canny uses approximation gradient magnitude and direction computations for the purpose of increasing pipelining efficiency. In [6][3][9], distributed Canny attemps to spare the on-chip memory at the expense of approximation orientated calculations and maintaining constant throughput through performing the computations at block level. In [5], area optimized Canny is mainly focued on a new architecture of the non-maximum suppression and approximation framed hysteresis thresholding modules. In modified Canny [4] [7], the image is partitioned into q sub-images and each sub-image is further divided into p m×m blocks. Approximation is applied to gradient magnitude and direction module, as well as the calculation of high and low thresholds.

The objective of the paper is to thoroughly analyze the approximation patterns employed in the FPGA based Canny calculations. The task is to scrutinize the computational results which are close but do not coincide with the exact numbers obtained through using the standard mathematics in terms of: types, distribution by Canny's modules, attitude towards speed optimization, effect on the utilized FPGA area, influence on the precision of detected contours, impact on speed and accuracy relation with respect to FPGA's advanced capabilities. It is the Altera FPGAs that are referred to throughout this study. Relevant to the conducted analyses and conclusions arrived at are only gray scale images.

2. Types of approximation

Canny edge detection includes five modules: Gaussian smoothing, computing the orthogonal gradients, computing gradient magnitude and direction, non-maximum suppression, hysteresis

 ISSN 1310 - 8271

thresholding. With respect to its utilization and effect on the calculations in FPGA orientated realization of Canny, approximation is to be presented within four types.

2.1. Mandatory approximation

It is imposed by the requirements of the discrete domain in digital image processing. In that respect, actually two out of the five computational modules in the Canny algorithm are entirely based on approximation.

Any weighted average filter employed in the Gaussian smoothing module is a discrete approximation of the 2-dimensional function

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad , \quad (1)$$

where

G is the Gaussian mask weight with coordinates x and y,

σ is the standard deviation of Gaussian distribution; it determines the filter's size,

$\frac{1}{2\pi\sigma^2}$ is normalization constant.

On FPGA, the filter performs a ZxZ neighborhood operation (Z is an odd number and Z $\geq$ 3)

$$v(m,n) = \frac{1}{s} \sum_{k=-w}^{w} \sum_{l=-w}^{w} g(k,l)u(m+k,n+l) \quad (2)$$

where

u(m,n) is the input image,

v(m,n) is the filtered image,

$w = \frac{z-1}{2}$, w = {1,2,..},

g(k,l) is the Gaussian filter ,

s is the sum of all coefficients in the mask; the coefficients are integers based on the binominal series.

The Sobel approach for computing the orthogonal gradients is actually the result of a two stage approximation:

1) the central difference approximation

$$\Delta f(x) = [\, f(x+1) - f(x-1)\,] * 0.5 \quad (3)$$

to the derivative of a continuous function

$$f' = \frac{df}{dx} = \lim_{\Delta x \to 0} \frac{f(x) - f(x-\Delta x)}{\Delta x} \quad (4)$$

for Δx = 1 as the smallest possible value of Δx in the discrete domain of digital image pixels;

2) approximation to (3) for the purpose of defining two 1-D masks: -1 0 1 and -2 0 2. They are the platform of the two 2-D filters employed in the calculation of the vertical and horizontal gradients (Fig. 1.)

-1	0	1
-2	0	2
-1	0	1

Gx

-1	-2	-1
0	0	0
1	2	1

Gy

Fig. 1. Sobel filters for x and y

2. 2. Targeted approximation

Precise and reliable, Canny is computationally complicated, extensively demanding in terms of hardware implementation, and relatively slow algorithm. Thus, approximation is deliberately employed to address different aspects of performance being optimized.

2.2.1. Speed enhancing approximation

Accomplishing the goal of a fast FPGA based Canny is favoured by a set of approaches related to approximation.

1) Reducing and exploiting only the fastest integer arithmetic needed to fulfil a particular computational task. With respect to both the highest frequency and clock cycles taken for execution, the most speed advanced on FPGA are multiplication, addition and subtraction [1] [2]. In terms of speed, division is definitely a weak point. Using a divider by employing the Altera LPM_DIVIDE function [2] will immensely erode the speed parameter of Canny's algorithm execution. Consequently, the conventional integer division is typically omitted. Thus, in computing the orthogonal gradients the normalization division is dropped out, the gradient magnitude calculation is modified to rely on a single addition, and the gradient direction value is the result of utilizing only the signs of the vertical and horizontal gradients.

2) Avoiding those Altera LPM functions [1] [2] which, despite being handy, take more than two clock cycles to output a correct result and lead to speed eroding delays. The focus here is on omitting the square root function for the gradient magnitude and the division in the gradient direction computations.

3) Keeping the sequential logic input data width optimally within 8 bits for executing the FPGA integer arithmetic at the highest frequencies [1] [2]. The impacted values are those of the vertical and horizontal gradients which need to be confined to the interval [-255, 255].
4) Replacing a slow computational procedure with values determined in advance. To this point, fixed values for the high and low thresholds tend to be used instead of assessing the entire image statistics in the non-maximum suppression module.

2.2.2. FPGA hardware sparing approximation

Typically, the less the utilized FPGA area, the more efficient the Canny implementation. This aspect of approximation in the FPGA based Canny, despite sometimes being colateral and far from particularly intended, is important in two ways. It is an implicit indicator of the efficiency of a computational mechanism, thus presenting tangible measurement for the practical applicability of any technique or algorithm aimed at optimizing the organization of Canny calculations to be implemented on FPGA. The hardware sparing approximation can also serve as guidelines for assessing the extent a proposed computational approach is functionally advanced.

This approximation type is present mainly in the calculations of the gradient magnitude and direction wherein, instead of the very expensive hardware realization of exact mathematical algorithms, the results are achieved through relying on addition and comparison. It is also appropriate for sparing the limited in quantity on-chip memory in terms of using predefined low and high threshold values to avoid computations involving the entire image statistics after executing the non-maximum suppression.

2.3. Complete approximation

This approximation type is focused on replacing a whole computational algorithm and its territory is the gradient magnitude and direction module. Mathematically, gradient magnitude G_M is calculated using

$$G_M = \sqrt{G_x^2 + G_y^2} \tag{5}$$

where

Gx and Gy are the x- and y-gradients.

In the FPGA based Canny, Pythagoras is generally replaced with

$$|G_M| = |G_X| + |G_Y| . \tag{6}$$

The accurate calculation of gradient direction G_D is

$$G_D = \tan^{-1}\left(\frac{G_y}{G_x}\right) \tag{7}$$

where

Gx and Gy are the x- and y-gradients; $G_y \in [-255,255]$; $G_x \in [-255,255]$.

Calculating the gradient direction by applying expression (7) has the following inauspicious aspects:
1) The FPGA implementation is computationally expensive, and therefore leads to substantial delays.
2) The divisor G_X can be equal to 0. In that case calculating the gradient direction is practically irrelevant, and, consequently, additional computational conditions should be introduced to ensure the correctness of results.

For hardware implementation, (7) is typically replaced with the sign relations of the gradients.

$$\begin{array}{l} Gy > 0 \ \ \& \ \ Gx > 0 \\ Gy < 0 \ \ \& \ \ Gx < 0 \\ Gy > 0 \ \ \& \ \ Gx < 0 \\ Gy < 0 \ \ \& \ \ Gx > 0 \ \ . \end{array} \tag{8}$$

2. 4. Partial approximation

This approximation type deals with a limited portion of the algorithm's computations and has its indispensible application in the calculations for determining the angular values 0, 45, 90, 135 which define the four axes of gradient direction. Two values are utilized to optimally approximate the two angles of crucial importance to the mathematical accuracy of gradient direction calculation.

Taking into account that four intervals are needed to contain all results of computing the Gy and Gx ratio, two referece points are employed. One of them should represent an angle of 22.5°, and the other - 67.5°. In the integer domain of FPGA based Canny, the most appropriate approximation to angle 22.5° is the fraction

$$\frac{2}{5} \quad \left(\tan^{-1}\left(\frac{2}{5}\right) = 21.8014094°\right) \tag{9}$$

and to angle 67.5° - the fraction

$$\frac{5}{2} \left(\tan^{-1}\left(\frac{5}{2}\right) = 68.1985905\right) . \tag{10}$$

3. Assessment of approximation

3.1. Approximation in terms of impact on computational results

Inasmuch as the weighted average filter approximation has steadily defined its presence in the Gaussian smoothing, the appropriate selection of the fiter's size is of significant importance as a tool to compensate for the omission of not to be underestimated subtleties imposed by the discrete domain of digital image processing. Thus, taking into account the numerical characteristics of the bell shaped curve, the most suitable variants prove to be sizes 5x5 and 7x7. With respect to the fact that the larger the size the worse the localization, and taking into account the amount of computations required for the hardware implementation, the optimal option is filter of size 5x5.

As gradient filters, the two Sobel matrices have positive and negative coefficients. If in a 3x3 neighborhood C(*x,y*) is the central pixel, and the numbers of neighboring pixels are as shown (Fig. 2.) ,

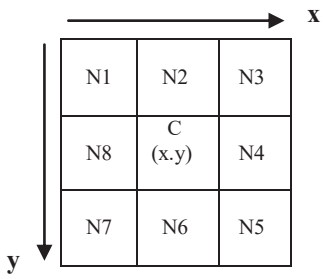


Fig. 2. 3x3 neighborhood pixels

the exact equations to calculate the *x* gradient and the *y* gradient for C(*x,y*) are:

$$G_{\dot{x}} = \frac{(N_3 + 2N_4 + N_5) - (N_1 + 2N_8 + N_7)}{4} \quad (11)$$

$$G_y = \frac{(N_7 + 2N_6 + N_5) - (N_1 + 2N_2 + N_3)}{4} \quad . \quad (12)$$

On FPGA, the approximation in computing the orthogonal gradients practically reduces the necessary integer arithmetic to addition and subtraction with division by the power of 2 being dropped out:

$$Gx = [(N_3 + 2N_4 + N_5) - (N_1 + 2N_8 + N_7)] \quad (13)$$

$$Gy = [(N_7 + 2N_6 + N_5) - (N_1 + 2N_2 + N_3)] \quad . \quad (14)$$

This approach presents two flaws. On the one hand, omitting division leads to lack of additional averaging, thus impeding a certain extra noise suppression to be applied to the image. On the other hand, taking into account that computing the gradient magnitude refers to the Pythagorean theorem, and magnitude values larger than 255 are irrelevant to the calculations of local maxima in the next Canny module, all the results equal or exceeding 2^8 are scaled down to the appropriate 255. Consequently, the division's being omitted requires that all the results larger than 255 should be cut down to fit the maximum value of a gray scale image pixel, and this entails serious disproportionality among this module's output results.

The gradient magnitude and direction module is of capital importance for the quality of detected contours. In computing the gradient magnitude, replacing (5) with (6) leads to calculated numbers which differ from the exact mathematical results in the range of 0.2-30 per cent. In terms of quantity, this lack of precision is irregularly distributed across the calculated gradient magnitudes, and any computational mechanism aimed at offsetting the inaccuracy is practically irrelevant. The largest differences are calculated for pixel values that are among the most frequently to be met in a gray scale image. In view of the fact that it is the gradient magnitude results that are compared in the non-maximum suppression module to perform edge thinning, the overall effect is poor localization and lack of plausibility, thus resulting in detected contour's getting unreliable for production line applications wherein taking correct measurements is of crucial importance. Depending only on (8) to compute gradient direction instead of (7) halves down the accuracy by reducing the required intervals from 8 to 4. The consequence is decreased correctness in the results of the algorithm involved in determining the local maxima, with eroded precision and gaps in the mapped contours.

The values calculated for the gradient magnitude and direction are straightforwardly used in the non-maximum suppression module. Utilizing (13), (14), and (6) can result in computing values exceeding 2^8 As long as only comparison of values limited of up to 255 is relevanly employed, keeping within 8 bits has no alternative.

Convenient as it is, utilizing two static values instead of a costly algorithm employing the entire image statistics to compute the high and low

thresholds results in both omitting real edges and defining noise as contours, the latter being capable of considerably compromising the output of Canny in terms of veracity, particularly with respect to uses orientated towards machine-to-machine interface.

Inasmuch as every pixel in the processed image is surrounded by 8 other pixels, the approximation in (9) and (10) is actually characterized by accuracy that is enough to provide computational results commensurable to those achieved through utilizing the conventional mathematical expression in (7).

3.2. Approximation in terms of speed vs. accuracy

Overtly or in the background, the analysis of approximation and its impact on the FPGA based Canny calculations eventually comes down to the speed vs. accuracy model. In this respect, there are two important facts.

Computationally expensive and slow as Canny is, addressing the accomplishment of speed optimization goal should peremptorily focus the proposed approaches on the framework of detected contours' being optimally plausible. Achieving reliability of the edge detected image especially for more demanding production line applications is feasible on the basis of using mathematically accurate algorithms to realize the calculations in the Canny modules. This being taken into account, approximation is a tangible technique as long as it does not violate the pattern of exactness and veracity at each successive stage in the flow of computations. Approximation for speed enhancement purposes which functions as a hurdle to the accuracy of results is unacceptable.

Even though being hard to implement on FPGA, Canny can afford avoiding approximation that compromises accuracy through utilizing the programmable logic's typical characteristic - executing computations in parallel. Whatever the complexity of an algorithm in a Canny's module, the simultaneity of calculations is quite an efficient tool to accelerate calculations without having to rely on accuracy decresing approximation. The additional requirement to be satisfied here is for the parallel computations to utilize the fastest integer arithmetic on FPGA.

4. Conclusions

Presented in this paper are four general types of approximation used in the FPGA based Canny edge detection computations: mandatory, targeted, complete, and partial. Mandatory approximation is imposed by the requirements of the discrete domain in digital image processing. In that respect, Gaussian smoothing and orthogonal gradients are entirely based on approximation. Accomplishing the goal of a fast FPGA based Canny is favoured by a set of approaches related to approximation. Complete approximation is focused on replacing a whole computational algorithm and its territory is the gradient magnitude and direction module. Partial approximation deals with a limited portion of the algorithm's computations and has its indispensible application in the calculations for determining the gradient direction. Two values are utilized to optimally approximate the two angles of crucial importance to the mathematical accuracy of gradient direction Each approximation type is analyzed and assessed in terms of: Canny's module it is used in, characteristics, expediency, departure from the mathematical exactness of results, impact on the reliability of calculations, relevance to the quality of detected contours. The attitude of approximation towards speed and accuracy relation with respect to FPGA's advanced functionalities is scrutinized.

REFERENCES

1. Altera Corporation. Cyclone II Device Handbook, Volume 1, 2014
2. Altera Corporation. Integer Arithmetic IP Cores User Guide, 2014
3. Aravindh G., C. S. Manikandababu, Algorithm and Implementation of Distributed Canny Edge Detector on FPGA, *ARPN Journal of Engineering and Applied Sciences,* Vol. 10 (7), 2015, pp. 3208-3216
4. Chandrashekar N .S. and Nataraj K. R., Design And Implementation of a Modified Canny Edge Detector based on FPGA, *International Journal of Advanced Electrical and Electronics Engineering, (IJAEEE),* Vol.2, (1), pp. 17-21
5. Chandrashekar N.S. and Nataraj K.R., NMS and Thresholding Architecture used for FPGA based Canny Edge Detector for Area Optimization, *Proceeding of International Conference on Control, Communication and Power Engineering,* 2013, pp. 80-84
6. Divya. D. and Sushma P. S., FPGA Implementation of a Distributed Canny Edge Detector, *International Journal of Advanced Computational Engineering and Networking,* Vol. 1, (5), 2013, pp. 46-51
7. Ramgundewar, Pallavi, S. P. Hingway, K. Mankar. Design of Modified Canny Edge Detector based on FPGA for Portable Device, *Journal of The International Association of Advanced Technology and*

Science. Vol. 16 (2), 2015, pp. 210-214
8. Shamlee V. and Jeyamani, A Split Canny Edge Detection: Algorithm and its FPGA Implementation, *International Journal of Science and Research (IJSR),* Vol. 3 (12), 2015, pp. 1198-1205
9. Veeranagoudapatil, Chitra Prabhu. Distributed Canny Edge Detector: Algorithm & FPGA Implementation, *Internationa Journal for Research in Applied Science & Engineering Technology* (IJRASET), Vol. 3 (5), 2015, pp. 586-588
10. Xianghua Hou, Honghai Liu, Welding Image Edge Detection and Identification Research Based on Canny Operator, *International Conference on Computer Science and Service System,* 10-13 Aug. 2012, pp. 602-610

Plovdiv University Paisii Hilendarski
Plovdiv 4000
24 Tzar Asen Street
e-mail: dkromichev@yahoo.com

© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 23, 2017

INTRUSION DETECTION AND PREVENTING SYSTEM

KOLYO RAYCHINOV, HRISTO VALCHANOV

Abstract: *The attacks against computer and network security are a real threat in the modern Internet world. Detection of attacks is a process of monitoring of computer systems or networks and analyzing them for signs of possible events that are violations or threats. Developing of intrusion detecting and preventing systems (IDPS) requires the availability of resources that are needed to be integrated within the protected network infrastructure. This paper presents the architecture and functionality of IDPS, which offers affordable, flexible and efficient solution for building such as systems in different fields.*

Key words: *Network Attacks, Intrusion Detecting and Preventing System, Raspberry Pi*

1. Introduction

Besides the advantages offered by the Internet technologies, they hide many risks that professionals in the IT field have to deal. Cybercrime is a growing criminal activity, and the convenience of fast connections to multiple resources connected to Internet makes hackers to invent increasingly clever and high-tech ways to exploit vulnerabilities in the systems. The attacks are a real threat for the information, which is a valuable resource in today's world, both for individuals and for businesses. The corporate sector often falls under similar attacks because of financial reasons - companies own resources and information that can be stolen or to speculate with them. For this reason the companies pay special attention on used hardware and software - they are usually of significantly higher class and price compared to widely used ones. There is need for effective, smart and cheap solutions for threat recognition and preventing.

This paper presents the basic architecture and functionality of intrusion detection and preventing system, which offers simple, flexible and efficient solution for building similar systems in different fields.

2. Attacks detection and prevention

Detection of attacks is the process of monitoring of events in computer systems or networks and analyzing them for signs of possible violations or threats of computer security policies [1]. The events may be due to many reasons such as malware (spyware and worms) which acquire unauthorized access to the system through the Internet, or intrusion by unauthorized people who are using incorrect privileges.

The system for detecting attacks (Intrusion Detection System - IDS) is software that automates the process of detecting the attacks. The system to prevent attacks (Intrusion Prevention System - IPS) is software that has all the capabilities of IDS and can take to deter possible incidents. IDS and IPS technologies offer many similar features and administrators can usually exclude options for the prevention of IPS products, forcing them to function as IDS. Today in cyber security is widely used the term "Intrusion Detection and Prevention Systems" (IDPS), which combines two technologies [2].

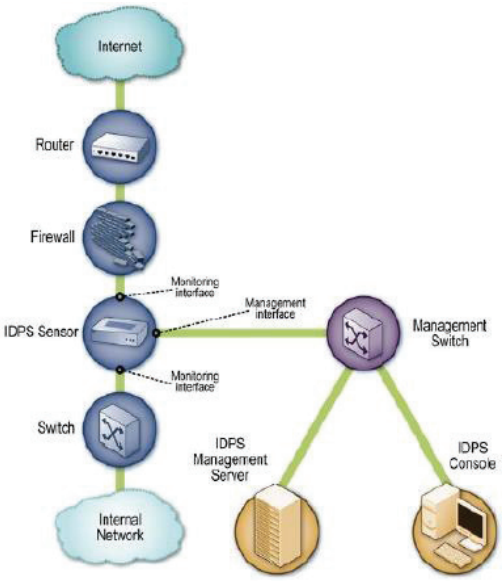


Fig. 1. *IPDS infrastructure*

 ISSN 1310 - 8271

The implementation of IDPS requires certain resources that need to be integrated within the protected network infrastructure (Figure 1). Typical components are as follows:

- Sensor or agent. They monitor and analyze activity. The term "sensor" is typical for the IDPS who monitor networks and the term "agent" - for those tracking only one device.
- Management Server. It is a device that receives information from sensors or agents and manages them.
- Database server. It is a repository of information regarding the events detected by sensors, agents and management servers.
- Console. This is a program that provides an interface for access to IDPS of users and administrators.

The building of such infrastructure requires substantial costs, both from financial and organizational point of view. A number of manufacturers offer complete solutions on the market, including special devices and software that are however acceptable only for large companies.

The IPDS technology uses several basic methods for detection of attacks. Anomaly based detection is the process of comparing definitions of what activity is considered normal for watched events and then to recognize large deviations. Such a system has profiles that represent the normal behavior of elements such as users, devices, network connections or applications. The profiles have been developed by monitoring the characteristics of a typical activity for a given period of time. A major problem in the generation of profiles is that it can be quite difficult in some cases to be made precise, because of complexity of the computer activity.

Detection based on state of the protocols is the process of comparing the pre-set common definitions of normal actions for each protocol's state with observed events to identify deviations. Unlike anomalies based detection that uses network or host-based profiles, this analysis relies on universal profiles developed by manufacturers that describe how each protocol should or should not be used. The main disadvantage of this type of analysis is that they require a great intensity of resources, because of the complexity and the additional load during the simultaneous tracking of multiple sessions.

A more efficient method is the detection based on the signature. This is the process of comparing the signature and the observed events to be identified a possible attack. This approach is very effective for detection of known attacks. In terms of implementation and use of resources, this is the simplest method because there is only a comparing the current element of activity as a package or a log entry with a list of signatures using strings compare operations.

3. Architecture of the system

The architecture of the presented system is modular (Figure 2). It comprises a multiple modules, each having a specific functionality. This allows easy modification of system functionality in future extensions.

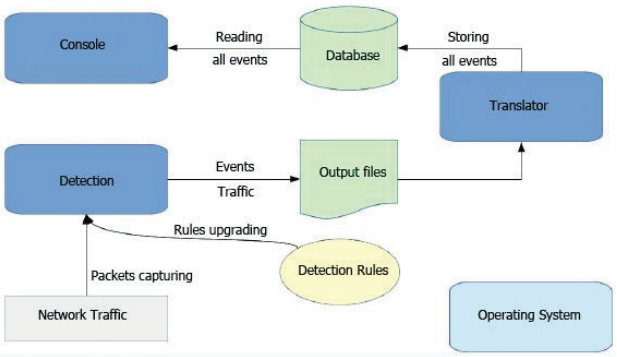


Fig. 2. *Architecture of IPDS*

The system consists of the following components:

- Operating System under which runs the software implementation.
- Software performing detection of attacks.
- Events translator.
- Console for showing the results of the analysis of the network traffic.

The main idea of the developed system is its easy integration into any network infrastructure. For this purpose is proposed a simple and efficient solution using a single board computer Raspberry Pi [3]. A characteristic feature of this system is its integrity - the concentration of all software components on the powerful hardware, at the same time with small size and low cost of ownership.

There are possible alternative configurations of the presented architecture. One configuration is a simplified version of that where IDPS records the output messages directly in the database, skipping the binary file and the translator. In this case the performance of the software is low, as it concentrates more functionality. In another architecture a file for direct analysis at a later stage is generated (for example, *csv* or *tcpdump*) and completely are absent translator, database and front-end interface. The proposed system takes account of the convenience and functionality it provides the graphical console for the analysis of the detected

threats. Thus, it is realized an architecture with high efficiency.

As an operating system is used Debian Linux based system - Raspbian [4] which is one of the officially supported according to the Raspberry Pi Foundation and provides opportunities to meet the requirements of current implementation.

Under the operating system is running another important component for the IDPS – the software for control of the traffic that passes through the device. For fulfilling this role is used Snort, which is a fundamental component in the network IPDS [5]. In this implementation the IDPS interacts with network traffic through a transparent bridge through which the traffic passes.

For greater efficiency of Snort, the data from the system are generated in binary file in *unified2* format. This file is processed by translator - it parses recorded events and particular traffic and redirects them to a database. The used translator is Barnayrd2, running parallel with Snort and redirecting new events from the binary file to the database in real time [6]. As a database management system is used MySQL [7].

Key for each signature-based IDPS are lists of rules according which the events are defined as having possible degree of threat or not. These signatures are stored by Snort in several types of local files. Because of the dynamics of the security systems, the information about each newly discovered vulnerability or malicious code must reach the protected systems as soon as possible. To satisfy this requirement, in the system is implemented a tool which automates the process. The company, which develops Snort, supports a repository with signatures for detection of attacks. The updates are performed by the tool PulledPork, which connects to the central repository, downloads the new rules and applies them in IDPS [8].

4. Experimental study and results

Testing of the presented system is done in a real network infrastructure (Figure 3).

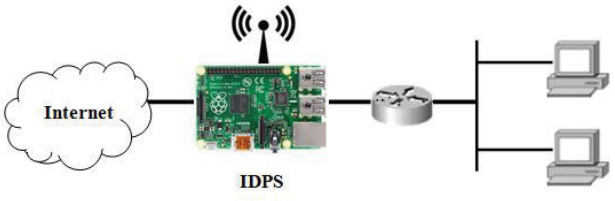


Fig. 3. *Integration of IPDS into a network*

An important issue is the management of the system. In order to avoid physical accessibility to Razpbery Pi (keyboard or console terminal) there is applied a solution by adding third WiFi network interface. The standard model of Raspberry Pi rev.2 has only one Ethernet interface, but for the realization of the system is required at least two (in the best case and three network communication interfaces). For this purpose, additional external USB modules are added - an USB-to-Ethernet adapter, an Ethernet card and a single USB dongle for wireless WiFi connection. Thus achieves the intended functionality via two interfaces through which to pass network traffic checked for problems and an interface to access the system for its management. Both the Ethernet interfaces have no IP addresses and are connected as a bridge running on the link layer of the OSI model, which is transparent to the upper layers.

The network system for detecting attacks is situated in a part of the network that carries traffic between internal and external to the organization networks. The Snort sensor must be behind the devices communicating through encrypted connection (VPN routers), so watched traffic is not encrypted. At the same time, it must be located behind a firewall that prevents any attempts to attack. Thus, in a security breach, the system will detect the incidence and will notify by messages the responsible personnel.

There are simulated Denial of Service (DoS) attacks. Large amount of requests (hundreds of thousands) from spoofed sources are generated for the shortest period of time. This results in flooding of network. The attack is implemented by sending only SYN packets, as the attack is directed to the router in the network.

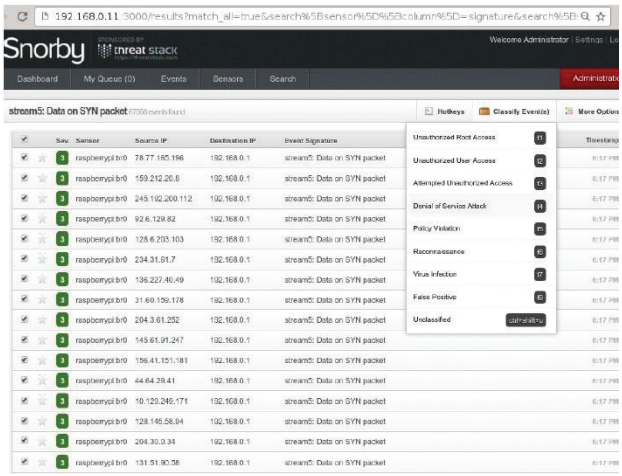


Fig. 4. *Events classification*

Over 65,000 messages (Figure 4) for unusual behavior in the network are generated. Its number continues to grow because the translator continues to read from the binary file and fill the database with new messages, although the simulated attack is over. The reason for detected by the system anomaly are data in SYN packets, intercepted by Stream preprocessor of Snort.

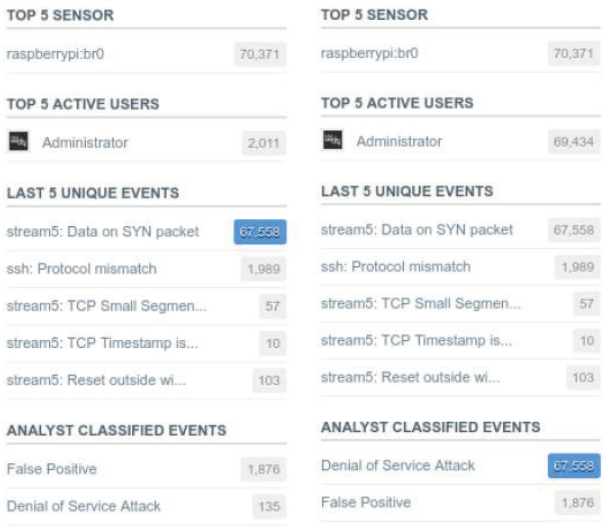


Fig. 5. *Events before and after classification*

After the detection of the events it is necessary to classify them. In this case, they are defined as DoS attack due to the amount of requests from multiple addresses without completed TCP dialogue. This classification is done through the management console - Snorby and is illustrated on Figure 5. After the process of mass classification of new events they are marked as a DoS attack.

5. Conclusions

The computer security will continue to be a serious problem in the coming years. However, with tools such as systems for attacks detection, countering threats would be an achievable goal. In this aspect, network security products developed for small organizations have great potential value.

This paper presents the architecture and basic functionality of an affordable, portable and easy to use system for detection and prevention of network attacks. The system is based on open source code, allowing tremendous flexibility in need of modifications and further developments.

The proposed IPDS is designed to be integrated easily into any existing network infrastructures. These can be both small business networks and organizations, as well as large corporate networks. The compactness of the system allows its use in different segments of the network, without the need for a physical change of its topology. This is especially effective for networks where is very important their operations to be uninterruptable, such as bank branches, hospitals, etc. The combination of cheap hardware and open source software is a possibility for widespread use of the system.

The goal of future work is the development of architecture by using multiple Raspberry Pi sensors distributed in various key network nodes to redirect the output data to a centralized database, thereby optimizing use of the IPDS system.

REFERENCES

1. Sanders C., Smith J. (2014). *Applied Network Security Monitoring: Collection, Detection, and Analysis*, Syngress.
2. Richard B. (2013). *The Practice of Network Security Monitoring: Understanding Incident Detection and Response*. No Starch Press.
3. RaspberryPi Project. http://www.raspberrypi.org.
4. Raspbian OS. http://www.raspbian.org.
5. Snort. http://www.snort.org.
6. Barnyard2. http://www.github.com/firnsy/barnyard2
7. MYSQL. http://www.mysql.org.
8. PulledPork. http://www.github.com/shirkdog/pulledpork/.

Contacts
Hristo Valchanov, Kolyo Raychinov
Technical University of Varna
9010, Varna, 1 Studentska Str.
phone: +359 52 383 278
E-mail: hristo@tu-varna.bg

© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 23, 2017

SIMULATION FRAMEWORK FOR REALIZATION OF PRIORITY-BASED LTE SCHEDULER

VENETA ALEKSIEVA, AYDAN HAKA

Abstract: *The LTE Technology provides simultaneously voice, data and video with different priority on networks. LTE applies the QoS bearers technique that provides high performance in packet delivery based on prioritization of the traffic. In this paper is proposed a simulation framework for LTE technology, which realized a priority based algorithm for LTE Scheduler, which reorders packets, based on classification mechanism.*

Key words: *LTE, Scheduler, prioritization, QoS*

1. Introduction

In the middle of 2016 the European Commission in [1] presents coordinated designation and authorization of the 700 MHz band for wireless broadband by 2020 and coordinated designation of the sub-700 MHz band for flexible use which safeguards the provision of audiovisual media services to mass audience, as well as investments into more efficient technologies, which are needed in order to vacate the current use of the 700 MHz band by digital terrestrial television. The prognosis of these estimates require the search of optimal 4G solutions in terms of QoS offered by telecom providers. According to published from Ericsson paper [2], suggests that by 2020 it is expected a growth of LTE subscriptions up to 3.7 billion. In February 2015 Cisco System published, in turn, prognosis for the period 2014-2019 [3], where the Global mobile data traffic will grow three times faster than global fixed IP traffic from 2014 to 2019. The Mobile data traffic will grow 10-fold from 2014 to 2019, a compound annual growth rate of 57%, and it will reach an annual run rate of 291.8 Exabytes by 2019, up from 30.3 Exabytes in 2014.

Requirements of IMT-Advanced [4] are 1 Gbit/s speeds for fixed and 100 Mbit/s for mobile users. Up to 2016 under these requirements, providers of 4G services choose between two advanced wireless technologies - LTE [5] or WiMAX [6], but since 2016 widely used 4G technology is LTE and many vendors implement only this technology in their end devices.

Solutions are needed to improve QoS in terms of delays in larger loads and any packet loss. For communications to be successful, it is also essential to focus on network traffic prioritizes for different types of communication streams.

2. Priority in LTE networks

After the first implementations of LTE the focus in the allocation of resources is shifting towards to the profit maximization and user satisfaction [7]. Even with the developing of LTE technology the QoS for uplink is discussed by many authors [8,9,10,11,12].

In 3GPP, the QoS Class Indicator (QCI) consist of basic classes, which are defined as "default", "expedited forwarding", and "assured forwarding". It means: expedited forwarding is used for "strict" priority (video and voice), and "assured forwarding" is used for business differentiation (e.g., weighted-fair priority).

In LTE network QoS is between end-user devices and Packet Data Network (PDN) Gateway applying the "bearers". "Bearers" is a set of network configurations to provide a special handling of traffic to its set prioritization. Their hierarchy is presented in Table 1. Default bearer is established when the user equipment (UE) connects to the LTE network, while Dedicated bearer is established whenever must be set QoS for a specific traffic type (service) as VoIP, video and etc.

GBR (Guaranteed Bit Rate) provides guaranteed bandwidth and monitors two parameters in directions uplink and downlink:

- GBR- minimum GBR for EPS (Evolved Packet switched System) bearer,
- MBR- maximum GBR for EPS bearer,

Non-GBR bearer does not provide guaranteed bandwidth and also monitors two parameters in directions uplink and downlink: A-

 ISSN 1310 - 8271

AMBR-general maximum speed permitted for the entire non-GBR throughput for specific APN (Access Point Name) and UE -AMBR- overall maximum speed permitted for the entire non-GBR throughput for all of APN particularly UE.

Table 1. *LTE QoS*

Dedicated Bearer		Default Bearer
Non-GBR	GBR	Non GBR
QCI 5-9	QCI 1-4	QCI 5-9
APN-AMBR UE-AMBR TFT ARP L-EBI	GBR MBR TFT ARP L-EBI	APN-AMBR UE-AMBR APN IP Address ARP

In LTE networks for differentiation of QoS same as in WiMAX are applicable classes which here are called QoS Class of Identifier (QCI). They define the basic characteristics of the IP packet level, as presented in Table 2.

Table 2. *QCI classes in LTE*

QCI	Bearer Type	Priority	Delay of the Packet	Packets Loss	Example of Traffic Type
1	GBR	2	100ms	10^{-2}	VoIP
2		4	150ms	10^{-3}	Video call
3		3	50ms	10^{-3}	Real time games
4		5	300ms	10^{-6}	Video stream
5	Non-GBR	1	100ms	10^{-6}	IMS Signaling
6, 8, 9		6, 8, 9	300ms	10^{-6}	TCP based services – chat, ftp...
7		7	100ms	10^{-3}	Voice, video, interactive games

Then in the base station (eNodeB) is applied a preemption algorithm, which allows high priority requesting bearers to displace low priority connected bearers in order to reduce the cell load. This algorithm coupled with a priority-based admission control can achieve low dropping and blocking probabilities.

3. Proposed algorithm for prioritization of UEs in the LTE

The aim of the proposed algorithm is to achieve keeping the network throughput as high as possible at a small price of only a bit more handovers. The functions for management of QoS in access networks are responsible for the efficient allocation of resources in a wireless interface. They are generally defined as the control algorithms of radio resources and incorporate power management, control of the transfer connection, access control , load control and the management packet , but directly related to QoS level cell are the last three. They are used to ensure a maximum throughput for individual services.

LTE uses multiple access technology (OFDMA), where the total bandwidth is divided into Resource Blocks (RBs) in the frequency domain. The Data is transmitted in the Transport Blocks (TB) in one transmission time interval (TTI) for 1ms. Each RB consists from 12 subcarriers (each of them is 15kHz). The frame is 10ms and divides into 10 equal subframes. Each subframe contains 2 slots*0.5ms. Each RB is related to one slot in time. One TB is related to 1 subframe and it is the minimum unit to schedule. The serve rule is to find first space that can fit the TB. If there are not enough RBs in the current TTI, the scheduler tries to find resources in the next TTI. This strategy minimizes the response latency, which is the best practice for delay sensitive traffic.

But in 1 Timeslot number of RB depends from frequency. It is presented in the Table 3.

Table 3. *Number of RB in LTE*

frequency	RBs for users traffic	RB for overhead
25MHz	4 RB	2 RB
2. 5MHz	10 RB	2 RB
5MHz	23 RB	2 RB
10 MHz	48 RB	2 RB
15MHz	73 RB	2 RB
20 MHz	98 RB	2 RB

But this procedure is not applicable for beacon transmissions (it is sent among devices each 100ms), because of emergency information it conveys, therefore the reserved resource blocks exist to accommodate the temporary overload.

The present paper offers an algorithm for UEs service in the distribution of resources in the uplink of LTE network as composed of two modules - by a control mechanism for admission (admission control) and Scheduler.

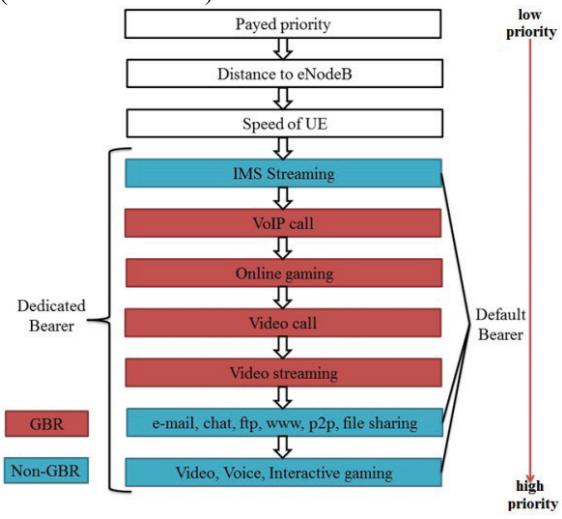


Fig. 1. *Traffic Prioritization in the Scheduler*

According to the network load, the admission control for the reception of orders manages the number of UEs, which can enter into

the Scheduler, in order to avoid overloading the system with too many UEs.

The Scheduler allocates RBs among UEs according to UEs needs. Resource allocation in the Scheduler is based on the priority, which is presented on the Figure 1.

4. Simulation Framework of LTE Scheduler

In this approach a simulation environment was established for implementation and exploration of the proposed algorithm. Used software tool is Visual Basic 2010. The architecture of the simulator is presented on the Figure 2.

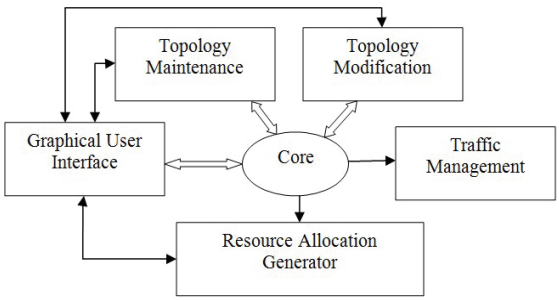


Fig. 2. *The LTE simulator's architecture*

The modules "Topology Maintenance" and "Topology Modification" are realized with classes "Form1" and "Form2". The classes contain methods for adding parameters of eNodeB and related UEs. A database for storing data from individual experiments for each eNodeB and its connected UEs is created. The tables from database with parameters for coordinator and related devices are presented on the Figure 3.

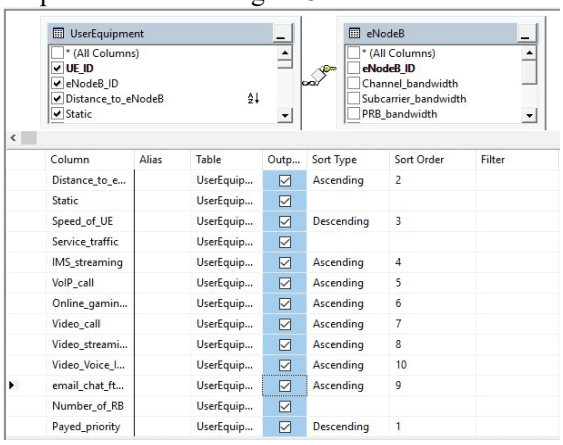


Fig. 3. *Database of LTE Prioritization in the eNodeB*

The module "Traffic Management" is realized with class "Form3". It loads data into the 'UserEquipment' table and visualizes the chart of the timing diagram.

The module "resource Allocation Generator", based on class "eNodeBdata", realizes the proposed algorithm for priority. The class contains methods for sorting UEs, adding it's data in array and arranging them.

Thus, the input of data for each device starts from initial parameters for the eNodeB. This is represented in the example of Figure 4. Another eNodeB (the next in the order) receives serial number. After that the user must type the number of UEs that will participate in the network of this base station, according to the limits imposed by the standard. The choice of bandwidth sets limits in the cell radius of the eNodeB and the transmission speed (bandwidth).

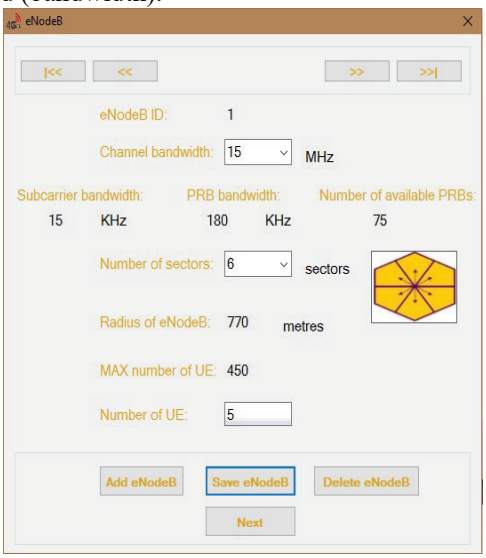


Fig. 4. *Input form of the parameters of eNodeB*

The tabular presentation of base stations and the information for UE (Figure 5) is suitable for displaying of multi aspected information that can be edited (Figure 6).

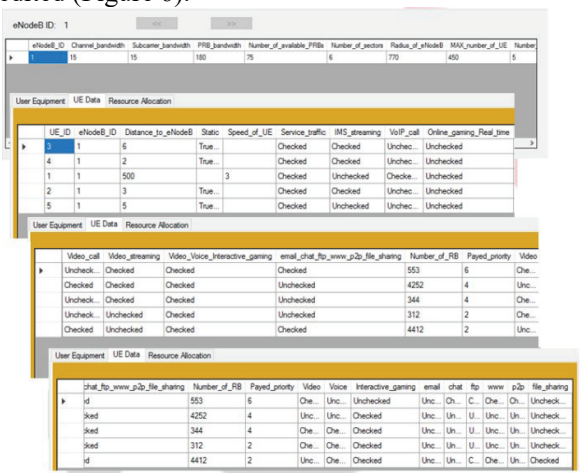


Fig. 5. *Database with information of eNodeB and related UEs*

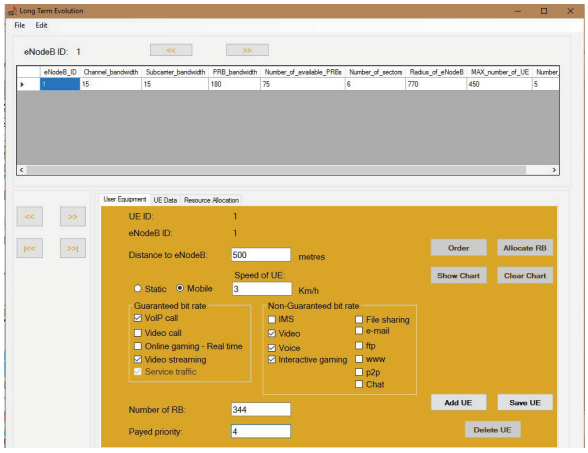


Fig. 6. *Data of UE, connected to the eNodeB*

One example of resource allocation and transmission matrix is presented on figure 7.

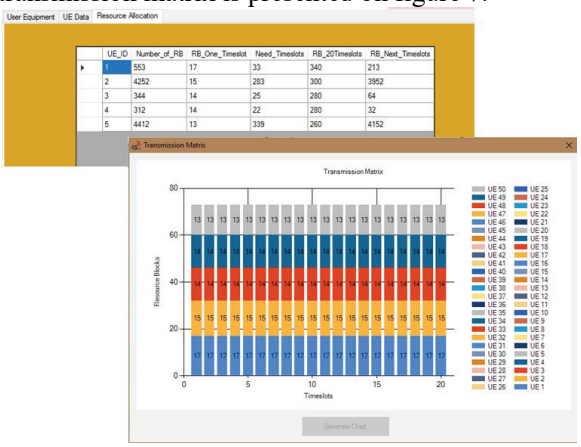


Fig. 7. *Resource Allocation and Transmission Matrix*

5. Experimental Results

The data from different experiments are send in .xls format to the next estimation.

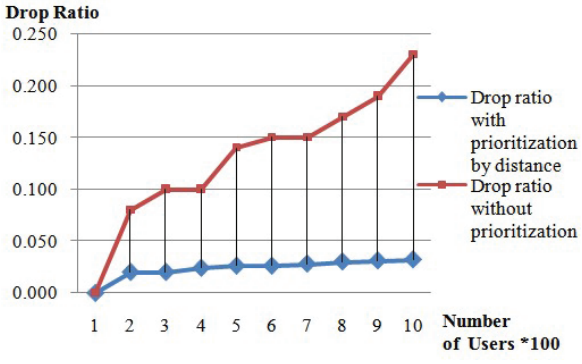


Fig. 8. *Drop Ratio*

On Figure 8 is presented the Drop ratio. It is possible to compare the number of drops when it is applied the prioritization, based on distance between the user and Base station and drops without prioritization. When the number of users grows up, the numbers of rejected beacons grow up too. This is the reason to increase the drop ratio when the number of the users is increased. The observed parameters degradation when connecting more users is related to the priority implemented scheduler, in which case the less priority queues may not be served in the case of network overload or congestion.

In assessing the performance of proposed algorithm of LTE Scheduler it also is assumed that UEs are evenly distributed within the range of the cell according to Section 7.1.6.1 of [13]. Each UE sends a signal periodically to know the channel condition of the UE for each TTI period of 1ms.

Figure 9 shows the dependence of the average time to establish a connection from UE to the eNodeB at different average intensity of applications when under area of the eNodeB are 100 UEs. An increase of the time for establishing a connection from the UE to eNodeB when the intensity of the requests is increased, as for the first 10% of the increase is from 10ms to 17ms, then the increase is significantly smoother and it amended in the range of 18ms to 23ms.

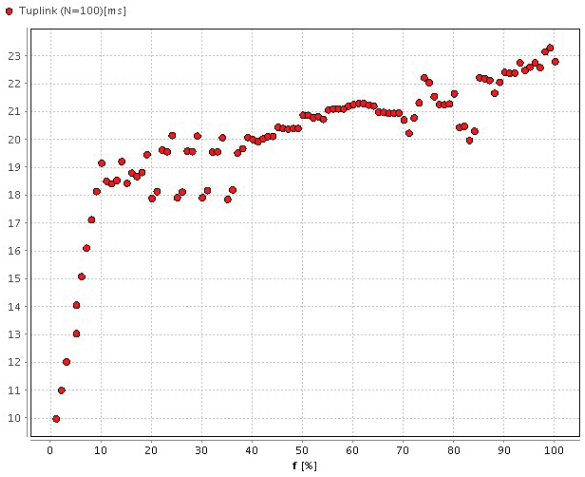


Fig. 9. *Average time for connection establishment with 100 UEs*

The results give reason to conclude that the presented algorithm for admission control in the Scheduler for LTE network can be applied successfully in the number of UEs under 100, because regardless of the intensity of the requests of the active UEs the average connection time is under 25ms-time, fully satisfying the requirements of [13].

In the Figure 10 it presents the dependence of the average time to establish a connection from the UE to eNodeB depending on the number of the UE, which are within the scope of the eNodeB for three different values of intensity of applications - 10%, 50% and 100%.

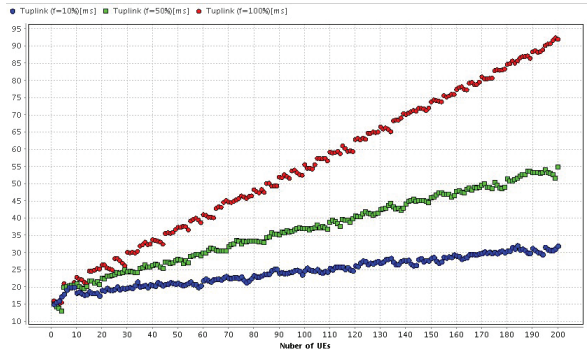


Fig. 10. *Average time for connection establishment with requests intensity of 10%, 50% and 100%*

It is easy to see a trend of increase in the time for establishing connection with an increase in the number of UE within the eNodeB, such as at a higher intensity time is significantly larger. If number of UEs is 200 and 100% intensity is reached 92ms, while number of UEs is 200 and 10% intensity the average time is 3 times less - 32 ms.

6. Conclusion

In this paper is proposed a simulation framework for LTE technology, which realized an efficient method for QoS for LTE service classes. In this framework is realized an algorithm for Scheduler to prioritize users in order to fit the bandwidth requirements, which are satisfying the application needs, based on the number of users and the prioritization of users. Simulation's results show that the proposed mechanism improves QoS, but the observed parameters degrade when the use of more subscribers is related to the priority implemented Scheduler, in which less priority queues may not be served in the case of network overload or congestion. There are presented drop ratio with prioritization and without prioritization. It was always assured a minimum transmission for all the service classes, although with different performances due to prioritization.

REFERENCES

1. EC (2016). Use of the 470-790 MHz frequency band in the Union, http://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1454410061980&uri=COM%3A2016%3A43%3AFIN, June 2016.
2. P. Cerwall and others (2015). Ericsson Mobility Report-On the Pulse of the Networked Society, http://www.ericsson.com/res/docs/2015/ericsson-mobility-report-june-2015.pdf, June 2015.
3. Cisco System (2016). VNI Mobile Forecast Highlights 2016-2021, http://www.cisco.com/c/dam/assets/sol/sp/vni/forecast_highlights_mobile/index.html#~Region, February 2015.
4. (2008). Document IMT-ADV/1-E, Background on IMT-Advanced, ITU Radiocommunication Study Groups, ps. 3, 7 March 2008.
5. (2014). LTE: capacity and cell-edge performance improvements, February 2014, http://www.3gpp.org/.
6. Ramadas K., R. Jain, (2008). WiMAX System Evaluation Methodology v2.1, WiMAX Forum, July 2008.
7. Abu-Ali N., Abd-Elhamid M. Taha, Mohamed Salah, and Hossam Hassanein (2014). Uplink Scheduling in LTE and LTE-Advanced: Tutorial, Survey and Evaluation Framework, IEEE Communications Surveys & Tutorials, vol. 16, Issue:3, pp. 1239-1265, Aug. 2014.
8. Al-Rawi M., R. Jantti, J. Torsner, and M. Sagfors (2008). On the Performance of Heuristic Opportunistic Scheduling in the Uplink of 3G LTE Net- works, in Personal, Indoor and Mobile Radio Communications, 2008. PIMRC 2008. IEEE 19th International Symposium, pp. 1 –6, 15-18 2008.
9. Calabrese F., C. Rosa, M. Anas, P. Michaelsen, K. Pedersen, and P. Mogensen (2008). Adaptive Transmission Bandwidth Based Packet Scheduling for LTE Uplink, in Vehicular Technology Conference, 2008. VTC 2008- Fall. IEEE 68th, pp. 1 –5, 21-24 2008.
10. Calabrese F., C. Rosa, K. Pedersen, and P. Mogensen (2010). Performance of Proportional Fair Frequency and Time Domain Scheduling in LTE Uplink, pp. 271 –275, May. 2009.
11. Liu F., X. She, L. Chen, and H. Otsuka (2010). Improved Recursive Maximum Expansion Scheduling Algorithms for Uplink Single Carrier FDMA System, pp. 1 –5, may. 2010.
12. Ruiz de Temino L., G. Berardinelli, S. Frattasi, and P. Mogensen (2008). Channel-Aware Scheduling Algorithms for SC-FDMA in LTE Uplink, in Personal, Indoor and Mobile Radio Communications, PIMRC 2008. IEEE 19th International Symposium, pp. 1–6, 15-18 2008.
13. 3GPP (2015). 3GPP, Tech. Specif. Group Radio Access Network; Physical layer procedures, 3GPP TS36.213, ver.12.4, 2015.

Veneta Aleksieva, Aydan Haka
Organization: Technical University of Varna
Address:Str. Studentska 1, Varna , 9010
Phone (optional): +35953383439
E-mail: valeksieva@tu-varna.bg, aydin.mehmed@abv.bg

© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 23, 2017

BENCHMARKING HASH FUNCTIONS

SVETOSLAV ENKOV, TONY KARAVASILEV

Abstract: *This paper presents the most effective use cases of hash functions. The purpose of the developed practical tests is evaluating hash functions by speed and security. They show how the different cryptographic algorithms, application level hashing, database layer hashing and even disk encryption can impact both the overall performance of a system and are crucial for the security of the data implied in it. The results of this experimental research are presented in this paper.*

Key words: *hash functions, cryptographic algorithms, security, performance, encryption*

1. Introduction

Hash functions are any functions that can be used to map data of various size to data of fixed size. The value returned by a hash function is called a digest, hash value or just hash and are commonly represented as a hexadecimal number. [1]

There is a special class of cryptographic hash functions that can map data with arbitrary size to a fixed bit string. This class is characterized by the one-way mathematical design used in the generation of digests and can also be referred as a digital fingerprint of the input data. Said in others words, every input should have an associated unique digest output without a chance of finding duplicates. [2]

Hash functions are used a lot in computer science and have many information-security applications, such as:

- Identifying duplicated files or substrings;
- Detection of accidental data corruption;
- Providing forms of authentication;
- Secure storage of sensitive data;
- Fingerprinting of data;
- Digital signatures;
- Caching identifiers.

Nowadays every modern programming language, database script language, software framework or plugin library has a set of cryptographic functions that comes preinstalled with it. Also, every hash function can have a different kind of complexity, amount of required resources, security level, performance grade and realization for the its algorithm.

The most frequently used functions in software development currently are: [1] [3]

- MD5 – not suitable for securing data [4], but usable for caching, verify data integrity, etc.
- SHA-1 – not suitable for protecting data [5], but usable for detection of corruption, identifying duplicates, etc.
- SHA-2 – a set of six hash functions that are cryptographically secure and can be used for all purposes. Every one of them has a different output size and level of security. They produce 224, 256, 384 or 512 bit hash values. Those algorithms are SHA-224, SHA-256, SHA-384, SHA-512, SHA-512/224, SHA-512/256. The last two are rarely used.

The main purpose of this article is testing the performance of different algorithms when used in various ways. This way we can define the correct use cases for each algorithm and find its best use. The main tests are focuses on the hashing of 100000 strings in three lengths:

- Password type size – 20 symbols;
- Identifier type size – 50 symbols;
- Comment type – 100 symbols.

The hashing is tested both from an application point with PHP web pages and from database level point with MySQL stored procedures. For both PHP and MySQL, we used the internal realizations for the algorithms MD5, SHA-1, SHA-244, SHA-256, SHA-384 and SHA-512, being the mostly used hash functions. We have performed two tests for each layer:

- Read and hash the values;
- Read, hash and update the values.

Besides that, this paper also includes an overview of the time that would be needed when combining both application and database hashing

 ISSN 1310 - 8271

for higher security, other problems that may change the software climate and how can using disk encryption effect the performance and resources of the software system, while boosting protection.

All the practical tests in this article have arisen from the need of implementing both the protection of sensitive data, passwords verification and generating caching identifiers for an improved second version of a private online website.

2. Testing environment specification

For the results to be adequate we have chosen to run the test on a virtual machine, created with Oracle VM VirtualBox version 5.1.14, that has a typical web development stack installed on it. The specification of the allocated resources for the virtual machine and the software on it are shown on Table 1.

Table 1. *Virtual machine specification*

	Detail
CPU	Intel i7-6700HQ, 2 cores, 2.59GHz
RAM	DDR4, 2048 MB, 2.40 MHz
GPU	Intel HD Graphics 530, 32 MB, 2.40 MHz
HDD	20GB, 7200 RPM, 32 MB cache
OS	Ubuntu Server 16.04.2 LTS, Kernel 4.4.0
LAMP	Apache 2.4, PHP 7.0.15, MySQL 5.7.17

The virtual machine has installed all available updates, kernel drivers and virtualization needed packages. All settings for the LAMP stack (Linux Apache MySQL PHP) are by default, with the exception of boosting the values for maximum memory usage by both PHP and MySQL to allow the allocation of all available random-access memory (RAM).

The tests times for PHP are just for the section of the program that does iteration, hashing and updating of values. The time needed for generating 100000 strings with the various length is explicitly excluded.

The MySQL database has three tables having per 100000 strings, for each of the three length types tested. The time for generating digests and running the select or update statements is only taken under consideration.

Each single experiment is executed 10 times and the average time of those runs is taken as final. All the results will be in seconds with 6-digit precision after the decimal point.

3. Application vs. Database level hashing

Most of the developers are feed up with the dilemma where to put the encryption logic and how much will it cost the system. The two daily operations that every kind of cryptography logic needs to imply are read-hash-use and read-hash-update. The next two sections show how PHP and MySQL handle those two with different string lengths.

3.1. Read and hash experiment

This experiment will test the situations when you need to get some string data and hash it so that you can do some sort of verification, comparison or uniquely map the data. For PHP, we used an array with the strings to iterate and hash them. While for MySQL, a table with the data in it, so that we could do a select statement with an alias that calls the current hash function and returns the result for all rows.

3.1.1. Password length results

For the password type size of 20 symbols experiment, the results for application and database hashing times are shown on Table 2.

Table 2. *20 symbols hashing results*

	PHP	MySQL
MD5	0.026375	0.050560
SHA-1	0.033895	0.058299
SHA-224	0.059067	0.078799
SHA-256	0.059711	0.081613
SHA-384	0.072731	0.096270
SHA-512	0.075571	0.113985

As we can see from the results, the application level hashing is faster than the database one. The MD5 and SHA-1 can be easily used for generation of identifiers from PHP, but not for passwords. The SHA-2 algorithms performance is pretty close on both sides, so you can easily use it for protecting data and moving some of the pressure away from the application code to the database.

Because this test is overviewing sensitive data hashing, stick with the SHA-2 algorithms for better security.

3.1.2. Identifier length results

When using hashing for caching purposes, you would probably need an identifier for data of size about 50 symbols. The performance of all the chosen hashing algorithms can be seen on Table 3.

Table 3. *50 symbols hashing results*

	PHP	MySQL
MD5	0.026706	0.049312
SHA-1	0.034188	0.055153
SHA-224	0.059884	0.082104
SHA-256	0.060874	0.094240
SHA-384	0.072884	0.107898
SHA-512	0.076474	0.122992

The final results show that it is better to use application layer hashing when using multiple time generation of identifiers and that the MD5 and SHA-1 are the fastest to compute. However, if you are using it for one-time generation for some third party in memory table or MySQL memory table, it probably will not slow the software system much.

This type of strings is probably used for caching or mapping and this means security is only needed for the cached data, but not for the identifier. This means you can use SHA-1 for balancing performance with ease.

3.1.3. Comment length results

Sometimes the developer needs to hash longer data, for example finding duplicate comments that can be about 100 long and above. This is why we will use string length of 100 symbols for this experiment. The performance of the hash functions for this experiment is shown on Table 4.

Table 4. *100 symbols hashing results*

	PHP	MySQL
MD5	0.037308	0.063266
SHA-1	0.048937	0.075928
SHA-224	0.094610	0.128636
SHA-256	0.100509	0.130748
SHA-384	0.072727	0.105875
SHA-512	0.076085	0.106118

The length of input data seems to be affecting every algorithm in a way. The first thing that can be notice is that SHA-224 and SHA-256 computation time is slower that the more secure algorithms SHA-384 and SHA-512. The second thing is that you can use SHA-512 hashing from the database side and be as fast as using SHA-256 from the application layer.

Comments sometimes may be sensitive information and you should be careful what you are comparing, use SHA-2 algorithms. You can drop some of the computation pressure by using SHA-512 from the database side or stick to the faster SHA-512 from application level.

3.2. Read, hash and update experiment

The second test will apply to the scenarios where you are obliged to save the computed hashes. It is way normal of saying that the application side results will be faster here. This is only because resaving data for use there will happen in the random-access memory which is faster than any available storage drives today, but is only temporary media. We will talk about combining application and database hashing for more security later in this paper. The point of this test is mostly for identifying the generation times differences on each side.

For PHP, we used an array of arrays to contain both the input data and the computed hash in the same subarray. As for MySQL, we used a table with the data in it and added a second column for the computed digests. This way we could easily do an update statement that reads the first column, computes the hash of each value and saves it to the second column.

It should also be noted that when using hash functions with longer digest output size, you should prepare for bigger storage sizes.

3.2.1. Password length results

For the password type size of 20 symbols experiment, the final results are shown on Table 5.

Table 5. *20 symbols hashing results*

	PHP	MySQL
MD5	0.028027	0.317091
SHA-1	0.035334	0.363700
SHA-224	0.057294	0.398448
SHA-256	0.061983	0.441395
SHA-384	0.075343	0.450580
SHA-512	0.079736	0.592357

Since MD5 and SHA-1 algorithms are security broken, it is recommended to use the SHA-2 ones. As we can see from the test, the stronger the algorithm, the more computation time is needed.

Having in mind that passwords are always sensitive data, you would probably want to use application data encryption to reduce the time of insert or update in SQL. Still, adding a second layer of encryption from the database side will boost security a lot, but will cost you only a bit of performance.

3.2.2. Identifier length results

When needing to map unique data and save it somewhere for further use, you would need an identifier generation operation. For this test, the needed data to map should be at least 50 symbols. The computation and save time can be seen on Table 6.

Table 6. *50 symbols hashing results*

	PHP	MySQL
MD5	0.028937	0.343682
SHA-1	0.034090	0.375937
SHA-224	0.061613	0.374757
SHA-256	0.063669	0.381966
SHA-384	0.075568	0.434025
SHA-512	0.078516	0.459434

For this data length, using a stronger algorithm needs more time. When generating identifiers, you can easily choose MD5 or SHA1.

If you do want to use another algorithm, have in mind that you would need more space for the output digests. You would probably need to map the data once and only add some new records over time, but you should definitely use the application side computing to save time. Security is not a factor for the identifier, only for the underlining data that it maps to.

3.2.3. Comment length results

When needing to map comments or having their hash for faster comparison, the developer needs to compute digests for longer strings. We will use a 100 symbols length for this experiment. The computation results are shown on Table 7.

Table 7. *100 symbols hashing results*

	PHP	MySQL
MD5	0.038966	0.332903
SHA-1	0.050947	0.359913
SHA-224	0.102294	0.381809
SHA-256	0.104354	0.401970
SHA-384	0.075037	0.424538
SHA-512	0.078616	0.488914

The results show that SHA-384 and SHA-512 behave better with bigger strings on the application side, even faster than SHA-224. The database level hashing time seems to be affected only by the complexity of the current algorithm.

If you will be using the hashing for comparison only, it is not a problem using even MD-5 or SHA-1. In any other situation stick to SHA-2. Because comments can be edited often, the smartest thing is avoiding database computations.

3.3. Disk Encryption aftermath

We cannot skip that having a full software disk encryption is a big factor in securing your data. So we did a research on the performance of disk encryption and retested the above results on a fully encrypted virtual machine, equivalent to the machine we tested before on.

The interesting thing that we found out from both theory and practice is that it will not affect computation time in any way. This is because modern processors have a special set of instructions called AES-NI [6] that allows encryption of up to 1 Gbit of data per second fully transparently.

The process of encrypting and decrypting is directly in the random-access memory and this creates the so called transparent encryptions effect, only because of the high speed of RAM. When you need to write something on your disk device, the central processor will first encrypt it and after that send it to the storage device. [7]

If you are a professional, using this kind of encryptions can only boost your security status and it will not affect your performance in any fatal way. The only risk is that when rebooting you need to input your encryption password to boot into the operating system and it would be fatal if you have lost it or forgot it somehow. If you get any resource shortage, switch to partition encryption instead.

In any way, using disk encryption will protect your data, but can also be a double-edged sword. Use it with caution and do not try to optimize things before they need to.

4. Coping with the combined approach and real world problems

When dealing with sensitive and top secret data in the real world some other needs, specifications and problems may occur. Even when money and people power are practically unlimited, you can get bad results if you do not give professionals the time needed to get rid of all your problems at once.

The next sections show the most frequent situations and the extra needs that may occur.

4.1. Combined approach

In a software system that deals with private and security information, an extra need for multilayer security may apply. In this case, combining hashing from both application and database levels is a good choice.

When faced with this challenge, it is better to lay the complex encryption algorithm in application level and just send it to the database for a one-time hashing before saving. As said before, while both sides do calculations in the random-access memory, the database side will need more time because it saves data on a disk for further use.

An overview of a real world situation would be where PHP generates a digest, sends it to the database for MySQL to generate a hash value of it and saves the final hash to disk. Using the data from the experiments and calculating an average of the time needed for application generations, database computations and saving the values to storage, the final results per algorithm can be seen on Figure 1.

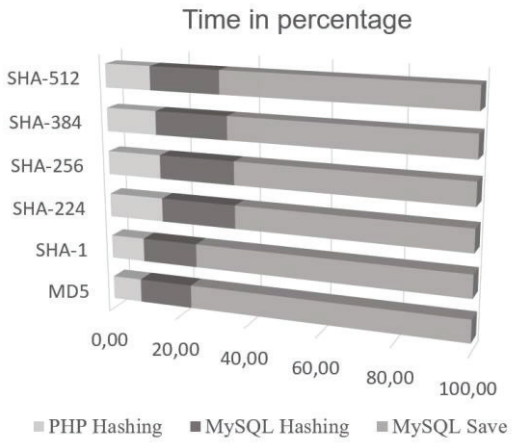


Fig. 1. *Combined approach times*

It is easy то conclude that the heaviest part of this approach is the process of saving data into the database. Application and database hashing are not the slowest operations. Having the results for each algorithm, we can easily calculate in percentage the average time taken from the three subprocesses. The result is shown on Figure 2.

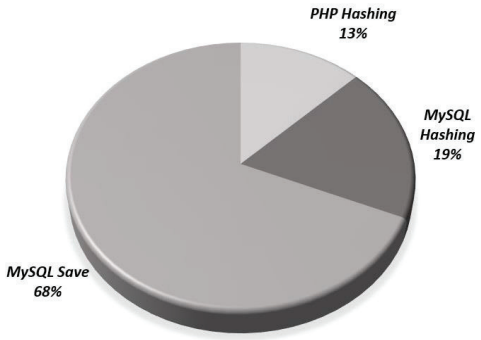


Fig. 2. *Average time for each part of the combined approach work*

When using both side calculations, we can conclude that the two most important parts of any kind of combined solution are:

- Choosing a number of secure algorithms that are proven to be effective and distributing them as needed. This is done in order to have a real point of using multilayer encryption;
- Profiling your application before and after realization, just to see if your software is dealing with more pressure on application or database level and double check final results before you shipping it.

Be cautious when having a security paranoia, because if things are not done right, this may damage your overall software performance. Finding equilibrium is the price for success.

4.2. Real world complications

Sometimes reality surprises us and changes our needs in an unpredicted way. That said: “If you believe everything you read, better not read.” - Japanese Proverb.

For an example, let say we have a fast application that needs 3 seconds for application computations and 1 second for database querying. You decide to add four time hashing in the application level and one-time digest generation at the database layer.

Your expectations should be that the whole software will load for maximum of 5 seconds. You profile it and it turns out to loads for 7 seconds. But how can this be possible? This is only because you have not taken these several factors:

- Network time – when the connection between the application and the database is over a network or a socket, you may get some extra delay for networking.
- Domain Name System resolution time – this is one of the most underestimated threats. After you ship your application online, some of your clients may notice a bigger delay than others. This may be because of the amount of request you do involving DNS queries.
- Database driver – have in mind that your database driver may behave differently on bigger amounts of data or because of your local network quality;
- Virtualization – although it can give you network isolation and less physical devices in your server room, sometimes the virtualization drivers cause extra slowness and may behave irradically. Placing software on top software may only slow your machine, but placing hardware on top of hardware is the real deal;
- Slow hardware – your oldest enemy. This may break your clients experience and security. Do not save money for hardware or system administrators and you will have less problems.

The above and other unknown problems may cause your system to get more than two time slower than you have expected.

Main point is, before you do anything, test your code in a real world situation and try to simulate the worst case scenarios before shipping it to the real world. Keep the balance between security and performance or you will slowly start losing your clients.

4.3. Professional data protection side notes

When taking care of sensitive data, sometimes just hashing your passwords may not be enough. When your information is strictly

confidential, just combine compression, multilayer hashing, symmetrical encryption, secure connections, network isolation, firewalls, password policies and double side application validations. [1]

Do not assume that you are safe. Include thread assessment and insert security measures in every single part of your development cycle. Starting from the design phase, all the way to release and maintenance. This is the professional way of creating secure applications and not just adding security features on top of by design vulnerable software. [3]

We are not going to drill more into this topic, because the point of this paper is focused on hash functions.

5. Best hashing use case scenarios

Based on the results of this paper, we can say there are three main uses for hash function and the best ways of choosing algorithms for each one are summarized in the next sections.

5.1. Generating identifiers

Every developer comes to the need of mapping data and computing unique identifiers to associate the data with. Using MD5 or SHA-1 seems to be the smartest choice only because they the fastest and do not have a great chance of duplicated digests results. Of course using either SHA-224 or SHA-256 would not hurt your performance much.

5.2. Data verification

For verifying strings or easily comparing them, using any of the tested algorithms will do the trick for small amounts of data and be fast enough. However, we would suggest SHA-256, because it is currently used for professional application and digital signatures. When dealing with longer data, switch to SHA-384 instead.

5.3. Securing data

When using hash functions for protecting sensitive data, it is recommended to use only the SHA-2 algorithms. Choosing between SHA-224, SHA-256, SHA-384 or SHA-512 are the best choices for secure storage and data transfers. Of course in a real software system you would probably want more than one type of cryptographic algorithm involved.

6. Conclusion

This paper has tested the performance of all the most used and standardized hash algorithms from both application and database realization sides. It maps the best uses of each algorithm and shows the consequences a developer may be faced with by using them.

The most interesting results from the experiments are:

- Application layer hashing is faster that database, but are pretty close;
- SHA-384 and SHA-512 seem to deal faster with longer input data than others;
- Disk encryption does not affect your application speed, but requires more computation power;
- Using just hashing may not be enough for the security of your data.

REFERENCES

1. Katz, J., and Lindell, Y. (2007). *Introduction to Modern Cryptography: Principles and Protocols*. ISBN: 9781584885511.
2. Menezes, A., P. van Oorschot, and Vanstone, S. (1996) *Handbook of Applied Cryptography*. ISBN: 9781439821916.
3. Howard, M., and LeBlanc, D. (2004) *Writing Secure Code: Practical Strategies and Proven Techniques for Building Secure Applications in a Networked World.* ISBN: 9780735617223
4. Wang, X., and Yu, H. (2005) How to Break MD5 and Other Hash Functions. *Advances in Cryptology – EUROCRYPT 2005*. ISBN: 9783540259107
5. Wang, X., Yin, Y.L, and Yu, H. (2005) Finding Collisions in the Full SHA-1. *Advances in Cryptology – EUROCRYPT 2005*. ISBN: 9783540259107
6. http://www.intel.com/content/dam/doc/white-paper/enterprise-security-aes-ni-white-paper.pdf
7. https://wiki.archlinux.org/index.php/disk_encryption

Contacts:
UNIVERSITY OF PLOVDIV PAISII HILENDARSKI
24 TZAR ASEN
PLOVDIV
E-mail: enkov@uni-plovdiv.bg
E-mail: tony.karavasilev@gmail.com

© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 23, 2017

CLUSTERING TECHNIQUES FOR ANALYSIS OF LARGE DATASETS

MILENA ANGELOVA

Abstract: *Constant advance in science and technology makes collection of data and storage much easier and very inexpensive than before. This led to the information of enormous datasets in science, government and industry, which should be processed or sorted to get useful information. The goal of this survey is to provide a comprehensive review of different techniques in data mining and how they can deal with large data sets.*

Keywords: *traditional clustering techniques, data stream clustering, incremental clustering*

1. Introduction

Constant advance in science and technology makes collection of data and storage much easier and very inexpensive than ever before. This led to the information of enormous datasets in science, government and industry, which should be processed and/or sorted to get useful information.

For example, if we consider the results generated by a search engine for a particular query, user has to sift through the long lists and find the desired solution. But this job can be very difficult for the user if there are millions of web pages displayed as solutions for a given query. Thus clustering techniques can be very useful in grouping the closely related solutions of a given query and displaying the results in the form of clusters so that the unrelated documents can be avoided even without taking a glimpse at them.

The main idea behind clustering any set of data is to find inherent structure in the data, and interpret this structure as a set of groups, where the data objects within each cluster should show very high degree of similarity known as intra-cluster similarity, while the similarity between different clusters should be reduced. There exist several clustering algorithms which can be used for large datasets. They can be separated in three different groups: partition methods, hierarchical and density based clustering algorithms. They can be called traditional clustering algorithms.

The main problems associated with the traditional clustering algorithms are handling multidimensionality and scalability with rapid growth in size of data. The increase in size of data increases the computational complexities which have a devastating effect on the runtime and memory requirements for large applications. There exist two approaches which deal with these issues: data stream clustering and incremental clustering.

Data stream clustering has recently attracted attention for emerging applications that involve large amounts of streaming data. Data stream model has some constraints. Data from the streams cannot be stored due to the large volume, therefore only summaries should be processed and stored. The speed of arrival is fast thus each item has to be processed in real time or only once. Data from past might become irrelevant and can even affect negatively the current summaries.

Incremental clustering, sometimes called data stream clustering, is a clustering problem, where the data to be clustered arrive continuously. In data stream clustering, the data is expected to be finite, but for the incremental clustering we have to expect the data stream to be infinite. Therefore, incremental clustering supposes two approaches online and offline clustering of streaming data.

In this work, first we present all the major traditional clustering techniques in brief and then we discuss how they can deal with a large data sets. Then we analyze how the data stream and incremental techniques can deal with issues from traditional techniques and large data sets.

2. Traditional techniques for clustering

Clustering can be done in many different ways. There exists different kind of clustering techniques deal with data sets. Some take input parameters from the user like number of clusters to be formed etc., but some decide on the type and amount of data given. Clustering algorithms can be classified in five distinct types: partitioning,

 ISSN 1310 - 8271

hierarchical, model-based, density based methods and grid-based methods. In this section we have described partitioning, hierarchical and density based methods and how they can deal with large data sets.

2.1. Partitioning clustering

Let a database containing n data objects is given, then a partitioning method constructs *k* clusters of the data where $k <= n$ and *k* is the input parameter provided by the user. That is, it classifies the data into *k* groups (clusters) which should satisfy the following conditions: (1) each group must contain at least one data object and (2) each data object should belong to only one group. The most well-known and commonly used partitioning methods are k-means proposed by Mac Queen [1] and k-medoids proposed by Kaufman and Rousseeuw [2].

2.1.1. The k-means method

The k-means algorithm takes input k from the user and partitions n data objects into k clusters so that the resulting intra-cluster similarity is very high and inter-cluster similarity is very low. The cluster similarity is calculated based on the mean value of the objects in the cluster. First, it randomly picks k data objects as the mean or centroid points. For each of the remaining objects, an object assigned to the centroid to which it is most similar based on distance between the object and the cluster. This process iterates till good clusters are formed. Typically, squared root function is used for this which can be defined as

$$\boldsymbol{E} = \sum_{i=1}^{k} \sum_{x \in C_i} |\boldsymbol{x} - \boldsymbol{m_i}|^2, \qquad (1)$$

where *x* is the point in space representing the given object, and m_i is the mean of cluster C_i. This function tries to make the clusters as separate as possible.

The method is relatively scalable and efficient in handling large data sets because the computational complexity of the method is O(*nkt*), where *n* is the total number of iterations, k is the number of clusters and t is the number of iterations. Normally $k<<n$ and $t<<n$ so, the method often ends up at local optimum. However, this method has some drawbacks: 1) it can be applied only when mean of a cluster is defined, but when data with categorical attributes is involved it cannot be the case; 2) the user should specify the number of clusters k in advance and 3) it is sensible to noise and outlier data points.

2.1.2. The k-medoids method

K-medoids algorithm was developed to overcome the drawbacks of k-means which is very sensitive to outliers. For example, an object with some extremely large value may substantially distort the distribution of data by k-means method. Therefore, instead of taking the mean value of objects in a cluster as a reference point, an object that is most centrally located in the cluster can be taken as a representative object, called as medoid. Thus the partitioning method can be performed by minimizing the sum of dissimilarities between each object and with its corresponding reference point.

The algorithm creates *k* partitions for *n* given objects. Initially k-medoids are selected which are located more centrally in each cluster. Then the algorithm repeatedly tries to make a better choice of medoids by analyzing all the possible pairs of objects.

The k-medoids method is more robust than k-means because it is less influenced by outliers or other extreme values than mean. But its processing is very costly than k-means method and it also has the drawback of user providing the input parameter k.

The main advantage of k-means and k-medoids is that their complexity is linear. Therefore, their execution time is proportional to the number *n* of data objects, so that they can be used with large volumes of data. The second advantage is that it is possible to detect outliers, which appear in the form of singleton clusters. There is also a third advantage. Unlike hierarchical methods, in which the clusters are not altered once they have been constructed, the reassignment algorithms constantly improve the quality of the clusters, which can thus reach a high level when the form of the data is suitable.

The first disadvantage is that the final partition depends greatly on the more or less arbitrary initial choice of the centers c_i. Consequently, we do not have a global optimum, but simply the best possible partition based on the starting partition. The second drawback of partitioning methods is that the number of clusters, *k*, is fixed in these methods, and it is not less than k unless certain clusters are empty. If this number does not correspond to the actual configuration of the cloud of data objects, the quality of the clustering may be adversely affected. The third drawback of these methods is that they are only good at detecting spherical forms. Even convex forms such as ellipses cannot be detected well if they are not sufficiently separated.

2.2. Hierarchical clustering algorithms

Hierarchical clustering, also known as connectivity based clustering, is based on the core idea of objects being more related to nearby objects that to objects farther away. These algorithms connect “objects” to form “clusters” based on their distance. A cluster can be described largely by the maximum distance needed to connect parts of the

cluster. Strategies for hierarchical clustering generally fall into two types:

- **Agglomerative clustering** is a "**bottom-up**" approach: each observation starts in its own cluster, and pairs of clusters are merged as one moves up the hierarchy.
- **Divisible clustering** is a "**top-down**" approach: all observations start in one cluster, and splits are performed recursively as one moves down the hierarchy.

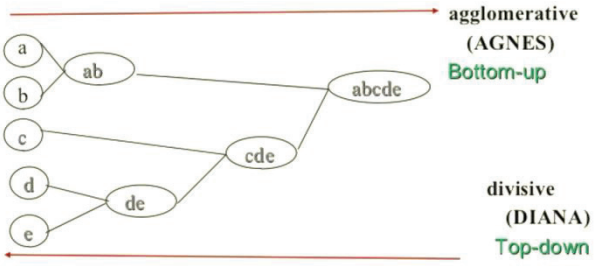


Fig. 1. *Examples of agglomerative "Bottom-up" and divisive "Top-down" clustering*

In general, the merges and splits are determined in a greedy manner. The results of hierarchical clustering are usually presented in a dendrogram. The complexity of agglomerative clustering is $O(n^2\log(n))$ [3], which makes them too slow for large data sets. Divisible clustering with an exhaustive search is $O(2^n)$, which is even worse. However, for some special cases, optimal efficient agglomerative methods (of complexity $O(n^2)$) are known: SLINK [4] for single-linkage and CLINK [5] for complete-linkage clustering.

The agglomerative clustering does not suffer from the two major drawbacks of the moving centers method, namely its dependence on the choice of initial centers and the fixed number of clusters chosen in advance. The second advantage is that it can detect clusters of different shapes, according to the distance chosen. The best possible distance can be chosen according to previous knowledge of the shape of the clusters to be detected, but the choice is mainly made by carrying out a number of trials and observing the results. It is possible to use a given distance because the resulting clusters are more easily described or because the number of clusters appears more natural with respect to the graphic indicators. The third advantage is that it enables us to cluster data elements, variables or centers of clusters obtained by using a moving centers algorithm.

The main disadvantage is its algorithmic complexity, which is non-linear: in order to move from $k + 1$ clusters to k clusters, we must calculate $(k + 1)k/2$ distances and combine the two closest clusters. If n is the number of data objects (elements) to be clustered, the complexity of the basic algorithm is of the order of n^3, and it will soon exceed the capacity of even a powerful computer. The difference from the moving centers method is that we are comparing individuals with each other instead of comparing them with the centers of the clusters only. The situation can be improved to a certain extent by the nearest-neighbor algorithm, which reduces a complexity of n^3 to n^2, by a judicious combination of more than two observations on each iteration. It can be used with the Ward, average linkage, single linkage and complete linkage methods. A second drawback is that, at each step, the partitioning criterion is not global, but depends on the clusters obtained so far: two individuals placed in different clusters will no longer be compared. In other words, this type of clustering into n clusters in not necessarily the best possible outcome, but only the best of those obtained by combining the clusters of a clustering procedure into $n + 1$ clusters. Thus some natural clusters may be hidden by an earlier branching.

2.3. Density based clustering algorithms

Density-based clustering [7] methods are based on a local cluster criterion. Clusters are assumed as regions in the data space in which the objects are dense and the clusters are separated by regions of low object density. These regions have an arbitrary shape and the data points inside a cluster may be arbitrarily distributed.

The most popular [8] density based clustering method is DBSCAN [9]. In contrast to many newer methods, it features a well-defined cluster model called "density-reachability". Similar to linkage based clustering, it is based on connecting points within certain distance thresholds. However, it only connects points that satisfy a density criterion, in the original variant defined as a minimum number of other objects within this radius. A cluster consists of all density-connected objects (which can form a cluster of an arbitrary shape, in contrast to many other methods) plus all objects that are within these objects' range. Another interesting property of

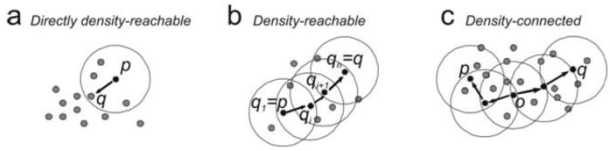


Fig. 2*. Examples of (a) directly density-reachable, (b) density-reachable, (c) density-connected in density-based clustering.*

DBSCAN is that its complexity is fairly low – it requires a linear number of range queries on database

– and that it will discover essentially the same results in each run, therefore there is no need to run it multiple times. OPTICS [10] is a generalization of DBSCAN that removes the need to choose an appropriate value for the range parameter ε, and produces a hierarchical result related to that of linkage clustering. DeLi-Clu, [11] Density-Link-Clustering combines ideas from single-linkage clustering and OPTICS, eliminating the ε parameter entirely and offering performance improvements over OPTICS by using an R-tree index.

The key drawback of DBSCAN and OPTICS is that they expect some kind of density drop to detect cluster borders. Moreover, they cannot detect intrinsic cluster structures which are prevalent in the majority of real life data. A variation of DBSCAN, EnDBSCAN, [12] efficiently detects such kinds of structures. On data sets with, for examples, overlapping Gaussian distributions – a common use case in artificial data – the cluster borders produced by these algorithms will often look arbitrary, because the cluster density decreases continuously. On a data set consisting of mixtures of Gaussians, these algorithms are nearly always outperformed by methods such as EM clustering that are able to precisely model this kind of data.

3. Data Stream and Incremental clustering algorithms

Traditional clustering techniques are not suitable for large data sets. Many applications have massive amount of data which increases every day. This causes limitation in data storage capacity and processing time. In this section, we have described two techniques data stream and incremental clustering which can be used to deal with a big amount of data.

3.1. Data stream clustering

Data stream clustering algorithms have been developed as an adaption of traditional clustering algorithms to fit the streaming model and comply to its constraints. The data stream problem has been researched in recent years because of the large number of relevant applications [13, 14, 15, 16, 17]. Furthermore, data stream clustering has recently attracted attention for emerging applications that involve large amounts of streaming data. In the literature, one of the first results for data streaming appeared in 1980 [18] but the model was formalized in 1998 [19].

A formal definition for data stream S is that of a sequence of arriving data objects $x_1, x_2, \ldots, x_n$, rendering $S = \{x_i\}$, where $i = 1, 2, \ldots, m$ and $m \rightarrow \infty$. Each data object is an m-dimensional attribute vector $x_i = [\mathbf{x_i^j}]$ where $j = 1, 2, \ldots, n$, belonging to an attribute space that can be continuous, categorical, or mixed.

The main disadvantage of data streaming algorithms is to fit a large data into main memory. It is not possible for the algorithm to "remember" too much of the data scanned in the past. This scarcity of space necessitates the design of a novel kind of algorithm that stores only summary of past data, leaving enough memory for the processing of future data [16, 17]. Researchers in [31] are developed a clustering method which is called CluStream. In the literature, CluStream is defined as algorithm which summarizes streams into micro-clusters to deal with memory limitations. The CluStream model has a wide functionality in characterizing data stream clusters over different time horizons in an evolving environment. They use a pyramidal time window assures that the essential statistics of evolving data streams can be captured without sacrificing the underlying space and time efficiency of the stream clustering process. [31] As advantages, CluStream is very flexible for real time transactions and pyramidal time window guarantees efficient time and space.

Another approach for these constraints have a direct impact in memory usage and data "age" relevance. Some techniques have been developed to deal with these problems such as sliding window combined with low memory functions. Sliding window is a technique where a window of size W keeps the last W items that arrived and run a desired algorithm on these items. The summary is then kept and the items are then discarded. In this way an up to data statistics is always computed. These statistics are usually kept by updating the model; in case of classification is could be the decision tree and in clustering could be the final clustering centroids.

Advantage of data stream algorithms instead of k-partitioning clustering methods is that we do not need to process the data multiple times to receive final clustering decision. Different techniques were created to deal with evolving data such as one pass processing and summarization. In literature, researches in [16] give a constant-factor approximation algorithms for the k-means problem in the data stream model of computation in a single pass.

As a conclusion of this section, clustering data streams is gaining a lot of importance and a lot of research is done on improving these algorithms to apply them on real time applications.

3.2. Incremental clustering

Incremental clustering is based on the assumption that it is possible to consider instances one at a time and assign them to existing clusters. Here, a new instance is assigned to a cluster without significantly affecting the existing clusters. Only the cluster representations are stored in the main memory to alleviate the space limitations.

One of the most famous example is the COWEB[20]. It keeps a classification tree as its hierarchical clustering model. Then the algorithms place new points in a top-down fashion using a category utility function.

Another incremental hierarchical clustering algorithm that works in a bottom-up fashion is described in [21].

SLINK [22] is the most time-wise effective implementation of single linkage hierarchical clustering. It works incrementally, building several linear indexes.

BIRCH [23] also uses a hierarchical clustering, but the hierarchy is built on so called clustering features. A clustering feature statistically summarizes the underlying data.

DBSCAN [24] searches for its nearest neighbors when placing a new point. If there are sufficiently enough points under a minimal distance of the new point, such point is then added into the respective cluster of the nearest nodes. A generalization of DBSCAN called OPTICS [25] work with a varying density of clusters.

CURE [26] uses yet another approach to clustering. It lies between BIRCH and SLINK, as it uses hierarchical clustering, but instead of representing the cluster with once center as BIRCH, or considering all points as SLINK, it chooses only several representatives of the cluster that are then moved closed to the center.

Algorithms such as DBSCAN, OPTICS and DENCLUE, STING, CLIQUE, Wave-Cluster and OPTIGRID do not optimize the k-means objective. An overview article of incremental data stream algorithms [27, 28, 29].

Order-independences is an important property of clustering algorithms. An algorithm is order-independent if it which the data is presented, otherwise, it is order-dependent. Most of the incremental algorithms presented above are order-dependent.

The major advantage for incremental algorithm is that, it can operate without having to revisit old data it is more applicable to situations with very large data sets that cannot fit in memory, and streaming data sets which are being produced for long periods of time. Additionally, these algorithms are better suited to data whose characteristics might be evolving over time.

One of the first incremental clustering algorithms was the leader algorithm describes by Hartigan [30]. This algorithm attempts to partition a set of data samples into a number of disjoint clusters. The desired number of clusters is not specified but a distance metric and maximum threshold must be defined.

The algorithm gets its name from the fact that cluster coordinates are set equal to the input used to create the cluster. This input is known as the leader. Once, the cluster is established new members can be assigned to it, but its coordinates are never adjusted from those of the leader. The decision to assign an input to an existing cluster instead of creating a new cluster is made based on distance. If the current input's distance is within a threshold from an existing cluster then it is assigned to that cluster, otherwise a new cluster is created. It should be noted that the algorithm assigns inputs to the first cluster that satisfies the threshold requirement rather than searching for the closest cluster.

The leader algorithm requires only one pass through the sample data so it has speed advantage over non-incremental techniques. A second advantage is that it does not require prior knowledge of the number of expected clusters. A drawback, however, is that the clustering is highly dependent on the other of the input samples. This dependence exists because samples are assigned to the first cluster that meets the threshold requirement rather than finding the nearest cluster.

It exists many different incremental clustering algorithms. Each of them attempts to improve clustering accuracy by adjusting the algorithm to the specific characteristics of the desired application. Though the overall flow of most incremental clustering algorithms is similar, they differ in their initialization of their conditions for creation of deletion of clusters.

4. Conclusion

This paper has presented a short review in most of all clustering techniques which can be deal with large data sets which requires a large memory, arrive continuously, and might change over time. The tradional clustering methods do not scale very large data sets since they either need several passes over the data or they create data structures that do not scale lineraly with the number of objects. Incremental and data stream clustering methods are gaining a lot of research which is done on improving these algorithms to apply them on real time applications.

In the future work, we are planning to research in this area, to seek new clustering algorithms which can deal with a large amount of data.

REFERENCES

1. MacQueen, J. B. (1967). Some Methods for classification and Analysis of Multivariate Observations. *Proceedings of 5th Berkeley*

Symposium on Mathematical Statistics and Probability. volume (1), 281–297.
2. Kaufman, L., and Roussew, P. J. (1990). Finding Groups in Data - An Introduction to Cluster Analysis. A Wiley-Science Publication.
3. Rokach, L., and Maimon, O. (2005). Clustering methods. *Data mining and knowledge discovery handbook*, 321-352, Springer US.
4. Sibson R. (1973). SLINK: an optimally efficient algorithm for the single-link cluster method. *The Computer Journal. British Computer Society*. 16 (1), 30–34.
5. Defays D. (1977). An efficient algorithm for a complete-link method. *The Computer Journal. British Computer Society.* 20 (4), 364–366.
6. Steinbach, M., Karypis, G., and Kumar, V. (2000). A comparison of document clustering techniques, 525-526. In KDD workshop.
7. Kriegel, H., Kröger, P, Sander, J. and Zimek, A. (2011). Density-based Clustering. *WIREs Data Mining and Knowledge Discovery.* 1 (3), 231–240.
8. Microsoft academic search: most cited data mining articles: DBSCAN is on rank 24.
9. Ester, M., Kriegel, H., Sander, J., and Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise, 226-231.
10. Ankerst, M., Breunig, M.M., Kriegel and H., Sander, J. (1999). OPTICS: Ordering Points To Identify the Clustering Structure, 49-60.
11. Achtert, E., Böhm, C., and Kröger, P. (2006). DeLi-Clu: Boosting Robustness, Completeness, Usability, and Efficiency of Hierarchical Clustering by a Closest Pair Ranking, 119-128.
12. Roy, S. and Bhattacharyya, D. K. (2005). An Approach to find Embedded Clusters Using Density Based Techniques. *Springer Verlag.* volume (3816), 523–535.
13. Aggarwal, C. C. (2003) A Framework for Diagnosing Changes in Evolving Data Streams. ACM SIGMOD Conference.
14. Babcock, B. and Other, (2002). Models and Issues in Data Stream Systems, ACM PODS Conference.
15. Domingos, P. and Hulten, G. (2000). Mining High-Speed Data Streams. ACM SIGKDD Conference.
16. Guha, S., Mishra, N., Motwani, R., O'Callaghan L. (2000). Clustering Data Streams. IEEE FOGS Conference.
17. O'Callaghan, L. and Other, (2002). Streaming-Data Algorithms for High-Quality Clustering. ICDE Conference.
18. Munro, J., and Paterson, M. (1980). Selection and Sorting with Limited Storage. *Theoretical Computer Science*, 315-323.
19. Henzinger, M., Raghvan, P., and Rajagopalan, S. (1998). Computing on Data Streams. *Digital Equipment Corporation*.
20. Douglas H Fisher. (1987). Knowledge acquisition via incremental conceptual clustering. *Machine learning*, 139-172.
21. Dwi H Widyantoro, Thomas R Ioerger, and John Yen. (2002). An incremental approach to building a cluster hierarchy. *In Data Mining ICDM.*, 705–708.
22. Robin Sibson. (1973). Slink: an optimally efficient algorithm for the single-link cluster method. *The Computer Journal*, 30–34.
23. Tian Zhang, Raghu Ramakrishnan, and Miron Livny. (1996). Birch: an efficient data clustering method for very large databases. *In ACM SIGMOD Record*, volume 25, 103–114.
24. Martin Ester, Kriegel Hans-Peter, J¨org Sander, and Xiaowei Xu. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In KDD, volume 96, 226–231.
25. Mihael Ankerst, Markus M Breunig, Hans-Peter Kriegel, and J¨org Sander. (1999). Optics: Ordering points to identify the clustering structure. *In ACM SIGMOD Record*, volume 28, 49–60.
26. Sudipto Guha, Rajeev Rastogi, and Kyuseok Shim. (1998). Cure: An efficient clustering algorithm for large databases. *SIGMOD Rec.*, volume, 27(2), 73–84.
27. Mohamed Medhat Gaber, Arkady Zaslavsky, and Shonali Krishnaswamy. (2005). Mining data streams: A review. *SIGMOD Rec.*, volume 34(2), 18–26.
28. Sudipto Guha, Nina Mishra, Rajeev Motwani, and Liadan O'Callaghan. (2000). Clustering data streams. *In Foundations of computer science*, 359–366.
29. Moses Charikar, Chandra Chekuri, Tom´as Feder, and Rajeev Motwani. (1997). Incremental clustering and dynamic information retrieval. ACM symposium on Theory of computing, 626–635.
30. Hartigan J. A. (1975) Clustering Algorithms. New York: John Wiley and Sons.
31. Aggarwal C. and Other, (2003). A framework for Clustering Evolving Data Streams. Proceedings of the 29th VLDB Conference.

Authors' contacts:
E-mail: mangelova@tu-plovdiv.bg
University: TU – Sofia, Plovdiv branch
Department: Computer Systems and Technologies

© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 23, 2017

DEVELOPMENT OF CLOUD COMPUTING BASED SCADA IN ELECTRICAL POWER SYSTEMS

DIMITAR GROZEV, GRISHA SPASOV, NIKOLAY KAKANAKOV, GALIDIYA PETROVA

Abstract: *The report presents the development of an experimental platform for the implementation of Cloud Computing in SCADA systems for power systems. The general idea of the experimental platform is to realize the basic modules of SCADA systems for measuring electrical power on cloud platforms. The logical network structure of the SCADA Management Center is presented. Virtualization of the different types of servers in the Control Center is considered together with the physical network solution for the implementation of the cloud platform.*

Key words: *Cloud computing, SCADA, Power systems, Cross vC NSX, SRM*

1. Basics

An electric power system typically involves generation, transmission, transformation, distribution, consumption, and dispatching processes. Electric power is generated and consumed simultaneously and is not available for mass storage or transportation. Thus, as the key to the electric power system, monitoring and dispatching guarantee the reliability and security of electric power generation, transmission, distribution, and consumption, and play a key role in providing high quality and economic electric power. An electric power dispatching and monitoring system is often called a supervisory control and data acquisition (SCADA) system, energy management system (EMS), or distribution management system (DMS) globally. In this article, we call it the electric power SCADA system. In the early stage, the electric power SCADA system mainly used multi computer architecture, consisting of single server and two server cluster systems. Currently, the system uses computer systems with a distributed open architecture [1].

The electric power SCADA system provides SCADA, automatic generation control (AGC), automatic voltage control (AVC), EMS, DMS, dispatcher training system (DTS), geographic information system (GIS), and other useful functions. The software application system is constructed and developed by using the standard CIM model, and software architecture has evolved from the Client/Server architecture to the current Browser/Server architecture. The new generation Smart Grid [2] dispatching and control system uses multi-core computer cluster technology to improve system reliability and processing capacity, and uses a service oriented architecture (SOA) to enhance system interoperability and achieve horizontal integration and vertical interconnection of power grid dispatching services. Hardware infrastructure of these application systems involves high performance servers, complex high speed computer networks, high performance and highly reliable data storage systems, and workstations. The SCADA software is developed based on the platform which consists of Windows, UNIX, and Linux operating systems, and is based on relational databases.

The whole system is connected through computer networks for data exchange and sharing, and application programs share information through an enterprise service bus (ESB). With the development of the traditional IT application systems, IT-based applications have been expanding deeply to another industry field, and encountered various problems and bottlenecks. The same is true for the electric power SCADA system, which is a professional IT application system. The future electric power dispatching center should have high computing capacity, and powerful information acquisition, integration, and analysis functions especially with the advancement of Smart Grid applications.

Existing centralized computing platforms of electric power systems can hardly meet the above requirements, which have become one of the major bottlenecks in the Smart Grid. Some of the major disadvantages of these platforms are as follows [3]:

- Low basic resource utilization and poor scalability. A large amount of basic

 ISSN 1310 - 8271

computing resources to meet the demands in peak hours are idle during off peak hours. To ensure reliability, lots of resources are redundant, and cannot be fully utilized. Contradictions are growing between energy demands and conservation policies. Due to the upgrade of business, the existing IT infrastructure cannot be reused. Currently, analysis and computing in the electric power system rely on the centralized computing platform in the dispatching center. Due to limited computing capacity, poor scalability, and high upgrade costs, large scale power systems suffer from insufficient data storage and analysis capabilities.

- Poor system interoperability leading to information islands. Parallel application systems have their own architectural features, and therefore resources cannot be exchanged or reused.

Increasing management costs and risks, and decreasing equipment utilization.

2. SCADA System Components

The main components of a traditional SCADA system applicable to power systems and the information connectivity between them are shown in Figure 1.

.

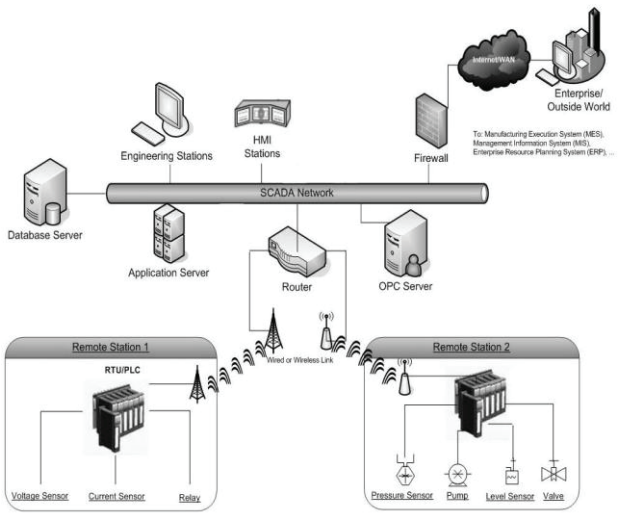


Fig. 1. *The traditional SCADA architecture.*

- RTU – remote terminal units are connected to sensors. They have telemetry hardware capable of sending digital data to the supervisory system, as well as receiving digital commands from the supervisory system.
- PLCs – programmable logic controllers are also connected to sensors. PLCs have more sophisticated embedded control capabilities than RTUs. PLCs do not have telemetry hardware, although this functionality is typically installed alongside them. PLCs are sometimes used in place of RTUs as field devices because they are more economical, versatile, flexible, and configurable
- Telemetry system is used to connect RTUs and PLCs to control center or data warehouse systems.
- A data acquisition server is a software service which uses industrial protocols to connect software services, via telemetry, with field devices such as RTUs and PLCs. It allows clients to access data from these field devices using standard protocols.
- HMI is the apparatus or device which presents processed data to a human operator, and through this, the human operator monitors and interacts with the process.6. A Historian is a software service which accumulates time-stamped data in a database which can be queried or used to populate graphic trends in the HMI. The historian is a client that requests data from a data acquisition server.
- A Historian is a software service which accumulates time-stamped data in a database which can be queried or used to populate graphic trends in the HMI. The historian is a client that requests data from a data acquisition server.
- A supervisory (computer) system, gathering (acquiring) data on the process and sending commands (control) to the SCADA system.
- Communication infrastructure.

In the implementation of modern SCADA systems the main accents are focused on the development of the communication infrastructure, on the other hand - on the development of the computer technologies and their integration in the sensor networks.

3. Cloud Computing Technologies For Electric Power Industry

Cloud computing can be used to solve various problems occurred during application deployment, use, and innovation process of the electric power SCADA system. Nowadays most of companies have their own cloud platform and virtual infrastructure, where they run business critical application and store data.

Private cloud advantages [4]:

- It is reliable and scalable. All resources are virtualized and in case of demand more storage or computing without downtime or impact can be added. With latest technology, virtual infrastructure can easily span across multiple data centers, without need to change/mapping IP addresses, FW

rules, or deploy sophisticated clusters. Enhanced disaster recovery is supported.

- Fast provisioning. Using techniques like templates can deploy thousands of machines with few clicks.
- Automation. Pretty much everything can be automated using powercli, REST API, powernsx, ansible and more. Common user interface: decoupling the computation infrastructure and the input system, enables multiple user interfaces to exist side by side allowing user-centric customization.

Cost effective.

4. PRIVATE CLOUD INFRASTRUCTURE

For the purposes of proposed experimental study, we use two virtual environments. Primary site - private cloud infrastructure in Technical University of Sofia, branch Plovdiv and secondary site - private cloud infrastructure in Technical University of Sofia. Physical topology of two sites is similar. Difference is in used hardware. In Primary site, we use Enterprise class storage, switches (Ethernet and FC) and servers, while in secondary site we have micro computers Intel NUC and Synology storage.

Figure 2 (2a and 2b) shows secondary and primary site physical topology.

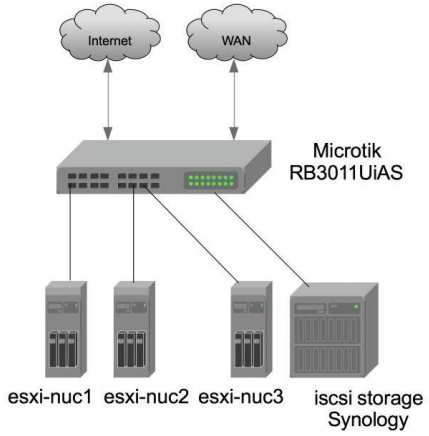


Fig. 2a. Secondary site cloud physical topology

Primary site private cloud in use is shared by the university departments and students. Everyone can request computing capacity and after an approval process one could receive VMs with requested parameters. Secondary site is dedicated only to SCADA system and it's used to cover HA requirements implementing disaster recovery site.

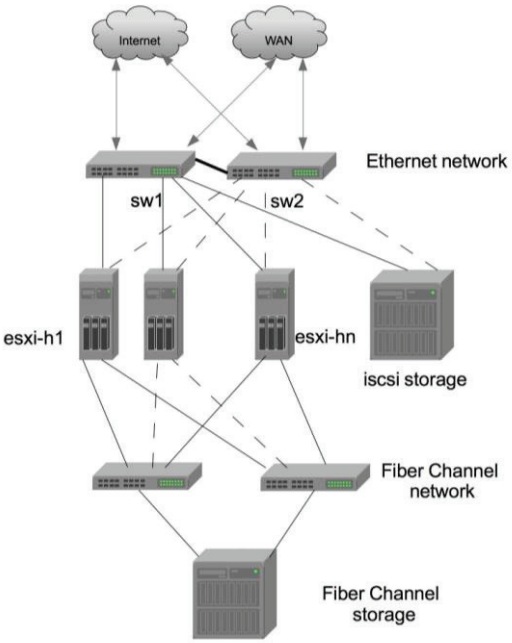


Fig. 2b. *Primary site cloud physical topology*

The following types of traffic are recognized:
Local in each site:

- Management – this include all traffic generated by virtual infrastructure: vMotion, iSCSI, NFS, fault tolerance, web-based graphical user interface, etc.
- VM traffic – this is traffic generated from all VMs connected to VLAN based port groups – various types, generated by different applications.
- VXLAN traffic – this include traffic generated from all VMs connected to logical switches - various types, generated by different applications.
- SCADA traffic – this include traffic generated from SCADA system servers

Between sites using WAN connectivity:

- Traffic generated by VMware Site Recovery Manager (SRM)
- VMs and SCADA traffic

Because of the importance and sensitivity of the SCADA traffic we implement partial application failover disaster recovery scenario using VMware site recovery manager (SRM) and Cross vCenter NSX deployment [5] [8]. All SCADA components are protected by SRM. Cross vCNSX technology solves issues like IP address mapping and FW policy. Universal Distributed FW allows deployment of security policy that span between primary and secondary site. Universal DLR along with universal distributed virtual switch span across sites too and allow same universal SCADA virtual switch on both locations. This way SCADA servers will keep their IP settings. (figure 3).

T

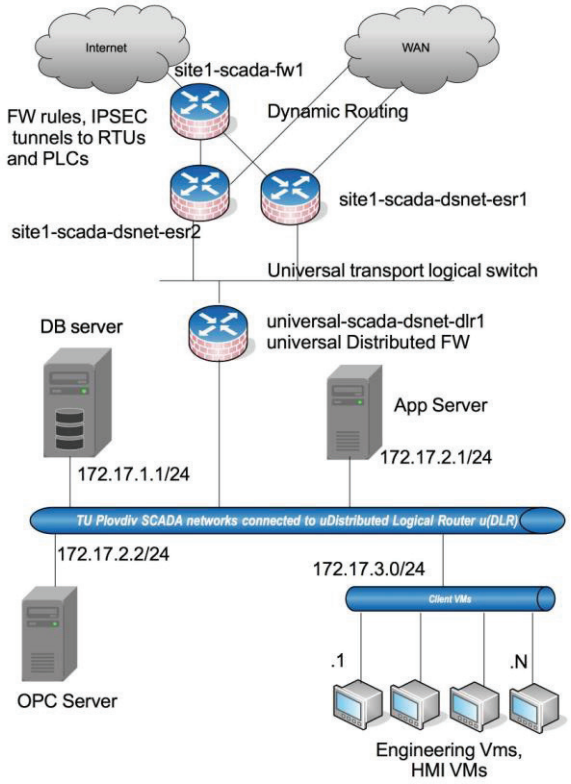


Fig. 3. *Private cloud network topology with universal objects: uDLR, uDVS*

WAN entry point of SCADA network aresite1-scada-dsnet-esr1 and esr2 - ECMP configured edges. They run dynamic routing protocol BGP and send updates for SCADA networks. Manipulate different BGP attributes in order to control SCADA networks entry point. Same pair of ECMP edges we have in both sites. Site1-scada-fw1 is NSX router connected to internet that has multiple features enabled: firewall, IPSec, SSL VPN support and load-balancing. It's an entry point for all internet accessible SCADA components like some RTUs and PLCs. To keep traffic private, we use IPSec and SSL tunnels. Through this secure channel data for controlled processes are collected. Universal-scada-dsnet-dlr1 is a universal distributed logical router (uDLR) [6][9] that optimize routing traffic between different SCADA components. Universal distributed FW [7][9] is used to control the communication.

5. Conclusion

The paper presents an implementation of a cloud-based SCADA system for use in electrical power control systems. The private cloud infrastructure in use is based on VMware platform. Deploying SCADA in cloud infrastructure automatically inherits all cloud advantages:

1. Most of SCADA components are virtual machines: DB server, Application Server, OPC server, Engineering and HMI VMs. This means that we can easily create new ones, clone, backup, take snapshots, HW upgrade - change parameters like CPUs, RAM, HDD, software upgrades.
2. All SCADA components can be protected by technologies like HA and FT [7].
3. All communication nodes are virtual nodes and thus are also protected by HA - scada-dsnet-esr1 and scada-dsnet-dlr1.
4. VMware Site Recovery Manager and cross vC NSX technology is used to ensure SCADA service is available even if we have Data Center or site failure.

The presented test bed platform will allow various types of information flows and configurations to be investigated and analyzed and the best set-up for the production services to be selected. Based on the experiments with different configuration of network services in the cloud the traffic flows will be optimized for best performance in terms of delay and resource usage.

Glossary

vSphere - VMware vSphere is the brand name for VMware's suite of virtualization products.

ESXi/ESX – Elastic Sky X. An enterprise-class, type-1 hypervisor developed by VMware for deploying and serving virtual computers.

VXLAN (Virtual Extensible LAN) – A Virtual Network that emulates an Ethernet broadcast domain.

vMotion– VMware vSphere live migration allows you to move an entire running virtual machine from one physical server to another, without downtime.

NSX - VMware NSX is the network virtualization platform for the Software-Defined Data Center (SDDC).

DRS (Distributed Resource Scheduler) – technology for balancing the computing capacity by cluster to deliver optimized performance for hosts and virtual machines.

VDS – VMware vSphere Distributed Switch (VDS) provides a centralized interface from which you can configure, monitor and administer virtual machine access switching for the entire data center.

DLR – Distributed Logical Router. It separates Control and Data plane. Control plane is a VM, but data plane is part of hypervisor kernel.

ECMP – Equal Cost Multi Pathing. Routing data across multiple links with same metric.

6. Acknowledgments

The presented work is supported by the National Science Fund of Bulgaria project "Investigation of methods and tools for application of cloud technologies in the measurement and control in the power system" under contract E02/12 (http://dsnet.tu-plovdiv.bg/energy/).

REFERENCES

[1] Arghira N., et al. Modern SCADA philosophy in power system operation – a survey, Scientific Bulletin, University POLITEHNICA Bucharest, Series C: Electrical Engineering, vol. 73, no. 2, pp. 153-166, 2011.

[2] Fang Xi, MisraSatyajayant, XueGuoliang, and Yang Dejun. Smart Grid – The New and Improved Power Grid: A Survey, IEEE Communications Surveys & Tutorials 2012, vol. 14, no. 4, pp. 944 – 980, 2012.

[3] Gang Tan, Li Zhao, Yaduan Wang, Jingwei Xu, and Honggang Liu. A Safety Design of Electric Cloud Computing Platform, In Proceedings of the Fourth International Conference on Computational and Information Sciences (ICCIS '12). IEEE Computer Society, Washington, DC, USA, pp. 868-871, 2012.

[4] Alamri A, Ansari WS, Hassan MM, Hossain MS, Alelaiwi A, Hossain MA. A survey on sensor-cloud: architecture, applications, and approaches, International Journal of Distributed Sensor Networks, Article ID 917923, Feb 2013.

[5] VMware NSX Technical Product Management Team. VMware® NSX for vSphere Network Virtualization Design Guide ver 3.0, Aug 21, 2014, online [04.02.2016]:

www.vmware.com/files/pdf/products/nsx/vmw-nsxnetwork-virtualization-design-guide.pdf

[6] B. Hedlund. Distributed virtual and physical routing in VMware NSX for vSphere, The Network Virtualization Blog, November 25, 2013. Online: [04.02.2016]:

https://blogs.vmware.com/networkvirtualization/2013/11/distribut ed-virtual-and-physical-routing-in-vmware-nsx-for-vsphere.html

[7] DebashisBasak, RohitToshniwal, Serge Maskalik, and AllwynSequeira.Virtualizing networking and security in the cloud. SIGOPS Oper. Syst. Rev. 44, 4 (December 2010), pp. 86-94, 2010.

[8] VMware NSBU team.Recovering NSX backed Data Ceneters utilizing SRM (Site Recovery Manager). Online [21.4.2016]:

https://blogs.vmware.com/networkvirtualization/2016/04/enhanced-disaster-recovery-with-nsx-and-srm.html#.WKWv6hD4Eg0

[9] VMware NSX for vSphere 6.2 Documentation Center, online [02.02.2017]:

https://pubs.vmware.com/NSX-62/index.jsp

Department of Computer Systems and Technologies, Department of Electronics, Faculty of Electronics and Automation
Technical University of Sofia, branch Plovdiv
25TsankoDiustabanov Str., 4000 Plovdiv, Bulgaria
E-mail: dimitargr@gmail.com, {gvs, kakanak, gip}@tu-plovdiv.bg

© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 23, 2017

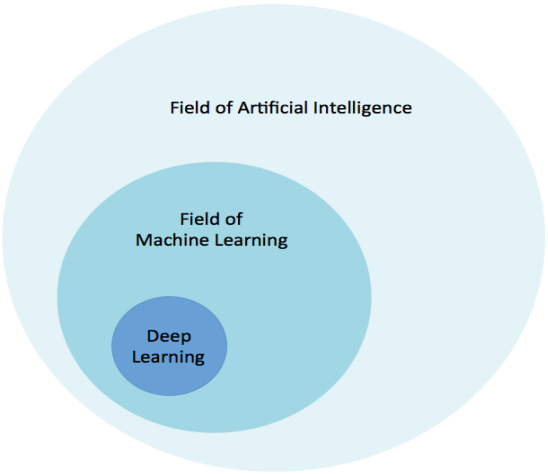
ALGORITHMS IN DEEP LEARNING

T. HRISTEVA, M.MARINOVA

Abstract: *Deep Learning is a software tool for learning in neural networks that utilizes multiple layers of abstraction. Today it has impact on many research areas. There are two broad types of machine learning – supervised learning and unsupervised learning. The supervised learning uses labeled datasets to produce the desired results. The supervised learning is useful for the classification and the regression. In the unsupervised learning the main purpose is to find out hidden models of data. The unsupervised learning is used for clustering and associative tasks.*

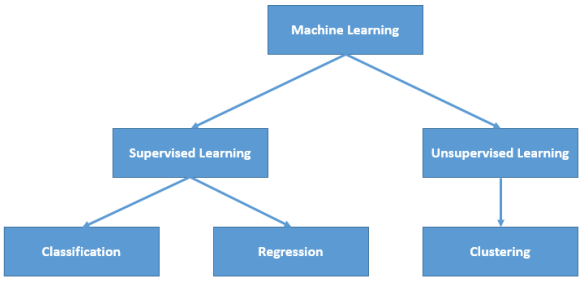
Key words: *deep learning, algorithm, neural network*

INTRODUCTION



Credit: http://blog.algorithmia.com

Deep learning is impacting everything from healthcare to transportation to manufacturing, and more. Companies are turning to deep learning to solve hard problems, like speech recognition, object recognition, and machine translation. It might seem like this technology is still years away, but we are beginning to see its commercial use. Such is the case with self-driving cars. Companies like Google, Tesla, and Uber are already testing autonomous cars on the streets.



Credit: https://cuteprogramming.wordpress.com

Supervised Learning

In supervised learning we have input variables (x) and an output variable (Y) and you use an algorithm to learn the mapping function from the input to the output(Y = f(X))

The goal is to approximate the mapping function so well that when you have new input data (x) that you can predict the output variables (Y) for that data.

It is called supervised learning because the process of an algorithm learning from the training dataset can be thought of as a teacher supervising the learning process. We know the correct answers, the algorithm iteratively makes predictions on the training data and is corrected by the teacher. Learning stops when the algorithm achieves an acceptable level of performance. Supervised learning problems can be further grouped into

 ISSN 1310 - 8271

regression and classification problems. Classification: A classification problem is when the output variable is a category, such as “red” or “blue” or “animal” and “human”

Regression: A regression problem is when the output variable is a real value, such as “dollars” or “weight”. Some common types of problems built on top of classification and regression include recommendation and time series prediction respectively.

A popular example of supervised machine learning algorithms is support vector machines for classification problems

Unsupervised Machine Learning

Unsupervised learning is where you only have input data (X) and no corresponding output variables.The goal for unsupervised learning is to model the underlying structure or distribution in the data in order to learn more about the data. These are called unsupervised learning because unlike supervised learning above there is no correct answers and there is no teacher. Algorithms are left to their own devises to discover and present the interesting structure in the data.

Unsupervised learning problems can be further grouped into clustering and association problems.

Clustering: A clustering problem is where you want to discover the inherent groupings in the data, such as grouping customers by purchasing behavior.

Association: An association rule learning problem is where you want to discover rules that describe large portions of your data, such as people that buy X also tend to buy Y.

An examples of unsupervised learning algorithms are: k-means for clustering problems.

MOST IMPORTANT ALGORITHMS

Supervised Learning

Decision Trees

A decision tree is a decision support tool that uses a tree-like graph or model of decisions and their possible consequences, including chance-event outcomes, resource costs, and utility.A decision tree is the minimum number of yes/no questions that one has to ask, to assess the probability of making a correct decision, most of the time. As a method, it allows you to approach the problem in a structured and systematic way to arrive at a logical conclusion.

Decision trees have many advantages.They are simple to understand and interpret. People are able to understand decision tree models after a brief explanation, allow the addition of new possible scenarios easily.Help determine worst, best and expected values for different scenarios, can be combined with other decision techniques.As a disadvantages can be given that the calculations can get very complex if many values are uncertain or if many outcomes are linked.

Logistic Regression

Logistic regression is a powerful statistical way of modeling a binomial outcome with one or more explanatory variables. It measures the relationship between the categorical dependent variable and one or more independent variables by estimating probabilities using a logistic function, which is the cumulative logistic distribution.

Regressions can be used in applications such as:

- Credit Scoring
- Measuring the success rates of marketing campaigns
- Predicting the revenues of a certain product

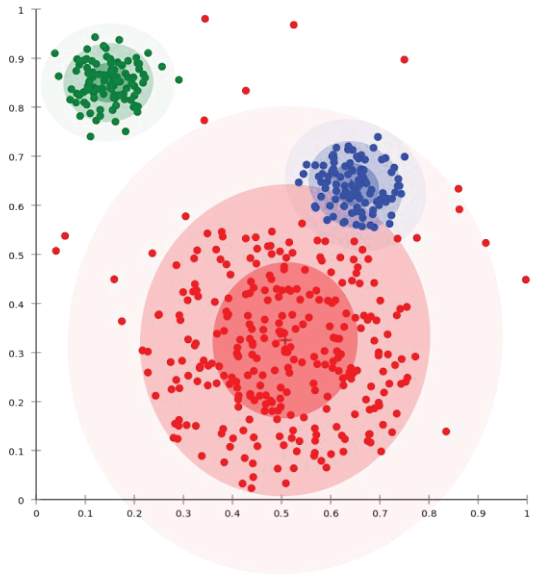
Support Vector Machines

SVM is binary classification algorithm. Given a set of points of 2 types in N dimensional place, SVM generates a (N—1) dimensional hyperlane to separate those points into 2 groups. Say you have some points of 2 types in a paper which are linearly separable. SVM will find a straight line which separates those points into 2 types and situated as far as possible from all those points.

In terms of scale, some of the biggest problems that have been solved using SVMs (with suitably modified implementations) are display advertising, human splice site recognition, image-based gender detection, large-scale image classification, etc.

Unsupervised Learning

Clustering Algorithms



Credits: http://www.machinelearningtutorial.net/

Clustering is the task of grouping data objects in such a way that objects in the same group (cluster) are more similar to each other than to the object in other groups.

Every clustering algorithm is different, and here are a couple of them:

- **Centroid-based algorithms**

In centroid-based clustering, clusters are represented by a central vector, which may not necessarily be a member of the data set. When the number of clusters is fixed to k, k-means clustering gives a formal definition as an optimization problem: find the {\displaystyle k} k cluster centers and assign the objects to the nearest cluster center, such that the squared distances from the cluster are minimized.

- **Connectivity-based algorithms**

Connectivity based clustering, also known as hierarchical clustering, is based on the core idea of objects being more related to nearby objects than to objects farther away. These algorithms connect "objects" to form "clusters" based on their distance. A cluster can be described largely by the maximum distance needed to connect parts of the cluster.

- **Density-based algorithms**

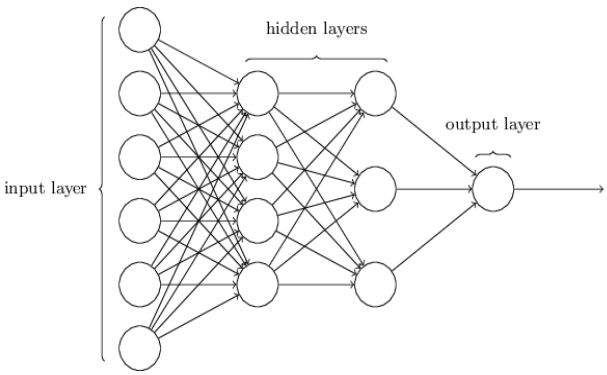
In density-based clustering, clusters are defined as areas of higher density than the remainder of the data set. Objects in these sparse areas - that are required to separate clusters - are usually considered to be noise and border points. The most popular density based clustering method is DBSCAN. In contrast to many newer methods, it features a well-defined cluster model called "density-reachability". Similar to linkage based clustering, it is based on connecting points within certain distance thresholds.

- **Neural networks.**

Neural networks are inspired by the structure of the cerebral cortex. At the basic level is the perceptron, the mathematical representation of a biological neuron. Like in the cerebral cortex, there can be several layers of interconnected perceptrons. The first layer is the input layer. Each node in this layer takes an input, and then passes its output as the input to each node in the next layer. There are generally no connections between nodes in the same layer and the last layer produces the outputs.

The middle part is called the hidden layer. These neurons have no connection to the outside (e.g. input or output) and are only activated by nodes in the previous layer.

Neural networks have the ability to derive meaning from complicated or imprecise data and can be used to extract patterns and detect trends that are too complex to be noticed by other computer techniques. A trained neural network can be thought of as an "expert" in the category of information it has been given to analyse.



Credit: Michael A. Nielsen, “Neural Networks and Deep Learning”

Deep learning can be described as the technique for learning in neural networks that utilizes multiple layers of abstraction to solve pattern recognition problems.It is a specific

approach used for building and training neural networks, which are considered highly promising decision-making nodes. An algorithm is considered to be deep if the input data is passed through a series of nonlinearities or nonlinear transformations before it becomes output. In contrast, most modern machine learning algorithms are considered "shallow" because the input can only go only a few levels of subroutine calling.

Deep learning removes the manual identification of features in data and, instead, relies on whatever training process it has in order to discover the useful patterns in the input examples. This makes training the neural network easier and faster, and it can yield a better result that advances the field of artificial intelligence.

CONCLUSION

The digital era has brought un important amount of data in all forms and from every region of the world. This data, known simply as Big Data, is gotten from sources like social media, internet search engines, e-commerce platforms, online cinemas, etc. This enormous amount of data is readily accessible and can be shared through fintech applications like cloud computing. However, the data, which normally is unstructured, is so vast that it could take decades for humans to comprehend it and extract relevant information. Companies realize the incredible potential that this information has and are increasingly adapting to systems for automated support. This is where deep learning fits perfectly.In contrast to more conventional machine learning and feature engineering algorithms, deep learning has an advantage of potentially providing a solution to address the data analysis and learning problems found in massive volumes of input data. More specifically, it aids in automatically extracting complex data representations from large volumes of unsupervised data. This makes it a valuable tool for Big Data Analytics, which involves data analysis from very large collections of raw data.

BIBLIOGRAPHY

1. https://www.udacity.com/course/machine-learning--ud262
2. http://machinelearningmastery.com/supervised-and-unsupervised-machine-learning-algorithms/
3. http://blog.algorithmia.com/introduction-to-deep-learning-2016/
4. http://www.kdnuggets.com/2016/08/10-algorithms-machine-learning-engineers.html
5. http://www.iro.umontreal.ca/~pift6266/H10/notes/mlp.html
6. http://neuralnetworksanddeeplearning.com
7. http://deeplearning.net
8. Ian Goodfellow, Yoshua Bengio, Aaron Courville-Deep Learning,MIT Press
9. Deep Learning Tutorial,LISA lab, University of Montreal
10. Deng,Dong Yu, Deep Learning: Methods and Applications

Teodora Hristeva
Technical University - Sofia, Branch Plovdiv
Telephone:+359 899 862 656
Email: thisteva@gmail.com
E-mail: thristeva@gmail.com

© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 23, 2017

NUMERICAL ANALYSIS OF BI-METALLIC PARTS USING A FINITE ELEMENT SOFTWARE

CENK MISIRLI, ISIK CETINTAV, YILMAZ CAN, CEM MISIRLI

Abstract: *Numerical estimation of upsetting processes have very important role in manufacturing world. In this investigation, bimetallic cylindrical parts have an outside material and a different solid material inside has been upset in one operation and behavior of this component during the operation investigated. Two bimetallic component models were considered. Each model has a ring material as steel. Model 1 has a solid core material as brass and Model 2 has a solid core material as copper. Parts were upset to different heights by using parallel flat dies in press machine. Press load was recorded during the process. Loads and metal flows have been compared and shown reasonable results for two models. Different finite element software calculated and run very close results with th experiment.*

Key words: *bimetallic, upsetting, finite element*

In metal forming world, open die forging is the most important and fundamental technique. Over decades, about open die forging of single materials, such as steel, copper, aluminum, lots of works were studied both theoretically and experimentally. Vilotic et al.[1] have given particular attention to process of upsetting of cylindrical billets. K. Baskaran and R. Narayanasamy investigated study the bulging characteristics of preformed irregularly shaped aluminium billets (Elliptical billets) during cold upset forging using lubricanton both sides of elliptical billets of pure commercial aluminium under different stress state conditions namely uniaxial, plane and triaxial stress states [2]. S. Malayappan et al. studied experimentally the bulging effect of aluminium solid cylinders by introducing an extrusion die at one end [3]. Aksakal et al. investigated upsetting of polygonal samples by using dual stream functions theoretically and experimentally [4]. Thaheer and Narayanasamy[5] studied barrelling effect of truncated cone billets of various materials. Also Thaheer and Narayanasamy explained barrelling effect of truncated cone billets of various materials [6]. Chang and Bramley[7] investigate the determination of the heat transfer coeffcient at the workpiece-die interface for the forging process. Dyja et al. [8] worked theoretical and experimental analysis of rolling of bimetallic Cu-Al and Cu- Steel rods. Kocanda et al.[9] examined contact pressure distribution in upsetting of compound cylindrical and cubic metals. Senthilkumar and Narayanasamy[10] evaluates some of the cold-forging features of composite steel preforms of varying titanium carbide contents during cold upsetting under triaxial stress state conditions.

Nowadays, compared to single components, especially in marine industry, chemical industry and electrical equipment, bimetallic components are mostly preferred in industry due to their some better properties. For example, bimetallic parts especially provide corrosion resistance and better conductivity. Bimetallic components have an advantage due to each two materials consist of different chemical compositions. For example, impurities and different grain size in materials show different plastic deformation characteristics. But bimetallic components are not composite materials, because in composite materials reinforcing elements spread uniformly in matrix materials.

In recent years, plastic forming of bimetallic parts have studied by a lot of researchers. Jingcai Wang et al.[11] investigated hot forging of multi-material cladded work pieces using upsetting tests, Yang et al. [12] studied by using upper bound method. Eivani and Tahiri [13] used equal channel angular extrusion process to produce bimetallic rods and compared to one produced by general extrusion process. Plancak et al. [14] studied experimental analysis of joining of two bimetallic axisymmetric components from various materials by upsetting bimetallic components in a closed die. Plancak et al. [15] examined the behaviour of bimetallic components which consist of an outer ring of mild

 ISSN 1310 - 8271

steel C45 and inner core of softer C15 steel during the cold upsetting process. Plancak et al. [16] then extended to this research and compared all the experimental data to one obtained by using finite element method.

Fig.1. Plancak Bi-metallic Model [16]

In this work, the open die forging of bimetallic cylindrical components producing using different materials designed as two models was examined both experimentally and analytically. This process may provide many parameters for metal forming operations, such as formability and barreling. The aim of this study was to examine and compare the metal flow and load requirement for these two models when upsetting of bimetallic cylindrical components.

2. EXPERIMENTAL WORK

For the interaction between different metals, two bimetallic models have been considered for this work. Material combinations and properties of these models have shown in Table 1. Basic design of Model 1 and Model 2 are shown in Fig. 2.

Table 1. Combinations of Model 1 and Model 2

Model 1: Ring material AISI 1020 STEEL
Solid inner cylinder material BRASS

Pb(%)	Al(%)	Fe(%)	Ni(%)	Sn(%)	Zn(%)	Cu(%)
1,6-2,5	0,05	0,3	0,3	0,3	Rest	57-59

Model 2: Ring material AISI 1020 STEEL
Solid inner cylinder material COPPER

O(%)	Cu(%)
0,04	99,9

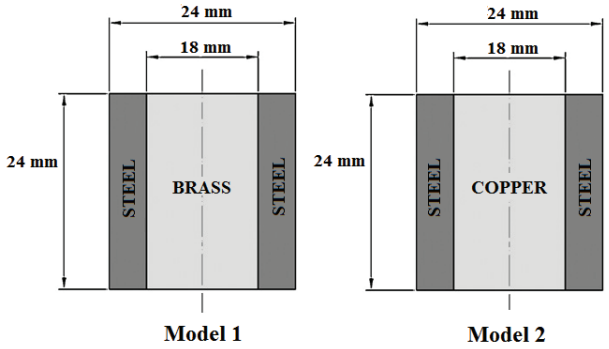


Fig.2. Design of Bimetallic Models

After prepared all the metals in lathe machine with appropriate dimensions, inner solid cylinder was pressed by low pressure into outer ring material in order to produce these bimetallic components. Produced bimetallic components are shown in Fig. 3.

Fig. 4 shows the 150 tons' hydraulic press that the experiments were carried out. It has a constant speed of 5 mm/s. The hydraulic press was equipped with a pressure current transducer in order to measure and record the experimental load. The pressure or load value was read by an I/O card of the personal computer during the experiment.

This value was converted into a resistance value, then this resistance value, consequently the pressure value was provided as an input value to the software.

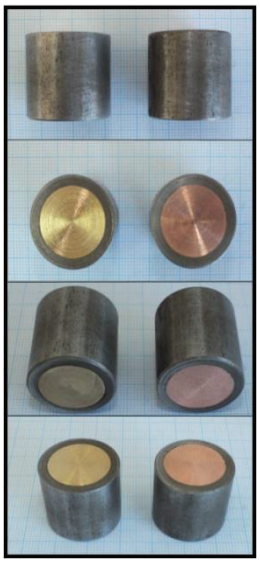


Fig.3. Produced Bimetallic Components

There were no special lubricants used but specimens and flat dies were sanded before the experiments. The fixation of components to hydraulic press is also shown in Fig. 4. Then bimetallic components were forged with reduction ratios %15, %30 and %40.

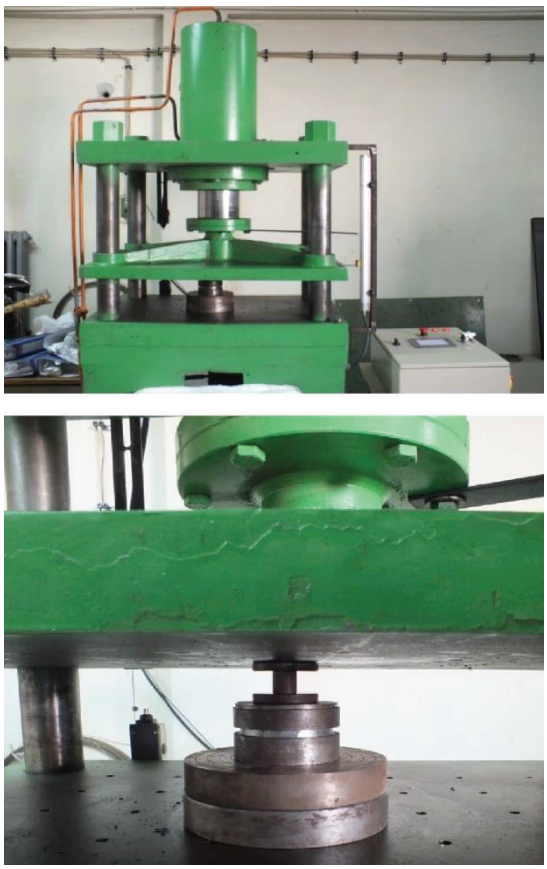


Fig.4. Hydraulic Press and fixation of parts

3. FE SIMULATION

In order to validate experimental work, a simulation model was designed using finite element softwares-DEFORM and ABAQUS. This gives more insight and helps to compare with experimental data of open die forging of bimetallic parts, also provides conducting further analyses about this work.

First, bimetallic parts were designed in Deform software as 2D model and due to the symmetry only one half of the parts were performed. All the metals properties were selected from Deform software library and they were assigned to the designed parts. For meshing, 8 node axisymmetric elements were used. For all the FE models, at the part-die interfaces a friction coefficient 0.1 is considered, but between the inner solid cylinder and the outer ring interface zero friction was assumed. Surface-to-surface contact was applied to all the friction locations. Fig.5 shows meshed component and interactions between all the contact surfaces.

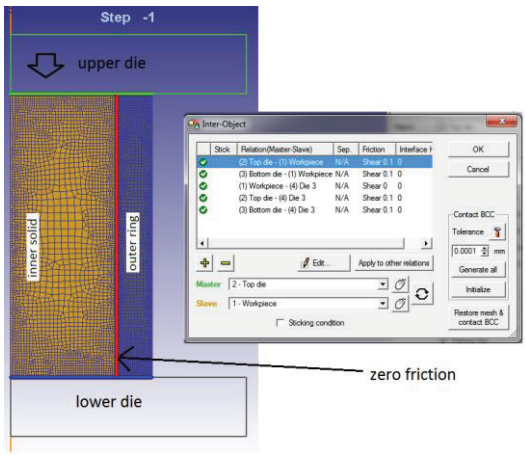


Fig.5. Meshed component and interactions

4. RESULT AND DISCUSSION

Open die forging of bimetallic cylindrical parts producing using different metals was studied experimentally and analytically for two different models. Three different metals have been used in order to produce bimetallic part models that have a ring and a solid cylinder. Model 1 has a ring material brass, and Model 2 has copper, each of these has steel solid cylinder. These parts were forged in an open die with reduction ratios %15, %30 and %40 experimentally and FE simulation model has run for the same conditions.

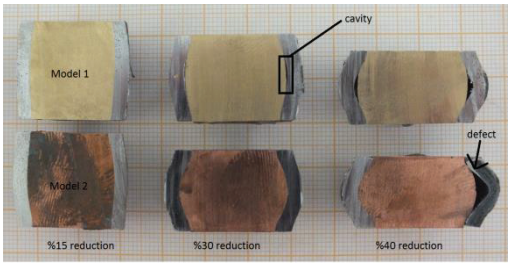
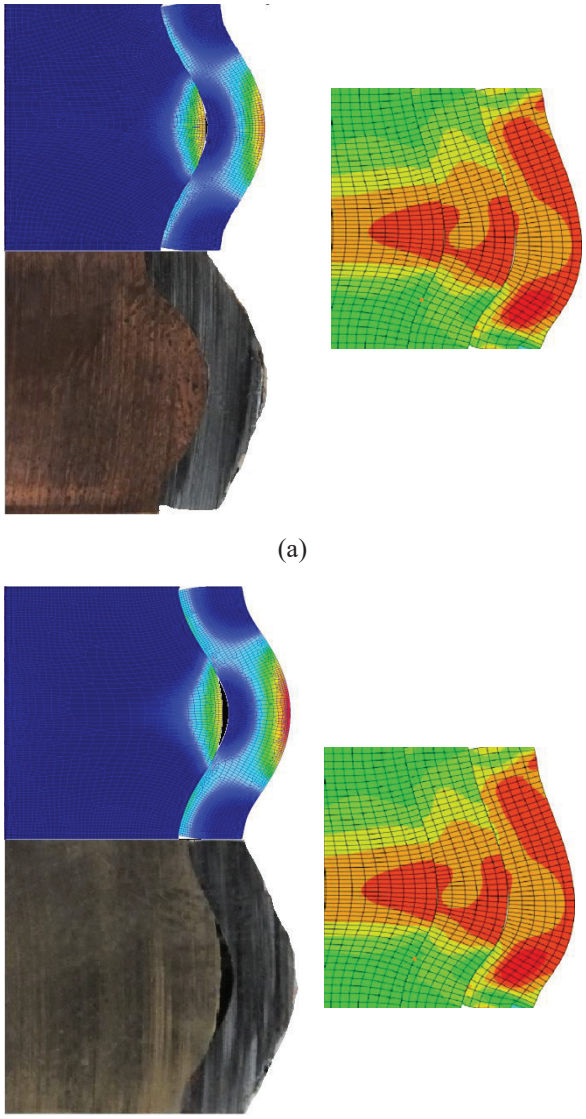


Fig.6. Meridional cut after %15, %30 and %40 reductions

After reduction, samples were cut and barrel profiles of the inner solid cylinder and outer ring were observed. Fig.6 and Fig. 7 shows that deformations and barreling of bimetallic parts after %15, %30 and %40 reductions. Clearly it is shown that Model 2 has much better material flow and zero cavity between the inner and outer interface. It is observed that the meridional cross section of inner cylinder material flows very fast and occurred a barrel profile is different from a normal barrel arc. Inner cylinder material barrel profile is not the same as outer ring material barrel profile. It can be also seen that in Model 1 there is no contact between inner and outer material and a cavity occurred in the inner meridian part of inner material. This cavity can be explained that inhomogeneous deformation

occurs in the centerline and this cavity is attributed to a state of hydrostatic stress. At %40 reduction Model 2 has a defect on the right side. It is guessed that high speed of hydraulic press and the close flow curves of copper and steel may cause this fracture.



(a)

(b)

Fig.7. Comparison of FE simulation and experimental work after %40 reduction, (a) Model 1 (b) Model 2

Fig.7 shows that comparison of FE simulation and experimental work after %40 reduction. FE simulation has really similar agreement with the experimental results.

5. CONCLUSIONS

- Obtained experimental results showed that metal flow has inhomogeneity for these two models especially for solid cylinder materials. The shift of ring and solid cylinder materials also affects metal flow and load requirement.
- A defect occurred at %40 reduction of Model 2 as seen in Fig.6. It is considered that due to the high speed of hydraulic press this defect has taken place. In the future works about this topic experimental works will be carry on lower hydraulic press speed.
- As considered the load-stroke and load-reduction graphs, Model 2 or Steel-Copper bi-metallic sample needs lower load to produce or use. Also it can be seen that Model 2 has better material flow than Model 1. Copper and steel are more useful to produce bi-metallic materials.
- This work has shown that under certain conditions, in a range of tolerances, bimetallic cylindrical components may be forged while carrying on a good interface contact between the two components. This establishes the possibility of producing bi-metallic components via this process either as a finished component or, more likely, as a pre-forming operation prior to further processing.
- In future works, flow curves of single metals copper, brass and steel will be obtain and investigate how effect the flow curves to form of bi-metallic samples.
- In order to obtain further information for producing different bimetallic components for different materials, more FEM simulation and more experimental work are planning.

6. ACKNOWLEDGEMENT

The authors would like to thank to Trakya University-Turkey (TUBAP-2011-193) for their help in experimental part of the study.

7. REFERENCES

[1] Vilotic, D.; Vujovic, V. & Plancak M. (1992). *Analysis of Upsetting of Cylinder by Conial Dies.* Journal for Technology of Plasticity. Vol., 17 (1-2), pp. 67-72

[2] K. Baskaran, R. Narayanasamy (2008). *Some aspects of barrelling in elliptical shaped billets of*

aluminium during cold upset forging with lubricant. Materials and Design 29, 638–661

[3] S. Malayappan, R. Narayanasamy, K. Kalidasamurugavel (2007). *A study on barrelling behaviour of aluminium billets during cold upsetting with an extrusion die constraint at one end.* Materials and Design 28, 954–961

[4] Aksakal, B.; Sezek, S. & Can Y. (2005). *Forging of Polygonal discs using the dual stream functions*. Materials & Design. Vol., 26, pp 643-654

[5] Thaheer, A. S. A. & Narayanasamy, R. (2007). *Barrelling in truncated lubricated zinc cone billets during cold upset forging*. Materials & Design. Vol., 28, pp. 434-440

[6] Thaheer, A. S. A. & Narayanasamy, R. (2007). *Barrelling in truncated lubricated zinc cone billets during cold upset forging*. Materials & Design. Vol., 28, pp. 434-440

[7] Zhang, H.G., Dean, T.A.,(1995). *Interaction of process and machine in multi-stage forging on a mechanical press.* Part B: Journal of Engineering Manufacture. Proc Instn Mech Engrs Vol 209, 19-32. [8] Dyja, H.; Mroz, s. & Milenin, A. (2004). *Theorical and experimental analysis of the rolling process of bimetallic rods Cu-steel and Cu-Al*. Journal of Materials Processing Technology. Vol. 153-154, pp. 100-107

[9] Kocanda, A.; Presz, W.; Adamczyk, G.; Czyzewski, P. & Mazurek, W. (1996). *Contact pressure disrtubition in upsetting of compound metals.* Journal of Materials Processing Technology. Vol., 60, pp. 343-348

[10] Senthilkumar, V., Narayanasamy, R., (2008). *Influence of titanium carbide particles addition on the forging behaviour of powder metallurgy composite steels.* Proc. IMechE Vol. 222 Part B: J. Engineering Manufacture, 1333-1345.

[11] Jingcai Wang, Laurent Langlois, Muhammad Rafiq, Régis Bigot, Hao Lu (2014). *Study of the hot forging of weld cladded work pieces using upsetting test.* Journal of Materials Processing Technology. 214, 365– 379

[12] Yang, D. Y.; Kim, Y. G. & Lee, C. M. (1991). *An upper-bound solution for axisymmetric extrusion of composite rods through curved*. International Journal of Machine Tools and Manufacture. Vol. 31, pp. 565-575

[13] Eivani, A. R. & Taheri, A. K. (2007). *A new method for producing bimetallic rods*. Materials Letters. Vol. 61, pp. 4110-4113

[14] Miroslav PLANČAK, Zlatan CAR, Dragiša VILOTIĆ, Dejan MOVRIN, Marko KRŠULJA (2011). *Upsetting of bimetallic components in a closed die*. Engineering Review Vol. 31, Issue 2, 105-109

[15] Miroslav PLANCAK, Igor KACMARCIK, Dragiša VILOTIC, Marko KRSULJA (2012). *Compression of bi-metallic components- analytical and experimental investigation*. International Journal of Engineering. 1584-2665

[16] K. Essa, I. Kacmarcik, P. Hartley, M. Plancak, D. Vilotic (2011). *Upsetting of bi-metallic ring billets*. Journal of Materials Processing Technology. 212 817– 824

© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 23, 2017

COMPUTER AIDED RESEARCH OF THE SHAFT OF A ROTOR CRUSHER OF AN ECCENTRIC TYPE

ZHIVKO ILIEV, GEORGI DINEV

Abstract: *The present article reviews the basic classification of a rotor crusher of an eccentric type as well as its usage in the field of mining industry. The analytical active forces of the eccentric shaft have been determined as well as the reactions in the bearing supports have been calculated. A 3D model of the crusher has been made as it has been used for the purpose of computer aided simulation research of one of the most stressed details, namely, the eccentric shaft. The obtained results have been analyzed as well as some recommendations have been given as they concern the possibilities for prolongation of the durability of the researched unit.*

Key words: ***mine transport,*** *rotor eccentric crusher, finite element method*

1. Basic classification

According to their pattern of action, the rotor crushers of eccentric origin are classified somewhere between the conic types and the jaw types of crushers. The process of grinding for this type of crushers is conducted by the use of two steep working chambers that shrink up and down as they are placed symmetrically in accordance with the vertical plane (fig 1). The vertical plane is the symmetrical plane of the machine itself. The operation of that type of a crusher is driven by an eccentric type of mechanism. There are four basic cinematic groups: grinding jaws, a rotor, an eccentric type of a shaft and a support protection system. Fig. 2 and fig.3 show cinematic sketches of an eccentric type of a rotor crusher along with a hydraulic type of a support protection system and an outlet that is hydraulically controlled.

The grinding chambers [1, 4, 16, 17] are placed between the curved operational surfaces (fig. 1) of the immobile jaws 1 and the cylindrical mobile rotor 3 that is mounted between them along with the shaft that is placed horizontally. Rotor 3 is located into the eccentric part of shaft 2 which is located in accordance with the basic pins. The shaft is fixed within the frame of the crusher in accordance with the pins. The crushing jaws 4 are located next to the both sides of the rotor as they are symmetrical to the basic pins of the eccentric shaft. They are mounted on the upper part as they hang fixed on axes 5 that are fixed within the frame of the crusher. Each jaw has hinged connection with the hydraulic cylinders 1 in its lower part as the jaws are fixed within the frame of the crusher by the use of special cylindrical pins.

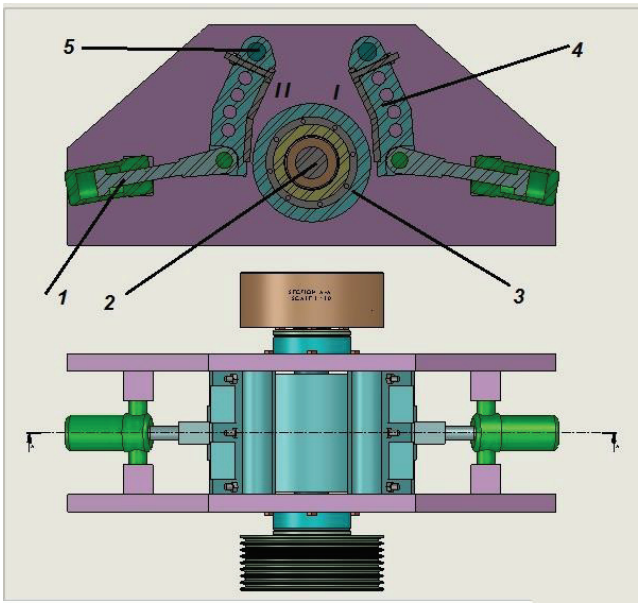


Fig.1 *Common look of a rotor crusher of an eccentric type*

 ISSN 1310 - 8271

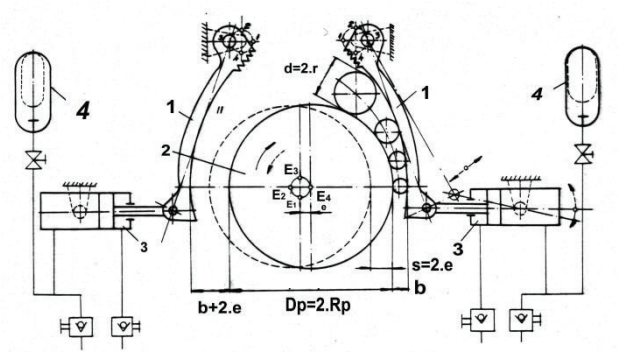


Fig. 2 *A cinematic sketch of a rotor crusher of an eccentric type and a hydraulically controlled outlet*

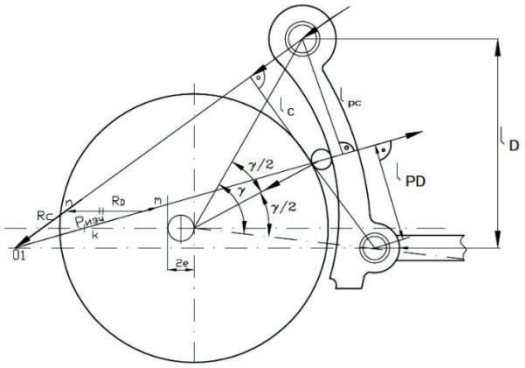


Fig. 3 *Forces of grinding of the rotor crusher of an eccentric type*

Figure 2 depicts a cinematic sketch of a rotor crusher of an eccentric type with hydraulic control of its outlet as it has the following composite details: 1 a grinding jaw, 2 mobile brake of a rotor, 3 hydro cylinders, 4 pneumatic-hydro accumulator.

Fig.3 illustrates a sketch of the forces of crushing of the crusher that shows $\theta = 10^0$ and $\gamma = 25^0$ – angles which combine forces Q and Pe along with axis *OY*. The forces of the belt gearing are depicted on Fig 4.

Force Q = 7452 N is the maximum force which is applied upon the shaft as a result of the forces of the belt gearing [3, 21].

2. Determination of the reactions in the bearing supports of the eccentric shaft.

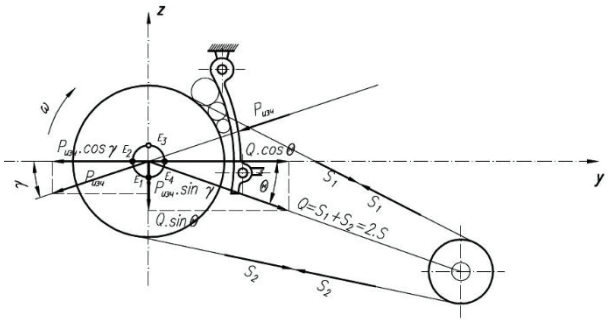


Fig. 4 *Forces within the belt gearing*

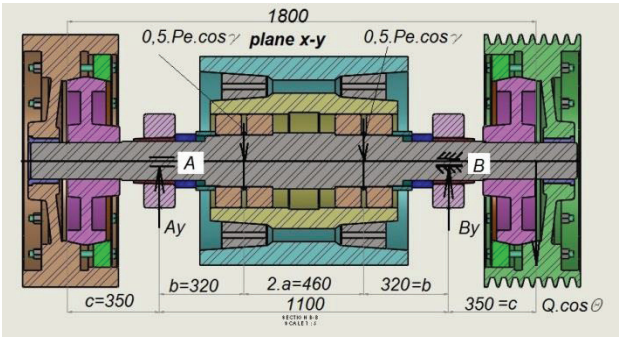


Fig. 5 *A sketch that is used for determination of the reactions of the bearing supports within the plane of X-Y*

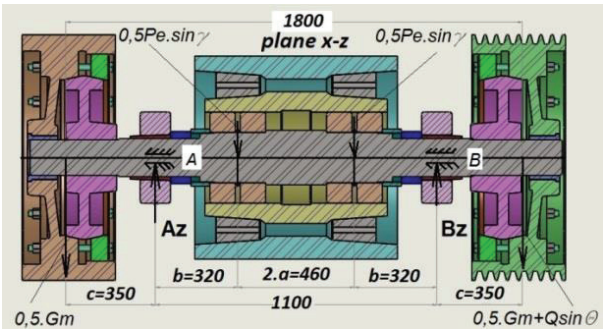


Fig. 6 *A sketch that is used for determination of the reactions in the bearing supports within the plane of X-Z*

The reactions Az and Bz within the bearing supports (Fig. 4) are to be defined by the moment equations:

Regarding plane ***X-Z*** *(Fig.6)* we have the following:

$$\sum M_B = 0; \qquad (1)$$

$$0{,}5.G_m(c+b+2.a+b) - A_z(2.a+2.b) + \\ +0{,}5.P_e.sin\gamma.(2a+b) + 0{,}5.P_e.sin.b - \\ -(0{,}5.G_m + Q.sin\theta).c = 0;$$

$$A_z = 53\,504, N$$

where ***2.a=460 mm; b=320mm; c=350mm*** is the geometrical size of the rotor head along with the shaft

$$\sum M_A = 0; \qquad (2)$$

$$0{,}5.G_m.c - 0{,}5.P_e.sin\gamma.b - 0{,}5.P_e.sin\gamma.(2.a+b) + B_z.(2.a+2.b) - \\ -(0{,}5.G_m + Q.sin\theta).(2.a+2.b+c) = 0;$$

$$B_z = 55\,620, N$$

where $G_m = 6800N$ –is the weight of the flywheel [2,3]

$$P_e = \frac{1000.\eta_m.N_{eng}}{4.e.z} = 239\,062N; \qquad (3)$$

Force P_e is perpendicularly [2] applied according to the operational plane of the grinding jaw (Fig.3).
where $\eta_m = 0{,}85$ – a mechanical К.П.Д.;
- $N_{eng.} = 90kW$ – power of the drive;
- e =10 mm – eccentricity of the shaft;
- z = 8 Hz –number of swaying movements of the shaft;

Regarding plane ***X-Y*** (Fig. 5) we have the following:

$$\sum M_A = 0; \qquad (4)$$

$$B_y.(2.a + 2.b) - 0{,}5.P_e.cos\gamma.b - 0{,}5.P_e cos\gamma(2a + b) + Q.cos\theta(2.a + 2.b + c) = 0$$

$$B_y = 206\,637, N$$

$$\sum M_B = 0; \qquad (5)$$

$$A_y.(2.a + 2.b) - 0{,}5.Pe.cos\gamma.(2.a + b) + +0{,}5.P_e.cos\gamma.b + Qcos\theta.c = 0$$

$$A_y = 105\,997, N$$

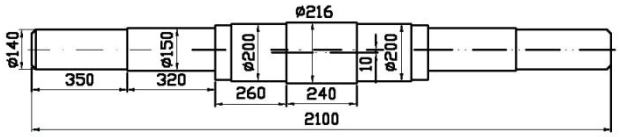


Fig.7 *A drawing of the eccentric shaft*

2.1. Methodology of researching

In order to increase the values of the exploitation indexes throughout the process of designing of the machine elements, one can use the capabilities of contemporary software products.

The first stage of execution of the research is creation of a 3D CAD model of the eccentric shaft of the rotor crusher *(fig. 7)*. The instance applies to a detail that is shown on *fig. 8*. which has been drawn by usage of Solid Works. Limitation with zero displacements has been imposed using the option of 'fixed' as it is required for the purpose of execution of the computer aided simulation. The consequent reactions within the bearing supports will affect the stress upon the shaft significantly. Concerning that fact, one has to assume the greater values, namely the results from plane X-Y, as they represent a more serious case of operation of the shaft.

A computer aided simulative and structural analysis has been conducted in order to determine size and distribution of the equivalent stresses.

The material that is assumed to be used for the explored detail is steel type S 235 J2G3 БДС EN 10025, as it sustains allowable stresses of approximately $[\sigma_{AL}] = 160MPa$.

2.2. Results from the static analysis that was conducted.

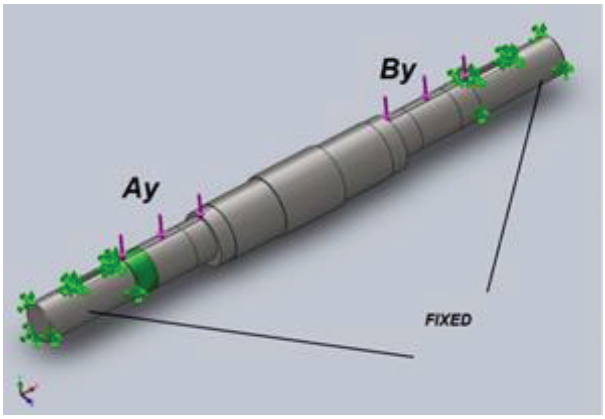


Fig.8 *Loading applied upon the eccentric shaft*

The capabilities for MKE investigation of the software product will be applied throughout the process of researching by simulation. These are as follows:

- Equivalent stress according to the Theory of von Mises [5, 6, 18, 19, 20];
- Reliability (Safety) coefficient according to the Theory of Mises-Henky

According to the Theory of von Mises-Henky, the limit of protraction is to be defined by the correspondence between the amount of equivalent stress according to von Mises σ_{VON} and the amount of allowable stress $[\sigma_{AL}]$ [7]

$$\sigma_{VON} \geq [\sigma_{AL}]$$

The stress types, according to von Mises, can be defined by the three main types of stress in correspondence with dependency [6, 11, 12, 13].

$$\sigma_{VON} = \sqrt{\frac{(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_1 - \sigma_3)^2}{2}} \qquad (6)$$

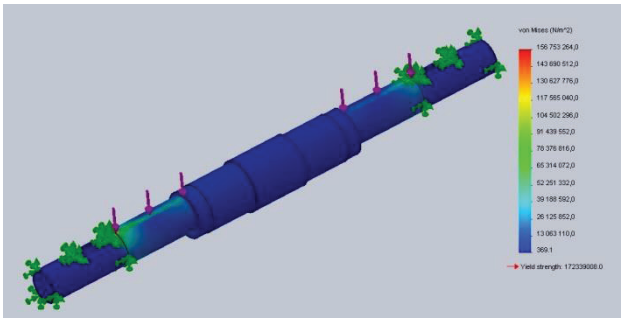


Fig. 9 *Distribution of the equivalent types of stress within the eccentric shaft*

3. Application of Expert system for design filet curve.

Production rules are a formalism used in design of program languages before it was used in psychological modeling [7] and in Expert systems [8],[9].

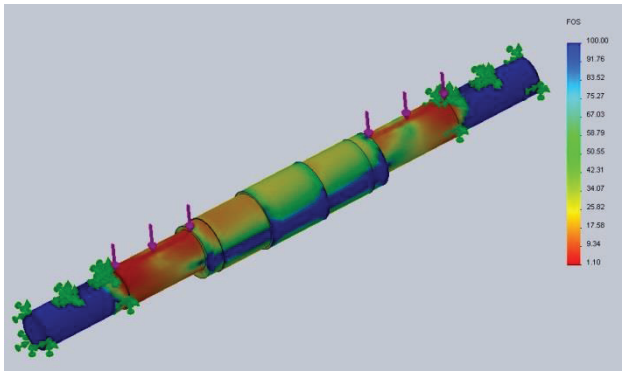


Fig. 10 *Distribution of the Factor of Reliability (Safety) within the eccentric shaft*

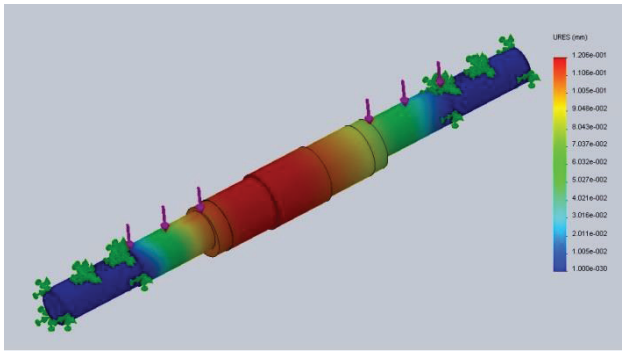


Fig. 11 *Distribution of the equivalent displacements within the eccentric shaft*

Some of most popular model from "production rules" include: Rules from kind "IF" and "THEN"; Different kinds of network; Logistic equations for coding of facts and principles"[10].

In the literature concerning expert systems sometime they are call rules from type "condition-action" or "situation-action". This is so because in general they are used to code empiric relations between models of data presented to the system and actions, which the system must execute as a consequence.

3.1. Solving of conflicts

The strategy for solving conflicts is of particularly importance for the work of production system and for this purpose must be chosen precisely. The mechanisms for conflicts solving varied but three of them are very popular and often are used in combination thus forming a global direction regime.

Unity. One rule cannot be used more than once for one and the same data.

Actuality. The elements in operational memory in OPS5 contain a field for time, thus it is known on the period of wich cycle they are added to operational memory.

Specificity. The concrete elements, drawn from more specific rules, for example rules with greater number of condition, are more difficult to be satisfied and therefore are prefered before more general rules with smaller number of conditions.

3.2 Production Ruls

In the process of designing the shaft designer may use the production rules by carrying out simulation analysis by the finite element method searching the optimum value of the filet radius from the standpoint of stress concentration.

This is illustrated by the following algorithm:

If: 1. The radius of curve is 0,5 mm;
2. The stress $\sigma > 0{,}260E9.MPa$
Then: To change the radius of 1,5 mm.
If: The stress $\sigma > 0{,}200E9.MPa$
Then: To chose Steel grade 41Cr4
If: The stress < 0,200E9.MPa.
Then: The concentration of stress are lucky.

In next *Fig.12* are shown the results from investigation of shaft with Finite Element Methods-ANSYS by 3D geometrical model and determined theoretical distribution of stresses in character points and surfaces concentrators. The designer can to used *expert system* included some *algorithms*. For example let to illustrate choice of fillet curve of shaft.

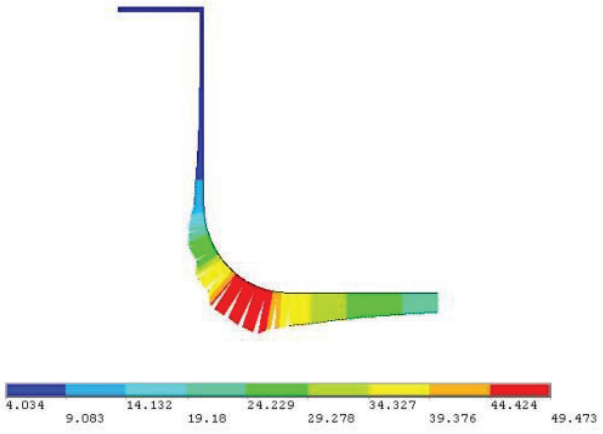


Fig. 12 *Distribution of stresses on fillet curve of shaft*

The results shown that maximum value for stresses are in left and right zone of fillet curve.

This value are from *49 MPa to 53 MPa.*in the both sides of shaft. The designer can to used the same algorithm.

4. Analysis of the obtained results and conclusions

Fig. 9 shows clearly that, according to Von Mises, the maximum amounts of stress within the shaft are of the value of *156,7 MPa.* All types of stress are concentrated within the bearing supports. In *Fig. 10*, it is evident that the minimum factor of reliability (Safety) *FOS = 1,1* is quite noticeable at the place of fixation. However, the maximum of equivalent displacements appears at the middle shoulder of the shaft (Fig. 11). This is due to the resilience of the detail. The maximum of displacements is of a value around 0, 12 mm. These results are normal and quite anticipated. It is possible that the factor of reliability (Safety) could be raised by constructive changes. An example of such changes is enlargement of the diameter of the vulnerable shoulders.

In the process of constructing machine parts using the capabilities of expert systems and the method of finite elements is improved the concentration of stress in their operation.

REFERENCES

1. Iliev Zh., Perenovski N. (2014) *"Opportunities for using the finite element method by computer research on rotary crusher of the eccentric type 800D x75"*, International multidisciplinary scientific symposium "Universitaria Simpro 2014", Petrosani, Romania..
2. Цанев Ц. (1973) „*Върху някои кинематични и конструктивни особености на роторна трошачка ексцентиков тип", сп.* Рудодобив, бр. 9, София.
3. Цветков Хр., Цанев Ц. (1976) „*Кинематика на роторна трошачка ексцентриков тип при задействане на предпазната й система."* Годишник на ВМГИ, София.
4. Dinev G, Zh. Iliev (2016) *"About computer design of machine and equipment the general machine building"*, International Journal "MACHINE, TECHNOLOGIES AND MATERIALS".
5. Lawrence, K.,(2005) *ANSYS Workbench Tutorial, SDC Publications*, Texas
6. Iliev Zh., (2016) " *Computer modal analysis of a eccentric shaft of a complex pendulum jaw crusher""5- National scientific and technical conference with international participation"*, "Technologies and Practices in Underground Mining and Mine Construction", Devin, Bulgaria.
7. Dinev G. *An approach for simulation design of mechanical assembled unit*, Journal Advanced Material Research , Vol. 463-464, Trans Tech Publication, Switzerland, (2012), pp.1085-1088.
8. Jackson P. *Introduction to expert systems. Allison Wasley Longman*, 1998.
9. Dinev G. *Expert systems used throughout the training of master engineering in the field of industrial engineering*, International virtual journal for science, technics and innovations for the industry, MTM, Year IX Issue 1/2015, ISSN 1313-0226, pp. 25-28.
10. Dinev G. *Course on principles of mechanical design.* In: AVANGARD PRIMA, 2011.
11. Iliev Zh., Dinev G.(2016) "*Improving the maintance of drive sheave of hanging ropeway",* International Conference on Engineering, Technologies and Systems, TECHSYS 2016, Technical University – Sofia, Plovdiv branch, Bulgaria.
12. Iliev Zh., Dinev G.(2015) *„About an education of engineering specialists for a computer design of the technical products in mining production",* Technologies and Systems, TECHSYS 2015, Technical University – Sofia, Plovdiv branch, Bulgaria.
13. Iliev Zh.,(2012) *"Opportunities for using finite element method by computer research of three roll directive devices of mine skip winding machines"* 12th INTERNATIONAL SCIENTIFIC CONFERENCE VSU'.
14. Iliev Zh.,(2012) "*Компютърно моделиране по метода на крайните елементи на двуролково направляващо устройство на подемни съдове за руднична клеткова уредба",* "Technologies and Practices in Underground Mining and Mine Construction", Devin, Bulgaria.
15. Iliev Zh., Perenovski N. (2016) *"Computer modeling and investigation of the operating parts of the screw conveyor by finite element method",* "Technologies and Practices in Underground Mining and Mine Construction", Devin, Bulgaria.
16. Iliev Zh., Bogdanov Iv.(2015)" *Analysis of the vibration state of the eccentric shaft with the bearings of a complex pendulum jaw crusher",* XVI Balkan mineral processing congress, Belgrad, Serbia.
17. Tsvetkov Kr., Iliev Zh., Perenovski N.,(2013)"*Power of electric motor and forces of breakage in jaw crushers with direct drive of the*

movable jaw." International Scientifis Conference, UNITECH 2013, Gabrovo, Bulgaria.
18. Iliev Zh., Perenovski N. (2015) *"Computer simulation research of hydrostatic thrust bearings of cone crusher type "kubria" by finite element method"*, Scientific journal 'Mechanics Transport Communications': 3 / 2015.
19. Iliev Zh., Perenovski N. (2014) *"Using finite element method by design of impact crusher for average crushing",* Scientific journal 'Mechanics Transport Communications': 3 / 2014 Faculty scientific conferences.
20. Radlov K., Iliev Zh. (2011) *"Opportunities for using finite element method by design of disk feeder for mining and building materials",* International Scientific Conference, UNITECH 2011, Gabrovo, Bulgaria.
21. Цанев Ц. (1976), *"Изследване на функционалната връзка между основните технологични и технически параметри на роторна трошачка ексцентриков тип* ",Дисертация, ВМГИ-София,

Faculty of Mining Electromechanics
University of Mining and Geology"St.Ivan Rilski", Sofia
Prof. Boyan Kamenov St.
1700 Sofia
BULGARIA
E-mail: halkopirit@mail.bg

Faculty of Mechanical Engineering
Technical University of Sofia
Bvd. Kliment Ochridski-8
1000 Sofia
BULGARIA
E-mail: gdinev@tu-sofia.bg

© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 23, 2017

COMPARATIVE STUDY OF THE MECHANICAL PROPERTIES OF TI AND CR BASED HARD COATINGS DEPOSITED BY CLOSE FIELD UNBALANCED MAGNETRON SPUTTERING

VASILIY CHITANOV, LILYANA KOLAKLIEVA, ROUMEN KAKANAKOV, TETIANA CHOLAKOVA

Abstract: *Two quaternary hard coating structures Ti/TiN/TiCrAlN and Cr/CrN/CrTiAlN, were deposited on HSS substrates by close field unbalanced magnetron sputtering (CFUBMS). A comparative study of their mechanical characteristics, nanohardness, a coefficient of friction and adhesion was carried out. The effect of the nitrogen flow controlled by an Optical Emission Monochromator (OEM) on the adhesion and transition layers was investigated. Among both investigated structures, the Cr-based coating showed better mechanical properties pronounced by higher nanohardness of 27 GPa, a lower coefficient of friction of 0.089 and excellent adhesion up to 30N loading.*

Key words: *coatings, nanohardness, adhesion, CFUBMS*

1. Introduction

Chromium nitride (CrN) and titanium nitride (TiN) coatings have well established background in advance machining operations [1]. The subsequent ternary nitride coatings like TiAlN [2] and CrAlN [3] exhibit improved tribological, mechanical, thermal and oxidation properties compared to the binary nitride based coatings. Besides, CrTiN coatings show enhanced mechanical behaviour [4]. The development of TiAlCr alloys allowed applications at temperatures as high as 1000 ^{o}C [5]. The CrTiAlN coatings obtained by Close Field Unbalanced Magnetron Sputtering (CFUMS) exhibit a high hardness and strong adhesion between the coating and substrate [6]. Furthermore, this technology allows deposition of compositionally graded structures, which improves the tribological characteristics of the coatings [7]. In this study, Ti/TiN/TiCrAlN and Cr/CrN/CrTiAlN graded coatings deposited by CFUBMS were comparatively studied in respect to their mechanical properties, a nanohardness and adhesion.

2. Experimental

The quaternary coating structures, Ti/TiN/TiCrAlN and Cr/CrN/CrTiAlN, named Ti-based and Cr-based, respectively, were deposited on high speed steel (HSS) specimens by close field unbalanced magnetron sputtering in UDP 800-4 equipment (Teer Coatings, UK). Figure 1 presents a scheme of the CFUBMS vacuum chamber used for deposition of hard coatings. Two Al targets, one Cr and one Ti targets were used for the coatings deposition.

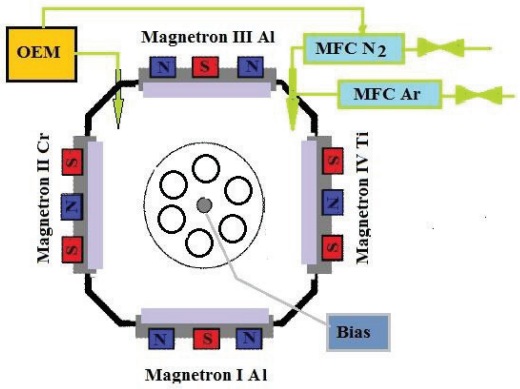


Fig. 1. *A vacuum chamber scheme of the UDP 800-4 equipment.*

Before the coating deposition the specimens were ultrasonically cleaned in an alkaline solution, rinsed in deionised water and dried at temperature of 130 °C. After that the substrates were loaded and the vacuum chamber was evacuated to $1.5x10^{-5}$ torr. The ion cleaning in an Ar plasma and at a bias voltage of − 500 V was performed for 30 min immediately before the coating deposition. To obtain better adhesion with the substrate an adhesion layer of Ti or Cr was

 ISSN 1310 - 8271

deposited in an Ar plasma. After that a transition layer of TiN or CrN was grown by a gradiently increased N_2 flow. The internal stress between the transition layer and the active coating layer was reduced with deposition of a gradient TiCrAlN or CrTiAlN layer. In this step, both parameters, the target power and nitrogen flow were increased to values optimal for the active top layer of the coating. The active coating CrTiAlN and TiCrAlN layers were deposited at constant values of the regime parameters. The total thickness of the Ti-based coating is 2.0 μm, while of the Cr-based it is 2.2 μm. During deposition, the nitrogen flow was controlled by an OEM monochromator, which was tuned to the Ti peak at 501 nm or to the Cr peak at 421nm. The target's current was kept the same and the OEM set points were chosen to fix the same N_2 flow during the graded layer deposition. The sputtering process was adjusted by the power and was controlled by the current. It was found out that best mechanical parameter were achieved at a current of 5 A, 8 A and 3 A of the Cr, Ti and both Al targets, respectively. The currents of the different material targets differ due to the different sputtering material rates and their reaction activity to N_2 [8]. The argon (Ar) flow was controlled by a Mass Flow Controller (MFC) and was kept at 25 sccm during the whole process. The pulsed bias voltage was maintained at −70 V during deposition. The working preassure during the process was between $1.6x10^{-3}$ and $1.9x10^{-3}$ torr. The carousel was rotating with 5 rpm.

The mechanical parameters were determined with a Compact platform CPX- MHT/NHT – CSM Instruments, Anton Paar, Austria. The nanohardness was measured with a nanoindentation module. A microindentation module with a Rockwell indenter was used performing a scratch test in order to study the adhesion and to determine the coefficient of friction against the diamond indenter. The nanoindentation module has a diamond Berkovich indenter that allows indentation in the interval 0.01-500 mN. The equipment software automatically calculates the nanohardness and elastic module depending on the penetration depth. The Oliver and Pharr methodology is implemented for the calculations. An indentation with a load of 200 mN was made for characterization the complex nanohardness including the coating and substrate influence. Besides, a loading of 15 mN was performed to eliminate the substrate effect keeping the well-established 10% rule of penetration in the coating. The scratch module has a certified diamond spherical Rockwell indenter with a radius of 200 μm and cone angle of 120°. The equipment allows loading in the interval of 0.1-30 N. The precision of the measurement of the loading is 0.3 mN. The maximum loading depth is – 200 μm. The software allows display of the penetration depth, friction and load forces, acoustic emission and coefficient of friction. The scratch tests were made with a progressive load from 1 to 30 N and a length of 1 mm for each investigated sample. The scratch velocity was 0.5 N/min.

3. Mechanical properties

The load–displacement curves of the Ti-based coating at indentation loads of 15 mN and 200 mN are shown in Fig. 2.

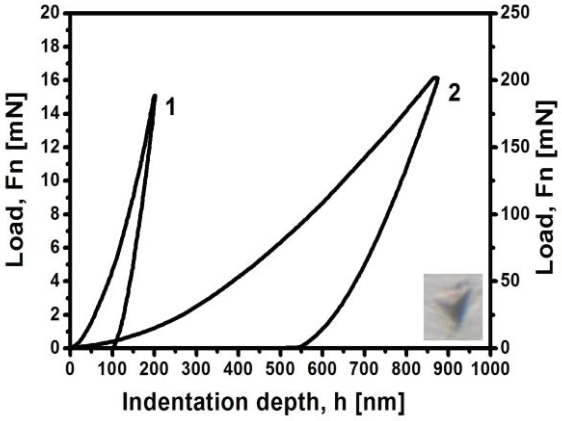


Fig. 2. *Load-displacement curves of the Ti/TiN/TiCrAlN coating measured at: 1 - 15 mN, 2 - 200 mN.*

The nanohardness H, elastic module E and maximum indentation depth h_{max} are summarized in Table 1. A highest nanohardness of 23 GPa was measured at h_{max}=202 nm, which value evidences that the influence of the substrate could be accepted negligible according to the 10 % rule. The corresponded module of elasticity was 271 GPa.

Table 1. *Nanohardness and elastic module*

Fn [mN]	**H [GPa]**	**E [GPa]**	**h_{max} [nm]**
Ti/TiN/TiCrAlN			
15	23.3	271	202
200	16.5	292	875
Cr/CrN/CrTiAlN			
15	27	385	185
200	15.4	305	889

The are few published results about the nanohardness of the TiCrAlN coatings deposited with CFUBMS. A value of 32 GPa is reported for the same coating structure obtained by cathodic arc evaporation [9]. This technique allows achievement of high ionisation rates and deposition of dense coatings with an increased nanohardness, but in a combination with formation of macro particles.

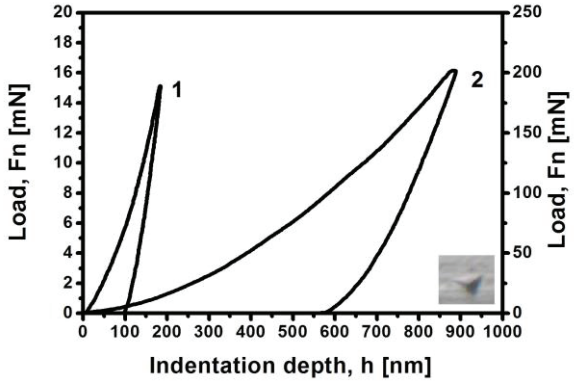


Fig. 3. *Load-displacement curves of the Cr/CrN/CrTiAlN coating measured at: 1-15 mN, 2- 200 mN.*

The results from measurements of the Cr-based structure are presented in Fig. 3 and Table 1. The inserted photo presents the residual indent on the coating surface after indentation with 200mN.

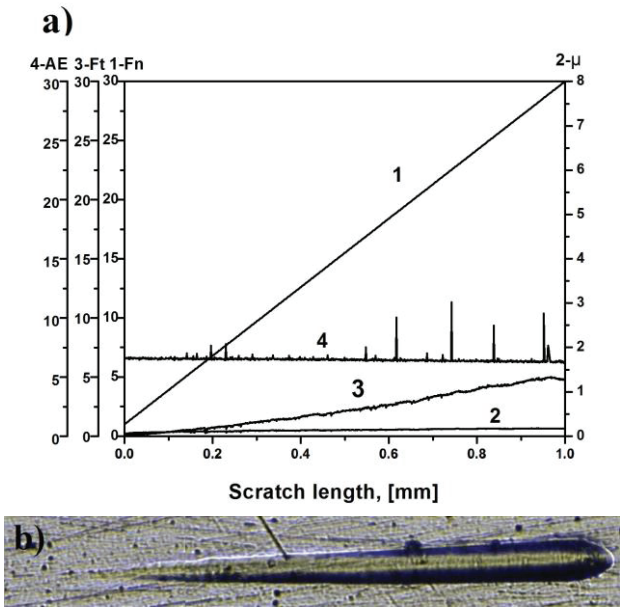


Fig. 4. *Scratch test results for the Ti/TiN/TiCrAlN coating:*
a)Graphic trends of the test process, 1-Load force, 2- Coefficient of friction, 3-Friction force, 4-Acustic emission;
b) A photo of the scratch test track;

For the Cr based coating, the measured nanohardness at the maximum indentation depth of 185 nm was 27 GPa and the elastic module was 385 GPa. The obtain result for the nanohardness is comparable with that reported in other publications for CFUBMS [10, 11] but still could be improved by an optimisation of the N_2 partial pressure [12].

The scratch test trends of the Ti based coating are presented in Fig. 4. Four signals are presented in the graph: the load force Fn increased from 1 to 30 N, coefficient of friction μ, friction force Ft and acoustic emission AE. As it is seen, there is a very good adhesion between the coating and the substrate. There is no delamination of the coating. The AE signal has several peaks, which correspond to rare small and short semi-circular and angular cracks, which are not considerable for the very good resistance of the coating during the scratch test. For the whole test loading up to 30N, coating chipping and spalling do not appear. The measured coefficient of friction at the end of the test was 0.167.

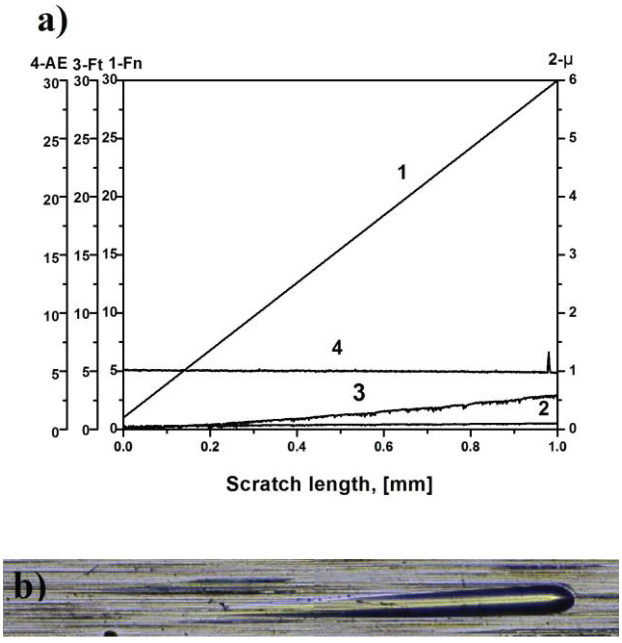


Fig. 5. *Scratch test results for the Cr/CrN/CrTiAlN coating:*
a) Graphic trends of the test process, 1-Load force, 2- Coefficient of friction, 3-Friction force, 4-Acustic emission;
b) A panoram photo of the scratch test track;

The scratch test results of the Cr/CrN/CrTiAlN coating are given in Fig. 5. This coating also have not delamination. The AE signal is clear and no cracks were found during the test. The only one pick at the end of AE signal could

be attributed to the unloading process of the indetor or defect in the substrate and it is not corresponding to a crack in the scratch track. The resistance of the coating during the scratch was excellent, i.e. the adhesion is excellent. The coefficient of friction at the end was 0.089.

The deposition of Ti–based and Cr-based quaternary structure at the same technological conditions allows the comparative study of the mechanical behavior resulting from the interactions between the Ti and Cr layers and the HSS substrate, as well as from the influence of TiN/ TiCrAlN and CrN/CrTiAlN transition layers. The N_2 flows during the main coating layers deposition were 21.2 sccm and 21 sccm for Ti-based and Cr-based coating, respectively. Thus, it can be consider that these layers were deposited at the same nitrogen flow.

The dependence of the hardness and elastic modulus on the loading is presented in Fig. 6. The mechanical parameters were measured in two characteristic points 15 mN and 200 mN, corresponding to the depth penetration giving information about the coating hardness itself and the hardness influenced by the substrate. It is seen that no significant difference of the coating hardness were observed at 15 mN loading. It is expectable because there are no differences in the technological parameters at which these layers were deposited. With the increase the load the composite hardness decreases due to the increased influence of the substrate.

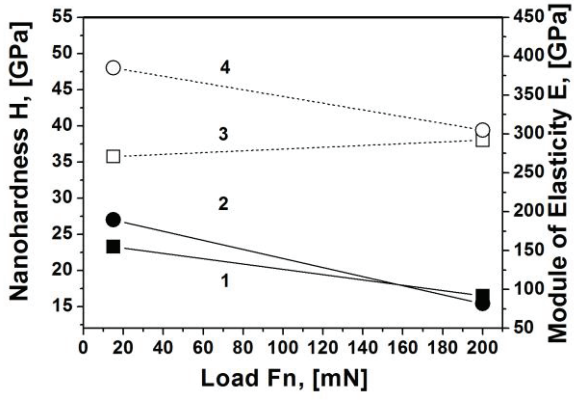


Fig. 6. *Dependence of the nanohardness and elastic modulus of the Ti- and Cr-based coatings:*
1- Nanohardness of Ti/TiN/TiCrAlN coating,
2- Nanohardness of Cr/CrN/CrTiAlN coating,
3- Module of elasticity of Ti/TiN/TiCrAlN coating,
4- Module of elasticity of Cr/CrN/CrTiAlN coating.

The scratch parameters for both coatings are presented in Table 2. It can be seen that the Cr-based coating has a better coefficient of friction and excellent resistance to the progressive load from 1 to 30 N, while the coefficient of friction is higher and the scratch resistance is very good with small non-essential cracks of the Ti-based coating. This difference in the scratch resistance is because of the better adhesion strength and coefficient of friction of the CrN layer than the TiN one. No delamination was observed in both scratch tracks, which implies the good adhesion to the substrates.

Table 2. *Comparison of the Scratch results*

Layer	Coefficient of friction	Scratch resistance to progressive load	Adhesion
TiCrAlN	0.167	Very good	yes
CrTiAlN	0.089	Excellent	yes

4. Conclusion

The investigated quaternary Ti- and Cr-based coatings consist of an adhesition, transition and main layers. The technological parameters during the main layer deposition were set the same for both coating structures.

The study of the mechanical properties of both structure types showed that the same deposition conditions affect the same coating hardness. However, the different composition of the adhesion and transition layers lead to different adhesion to the substrate.

Because of the better scratch properties of the Cr/CrN interlayers the Cr/CrN/CrTiAlN structure has lower friction coefficient of 0.089 and excellent scratch resistance to loads up to 30 N, which defines it as more suitable for industrial applications.

REFERENCES

1. Cakan, A., Ozkaner V. and Yildirim M. M. (2008). Quantifying Cutting and Wearing Behaviors of TiN- and CrN Coated AISI 1070 Steel. *Sensors*, volume(*8*), 6984-6998.
2. Chauhan, K. V. and Rawal, S. K. (2014). A review paper on tribological and mechanical properties of ternary nitride based coatings, *Procedia Technology*, volume (14), 430-437.
3. Chim Y.C., Ding X. Z., Zeng X. T. and Zhang S. (2009). Oxidation resistance of TiN, CrN, TiAlN and CrAlN coatings deposited by

lateral rotating cathode arc, *Thin Solid Films*, volume (517), 4845–4849.
4. Zhanga, G.A., Yana, P.X., Wangb P., Chenb, Y.M. and Zhang J.Y. (2007). The structure and tribological behaviors of CrN and Cr–Ti–N coatings. *Applied Surface Science*, volume (517), 7353–7359.
5. Lee, J.K., Oh, M.H. and Wee D.M. (2001). Long-term oxidation properties of Al–Ti–Cr two-phase alloys as coating materials for TiAl alloys. *Intermetallics*, volume (10), 347-352.
6. Bai L., Zhu X., Xiao J. and He J. (2007). Study on thermal stability of CrTiAlN coating for dry drilling. Surface & Coatings Technology, volume (201), 5257–5260.
7. Kabir, M.S., Munroe, P., Zhou, Z. and Xie Z.(2016), Structure and mechanical properties of graded Cr/CrN/CrTiN coatings synthesized by close field unbalanced magnetron sputtering, *Surface & Coatings Technology*, volume (309), 779-789.
8. Danek M., Fernandes F., Cavaleiro, A., Polcar T., (2017) Influence of Cr additions on the structure and oxidation resistance of multilayered TiAlCrN films, *Surface & Coatings Technology*, volume (313), 158-167.
9. Saleem H., (2014) Thermal stability of cubic and nanocrystalline arc evaporated TiCrAlN coatings, Master's Thesis, Department of Physics, Chemistry, and Biology, Linköpings universitet.
10. Wu W., (2010) Development and characterization of novel low-friction wear-resitant multilayer nanocomposite CrAlTiCN coatings. Ph.D. thesis, University of Birmingham, 2010.
11. Zhou Z., Tam P., Shum P., Li K., (2009) High temperature oxidation of CrTiAlN hard coatings prepared by unbalanced magnetron sputtering, *This Solid Films*, volume (517), 5243-5247.
12. Kim Y., Lee H., Kim S.and Han J., (2007) Microstructural and Mechanical Properties of CrTiAlN Nanocomposite Thin Films Synthesized by Closed Field Unbalanced Magnetron Sputtering, *Journal of the Korean Physical Society*, volume (51), 1183-1186.

Organization:
Central Laboratory of Applied Physics,
Bulgarian Academy of Sciences
Address: Bulgaria, Plovdiv, 4000,
61, Sankt Petersburg Bld.
Phone: 00359(0) 893611035
00359 (0) 32 265515
E-mails: vchitanov@gmail.com,
ohmic@mbox.digsys.bg,
ipfban@mbox.digsys.bg,
ipfban-dve@mbox.digsys.bg

© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 23, 2017

SENSORS MASS INFLUENCE ON THE NATURAL FREQUENCY OF A CANTILEVER BEAM

SVETLIN STOYANOV

Abstract: *The influence of relation between the sensor mass and the beam mass to the natural frequencies of a cantilever beam is investigated theoretically. This influence is not simple, because it depends from the mode number, and a three dimensional plot is used to present the results.*

Key words: *natural frequency, cantilever beam, modal analysis, Fourier transform, Matlab.*

1. Introduction

Over recent decades, modal analysis has been widely used in aerospace, mechanical and civil engineering. One of its advanced applications is for damage detection in beams [1].

The **experimental modal analysis** makes use of input (excitation) and output (response) measurements to estimate modal parameters: natural frequencies, mode shapes, and damping ratios [2]. Usually, the input is applied through an impulse hammer.

The **operational modal analysis** carries out in operational environment. In this case, structures are exited naturally due to external dynamic forces or displacements. Therefore, the information about these input excitations is absent or incomplete [2, 3, 4].

The **theoretical modal analysis** tries to predict the values of the modal parameters by solving the eigenvalue problem of structures [3].

In [5], theoretical and experimental modal analysis of a cantilever beam is conducted and the results are compared. It is raised the problem about the influence of sensor mass on the beam natural frequencies. This justifies **the aim of this work** – to investigate theoretically the influence of relation between the sensor mass and the beam mass to the natural frequencies of a cantilever beam.

2. Theoretically obtaining of the natural frequencies

The frequency equation of a cantilever beam without a lumped mass added (i.e. without sensor mass accounting) is [6]

$$\cos(kL)\cosh(kL)+1=0\,, \tag{1}$$

where L is the beam length. The values of k for which this equation is satisfied are his characteristic numbers. Each of the characteristic numbers corresponds to one natural frequency ω:

$$\omega_n=(kL)^2\sqrt{\frac{EI}{mL^4}}\,, \tag{2}$$

where m is the mass of the one meter of the beam.

The acceleration sensor acts as lumped mass in the end of the cantilever beam. As a result of this, the cantilever beam is loaded with a constant force – the gravitational force of the sensor. Therefore, the differential equation of the forced vibration of the beam is used for the frequency equation derivation. The frequency of the distortion force is set to be zero and the frequency equation becomes [7, 5]:

$$\begin{vmatrix} A(kL)+kL\frac{m_1}{mL}(kL) & B(kL) \\ D(kL)+kL\frac{m_1}{mL}C(kL) & A(kL) \end{vmatrix}=0\,, \tag{3}$$

where:

$$A(kL)=\frac{1}{2}(\cosh(kL)+\cos(kL))\,, \tag{4}$$

$$B(kL)=\frac{1}{2}(\sinh(kL)+\sin(kL))\,, \tag{5}$$

$$C(kL)=\frac{1}{2}(\cosh(kL)-\cos(kL))\,, \tag{6}$$

 ISSN 1310 - 8271

$$D(kL)=\frac{1}{2}(\sinh(kL)-\sin(kL)) \quad (7)$$

are the Krilov's functions and m_1 is the mass of the sensor.

3. Experimental evidence of sensor mass influence

In [5], an experimental setup is created (fig. 1) and the acceleration of the free vibration of the end point of a cantilever beam is measured as a function of the time (fig. 2). From this time signal, the spectrograms are calculated using Fourier transform (fig. 3).

Fig. 1. *A photography of the experimental setup*

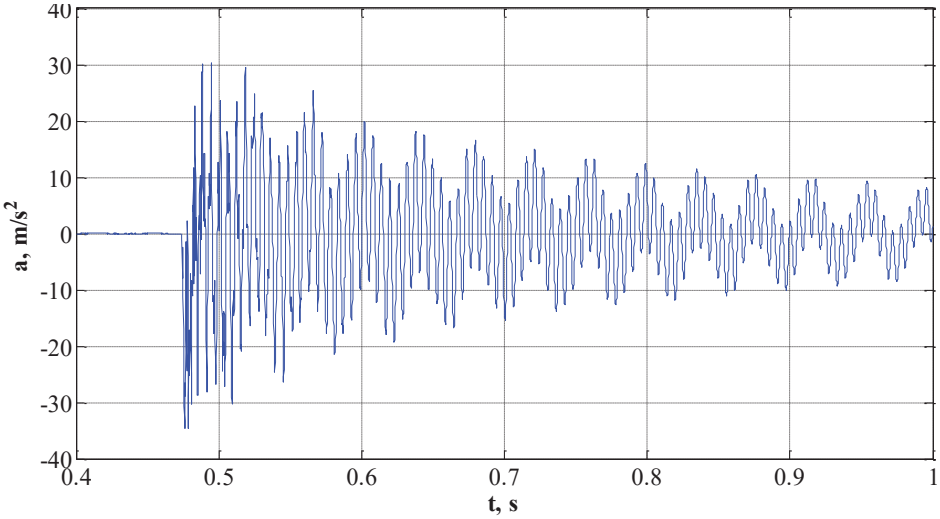


Fig. 2. *Time-diagram of the experimentally obtained acceleration of the end point of the beam*

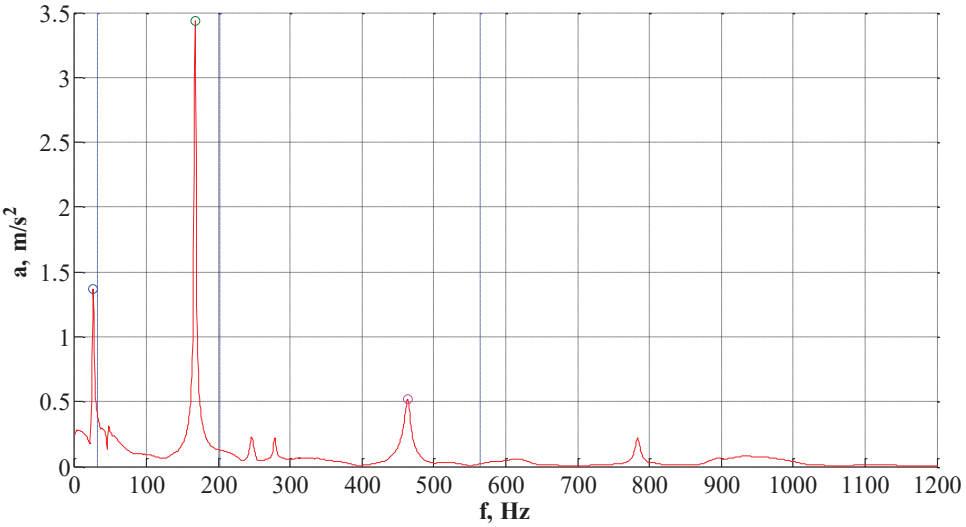


Fig. 3. *The experimentally obtained acceleration (red line) and the theoretically predicted natural frequencies without sensor mass accounting (blue marker lines)*

The peak points of the red line on Fig. 3 indicate the frequencies f_{exp} of the harmonics contained in the measured acceleration. These frequencies are shown in the first row of Table 1. The blue straight lines marks frequencies f_t, this is theoretically calculated without sensor mass accounting, according to Eq. 1 and Eq. 2. These frequencies are shown in the second row of Table 1. The third table row presents the relative difference.

Similarly, Fig. 4 and Table 2 show the corresponding frequencies for the case with sensor mass accounting, according to Eq. 3. One can observe significant better results in this case. This justifies the need for a detailed investigation.

Table 1. *Numerical comparison of the case without sensor mass accounting*

f_{exp}, Hz	26.25	167.85	463.26
f_t	32.173	201.85	564.562
$(f_{exp}-f_t)/f_{exp}$, %	-22.6	-20.1	-21.9

Table 2. *Numerical comparison of the case with sensor mass accounting*

F_{exp}, Hz	26.25	167.85	463.26
f_t	27.55	178.64	510.89
$(f_{exp}-f_t)/f_{exp}$, %	-5.0	-6.4	-10.3

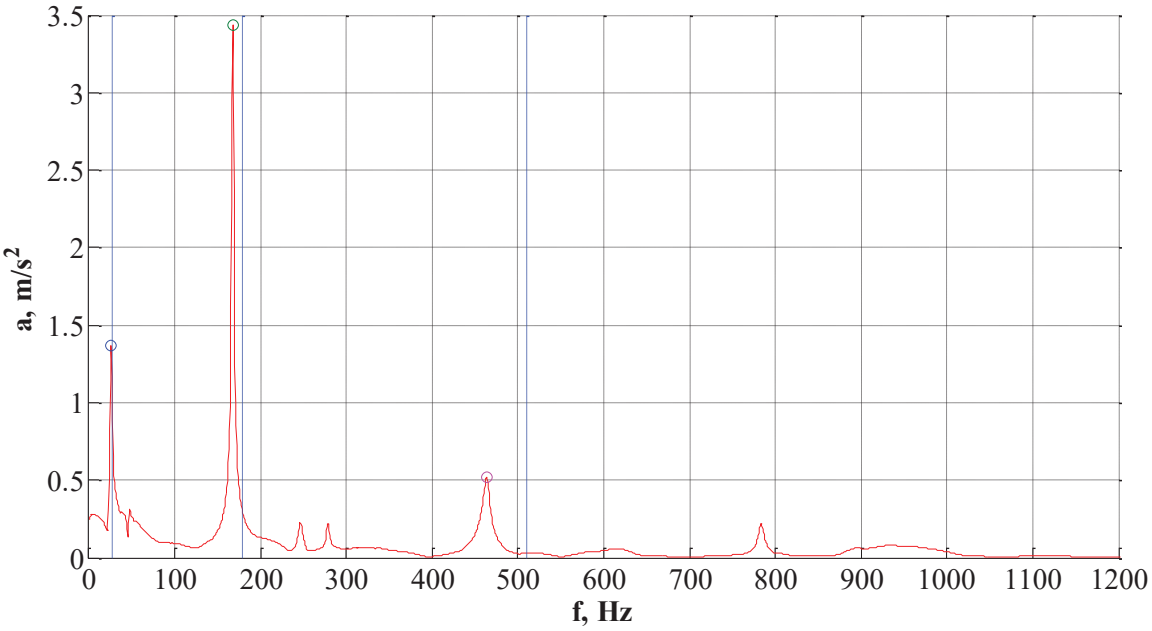


Fig. 4. *The experimentally obtained acceleration (red line) and the theoretically predicted natural frequencies with sensor mass accounting (blue marker lines)*

4. An investigation of sensor mass influence

To evaluate the influence of sensor mass on natural frequencies, two quantities are defined as follows:

- mass relation factor

$$\delta = \frac{m_1}{m} \qquad (8),$$

- frequency relation factor

$$\lambda = \frac{f_1}{f} \qquad (9),$$

where f_1 is the natural frequency for the case of sensor mass accounting (according to Eq. 3) and f is the natural frequency for the case without sensor mass accounting (according to Eq. 1 and Eq. 2)

Fig. 5 presents the influence of the mass relation factor to the frequency relation factor for the first three natural frequencies. As can be seen from this figure, the influence is the greatest for the first mode (red line). For the second mode and for the third mode, when the mass relation factor is greater than about one, the relation is close to a straight line, i.e. the change of the frequency relation factor can be neglected.

Fig. 6 presents the influence of the mode shape number to the frequency relation factor for six values of the mass relation factor. One can observe that the influence grows with sensor mass increasing.

The relations from Fig. 5 and Fig. 6 are combined together in a three dimensional plot – Fig. 7.

5. Conclusion

An experimental evidence of significant sensor mass influence on the cantilever beam natural frequencies is presented. This justifies the necessary to investigate theoretically this influence. The investigation is done through the frequency equation describing the eigenvalue problem.

The frequency equation and the equation of the forced vibrations of the beam are solved numerically by the help of the software system Matlab. The results obtained are presented and discussed. The influence of the sensor mass is not simple, because it depends from the mode number, and a three dimensional plot is used to present the results. It was found that the influence is the greatest for the first mode. Furthermore, it is of interest to obtain theoretically and experimentally the natural frequencies of a truss structure.

 ISSN 1310 - 8271

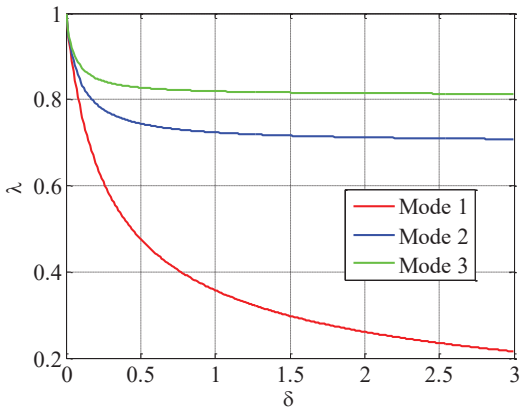


Fig. 5. *Dependences of the frequency relation factor on the mass relation factor*

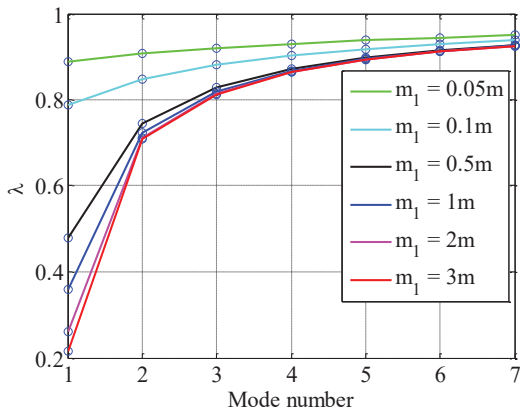


Fig. 6. *Dependences of the frequency relation factor on the mode number*

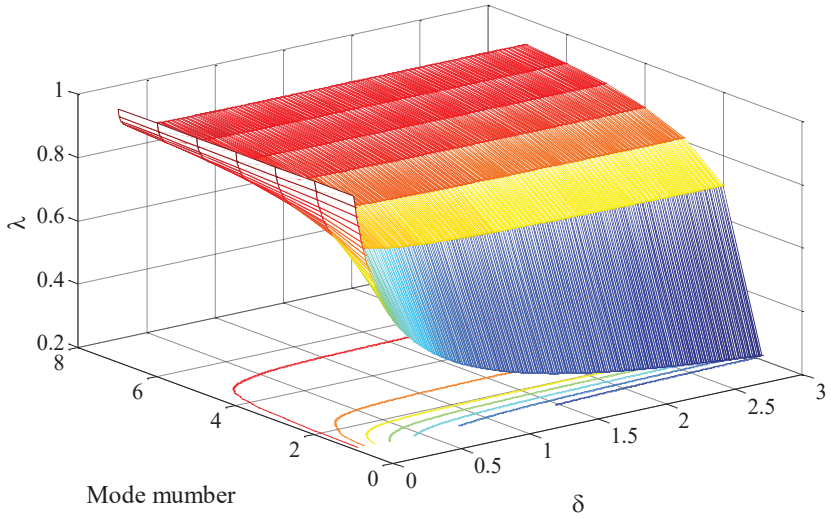


Fig. 7. *Three dimensional representation of the frequency relation factor dependence*

REFERENCES

1. Yang Z., Radzienski M. and Kudela P., Ostachowicz. (2017) 'Fourier spectral-based modal curvature analysis and its application to damage detection', *Mechanical Systems and Signal Processing*, vol. 84, pp. 763-781.
2. De Vivo A., Brutti C. and Leofanti J.L. (2013) 'Modal shape identification of large structure exposed to wind excitation by operational modal analysis technique', *Mechanical Systems and Signal Processing*, vol. 39, pp. 195–206.
3. Aenlle M.L. and Brincker R. (2013) 'Modal scaling in operational modal analysis using a finite element model', *International Journal of Mechanical Sciences*, v76, pp. 86-101.
4. Dziedziech K., Staszewski W.J. and Uhl T. (2015) 'Wavelet-based modal analysis for time-variant systems', *Mechanical Systems and Signal Processing*, vol. 50-51, pp. 323–337.
5. Stoyanov, S. (2016) 'Experimental setup for the determination of the natural frequencies of a cantilever beam', *Scientific proceedings of University of Ruse*, Ruse, ISSN 1311-3321. (In Bulgarian. In press.)
6. Zannon M. (2014) 'Free Vibration of Thin Film Cantilever Beam', *International Journal of Engineering and Technical Research (IJETR)*, vol. 2, pp. 304-314.
7. Repetto S. E., Roatta A. and Welti R.J. (2012) 'Forced vibrations of a cantilever beam', *European Journal of Physics,* vol. 33, pp. 1187-1195.

Authors' contacts
Address: 36 Zahari Stoianov str.,
POB 7005, Ruse, Bulgaria
Phone: 887-549-965
E-mail: sstoyanov@uni-ruse.bg

© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 23, 2017

THE APPLICATION OF THE MODAL ANALYSIS IN DETERMINING THE DYNAMIC MODULE OF ELASTICITY OF POLYMER-CONCRETE COMPOSITES

ILIA POPOV, SABI SABEV

Abstract: *This article presents the quantitative values of the dynamic modulus of elasticity (Young's modulus) of gamma polymer-concrete composites. Experimental studies are based on the experimental modal analysis method. Using the similarity and comparison methods, the results of the experimental determination of the dynamic modulus of elasticity of samples of gray cast iron GF20 and C45 steel with the same dimensions as PC samples are presented.*

Keywords: *Polymer-concrete Composites, Dynamic Modulus of Linear Deformations, Experimental Modal Analysis.*

1. Introduction

Knowing the material constants of structural engineering materials is a must when defining the strength deformation behaviour of the parts and structures produced by these materials.

In this aspect, the study and the determination of the quantitative values of the elastic characteristics (such as the modulus of the linear deformations E, the modulus of the angular deformations G, and the Poisson μ coefficient, which are in simple algebraic relations) of the gamma polymer-concrete composites is an up-to-date engineering task. These features, as well as the density of the material, are input data for any software product for static or dynamic engineering analysis (CAD, CAM CAE), which greatly assists in designing and enables reliable results to be obtained quickly.

The application of the polymer-concrete (PC) composites as an alternative non-metallic structural material for the production of bodies and body parts (B and BP) of the support system of the production equipment is an innovation activity which creates prerequisites for improving the technical and operational qualities of the manufactured machines such as:

- ✓ static stability;
- ✓ dynamic sustainability;
- ✓ thermal behaviour.

The article has an experimental-research character and its main objective is: Quantification of the dynamic elastic modulus E_d of gamma PC composites, created and realized in the testing and research laboratory of Metal cutting-machines at TU-Sofia, branch Plovdiv.

2. Theoretical prerequisites

Theoretically, natural beam frequencies with a rectangular cross section are determined by the dependence [2]:

$$\omega = \beta_n{}^2 \sqrt{\frac{E_d I}{\rho A}} = (\beta_n l)^2 \sqrt{\frac{E_d I}{\rho A l^4}} \qquad (1)$$

From equation (1) E_d is expressed and the following is obtained:

$$E_d = \frac{\omega^2 \rho b h l^4}{(\beta_n l)^4 I} \qquad (2)$$

Where

ω - is the frequency of vibration
b - beam thickness
h - height of the material
ρ - density of the material
n - density of the material
I - beam moment of inertia around y
$(\beta_n l)$ - depends on the boundary conditions of the beam.

$$(\beta_1 l = 4.730;\ \beta_2 l = 7.853;\ \beta_3 l = 10.996)$$

 ISSN 1310 - 8271

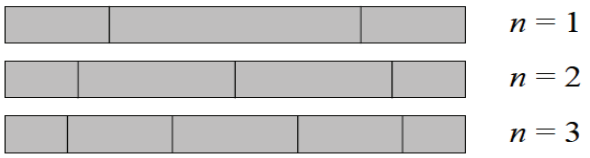


Fig. 1.1. *Forms for bending a rectangular beam with designated detail lines*

3. Methodology of the experiment

By impulse (shock) excitation of the system with a subsequent analysis of free damping oscillations. The impulse method is realized with an unattached exciter - an impact (power) hammer. Thus, the generated energy is distributed over a wide frequency spectrum and excites all the mods in the considered frequency range of the sample [4,5].

For the quantification of the dynamic elastic modulus, the following experimental approach was used: - A test piece of dimensions 30x30x350 of PC is suspended freely in strings, fig.1.2. At one end of the beam the impulse dynamic effect is applied and at the opposite side at the other end the repercussion is measured by the microphone.

The method is easy to implement, and only requires some basic components: a computer with a suitable sound card and a microphone.

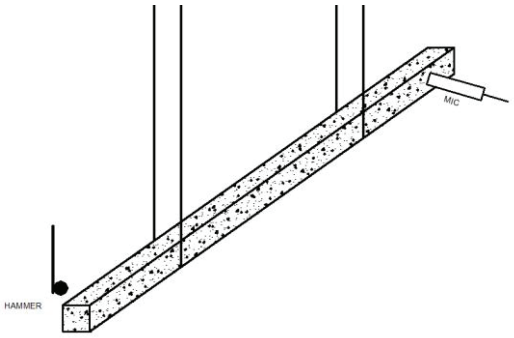


Fig. 1.2. *Transverse vibration study*

The assembly of measuring equipment includes:

- ✓ Microphone of the company "Audio Technica" - AT2031;
- ✓ Sound card M-Audio Audiophile 192
- ✓ Specialized software "Spectra PLUS".

Experimental results for the dynamic modulus of linear deformations E_d of PC samples are obtained by pulse excitation, Fig. 1.2 and processed with the specialized software "Spectra PLUS".

SpectraPLUS is a powerful 2-channel spectral analyser. Software interfaces provide real-time spectral analysis. Spectrum analyser is a tool used to convert a time domain signal (amplitude-time) into a frequency domain (amplitude-frequency). The Audio Spectrum Analyser, by definition, is limited in signal bandwidth processing. The limit of the analysed frequency is determined by the capabilities of the sound card used.

The software uses Fast Fourier Transform (FFT), which converts the signal from the time domain into the frequency domain.

Fig.1.3*. Test bench*

Table 1 *Experimental results*

Sample	f [Hz]	**r** [m^3]	**ω**[rad]	E_d[Pa]
1	953	2381	5985	3,61E+10
	2537,5		15936	3,37E+10
	4751		29836	3,07E+10
2	866	2339	5438	2,93E+10
	2311		14513	2,74E+10
	4294		26966	2,46E+10
3	977	2365	6136	3,77E+10
	2576		16177	3,44E+10
	4817		30251	3,13E+10
4	899	2355	5646	3,17E+10
	2385		14978	2,94E+10
	4456		27984	2,67E+10
5	815,6	2268	5122	2,52E+10
	2165		13596	2,33E+10
	4044		25396	2,12E+10
6	761,7	2181	4783	2,11E+10
	2027		12730	1,97E+10
	3792		23814	1,79E+10
7	767	2258	4817	2,22E+10
	2129		13370	2,25E+10
	3398		21339	1,49E+10
8	799	2216	5018	2,36E+10
	2137		13420	2,22E+10
	3958		24856	1,98E+10
9	839	2300	5269	2,70E+10
	2184		13716	2,41E+10
	4090		25685	2,20E+10
10	864	2348	5426	2,92E+10
	2319		14563	2,77E+10
	4344		27280	2,53E+10
11	815	2252	5118	2,49E+10
	2184		13716	2,36E+10
	4068		25547	2,13E+10
12	785	2280	4930	2,34E+10
	2114		13276	2,24E+10
	3980		24994	2,06E+10
13	858	2342	5388	2,88E+10
	2281		14325	2,67E+10
	4250		26690	2,42E+10
14	939	2387	5897	3,51E+10
	2504		15725	3,29E+10
	4688		29441	3,00E+10
15	788	2245	4949	2,33E+10
	2105		13219	2,18E+10
	3937		24724	1,99E+10
FG20	1001	7203	6286	1,20E+11
	2643		16598	1,10E+11
	4891		30715	9,84E+10
C 45	1262	7803	7925	2,07E+11
	3342		20988	1,91E+11
	6225		39093	1,73E+11

4. Experimental results

The obtained experimental results for the dynamic modules of the linear deformations E_d of the PC composites are presented in Table 1. In Figure 1.4 these are represented graphically.

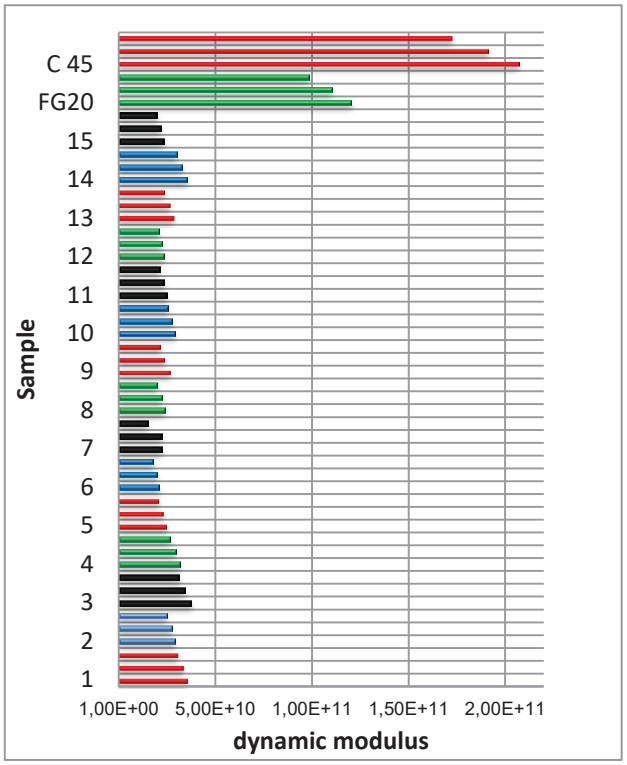


Fig.1.4.. *Experimental results for* E_d

We obtain the values, in which we are interested, of the sampling frequencies of the experimental PC composites from the auto-spectral characterization, fig.1.5 and 1.6.

Fig.1.5. *Autospectrum of sample 1*

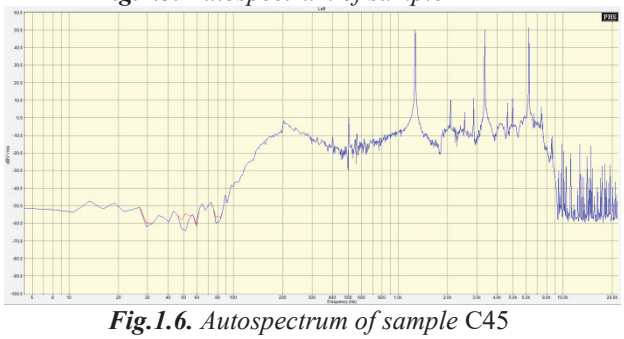


Fig.1.6. *Autospectrum of sample* C45

The theoretical verification of the first three forms (own frequencies) of the lateral oscillations of the 15 composites was performed using the dynamic simulation analysis module of SolidWorks 2014, fig.1.7. The results are shown in Table 2.

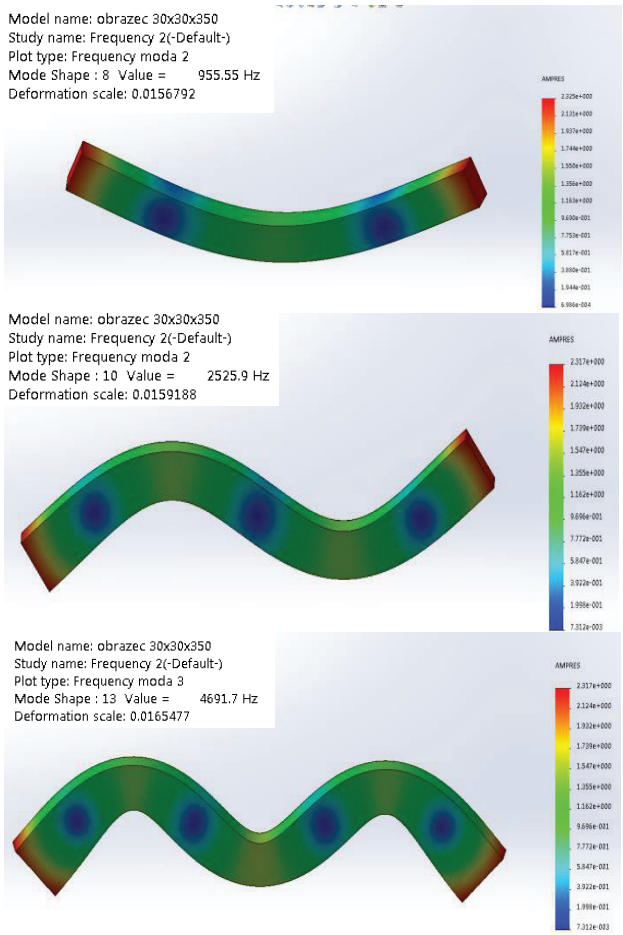


Fig. 1.7 *Graphics of the first three own forms of vibration for the transverse vibrations of sample 1.*

Table 2 *Modal frequencies*

Sample number	Measured values			Simulated values			Ratio to Percentage		
	$f_{\beta 1}$	$f_{\beta 2}$	$f_{\beta 3}$	f_{S1}	f_{S2}	f_{S3}	$f_{\beta 1}/f_{S1}$	$f_{\beta 2}/f_{S2}$	$f_{\beta 3}/f_{S3}$
	[Hz]	[Hz]	[Hz]	[Hz]	[Hz]	[Hz]	[%]	[%]	[%]
1	953	2538	4751	956	2526	4692	100%	100%	101%
2	977	2311	4294	868	2296	4268	113%	101%	101%
3	899	2576	4817	980	2591	4814	92%	99%	100%
4	870	2385	4456	900	2381	4428	97%	100%	101%
5	816	2165	4044	819	2166	4030	100%	100%	100%
6	762	2027	3792	764	2023	3766	100%	100%	101%
7	797	2129	3398	770	2039	3795	104%	104%	90%
8	799	2137	3958	801	2122	3946	100%	101%	100%
9	839	2184	4090	842	2227	4144	100%	98%	99%
10	864	2319	4344	866	2292	4265	100%	101%	102%
11	815	2184	4068	816	2159	4013	100%	101%	101%
12	785	2114	3980	786	2080	3869	100%	102%	103%
13	858	2281	4250	859	2275	4232	100%	100%	100%
14	939	2504	4688	938	2478	4600	100%	101%	102%
15	788	2105	3937	786	2081	3871	100%	101%	102%
FG20	1001	2643	4891	998	2636	4889	100%	100%	100%

5. Conclusion

The results of this work are limited to:

• Quantitative values for the dynamic elastic modulus of 15 different PC composites were obtained experimentally.

• By increasing the sequence of the sample's own resonance frequency, the module decreases.

• The method of experimental modal analysis is used with priority in the quantitative determination of the dynamic modulus of elasticity of the range of PC composites.

• The dynamic elasticity modulus of PC composites is determined on the basis of the dynamic response (response) of the free damping oscillations of the impulse excited dynamic system of the experimental samples.

• Opportunity to obtain reliable and realistic information about simulation shapes and own oscillation frequencies of each PC composites through SolidWork Simulation Module has been demonstrated.

• By a comparative analysis of the conducted tests and the results obtained for the elastic characteristics of the samples of gray cast iron BS20, the steel C45 and the PC composites it was found that the modulus of linear deformations is one level lower.

REFERENCES

1. S. Timoshenko, D. H. Young, and W. Weaver, Jr., Vibration Problems in Engineering (4th ed.), Wiley, New York, 1974.
2. Singiresu S. Rao, Mechanical Vibrations 5th Edition, University of Miam, 2011.
3. L. Klein, Transverse vibrations of non-uniform beams, Journal of Sound and Vibration,Vol. 37, 1974.
4. Ulbrich H. Vibration, Control of Nonlinear Mechanism and Structure. Technical University, Munich, 2005.
5. Попов Ил., „Изследване и оптимизиране на полимербетонни състави по отношение на демпферните им характеристики”, Докторска дисертация, (2013).

Адрес : ТУ-София, филиал Пловдив
E-mail: ilgepo@abv.bg
E-mail: sabi_sabev@abv.bg

© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 23, 2017

DETERMINATION OF PROPELLER CHARACTERISTICS FOR MULTIROTOR DRONE DESIGN

HRISTIAN PANAYOTOV, STANIMIR PENCHEV

Abstract: *A methodology for determination of multirotor drone propeller characteristics is described. A special formula with correction coefficients is derived to calculate propeller thrust and power at different flight modes. The calculated theoretical results are compared with experimental data and thorough analysis is carried out.*

Key words: *multirotor drone, propeller, characteristics*

1. Basic theory

Like helicopter and airplane propellers, multirotor drone propellers are well described by momentum and blade element theory. However in multirotor drone design it is often necessary to assess the main propeller characteristics such as propeller thrust, power and efficiency in hover and translational non-axial flight. Hence the aim of the present study is to derive some approximate dependencies that could be useful in early multirotor design stage for determination of propeller characteristics.

For propellers in hover mode the main assumptions for the flow field are shown in Fig. 1.

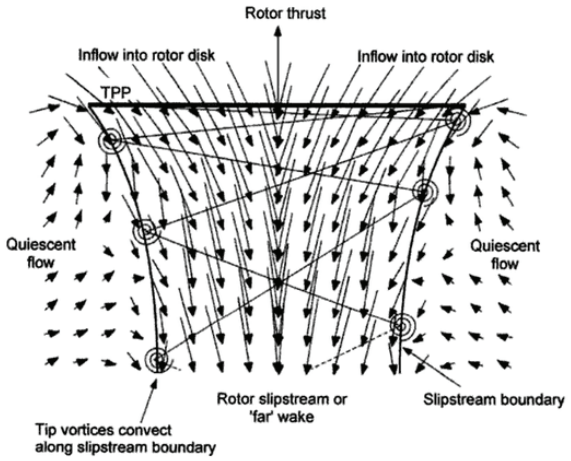


Fig. 1. *Propeller flow field*

Far upstream the air flow is undisturbed and passing through the propeller disc the flow induces speed – V_i. The airflow is further accelerated until it reaches the far wake at slipstream velocity or the jet velocity V_j [1]. The described control volume is shown in Fig. 2.

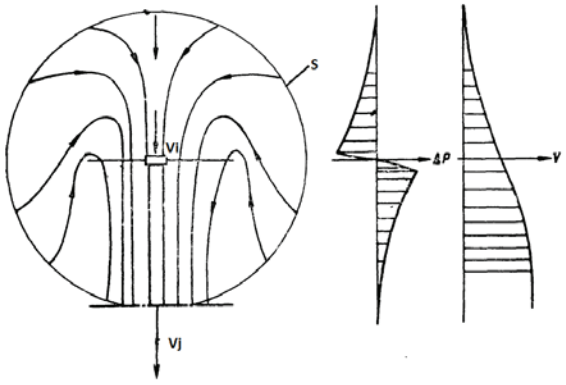


Fig. 2. *Propeller control volume*

Applying the governing conservation laws for the fluid mass and momentum to this control volume yields the following equations [2]:

$$\iint_S \rho\vec{V}.d\vec{S} = 0 \qquad (1)$$

$$\vec{F} = \iint_S p d\vec{S} + \iint_S (\rho\,\vec{V}.d\vec{S})\vec{V}\,, \qquad (2)$$

where $\vec{V}$ is the local velocity, ρ is the density of the fluid across the surface S, and $\vec{F}$ is the net force.

According to eq. (1) for the control volume in Fig. 2 the air mass flow at the station of the propeller disc is:

$$\dot{m} = \rho A V_i, \qquad (3)$$

where $A = \pi R^2$ is the propeller disc area provided the propeller has a radius of R.

 ISSN 1310 - 8271

Applying eq. (2) the thrust of the propeller is:

$$T = \dot{m}V_j = \rho A V_i V_j. \quad (4)$$

For hover mode (Fig. 2) the theory [1] yields that $V_j = 2V_i$, so finally:

$$T = 2\rho A V_i^2. \quad (5)$$

Additionally the power exerted to the flow from the propeller is [1]:

$$P = TV_i = T\sqrt{\frac{T}{2\rho A}} = \frac{T^{3/2}}{\sqrt{2\rho A}}. \quad (6)$$

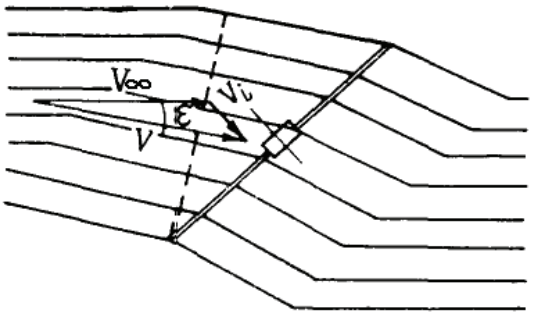


Fig. 3. *Propeller in translational flight*

In the case of translational non-axial flight the flow field assumption is given in Fig. 3. The disc is inclined at some angle of attack A and the resultant velocity across the propeller disc is $\vec{V} = \vec{V_\infty} + \vec{V_i}$. Hence the mass airflow in that case is $\dot{m} = \rho A V$ and propeller thrust in translational flight will be:

$$T = 2\rho A V V_i. \quad (7)$$

Alternatively from eq. (5) provided the thrust of the propeller is known the induced speed can be calculated:

$$V_i = \sqrt{\frac{T}{2\rho A}}. \quad (8)$$

2. Wind Tunnel Propeller Test Bed

The Department of Transport and Aviation Engineering in the Technical University of Sofia, Plovdiv Branch is equipped with state-of-the-art laboratory wind tunnel used for aerodynamic measurements both for fix-wing and rotary-wing aircraft. There is a specially developed propeller test bed, mounted in the open test area of the tunnel (Fig. 4) that allows orientation of the axis of the propeller at arbitrary inclination towards the vector of undisturbed velocity flow (the angle of attack of the propeller – A). The test bed is designed so that at a certain flow velocity, rotational speed and angle of attack of the propeller the measured properties are: propeller thrust, power and efficiency. Also at steady conditions with no wind tunnel flow the slipstream velocity of the jet (V_j) could be measured using calibrated pitot-static tube. According to eq. (4) and (5) the induced speed (V_i) could be calculated as well.

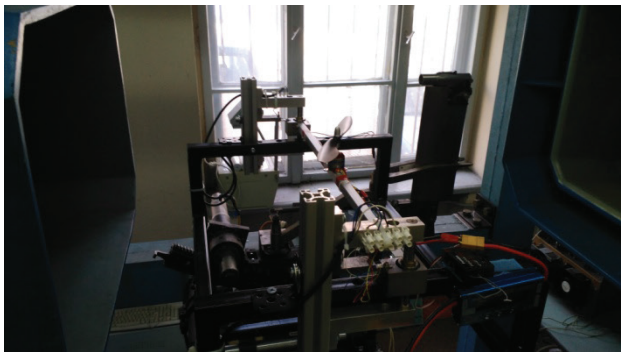


Fig. 4. *Propeller test bed*

Firstly the properties of a propeller are measured to validate the theoretical model from eq. (1) to (8). The chosen drone propeller is well known APC 10''x4,5'' (Diameter=10'', Pitch=4,5''). The propeller is shown in Fig. 5.

Fig. 5. *Propeller APC 10x4.5*

The results from the experiment are given in Table 1. Here the angle of attack of the propeller disc is zero ($A = 0$). The variables are propeller rotational speed in RPM and flow velocity. This represents the hover mode and translational non-axial flight. In hover mode ($V = 0$) thrust T, consumed electric power P_{el}, and jet slipstream velocity V_{jexp} are measured. Additionally using eq. (8) the induced speed (V_i) is calculated, based on the measured thrust. This is to compare the induced

speed based on momentum theory and actual measured speed.

Table 1. *Propeller experimental data*

A=0 \ n, rpm		4000	6000	7100
V=0	T, N	3,15	7,20	9,45
	P_{el}, W	34,885	100,41	162,41
	$\frac{T}{P_{el}}, \frac{daN}{W}$	9,03	7,17	5,82
	$V_i, m/s$	5,07	7,66	8,78
	$V_{jexp}, m/s$	9,17	14,17	17,22
	$V_{iexp}, m/s$	4,58	7,08	8,61
V=10 m/s	T, N	4,05	9,00	10,35
	P_{el}, W	34,10	99,86	163,35
	$\frac{T}{P_{el}}, \frac{daN}{W}$	11,88	9,01	6,34
V=20 m/s	T, N	5,85	12,60	13,95
	P_{el}, W	34,78	102,41	164,80
	$\frac{T}{P_{el}}, \frac{daN}{W}$	16,82	12,30	8,46

For the other non-hover mode cases at flow velocity 10 m/s and 20 m/s (Table 1) the induced and jet velocities are not measured due to obvious difficulties.

It is apparent from Table 1 for the hover mode that the errors between the measured and calculated induced speeds from the propeller are acceptably small, and the assumptions from the momentum theory are reliable and applicable to early design phases.

3. Methodology for determination of propeller characteristics.

In Table 1 induced velocities in hover mode are easily measured and calculated provided the propeller thrust is a priori known. However the calculation of the induced velocities in translational flight at some angle of attack is a difficult task. In the case of translational flight eq. (7) is used to calculate propeller thrust. If we have correct measurements of the propeller characteristics in hover mode then the induced speed can be calculated using eq. (8) or measured directly. If we assume that in translational flight the induced speeds for a certain rotational speeds remain the same then eq. (7) should yield the propeller thrust for a given airspeed. However two problems arise that influence the thrust and power of the propeller in translational flight. First at a given airspeed at zero angle of attack the flow field around the rotor is not axisymmetric. There is a region with reversed flow defined by the regime coefficient:

$$\mu = \frac{V_\infty cosA}{\omega R} \quad , \tag{9}$$

where ω is the rotational speed in rad/s. The diameter of the reversed flow region is $d = \mu R$ [2]. So it is as if at higher airspeed the actual radius of the disc is reduced. This phenomenon has to be taken in account.

The second problem is that if the propeller disc were at a given angle of attack towards the undisturbed flow then the axial component of the resultant flow velocity decreases the angle of attack of the propeller blade elements, which reduces the actual induced speed. So if formula (7) is used to calculate propeller thrust in translational flight the abovementioned phenomena should be considered. As we can see from Table 1 with the airspeed increased the resultant speed V is increased and so the thrust is increased, however the reverse flow area is also increased and if the propeller is at some angle of attack (negative for forward flight) then the induced velocity should be decreased.

We can take advantage of the blade element theory in order to take in account the abovementioned problems. The propeller thrust and power are usually defined in terms of thrust and power coefficients – respectively C_T and C_P:

$$T = C_T \rho n^2 D^4 \qquad (10)$$
$$P = C_P \rho n^3 D^5.$$

If we apply the first assumption that the effective diameter of the propeller is reduced due to reverse flow region with diameter μR and bear in mind that the momentum theory assumes even distribution of the induced velocity across the propeller disc then the following is true:

$$\frac{T\prime}{T} = \left(\frac{D-\mu R}{D}\right)^4 = \left(1-\frac{\mu}{2}\right)^4 \equiv k_\mu, \tag{11}$$

where T' is the thrust of the propeller due to reverse flow region. In other words the thrust in eq. (7) has to be corrected with coefficient k_μ equal to the ratio in eq. (11).

The second problem – the reduction of the induced velocity with the increase of axial component of the resultant speed is connected with the blade element angle of attack or with the thrust coefficient. If we assume that due to increased axial

velocity the initial blade angle of attack at hover mode (φ_H) is decreased by $\Delta\varphi$ then:

$$\frac{T'}{T} = \frac{C_T'}{C_T} = \frac{C_{T_\varphi}\varphi'}{C_{T_\varphi}\varphi_H} = \frac{\varphi_H - \Delta\varphi}{\varphi_H} = 1 - \frac{\Delta\varphi}{\varphi_H}, \quad (12)$$

where C_{T_φ} is the derivative $\frac{\partial C_T}{\partial \varphi}$, which for a certain airfoil is accepted to be a constant.

The decrease of the blade angle of attack can be found if the velocity triangle for a blade element is viewed:

$$\tan\Delta\varphi \approx \Delta\varphi = \frac{\Delta V_i}{\omega R} = \frac{V_\infty \sin A}{\omega R} = \mu \tan A. \quad (13)$$

Finally the ratio in eq. (12) will take the form:

$$\frac{T'}{T} = 1 - \frac{\Delta\varphi}{\varphi_H} = 1 - \frac{\mu \tan A}{\varphi_H} \equiv k_A. \quad (14)$$

In other words provided we use the induced speed in hover mode to calculate the thrust in translation flight the formula in eq. (7) has to be corrected with k_μ *and* k_A and the corrected thrust will be:

$$T' = 2k_\mu k_A \rho A V V_i \quad (15)$$

Regarding to the propeller power from Table 1 it is obvious that the required electric power for all flight regimes including hover is almost identical so with some precision the power of the propeller in translational flight could be considered approximate to the power at hover mode for a given rotational speed (Fig. 7).

4. Results and discussion

To asses the accuracy of the hereby methodology, propeller characteristics can be evaluated using formula (15) with the corrections (11) and (14). Figure 6 shows experimental data for the thrust from Table 1 (solid lines) compared to momentum theory results, corrected with k_μ (dashed lines). The angle of attack of the propeller is zero and the k_A coefficient is zero respectively. It can be seen (Fig. 6) that with the increase of the flow velocity the thrust generated by the propeller is increased due to the increased mass airflow through the propeller disc. If uncorrected the predictions for the thrust are too high. In this case this is a satisfactory result.

Figure 7 shows the power of the propeller – both experimental (electric) and theoretical – according to momentum theory, calculated from the induced speeds, corrected with k_μ. The measured electric power (solid lines) is far bigger than the theoretical, which gives the time rate of change of the mechanical work exerted to the flow. In fact the ratio of the mechanical power to electric power gives the efficiency of the propulsion (propeller, electric motor and ESC) that is shown with the dotted line in Fig.7. In that manner the described methodology is capable to predict the overall efficiency of the propulsion.

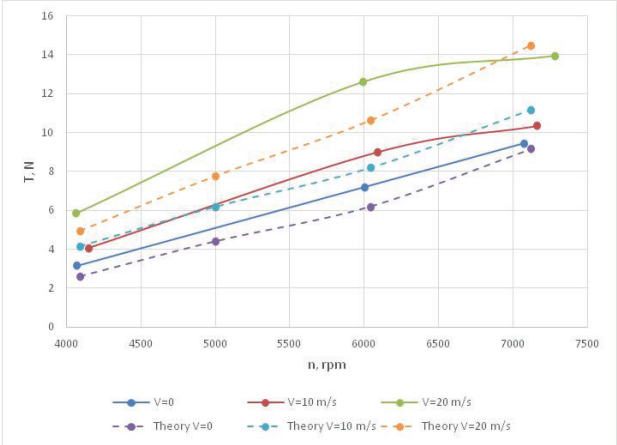


Fig.6. *Thrust vs RPM & Airspeed*

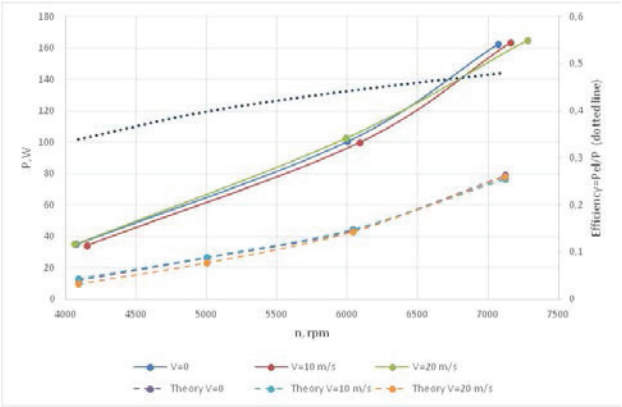


Fig.7. *Power vs RPM & Airspeed*

Once again Fig. 7 shows that the power consumed by the propeller at different flight modes is approximately the same.

So far the angle of attack of the propeller disk was considered to be zero. In order to check the assumption made in derivation of (14) additional experimental and theoretical research is done. Propeller characteristics are measured at angle of attack $A = -25\,deg$ (forward flight). The experimental data are compared to theoretical (Table 2). The corrected thrust, calculated using eq. (15) is shown in the last row, compared to the experimental in the first row. Here the hover induced speed are used (V_{ih}) to calculate the thrust. This is also plotted in Fig. 8. The results show acceptable convergence and accuracy.

Table 2. *Propeller experimental data at A=-25 deg*

A=-25 deg \ n, rpm		4000	5000	6000	7100
V=10	T_{exp}, N	2,70	4,05	4,95	8,55
	μ	0,18	0,14	0,12	0,10
	k_μ	0,69	0,74	0,78	0,81
	k_A	0,57	0,65	0,71	0,76
	$V_{ih}, m/s$	4,58	5,97	7,08	8,61
	T, N	6,22	8,59	10,71	14,03
	T', N	2,43	4,17	5,96	8,75

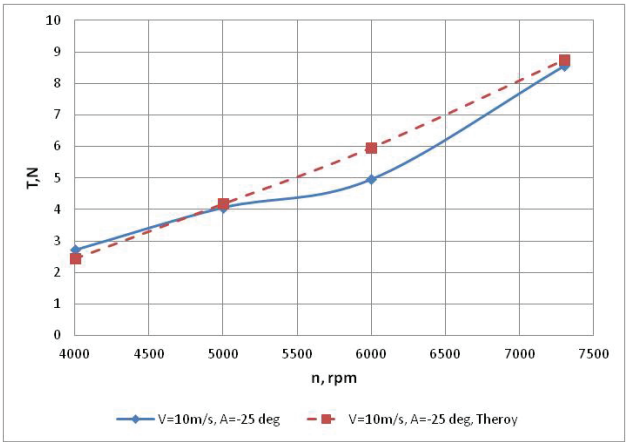


Fig. 8. Propeller thrust at V=10 m/s & A=-25 deg

5. Conclusion

The present paper deals with a methodology for determination of a multirotor drone propeller characteristics at different flight modes. Specially derived coefficients are proposed to correct the results from momentum theory. Theoretical results are compared with experiment.

LITERATURE

1. Leishman J. Gordon (2006), *Principles of Helicopter Aerodynamics*, 826 p., Cambridge Aerospace Series, USA.
2. W.Z. Stepniezki, C.N. Keys (1979). *Rotary-Wing Aerodynamics, Vol.1*, 380 p., NASA, Philadelphia, USA.
3. Homiao Huang, Gabriel Hofmann et al. (2008), *Aerodynamics and Control of Autonomous Quadrotor Helicopters in Aggressive Maneuvering,* Stanford, California

Contacts:
Hristian Panayotov
TU - Sofia, Plovdiv Branch
Address: 25 Tsanko Dyustabanov Str, Plovdiv, Bulgaria
Tel.:+359 32 659 518
E-mail: hristian@tu-plovdiv.bg

Stanimir Penchev
TU - Sofia, Plovdiv Branch
Address: 25 Tsanko Dyustabanov Str, Plovdiv, Bulgaria
Tel.:+359 32 659 632
E-mail: spenchev@tu-plovdiv.bg

© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 23, 2017

PROBLEMS IN RECRUITING HUMAN RESOURCES FROM BULGARIAN INDUSTRIAL ENTERPRISES

TONI MIHOVA, VALENTINA NIKOLOVA - ALEXIEVA

Abstract: *The report is an attempt to analyze the problems associated with finding suitable human resources by industrial enterprises in Bulgaria. For this purpose, the authors were based on two national surveys conducted by ManpowerGroup and Bulgarian Industrial Association, as well as its own investigation on the causes of this problem. Have been formulated conclusions and guidelines for improving the activity of selection of human resources in Bulgarian industrial enterprises.*

Keywords: *human resources, recruitment, motivation, remuneration, training and development*

1. Introduction

Modern industrial enterprises operate in an extremely dynamic development and application of new equipment, technology, communication and information systems. Nevertheless, the human factor occupies a leading position as the main productive and creative force in the enterprise.

The management of human resources is directly related to the competitiveness and efficiency of each company. Activities on recruitment and selection of staff play a leading role in attracting candidates who meet the requirements for education, experience, training, contributing to the achievement of organizational goals. This determines the **topicality of the theme** of the report.

The **subject** of research are the difficulties and problems faced by industrial enterprises in selecting the most appropriate human resources. The authors base their findings on two national studies as well as own research conducted in industrial enterprises in the Plovdiv region.

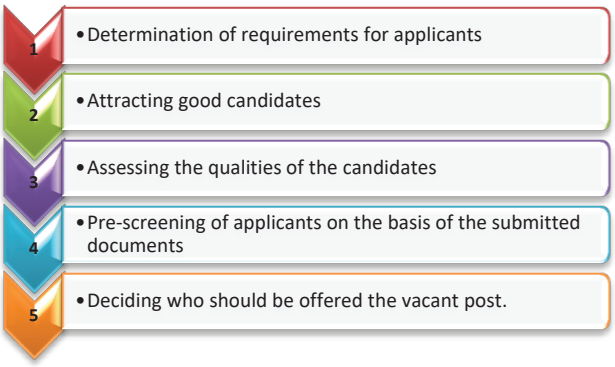
2. Definition of the recruitment and selection of human resources

In connection with the objective to understand what the problems and difficulties in recruitment and selection of human resources will be clarify in a systematic way the essence of this process. In the specialized literature has several publications about the nature of recruitment and selection of human resources. Extremely a thorough studies on this issue have been issued by Dimitar Shopov, Margarita Atanasova, Tatiana Hristova Dimitar Kamenov and Jordan Bliznakov.

According to the authors, definition that most accurately reflects the nature of the selection of human resources is: "The selection of human resources in the enterprise be an activity to attract suitable candidates for a post evaluation of their merits in terms of job requirements and then deciding whom to make a job offer. [1]

The main objective of selection is to ensure "the best man" for the job. It depends on the objective and accurate forecast of future job performance of candidates.

The selection process of candidates includes a number of activities that can be defined as its major stages:



 ISSN 1310 - 8271

Basic methods and tools for contacting potential candidates for vacancies in the enterprise are:

- Advertise the vacancy;
- Oral advertising by relatives and friends;
- Attracting good candidates with the help of employees in the enterprise;
- Contacts with employment exchanges;
- Contacts through social networks;
- Contacts with students in their respective schools.

Business practices shows that this is an extremely complex and responsible process, which depends on the overall performance of the enterprise. Therefore it is necessary human resources specialists have appropriate training and competence to be able to effectively carry out activities in the selection and recruitment.

3. Contemporary tools in recruitment and selection of human resources

Besides the necessary training and competence recruiters also need to monitor emerging technological novelties on the process of recruitment and selection of staff. One such new challenge is recruitment through the use of so-called „social and mobile channels“. According to British electronic edition HRmagazine, "the market already has many social tools that allow you to quickly find people - this is not a problem" [2].

In this sense, the question is not only which channels to select are the most suitable for the needs of industrial enterprise, but also how to be used in the most effective manner. For example, according to marketing firm 4MAT, currently between 20% and 30% of traffic to websites for careers and recruitment is generated by users who visit mobile versions of the relevant pages [2].

We investigated various technologies and tools for selection, which we believe have the potential to change the face of the HR industry and provide a rating as follows:

- **Systems for tracking applicants (applicant tracking systems, ATS)**

The first such instruments appear early 90s of the last century and are among the earliest opportunities for online selection. At the very beginning their functionality is limited to tracking incoming company CV-s, but lately it offers opportunities to expand in various areas, including for example compliance checks skills, Analysis of CV-th and posting job advertisements through various communication channels, including social media. In practice, modern ATS is a platform that allows HR professionals to connect various tools and services used in the selection process, and thus streamline operations.

- **Instruments for mobile selection**

Rising sales of powerful smartphones and tablets will probably lead to a time when their use will overtake those of desktops. This in turn brings to the fore the question of mobile selection, and expectations are to become a major recruitment channel in the future.

- **Gamification**

According to one of the leading analytical companies in the world - Gartner, the so-called *gamification* is the trend of using different gaming mechanisms in the *non-gaming* situations, such as the selection of personnel. The idea is thus to increase the motivation of the people or to provoke change in their behavior. Gartner forecast that in the next five years the use of such methods will become "a pretty significant trend." Factor for this is the growing market of developers, offering opportunities for the use of game elements in processes like recruitment and management of staff. The use of such methods for example can allow more rapid identification of candidates who possess the necessary skills for a position. Game elements can be used also for better communication and build relationships between talent within an organization.

- **Collaboration tools**

Despite its great potential for sharing and exchanging information currently professional social networks are not as used by recruiters. However, if used effectively, this type of networks can contribute to a better integration of the processes of selection and management of human resources with other departments and activities within an organization, which ultimately contribute to increasing the efficiency of the selection process .

- **Technology to determine the location**

The idea of using this type of tool is that it allows both employers and HR professionals in an easy and inexpensive way to target jobseekers who are often on the move. The use of technology to determine the location (most often through the IP address of the mobile phone or the computer of a user) enables companies to provide potential employees information that is relevant to the area in which they are located. "When you ask people looking for work, what is important to them three things that always stated are salary, position and location",explains John Salt from the TotalJobs.com. According to him the possibility of receiving targeted ads depending on their location is already something jobapplicants seeking and HR professionals should be aware.

- **Video interviews**

In 2016, according to the head hunting company OfficeTeam UK, 45% of recruiters have increased their use of video to conduct job interviews compared to 2013. While it should not be taken entirely as an alternative to actual face to

face meetings, interest in such tools grows and factors are speed and convenience they offer. Furthermore, the video channels are a natural environment for the majority of the representatives of the younger generations, which means that HR professionals will have to accept their increasing importance in the selection process.

- **Internet sources**

For a number of recruiters using social networks like LinkedIn, Twitter, Facebook etc. is now living part of their work. At a later stage in this category includes the use of a wider range of sources and head-hunting techniques for more difficult to detect talents. The challenge for HR professionals in this case is rather in how to optimize the use of such sources, so to maximize efficiency. Possible solutions to this effect are as applications that enable semantic search, where results are based on the meaning of search keywords rather than specific matches word for word. For now, however the use of such tools is not yet extensive, indicated by HRmagazine [2].

4. Results from research

To realize the goal of the report and to reveal the main difficulties and problems in the recruitment and selection of human resources from industrial enterprises, we can use three surveys.

The first study is a national survey of ManpowerGroup [3]. According to him, nearly two thirds (62%) of employers in Bulgaria have difficulties to fill their vacancies in 2016. In comparison, the percentage of Bulgarian employers shared these difficulties was 50% in 2015. About one-fifth (24%) of respondents consider that they are more difficult to find suitable candidates for the job in 2016. The situation has facilitated only in 6% of Bulgarian employers.

The main reasons for difficulties in filling the vacancies, cited by employers, are:

- first, with 32% - lack of so-called required. hard skills (eg qualifications relating to certain tasks, IT skills, or language and mathematical skills);
- secondly with 30% - lack of candidates;
- third with 13% - lack of so-called required soft skills (professionalism, enthusiasm, interpersonal skills, flexibility / adaptability);
- 12% of employers indicate lack of experience and talent in candidates;
- 11% of employers indicate higher expectation than the remuneration offered.

The results of this study also indicate that:

Employers in Bulgaria suffer most from a lack of skilled workers, engineers and drivers.

To deal with difficulties in finding people Bulgarian employers use several strategies:

- the majority - 50% - provide additional training and development of current staff;
- second, employers indicated hiring people outside the traditional range of talent - older or younger than usual employees;
- one-third of employers (33%) use outsourcing activities;
- 30% of surveyed employers offer a larger salary to employees, and 29% - offer additional motivational packages.
- 22 percent say changing existing business models - such as flexible or teleworking;
- only 9% of respondents said they did not follow any strategies, and 2% said they did not know.

In 2016 most sought after by employers in Bulgaria are skilled workers - they change their place in the rankings with engineers who were most scarce last year. In 2016 IT talents are in 10th place to fourth position in 2015. Sales representatives and project managers were replaced by laborers and operators proceedings. In 2016 remained large imbalance between demand and supply of skills.

The second study was conducted by the Bulgarian Industrial Association (BIA) in 2015, whose results were announced at a round table with the participation of the World Bank and the "Open Society" Institute [4]. The conclusions from it are:

- Businesses looking for staff with interdisciplinary knowledge and skills hybrids - combinations of technical capabilities and marketing flair and entrepreneurial attitudes;
- Labor market is valued also skills in development, creation and implementation of new technologies, willingness to learn and improve;
- Required skills for teamwork, leadership, decision making, problem solving, communication, emotional intelligence, commitment to change;
- The advantages are quickness, flexibility, adaptability, organization, personal effectiveness, courage to assume responsibility, reliability, attention to detail and focus on the customer. Important to many areas are also ability to work under pressure and stress, self-control and time management.

According to this study do not reach professionals in the software industry, nanotechnology, mechatronics, mobile communi-cations, applied sciences, energy, chemistry, construction, mechanical engineering.

The third study was conducted by the authors of the report during the period February 2016 - December 2016 into industrial enterprises in the Plovdiv region. In particular industrial enterprises are from the "Trakia economic zone" and the collection of empirical material is accomplished with the help of students from

Technical University - Sofia, Plovdiv branch and the University of Food Technologies and pupils from two vocational schools in Plovdiv. It covers 65 managers (from middle and operational level), 587 employees (121 - 466 engineers and skilled workers) and 332 students (175 students and 157 pupils).

The methods used are interview and questionnaire. We'll point out some highlights from the survey.

The main problem that indicate 92% of the managers is the difficulty in finding quality and skilled human resources. Only 3% do not think there is a problem to find what specialists, and 5% did not answer the question of selection of human resources (fig. 1).

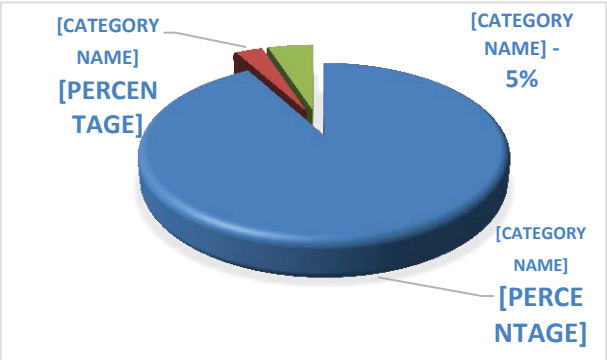


***Fig. 1**. "What kind of problems are you experiencing in the staff selection?"*

On question of job satisfaction, the results are as follows:

- 78% of engineering professionals are largely motivated for the most effective job performance. They are satisfied with the remuneration, training opportunities and career development and working conditions;
- 16% are not motivated due to stress at work, poorly constructed teams and ineffective communication;
- 6% of respondents did not answer the question.

Highly skilled workers meet the same issues as follows:

- 67% are highly motivated to work effectively. Satisfied are from the salary, benefits, training opportunities and working conditions.
- 24% have expectations for higher remuneration;
- 9% did not answer the questions in the survey.

Study analysis of the results conducted with students enables to point out the following:

- 17% would apply for a job at the declared vacancy into industrial enterprises from "Trakia economic zone";
- 53% do not believe they will apply for jobs in these companies (do not feel sufficiently prepared; have other plans for their development; do not like the location of businesses; they do not think that they will receive the expected salary);
- 19% do not have information about these companies and cannot decide;
- 11% did not answer the survey questions.

On the same question (fig.2), students from vocational schools have responded as follows:

- 13% would apply for a job at the declared vacancy;
- 61% share the opinion that it would apply (they have other plans for their development; they are not convinced that will handle the job; have expectations for higher remuneration from the alleged);
- 21% are unaware of these enterprises;

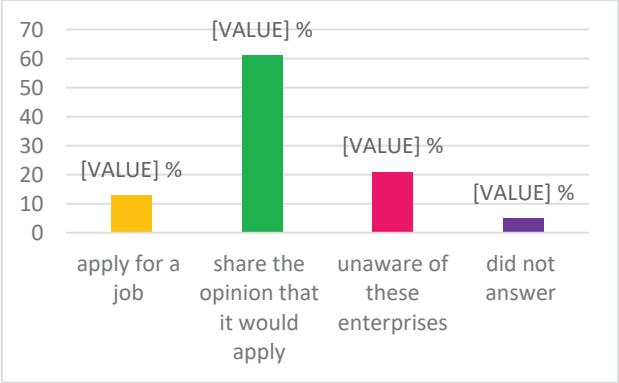


***Fig. 2**. "Are you interested to work in the industrial enterprises from "Trakia industrials zone"?"*

- 5% did not answer the poll question.

What are the specific findings of the present studies?

- major problem for industrial companies is finding quality and skilled human resources;
- for the most part, employees are motivated the most effective job performance and are satisfied with the remuneration, opportunities for training and career development and working conditions in industrial enterprises;
- the majority of potential job applicants (students from vocational schools and students of universities) are not sufficiently confident in their professional employability in the industrial enterprises, and they lack the necessary information about.

The findings are reason enough to formulate some guidelines to solve the problem - the difficulty of finding the most appropriate human resources.

Above all, it is necessary to build mechanisms for extremely strong connection between business on the one hand and vocational schools and universities - on the other. The training of engineers and technical specialists creates a

serious potential for recruitment for industrial enterprises in the Plovdiv region.

The problem with insecurity of the students in their preparation is solvable also with the direct involvement of business representatives who can actively participate in updating curricula, visits enterprises to conduct practices, offering training programs, organizing various joint events to increase the practical orientation of the training.

Moreover, this relationship between the two parties will produce information for industrial enterprises will construct attitude and willingness to work in them.

The other direction is the activities of HR specialists who can use more advanced and modern techniques to attract staff as discussed in the previous section of this report.

The use of social networks is also an important factor in relation to students and students who had been thoroughly studied in contemporary publications [5].

Last but not least is the use of various motivational techniques by the management of industrial enterprises detained successful staff – more attractive payment, flexible system of bonuses for stimulation, opportunities for training and career development. The analysis of the relationship between motivation of human resources and commitment to the business organization is thoroughly investigated problems in modern scientific works. [6]

This is a summary of some of the basic guidelines for dealing with problems with finding the most appropriate human resources that the effective functioning of industrial enterprises will depend on.

5. Conclusion

The report is an attempt to explain the essence of one of the major problems of industrial enterprises, the difficulty in finding qualified and quality human resources.

The dynamic development of these enterprises will make this problem even more acute and limiting the activity in the event it does not prepare adequate strategies, policies and tactics of human resources management.

The analysis of publications shows that the management of these companies is making effective efforts in the right direction. Proof of this are the meetings at universities and vocational schools, as well as the formation of a joint partnership with the economic cluster and educational board between Plovdiv Municipality, Ministry of Education, "Trakia Economic Zone" and "Industry Watch".

This will create conditions for young people to see alternative for its development as a highly qualified human resources in Bulgarian industrial enterprises.

Bibliography

1. Shopov, D. and co., How to manage human resources in the enterprise, Sofia, 2013
2. Seven tool selection that you need, https://www.capital.bg/
3. Firms increasingly difficult to find qualified staff, https: //technews.bg/article-94426.html
4. Businesses seeking employees with hybrid skills, https://technews.bg/article-80429.html
5. Tepavicharova, M., Boykova, L., Social networking in the context of increasing student motivation in Modern Education, Scientific Works of the Union of Scientists in Bulgaria - Plovdiv, Serie A: Social Sciences, Arts and Culture, Volume II, 2015
6. Tepavicharova, M., Analysis of correlation between the motivation of human resources and their commitment to business organizations: Scientific papers of the Union Scientists, Volume II, Serie A, Plovdiv, 2016

Contacts
Name of the authors:
Toni Bogdanova Mihova
Valentina Nikolova-Alexieva
Organization:
Technical University – Sofia, branch Plovdiv
University of Food Technologies - Plovdiv
Address:
4000 Plovdiv, ul. "Tcanko Dyustabanov"25
4000 Plovdiv, bul. "Maritsa" № 29
Phones:0893 69 06 55 и 0885 696 696
E-mail:
expert2009@abv.bg
valentina_nikolova@abv.bg

© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 23, 2017

CHALLENGES AND OPPORTUNITIES FOR FLEXIBLE CREDITING OF SMALL AND MEDIUM-SIZED ENTERPRISES IN BULGARIA

MINA ANGELOVA, DANIELA PASTARMADZHIEVA

Abstract: *The small and medium-sized enterprises (SMEs) play an important role in the economic processes and in this sense the optimization of their functioning is extremely important for raising the competiveness of the Bulgarian economy as a whole. The purpose of this study is to analyze, generalize and systematize challenges and opportunities for flexible financing of SMEs by European funds. A particular focuses of the study are challenges related to political instability and opportunities coming from European Union (EU) funding. The major goal of this study is subordinate to a major project, namely to create a useful model of an information platform that enables SMEs to get information about the best options for European funding in a quick and accurate manner.*

Key words: *business financing, credit instruments, small and medium-sized enterprises, smart information platform, human resources, European financial programs, political instability*

1. Introduction

The maturing of markets and market relations lead to intense competition and increasing consumer requirements, turning the quality into a factor for the success and survival of organizations. The significance of the studied topic is central to the discussions, research and organizational activities in the field of manufacturing and services, because the opportunities for flexible financing of organizations are preconditions for growth of quality and competitiveness.

The topicality of the study is grounded in the current situation in the global market. Dynamically changing economic and political environment requires more flexibility in the companies and willingness to change in accordance with the new conditions. Organizations face the challenge to "fight for survival" in a rapidly changing and uncertain environment. This process is enhanced in terms of the continuing economic crisis.

Optimizing the performance and opportunities for flexible financing of SMEs is extremely important to increase the competitiveness of the organizations and hence to the Bulgarian economy as a whole.

Despite the circumstances mentioned above, the development of small and middle sized enterprises (SMEs) has been an important area of EU's policies [1]. The recent measures under "Juncker plan" (such as COSME) and the reaffirmation of SMEs as priority in the Rome declaration from 26 March 2017 [2] give ground for further examination of capabilities for optimizing their role and functioning and the effect of political instability on their growth.

In these terms, the objects of the analysis are SMEs and the particular focus is on main theoretical concepts regarding SMEs and the conceptualization of correlation political instability-economic growth.

The purpose of this study is to analyze, generalize and systematize challenges and opportunities for flexible financing of SMEs by European funds. But the existence of opportunities is not enough for the growth of SMEs. The latter need accurate, accessible and easy to use information in order to put the possibilities into effect. Thus, the major goal of this study is subordinate to a project, namely to create a useful model of an information platform that enables SMEs to get information about the best options for European funding in a quick and accurate manner.

In this sense, the research tasks of this analysis are:

1) To systematize the theoretical concepts regarding SMEs;

2) To conceptualize the effect of political instability on the growth of SMEs;

 ISSN 1310 - 8271

3) To study empirical data related to access to finance of SMEs;

4) To list some general options for financing from EU.

2. Role of SMEs in the national economy

According to Commission Recommendation 2003/361/EC, as published in the Official Journal of the European Union L 124, p. 36 of 20 May 2003 "The category of micro, small and medium-sized enterprises (SMEs) is made up of enterprises which employ fewer than 250 persons and which have an annual turnover not exceeding EUR 50 million, and/or an annual balance sheet total not exceeding EUR 43 million." [3]

This definition was transposed in Bulgarian legislation and in particular in article 3 of Bulgarian Law on Small and Medium-Sized Enterprises.

At European level SMEs are seen as drivers of growth, employment and innovation. According to European Investment Bank they represent over 90% of businesses in the EU and also two thirds of the active working population is employed in them. [4]

SMEs play an essential role in economic processes and therefore optimization of their operation is extremely important to enhance their competitiveness and sustainable development. They are not only the backbone, but they are nearly the entire economy of Bulgaria, as according to NSI data for 2015 they represent 90.3% of all business organizations [5]. SMEs are a major source of added value and the largest employer in the country, but at the same time are experiencing serious difficulties in many areas. Undoubtedly, management must focus on the opportunities for flexible business financing, and access to various financial instruments that could help to overcome the main difficulties and discover horizon for innovation and investment.

According to Bulgarian National Strategy for Small and Medium-sized Enterprises 2014-2020 a main weakness regarding the state of SMEs in Bulgaria is underdeveloped information environment in terms of opportunities to provide affordable services and project financing. This also leads to another weakness - support programs for SMEs funded by the EU, remain largely underutilized [6]. In this sense, the increase of awareness among SMEs about EU funding should be a priority and specific measures should be taken.

3. Political instability as a challenge faced by SMEs

There are number of studies dedicated to the effects of political instability on economic growth and particularly on the growth of SMEs. Considering the unstable political environment in many European countries and in European Union at large it is an up to date issue for Bulgaria as well.

According to a recent analysis in the opinion to the companies worldwide there are five main obstacle of their growth: access to finance, electricity, political instability, competition and tax rate [7]. The relationship between political instability and economic growth has been examined in numerous scientific studies over a long period of time. This correlation is bilateral. If an economy works poorly this may cause "government collapse and political unrest". And when a political environment is unstable usually this leads to decrease of "investments and the speed of economic development" [8]. In fact national political instability may cause some kind of positive effect. It can be a factor for internationalization of SMEs, because when national political environment is unreliable business is motivated to go abroad. [9]

Some major factors of political instability are "political motivated violence", "mass civil protest", "instability *within* the political regime", and "instability *of* the political regime" [10]. Nevertheless, it may be caused by various factors but Alesina et al measure political instability as "propensity of government changes" [8]. The latter is a major factor of political instability in Bulgaria in recent years. In the period January 2013 – March 2017 there have been six governments, three of which were caretaker governments. The frequent change of governments has affected the economic environment and caused difficulties in financing of the SME's.

4. EU instruments supporting SMEs

According to analysis commissioned by Bulgarian Small and Medium Enterprises Promotion Agency (BSMEPA) the access to finance is a major problem for SMEs in Bulgaria.

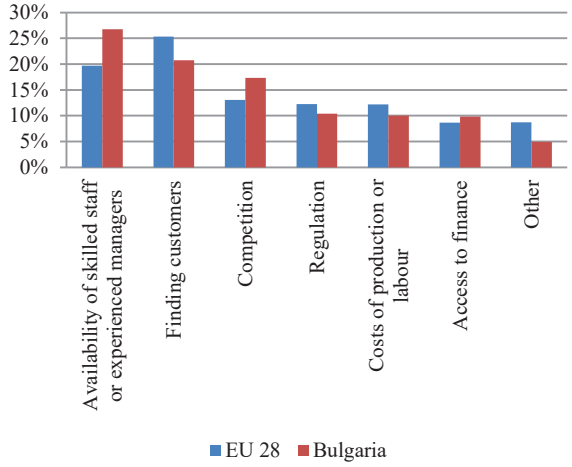


Fig. 1. *Most important problems for the firms, 2016*

Most commonly utilized sources of funding are banking instruments, government funding and international programs to support SMEs and the use of resources of the owner's family and his/her relatives. [11] Access to finance is a major issue at European level as well. Empirical data from Survey on the Access to Finance of Enterprises shows that in 2016 "access to finance" is challenge for 8,7% of SMEs at European level and for 9,8% companies in Bulgaria (Fig. 1.). [12]

The results on Fig. 2 indicate that there is a decrease in the significance of this problem in the period 2013-2015, but in 2016 there is a slight increase. [12]

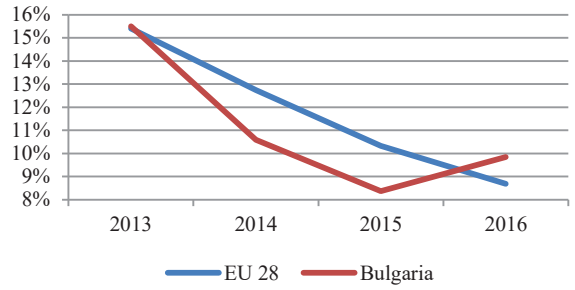


Fig. 2. *Dynamics in "access to finance" as a problem to the firms*

The European Union provides many opportunities for grants, loans and guarantees available for small and medium enterprises. The financing options often are not direct funding. Usually various national and sub-national institutions are intermediaries in this process. These can be authorities or financial institution such as banks. [13]

Two major programs under "Juncker plan" can be mentioned. One of them is Program for the Competitiveness of Enterprises and small and medium-sized enterprises (COSME) (2014 - 2020). It aims at "strengthening the competitiveness and sustainability" of the companies in EU with emphasis on SMEs. The program has potential to support the access to finance for the firms but also to widen their access to foreign markets (in EU and outside). Alongside COSME is expected to encourage the formation of new SMEs and they can assure economic growth in the European Union. [14]

COSME is a mechanism for guaranteeing loans financed by the European Union and managed by the EIF. It can contribute to reducing structural weaknesses in the financial market for SMEs and to become more diverse and accessible. Financial intermediaries in Bulgaria for loan/guarantee under COSME programme are CIBANK and Raiffeisenbank.

Another option is InnovFin SME Guarantee Facility. This program is specially designed for research-based and innovative SMEs and Small Mid-caps. It also is carried out through intermediaries who can be banks, leasing companies, guarantee institutions, etc. The type of support that SMEs can receive is loan/guarantee. Bulgarian partners of the programme are DSK Bank, CIBANK, ProCredit Bank, Piraeus Bank, Eurobank Bulgaria, UniCredit, Raiffeisenbank, United Bulgarian Bank, Deutsche Leasing and European Investment Bank [15]

The options shortly reviewed above can only be effective if the SMEs in Bulgaria are informed about them in an accurate and accessible manner.

5. Expected results and effects

The study of opportunities for flexible crediting of small and medium-sized enterprises in Bulgaria can generate new knowledge of the studied economic and socio-political phenomena and the relationship between socio-political environment and funding opportunities for SMEs. Challenges that might be faced are associated with the ability to propose alternative, improved versions of existing information tools concerning opportunities for European funding organizations. The research interest was provoked by the possibility of the application of new methods and technologies to detect dependencies between unexplored phenomena of economic, social and political spheres.

The expected results of the continuing research work this area are as follows: creating lasting and sustainable relationships with businesses and construction of model of an innovative information platform for flexible financing of SMEs. The results can be used in practice: management and control of business and public organizations; financial investments; development of marketing campaigns and social programs; business decisions with incomplete and inaccurate information and more.

6. Conclusion

In order to continue to carry out its role as the backbone of the economy SMEs need a special support. The changing international and national environment constitutes an obstacle to their access to finance. Although European Union provides various measures for promoting creation of SMEs and for support of their growth, the existence of

such is not enough. First of all companies need to be inform about them.

A study on the topic will increase knowledge of mechanisms for financing SMEs. In this case the general purpose should be to develop an innovative information platform. It will work in favor of the business to quickly and easily orient them to the most appropriate financing option.

The research has the potential to enhance and develop the scientific knowledge in the following three areas:

1) Displaying the challenges facing SMEs in the Bulgarian context, providing opportunities to enhance their competitiveness and sustainable development and positioning in the global market.

2) Systematization of conceptual framework concerning the relationship and mutual influence of political instability and certain economic indicators concerning SMEs. In addition to systematization of scientific knowledge on the subject an empirical evidence based on Bulgarian political process and economic development in the period from 2013 onwards should be provided.

3) Enrichment of knowledge of Information and Communication Technology can be applied in favor of SMEs.

Future research effort will be focused on building innovative information platform for flexible financing SMEs to orient organizations in funding opportunities, grants, guarantees, etc. Furthermore, a "bank" with information on potential future employees in organizations (including CVs, experience, key competencies) can be integrated in such platform.

The tasks that should be fulfilled in order to realize such objective are:

1) to further analyze, summarize and systematize opportunities for flexible SME financing from European funds;

2) to conduct a survey among representatives of small and medium-sized organizations and to identify problem areas in connection with flexible financing;

3) to study the experience of EU Member States in building such information platforms.

REFERENCES

1. https://ec.europa.eu/growth/smes_en
2. http://www.consilium.europa.eu/en/press/press-releases/2017/03/25-rome-declaration/
3. European Commission (2015) *User Guide for the SME Definition*, Luxembourg: Publications Office of the European Union;
4. http://www.eib.europa.eu/projects/priorities/sme/index.htm
5. http://www.nsi.bg/
6. National Strategy for Small and Medium-sized Enterprises 2014-2020 - Small Business Act, adopted by Decision № 37 of the Council of Ministers of 23 January 2014;
7. Wang, Y. (2016) What are the biggest obstacles to growth of SMEs in developing countries? e An empirical evidence from an enterprise survey, Borsa Istanbul Review 16-3, pp. 167-176
8. Alesina, Alberto, Sule Ozler, Nouriel Roubini, and Phillip Swagel. (1996) Political instability and economic growth. Journal of Economic Growth, 1(2), pp. 189-211
9. Shirokova G. and T. Tsukanova (2014) *Internationalization of SMEs from transition economies: institutional perspectives*, In: Etemad, H, T. K. Madsen, E. S. Rasmusen, P. Servais (eds.) Current Issues in International Entrepreneurship, Cheltenham: Edward Elgar Pub;
10. Jong-A-Pin, R. (2008). Essays on political instability: Measurement, causes and consequences Enschede: Print Partners Ipskamp B.V., Enschede, The Netherlands, pp. 4-5
11. Vladimirov, Zh, K. Ganev, R. Simeonova-Ganeva (eds) (2013) *Study of entrepreneurship and development prospects of innovation in SMEs (2012-2013)*, Sofia: Bulgarian Small and Medium Enterprises Promotion Agency (BSMEPA);
12. http://ec.europa.eu/growth/access-to-finance/data-surveys
13. https://ec.europa.eu/growth/access-to-finance_en
14. http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32013R1287
15. http://www.eif.org/what_we_do/guarantees/single_eu_debt_instrument/innovfin-guarantee-facility/

Mina Angelova, Daniela Pastarmadzhieva
University of Plovdiv Paisii Hilendarski
24 Tzar Asen
4000, Plovdiv
00359887461272, 00359884727096
mina.marinova@abv.bg,
daniela.pastarmadjieva@gmail.com

© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 23, 2017

RIEMANNIAN ALMOST PRODUCT MANIFOLDS GENERATED BY A CIRCULANT STRUCTURE

DOBRINKA GRIBACHEVA, DIMITAR RAZPOPOV

Abstract: *A 4-dimensional Riemannian manifold equipped with a circulant structure, which is an isometry with respect to the metric and its fourth power is the identity, is considered. The almost product manifold associated with the considered manifold is studied. The relation between the covariant derivatives of the almost product structure and the circulant structure is obtained. The conditions for the covariant derivative of the circulant structure, which imply that an almost product manifold belongs to each of the basic classes of the Staikova-Gribachev classification, are given.*

Key words: *Riemannian metric, circulant matrix, almost product structure.*

Mathematics Subject Classification (2010). *Primary 53B20, 53C15, Secondary 15B05.*

1. Introduction

The circulant matrices and the circulant structures have application to Vibration analysis, Graph theory, Linear codes, Geometry (for example [1], [2] and [3]). Riemannian manifolds equipped with a circulant structure, whose fourth power is the identity were considered in [4] and [5]. In particular case, such manifolds could be Riemannian almost product manifolds. The systematic development of the theory of Riemannian manifolds M with a metric g and an almost product structure P was started by K. Yano in [6]. In [7] A. M. Naveira classified the almost product manifolds (M,P,g) with respect to the covariant derivative of P. The Riemannian almost product manifolds (M,P,g) with zero trace of the structure P were classified with respect to the covariant derivative of P by M. Staikova and K. Gribachev in [8]. The basic classes in this classification are W_1, W_2 and W_3. The class $W_0 = W_1 \cap W_2 \cap W_3$ was called the class of Riemannian P-manifolds. Our purpose is to obtain characteristic conditions for each of these classes according to the circulant structure. In the present paper we consider a 4-dimensional differentiable manifold M with a Riemannian metric g and a circulant structure Q, whose fourth power is the identity and Q acts as an isometry on g. This manifold we will denote by (M,Q,g). We study the Riemannian almost product manifold (M,P,g) where $P = Q^2$. The paper is organized as follows. In Sect. 1, some necessary facts about considered manifolds (M,Q,g) and (M,P,g) are recalled. In Sect. 2, the relation between the covariant derivative of P and the covariant derivative of Q is obtained. In Sect. 3, the conditions for the covariant derivative of Q, which imply that (M,P,g) belongs to each of the basic classes of the Staikova-Gribachev classification, are given.

2. Preliminaries

Let M be a 4-dimensional Riemannian manifold equipped with a metric g and an endomorphism Q in the tangent space T_pM at an arbitrary point p on M. Let the coordinates of Q with respect to some basis $\{e_i\}$ of T_pM form the circulant matrix

$$Q = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}. \tag{1.1}$$

Then Q satisfies the equalities

$$Q^4 = id,\ Q^2 \neq \pm id.$$

Let the structure Q be compatible with the metric g, i.e.

$$g(Qx,Qy) = g(x,y). \tag{1.2}$$

 ISSN 1310 - 8271

Here and anywhere in this work x, y, z, u will stand for arbitrary elements of the algebra of the smooth vector fields on M or vectors in T_pM. The Einstein summation convention is used, the range of the summation indices being always $\{1, 2, 3, 4\}$.

Further, we consider a manifold (M, Q, g) equipped with a metric g and a structure Q, which satisfy (1.1) and (1.2). This manifold is studied in [4] and [5].

We denote $P = Q^2$. In [4] it is noted that the manifold (M, P, g) is a Riemannian manifold with an almost product structure P, because $P^2 = id$, $P \neq \pm id$ and $g(Px, Py) = g(x, y)$. Moreover $trP = 0$. For such manifolds is valid the Staikova-Gribachev classication given in [8]. This classification was made with respect to the tensor F of type $(0,3)$ and the Lee form α, which are defined by

$$F(x, y, z) = g((\nabla_x P)y, z),$$
$$\alpha(x) = g^{ij} F(e_i, e_j, x). \qquad (1.3)$$

Here ∇ is the Levi-Civita connection of g, and g^{ij} are the components of the inverse matrix of g with respect to $\{e_i\}$.

The basic classes of the Staikova-Gribachev classification are W_1, W_2 and W_3. Their intersection is the class of Riemannian P-manifolds W_0. A manifold (M, P, g) belongs to each of these classes if it satisfies the following conditions:

$$W_0: \quad F(x, y, z) = 0, \qquad (1.4)$$

$$W_1: \quad F(x, y, z) = \frac{1}{4}((g(x, y)\alpha(z) + g(x, z)\alpha(y) - g(x, Py)\alpha(Pz) - g(x, Pz)\alpha(Py)), \qquad (1.5)$$

$$W_2: \quad F(x, y, Pz) + F(y, z, Px) + F(z, x, Py) = 0, \quad \alpha(z) = 0, \qquad (1.6)$$

$$W_3: \quad F(x, y, z) + F(y, z, x) + F(z, x, y) = 0. \qquad (1.7)$$

It is well known that ∇ satisfies the equalities:

$$(\nabla_x Q)y = \nabla_x Qy - Q\nabla_x y \qquad (1.8)$$

$$(\nabla_x P)y = \nabla_x Py - P\nabla_x y \qquad (1.9)$$

Let the structure Q of a manifold (M, Q, g) be the covariant constant, i.e. $(\nabla_x Q)y = 0$. Then, from (1.8) we obtain successively $\nabla_x Qy = Q\nabla_x y$, $\nabla_x Q^2 y = Q\nabla_x Qy = Q^2\nabla_x y$, thus we get $\nabla_x Py = P\nabla_x y$.

Therefore, from (1.9) it follows $(\nabla_x P)y = 0$.

By using the latter equality and (1.3) we find (1.4). Hence the next theorem is valid.

Theorem 1.1. If the structure Q of the manifold (M, Q, g) satisfies $\nabla Q = 0$, then (M, P, g) belongs to the class W_0.

As it is known the curvature tensor R of ∇ is determined by

$$R(x, y)z = \nabla_x \nabla_y z - \nabla_y \nabla_x z - \nabla_{[x,y]} z.$$

The corresponding tensor of type $(0,4)$ is defined as follows $R(x, y, z, u) = g(R(x, y)z, u)$.

Proposition 1.2. [5] If the structure Q of the manifold (M, Q, g) satisfies $\nabla Q = 0$, then for the curvature tensor R it is valid

$$R(x, y, Qz, Qu) = R(x, y, z, u).$$

We substitute Qz for z and Qu for u in the latter equality, and using Theorem 1.1, we obtain

Corollary 1.3. If the manifold (M, P, g) belongs to W_0, then the curvature tensor R satisfies,

$$R(x, y, Pz, Pu) = R(x, y, z, u).$$

i.e. R is a Riemannian P-tensor.

3. Relation between F and $\overline{F}$

We consider manifolds (M, Q, g) and (M, P, g), where $P = Q^2$. We define a tensor $\overline{F}$ of type $(0,3)$, as follows

$$\overline{F}(x, y, z) = g((\nabla_x Q)y, z),$$
$$\overline{\alpha}(x) = g^{ij} \overline{F}(e_i, e_j, x), \qquad (2.1)$$

where $\overline{\alpha}$ is the Lee form associated to $\overline{F}$.

Theorem 2.1. For the tensors F on the manifold (M,P,g) and $\overline{F}$ on the manifold (M,Q,g) the following equalities are valid:

$$\overline{F}(x,y,z)+\overline{F}(x,Qy,Qz)=F(x,y,Qz), \qquad (2.2)$$

$$\overline{F}(x,y,Q^3z)+\overline{F}(x,Qy,z)=F(x,y,z). \qquad (2.3)$$

Proof. From (1.3) and (1.9), due to $P=Q^2$, we get

$$\begin{aligned} F(x,y,z) &= g(\nabla_x Py - P\nabla_x y, z) \\ &= g(\nabla_x Q^2 y - Q^2\nabla_x y, z), \end{aligned}$$

i.e. $F(x,y,z)=g(\nabla_x Q^2 y - Q^2\nabla_x y, z)$.
Then

$$F(x,y,Qz)=g(\nabla_x Q^2 y, Qz)-g(Q^2\nabla_x y, Qz),$$

from which, because of (1.2) we have

$$\begin{aligned} F(x,y,Qz) &= g(\nabla_x Q^2 y, Qz) \\ &\quad - g(Q\nabla_x y, z), \end{aligned} \qquad (2.4)$$

From (1.8) and (2.1) we obtain

$$\begin{aligned} \overline{F}(x,y,z) &= g(\nabla_x Qy, z) \\ &\quad - g(Q\nabla_x y, z), \end{aligned} \qquad (2.5)$$

and consequently

$$\begin{aligned} \overline{F}(x,Qy,Qz) &= g(\nabla_x Q^2 y, Qz) \\ &\quad - g(\nabla_x Qy, z). \end{aligned} \qquad (2.6)$$

Taking the sum of (2.5) and (2.6) we get

$$\begin{aligned} &\overline{F}(x,y,z)+\overline{F}(x,Qy,Qz) \\ &\qquad = g(\nabla_x Q^2 y, Qz) - g(Q\nabla_x y, z). \end{aligned}$$

Then, having in mind (2.4), we find (2.2). Now we substitute Q^3z for z into (2.2) and using (1.2), we find (2.3). Hence the next theorem is valid.

Theorem 2.2. The manifold (M,P,g) belongs to W_0 if and only if Q satisfies

$$(\nabla_x Q)Qy = -Q(\nabla_x Q)y. \qquad (2.7)$$

Proof. Let $(M,P,g)\in W_0$, i.e. $F=0$. Then, due to (2.2), it follows

$$\overline{F}(x,Qy,Qz)=-\overline{F}(x,y,z).$$

The latter equality and (2.1) imply (2.7).

Vice versa. According to (2.1) and (2.7) we find $\overline{F}(x,y,z)+\overline{F}(x,Qy,Qz)=0$.
Then, due to (2.2), it follows $F=0$, i.e. the manifold $(M,P,g)\in W_0$.

4. Properties of $\overline{F}$

Theorem 3.1. For the tensor $\overline{F}$ on (M,Q,g) the following equalities are valid:

$$\begin{aligned} &\overline{F}(x,y,Q^3z)+\overline{F}(x,Qy,z) \\ &= \overline{F}(x,z,Q^3y)+\overline{F}(x,Qz,y), \end{aligned} \qquad (3.1)$$

$$\begin{aligned} &\overline{F}(x,y,z)+\overline{F}(x,Qy,Qz) \\ &+\overline{F}(x,Q^2y,Q^2z)+\overline{F}(x,Q^3y,Q^3z)=0, \end{aligned} \qquad (3.2)$$

$$\overline{F}(x,y,Qz)=-\overline{F}(x,z,Qy), \qquad (3.3)$$

$$\overline{F}(x,y,Q^3z)=-\overline{F}(x,Q^2z,Qy). \qquad (3.4)$$

Proof. It is known that the tensor F determined by (1.3) has the properties:

$$F(x,y,z)=F(x,z,y), \qquad (3.5)$$

$$F(x,Py,Pz)=-F(x,y,z). \qquad (3.6)$$

Equalities (2.3) and (3.5) imply (3.1).
From (2.3) and (3.6) we get

$$\begin{aligned} &\overline{F}(x,Qy,z)+\overline{F}(x,y,Q^3z) \\ &+\overline{F}(x,Q^3y,Q^2z)+\overline{F}(x,Q^2y,Qz)=0. \end{aligned}$$

In the latter equality we substitute Qz for z, and we obtain (3.2). Further, we substitute Qz for z into (2.5) and we have

$$\begin{aligned}\overline{F}(x,y,Qz) &= g(\nabla_x Qy, Qz) - g(\nabla_x y, z)\\ &= xg(Qy,Qz) - g(Qy, \nabla_x Qz)\\ &- xg(y,z) + g(y, \nabla_x z)\\ &= -g(\nabla_x Qz, Qy) + g(\nabla_x z, y)\\ &= -g(\nabla_x Qz, Qy) + g(Q\nabla_x z, Qy)\\ &= -g(\nabla_x Qz - Q\nabla_x z, Qy) = -\overline{F}(x,z,Qy).\end{aligned}$$

Therefore we get (3.3). From (3.3) directly follows (3.4). Using (1.5), (2.1) and (2.2) we obtain the following

Theorem 3.2. The manifold (M,P,g) belongs to W_1 if and only if the tensor $\overline{F}$ on (M,Q,g) satisfies the following conditions:

$$\begin{aligned}&\overline{F}(x,y,Q^3z) + \overline{F}(x,Qy,z)\\ &= \frac{1}{4}(g(x,y)\alpha(z) + g(x,z)\alpha(y)\\ &+ g(x,Q^2y)\alpha(Q^2z) + g(x,Q^2z)\alpha(Q^2y),\end{aligned}$$

$$\overline{\alpha}(Q^3z) + g^{ij}\overline{F}(e_i, Qe_j, x) = \alpha(z).$$

We apply (2.2) and (2.3) into (1.6) and we find

$$F(x,y,Pz) = \overline{F}(x,y,Qz) + \overline{F}(x,Qy,Q^2z).$$

Therefore we arrive at the following

Theorem 3.3. The manifold (M,P,g) belongs to W_2 if and only if the tensor $\overline{F}$ on (M,Q,g) satisfies the following condition

$$\begin{aligned}&\overline{F}(x,y,Qz) + \overline{F}(x,Qy,Q^2z)\\ &+ \overline{F}(y,z,Qx) + \overline{F}(y,Qz,Q^2x)\\ &+ \overline{F}(z,x,Qy) + \overline{F}(z,Qx,Q^2y) = 0.\end{aligned}$$

We apply (2.3) into (1.7) and we have

Theorem 3.4. The manifold (M,P,g) belongs to W_3 if and only if the tensor $\overline{F}$ on (M,Q,g) satisfies the following condition

$$\begin{aligned}&\overline{F}(x,y,Q^3z) + \overline{F}(x,Qy,z)\\ &+ \overline{F}(y,z,Q^3x) + \overline{F}(y,Qz,x)\\ &+ \overline{F}(z,x,Q^3y) + \overline{F}(z,Qx,y) = 0.\end{aligned}$$

This work was partially supported by project FP17-FMI-008 of the Scientific Research Fund, Paisii Hilendarski University of Plovdiv, Bulgaria.

References

1. S. I. R. Costa, J. E. Strapasson, M. M. S. Alves, T. B. Carlos: Circulant graphs and tessellations on at tori, *Linear Algebra Appl.*, 432(1) (2010) 369-382.
2. A. Kaveh, H. Rahami: Block circulant matrices and application in free vibration analysis of cyclically repetitive structures, *Acta Mech.*, 217 (2011) 51-62.
3. R. M. Tanner, D. Sridhara, A. Sridharan, T. Fuja, D. J. Costello: LDPC Block and convolutional codes based on circulant matrices, *IEEE Trans. Inf. Theory,* 50(12) (2004) 2966-2984
4. I. Dokuzova: Curvature properties of 4-dimensional Riemannian manifolds with a circulant structure. *J. Geom.* (2016). doi:10.1007/s00022-016-0356-9
5. D. Razpopov: Four-dimensional Riemannian manifolds with two circulant structures, In: *Proc. of 44-th Spring Conf. of UBM*, SOK "Kamchia", Bulgaria (2015) 179-185.
6. K. Yano: Differential geometry on complex and almost complex spaces. *Pergamont Press,* Oxford, (1965)
7. A. M. Naveira: A classification of Riemannian almost product manifolds, *Rend. Math.* 3 (1983) 577-592.
8. M. Staikova, K. Gribachev: Cannonical connections and conformal invariantson Riemannian almost product manifolds. *Serdica Math. J.* 18 (1992) 150-161.

Dobrinka Gribacheva
Department of Algebra and Geometry
Faculty of Mathematics and Informatics,
University of Plovdiv Paisii Hilendarski
24 Tzar Asen, 4000 Plovdiv, Bulgaria
E-mail address: dobrinka@uni-plovdiv.bg

Dimitar Razpopov
Department of Mathematics, Informatics and Physics Faculty of Economics,
Agriculture University Plovdiv
12 Mendeleev blvd,
4000 Plovdiv, Bulgaria
E-mail address: razpopov@au-plovdiv.bg

© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 23, 2017

QUADRATURE ITERATIVE METHOD FOR A NUMERICAL SOLUTION OF A NONLINEAR HAMMERSTEIN FUZZY FUNCTIONAL INTEGRAL EQUATION

ALBENA PAVLOVA

Abstract: *In this paper, we obtain error estimation of the iterative method using a fuzzy trapezoidal quadrature formula to solve nonlinear Hammerstein fuzzy functional integral equations. Error estimation of the proposed method is given in terms of uniform and partial modulus of continuity. Finally, an illustrative numerical experiment confirms the theoretical results and demonstrates the accuracy of the method.*

Key words: *Hammerstein fuzzy functional integral equations, approximation solution, fuzzy trapezoidal quadrature formula, modulus of continuity.*

1. Introduction

The study of fuzzy integral equations begins with the investigations of Kaleva [1] and Seikkala [2]. The numerical methods for solving fuzzy integral equations involve various techniques. The method of successive approximations and iterative methods are applied in [3,4].

In this paper, we investigate the nonlinear Hammerstein fuzzy functional integral equation

$$x(t) = g(t) \oplus f\left(t, x(t)\right) \oplus (FR)\int_a^b k(t,s) \odot H\left(s, x(s)\right) ds\,,\ a \le s \le t \le b\,, \tag{1}$$

where $k(t,s)$ is a positive kernel function for $s,t \in [a,b]$. The fuzzy-number-valued functions g, f, H such that $g:[a,b] \to E^1$ and $f, H:[a,b]\times E^1 \to E^1$ are supposed to be continuous.

The convergence of the iterative numerical method proposed in [5] is based on error estimation in approximation of the solution of (1) that was derived using supplementary Lipschitz continuous for g, f, k and H.

The error estimate obtained in this paper is expressed in terms of the modulus of continuity for g and k.

2. Preliminaries

Definition 1. [6] *A fuzzy number is a function* $u:\mathbb{R} \to [0,1]$ *satisfying the following properties*

1. u *is upper semicontinuouson* $\mathbb{R}$,

2. $u(x) = 0$ *outside of some interval* $[c,d]$,

3. there are the real numbers a *and* b *with* $c \le a \le b \le d$, *such that* u *is increasing on* $[c,a]$, *decreasing on and for each* $[b,d]$, *and* $u(x) = 1$ *for each* $x \in [a,b]$,

4. u *is fuzzy convex set (that is* $(\lambda x + (1-\lambda)y) \ge \min\{u(x), u(y)\}$, *for all* $x, y \in \mathbb{R}, \lambda \in [0,1]$).

 ISSN 1310 - 8271

The set of all fuzzy numbers is denoted by E^1. The neutral element with respect to in E^1 is denoted by $\tilde{0} = \chi_{\{0\}}$. For any $0 < r \le 1$ an arbitrary fuzzy number is represented, in parametric form, by an ordered pair of functions $(\underline{u}(r), \overline{u}(r))$, which satisfies the following properties

1. $\underline{u}(r)$ is bounded left continuous non-decreasing function over $[0,1]$,
2. $\overline{u}(r)$ is bounded left continuous non-increasing function over $[0,1]$,
3. $\underline{u}(r) \le \overline{u}(r)$.

Definition 2. [7] *For arbitrary fuzzy numbers* $u = (\underline{u}(r), \overline{u}(r))$, $v = (\underline{v}(r), \overline{v}(r))$, *the quantity* $D(u,v) = \sup_{r \in [0,1]} \max\{|\underline{u}(r) - \underline{v}(r)|, |\overline{u}(r) - \overline{v}(r)|\}$ *is the distance between* u *and* v *and also the following properties hold*

1. (E^1, D) *is a complete metric space,*
2. $D(u \oplus w, v \oplus w) = D(u,v)$, *for all* u, v, $w \in E^1$,
3. $D(u \oplus v, w \oplus e) \le D(u,w) + D(v,e)$, *for all* $u, v, w, e \in E^1$,
4. $D(u \oplus v, \tilde{0}) \le D(u, \tilde{0}) + D(v, \tilde{0})$, *for all* $u, v \in E^1$,
5. $D(k \odot u, k \odot v) = |k| D(u,v)$, *for all* $u, v \in E^1$, *for all* $k \in \mathbb{R}$,
6. $D(k_1 \odot u, k_2 \odot u) = |k_1 - k_2| D(u, \tilde{0})$, *for all* $k_1, k_2 \in \mathbb{R}$ *with* $k_1 k_2 \ge 0$ *and* $u \in E^1$.

Definition 3. [8] *A fuzzy-number-valued function* $f : [a,b] \to E^1$ *is said to be continuous at* $t_0 \in [a,b]$ *if for each* $\varepsilon > 0$ *there is* $\delta > 0$ *such that* $D(f(t), f(s)) < \varepsilon$, *whenever* $t \in [a,b]$ *and* $|t - s| < \delta$. *We say that* f *is fuzzy continuous on* $[a,b]$ *if* f *be continuous for each* $t_0 \in [a,b]$, *and denote the space of all such functions by* $C([a,b], E^1)$.

On the set $C([a,b], E^1) = \{ f : [a,b] \to E^1 ; f \text{ is continuous } \}$, we define $D^*(f,g) = \sup_{t \in [a,b]} D(f(t), g(t))$, for all $f, g \in C([a,b], E^1)$ and $D^*(., \tilde{0})$ is denoted by $\|.\|_3$. It is obvious that $(C([a,b], E^1), D^*)$ is a complete metric space.

Definition 4. [4] *Let* $f : [a,b] \to E^1$, *be a bounded mapping, then the function* $\omega_{[a,b]}(f, \delta) = \sup\{ D(f(t), f(t_0)) : t, t_0 \in [a,b], |t - t_0| \le \delta \}$ *is called the modulus of oscillation of* f *on* $[a,b]$. *In addition if* $f \in C([a,b], E^1)$, *then* $\omega_{[a,b]}(f, \delta)$ *is called uniform modulus of continuity of* f.

According to [4] the following properties hold

1. $D(f(t), f(s)) \le \omega_{[a,b]}(f, |t - s|)$ for any $t, s \in [a,b]$,
2. $\omega_{[a,b]}(f, \delta)$ is a non-decreasing mapping in δ,
3. $\omega_{[a,b]}(f, 0) = 0$,
4. $\omega_{[a,b]}(f, \delta_1 + \delta_2) \le \omega_{[a,b]}(f, \delta_1) + \omega_{[a,b]}(f, \delta_2)$ for any $\delta_1, \delta_2 \ge 0$,
5. $\omega_{[a,b]}(f, n\delta) \le n \omega_{[a,b]}(f, \delta)$ for any $\delta \ge 0$ and $n \in \mathbb{N}$,
6. $\omega_{[a,b]}(f, \lambda\delta) \le (\lambda + 1) \omega_{[a,b]}(f, \delta)$ for any $\delta, \lambda \ge 0$,
7. If $[c,d] \subseteq [a,b]$, then $\omega_{[c,d]}(f, \delta) \le \omega_{[a,b]}(f, \delta)$ for all $\delta \ge 0$.

We consider the integral equation (1) and assume that k is continuous and therefore it is uniformly continuous and there exists $N_K > 0$ such that $N_K = \max\{ |k(t,s)| : t, s \in [a,b] \}$.

We introduce the following conditions:

(i) $g \in C([a,b], E^1)$, $f \in C([a,b] \times E^1, E^1)$, $H \in C([a,b] \times E^1, E^1)$;

(ii) there exist α_f, $\gamma_f \ge 0$ such that $D(f(t_1, u), f(t_2, v)) \le \gamma_f |t_1 - t_2| + \alpha_f D(u,v)$ for all $t_1, t_2 \in [a,b]$, $u, v \in E^1$;

(iii) there exist $\alpha_H, \gamma_H \geq 0$ such that $D\big(H(t_1,u),H(t_2,v)\big) \leq \gamma_H |t_1 - t_2| + \alpha_H D(u,v)$ for all $t_1,t_2 \in [a,b], u,v \in E^1$;

(iv) $B = \alpha_f + N_K \Delta \alpha_H < 1$, where $\Delta = b - a$.

Theorem 1. [5] *Let the conditions* (i)-(iv) *are fulfilled. Then the integral equation* (1) *has unique solution* $x^* \in C([a,b],E^1)$ *and the sequence of successive approximations* $\{x_m\}_{m\in\mathbb{N}} \in C([a,b],E^1)$

$$x_m(t) = g(t) \oplus f\big(t, x_{m-1}(t)\big) \oplus (FR)\int_a^b k(t,s) \odot H\big(s, x_{m-1}(s)\big)ds,\ t \in [a,b],\ m \in \mathbb{N} \tag{2}$$

convergences to x^* *in* $C([a,b],E^1)$ *for any choice of* $x_0 \in C([a,b],E^1)$. *In addition, the following error estimates hold*

$$D(x^*(t),x_m(t)) \leq \frac{B^m}{1-B} D(x_1(t),x_0(t)) \text{ for all } t \in [a,b],\ m \in \mathbb{N} \tag{3}$$

$$D(x^*(t),x_m(t)) \leq \frac{B}{1-B} D(x_m(t),x_{m-1}(t)) \text{ for all } t \in [a,b],\ m \in \mathbb{N}. \tag{4}$$

Choosing $x_0 \in C([a,b],E^1)$, $x_0 = g$ *the inequality* (3) *becomes*

$$D(x^*(t),x_m(t)) \leq \frac{B^m}{1-B}\big(M_0 + N_0 N_K (b-a)\big) \text{ for all } t \in [a,b],\ m \in \mathbb{N}, \tag{5}$$

where $D(f(t,g(t)),\tilde{0}) \leq M_0$ *and* $D(H(t,g(t)),\tilde{0}) \leq N_0$.

Remark 1. Let $\| f \|_{\mathbb{3}} = \sup_{t\in[a,b]} D(f(t,\tilde{0}),\tilde{0})$ and $\| H \|_{\mathbb{3}} = \sup_{t\in[a,b]} D(H(t,\tilde{0}),\tilde{0})$ then the inequality (5) has the form

$$D(x^*(t),x_m(t)) \leq \frac{B^m}{1-B}\big(B \| g \|_{\mathbb{3}} + \| f \|_{\mathbb{3}} + N_K \Delta \| H \|_{\mathbb{3}}\big) \text{ for all } t \in [a,b],\ m \in \mathbb{N}. \tag{6}$$

Theorem 2. [4] *Let* $f : [a,b] \to E^1$ *be a continuous fuzzy –number-valued function. Then*

$$D\left((FR)\int_a^b f(t)\,dt, \frac{b-a}{2} \odot \big(f(a) \oplus f(b)\big)\right) \leq \frac{b-a}{2}\,\omega_{[a,b]}\left(f, \frac{b-a}{2}\right).$$

3. Successive approximations and the iterative algorithm

Now, we introduce a numerical method for solving the fuzzy integral equation (1). We consider the uniform partition of the interval $[a,b]$, $a = t_0 < t_1 < \ldots < t_{n-1} < t_n = b$ with $t_i = a + ih$, $i = \overline{0,n}$, where $h = \frac{b-a}{n}$. Then the following iterative procedure gives the approximate solution of equation (1) on the point t

$$\begin{aligned} &\tilde{x}_0(t) = g(t), \\ &\tilde{x}_m(t) = g(t) \oplus f\big(t, \tilde{x}_{m-1}(t)\big) \oplus \sum_{j=0}^{n-1} \frac{h}{2} \odot \Big[k(t,t_j) \odot H\big(t_j, \tilde{x}_{m-1}(t_j)\big) \oplus k(t,t_{j+1}) \odot H\big(t_{j+1}, \tilde{x}_{m-1}(t_{j+1})\big)\Big], \end{aligned} \tag{7}$$

for $m \geq 1$.

Lemma 1. *Under the conditions* (i)-(iv) *we have*

a) $\omega_{[a,b]}(H(s,x_m(s),h) \leq \gamma_H h + \alpha_H \omega_{[a,b]}(x_{m-1}(s),h)$,

b) $\omega_{[a,b]}(x_m(s),h) \leq \frac{1}{1-\alpha_f}\omega_{[a,b]}(g,h) + \frac{\gamma_f}{1-\alpha_f} h + \Delta\left(\frac{\| H \|_{\mathbb{3}}}{1-\alpha_f} + \frac{\alpha_H B^m \| g \|_{\mathbb{3}}}{B-\alpha_f} + \frac{\alpha_H M}{(1-B)(1-\alpha_f)}\right)\omega_1$,

where $M = \| g \|_{\mathbb{3}} + \| f \|_{\mathbb{3}} + N_K \Delta \| H \|_{\mathbb{3}}$ *and* $\omega_1 = \omega_{[a,b]}(k,\delta) = \sup_{t_1,t_2\in[a,b]} \{|k(t_1,s) - k(t_2,s)| : |t_1 - t_2| \leq \delta\}$

for all $\delta > 0$.

Proof. a) For $s_1, s_2 \in [a,b]$ with $|s_1 - s_2| \le h$ and (iii)

$D(H(s_1, x_m(s_1)), H(s_2, x_m(s_2))) \le \gamma_H \mid s_1 - s_2 \mid + \alpha_H D(x_m(s_1), x_m(s_2)) \le \gamma_H h + \alpha_H \omega_{[a,b]}(x_m, h)$.

b) For $t_1, t_2 \in [a,b]$ with $|t_1 - t_2| \le h$, by using Theorem 2 and (2) we obtain

$$D(x_m(t_1), x_m(t_2)) \le D(g(t_1), g(t_2)) + D(f(t_1, x_{m-1}(t_1)), f(t_2, x_{m-1}(t_2))) +$$

$$+ D\left((FR)\int_a^b k(t_1,s) \odot H(s, x_{m-1}(s)\, ds, (FR)\int_a^b k(t_2,s) \odot H(s, x_{m-1}(s)\, ds \right) \le$$

$$\le \omega_{[a,b]}(g,h) + \gamma_f h + \alpha_f D(x_{m-1}(t_1), x_{m-1}(t_2)) + \int_a^b \mid k(t_1,s) - k(t_2,s) \mid D\left(H(s, x_{m-1}(s)), \tilde{0}\right) ds \le$$

$$\le \omega_{[a,b]}(g,h) + \gamma_f h + \alpha_f D(x_{m-1}(t_1), x_{m-1}(t_2)) + \omega_1 \int_a^b \left(D\left(H(s, x_{m-1}(s)), H(s, \tilde{0})\right) + D\left(H(s, \tilde{0}), \tilde{0}\right)\right) ds \le$$

$$\le \omega_{[a,b]}(g,h) + \gamma_f h + \alpha_f D(x_{m-1}(t_1), x_{m-1}(t_2)) + \omega_1 \Delta\left(\alpha_H \parallel x_{m-1} \parallel_{\mathfrak{B}} + \parallel H \parallel_{\mathfrak{B}}\right) \le$$

$$\le \alpha_f D(x_{m-1}(t_1), x_{m-1}(t_2)) + \omega_1 \Delta \alpha_H \parallel x_{m-1} \parallel_{\mathfrak{B}} + P, \text{ where } P = \omega_{[a,b]}(g,h) + \gamma_f h + \omega_1 \Delta \parallel H \parallel_{\mathfrak{B}}.$$

So, we have

$D(x_m(t_1), x_m(t_2)) \le \alpha_f D(x_{m-1}(t_1), x_{m-1}(t_2)) + \omega_1 \Delta \alpha_H \parallel x_{m-1} \parallel_{\mathfrak{B}} + P$

$D(x_{m-1}(t_1), x_{m-1}(t_2)) \le \alpha_f D(x_{m-2}(t_1), x_{m-2}(t_2)) + \omega_1 \Delta \alpha_H \parallel x_{m-2} \parallel_{\mathfrak{B}} + P$

. . .

$D(x_1(t_1), x_1(t_2)) \le \alpha_f D(g(t_1), g(t_2)) + \omega_1 \Delta \alpha_H \parallel g \parallel_{\mathfrak{B}} + P$.

Multiplying these inequalities by $1,\ \alpha_f, \ldots, \alpha_f^{m-1}$, respectively, and summing them, we have

$$D(x_m(t_1), x_m(t_2)) \le \frac{1}{1-\alpha_f} \alpha_f^m \omega_{[a,b]}(g,h) + \frac{\gamma_f}{1-\alpha_f} h + \frac{\parallel H \parallel_{\mathfrak{B}} \Delta(b-a)}{1-\alpha_f} \omega_1 +$$
$$+ \omega_1 \Delta \alpha_H \left(\parallel x_{m-1} \parallel_{\mathfrak{B}} + \alpha_f \parallel x_{m-2} \parallel_{\mathfrak{B}} + \ldots + \alpha_f^{m-2} \parallel x_1 \parallel_{\mathfrak{B}} + \alpha_f^{m-1} \parallel g \parallel_{\mathfrak{B}} \right) \qquad (8)$$

Now we calculate $\parallel x_{m-1} \parallel_{\mathfrak{B}} + \alpha_f \parallel x_{m-2} \parallel_{\mathfrak{B}} + \ldots + \alpha_f^{m-2} \parallel x_1 \parallel_{\mathfrak{B}} + \alpha_f^{m-1} \parallel g \parallel_{\mathfrak{B}}$.

$$\parallel x_{m-1} \parallel_{\mathfrak{B}} = D(x_{m-1}(t), \tilde{0}) \le D(g(t), \tilde{0}) + D(f(t, x_{m-2}(t)), \tilde{0}) + D\left((FR)\int_a^b k(t,s) \odot H(s, x_{m-2}(s)\, ds, \tilde{0} \right) \le$$

$$\le \parallel g \parallel_{\mathfrak{B}} + \alpha_f \parallel x_{m-2} \parallel_{\mathfrak{B}} + \parallel f \parallel_{\mathfrak{B}} + N_K (b-a)\left(\alpha_H \parallel x_{m-2} \parallel_{\mathfrak{B}} + \parallel H \parallel_{\mathfrak{B}}\right) \le B \parallel x_{m-2} \parallel_{\mathfrak{B}} + M.$$

So, we have

$\parallel x_{m-1} \parallel_{\mathfrak{B}} \le B \parallel x_{m-2} \parallel_{\mathfrak{B}} + M$

$\parallel x_{m-2} \parallel_{\mathfrak{B}} \le B \parallel x_{m-3} \parallel_{\mathfrak{B}} + M$

. . .

$\parallel x_1 \parallel_{\mathfrak{B}} \le B \parallel g \parallel_{\mathfrak{B}} + M$

Multiplying these inequalities by $1,\ B, \ldots, B^{m-2}$, respectively, and summing them, we have

$$\parallel x_{m-1} \parallel_{\mathfrak{B}} \le B^{m-1} \parallel g \parallel_{\mathfrak{B}} + \frac{M}{1-B}.$$

Hence

$$\parallel x_{m-1} \parallel_{\mathfrak{B}} + \alpha_f \parallel x_{m-2} \parallel_{\mathfrak{B}} + \ldots + \alpha_f^{m-2} \parallel x_1 \parallel_{\mathfrak{B}} + \alpha_f^{m-1} \parallel g \parallel_{\mathfrak{B}} \le$$

$$\le B^{m-1} \parallel g \parallel_{\mathfrak{B}} + \frac{M}{1-B} + B^{m-2} \alpha_f \parallel g \parallel_{\mathfrak{B}} + \frac{\alpha_f M}{1-B} + \ldots + B \alpha_f^{m-2} \parallel g \parallel_{\mathfrak{B}} + \alpha_f^{m-2} M + \alpha_f^{m-1} \parallel g \parallel_{\mathfrak{B}} \le$$

$$\le \frac{B^{m-1}}{1 - \frac{\alpha_f}{B}} \parallel g \parallel_{\mathfrak{B}} + \frac{M}{(1-B)(1-\alpha_f)} = \frac{B^m}{B - \alpha_f} \parallel g \parallel_{\mathfrak{B}} + \frac{M}{(1-B)(1-\alpha_f)}.$$

From (8) we obtain

$$D(x_m(t_1),x_m(t_2)) \le \frac{1}{1-\alpha_f}\alpha_f^m \omega_{[a,b]}(g,h) + \frac{\gamma_f}{1-\alpha_f}h + \frac{\| H \|_{\mathbb{R}}\,\Delta}{1-\alpha_f}\omega_1 + \omega_1 \Delta \alpha_H \left(\frac{B^m}{B-\alpha_f}\| g \|_{\mathbb{R}} + \frac{M}{(1-B)(1-\alpha_f)} \right).$$

Hence,

$$\omega_{[a,b]}(x_m(s),h) \le \frac{1}{1-\alpha_f}\omega_{[a,b]}(g,h) + \frac{\gamma_f}{1-\alpha_f}h + \Delta\left(\frac{\| H \|_{\mathbb{R}}}{1-\alpha_f} + \frac{\alpha_H B^m \| g \|_{\mathbb{R}}}{B-\alpha_f} + \frac{\alpha_H M}{(1-B)(1-\alpha_f)} \right)\omega_1 . \quad \square$$

Lemma 2. *Under conditions* (i)-(iv) *we have* $\| \tilde{x}_m \|_{\mathbb{R}} \le B \| \tilde{x}_{m-1} \|_{\mathbb{R}} + M$.

Proof.

$$\| \tilde{x}_m \|_{\mathbb{R}} = D(\tilde{x}_m(t),\tilde{0}) \le D(g(t),\tilde{0}) + D(f(t,\tilde{x}_{m-1}(t)),\tilde{0}) +$$
$$+ D\left(\frac{h}{2}\sum_{j=0}^{n-1} k(t,t_j) \odot H\left(t,\tilde{x}_{m-1}(t_j)\right),\tilde{0} \right) + D\left(\frac{h}{2}\sum_{j=0}^{n-1} k(t,t_{j+1}) \odot H\left(t_{j+1},\tilde{x}_{m-1}(t_{j+1})\right),\tilde{0} \right) \le$$
$$\le \| g \|_{\mathbb{R}} + \| f \|_{\mathbb{R}} + \alpha_f \| \tilde{x}_{m-1} \|_{\mathbb{R}} + \Delta N_K \alpha_H \| \tilde{x}_{m-1} \|_{\mathbb{R}} + \Delta N_K \| H \|_{\mathbb{R}} \le B \| \tilde{x}_{m-1} \|_{\mathbb{R}} + M . \quad \square$$

Theorem 3. *Under conditions* (i)-(iv) *the iterative method* (7) *converges to a unique solution* x^* *of* (1), *and its error estimate is as follows*

$$D^*(x^*,\tilde{x}_m) \le \frac{B^m}{1-B}\left(B \| g \|_{\mathbb{R}} + \| f \|_{\mathbb{R}} + N_K \Delta \| H \|_{\mathbb{R}} \right) +$$
$$+ \frac{3\Delta N_K \alpha_H}{4(1-B)(1-\alpha_f)}\omega_{[a,b]}(g,h) + \frac{3\Delta N_K}{4(1-B)}\left(\gamma_H + \frac{\alpha_H \gamma_f}{1-\alpha_f} \right)h +$$
$$+ \frac{3\Delta^2 N_K \alpha_H}{4(1-B)}\left(\frac{\alpha_H B^{m-1}}{B-\alpha_f} + \frac{M + (1-B)\| H \|_{\mathbb{R}}}{(1-B)(1-\alpha_f)} \right)\omega_1 + \left(\alpha_H m B^{m-1} \| g \|_{\mathbb{R}} + \frac{\alpha_H M + (1-B)\| H \|_{\mathbb{R}}}{(1-B)^2} \right)\Delta\omega_2 ,$$

where $M = \| g \|_{\mathbb{R}} + \| f \|_{\mathbb{R}} + N_K \Delta \| H \|_{\mathbb{R}}$, $\omega_2 = \omega_{[a,b]}(k,\delta) = \sup\limits_{s_1,s_2 \in [a,b]} \{ | k(t,s_1) - k(t,s_2) | : | s_1 - s_2 | \le \delta \}$.

Proof. Considering iterative procedure (7), for all $t \in [a,b]$ we have

$$D(x_m(t),\tilde{x}_m(t)) \le D(g(t),g(t)) + D(f(t,x_{m-1}(t)),f(t,\tilde{x}_{m-1}(t))) +$$
$$+ D\left((FR)\int_a^b k(t,s) \odot H(s,x_{m-1}(s))\,ds, \frac{h}{2}\sum_{j=0}^{n-1}\left[k(t,t_j) \odot H\left(t,\tilde{x}_{m-1}(t_j)\right) \oplus k(t,t_{j+1}) \odot H\left(t_{j+1},\tilde{x}_{m-1}(t_{j+1})\right) \right] \right) \le$$
$$\le \alpha_f D(x_{m-1}(t),\tilde{x}_{m-1}(t)) +$$
$$+ \sum_{j=0}^{n-1} D\left((FR)\int_{t_j}^{t_{j+1}} k(t,s) \odot H(s,x_{m-1}(s))\,ds, \frac{h}{2}\left[k(t,t_j) \odot H\left(t,\tilde{x}_{m-1}(t_j)\right) \oplus k(t,t_{j+1}) \odot H\left(t_{j+1},\tilde{x}_{m-1}(t_{j+1})\right) \right] \right) \le$$
$$\le \alpha_f D(x_{m-1}(t),\tilde{x}_{m-1}(t)) +$$
$$+ \sum_{j=0}^{n-1} D\left((FR)\int_{t_j}^{t_{j+1}} k(t,s) \odot H(s,x_{m-1}(s))\,ds, \frac{h}{2}\left[k(t,s) \odot H\left(t,x_{m-1}(t_j)\right) \oplus k(t,s) \odot H\left(t_{j+1},x_{m-1}(t_{j+1})\right) \right] \right) +$$
$$+ \sum_{j=0}^{n-1} D\left(\frac{h}{2}\left[k(t,s) \odot H\left(t,x_{m-1}(t_j)\right) \oplus k(t,s) \odot H\left(t_{j+1},x_{m-1}(t_{j+1})\right) \right], \right.$$
$$\left. \frac{h}{2}\left[k(t,s) \odot H\left(t,\tilde{x}_{m-1}(t_j)\right) \oplus k(t,s) \odot H\left(t_{j+1},\tilde{x}_{m-1}(t_{j+1})\right) \right] \right) +$$
$$+ \sum_{j=0}^{n-1} D\left(\frac{h}{2}\left[k(t,s) \odot H\left(t,\tilde{x}_{m-1}(t_j)\right) \oplus k(t,s) \odot H\left(t_{j+1},\tilde{x}_{m-1}(t_{j+1})\right) \right], \right.$$
$$\left. \frac{h}{2}\left[k(t,t_j) \odot H\left(t,\tilde{x}_{m-1}(t_j)\right) \oplus k(t,t_{j+1}) \odot H\left(t_{j+1},\tilde{x}_{m-1}(t_{j+1})\right) \right] \right) \le$$

$$\leq \alpha_f D(x_{m-1}(t),\tilde{x}_{m-1}(t)) + N_K \frac{\Delta}{2}\omega_{[a,b]}(H(s,x_{m-1}(s)),\frac{h}{2}) + \Delta N_K \alpha_H D^*(x_{m-1},\tilde{x}_{m-1}) + \Delta\omega_2 D(H\left(t,\tilde{x}_{m-1}(t)\right),\tilde{0}) \leq$$

$$\leq BD^*(x_{m-1},\tilde{x}_{m-1}) + \frac{3\Delta N_K}{4}\omega_{[a,b]}(H(s,x_{m-1}(s)),\frac{h}{2}) + \Delta\omega_2\alpha_H \| \tilde{x}_{m-1} \|_{\mathfrak{F}} + \Delta\omega_2 \| H \|_{\mathfrak{F}}.$$

We denote $P = \frac{3\Delta N_K}{4}\omega_{[a,b]}(H(s,x_{m-1}(s)),\frac{h}{2}) + \Delta\omega_2 \| H \|_{\mathfrak{F}}$.

$D^*(x_m,\tilde{x}_m) \leq BD^*(x_{m-1},\tilde{x}_{m-1}) + \Delta\omega_2\alpha_H \| \tilde{x}_{m-1} \|_{\mathfrak{F}} + P$

$D^*(x_{m-1},\tilde{x}_{m-1}) \leq BD^*(x_{m-2},\tilde{x}_{m-2}) + \Delta\omega_2\alpha_H \| \tilde{x}_{m-2} \|_{\mathfrak{F}} + P$

. . .

$D^*(x_1,\tilde{x}_1) \leq \Delta\omega_2\alpha_H \| g \|_{\mathfrak{F}} + P$

Multiplying these inequalities by $1, B, \ldots, B^{m-1}$, respectively, and summing them, we have

$$D^*(x_m,\tilde{x}_m) \leq \Delta\alpha_H\omega_2\left(\| \tilde{x}_{m-1} \|_{\mathfrak{F}} + B \| \tilde{x}_{m-2} \|_{\mathfrak{F}} + \ldots + B^{m-1} \| g \|_{\mathfrak{F}}\right) + \frac{P}{1-B}.$$

We use Lemma 2 and get

$$\| \tilde{x}_{m-1} \|_{\mathfrak{F}} + B \| \tilde{x}_{m-2} \|_{\mathfrak{F}} + \ldots + B^{m-1} \| g \|_{\mathfrak{F}} \leq mB^{m-1} \| g \|_{\mathfrak{F}} + \frac{M}{(1-B)^2}.$$

Hence,

$$D^*(x_m,\tilde{x}_m) \leq \Delta\alpha_H\omega_2\left(mB^{m-1} \| g \|_{\mathfrak{F}} + \frac{M}{(1-B)^2}\right) + \frac{1}{1-B}\left(\frac{3\Delta N_K}{4}\omega_{[a,b]}(H(s,x_{m-1}(s)),\frac{h}{2}) + \Delta\omega_2 \| H \|_{\mathfrak{F}}\right).$$

From Lemma 1 we obtain

$$D^*(x_m,\tilde{x}_m) \leq \frac{3\Delta N_K\alpha_H}{4(1-B)(1-\alpha_f)}\omega_{[a,b]}(g,h) + \frac{3\Delta N_K}{4(1-B)}\left(\gamma_H + \frac{\alpha_H\gamma_f}{1-\alpha_f}\right)h +$$

$$+ \frac{3\Delta^2 N_K\alpha_H}{4(1-B)}\left(\frac{\alpha_H B^{m-1}}{B-\alpha_f} + \frac{M+(1-B)\| H \|_{\mathfrak{F}}}{(1-B)(1-\alpha_f)}\right)\omega_1 + \left(\alpha_H mB^{m-1} \| g \|_{\mathfrak{F}} + \frac{\alpha_H M+(1-B)\| H \|_{\mathfrak{F}}}{(1-B)^2}\right)\Delta\omega_2.$$

Considering the inequality $D^*(x^*,\tilde{x}_m) \leq D^*(x^*,x_m) + D^*(x_m,\tilde{x}_m)$ and inequality (6) we prove the theorem.□

REFERENCES

1. Kaleva, O. (1987). Fuzzy differential equations. *Fuzzy Sets and Systems,* vol. 24, pp.301-317.
2. Seikkala, S.(1987). On the fuzzy initial value problem. *Fuzzy Sets and Systems*, vol. 24, pp. 319-330.
3. Bica, A. M. (2008). Error estimation in the approximation of the solution of nonlinear fuzzy Fredholm integral equations. *Inf. Sci.*178, pp. 1279-1292.
4. Bede, B. and Gal, S. G. (2004).Quadrature rules for integrals of fuzzy-number-valued functions. *Fuzzy Sets and Systems,* vol. 145, pp.359-380.
5. Enkov, S.,Georgieva, A. and Nikolla, R. (2016).Numerical solution of nonlinear Hammerstein fuzzy functional integral equations. *AIP Conference Proceedings* 1789*,* 030006, http://doi.org/10.1063/1.4968452, pp. 1551-7616.
6. Goetschel, R. and Voxman W. (1986).Elementary fuzzy calculus. *Fuzzy Sets and Systems,* vol. 18, pp.31-43.
7. Wu, C.and Gong, Z.(2001). On Henstock integral of fuzzy-number-valued functions. *Fuzzy Sets and Systems,* vol. 120 , pp. 523-532 .
8. Congxin Wu and Gong Wu (1997). The supremum and infimum of these to fuzzy-numbers and its applications. *J. Math. Anal. Appl.*, 210, pp. 499-511.

Albena Pavlova
Department of MPC
Technical University-Sofia, Plovdiv Branch
25 Tzanko Djustabanov Str., 4000 Plovdiv, Bulgaria
E-mail: akosseva@gmail.com

JOURNAL

OF THE TECHNICAL UNIVERSITY - SOFIA
PLOVDIV BRANCH, BULGARIA
"FUNDAMENTAL SCIENCES AND APPLICATIONS"

publishes new and original results in the fields of MATHEMATICS, MECHANICS, PHYSICS, CHEMISTRY, ECONOMICS and their applications in technical sciences. They can further be expanded to be published elsewhere.

The manuscript should be sent to the Journal Editor-in-Chief
Prof. Michail Petrov, PhD
E-Mail: journal@tu-plovdiv.bg

Acceptance for publication will be based on a positive recommendation from the Journal Editorial Board.

Technical University - Sofia, Plovdiv Branch
25, Tsanko Dyustabanov St., Plovdiv 4000, BULGARIA

http://www.tu-plovdiv.bg **ISSN 1310-8271**