COMMUNICATIONS

COMMUNICATIONS

Dear Readers,

Technological progress in real life is forced by technology convergence and interworking among various areas of human life. Deep penetration can be seen not only between communications and power electrical systems, but these areas have strong influence in such areas as civil and mechanical engineering, medicine, transport and others. Focusing on interdisciplinary branches such as mechatronics, biomedical engineering, nanotechnologies and specializations oriented on information systems in power electrical systems, safety ICT, intelligent transportation systems, multimedia technologies, space communication etc. is also increasing.

This issue of the Communications Letters is focused on selected areas in power electrical systems and communications. It is a pleasure for the Editorial Board members to bring 33 selected and reviewed contributions.

Let us express our gratitude to the authors, colleagues, reviewers and all who contributed and prepared this issue of the Communications - Scientific Letters of the University of Zilina.

Pavol Rafajdus, Martin Vaculik

Galina Cherneva – Elena Dimkina *

SIMULATION AND EXAMINATION OF A SIGNAL MASKING CHAOTIC COMMUNICATION SYSTEM, BASED ON THE DUFFING OSCILLATOR

The study in this paper is focused on the applications of the chaotic Duffing oscillator to signal masking communication scheme. Using the concept of synchronized chaotic systems the signal masking approach is demonstrated with the Duffing oscillator implemented in both the transmitter and receiver. The chaotic masking signal is added at the transmitter to the information signal, and at the receiver, the masking is regenerated and subtracted from the received signal. Chaotic behavior and chaotic masking scheme are designed and simulated using Simulink/Matlab program. All the results are used to illustrate the effectiveness and the applicability of the Duffing oscillator in signal masking communication.

Keywords: *Chaotic signal, chaotic communication system, chaotic masking, Duffing oscillator, error signal.*

1. Introduction

During the last years the use of chaos in communications has become a very important topic of research. Chaos is applied in communications solutions especially in the context of channel encoding and modulation. Observing the ongoing research [1, 2], one observes that the vast majority of the schemes proposed so far concerns modulation in the first place, whereas only a few consider the channel encoding. Chaotic masking [3, 4, 5], switching [6, 7], parameter modulation [8, 9] are some of the methods, proposed in the literature to send binary messages.

Chaotic signal masking is an approach with application in the secure communication [2, 3, 7]. It is based on chaos synchronization [1, 3, 5, 10]. The synchronized motion in coupled oscillator systems is first discussed by Fujisaka and Yamada [1]. The most effective and widely studied approach is due to Pecora and Carroll's discovery [5, 8]. They have reported that certain chaotic systems possess a selfsynchronization property.

The Duffing non-autonomous oscillator [6, 11] is a classical example of a nonlinear dynamical system representing chaotic behavior. The Duffing oscillator has synchronization property and it can be used to the signal masking communication scheme.

In this paper synchronization between two chaotic systems, based on the Duffing oscillator is proposed. The chaotic systems are applied to masking approach in sending secret messages. The communication scheme is simulated using Simulink/Matlab program. The simulation results in this paper investigate the circumstances of application of the Duffing oscillator to drive and response chaotic systems in signal masking communication scheme.

2. Chaotic masking approach description

The chaotic signal masking approach is based on the identical synchronization, where the state of response system converges asymptotically to the state of the drive system [6].

Two continuous-time chaotical systems

$$x^{(1)} = f(x, t) \qquad (1)$$

and

$$x^{(2)} = f(\tilde{x}, t) \qquad (2)$$

are said to synchronize identically if

$$\lim_{t \to \infty} \left(x^{(2)} - x^{(1)} \right) = 0 \qquad (3)$$

for any combination of initial states [4, 7].

In the equations (1) and (2) x, $\tilde{x}$ are according the vectors of the state of the two systems, t – the time.

The system (1) is located in the transmitter and system (2) in the receiver (Fig. 1). If the same initial condition is chosen for the transmitter and the receiver, the both systems will evolve in a syn-

---

* **Galina Cherneva, Elena Dimkina**
Faculty of Telecommunications and Electrical Equipment in Transport, Kableshkov University of Transport, Sofia, Bulgaria,
E-mail: cherneva@vtu.bg

chrony in the sense that, $x^{(1)}$ will continue being equal to $x^{(2)}$ for all $t > 0$.

The idea of a chaotic masking approach is that the information signal is masked by directly adding a chaotic signal at the transmitter. Because chaotic signals are typically broadband, noise like, and difficult to predict, they can be used for masking information-bearing waveforms.

At the receiver which is synchronized to the transmitter the chaotic component is subtracted from the received signal to recover the original transmitted message.

The transmitter state evolution is given by the chaotic dynamics

$$F(x(t)). \tag{4}$$

A chaotic signal $u(t)$ which is a function of the transmitter state $x(t)$ is added to the information signal $m(t)$.

The transmitted signal $s(t)$ is governed by

$$s(t) = u(t) + m(t). \tag{5}$$

The evolution of the receiver state $x_r(t)$ dynamics is given by the dynamics

$$\frac{dx_r}{dt} = F\left(x_r(t), s(t)\right). \tag{6}$$

and $v(t)$ is the chaotic signal which is a function of the receiver state $x_r(t)$.

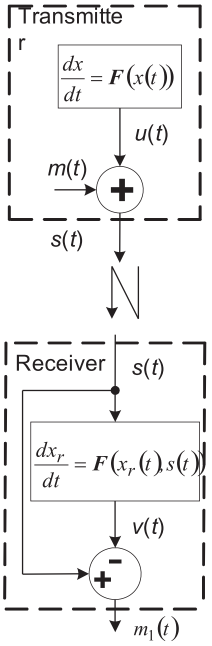


Fig. 1 Block diagram of a chaotic masking communication system

Then the combined signal (5), received by the receiver, can be used to regenerate the chaotic signal $u(t)$ through chaos synchronization. If the synchronization is successful, the $v(t)$ converge to $u(t)$ after a short initial transient and the original message can be recovered by subtracting the reproduced chaotic signal $v(t)$ from the received signal $s(t)$:

$$m_1(t) = s(t) - v(t) \approx s(t) - v(t) = m(t). \tag{7}$$

The masking scheme works only when the amplitudes of the information signals are much smaller than the masking chaotic signals

$$m(t) << |u(t)|. \tag{8}$$

The systems with chaotic masking works effectively in small disturbances in the channel (high ratio signal/noise).

3. Model construction

The Duffing oscillator is given by the second-order differential equation [4] with an external periodic drive term:

$$\ddot{x} + a\dot{x} - x + x^3 = A\sin\omega t, \tag{9}$$

or by an equivalent set of two first order non –autonomous equations

$$\left|\begin{array}{l} \frac{dx}{dt} = y \\ \frac{dy}{dt} = x - x^3 - ay + A\sin\omega t \end{array}\right., \tag{10}$$

where a, A and ω are parameters.

The transmitter state variables are denoted by the vector

$$d = (x, y) \tag{11}$$

and the receiver variables by the vector

$$r = (x_r, y_r) \tag{12}$$

Assume that the transmitter is described by the equation (10).
A response system that will synchronize to the chaotic signals at the transmitter (10) is described by

$$\left|\begin{array}{l} \frac{dx_r}{dt} = y_r - \delta(x - x_r) \\ \frac{dy_r}{dt} = x_r - x_r^3 = ay + A\sin\omega t \end{array}\right., \tag{13}$$

where δ is a coupling parameter to couple the receiver with the transmitter.

This system is obtained from the transmitter equations by renaming variables (x, y) to (x_r, y_r).

The synchronization between the transmitter (10) and receiver (13) has been proved by analyzing the stability of the error signal. It is derived from the difference between the output of the receiver system and that received from the transmitter:

$$e = d - r = (e_1, e_2) = (x - x_r, y - y_r) \tag{14}$$

The error signal is then used to modify the state of the receiver such that it can be synchronized with the transmitter. This signal is given by the system

$$\text{with} \left| \begin{array}{l} \dot{e}_1 = e_2 - \delta e_1 \\ \dot{e}_2 = e_1 - e_1\left(x^2 + x x_r + x_r^2\right) - a e_2 \end{array} \right. . \tag{15}$$

The characteristic equation of system (15) is

$$\lambda^2 + a_1\lambda + a_2 = 0 \tag{16}$$

where

$$a_1 = \delta + a, \tag{17}$$

$$a_2 = \delta a + x^2 + x x_r + x_r^2 - 1. \tag{18}$$

If $a_1 > 0$ and $a_2 > 0$, then

$$e_1 = e_2 = 0$$

is a stable fixed point [3] and the receiver will eventually synchronize with the transmitter.

If $a > 0$ and $\delta > 0$,

then

$$a_1 > 0.$$

Since

$$x^2 + x x_r + x_r^2 = \left(x + \frac{1}{2}x_r\right)^2 + \frac{3}{4}x_r^2 > 0,$$

so a_2 can also be positive if

$$\delta \geq \frac{1}{a} . \tag{19}$$

4. Simulation results

Chaotic signal masking communication scheme, based on the Duffing oscillator, is designed and simulated using Simulink/Matlab program. Figure 2 shows the simulation scheme. The information signal $m(t)$ is a sinus wave of amplitude 1 volts and of 5 kHz frequency.

Transmitter and receiver systems are identical. In the present case $a = 0.2$ is accepted.

As follows of (19) the transmitter and receiver are synchronized when $\delta \geq 5$.

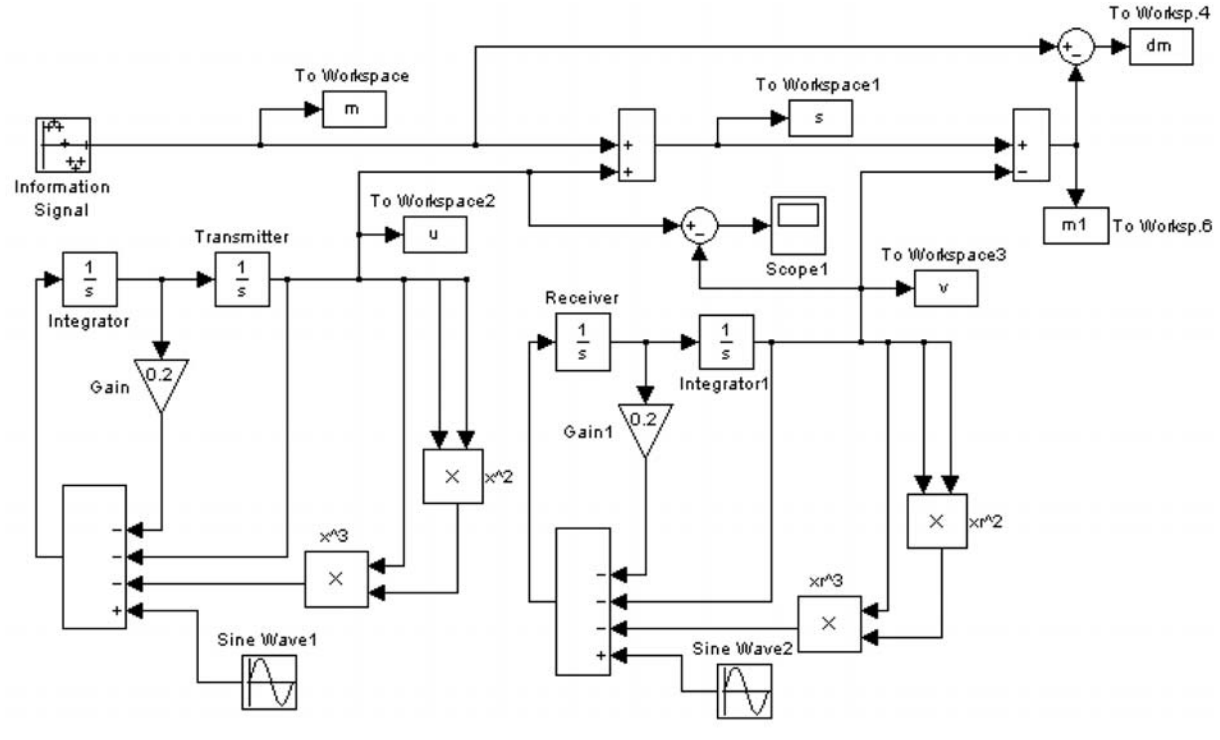


Fig. 2 Simulink modeling of chaotic masking scheme of the Duffing oscillator

In the following analysis the synchronization between the transmitter and the receiver is obtained with an absence of disturbance.

The information signal is added to the generated chaotic signal $u(t)$, and the signal

$$s(t) = m(t) + u(t)$$

is fed into the receiver.

The chaotic signal $v(t)$ is generated in receiver system and

$$m_1(t) = s(t) - v(t) \approx m(t),$$

if

$$u(t) = v(t).$$

The examination is based on a simulation of the systems in the transmitter and the receiver unto different initial conditions and comparison of the driving signal $u(t)$ with the response signal $v(t)$.

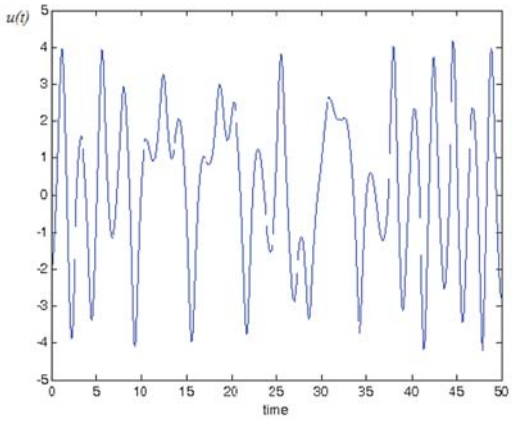


Fig. 3 Phase plane plot of the transmitter chaotic signal

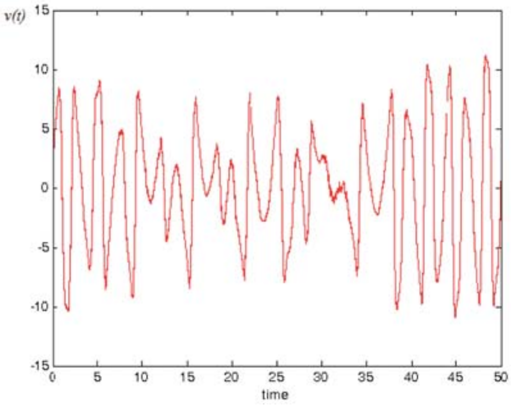


Fig. 4 Phase plane plot of the receiver chaotic signal

The phase plane of the transmitter and receiver chaotic signals is shown in Figs. 3 and 4.

Figure 5 shows the difference $u(t) - v(t)$ of scope.

It is seen that after a transient period the error signal of the transmitter and receiver systems goes to zero, showing that the two systems are synchronized.

One can see from the figures that the synchronization is satisfactory.

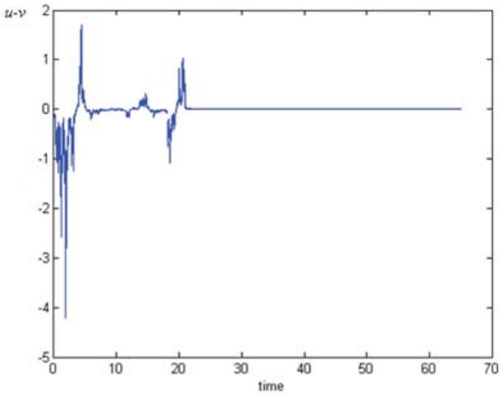


Fig. 5 Error state of the transmitter system and receiver system for $f = 5kHz$

When the frequency of the information signal is increased to 10 kHz , the error signal shown in Fig. 6 is much smaller indicating that the quality of the message signal recovery is greatly improved.

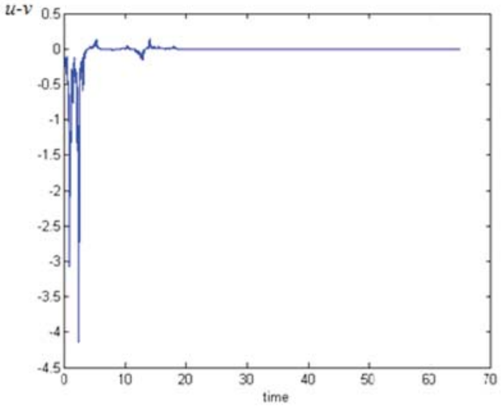


Fig. 6 Error state of the transmitter system and receiver system for $f = 10kHz$

The simulation results show that, when frequency increasing, the two chaotic systems come faster in synchronisation and the oscillation signal (u-v) is smaller. This is confirmed by the graph of Fig. 7, which shows the error signal for $f = 15$kHz.

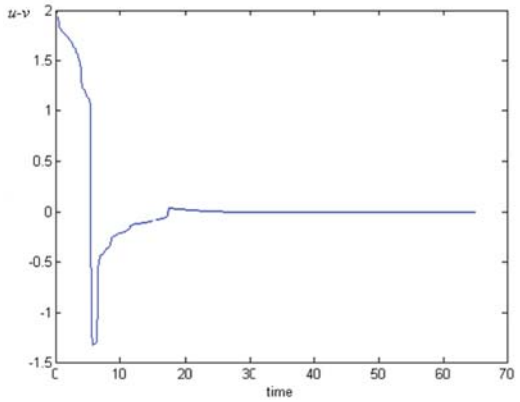


Fig. 7 Error state of the transmitter system and receiver system for f = 15kHz

5. Conclusions

In this paper, chaotic masking communication system of the Duffing oscillator is designed and simulated. The proposed simulation model of Simulink / Matlab allows to monitor the difference between chaotic signals in the transmitter and the receiver and the synchronization time of the two systems.

According to numerical simulations, by a choice of parameters the synchronization error states converge to zero and hence the synchronization between chaotic systems is achieved.

In this paper was investigated the dependence for error signal of the frequency, that with increasing frequency the chaotic systems enter more easily into synchronisation.

Simulation results are used to visualize and illustrate the effectiveness of Duffing chaotic system in signal masking.

The results of simulatition show that the Duffing chaotic system can be used for transmission of confidental information.

References

[1] FUJISAH, H., YARNADA, T.: Stability Theory of &Synchronized Motion in Coupied Oscillator Systems. *Prog. Theor. Phys. 69*, 32–47, 1953.

[2] CELIKOVSKY, S., CHEN, G.: Secure Synchronization of a Class of Chaotic Systems from a Nonlinear Observer Approach. *IEEE Transactions on Automatic Control 50*, 1, 2005.

[3] KOCAREV, L., PARLITZ, U.: General Afor Chaotic Synchronization with Applications to Communication. *Physical Review Letters 74,* 1995.

[4] JORDON, D.W., SMITH, P.: *Nonlinear Ordinary Differential Equation.* N. Y.: Oxford University Press, 1987.

[5] PECORA, L. M., CARROLL, T. L.: Synchronization in Chaotic Systems. *Phys. Rev. Lett. 64,* 8, 1990.

[6] CUOMO, K. M., OPPENHEIM, A. V.: Circuit Implementation of Synchronized Chaos with Applications to Communications. *Phys. Rev. Lett 71,* pp. 65–68, 1993.

[7] TENNY, R., TSIMRING, L. S., ABARBANEL, H. D. I., LARSON, L. E.: Security of Chaos-based Communication and Encryption. *Digital Communications Using Chaos and Nonlinear Dynamics (Institutefor Nonlinear Science).* Springer, 2006, pp. 191–229.

[8] PECORA, L. M., CARROLL, T. L.: Driving System with Chaotic Signals. *Phys. Rev. Lett. A44.* 1991.

[9] KOLUMBAN G., KENNEDY, M. P., CHUA, L. O.: The role of Synchronisation in Digital Communications Using Chaos – part II: Chaotic Modulation and Chaotic Synchronisation, *IEEE Transactions on Circuits and Systems – I: Fundamental Theory and Applications,* vol. 45, No. 11, 1998.

[10] BLEKHMAN, I.: *Synchronization in Science and Technology,* New York: ASME Press, 1988.

[11] JIMENEZ-TRIANA, A., WALLACE K., TANG, S., CHEN, G., GAUTHIER, A.: Chaos Control in Duffing System Using Impulsive Parametric Perturbations, *IEEE Trans. Circuits and Systems – Part II,* vol. 57, No. 4, pp. 305–309, 2010.

Jakub Wosyka – Pavel Pribyl *

DECISION TREES AS A TOOL FOR REAL-TIME TRAVEL TIME ESTIMATION ON HIGHWAYS

This paper presents a time travel estimation model based on a decision tree. Input parameter for travel time prediction is occupancy of detectors. The proposed model was tested on the most widely used arterial in Prague and also in the Czech Republic. This road section has many unmeasured inputs and outputs and with regards to only two detectors within a section it is difficult to estimate the travel time. A temporary installation of ALPR system is used for training the decision tree model which consequently provides reliable travel time estimation.

Keywords: *Travel Time estimation, occupancy, Decision Tree, FCD, ALPR.*

1. Introduction

Travel time (TT) is an important traffic system performance indicator and its great advantage is that it is easy to understand for drivers and generic users in contrast to the usual manner when the Level of Service is expressed as integer number 1 to 5 (“1” is free flow and “5” is congestion).

Travel time can be measured directly by a license plate matching system for every individual vehicle or indirectly estimated through many techniques, mostly using stationary traffic detectors as the most common type of measurement tool, ref. [1], [2]. However the automated licence plate recognition (ALPR) system is still rather expensive and the proposed method uses ALPR only for calibration of estimations of TT provided by static traffic detectors. The described method in this paper is based on a decision tree which estimates travel time using data from loop detectors which were being calibrated by ALPR. ALPR could be installed for a limited time, but it is highly recommended to share the wide range of traffic situations.

2. Literature review

There are many methods for travel time estimation. A trajectory method uses spot speed measures and transforms it to the estimate of the space mean speed. A first conversion formula was described as early as in 1952 by Wardrop, [3]:

$$\overline{v_\tau} = \overline{v_s} + \frac{\sigma_s^2}{\overline{v_s}}, \qquad (1)$$

where $\overline{v_s}$ is average space mean speed, $\overline{v_\tau}$ is average time mean speed and σ_s^2 is variance of the space-mean speed distribution. In the real traffic environment we can directly measure the speed at the spot (time-mean speed), but this formula enables to calculate it from space-mean speed, which is impossible to obtain by standard inductive loops.

Travel time is more meaningful traffic parameter than volume of traffic or occupancy of traffic detectors are. Travel time could be directly used in intelligent transport systems to inform drivers through information displays placed at highways or roads. Travel time is also measured for delay in a road network and could be used as input variable in traffic control systems. The typical example is published in paper [4]. The authors put forward new intersection control optimization method based on real-time delays estimation. The total delay is done by average delay caused by signal control, the additional delay caused by the queuing of previous vehicles and also by the stochastic delay. To obtain the total time delay, it is necessary to use analytical equations with many parameters that must be set in a complicated way. The presented method overcomes the necessity to apply analytical methods with many boundary conditions by the using of AI methods.

The neural network optimized by generic algorithm of models for the real time forecast of traffic flow is presented in [5]. Traffic flow is a better predicable parameter than travel time is. The reason is done by slowly varying curve of traffic flow which is “modulated” by pseudo stochastic randomly distributed clusters of vehicles. The average relative error of traffic flow forecast is 5.53%. However, the proposed method of estimating TT has not a larger deviation than

---

* **Jakub Wosyka, Pavel Pribyl**
Faculty of Transportation Sciences, CVUT Prague, Czech Republic, E-mail: xwosyka@fd.cvut.cz

6%, while it is relatively less complicated than neural network oriented method.

A. Trajectory based methods

Trajectory methods are the simplest and the most commonly used techniques to estimate travel time, [6]. Both methods consider the measured speed is constant until the next detector (or to the half distance on both sides). In half distance method the speed measured by a detector is applicable to half the distance on both sides.

$$TT = \frac{1}{2}\left(\frac{L}{v_1} + \frac{L}{v_2}\right) \tag{2}$$

In average speed method the speed is assumed to be the average speeds measured by detectors 1 and 2.

$$TT = \frac{2L}{(v_1 + v_2)} \tag{3}$$

The main disadvantage of the constant speed trajectory methods is the decreasing performance with increasing traffic congestion.

The author in paper [1] describes a method which uses raw (not time aggregated) data from loop detectors. The spot speed of an individual vehicle is considered constant until the next measurement of another available vehicle. The trajectory of one individual vehicle in time and space is constructed from the sequence of speed measurements obtained from the same vehicle and its few successors. The author evaluated the mean error of his method result between 7 – 10 % considering constant shock wave speed at 22.4 km/h. Unfortunately, the raw input data is essential for this method, but this data is not commonly available online in the Czech Republic.

B. Traffic flow theory based methods

Theoretical models have also been developed for the estimation of travel time from loop detector data based on traffic flow theory.

The Adaptive Kalman Filter (AKF) is used in the paper to improve the travel time estimate from inductive loop using a small number of probe vehicles. Unfortunately probe vehicles were simulated by microscopic simulation model PARAMICS. Because of the inaccuracy of the travel time estimation from point detectors, traffic data from other sources can be incorporated to improve the estimation. Due to the recent advances in probe vehicle technologies, such as Global Positioning Systems (GPS), Automatic Vehicle Identification (AVI), cellular phone positioning, and vehicle re-identification technology, probe vehicle has shown its potential to be another valuable real-time traffic data source according to the [2].

Both data inputs (loops and probe vehicles) do not truly reflect accurate section travel time especially under recurrent or non-recurrent traffic congestion condition.

The proposed method in [2] consists of two parts. First is based on the conservation or continuity equation:

$$\frac{\partial p}{\partial t} + \frac{\partial k}{\partial x} = 0 \text{ (or the traffic generation rate)} \tag{4}$$

where q is flow (vehicles/hour), k is density (vehicles/km), x is location, and t is time.

For a typical urban freeway section including one on-ramp and one off-ramp, the traffic flow passing the section during time period $(t - 1, t)$ can be estimated as:

$$q(t) = \alpha\left[q_u(t) + q_{on}(t)\right] + (1 - \alpha)\left[q_d(t) + q_{off}(t)\right] \tag{5}$$

where α is a smoothing parameter that is set to 0.5. $q_u(t)$ and $q_d(t)$ are traffic flows of the upstream and downstream boundaries within $(t - 1, t)$. $q_{on}(t)$ and $q_{off}(t)$ are total on-ramp and off-ramp traffic flows within $(t - 1, t)$.

Assuming that the traffic inside of the section is homogeneous, an intuitive estimation of the section travel time is:

$$TT(t) = \frac{\Delta x}{v(t)} = \frac{\Delta x}{q(t)}k(t) \tag{6}$$

where Δx is length of the section between upstream and downstream detectors. Section density, $k(t)$, which can be represented as a time series:

$$\begin{aligned} k(t) &= k(t - 1) + \\ &+ \frac{1}{L\Delta x}\left(q_u(t) + q_{in}(t) - q_d(t) - q_{out}(t)\right) \end{aligned} \tag{7}$$

where L is the number of lanes on the mainline freeway.

The Kalman filter is used for correction of section density, which is used for travel time estimation. Section density is treated as a state variable and the section travel time is treated as a measurement variable. The performance of the proposed algorithm is evaluated by Mean Absolute Percentage Error (MAPE). The author also simulated constant error patterns and time-varying error pattern on the loop detectors in order to evaluate how the proposed method can handle these kinds of common errors.

The AKF Algorithm using 5% probe vehicles in the traffic stream (simulated in PARAMICS simulation model) has MAPE in range 7.6 – 9.7% compared to both single data source methods up to 16.6%, which is a significant improvement. The author also stated that with the higher probe rate available, better performance can be expected from the AKF model. When the probe rate is higher than 20% no more improvement is observed.

3. Proposed model

The Travel time estimation model which improves travel time based only on data from the series of loop detectors is presented in this paper (see Fig. 1) and in [7]. The Automatic licence plate recognition (ALPR) system was temporarily used for training of decision tree model (DTM) by determining truth values of travel time on road sections of interest.

The length of the section where estimation of travel times was being tested is 3.6 kilometres and there are three profile measurements by traffic loop detector. The average distance between detectors is greater than one kilometre. The application of the analytical methods described in Chap. 2 was not successful and the difference between real measured TT and estimated time was often higher than 100 percent; depending on the actual traffic volume. The results were poorer for low or very high traffic volume. The reason is simple – the distance between detectors is too large. For instance, the recommended distance between fixed detectors on highways is about 500 m in the Netherlands.

A road without many exits is most appropriate for accurate TT estimation of TT, but it is not the case of our tested road, see Fig. 2. Jizni spojka is an arterial road in Prague that carries more than 100 000 personal cars connecting the southern and northern parts of Prague. Three significant exits and two petrol stations on this short tested section are present. Arriving and departing vehicles significantly influence accuracy of TT estimation. In order to reduce the inaccuracy to the lowest possible level the method based on artificial intelligence is probably best to use. Well-trained decision trees have the capacity to cover all traffic scenarios at this road.

A. Decision tree learning process

Decision trees (DT) can be used to discover features and extract patterns in large databases that are important for predictive models. Decision trees have an established position in artificial intelligence methods. DT are used as classification and prediction tool of travel time estimation based on input parameters which occupancy of detectors is.

Detector occupancy is the ratio between the time that the vehicle is over the detector, and the scanning period expressed in percentage.

The Decision tree learning process requires input variables which are traffic parameters (occupancy) and output values as travel time. Figure 2 depicts a specific part of Jizni spojka in Prague. It is possible to see the whole section enclosed by green (start) and red (finish) flags. Two video-detectors (2 and 3) lie within the road link section, a first one is located outside, in position -364 meters from the start. The ALPR system measured TT between point red flag and green flag and it is 3644 metres long.

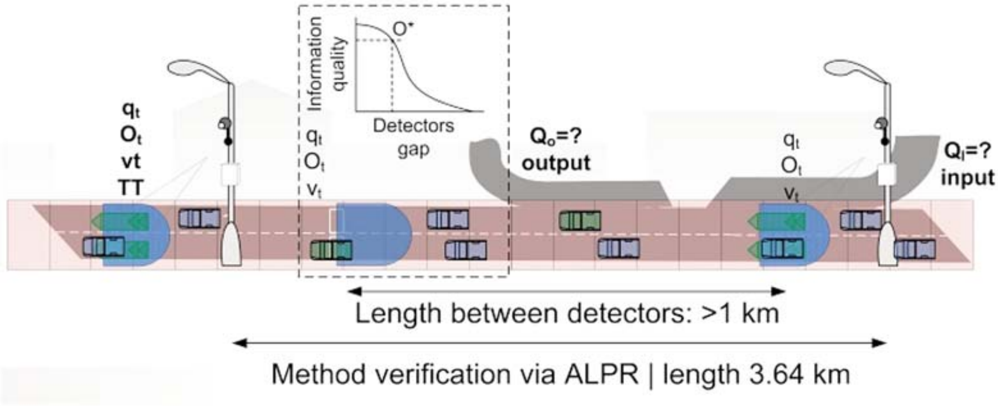


Fig. 1 Decision tree learning via automatic licence plate recognition (ALPR) system

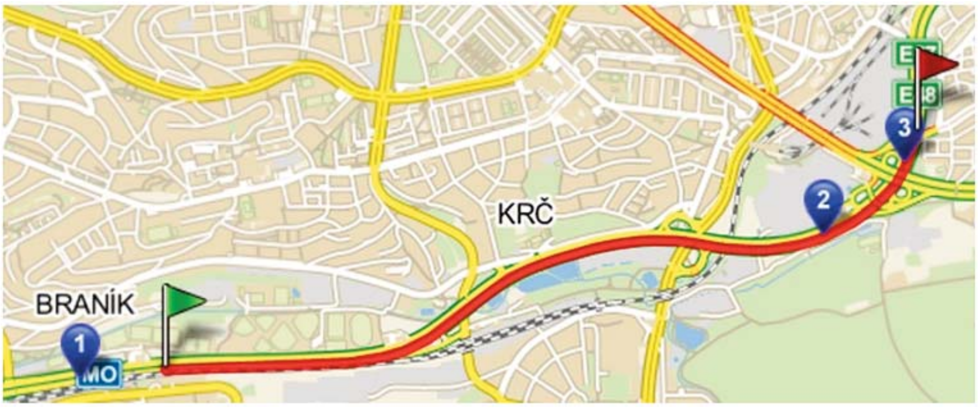


Fig. 2 Link road section in Prague (Jizni spojka)

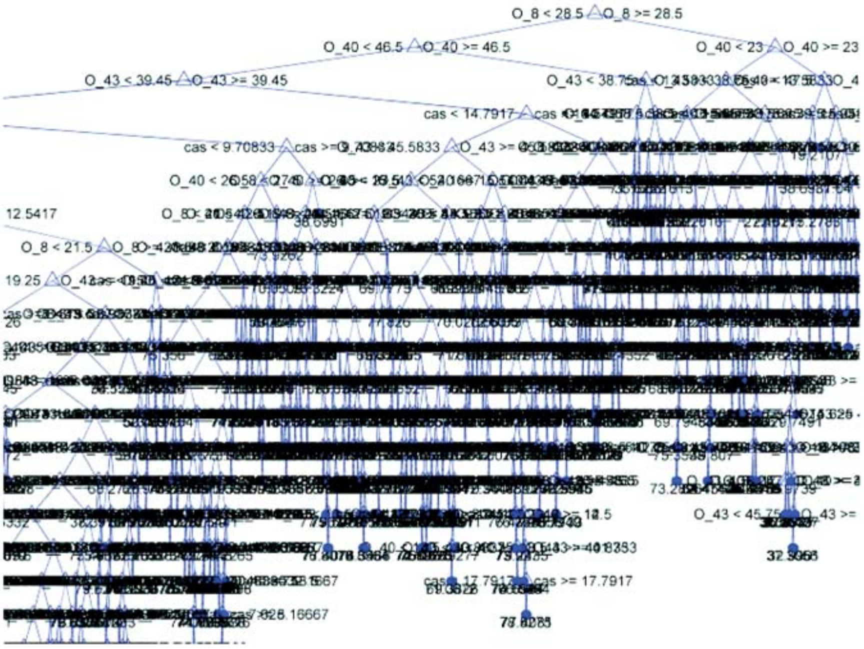


Fig. 3 Part of created decision tree for estimating travel time – Demonstration of complexity

Data recorded in one month contain both free flow and congested traffic states. 90% of data was used for learning decision tree model (training data). The rest was verification data. The preliminary DTM tests used traffic volumes as well as occupancy of detectors. Many experiments showed that the trees are too complex in this case and practically the same results can be achieved when using only occupancy.

The final model uses five minute aggregated occupancy values of all three video- detectors as the input variables. Based on these inputs the travel time estimation is provided in same time step (5 minute time step). Part of related decision tree for travel time estimation can be seen in Fig. 3, which shows how complex created DT could be.

Very fast evaluation of binary condition (if – then) is the significant advantage of decision trees for traffic engineering application. Hence the decision tree-based model can be easily implemented.

B. Model validation and performance evaluation

Consequently unused input data (about 10%) was used for model performance validation using the truth value of TT coming from ALPR system. Three performance indicators were used for evaluation purposes: MSE (mean square error) of average speed, MAE (mean absolute error) of average speed and SSE (sum of square error) of average speed

$$MSE = \frac{1}{n}\sum_{i=1}^{n} e_i^2 \quad MAE = \frac{1}{n}\sum_{i=1}^{n} \sqrt{e_i^2} \quad SSE = \sum_{i=1}^{n} e_i^2 \tag{8}$$

The performance indicators corresponding to twelve different days in May and June are elaborated in Table 1. It is possible to see that the travel time error (in seconds) corresponds to the mean absolute error of the speed on the road section.

Performance evaluation Table 1.

Date	MSE $[km/h]^2$	MAE [km/h]	SSE $[km/h]^2$	Mean time difference [sec.]	TT error
4. 5. 2009	33.46	4.51	6290	9	5%
5. 5. 2009	31.22	4.45	3903	9	5%
12. 5. 2009	35.09	4.66	5158	9	6%
13. 5. 2009	40.86	4.95	7722	9	6%
15. 5. 2009	41.77	4.97	7143	10	6%
18. 5. 2009	23.01	3.69	4304	7	4%
21. 5. 2009	33.27	4.55	6487	9	5%
22. 5. 2009	44.16	5.	7419	10	6%
28. 5. 2009	42.39	5.17	6316	10	6%
29. 5. 2009	37.7	4.74	4825	9	6%
5. 6. 2009	41.69	5.12	6837	10	6%
6. 6. 2009	15.75	2.97	2000	6	4%

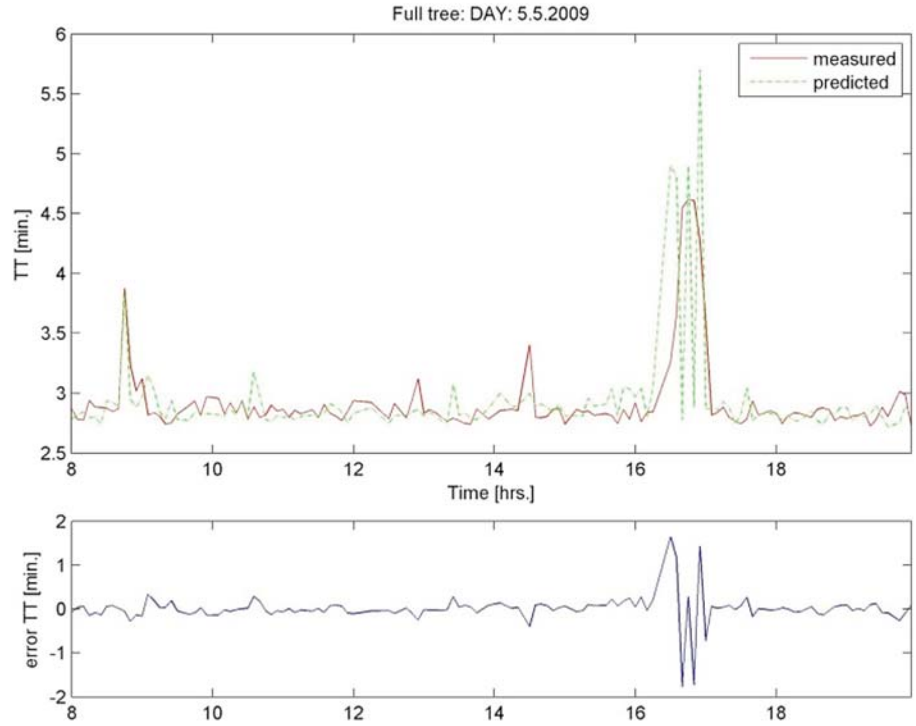
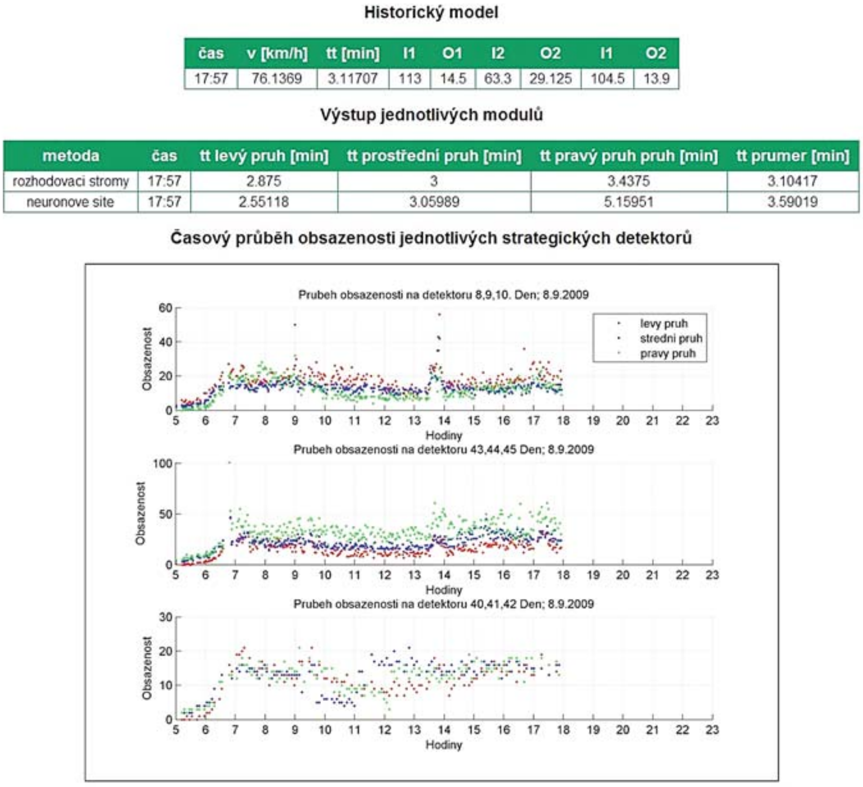


Fig. 4 Measured and predicted TT (above); residual value (bottom)



Historický model

čas	v [km/h]	tt [min]	I1	O1	I2	O2	I1	O2
17:57	76.1369	3.11707	113	14.5	63.3	29.125	104.5	13.9

Výstup jednotlivých modulů

metoda	čas	tt levý pruh [min]	tt prostřední pruh [min]	tt pravý pruh pruh [min]	tt prumer [min]
rozhodovaci stromy	17:57	2.875	3	3.4375	3.10417
neuronove site	17:57	2.55118	3.05989	5.15951	3.59019

Časový průběh obsazenosti jednotlivých strategických detektorů

Fig. 5 Web site opened for road users [7] (in Czech)

The diagram in Fig. 4 (bellow) is graphical presentation of real measured TT by ALPR (red curve) and predicted value by DTM (green curve). X-axis expresses the time. The residual value is the curve at the bottom.

4. Practical application

System for prediction of travel time based “only” on very limited set of input data was implemented as Czech web site of the City Hall of Prague to give a possibility for road users to check accuracy of the proposed system, Fig. 5.

Three different colours represent all three lanes measured on the tested link section (the legend translation is following: levy / stredny / pravy pruh = left / middle / right lane). The occupancy of three detectors is depicted in the figure above and the result of predicted travel time is shown in the table. The TT was predicted by decision trees and also neural network tools. The displayed values as outputs of both AI methods were mutually compared for the research reasons.

5. Conclusion and future work

The presented method is very effective because it enables us to estimate travel times relatively reliably, even when traffic detectors are far from each other. In comparison to the neural networks models, which were also tested in the frame of the project, the decision tree model is more controllable and well-arranged.

Nevertheless the temporary installation of ALPR system for TT measurement also has operational expenditures. The research continues in the direction to replace ALPR equipment by floating car data. The first measurements on the D1 highway proved a penetration of floating cars in the traffic stream volume is about 5 %. This volume of floating cars could very precisely calibrate travel times estimated by fixed detectors.

Acknowledgement

This project was supported by Ministry of Industry and Trade of the Czech Republic No. 2A_1TP1/023 and it has been processed in collaboration with Faculty of Electrical Engineering, University of Žilina.

References

[1] COIFMAN, B.: *Estimating Travel Times and Vehicle Trajectories on Freeways Using Dual Loop Detectors,* Ohio State University : Columbus, 2008

[2] CHU, L., OH, J., RECKER, W.: *Adaptive Kalman Filter Based Freeway Travel time Estimation*, TRB Annual Meeting, 2005

[3] WARDROP, J. G.: *Some Theoretical Aspects of Road Traffic Research,* Proc. of the Institute of Civil Engineers, vol. 1–2, 1952, pp. 325–378, 1952

[4] LIU, B., ZHANG, T., LI, Q., WANG, L.: An Traffic Control Method Based on the Real Time Detector Delay, *Communications – Scientific Letter of Zilina University,* 2, 2010, pp. 76–80, ISSN 1335-4205

[5] HAI-MIN, J., YU-LONG, P.: Neural Network Optimized by Generic Algorithm of Models for Real Time Forecast of Traffic Flow, *Communications – Scientific Letter of Zilina University,* 3, 2010, pp. 80–84, ISSN 1335-4205.

[6] SHEN, L.,: *Freeway Travel Time Estimation and Prediction Using Dynamic Neural Networks,* FIU Electronic Theses and Dissertations, Paper 17, 2008

[7] WOSYKA, J., PRIBYL, P.: *Satelitni lokalizace vozidel v dopravni telematice – SATEL [Vehicle Satellite Localization in Traffic Telematics],* Projekt MPO #. 2A_1TP1/023, VZ 322-2010-1930-EEG, Praha, 2010 (in Czech).

Jan Faber – Marek Stulrajter – Jan Vittek *

SELF-CALIBRATION OF THE RESOLVER SENSOR IN SERVO DRIVE APPLICATION

Resolver sensor is utilized as absolute position transducer in many industrial motion-control applications where robustness and ability to operate in harsh environment is required. In real system, the quality of position and speed measurement is badly affected by the resolver errors. In this paper, the software-based error-compensation technique is integrated into the control structure of the drive. Compensation scheme adapts online to errors in the resolver signals in order to improve accuracy of the measured position and speed. Performance of the compensation scheme and the effect on the drive operation is verified through experimental testing.

Keywords: *Resolver sensor, motion control, position measurement, velocity measurement, electric drives, self-calibration, accuracy.*

1. Introduction

In motion control, the quality of measured position and speed is one of the most important factors to achieve high performance. Although sensorless motor control finds its use in various applications [1], there are still applications where feedback signal of high quality is crucial. Resolver represents a robust position sensor that is suitable for heavy-duty applications like aerospace. In measurement system, the position and speed is extracted from the analog resolver signals that are in ideal case sine and cosine waves. However, the real resolver signals contain errors which results in position and speed inaccuracies. The major contribution to the position error is caused by different amplitudes, dc offsets and phase error of the resolver signals. Furthermore, the noise badly affects the effective resolution and obstructs the identification of the errors. Controller of the electric drive is implemented using Digital Signal Processor (DSP) with high computational power. This computational power can be utilized to improve the position and speed measurement by software-based approach.

There are several works that describe self-calibration and noise reduction methods for the resolver sensor. The ideal resolver signals plotted versus each other form a circle. In the real case, the circle is distorted and can be approximated by an ellipse. In [2], the ellipse fitting technique is used to correct resolver signals in servo drive application to improve accuracy of the position and speed measurement. A recursive least squares approach is used to fit the measured data to an ellipse. The method gives good results, but it is not working during the resting periods when rotor is not moving. In [3], the ellipse fitting technique is modified in order to solve the problem, but the computational costs are too high for the online calibration. Another approach of self-calibration was presented by Heydemann in 1981 [4]. It is also based on the least squares estimation. In order to reduce computational demands, the modified version of the algorithm has been presented so it can be implemented in DSP [5]. A method to suppress systematic errors of the resolver based on Fourier analysis is used in the paper [6]. In the technique, the resolver signals form one complex signal, and the systematic errors are removed by compensating for the non-principal harmonics in the second power of this complex signal. The method is very simple and effective and can be used for online calibration.

As mentioned before, distorted resolver signals cause position error which results in speed and current oscillations. Imperfections in resolver signals are induced by resolver itself with its limited accuracy, by the conditioning circuitry which adjusts these signals prior to AD conversion and also by the ADC errors. Coupling errors like eccentricity contributes to total position error. Another problem represents the imperfect demodulation of the resolver signals which can cause additional position error that is speed-dependent [7].

In this paper the self-calibration method is integrated into the variable-speed field-oriented control scheme of permanent magnet synchronous motor (PMSM) drive equipped with resolver in order to improve accuracy of the position and speed measurement. In following text, the block diagram of self-calibration algorithm is presented and individual parts are described in more details.

2. Servodrive application

Control structure of the PMSM servo drive is shown in Fig. 1 Special integrated circuit called resolver-to-digital converter is commonly used to convert the resolver signals into digital position and

---

* **Jan Faber, Marek Stulrajter, Jan Vittek**
Department of Power Electrical Systems, Faculty of Electrotechnical Engineering, University of Zilina, Slovakia, E-mail: Jan.Faber@fel.uniza.sk

speed (10–16 bits wide). In this case, the resolver-to-digital conversion is software-based, and it is implemented in a Micro-Controller Unit (MCU). Instead of using expensive resolver-to-digital converter, only the low-cost conditioning circuitry is needed to preprocess resolver sine and cosine signals prior to AD conversion. The excitation signal for the resolver is also generated by the interface board. Angle-tracking observer (ATO) is used to extract the position and speed information from the digitized resolver signals. The detail description of the ATO, together with the interface board schematics can be found in [8].

The resolver signals are first sampled by the ADC and then input to the correction block, where the dc offsets are canceled out, the amplitudes are attenuated and the phase error is compensated. The output of the correction block is applied to the ATO input to extract the position and speed from the corrected signals sin(θ), cos(θ). The error parameters A_{sin}, A_{cos}, B_{sin}, B_{cos} and α are calculated in the self-calibration block and fed-back to the correction block. This feedback loop is capable of online identification and compensation of the errors in resolver signals. Function of the individual blocks is described in more details in following text.

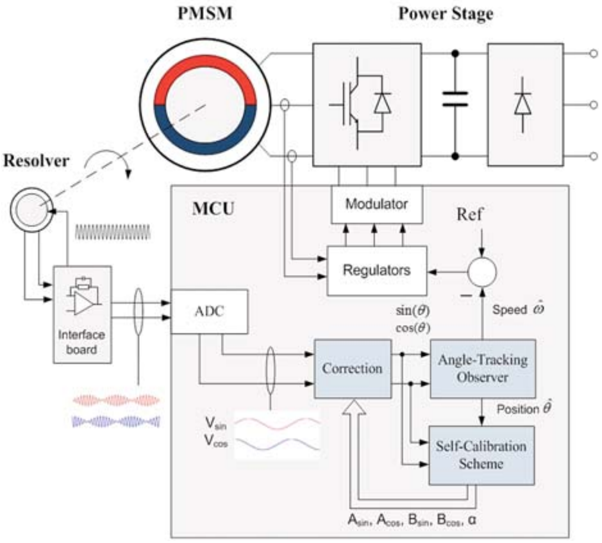


Fig. 1 Control structure of the drive supplemented with the self-calibration algorithm (blue)

A. Resolver Sensor

A resolver is a rotary transformer that houses three windings, primary excitation winding and two secondary sine and cosine windings. Excitation winding is placed in the rotor and the secondary windings are in the stator. The sine and cosine stator windings are mechanically displaced from each other by 90 degrees. The excitation voltage is supplied to the primary excitation winding through a rotary transformer (resolver transmitter type). The schematic of the brushless resolver sensor is shown in Fig. 2.

The rotor is energized with the high frequency voltage signal (excitation frequency $f_{exc} = 10$ kHz). The resolver generates two signals that are multiplied (modulated) by the sine and cosine of the mechanical angle θ. The equation (1) denotes these two signals that are, in ideal case, equal in amplitude, perfectly orthogonal and have no dc offsets.

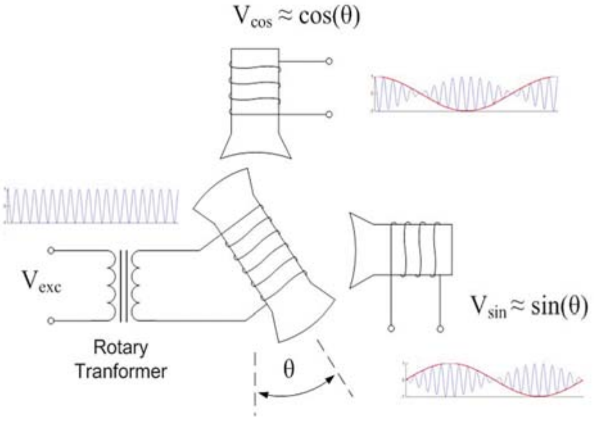


Fig. 2 Resolver schematics

$$\begin{aligned} V_{sin} &= A\sin(\theta)\cos(\theta_{exc}t) \\ V_{cos} &= A\cos(\theta)\cos(\theta_{exc}t) \end{aligned} \tag{1}$$

However, in the real system, resolver signals contain errors and noise. The real resolver sensor can be modeled using equations (2) where A_{sin}, A_{cos} represents the amplitudes, B_{sin}, B_{cos} are the offsets and α is the phase error.

$$\begin{aligned} V_{sin} &= A_{sin}\sin(\theta)\cos(\theta_{exc}t) + B_{sin} \\ V_{cos} &= A_{cos}\cos(\theta + \alpha)\cos(\theta_{exc}t) + B_{cos} \end{aligned} \tag{2}$$

B. Synchronous Demodulation

In software based resolver-to-digital conversion, the resolver signals have to be demodulated during the AD conversion. It is done by taking samples at the maximum of the sinusoids to extract envelope from the modulated signal. Demodulation removes high frequency components ω_{exc} form equations (2), see Fig. 3. The figure also shows the effect of the phase shift t_δ on the demodulation of the resolver signals. Samples of the signals, delayed by the phase shift, are not taken at the peak of sinusoids. In order to utilize full scale of the AD converter, the sampling instant has to be synchronized with the excitation signal of the resolver, otherwise the resolution will decrease.

To ensure perfectly synchronous demodulation is not easy because of the fact that excitation signal suffers from various delays in the conditioning circuitry and in the resolver itself. The long delay can highly distort calculated angle. Position error caused by the phase shift between the excitation signal and the sine and cosine signals that are input to ADC can be calculated as follows:

$$error = \frac{\omega}{\omega_{exc}}\delta \tag{3}$$

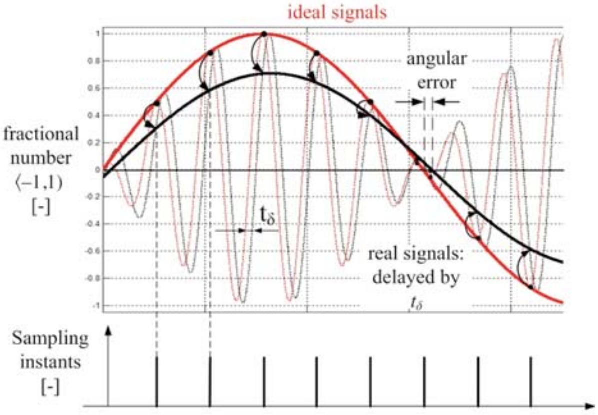


Fig. 3 Sampling of the resolver signal. Effect of the time-phase shift t_δ on the demodulation

According to (3), the error caused by total phase shift δ is lower with higher excitation frequency ω_{exc} and lower rotor speed ω. In high speed of the rotor or with the large δ, the position error can be significant [6].

In order to minimize the effect of the reference phase shift δ on demodulation, the sampling instant has to be attenuated to ensure synchronous demodulation of the resolver signals. The possible algorithm to lock both phases to the peaks of the sinusoids is in Fig. 4.

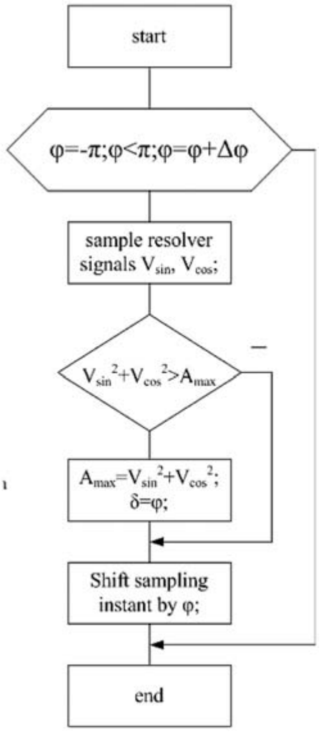


Fig. 4 Offline procedure to ensure synchronous demodulation of the resolver signals

The procedure takes place before the normal operation of the drive, during commissioning stage. Maximization of the sampled resolver signals V_{sin} and V_{cos} can be achieved by modification of the sampling instant relative to the excitation signal of the resolver.

Since both resolver signals are sinusoidal function of time, it is straightforward to determine their maximum by taking sum of squares, which remains constant over the time. Shifting the sampling instant φ by some $\Delta\varphi$ will affect the maximum of the samples. After several iterations, the sampling instant φ that corresponds to the highest maximum is used during the normal operation of the drive.

C. Correction of the Resolver Signals

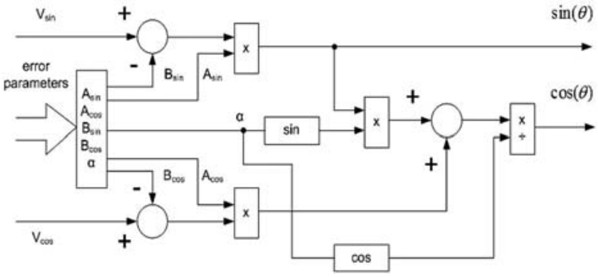


Fig. 5 Correction block

When the phase error α, amplitudes A_{sin}, A_{cos} and dc offsets B_{sin}, B_{cos} are determined, the correction can be done using equations (4).

$$\sin(\theta) = \frac{V_{\sin} - B_{\sin}}{A_{\sin}}$$
$$\cos(\theta) = \frac{\dfrac{V_{\cos} - B_{\cos}}{A_{\cos}} + \sin(\alpha)\sin(\beta)}{\cos(\alpha)} \tag{4}$$

Detail of the correction block is shown in Fig. 5.

D. Angle Tracking Observer

An observer that uses resolver sinusoidal signals to get smooth and accurate estimation of the rotor position and speed is known as Angle Tracking Observer (ATO) [8]. The block diagram is shown in Fig. 6.

The ATO is based on the same principle as Phase-Locked Loop. First, the observer error e is calculated by comparing corrected resolver signals $\sin(\theta)$, $\cos(\theta)$ with their corresponding estimations $\sin(\hat{\theta})$, $\cos(\hat{\theta})$:

$$e = \sin(\theta)\cos(\hat{\theta}) - \cos(\theta)\sin(\hat{\theta}) =$$
$$= \sin(\theta - \hat{\theta}) \approx (\theta - \hat{\theta}) \tag{5}$$

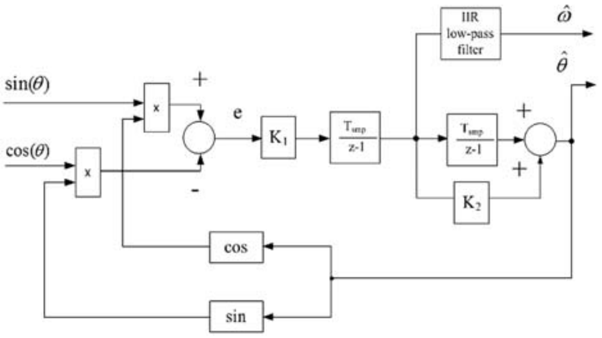


Fig. 6 Angle Tracking Observer

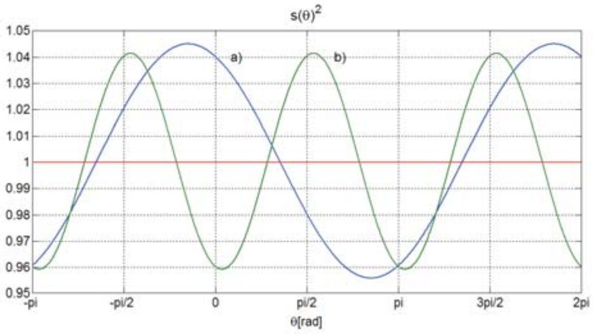


Fig. 7 Squared magnitude of the resolver signals. a) Offset errors $\bar{c} = 0.02 - j0.01$. b) Amplitude and phase errors $\bar{c}_{-1} = 0.02 + j0.01$

The error e is minimized by integrator and PS controller connected in series. ATO Transfer function is a second-order and has one zero. Natural frequency ω_0 [rad.s^{-1}] and damping factor ξ[-] are related to coefficients K_1 and K_2 and can be calculated using equations (6).

$$\omega_0 = \sqrt{K_1}$$
$$\xi = \frac{K_2\sqrt{K_1}}{2} \qquad (6)$$

IIR low-pass filter of the rotor speed is incorporated to achieve smoother waveform.

E. Self-Calibration Block

Purpose of the self-calibration block shown in Fig. 1 is to provide the correction block with new estimate of the error parameters of resolver signals: A_{sin}, A_{cos}, B_{sin}, B_{cos} and α. The self-calibration is based on the method that is presented in [5]. In this technique, the resolver signals form one complex signal $\bar{s}(\theta(t)) = V_{cos}(t) + jV_{sin}(t)$. The signal can be expanded into the Fourier series:

$$\bar{s}\big(\theta(t)\big) = \sum_{k=-\infty}^{\infty} \bar{c}_k e^{jk\theta(t)} \qquad (7)$$

If $k = \{-1, 0, 1\}$, the coefficient $\bar{c}_0$ represents the offset errors (Bcos+jBsin), represents the amplitude difference ($A_{sin} - A_{cos}$), and lm$\{\bar{c}_{-1}\}$ represents the phase error α. The second power of the complex signal $\bar{s}$ is analyzed in order to find the coefficients $\bar{c}_0$ and $\bar{c}_{-1}$. Waveform of the squared magnitude $(\bar{s})^2$ for the distorted resolver signals is shown in Fig. 7 where the red waveform represents the case with ideal resolver signals.

As it is shown in Fig. 7, the distorted resolver signals will introduce harmonics into the squared magnitude $(\bar{s})^2$. The cross correlation of the squared magnitude and the Euler formula is used to find the individual coefficients of the Fourier series, which contain the error parameters of the resolver signals:

$$\hat{\bar{c}}_k = c_1\bar{c}_k = \int_0^{n2\pi} \bar{s}(\theta)e^{j(k-1)\theta}d\theta \qquad (8)$$

However, for the online calculation it is more convenient to calculate the integral in feedback scheme instead of using feed-forward-oriented scheme, given by (7). The block diagram is shown in Fig. 8. The feedback loop includes the correction block and the ATO block as it is shown in Fig. 1. Squared magnitude is calculated from the corrected signals sin(θ), cos(θ).

The $(\bar{s})^2$ is compared with the reference value and fed to the amplitude controller. Actual rotor angle $\hat{\theta}$ is calculated in the ATO block.

Terms sin($2\hat{\theta}$) and cos($2\hat{\theta}$) are used according to (8) to get the amplitude difference ($A_{sin} - A_{cos}$) and the phase error α. The cross correlation for the offset errors B_{sin} and B_{cos} is performed simply with the corrected signals. Five I-regulators have time constant T_I that determines the dynamic behavior of the loop. The time constant has to be set sufficiently large because the error parameters have to be calculated over several periods.

3. Experimental results

Described self-calibration method has been integrated into the controller of a servo drive. The drive includes 3.4kW, 3000 rpm synchronous motor with permanent magnets (PMSM), the sinusoidal encoder, the power converter and the resolver under test. All calculations are performed using 32-bit floating-point microcontroller MPC5643L from Freescale. No additional hardware components are required. Sinusoidal encoder is used as a position reference. The accuracy of the encoder stated in the datasheet is

Application configuration Table 1

Excitation frequency of the resolver	10 kHz
Current loop sampling frequency	10 kHz
Current-loop bandwidth / attenuation	150Hz / 0.707
Speed loop sampling frequency	1 kHz
Speed-loop bandwidth / attenuation	20Hz / 1
Angle-tracking observer bandwidth / attenuation	150Hz / 0.707
Time constant TI of the self calibration loop	1s

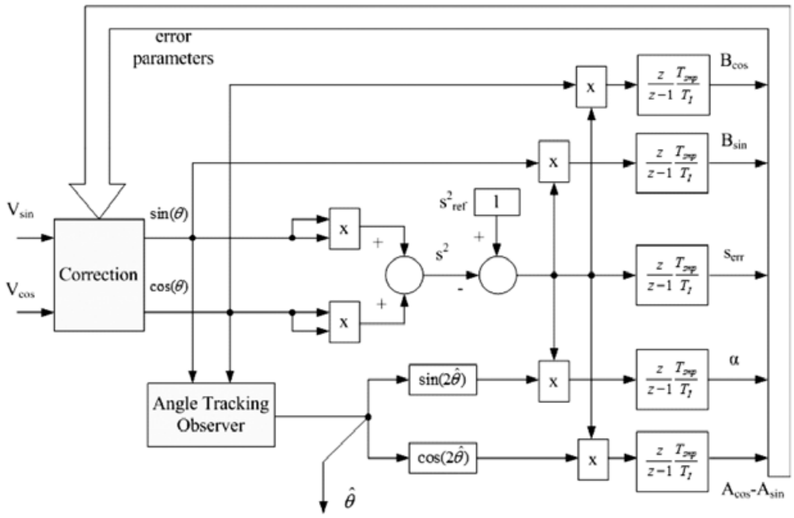


Fig. 8 Block diagram of the self-calibration loop

±20 arc seconds. Resolver under test is coupled to the shaft to provide position feedback for the FOC control algorithm. The accuracy of the resolver according to the datasheet is ±10 arc minutes. Some parameters of application configuration are summarized in Table I.

The results obtained at constant motor speed of 1000 rpm are illustrated in Fig. 9 and Fig. 10. Initially, the self-calibration is turned off and motor runs at constant speed. Oscillations in the speed and position signals are caused partly by the errors in the resolver signals and partly by the eccentricity between the resolver and the reference encoder. The moment when self-calibration is turned on, the observed error parameters start to approach their steady state values. Then these parameters are used for the correction of the resolver signals (Fig. 12), which leads to the reduction of the position error and speed oscillations. Offset error determined by the scheme is about 0.3% for both resolver signals. The difference between amplitudes is about 1.1% and the phase error is 0.06deg. Effect of the self-calibration on the position error and the speed is obvious. Fig. 11 shows comparison of the amplitude spectrum of the position error and speed before, and after the correction Total reduction of the position error is limited by the eccentricity between the resolver and reference encoder, which causes error that is still present in the position, even after applied correction.

Position error is periodic in nature. Close observation shows that there is an offset which is changing with respect to speed. The cause of this dc component being present in the position error is the phase shift δ described in section B. Although to ensure perfectly synchronous demodulation is not an easy task, it is not impossible. Another way to reduce the effect of imperfect demodulation is to increase excitation frequency of the resolver.

Variation of the shaft speed does not affect the dynamic behavior of the self-calibration loop. However, the time constant of the I-controllers (T_I) has to be sufficiently large because the error parameters have to be calculated over several periods of the resolver line signals.

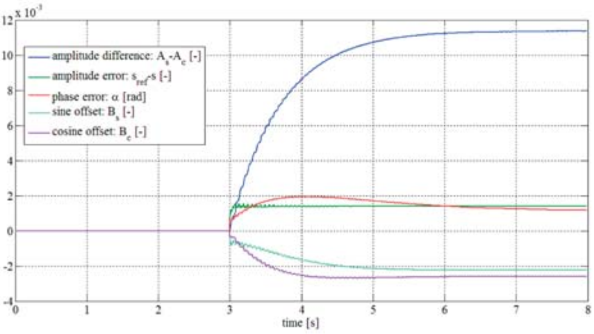


Fig. 9 Error parameters of the resolver signals estimated by Self-calibration loop. Self-calibration is turned-on at time t = 3s

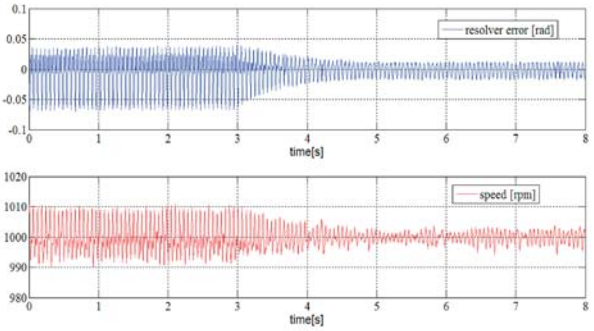


Fig. 10 Resolver error and speed of PMSM at 1000rpm. Self-calibration is turned-on at time t = 3s

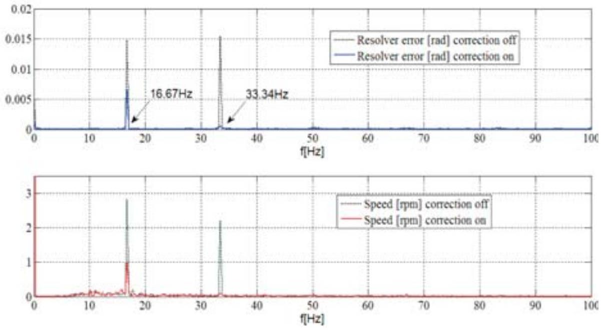


Fig. 11 Amplitude spectrum of the resolver error and speed at constant speed 1000rpm (16.67Hz)

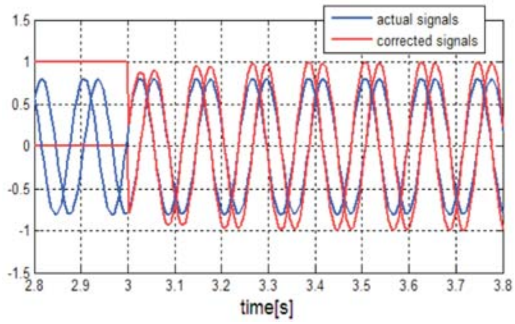


Fig. 12 Actual and corrected resolver signals

4. Conclusion

Self-calibration technique for the resolver sensor has been discussed. In the real system, the imperfections in resolver signals cause position error that introduces oscillations into the speed and current. To reduce these effects, the control structure of the drive is supplemented with self-calibration loop for online identification and compensation of the systematic errors in the resolver signals. Experimental results demonstrated online tuning ability of the method. Reduction of position error results in better running characteristics of the drive. Eccentricity between the resolver and the reference encoder limits additional reduction of position error and the speed oscillations. Furthermore, the imperfect demodulation of the resolver signals caused by phase shift results in additional position error that is proportional to speed and constant over the time.

In general, the described method represents effective and simple way to improve position measurement of the resolver sensor without the need of additional hardware equipment.

Acknowledgement

The authors wish to thank Slovak Grant Agency VEGA for funding the project No. V-11-023-00 and also the Freescale Technology Center in Roznov pod Radhostem for providing the equipment for experimental testing.

Refernces

[1] FILKA, R., BALAZOVIC, P., DOBRUCKY, B.: A Sensorless PM Synchronous Motor Drive for Electric Washers, *Communications - Scientific Letters of the University of Zilina*, 2007, pp. 24–32.

[2] FABER, J., VITTEK, J., MINARECH, P.: *Calibration of Resolver Sensor Using Ellipse-Fitting Technique,* Proc. of intern. Conference Transcom 2011, Zilina, pp. 29–33, 2011.

[3] HOSEINNEZHAD, R., BAB-HADIASHAR, A., HARDING, P.: Calibration of Resolver Sensors in Electromechanical Braking Systems: A Modified Recursive Weighted Least-Squares Approach, *IEEE Transactions on Industrial Electronics,* vol. 54, No. 2, 2007.

[4] L. M. HEYDEMANN, P.: *Determination and Correction of Fringe Measurement Errors in Interferometers,* Applied Optics, 1981.

[5] BALEMI, S.: *Automatic Calibration of Sinusoidal Encoder Signals,* Proc. of IFAC World Congress, Prague, 2005.

[6] BUNTE, A., BEINEKE, S.: High Performance Speed Measurement by Suppression of Systematic Resolve and Encoder Errors, *IEEE Transactions on Industrial Electronics,* February 2004.

[7] GROSS, G., TEXIDO, M., SUDRIA, A., BERGAS, J.: *All-digital Resolver-to-digital Conversion,* European Conference on Power Electronics and Applications, Dresden 2005.

[8] MIENKINA, M., PEKAREK, P., DOBES, F.: *56F80x Resolver Driver and Hardware Interface,* Freescale Semiconductor Application Note, August 2005.

Petr Polcar *

DESIGN, CONSTRUCTION AND EXPERIMENTAL VERIFICATION OF MAGNETORHEOLOGICAL BRAKE

This article deals with the principle, design and verification of the magnetorheological brake. Physical principles, possible designs and advantages of magnetorheological brakes are presented. A design is chosen, simulated and device is built. Simulations are compared with results obtained from measurements on the device, suitable magnetorheological fluid is chosen and torque characteristics of the brake are experimentally verified.

Keywords: *Controlled brake, dynamical characteristics, magnetorheological fluid.*

1. Introduction

Classical brakes possess several disadvantages. Friction of solid mechanical parts wears out the device, the need of hydraulic parts increases space requirements and the control of such brakes is difficult due to the passivity of conventional systems and due to the need of active elements. Magnetorheological brakes offer solution to these problems: they contain no solid to solid part joints, allow continuous changes of deceleration and they are direct electromechanical converters, which makes their control simple. Present-day industry uses magnetorheological brakes in several applications and many designs of magnetorheological brakes can be found nowadays [1, 2].

2. Properties of magnetorheological fluids

Magnetorheological fluids are fluids able to change their viscosity depending on the applied magnetic field. These fluids are thus an example of so called smart materials. They consist of micron sized metallic particles dispersed in a carrier liquid, most often oil. Without an external magnetic field, these particles are randomly scattered in carrier oil and do not influence the viscosity of the whole fluid. If an external magnetic field is applied, particles are organized in the direction of the field and resist possible movements thus increasing the viscosity of overall liquid, as seen in Fig. 1.

Highest viscosity change is measured in the movement direction perpendicular to the direction of the applied magnetic field. Devices working on the change of magnetorheological fluid viscosity principle are thus often geometrically designed to work in such a way.

Magnetorheological fluids are nonconductive and the size of their ferromagnetic particles is in micrometers. Such a small size means that eddy currents are not present in the fluid or particles, which makes their magnetization characteristics static [3].

3. Used Magnetorheological Fluid

Magnetorheological fluid MRHCCS4-B obtained from Liquids Research Ltc. was used as the braking liquid of proposed device.

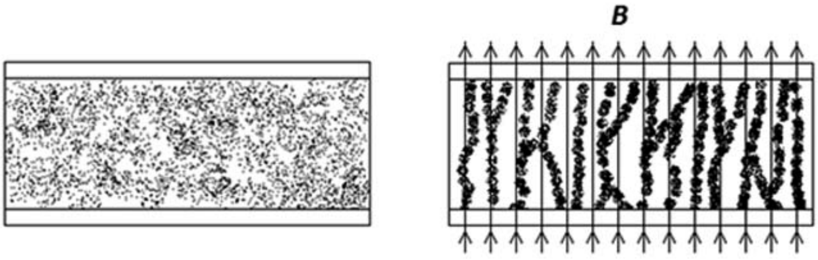


Fig. 1 Magnetic fluids viscosity change: Left – Magnetorheological fluid without external magnetic field; Right – Magnetic particles are organized in the direction of the applied magnetic field

---

* **Petr Polcar**
Faculty of Electrical Engineering, University of West Bohemia, Pilsen, Czech Republic, E-mail: paladin@kte.zcu.cz

Its magnetic properties were determined using our method for magnetic properties measurement mentioned in [4]. The relative permeability in the linear part of magnetization characteristics is $\mu_r = 1.61$. The change of viscosity depending on the perpendicular applied magnetic field was obtained from the University of Timisoara measurements [5]. Saturated magnetization of this fluid is 750 mT.

4. Brake Design

The physical principle of the magnetorheological brake design can be seen in Fig. 2. The brake disc is rotating in a gap of a magnetic circuit filled with the magnetorheological fluid. If coils are powered, the resulting magnetic field is led through the magnetic circuit and influences the fluid in its gap. By changing the viscosity of the fluid, the braking of the whole system can be continually controlled. No direct connection of metal parts is present, and braking is achieved by viscous losses in the fluid.

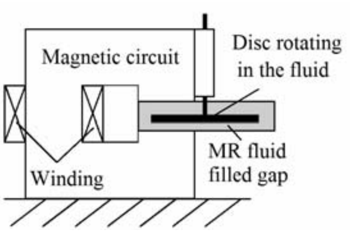


Fig. 2 Magnetorheological brake principle scheme

Most mechanically stressed part of the brake is the connection of the disc and the shaft. These parts must be properly fixed in order to withstand the resulting torque. The material the disc is made of has impact on the brake principle – if it is made of conductive material, the device works as an eddy current brake as well because of eddy currents inducted into the moving conductive material [6]. These currents generate magnetic field and forces in the opposite direction to the movement. Resulting drag torque is then given by the superposition of the drag torque generated by eddy currents and the drag torque generated by viscous losses. If the material is nonconductive, the device works as a pure magnetorheological brake, but nonconductive materials possess worse mechanical properties.

Significant disadvantage of this braking system is the initial viscosity of the fluid without the presence of an external magnetic field. The rotating disc is braked even without the applied magnetic field. Resulting energy is changed into heat and warms used fluid, which results in cooling requirement.

5. Mathematical Model of the Magnetic Field and Experimental Verification

Magnetic field in range $B = 0 \sim 400$ mT is needed in the brake gap to achieve demanded viscosity change of the magnetorheological fluid. Magnetic circuit was made and the problem was simulated to properly design the number of coils and to choose suitable power source for the brake coils. Magnetic field in the area of the magnetic fluid filled gap is observed. Because of known magnetic properties of the fluid ($\mu_r = 1.61$ in the linear part of the magnetization characteristics and the saturation magnetization of 750 mT) the problem can be treated as non coupled, the changes in magnetic fluid do not influence resulting magnetic field, and linear, the device operates in the linear part of the magnetization curve. The value of this field was computed using the mathematical model respecting the coil filling coefficient, then measured on the experimental device and results were compared.

Mathematical model of the problem was created using the FEM application Agros2D. The magnetic vector potential equation was solved in order to acquire the distribution of the magnetic field.

$$rot\frac{1}{\mu}rot\ A = J_{ext} \tag{1}$$

The value of the magnetic field B can be easily obtained from the formula

$$B = rot\,A\ . \tag{2}$$

The convergence of results was observed. The mathematical model with 7689 nodes and 3216 elements built in Agros2D software was found accurate and quick enough. Second order elements were used to ensure higher accuracy and faster computing of the model. Distribution of the resulting magnetic field can be seen in Fig. 3.

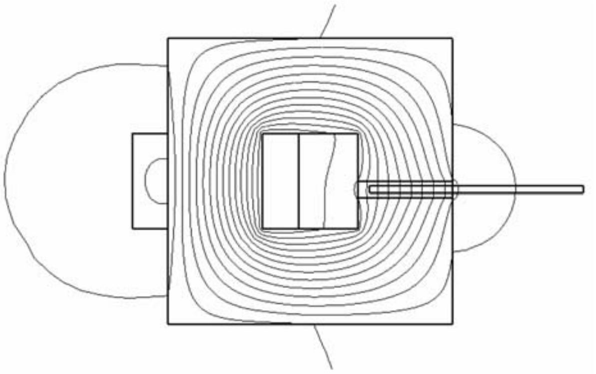


Fig. 3 Distribution of the magnetic field obtained from the numerical simulation of the brake in Agros2D software

Resulting magnetic field for different values of winding current was compared with values obtained by measuring on the magnetic circuit using the Elimag MP-1 fluxmeter with hall probe injected into the magnetic fluid filled gap. Results confirming the assumption that the device works in the linear part of the magnetization characteristics can be seen in Fig. 4.

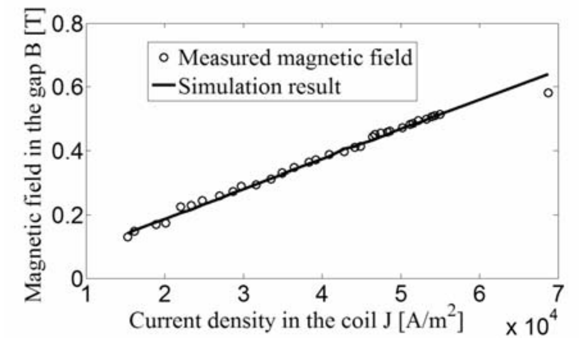


Fig. 4 Results of the magnetic field in the brake gap mathematical modeling compared with measured results

6. Experimental Device

An experimental device of magnetorheological brake was constructed and can be seen in Fig. 5. The magnetorheological fluid container is open to ensure better maintenance of the device and easy injection of measurement probes. This container would be sealed in a real application. Magnetorheological fluid MRHCCS4-B was used to fill it. The coils are powered with direct current so skin-effect is not present and construction steel is sufficient for the magnetic circuit material, there is no need for transformer plates. Current density is achieved by powering 500 coils made of 0.6 mm copper wire with values of direct current from up to 3A DC power source. The same values of current density can be achieved using more coils and fewer currents as well. The disc shaft and the disc itself are made of aluminum in order to obtain higher mechanical strength of disc to shaft connection.

The operation of the fluid in the device can be seen in Fig. 6.

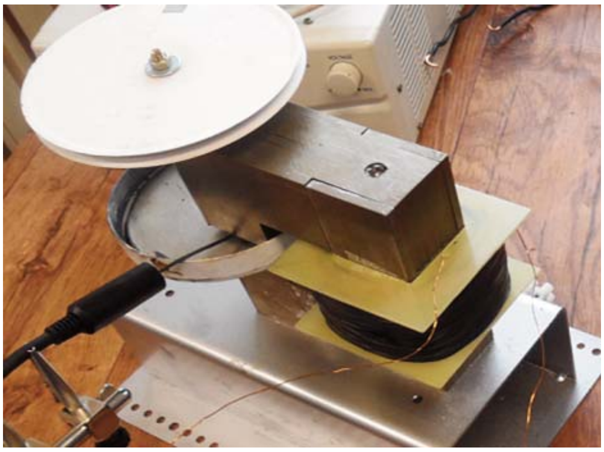


Fig. 5 Experimental magnetorheological brake with hall probe measuring the magnetic field in the brake chamber

7. Drag Torque of the Magnetorheological Brake

The viscous drag torque of a disc rotating in a fluid depends on the fluid viscosity η, revolution of the disc ω and the geometrical dimensions of the disc. This torque can be counted from the acquired values of the magnetic field and viscosity-magnetic field-disc revolution characteristics of the used fluid MRHCCS4-B [5]. As can be seen, the drag torque is present even when the fluid is not affected by magnetic field, the magnetorheological fluid possesses an initial viscosity causing viscous losses. When the viscosity rises, the drag torque changes as well. The total drag torque can be thus expressed as the superposition of drag torque caused by the initial viscosity, the drag torque caused by the increase of the viscosity and the drag torque caused by the eddy currents braking.

Fig. 6 Magnetorheological brake influences the viscosity of the fluid in the gap: Left – Coils are not powered; Right – Coils are powered with current density of 6 ·10^4 A/m^2

$$M = M_{rheo(I=0)} + M_{rheo}(I) + M_{eddy} \qquad (3)$$

The drag torque Meddy caused by eddy currents induced into the disc is given by the magnetic field and the square of the disc rotation and can be easily deduced from the elemental relations given by the electromagnetic field theory (see e.g. [7]). Because of the low values of the magnetic field in the brake gap (in order of hundreds of mT), the value of this torque is several orders lower than the values of the torques given by viscous losses and can be neglected. This assumption was experimentally verified by measuring the drag torque of the brake with an aluminum disc without the presence of the magnetorheological fluid.

Fig. 7 The setup for measuring the dynamical drag torque of the magnetorheological brake

The dynamical drag torque characteristics of the device were measured using the experimental setup that can be seen in Fig. 7.

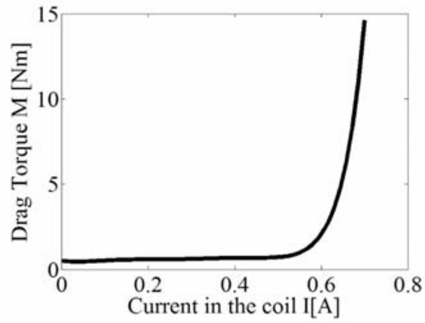


Fig. 8 Measured drag torque as a function of current powering the brake coil

The experimental brake was driven by an asynchronous machine FT4C52S. The mechanical revolution of the brake disc was measured with tachometer and the active power supporting the asynchronous machine was observed using a power network analyzer. The relation between the power and the disc revolution is given by:

$$P_{\kappa} = M_{\omega} \, . \qquad (4)$$

Where P stands for the active power, κ represents the machine efficiency and ω is the revolution of the disc. The drag torque characteristics depending on the current powering the brake coils can be seen in Fig. 8.

Drag torque Mrheo($I = 0$) caused by the initial viscosity of the magnetorheological fluid is not present in the characteristics, the graph shows the increase in the drag torque caused by the viscosity change of the magnetorheological fluid. When driven by the used motor, the revolution of the disc dropped under $\omega = 1\ \mathrm{s}^{-1}$ when powering the coil with current higher than $I = 0.7$ A. The disc stopped rotating when the coil was powered with the current of $I = 1$ A.

8. Conclusion

A design of a magnetorheological brake was presented and an experimental mechanism was built. This brake has several advantages: no direct solid to solid part connection causing negligible mechanical wear, direct electric control and low space requirements. Current density in order of $J = 10^4\ \mathrm{A/m^2}$ is enough to create such a viscosity change in the MRHCCS4-B fluid to stop the movement of the device driven by the used motor. Such values of current densities are quite easy to achieve, it was accomplished by 500 coils powered with 1 A in our case. The prime disadvantage of this brake design is the passive braking caused by viscous losses in injected fluid even without the applied magnetic field that predetermines the brake for the use in relatively slow running devices. Optimizing the magnetic circuit to decrease power requirements of the device and acquiring drag torque characteristics of the brake with different types of magnetorheological fluids is the theme of our future research in this field of interest.

Acknowledgment

This work was supported by the specific research project of the University of West Bohemia SGS-2012-039.

References

[1] LI, W. H., DU, W. H.: Design and Experimental Evaluation of a Magnetorheological Brake. *Intern. J. of Advanced Manufacturing Technology,* 21(7), pp. 508-515, 2003.

[2] WANG, J., MENG, G.: Magnetorheological Fluid Devices: Principles, Characteristics and Applications in Mechanical Engineering. *Proc. of the Institution of Mechanical Engineers Part L: J. of Materials: Design and Applications*, 215(3), pp. 165-174, 2001.

[3] MAYER, D., POLCAR, P. A Novel Approach to Measurement of Permeability of Magnetic Fluids. *Przeglad Elektrotechniczny,* 88(7 B), 2012, pp. 229–231.

[4] ODENBACH, S.: Ferrofluids – Magnetically Controlled Suspensions, Colloids and Surfaces, *Physicochemical and Engineering Aspects,* 2003, pp. 171–178.

[5] BALAN, C., BROBOANA, D., GHEORGHIU, E., VEKAS, L.: Rheological Characterization of Complex Fluids in Electro-magnetic Fields, *J. of Non-Newtonian Fluid Mechanics,* vol. 154, No. 1, September 2008, pp. 22–30, ISSN 0377-0257.

[6] JANOUSEK, L., MAREK, T., GOMBARSKA, D.: Eddy Current Non-destructive Evaluation of Conductive Materials. *Communications – Scientific Letters of the University of Zilina,* 8(1), pp. 29–33, 2006.

[7] DOLEZEL, I., KARBAN. P., SOLIN, P.: *Integral Methods in Low-frequency Electromagnetics.* Hoboken : John Wiley & Sons, 2009. p. 388, ISBN 978-0-470-19550-5.

Roman Radil – Ladislav Janousek *

YEAST GROWTH INFLUENCED BY PARALLEL COMBINATION OF TIME-VARYING AND STATIC LF EMF

Presented article deals with the mechanisms of low frequency electromagnetic field impact on biological structures. Special attention is paid to the theories describing the influence of applied electromagnetic field on microtubule structures and ions bound on the cell membrane surface. These theories are confronted with the experiments presented herein. For better reliability of the experiments, new exposure coil system has been designed and created. Using this exposure system, results are conducted on the eukaryotic cells represented by saccharomyces cerevisiae. Observed effects show a good correlation mainly with ion parametric resonance model predictions proposed by Lednev and are discussed in the scope of further investigation possibilities.

Keywords: *Low frequency electromagnetic field, ion parametric resonance, magnetic flux density, saccharomyces cerevisiae.*

1. Introduction

Investigations about non-thermal effect of low frequency electromagnetic field (LF EMF) applied on biological structures are still full of ambiguity and uncertainty. This is caused by the absence of a general theory about mechanisms on the cellular level that could be accepted by the whole scientific society. However, the research in this area has succeeded at least partially, because several theories have been established and published.

In this paper the two of the theories that seem to be most prominent in the research during last few years are briefly presented. Both of the theories are confronted with our experimental work, where the influence of LF EMF on the yeast strand VIVO is investigated. Findings and results are discussed within the conception of the mentioned theories, while some conclusions and recommendations for further work are presented.

2. Interaction mechanizms between LF EMF and biological structures

As it has been published in many studies before [1, 2], the EMF is naturally presented (at least in the form of transmembrane potential) in biological environment and aligned with many processes (i.e. communication between cells, proliferation processes and others). But most recent scientific attention is paid to phenomenon of non-thermal connection between biosystems (albeit unicellular in our case) and man-made electromagnetic fields, at least in the specific frequency range. The investigations in this field showed a multitude of possibilities to pursue in the process of quantifying and elucidating the mechanisms of actions observed over the course of considered experiments. However, for the purpose of this work, two mechanisms that are currently probably most prominent have been chosen.

A. Frohlich's theory of coherent oscilations

The first mechanism is based on notions by H. Frohlich [3] and explores highly polar structures within living cells and vibrations thereof, creating conditions for endogenous electromagnetic field generation. This notion was further explored and developed by authors at the Institute of Photonics, Prague, Czech Republic and was also used to quantify the mechanism of non-invasive cancer detection proposed by Vedruccio et al. and since commercially manufactured under the name TrimProbe [4]. The basis of this theory lies in the presence of microtubules, hollow cylinders made of highly polar protein pairs (α- and β-tubulin), Fig. 1.

Interestingly, these are most prominent during the division process, wherein they serve both as structural elements and locomotive helpers especially during the separation of nuclei of mother and daughter cells. Furthermore, experiments conducted by Pokorny et al. [5] showed significantly enhanced levels of electromagnetic activity during the budding phase of yeast cells and within synchronized yeast cells. Based on this theory it could be assumed that application of exogenous electromagnetic field within the frequency range employed by microtubules might lead to disruption of the underlying physiological function(s).

B. Lednev's ion resonance theories

The second mechanism lies in the interpretation of Lednev's Ion Parametric Resonance (IPR) model [6]. This model is based

---

* **Roman Radil, Ladislav Janousek**
Department of Electromagnetic and Biomedical Engineering, Faculty of Electrical Engineering, University of Zilina, Slovakia,
E-mail: roman.radil@fel.uniza.sk

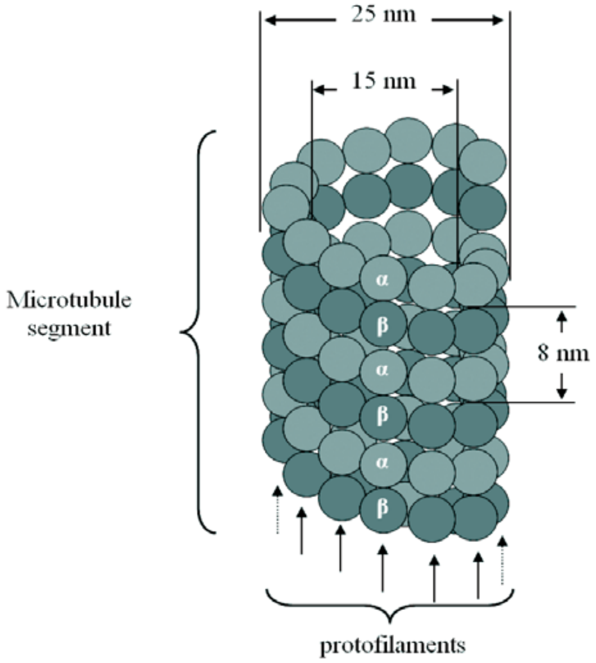


Fig. 1 Microtubule segment

on parallel combination of static and alternating magnetic fields applied to a biosystem, first described by McLeod and Liboff in [7]. The model is principally shown in Fig. 2 and can be interpreted as an adaptation of atomic spectroscopy to biological environment.

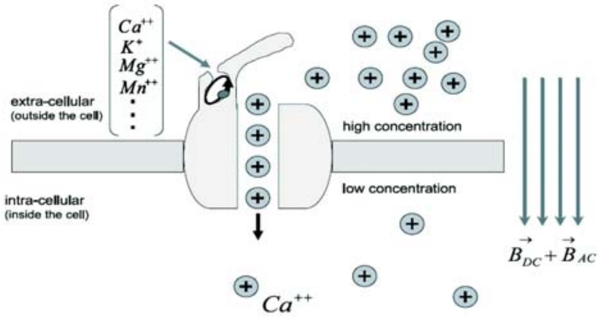


Fig. 2 Bound-ion polarization caused by a combination of static and time-varying magnetic field

The first version (1991) emanates from following assumptions:

1. The biosystem reaction onto the applied exogenous field is based on presence of Ca^{2+} ion bound to a specific protein location, enzymatic activity of which, is under control of such ion.
2. The bound ion is considered as isotropic, charged oscillator.
3. Oscillations are excited by thermal changes in biosystem.
4. Ca^{2+} ion's potential path is of spherical symmetry.
5. Degenerated vibrational level of the ion with its characteristic angular frequency ω is split into three sublevels (Zeeman effect)
6. Oscillations of magnetic sublevels ω_{E1} and ω_{E2} are coherent – amplitude and phase ratio is independent of time. The coherency allows interference interactions between the Ca^{2+} oscillator sublevels, which causes the biological response above the thermal noise level.
7. Oscillations are present only while the Ca^{2+} ion is bound to the specific protein location.

Based on these suggestions, Lednev in his works developed a model that could help with deduction of probability of specific ion transition to an energetically lower level, caused by application of external non-stationary magnetic field. This model was criticized by Adair [8] and defended by Engstrom [9]. The objections made by Adair were due to the spherical symmetry to enable the interference, and the lifetime of the excited ion state that must be of the order of Larmor precession; otherwise, the precession movement would be interrupted and the biological effect could not occur. Lednev addressed these objections by generalizing his theory to the case of continuous excitation and created a new model in 1996 [10], which led to some changes in assumptions from the previous model:

1. The ion displacement to the binding location on the protein is considered as the beginning of the excitation.
2. Magnetic fields (static, time-varying, combined) invoke precession of vibrational axes of Ca^{2+} oscillator depending on the orientation of the magnetic field. In certain combination of relaxing time of Ca^{2+} oscillator and applied field parameters, there is a possibility to achieve a considerable change of polarization level in a plane perpendicular to the applied magnetic field.
3. It is assumed that the probability of structure modification in Ca^{2+} bound location depends on time-average level of polarization of Ca^{2+} oscillator.

The resulting probability of energy density is given by:

$$p = \frac{\Psi x^2 - \Psi y^2}{\Psi x^2 + \Psi y^2}, \tag{1}$$

where Ψ represents the electromagnetic field intensity in the direction of the given axis.

This model was proposed on the basis of the experimental findings [8], where the magnetic flux density was lower than it was in previous cases. Spin movement of a particle placed into the exogenous magnetic field is described by Larmor precession ω_L:

$$\omega_L = \gamma B_0, \tag{2}$$

where γ is the value of gyromagnetic ratio of a nucleus and B_0 is the magnetic flux density of the static magnetic (DC) field. Even in this case, the combination of time-varying and static magnetic field causes frequency modulation of the resulting electromagnetic field, intensity of which is:

$$\Psi = \Psi_0 \Sigma J_n\left(\gamma\frac{B_1}{f}\right)e^{i(\omega_L + 2\pi nf)t}, \tag{3}$$

where J_n represents Bessel function of n-th order, f is frequency, B_1 represents magnetic flux density of time-varying (AC) magnetic field, which causes the modulation and t is the time. Statistically significant bio-effect was found during the application of fields in the μT range. Nowadays, this model is one of the most important subjects in many experimental investigations and its further development is highly probable.

3. Exposure system

For the purpose of our experiments, the incubator described in [11] was used. To obtain more exact results, it was decided that inhomogeneity should not exceed 5% of the applied LF EMF. That's why a new exposure coil system has to be designed. The new system proposal is based on numerical modeling. The model consists of three air coils. The role of the upper and lower side coils is to generate the LF EMF field of desired strength (mean value of 2.4 mT) and to complement the centre coil which further enhances the homogeneity of the generated EMF (max. 5 % variation within the exposed Petri dishes).

Number of turns for each of the side coils was determined to 105, wound in 14 layers. The stabilization coil was designed with 56 turns, wound in 4 layers. The results of numerical simulations, presented in Fig. 3 showed that the desired homogeneity in the exposure area was achieved at the 1A of excitation current of a sinusoidal shape. Minimum value of magnetic flux density in exposure area was 2.36 mT and maximum 2.47 mT.

As it is demonstrated in Fig. 3, the proposed coil system is of uncommon shape, so the atypical support system for its creation has to be designed.

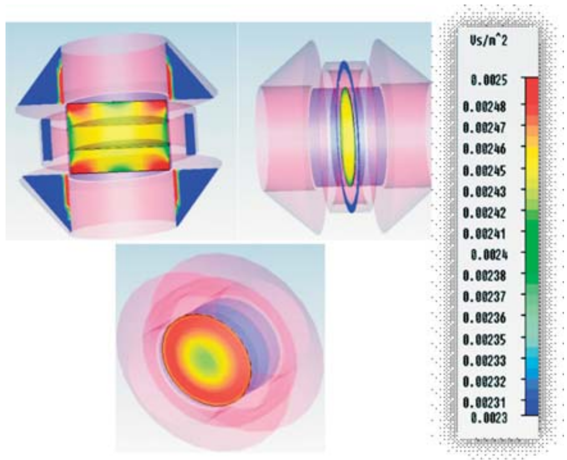


Fig. 3 Proposed model of exposure coil system–simulation results of B field; color range from 2.3 to 2.5 mT

This task was solved in cooperation with the Faculty of Mechanical Engineering, University of Zilina. Model of the coil strut was designed in the Autodesk Inventor software, sent to a 3D printer and printed. The comparison of the model and output from the 3D printer can be seen in Fig. 4.

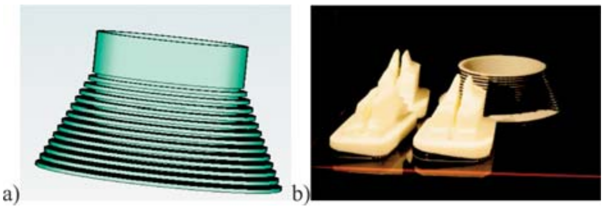


Fig. 4 a) Model of the coil strut b) Output from the 3D printer – the strut is in the background of the picture

The final creation of the proposed coil system was made manually using about 100 m of enameled copper winding wire with the diameter of 1.8 mm. The three coils were formed separately and connected together in series as shown in Fig. 5.

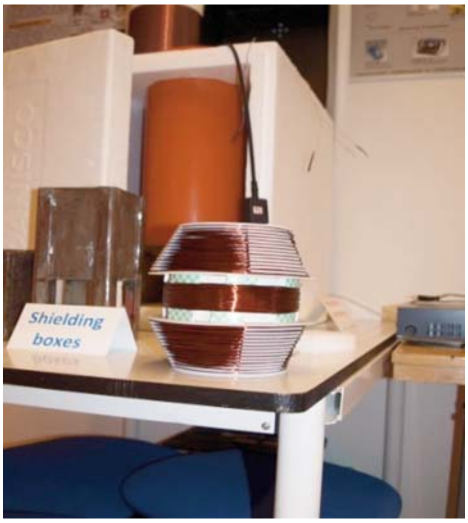


Fig. 5 The finalized exposure coil system in front of the incubator

Inductance of the whole coil system is $L = 5.346$ mH, measured by an LC-meter. The magnetic flux density(B)-field produced by the coil system was measured by flux-gate sensor at 0.1 A sinusoidal current and was recalculated to 1 A (because of linear dependence). The values from 2.37 to 2.49 mT confirmed that the constructed coil system is adequate to the simulation proposal. Due to the similar B with the coil used in [11] (B was 2.0 – 2.3 mT) the results from experiments presented herein could be compared with the results from previous experiments made by the authors.

4. Experimental work

As it has been mentioned in previous section of this article, this experiment is focused on the behavior of saccharomyces cerevisiae affected by exogenous time-varying LF EMF produced by the newly designed coil system. At the beginning of the experiment – three petri dishes (exposed samples) are placed to the proper position in the coil and three are arranged into the magnetic shielding box and inserted to the incubator, which is principally shown in Fig. 6.

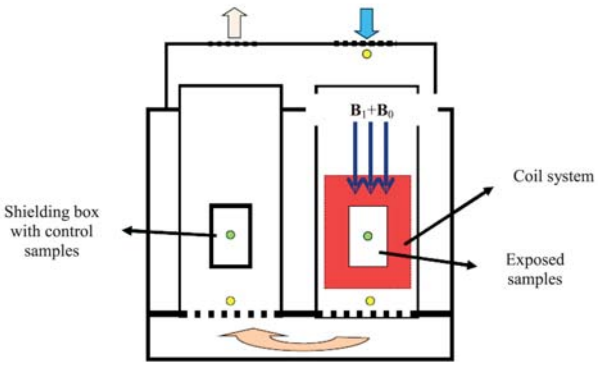


Fig. 6 Principal scheme of the incubator

Temperature of both exposed and unexposed samples is monitored by thermistors during the experiment. Frequency of applied field is 1.6 kHz at the current of sinusoidal shape with amplitude 1.01 A. The mentioned frequency was chosen to compare the results from [12], where statistically significant bio-effect occurred at this frequency.

Once the exposure system parameters are set up, the exposure has started. Time of the exposure is 66 hours according to [12, 13]. After the exposure, the petri dishes are scanned and evaluated using the application Petri scanner – software for counting of yeast colonies and growth (surface) area evaluation.

The first experiment was built-up according to the IPR theory [10], so the B vector of AC field produced by the coil system is parallel to the B vector of DC field represented by the Earth's magnetic field (measured as 47 μT). The second experiment was focused on the perpendicular combination of the applied electromagnetic fields. The whole incubator was placed on its side to the vertical plane, which is demonstrated in Fig. 7. Then another 66 hour lasting exposure started under the same conditions as in the previous case.

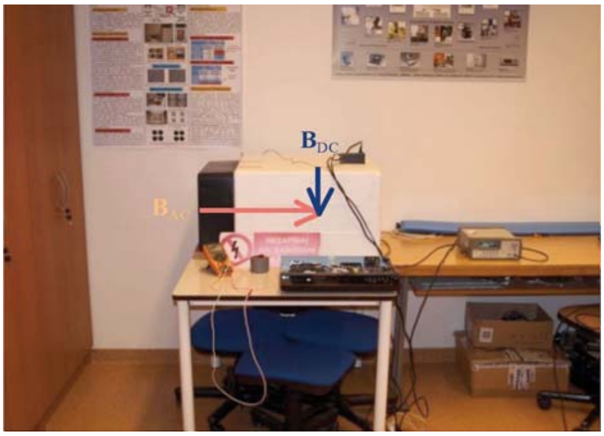


Fig. 7 Perpendicular combination of the applied EMFs

Comparison of results portrays an interesting picture, which is shown in Fig. 8. The ratio in the figures is given by the growth area ratio between the cell colonies of exposed and unexposed samples.

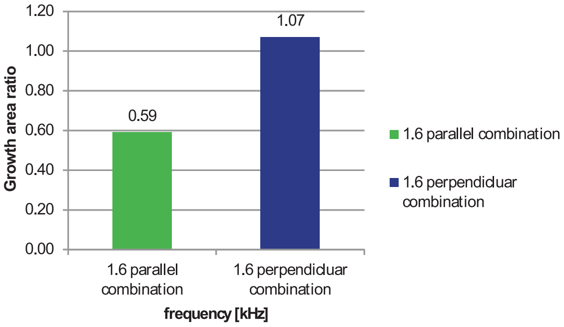
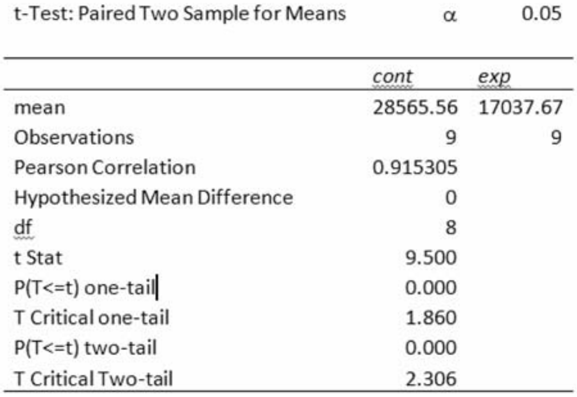


Fig. 8 Comparison between the influence of parallel and perpendicular combination of LF EMFs

For a better reliability each of the experiments is repeated under the same exposure conditions. According to Student's t-test, the obtained results are statistically comparative to the previous experiments, as it is demonstrated in Table 1.

Example of Student's t-test for first batch of experiments Table 1



t-Test: Paired Two Sample for Means α 0.05

	cont	exp
mean	28565.56	17037.67
Observations	9	9
Pearson Correlation	0.915305	
Hypothesized Mean Difference	0	
df	8	
t Stat	9.500	
P(T<=t) one-tail	0.000	
T Critical one-tail	1.860	
P(T<=t) two-tail	0.000	
T Critical Two-tail	2.306	

5. Conclusion

The presented article summarizes recent theories about the LF EMF influence to the biological structures. As a confirmation of previous experiments aligned with the presented theories, the new coil system was designed to achieve better homogeneity of the applied LF EMF. The new exposure system was used for the experiments focused on the biological effect of parallel and perpendicular combination of applied time-varying and static magnetic field.

The results from the presented experiments gave ground for the following conclusions. First, a biological effect was clearly observed at given frequency that confirms findings from [12, 13].

Second and probably foremost, the inhibitive proliferative response of the biological samples was observed during the application of parallel combination of AC and DC electromagnetic field, while the perpendicular combination showed no effect at given frequency. This finding could support mainly the Lednev's IPR theory about the influence of superposed AC and DC magnetic fields to the bound ions on the cell membrane surface. On the other hand, the anti-proliferative effect could be also explained by the Frohlich's theory. For further work it would be interesting to investigate the influence in wide frequency range of applied electromagnetic field. Especially the sub-harmonic frequencies of the base frequency presented in this work could provide considerable results especially verification of the IPR model predictions seems to be very interesting.

From afore mentioned it seems that both of the compared mechanisms could be used for explanation of the observed yeast behavior. But there are still some uncertainties, so the authors assume that the correct interpretation of the presented results could be achieved by synthesis of the described models. In advance, the presented findings give a reason to believe that such behavior could be exploited with advantage in therapeutic application targeting specific cell/tissue types.

References

[1] CONE Jr., C. D.: Unified Theory on the Basic Mechanism of Normal Mitotic Control and Oncogenesis. *J. of Theoretical Biology,* vol. 30, pp. 151-181, 1971.

[2] BARABAS, J., RADIL, R.: Investigation of Cellular Transmembrane Potential Variations via Exogenous Low Frequency Electromagnetic Fields. *Clinician and Technology J.,* vol. 40, pp. 41-45, 2010.

[3] WEBB, S. J., STONEHAM, M. E., FROHLICH, H.: Evidence for Non-thermal Excitation of Energy Levels in Active Biological Systems, *Physics Letters A,* vol. 63, pp. 407, 1977.

[4] POKORNY, J.: Cancer Physics: Diagnostics Based on Damped Cellular Vibrations in Microtubules. *Eur Biophys J.*, vol. 40, pp. 747-759, 2011.

[5] CIFRA M., POKORNY, J., JELINEK, F., HASEK, J.: Measurement of Yeast Cell Electrical Oscillations around 1 kHz. *PIERS Proceedings.* Cambridge, USA, pp. 780-784, 2008.

[6] LEDNEV, V. V.: Possible Mechanism for the Influence of Weak Magnetic Fields on Biological Systems. *Bioelectromagnetics.* vol. 12, No. 2, pp. 71-75, 1991.

[7] LIBOFF, A. R.: Cyclotron Resonance in Membrane Transport. Interactions between Electromagnetic Fields and Cells. *Plenum Press,* pp. 287-296, 1985.

[8] ADAIR, R. K.: Criticism of Lednev's Mechanism for the Influence of Weak Magnetic Fields on Biological Systems. *Bioelectromagnetics,* vol. 13, pp. 231-235, 1992.

[9] ENGSTROM, S.: Dynamic Properties of Lednev's Parametric Resonance Mechanism. *Bioelectromagnetics,* pp. 58-70, 1996

[10] LEDNEV, V. V.: Article in Russian, *Biofizika,* vol. 41, p. 815, 1996.

[11] BARABAS, J.: *Proliferative Response of Two Different Saccharomyces Cerevisiae Strands to Extremely Low Frequency Electromagnetic Fields.* POSTER 2011: 15th Intern. Student Conference on Electrical Engineering, 2011, Prague, ISBN/ ISSN978-80-01-04806-1.

[12] CAP, I., BARABAS, J.: Low Frequency Electromagnetic Fields: Friend or Foe? *Communications - Scientific Letters of the University of Zilina,* vol. 1, pp. 13-17, 2011.

[13] BARABAS, J., CAP, I.: S. *Cerevisiae Growth Regulation via Exogenous Low Frequency Electromagnetic Fields.* Proceedings Trends in Biomedical Engineering, pp. 141-144, 2011.

Peter Hockicko – Peter Bury – Francisco Munoz *

ANALYSIS OF DIELECTRIC PROPERTIES OF LITHIUM PHOSPHATE GLASSES

Dielectric relaxation spectra of lithium phosphate glasses are used to analyze the change of real and imaginary part of complex permittivity including loss tangent of ion conductive glasses. The experimental results getting from a.c. and d.c. electric measurements on Li phosphate glasses are analyzed, discussed and compared. Frequency dependence of the permittivity spectra is found to be consistent with the analysis of d.c. electrical conductivity. The universal a.c. conductivity law (universal dynamic response) for glass ionic conductor was confirmed. Arrhenius graphs constructed from a.c. and d.c. electrical measurements including loss tangent are used to characterize the relaxation processes and transport mechanisms connected with the mobility of conductive ions.

Keywords: *Li phosphate glasses, conductivity and dielectric relaxation spectroscopy, relaxation and transport processes.*

1. Introduction

Solid electrolytes present numerous potential advantages compared to liquid electrolytes, including absence of liquid containment and leakage problems, ability to operate with more reactive electrodes over a wider temperature range, and the possibility of miniaturization using thin-film-processing techniques. They are an important part of electrochemical devices such as electrochronic displays and sensors, solid-state batteries, electrochemical pumps, capacitors, time switches, etc. [1, 2].

General requirements for practical solid electrolytes are high ionic conductivity, stability with respect to thermal and electrochemical decomposition, suitable mechanical properties, ease of fabrication, and reasonable cost.

Lithium-ion conductive glasses have been widely studied due to their potential application as solid-state amorphous electrolytes in secondary batteries [1]. It was found that amorphous character of this material and the presence of nitrogen increase conductivity compared to that of the crystalline compound. The increase in conductivity is supposed to be related to the formation of P–N bonds which replace P–O ones leading to a more reticulated anionic network [2].

The optical and electrochemical properties of lithium phosphorous oxynitride (LiPON) thin films have been studied and successfully applied in lithium microbatteries [3]. LiPON exhibits a single, Li^+-ion conducting phase with an average conductivity of $(2.3 \pm 0.7) \times 10^{-6}$ S/cm at 25 °C and an average activation energy of Ea = (0.55 ± 0.02) eV [3].

In this contribution we present results obtained by electric investigation of lithium phosphate glasses with different Li/P ratios with compositions $xLi_2O \cdot (1-x)\, P_2O_5$ ($x = 55, 57.5$ mol %) and oxynitride phosphate glass prepared through the thermal treatment under ammonia atmosphere of the sample with initial composition ($x = 55$ mol %).

2. Theoretical principles

Dielectric relaxation spectroscopy is a powerful technique for the study of ion transport processes in fast ion conductive glasses. In general, using this spectroscopy, we can study molecular and ionic dynamics of charge carriers and dipoles [4], we can observe different systems from gases to solid substance.

The central quantities in this context are the dielectric displacement $\boldsymbol{D}$ and the electric field $\boldsymbol{E}$, with the dielectric permittivity ε_r using the SI system being defined by

$$\boldsymbol{D} = \varepsilon_r \varepsilon_0 \boldsymbol{E} = \varepsilon_0 \boldsymbol{E} + \chi \varepsilon_0 \boldsymbol{E} = \varepsilon_0 \boldsymbol{E} + \boldsymbol{P}, \qquad (1)$$

where $\varepsilon_0 = 8.854 \times 10^{-12}$ AsV^{-1}m^{-1} is the permittivity of vacuum, $\chi = \varepsilon_r - 1$ is the susceptibility and $\boldsymbol{P}$ is the polarization. It has been observed that the determining the imaginary part of the complex dielectric constant (ε'') compared to the real part (ε') play a crucial role in the study of fundamental properties of investigated materials [5], where $\varepsilon'(\omega)$ and $\varepsilon''(\omega)$ characterize the *refractive* and *absorptive* properties of the material, respectively. Using the functions

---

* **Peter Hockicko**[1], **Peter Bury**[1], **Francisco Munoz**[2]
[1] Department of Physics, Faculty of Electrical Engineering, University of Zilina, Slovakia, E-mail: hockicko@fyzika.uniza.sk
[2] Instituto de Ceramica y Vidrio (CSIC), Madrid, Spain

$$\varepsilon^*(\omega) = \varepsilon'(\omega) - j\varepsilon''(\omega), \tag{2}$$

is common in practice for presenting dielectric relaxation data for materials in which the motion of charges is dominated, where the position $\omega_0 = 1/\tau_0$ of a peak in the loss component $\varepsilon''(\omega)$ outlines a characteristic time scale of the orientational molecular mobility. For ionically conducting glasses no equivalent feature in $\varepsilon''(\omega)$ is found which immediately sets a time scale for ionic mobility [6]. In the ion-conducting materials the polarization and conduction are, therefore, integrated into a single, continuous process. The *loss tangent* defined by relation

$$\tan\delta(\omega) = \varepsilon''(\omega)/\varepsilon'(\omega), \tag{3}$$

can be also used for the relaxation spectroscopy that reflects the basic features of the relaxation processes of mobile ions. This quantity is related to the attenuation constant (or absorption coefficient) of an electromagnetic wave propagating in a material. The activation energies of the relaxation processes can be estimated from the plots of log f vs. $1/T_{max}$ were T_{max} can be found using the isochronal peaks of tan $\delta(\omega,T)$. These plots are straight lines in accordance with Arrhenius equation

$$f = f_0 \exp\left(-E_a^{\mathrm{tg}\delta(\omega)}/(k_B T_{max})\right), \tag{4}$$

where f is the frequency of the applied electrical field, f_0 is the pre-exponential factor. $E_a^{\mathrm{tg}\delta}$ is the activation energy associated to the dielectric loss, k_B is the Boltzmann constant.

The electrical conductivity σ of many solids including glasses, polymers and crystals has been shown to consist of a frequency independent and a strongly frequency dependent component [7]. Experimental data in a limited frequency region revealed that the overall frequency dependence of σ or the so called "universal dynamic response" (UDR) of ionic conductivity can be approximated by the following simple relation

$$\sigma(\omega) = \sigma_{dc} + A\omega^s \tag{5}$$

$$\sigma_{dc} = \sigma_0 \exp\left(-E_a^{dc}/(k_B T)\right), \tag{6}$$

where ω $(=2\pi f)$ is the angular frequency of measurement. For a typical ionic conductor containing moderate to high concentration of mobile ions, $s \approx 0.5$, and both σ_{dc} and A follow Arrhenius type strong temperature dependencies. In Eq. 6 σ_0 is the pre-exponential factor, E_a^{dc} is the activation energy of the ion transport through hopping processes which can be determined from the dc measurements.

3. Experimental details

The lithium phosphate glasses with composition $x\mathrm{Li_2O} \cdot (1 - x)\,\mathrm{P_2O_5}$ ($x = 55$, 57.5 mol %) were obtained by conventional melt-quenching technique. Stoichiometric amounts of reagent grade raw materials Li_2CO_3 (Aldrich, 99%) and $(NH_4)2HPO_4$ (Merck, 99%) were weighed and mixed. The batches were calcined in porcelain crucibles held in an electric furnace at a temperature up to 450 °C, and then melted in a gas furnace (propane/air) during 1 h at temperatures ranging from 800 °C to 1000 °C depending on composition. The compositions of the glasses are given by their molecular formula as $Li_{1.22}PO_{3.11}$ and $Li_{1.35}PO_{3.18}$ for Li_2O contents of 55 and 57.5 mol%, respectively [2].

The oxynitride lithium phosphate glasses were obtained through ammonolysis of base glasses in an Al_2O_3 gas-tight tube furnace at temperatures ranging from 600 °C to 750 °C and treatment times of 3 h. Base glasses were placed in graphite moulds acting as individual "crucibles" of 2 cm in diameter and 5 mm deep. The furnace was heated up to the treatment temperature at a constant heating rate of 10 K $\mathrm{min^{-1}}$ under N_2 flow. In the present work three glass compositions of prepared set of samples [2] are studied by dielectric relaxation and conductivity spectroscopy: $Li_{1.22}PO_{3.11}$ (B), $Li_{1.35}PO_{3.18}$ (C) and $Li_{1.22}PO_{2.80}N_{0.21}$ (BN).

The real and imaginary parts of the complex dielectric permittivities (ε', ε'') and loss tangent (tan δ) were calculated from the values obtained from the measurements in frequency ranges 50 Hz – 1 MHz by FLUKE PM 6306 impedance analyzer.

The sample C was chosen as representative of investigated glass system to present the experimental investigation by dielectric relaxation spectroscopy. The temperature dependence of both real and imaginary part of relative dielectric function $\varepsilon_r^*(\omega,T) = \varepsilon_r'(\omega,T) - j\varepsilon_r''(\omega,T)$, for all sample were investigated.

4. Results

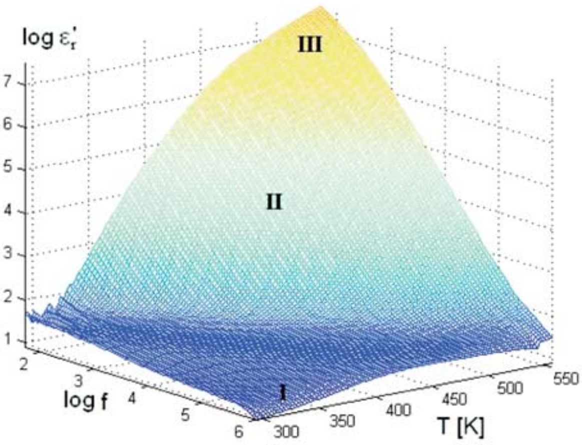


Fig. 1 Temperature and frequency dependence of real part of relative permittivity in 3D form

Three-dimensional (3D) plots of measured both real (ε_r') and imaginary part (ε_r'') of complex relative permittivity ε_r^* as a function of temperature and frequency are shown in 3D presentation (in Figs. 1 and 2), respectively. From these pictures can be iden-

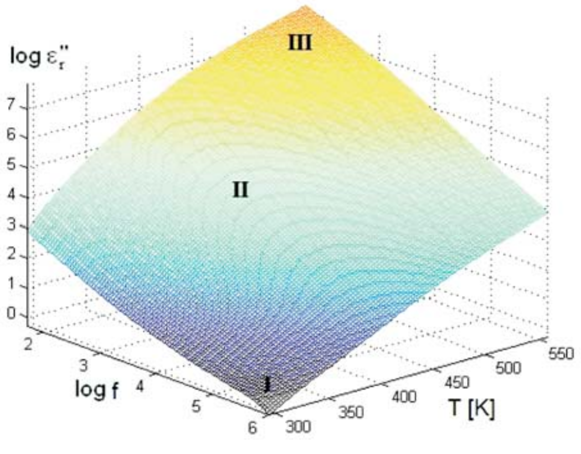


Fig. 2 Temperature and frequency dependence of imaginary part of relative permittivity in 3D form

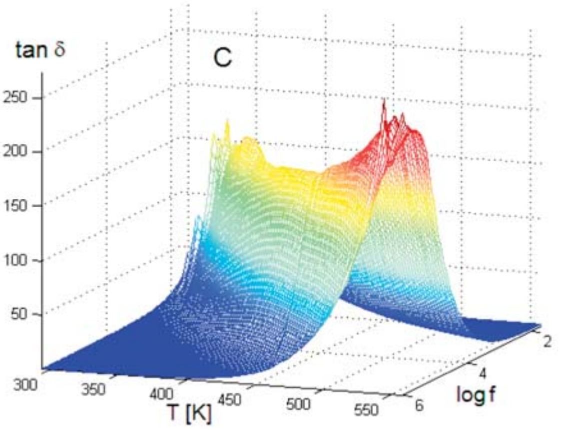


Fig. 3 Temperature and frequency dependence of loss tangent in 3D form

tify three district processes marked by I, II and III parts defined as follows:

(I) Low temperature processes,
(II) Mid temperature processes,
(III) High temperature processes.

(I) This process started over low temperature: 300 K and it can be appreciated by looking at ε_r' at higher frequencies and lower temperatures.
(II) That process extends over mid temperature region (about 350 – 450 K) and over low to moderate frequencies. These processes were observed for ε_r' and ε_r'', respectively.
(III) High temperature process is significant at higher temperatures ($T > 500$ K). Its amplitude increases rapidly with frequency for both real and imaginary part.

The better resolution is evident from Fig. 3, where the temperature and frequency dependences of loss tangent tanδ (f, T) in 3D presentation can be seen. It is seen that in the frequency range the only one broad peak was observed with maximum position (tanδ (T_{max})) shifted to higher temperatures with increasing frequency. Some differences can be seen in the temperature range at high and low temperature (two peaks are visible).

It is an established fact that the condition for observing a maximum in the dielectric losses is given by the relation

$$\omega_0 \tau_0 = 1, \qquad (7)$$

where $\omega_0 = 2\pi f_{max}$ (f_{max} is the frequency corresponding to the maximum of tanδ (f)) and τ_0 is the relaxation time. A maximum can be observed when hopping or jumping frequency of ions becomes approximately equal to the frequency of applied field. The shift of maxima in temperature dependencies of the dielectric loss tangent towards to high-frequency side could be due to high values of hopping probability per unit time in the sample.

Typical dependences of real part of the complex permittivity and conductivity at low temperature region ($T = 295$ K) are shown in Fig. 4.

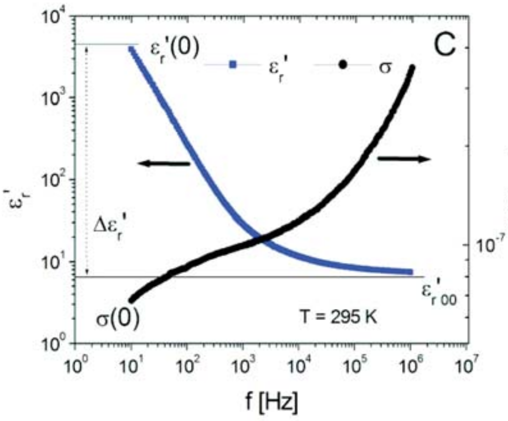


Fig. 4 The measured real part of complex relative permittivity and conductivity at low temperature region (T = 295 K)

Here we can see a considerable dependence of both quantities on the frequency f of the applied electric field. Starting at low frequencies, we found a large increase with decreasing frequency of the dielectric constant $\varepsilon_r'(f)$. It is supposed [8] that in the limit of a dc field, it attains a plateau value. This phenomenon results from the presence of so-called blocking (metallic) electrodes, which do not permit transfer of mobile ions into the external measuring circuit. At low frequencies, it was observed a short plateau in the conductivity $\sigma(0)$ [8]. This so-called dc conductivity represents the long-range diffusion of ions as they hop from site to site through the matrix. We observed a shoulder in the dielectric constant at low frequency range that suggests an incipient polarization occurring in this same frequency range. This polarization is understood to be associated with the growth and shrinkage of a dipole that occurs during the hop. This growth and shrinkage is in many respects similar to the rotational relaxation that occurs for a collection of noninteracting permanent dipoles of dipole moment p.

On right side of Fig. 4 we can see a dramatic increase in $\sigma(f)$ and a leveling off in $\varepsilon_r'(f)$. This leveling off of the dielectric constant can occur because, in addition to the mean-squared displacement of mobile ions, matrix contains atoms that have become elastically polarized under the influence of the applied field. The conductivity increases with increasing frequency in a roughly power-law mode. In the high frequency region the data appear to approach a linear dependence of frequency.

It was established [9] that at low frequencies, the dc conductivity implies that the mean-squared displacement is linear and this linear time dependence reflects the random diffusion of the ions which migrate from site to site through the disorded matrix. The time dependent mean-squared displacement was found in the classical random walk model of diffusion and is an identification of uncorrelated motions. At higher frequencies the mean-squared displacement becomes nonlinear and indicates that an ion motion is nonrandom or temporally correlated.

The frequency dependence of the real part $\varepsilon_r'(f)$ of the complex relative permittivity for sample C measured at 295 K approaches a constant value, $\varepsilon_{r\infty}'(f)$ at higher frequency, which probably results from rapid polarization processes occurring in the glasses [10]. With decreasing frequency, $\varepsilon_r'(f)$ increases significantly due to the electrode polarization arising from space charge accumulation at the glass-electrode interface and reaches a low-frequency plateau, ε_{rs}' (static permittivity). This the low-frequency plateau (in mHz frequency range) usually associated with the polarization effects of the mobile ions. It was proposed [10, 11] that the strength of the ionic relaxation determined as $\Delta\varepsilon_r' = \varepsilon_{rs}' - \varepsilon_{r\infty}'$ is the magnitude of the permittivity change due to ionic relaxation.

Several other alternative interpretations of the frequency dependences can be found [12]. For example, the electric modulus M^* is formally defined as the inverse of the complex permittivity

$$M^*(\omega) = 1/\varepsilon_r^*(\omega) = \varepsilon_r'/(\varepsilon_r'^2 + \varepsilon_r''^2) + j\varepsilon_r''/(\varepsilon_r'^2 + \varepsilon_r''^2). \quad (8)$$

Features of this equation are displayed in Fig. 5, where both the real and imaginary parts of the modulus are plotted against frequency for our investigated lithium phosphate glass at room temperature. As we can see the real part of the modulus exhibits a steplike increase with increasing frequency and the imaginary part displays a peaked function.

Complex impedance ($Z^* = Z' - jZ''$) plots for sample C at the temperatures $T = 295$ K is illustrated in Fig. 6. The Nyquist plot consists of a depressed semicircle and a spike at the lower frequencies suggesting electrode-electrolyte polarization. This plot was fitted by using an equivalent circuit model of one RC couple in parallel combination (representing the bulk resistance R and geometric capacitance C) in series with one constant phase element (CPE) (representing polarization at the electrode - electrolyte interface) (in the insert of Fig. 6a).

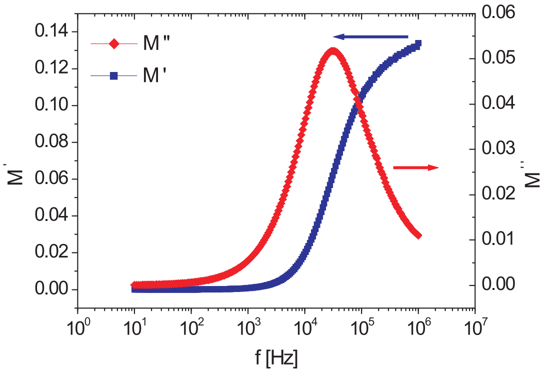


Fig. 5 The real and imaginary parts of the electric modulus for sample C at room temperature

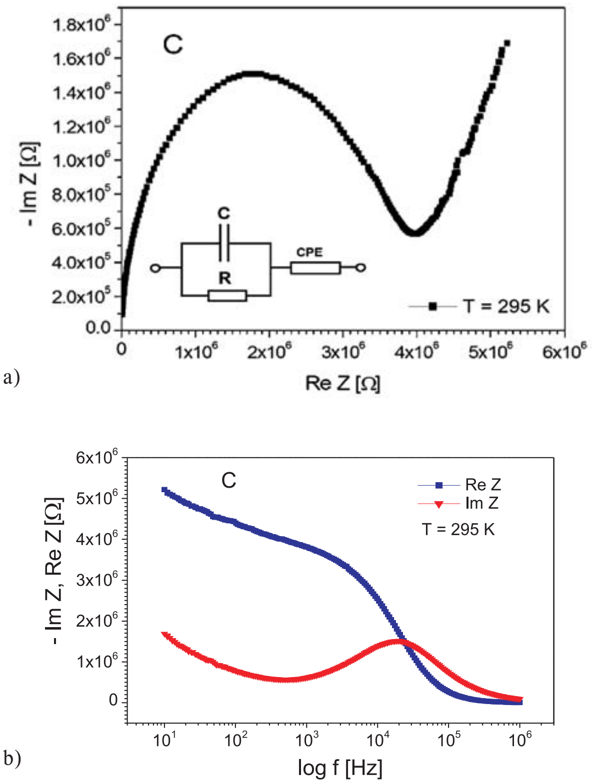


Fig. 6 The impedance plots at temperature T = 295 K for sample C, suitable equivalent circuit is shown inside (a). Plots of real (Re Z) and imaginary (Im Z) part of impedance as function of frequency (b)

The σ_{dc} at each temperature was calculated using the relation $\sigma_{dc} = d/(RA)$, where d is the thickness of the bulk glass piece and A is the area of cross section of the glass piece. The d.c. conductivity at 330 K is 2.6×10^{-7} S/cm which is an increase of one order as compared to $LiPO_3$ glass [13]. The value of d.c. conductivity calculated at the room temperature (303 K) for sample C was about 4.3×10^{-8} S/cm.

A representative Arrhenius plot for sample C is shown in Fig. 7. The temperature dependencies of the d.c. conductivity indicate two transport mechanisms with activation energies $E_{a1}^{dc} = 0.55$ eV and $E_{a2}^{dc} = 0.59$ eV corresponding to the higher and lower temperatures, respectively. (The activation energy of ionic conductivity determined from another measurements for the LiPON thin film was 0.57 eV [14, 15].)

The position of the break on the dc conductivity curves (Fig. 7) can imply association process [16].

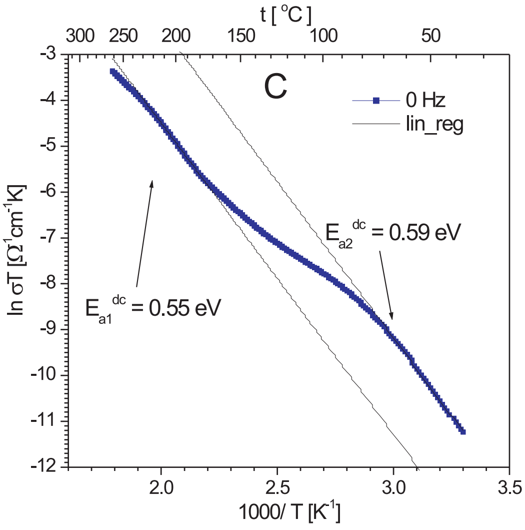


Fig. 7 Temperature dependence of d.c. conductivity

Plotting the frequency of a.c. electric field in logarithmic scale as a function of corresponding inverse temperature of maximum of loss tangent angle (Arrhenius plot), we can then obtain a value of $E_{a1}^{tg\delta} = 0.52$ eV and $E_{a2}^{tg\delta} = 0.62$ eV for activation energy of the dielectric losses.

The set of the frequency dependencies of a.c. conductivity measured at various temperatures (conductivity spectra) is illustrated for the glass sample C in Fig. 8. The obtained a.c. conductivity measurements correspond to the complete conductivity spectra of glassy samples. However, because of limited frequency range only two regimes (II and III) of the whole conductivity spectra [17] represented by a different slope of individual curves could be recognized, moreover the regime II only at low temperatures. The low frequency part in regime III and the dispersive regime II are due to the hopping motion of the mobile ions and can be obviously explained in terms of a jump relaxation model considering several kinds of sites. The slope of the brakes in the ac conductivity spectra indicates another transport hopping process. The jump from one value of the ac conductivity to another was registered at higher temperatures and lower frequencies, characterizing the transition between two kinds of hopping processes. The activation energies calculated from the Arrhenius plots of d.c. conductivity and dielectric measurements for all investigated glasses are summarized in Table 1.

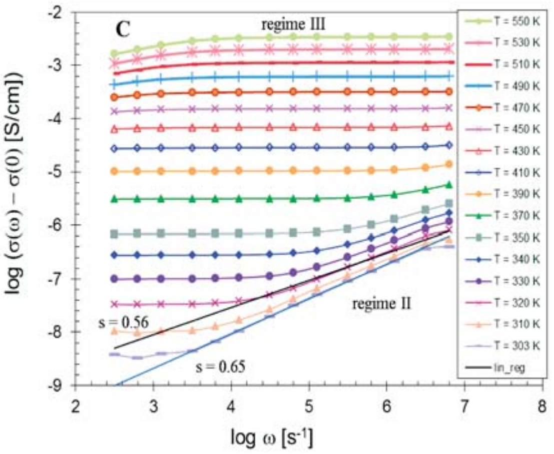


Fig. 8 Representative results obtained from a.c. electrical conductivity measurements for the glass sample $Li_{1.35}PO_{3.18}$ (C)

Summarized d.c. conductivities and activation energies calculated from Arrhenius plots ($\Delta E_a = \pm 0.01$ eV). Table 1.

Glass sample	d.c.			a.c.	
	σ_{dc} [S/cm] (T=330K)	E_{a1}^{dc} [eV]	E_{a2}^{dc} [eV]	$E_{a1}^{tg\,\delta}$ [eV]	$E_{a2}^{tg\,\delta}$ [eV]
B	1.5×10^{-7}	0.56	0.61	0.51	0.67
BN	2.2×10^{-7}	0.54	0.60	0.52	0.64
C	2.6×10^{-7}	0.55	0.59	0.52	0.62

Comparing the activation energies obtained from the conductivity and dielectric measurements the relation between d.c. conductivity and activation energy of dominant peak is evident, that means the decrease of activation energy with increasing conductivity.

5. Discussion

The both dielectric and conductivity spectra of investigated LiPON glasses suggest two main relaxation and/or transport mechanisms that are also in good agreement with previous investigation by acoustic spectroscopy [18]. The FTIR spectra of the $LiPO_3$ glasses of the system $50Li_2O \cdot 50P_2O_5$ [1] show characteristic peaks corresponding to different vibration modes of the PO_4 tetrahedra ($\nu(PO_4)^{3-}{}_{sym}$ and $\nu(PO_4)^{3-}{}_{asym}$) as well as those of the P-O-P bonds (ν(P-O-P)$_{sym}$, ν(P-O-P)$_{asym}$ and δ(P-O-P)) and the PO_{2-} groups

($\nu(PO_2)^-_{sym}$ and $\nu(PO_2)^-_{asym}$). Using Nuclear Magnetic Resonance it has been found [2] that the oxide glasses marked B, C (nitrogen-free) are composed of the PO_4 groups of Q^2 – type (metaphosphate composition with two bridging oxygen atoms bonded to neighboring phosphorous atoms, cyclic anions of exact $(PO_3^-)_n$) and Q^1 structural units (pyrophosphate composition with one bridging oxygen atom bonded to neighboring phosphorous atom, $P_2O_7^{4-}$ groups). Furthermore, the oxynitride glass marked BN contains also other groups: PO_3N and PO_2N_2, which result from nitrogen/oxygen substitution [19]. These groups increase the cross-linking density giving rise to a notable modification of the glass properties. A higher cross-linking density introduced by the nitrogen atoms should facilitate the lithium transfer between phosphate chains and the higher amount of non-bridging oxygen (NBO) generated by nitridation could increase the number of hopping positions available for Li^+ ions, creating conduction paths with lower activation energy.

The two main activated processes found from Arrhenius plots of d.c. conductivity and dielectric loss angle tangent maxima might be related to different structural units: Lithium ions bonded to non-bridging oxygen atoms present in Q^2 metaphosphate units and Q^1 pyrophosphate. From comparison with the IR spectra of LiPON glasses we can draw the assumption that all samples contain dimeric diphosphate oxoanions $P_2O_7^{4-}$ and $(PO_3^-)_n$ phosphate chain groups as the dominant structure elements of the glass. The local electric field around the dimeric diphosphate oxoanions $P_2O_7^{4-}$ is stronger than the local electric field around the phosphate chain groups $(PO_3^-)_n$ because the triphosphate anions have a smaller negative electric charge on the non-bridging oxygen as diphosphate anion atoms. This causes the electrostatic interactions between the mobile Li^+ ions and the diphosphate anions to be stronger than those between the Li^+ ions and the phosphate chain groups. Thus we can suppose that the relaxation processes with bigger activation energies can be connected with diphosphate anions and the processes with smaller energies can be connected with the low condensed triphosphate anions as well as some other polymeric structural units.

The effect of nitrogen on conductivity can be explained by the higher cross-linking density introduced by the nitrogen atoms, which should facilitate the lithium transfer between the phosphate chains. The higher amount of the non-bridging oxygen atoms generated by the nitridation could increase the number of hopping positions available for the Li^+ ions thus creating conduction paths with lower activation energy and the rise of the electrical conductivity as a consequence of this [2].

However, some differences can be caused by the different relaxation mechanisms connected with the ion hopping transport in a.c. and d.c. electric field [20].

6. Conclusion

The experimental investigation of the lithium phosphate and oxynitride phosphate glasses proved that the electrical conductivity and dielectric relaxation spectroscopy can be useful techniques for the study of relaxation processes and transport mechanisms in the Li^+ ion conductive glasses. Two main kinds of sites responsible for the ionic hopping motion were discovered and described. It was also verified that the nitridation increases the amount of the NBO atoms and the cross-linking density which increases the electrical conductivity of the oxynitride phosphate glasses and decreases the activation energy of conductivity.

Acknowledgment
The authors would like to thank to Mr. F. Cernobila for technical assistance. This study was supported by the Slovak Grant Agency KEGA through the projects No. 035ZU-4/2012, VEGA No. 1/0853/13 and R&D operational program Centre of excellence of power electronics systems and materials for their components II. No. OPVaV-2009/2.1/02-SORO, ITMS 26220120046 funded by European regional development fund (ERDF).

References

[1] MOREAU, F., DURAN, A., MUNOZ, F.: Structure and Properties of High Li_2O-containing Aluminophosphate Glasses, *J. of the European Ceramic Society* 29 (10), pp. 1895–1902, 2009

[2] MUNOZ, F., DURAN, A., PASCUAL, L., MONTAGNE, L., REVEL, B., RODRIGUES, A. C. M.: Increased Electrical Conductivity of LiPON Glasses Produced by Ammonolysis, *Solid State Ionics* 179 (15-16), pp. 574–579, 2008

[3] YU, X., BATES, J. B., JELLISON, G. E., HART, F. X.: A Stable Thin-Film Lithium Electrolyte: Lithium Phosphorus Oxynitride, *J. Electrochem. Soc.* 144 (2), pp. 524–532, 1997

[4] GUTTEN, M., KUDELCIK, J.: Methods of Humidity Analysis in Transformer Oil, *Communications – Scientific Letters of the University of Zilina* 12 (2), pp. 49–52, 2010

[5] VENKATESH, J., SIVASUBRAMANIAN, V., SUBRAMANIAN, V., MURTHYA, V.R.K.: Far-IR Reflectance Study on B-site Disordered Ba $(Zn_{1/3}Ta_{2/3})O_3$ Dielectric Resonator, *Materials Research Bulletin* 35 (8), pp. 1325–1332, 2000

[6] RICHTER, R., WAGNER, H.: The Dielectric Modulus: Relaxation Versus Retardation, *Solid State Ionics* 105 (1–4), pp. 167–173, 1998

[7] JONSCHER, A. K.: The 'Universal' Dielectric Response, *Nature* 267, pp. 673–679, 1977

[8] SIDEBOTTOM, D. L., ZHANG, J.: Scaling of the ac Permittivity in Ion-conducting Glasses, *Phys. Rev.* B 62(9) pp. 5505-5507, 2000
[9] SIDEBOTTOM, D. L.: Colloquium: Understanding Ion Motion in Disordered Solids from Impedance Spectroscopy Scaling, *Reviews of Modern Physics* 81, pp. 999-1014, 2009
[10] SIDEBOTTOM, D. L., ROLING, B., FUNKE, K.: Ionic Conduction in Solids: Comparing Conductivity and Modulus Representations with Regard to Scaling Properties, *Phys. Rev.* B 63 024301-024301-7, 2000
[11] SIDEBOTTOM, D. L.: Influence of Cat Ion Constriction on the ac Conductivity Dispersion in Metaphosphate Glasses, *Phys. Rev.* B 61, pp. 14507-14516, 2000
[12] SIDEBOTTOM, D. L., GREEN, P. F., BROW, R. K.: Anomalous-diffusion Model of Ionic Transport in Oxide Glasses, *Phys. Rev.* B 51(5), pp. 2770-2776, 1995
[13] PEVZNER, B., HEBARD, A. F., DRESSELHAUS, M. S.: Role of Molecular Oxygen and other Impurities in the Electrical Transport and Dielectric Properties of C60 Films, *Phys. Rev.* B 55, pp.16439-16449, 1997
[14] FLEUTOT, B., PECQUENARD, B., MARTINEZ, H., LEVASSEUR, A.: Thorough Study of the Local Structure of LiPON thin Films to better Understand the Influence of a Solder-reflow Type Thermal Treatment on their Performances, *Solid State Ionics* 206, pp. 72-77, 2012
[15] HAMON, Y., DOUARD, A., SABARY, F., MARCEL, C., VINATIER, P., PECQUENARD, B., LEVASSEUR, A.: Influence of Sputtering Conditions on Ionic Conductivity of LiPON thin Films, *Solid State Ionics* 177 (3-4), pp. 257-261, 2006
[16] KNAUTH, P., TULLER, H. L.: Solid-State Ionics: Root, Status, and Future Prospects, *J. of the American Ceramic Society* 85 (7) pp.1654-1680, 2002
[17] FUNKE, K., ROLING, B., LANGE, M.: Dynamics of Mobile Ions in Crystals, Glasses and Melts, *Solid State Ionics* 105 (1-4), pp. 195-208, 1998
[18] HOCKICKO, P., MUNOZ, F., BURY, P., SIDOR, P.: Analysis of Acoustic Attenuation Spectra of LiPON Glasses, *Communications - Scientific Letters of the University of Zilina* 12 (2), pp. 38-42, 2010
[19] HOCKICKO, P., BURY, P., SIDOR, P., JURECKA, S., JAMNICKY, I.: Mathematical Models for Acoustic Spectra Simulation, *Communications - Scientific Letters of the University of Zilina* 12 (1) pp. 44-49, 2010.
[20] HOCKICKO, P., BURY, P. MUNOZ, F.: Investigation of Relaxation and Transport Processes in LiPO(N) Glasses, *J. of Non-Crystalline Solids* 363 (1), pp. 140-146, 2013.

Frantisek Mach – Pavel Kus – Pavel Karban – Ivo Dolezel *

SEPARATION OF PLASTIC PARTICLES IN ELECTROSTATIC FIELD PRODUCED BY ELECTRODES OF OPTIMIZED SHAPE

Shape optimization of electrodes for the device for electrostatic separation of triboelectrically charged plastic particles is carried out. The objective function maximizes the efficiency of separation consisting in the highest possible number of particles falling down to the prescribed bins. Electric field in the system is solved numerically, using the fully adaptive higher-order finite element method. The movement of particles in the device influenced by the Coulomb force is determined by means of an adaptive Runge-Kutta-Fehlberg method with a time varying time step. The shape optimization is carried out using a technique based on genetic algorithms. The methodology is illustrated by an example whose results are discussed.

Keywords: *Charged plastic particles, triboelectric effect, higher-order finite element method, genetic algorithms, adaptivity.*

1. Introduction

Nowadays, an intensive research is conducted worldwide, aimed at the possibilities of recycling plastic materials, because in a lot of applications it may well replace the new ones. The necessary prerequisite of this reprocessing is a high-quality separation of particular kinds of plastics preliminarily ground into small sphere-shaped particles. One of the advanced techniques of separation of these particles is based on the triboelectric effect [1–3]. It is known that when electrically non-conducting particles of two different levels come into contact with electric charge, one of them becomes more positive (or negative) with respect to another one. And when such charged particles get into electric field, their movement is then mainly affected by the charge that they carry.

More accurately, the trajectories of the charged particles moving in electric field are influenced by the Coulomb force exerted on it by this field, gravity and drag aerodynamic forces. And these trajectories predetermine their impact points, in other words, the places where they fall down. It is clear from Fig. 1 showing a typical separator of this kind.

The separator consists of two electrodes, one of them being grounded. The voltage of the other electrode is used to be on the order of tens kV. The electrodes may be covered by Teflon or another insulating material that prevents recharging of the particles in case of the direct impact with them. The mixture of charged particles is delivered by the feeder. At the bottom of the device there are several recycle bins used for accumulating of particular levels of plastics. One of the principal demands is to tune the shape of the electrodes and widths of the bins so that the particles of different levels should fall down exactly to the corresponding bin.

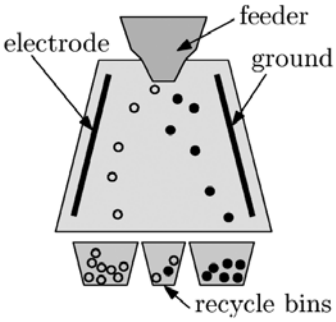


Fig. 1 Scheme of a typical separator

The problems of mapping electric field within the separator and trajectories of particles of plastics were dealt with by several authors (see, for example, [4–5]). Their approach, however, was based on low-order mapping techniques. The problem of the shape optimization of both electrodes (and also bins) was solved rather by a comparison of several different arrangements, without applying direct optimization techniques.

The paper represents an organic continuation of recent study [6] aimed at the shape optimization of the electrodes that was carried out using the conjugate gradient method. With respect to the discontinuousness of both functionals the method turned out to be rather disadvantageous for this application. Based on the results described in this paper we changed the formulation of the task in the sense of an appropriate change of both functionals and description of the shapes of the electrodes.

---

* **Frantisek Mach[1], Pavel Kus[1], Pavel Karban[1], Ivo Dolezel[2]**
[1] Faculty of Electrical Engineering, University of West Bohemia, Pilsen, Czech Republic, E-mail: fmach@kte.zcu.cz
[2] Czech Technical University, Faculty of Electrical Engineering, Prague, Czech Republic

2. Formulation of the problem

Let us start with a possible basic arrangement depicted in Fig. 2 (its principal dimensions being given in mm). The particles of plastics of charge Q and initial velocity v_0 get to the space between two electrodes, one of them being grounded. There they are deflected according to their charge and fall down into the recycle bins (presently, we neglect the Coulomb forces acting among the particles, so that their movement is driven mainly by the external electric field). The first task is to find their trajectories and evaluate the effectiveness of separation in this basic arrangement.

The second step is the optimization of the shape of the electrodes. Their basic shape is characterized by an efficiency defined as a ratio of particles of the given type that fall down to the corresponding bin versus the total number of particles of this kind. And the aim of this optimization is to increase this ratio on the highest value possible.

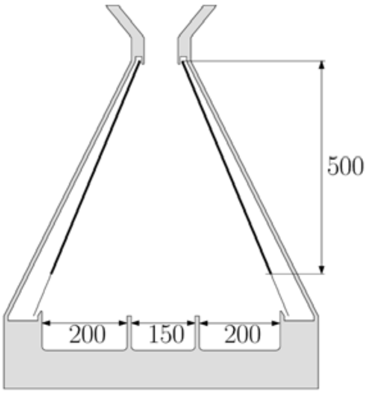


Fig. 2 Basic scheme of the separator (dimensions in mm)

3. Continuous mathematical model

Electric field in the working space of the separator is described by the equation for the electric potential φ (see [7-8])

$$\mathrm{div}(\varepsilon\,\mathrm{grad}\,\varphi) = 0, \tag{1}$$

where ε is the dielectric permittivity. The boundary conditions are given by the known values of the electric potential on the electrodes and the Neumann condition along the artificial boundary placed at a sufficiently distance from the device.

The movement of the particle obeys the equations for its velocity v and trajectory s in the forms

$$m\frac{dv}{dt} = F_e + F_a + F_g, \quad v = \frac{ds}{dt}, \tag{2}$$

where F_e is the Coulomb force acting on the particle. This force is given by the relation

$$F_e = QE = -\,\mathrm{grad}\,\varphi, \tag{3}$$

where E denotes the local value of the electric field strength.

Symbol F_a represents the aerodynamic resistance that is given by the formula

$$F_a = -v\frac{1}{2}\rho c S v, \tag{4}$$

c being the friction coefficient (depending on geometry of the particle), ρ denotes the density of ambient air, S is the characteristic surface of the particle and v stands for the module of its velocity. Finally,

$$F_g = mg, \tag{5}$$

where m denotes the mass of the particle and g is the gravitational acceleration.

The corresponding initial conditions read

$$v(0) = v_0, \; s(0) = s_0 \tag{6}$$

where s_0 is the entry position of the particle in the separator. Equation (2) is strongly nonlinear due to the first and second terms on the right-hand side.

4. Numerical solution

The above model (equations (1) and (2)) was solved numerically. For the field computations (1) we used our own code Agros2D [9], which is a powerful user's interface serving for pre-processing and post-processing of the problems solved. The code collaborates with the library Hermes [10] containing the most advanced algorithms for a fully adaptive solution of systems of generally nonlinear and nonstationary second-order partial differential equations (PDEs) based on the finite element method of higher order of accuracy.

Both codes written in C++ are freely distributable under the GNU General Public License. The most important and in some cases quite unique features of the codes follow:

- Solution of the system of PDEs is carried out monolithically, which means that the resultant numerical scheme is characterized by just one stiffness matrix. The PDEs are first rewritten into the weak forms whose numerical integration provides its coefficients. The integration is performed using the Gauss quadrature formulas.
- Fully automatic *hp*-adaptivity. When adaptivity is required, in every iteration step the solution is compared with the reference solution (realized on an approximately twice finer mesh), and the distribution of error is then used for selection of candidates for adaptivity. Based on sophisticated and subtle algorithms the adaptivity is realized either by a subdivision of the candidate element (*h*-adaptivity), by its description by a polynomial of a higher

order (*p*-adaptivity), or by a combination of both above options (*hp*-adaptivity).

- Each physical field can be solved on quite a different mesh that best corresponds to its particulars. This is of great importance, for instance, for respecting skin effect in the magnetic field, while the temperature field is usually smooth. Special powerful higher-order techniques of mapping are then used to avoid any numerical errors in the process of assembly of the stiffness matrix.
- In nonstationary processes every mesh can change in time, in accordance with the real evolution of the corresponding physical quantities.
- Easy treatment of the hanging nodes appearing on the boundaries of subdomains whose elements have to be refined. Usually, these nodes bring about a considerable increase of the number of the degrees of freedom (DOFs). The code contains higher-order algorithms for respecting these nodes without any need of an additional refinement of the external parts neighboring with the refined subdomain.
- Curved elements able to replace curvilinear parts of any boundary by a system of circular or elliptic arcs. These elements mostly allow reaching highly accurate results near the curvilinear boundaries with very low numbers of the DOFs.

The movement of the particles affected by electric field, gravity and drag aerodynamic resistance is modeled by an adaptive Runge-Kutta-Fehlberg method with a time-varying time step.

5. Illustrative example

Two kinds of particles were considered for the computations: polyethylentereftalate (PET) with positive charges and polyvinylchloride (PVC) with negative charges. Their charges and dimensions obey the normal distribution with parameters in Table 1. Their numbers $n = 250$ for each kind of material. The voltage between the electrodes was $U = 30$ kV.

Selected parameters of the particles Table 1

Type	Density	Radius		Charge	
	(ρ kg/m^3)	μ (mm)	σ (mm)	μ (C)	σ (C)
PET	1330	2	0.25	+0.25E−9	0.8E-10
PVC	1370	2	0.25	−0.5E−9	0.8E-10

Symbol μ denotes the median and σ is the variance.

For example, Figs. 3, 4 and 5 show the distributions of the radii and charges of the PET and PVC particles.

As the particles enter the separator in different places between the electrodes, Fig. 6 shows the distribution of their initial positions.

The goal of this work is not only to calculate the electric field in the device and use it to find a trajectory of the particles, but also to find such a design of the electrodes that the number of well-placed particles is maximized. To achieve this, both electrodes are divided into several segments. By changing relative angles of the neighboring segments, different shapes of electrodes may be obtained. The objective functional of the optimization is the ratio of correctly placed particles, which should be maximized. According to a discrete nature of the objective functional, it is not possible to use any gradient-based optimization algorithm, which might be the first choice otherwise. We used approach based on genetic algorithms instead. The vector containing angles of segments serves as a “genome”. An initial population of such genomes representing different designs is created. Then the operators of mutation and crossover are applied and designs with highest values of the objective functional are selected to the next generation. After sufficient number of iterations, the best design is selected as a product of the optimization.

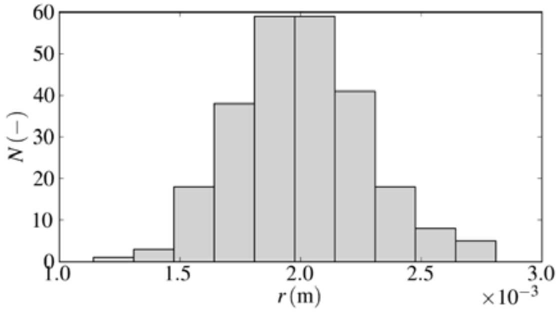


Fig. 3 Distribution of radii of the PET and PVC particles

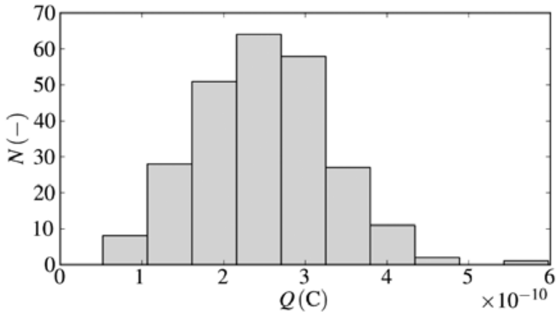


Fig. 4 Distribution of charges of the PET particles

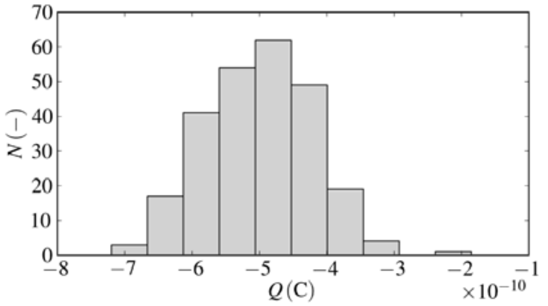


Fig. 5 Distribution of charges of the PVC particles

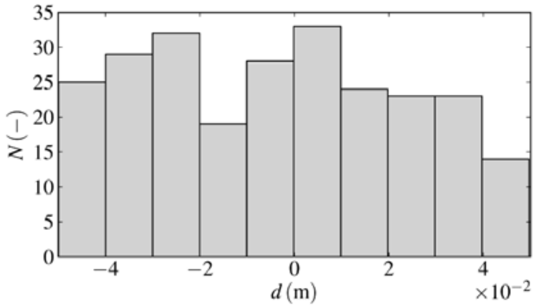


Fig. 6 Distribution of the initial positions of particular particles

More specifically, we tested two different functionals F and G with thirteen variable parameters. Every electrode was divided into four segments (three of them forming the electrode itself, the fourth one without voltage representing a barrier, see Fig. 7).

The border points are denoted by letters A_1, B_1, C_1, A_2, B_2 and C_2. These points can move in both directions within a prescribed range, which represents twelve variable parameters. The last parameter defines the place of the central bin and is given by the distance of the center of this bin from the center of the separator

The functional F describes how many particles fall into the correct bin, while the functional gives the sum of the average distances of the PET and PVC particles from the desired points (centers) of the corresponding bins.

The convergence of the algorithm was in every step evaluated from the best results obtained for both functionals ($F_{\max}$, $G_{\min}$). Its evolution is shown in Fig. 8. We can see there seemingly steady states (for example, in case of $F_{\max}$ between the sixth and tenth populations) caused by the effort of the algorithm to fill in the vacancies in the Pareto front. About 22 populations seem to be enough for obtaining the steady state.

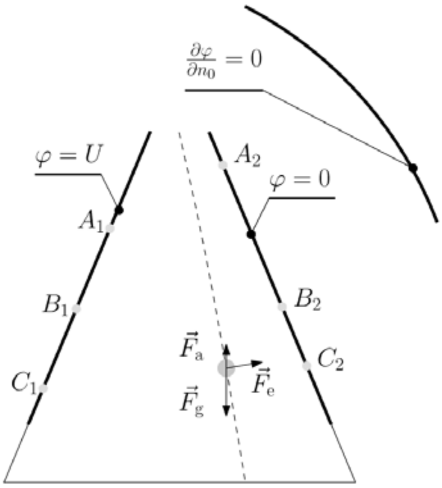


Fig. 7 Points A_1, ..., C_2 on the electrodes whose positions are to be optimized

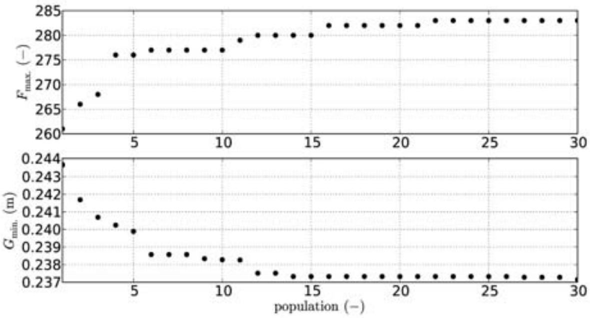


Fig. 8 Convergence of the functional s F and G

Figure 9 shows particular populations and their influence on both functionals. A detail of the front part is shown in Fig. 10.

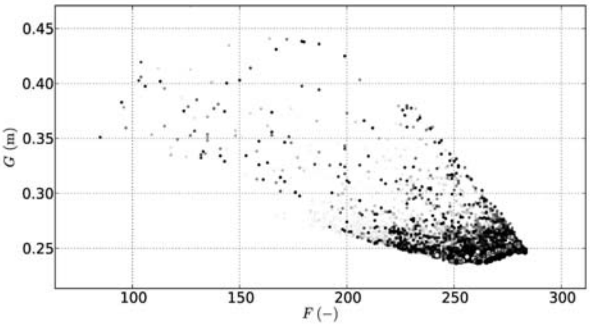


Fig. 9 Evolution of both functionals F and G with populations

The shape of the electrodes obtained after thirty populations is (for the variant of $F_{\max}$) depicted in Figs. 11 and 12 in common with trajectories of typical sets of the PET and PVC particles. In the entering part of the working chamber the particles are very close to both electrodes and the Coulomb forces acting on them reach the highest values. For the optimized variant, this fact is, moreover, supported by their shape. The vectors of electric force are here almost perpendicular to the trajectories of individual particles.

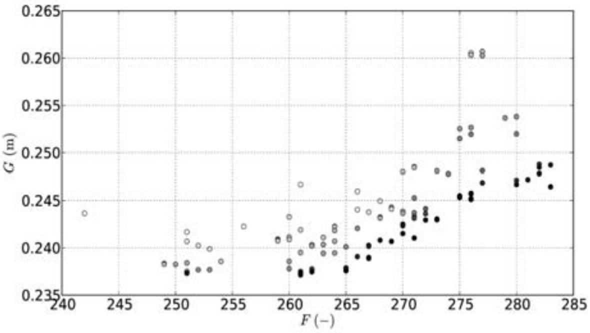


Fig. 10 Detail of the front part of Fig. 9

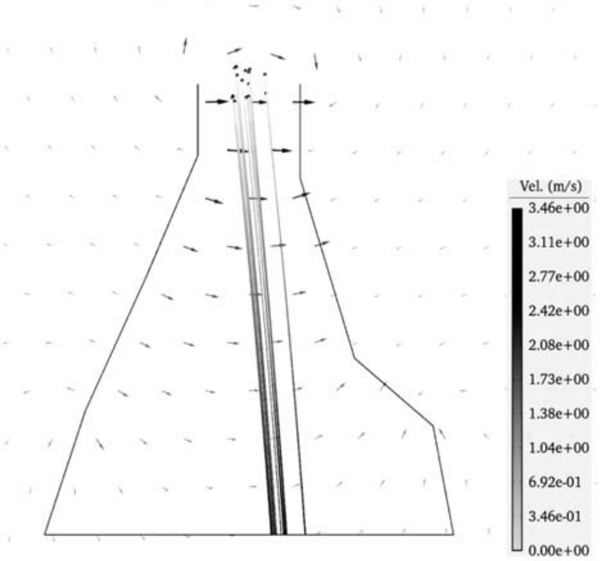


Fig. 11 Map of the trajectories and velocities of PET particles in the 30^{th} population

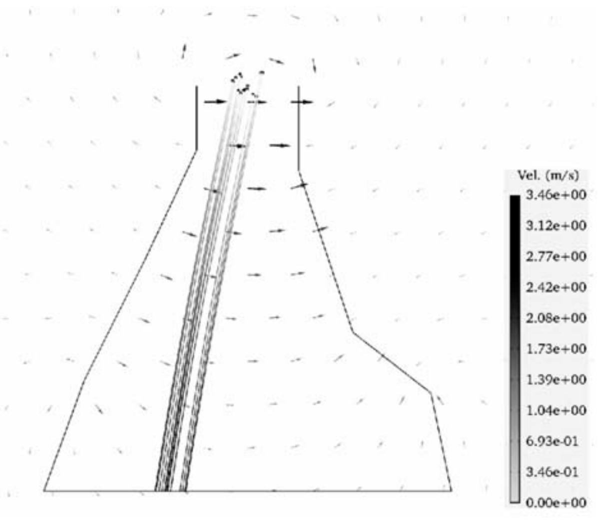


Fig. 12 Map of the trajectories and velocities of PVC particles in the 30^{th} population

With respect to a considerable difference in the charge of both types of particles (the charge of the PVC particles is almost twice as high as the charge of the PET particles), the shapes of both electrodes differ substantially. The source electrode acting on the PVC particles is turned out of the working chamber. On the other hand, the grounded electrode is bent inward the working chamber. In this way, the change of the distance of both electrodes from the axis of the chamber leads to the balance of their forces on particular trajectories of the particles.

6. Conclusion

A triboelectric separator of plastic particles was modeled and optimized. First, we proposed a device with symmetrically arranged electrodes, mapped the distribution of electric field it and investigated the trajectories of the particular levels of plastics. The grains of plastics were characterized by the normal distributions of both charges and dimensions. The efficiency of separation (evaluated from the viewpoint of the number of particles that fell down into the correct bins) was, however, low and did not exceed 47 %.

The second step was, therefore, to optimize the shapes of both electrodes in order to enlarge this quantity. After a thorough testing we suggested two objective functions to be extremized, which was realized by a suitable genetic algorithm. The modified shape of the electrodes leads to the increase of efficiency to 55.4 %.

Acknowledgment

This work was supported by the European Regional Development Fund and Ministry of Education, Youth and Sports of the Czech Republic (project No CZ.1.05/2.1.00/03.0094: Regional Innovation Centre for Electrical Engineering – RICE) and Grant project GACR P102/11/0498.

References

[1] SAEKI, M.: Triboelectric Separation of Three-Component Plastic Mixture, *Particulate Science and Technology*, vol. 26/5, pp. 494–506, 2008.

[2] MOESNER, F. M., TOSHIRO, H.: *Contactless Manipulation of Microparts by Electric Field Traps.* Proc. of SPIE's Intern. Symposium on Microrobotics and Microsystem Fabrication, Pittsburgh, vol. 3202, pp. 168–175, 1997.

[3] YANAR, D. K., KWETKUS, B. A.: Electrostatic Separation of Polymer Powders, *J. of Electrostatics,* vol. 36, No. 2–3, pp. 257–266, 1995.

[4] WEI, J., REALFF, M. J.: Design and Optimization of Free-Fall Electrostatic Separators for Plastics Recycling, *AIChE Journal*, vol. 49, No. 12, pp. 3138–3149, 2003.

[5] TILMATINE, A., MEDLES, K., YOUNES, M., BENDAOUD, A., DASCALESCU, L.: Roll-Type Versus Free-Fall Electrostatic Separation of Tribocharged Plastic Particles, *IEEE Trans. Industry Appl.,* vol. 46, No. 4, pp. 1564–1569, 2010.

[6] MACH, F., KUS, P., KARBAN, P., DOLEZEL, I.: *Optimized Arrangement of Device for Electrostatic Separation of Plastic Particles.* Proc. of conference ELEKTRO 2012, Zilina, 2012, CD-ROM.

[7] RUZICKOVA, M.: Discrete and Differential Equations in Applied Mathematics. *Communications - Scientific Letters of the University of Zilina,* No. 2, 2008.

[8] KUCZMANN, M., IVANYI, A.: *The Finite Element Method in Magnetics,* Akademiai Kiado, Budapest, 2008.

[9] Code Agros2D: http://agros2d.org.

[10] Code Hermes: http://hpfem.org/hermes.

Peter Sekerak – Valeria Hrabovcova – Matus Onufer – Pavol Rafajdus – Lukas Kalamen *

LOSSES, EFFICIENCY AND THERMAL BEHAVIOR OF THE SYNCHRONOUS MOTORS WITH DIFFERENT PM MATERIALS

The paper deals with losses, efficiency and thermal behavior of the synchronous machines with different permanent magnet (PM) materials. Three PMSM are presented in this paper: First, the PMSM with NdFeB PM, which is really constructed. Second, the NdFeB was replaced by ferrites PM. Third, the new design of synchronous motor with ferrites is presented. The volume of PM and the whole size are increased to keep the same power output. Equivalent circuit parameters, EMF, V-curves, losses, efficiency and thermal behavior of all three machines are investigated and compared.

Keywords: *NdFeB, ferrite, synchronous motor, efficiency, thermal behavior, V – curves.*

1. Introduction

Thanks to new PM materials with high energy product, such as NdFeB the synchronous motor with PM (PMSM) can be a very good choice for many industrial applications. Nowadays (2012) the cost of this type of PM has enormously increased. This situation forces the electric machine producers to restrict the PMSM production or to look for other PM materials such as Alnico alloys or ferrites, the main properties of which are shown in Table 1.

PM materials properties [1] and [2] Table 1.

	BH_{max} (kJm^{-3})	B_r (T)	H_c (kAm^{-1})	ρ_{PM} (Ωcm)
NdFeB	220 – 500	0.97 – 1.45	740 – 1000	$100 - 200 \cdot 10^{-6}$
SmCo	120 – 240	0.85 – 1.1	620 – 840	$86 \cdot 10^{-6}$
Ferrite	7 – 42	0.2 – 0.48	120 – 360	10^6
Alnico	10 – 35	0.6 – 1.16	40 – 120	$47 - 68 \cdot 10^{-6}$

It is interesting to compare properties of motors with various kinds of PM materials.

In [3] the authors investigated two PMSMs. One is with ferrite magnets and another is with NdFeB magnets. Both of them are interior type so the PMs are totally embedded in rotor. The authors did complete analysis of parameters, losses and efficiencies of both motors. The efficiency of a ferrite motor was about 1 % lower than in a motor with NdFeB if the volume of ferrites was five times higher than the volume of NdFeB.

In [4] the authors presented two PMSMs. One is with SmCo magnets and another is with ferrites. The efficiency of the ferrite motor was by 0.5 % lower at the same output power. The weight of ferrites increased 2.5 times in comparison with SmCo.

In [5] the authors presented PMSM with ferrites and with damper winding. The authors compare their machine with another at the constant PM volume. PMSM with ferrites shows 4 % improvement in efficiency.

This paper deals with synchronous motors with two kinds of PM materials. First, the original PMSM with NdFeB is presented (motor A). Second, the NdFeB is replaced by ferrite in the original PMSM construction (motor B), see Fig. 1. Third, the new design of PMSM with ferrites is presented (motor C), see Fig. 2. All three motors are investigated with regard to efficiency and thermal behavior.

2. PMSMS configurations and parameters

Permanent magnets of Motor A and B are totally immersed in rotor, which means they are interior type of the PMSM. Motor C has surface mounted magnets which provides higher magnetic flux density into the air gap. The aim was to achieve the same output power as in the case of Motor A. The parameters and dimensions of all three motors are in Table 2.

The motor C has bigger dimensions and PMs volume due to lower energy product of ferrite PMs. The parameters of PM materials used in the investigation of all motors are presented in Table 3.

---

* **Peter Sekerak, Valeria Hrabovcova, Matus Onufer, Pavol Rafajdus, Lukas Kalamen**
University of Zilina, Department of Power Electrical Systems, Zilina, Slovakia, E-mail: valeria.hrabovcova@fel.uniza.sk

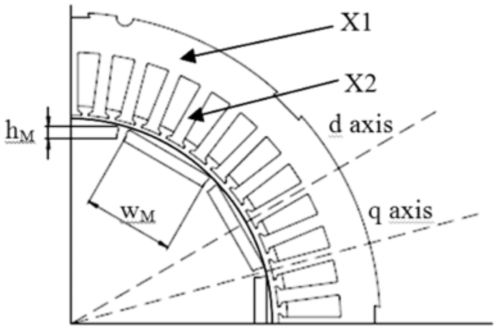


Fig. 1 Cross section area of a quarter of Motor A,B

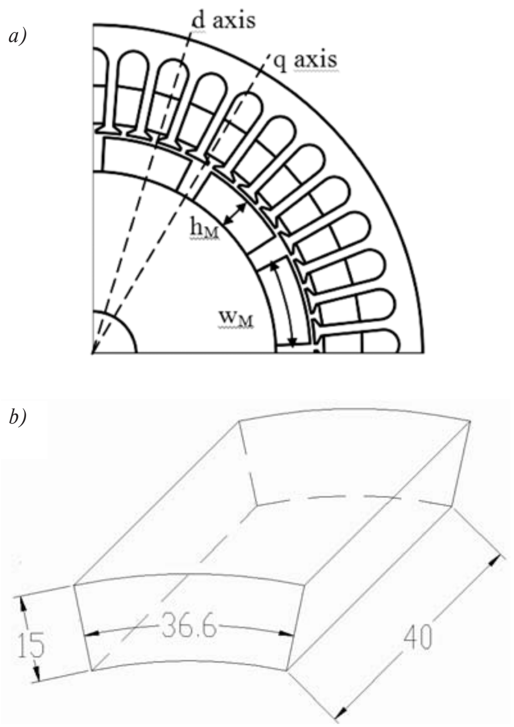


Fig. 2 a) Cross section area of a quarter of Motor C b) detail of PM shape

On the basis of analytical calculations and FEM simulations [6] the parameters of all three models were investigated. The parameters such as stator resistance R_s, stator leakage inductance L_σ, magnetizing inductance in d and q axis $L_{\mu d}$, $L_{\mu q}$ and flux linkage established by PMs ψ_{PM} are shown in Table 4.

Parameters and dimension of Motor A, B and C Table 2.

Parameter	Motor A	Motor B	Motor C
PM material	NdFeB	ferrite	ferrite
Stator phase voltage	230 V	230 V	230 V
Frequency	36 Hz	36 Hz	36 Hz
Pole pairs	6	6	6
Rated power	2 kW	0.5 kW*	2 kW
Outer rotor diameter	0.146 m	0.146 m	0.19 m
Outer stator diameter	0.22 m	0.22 m	0.288 m
Iron stack length	0.14 m	0.14 m	0.12 m
Permanent magnet width	4 mm	4 mm	15 mm
Permanent magnet length	32 mm	32 mm	36.6 mm
PMs volume	$215\ \text{cm}^3$	$215\ \text{cm}^3$	$794\ \text{cm}^3$
Stator winding turns	420	420	840
Stator slots	48	48	45
Slots per pole	4	4	3.75
Slots per pole per phase	1.33	1.33	1.25

* at Motor B it is power at rated current

PM materials properties [1] and [2] Table 3.

NdFeB		
Br	Hc	BHmax
1.15 T	$860\ \text{kAm}^{-1}$	$240\ \text{kJm}^{-3}$
Ferrite		
Br	Hc	BHmax
0.45 T	$340\ \text{kAm}^{-1}$	$40\ \text{kJm}^{-3}$

PMSM parameters Table 4.

	Motor A	Motor B	Motor C
R_s	3.93 Ω	3.93 Ω	6.047 Ω
L_σ	0.0289 H	0.0289 H	0.0933
$L_{\mu d}$	0.025 H	0.025 H	0.0297 H
$L_{\mu q}$	0.078 H	0.078 H	0.0297 H
ψ_{PM}	0.828 Vs	0.2 Vs	1.325 Vs
E_{PM}	132.5 V	34 V	211 V
U_{sph}/E_{PM}	1.73	6.76	1.09

3. PMSM properties

First, the PM electromotive force E_{PM} is studied. Its value has significant influence on all properties. Then the V -curves were investigated. They show an optimal operating point with minimal current and, consequently, with minimal value of Joule losses. Also thermal behavior of all motors was searched.

A. E_{PM}

E_{PM} is a very important parameter for PMSM operation. The ratio of the stator phase voltage U_{sph} over E_{PM} is defined. A very high ratio of the U_{sph}/E_{PM} means a big difference between the induced and terminal voltage which results in a high stator current $I_{as} = (U_{sph} - E_{PM})/X_d$. This is a base for high Joule losses.

The waveform of the air gap magnetic flux densities B_δ of Motors A, B and C are depicted in Figs. 3 a, b, c. Note a great difference in B_δ values of Motor B and its low fundamental harmonic in Fig. 4 where the Fourier series of B_δ of all the motors are shown.

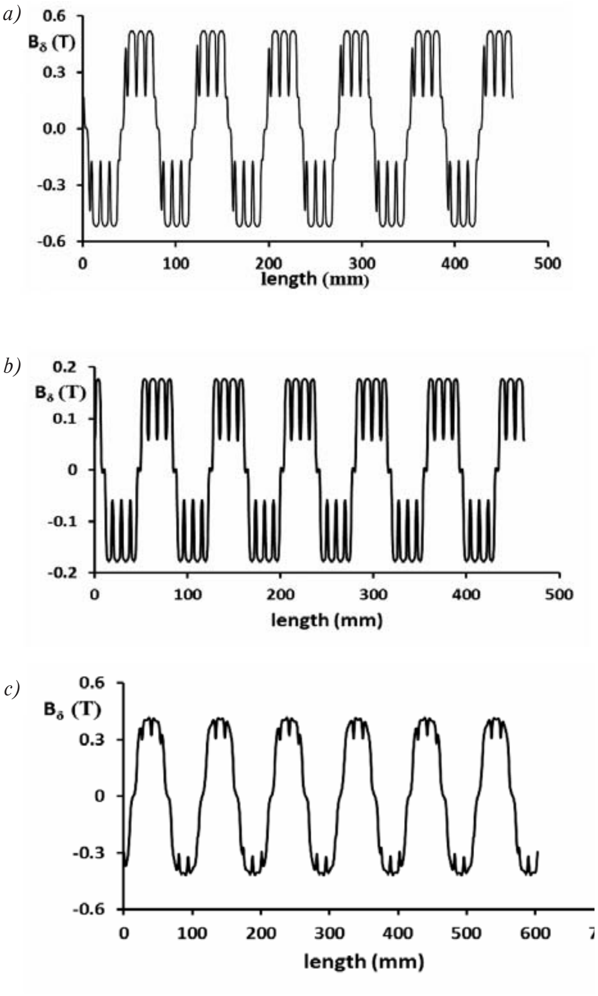


Fig. 3 Air-gap magnetic flux density of a) Motor A b) Motor B c) Motor C versus the whole air gap periphery (12 poles)

The induced voltage E_{PM} created in the stator winding by PM can be calculated for each ν harmonic component by the formula

$$E_{PM\nu} = \sqrt{2}\,\pi f_\nu \phi_\nu N k_{w\nu}\,, \tag{1}$$

where f_ν is the frequency, N is the number of stator winding turns, $k_{w\nu}$ is its winding factor and ϕ_ν is the magnetic flux calculated by

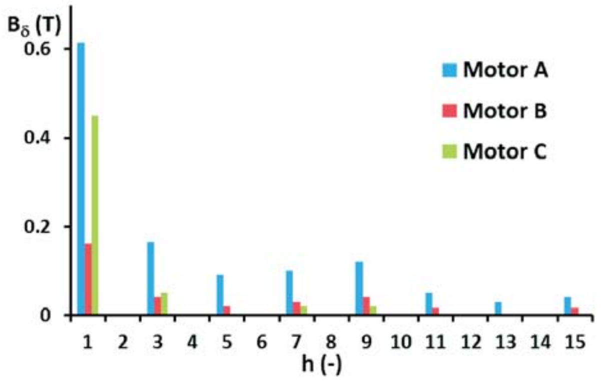


Fig. 4 Fourier series of air gap magnetic flux densities of Motors A, B and C

$$\phi_\nu = \frac{2}{\pi} B_{\delta\nu} \frac{\tau_p}{\nu} l_{Fe}\,, \tag{2}$$

where l_{Fe} is the iron stack length, τ_p is the pole pitch.

Fig. 5a shows the simulated induced phase voltage of Motor A gained by means of FEM and calculated by the sum of $E_{PM\nu}$ (1). In Fig. 5b is shown a real measured waveform of phase voltage of Motor A taken by an oscilloscope. The coincidence of measured and simulated waveforms is very good. This means that this procedure is applicable for Motors B and C. The RMS value of E_{PM} of Motor A is 132.5 V, Table 4. Ratio U_{sph}/E_{PM} is 1.73 at $U_{sph} = 230$ V.

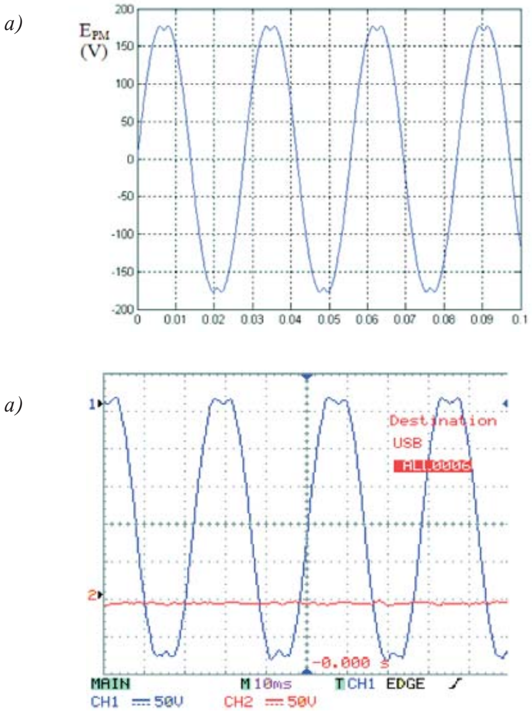


Fig. 5 E_{PM} waveform of Motor A a) simulated by means of FEM and calculated by (1), b) measured

The simulated E_{PM} of Motor B is shown in Fig. 6.

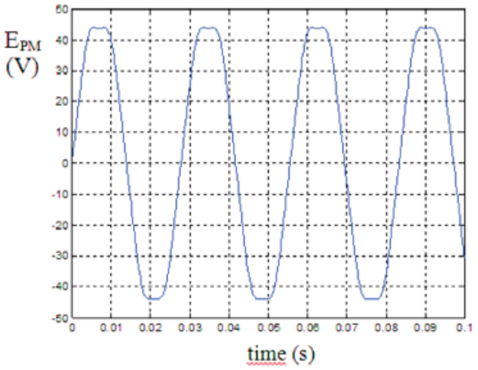


Fig. 6 E_{PM} waveform of Motor B – simulated by means of FEM

The simulated E_{PM} of Motor C is shown in Fig. 7.

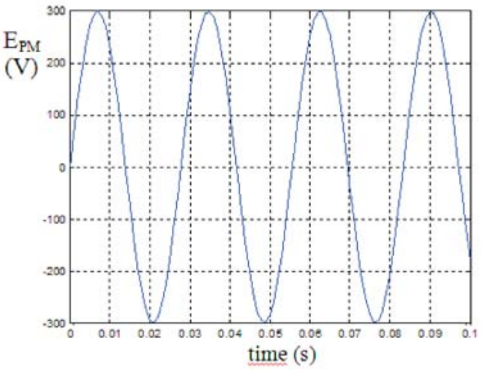


Fig. 7 E_{PM} waveform of Motor C – simulated by means of FEM

The RMS value of induced voltage E_{PM} of Motor B is $E_{PMrms} = 34$ V, see Table IV. This value gives the ratio $U_{sph}/E_{PM} = 6.76$ at $U_{sph} = 230$ V. In comparison with Motors A and C is this value enormously low because of its very low remanent magnetic flux density. Motor C gives $E_{PMrms} = 211$ V, see Table 4. The ratio U_{sph}/E_{PM} is 1.09 at $U_{sph} = 230$ V. This ratio is very close to theoretically recommended $U_{sph}/E_{PM} = 1.11$. These values and motors parameters will be put into the simulation model to determine their properties.

B. *V – curves*

The V – curves in the case of PM excitation, which is constant, can be plotted as stator current I_{as} versus ratio of U_{sph}/E_{PM} at different loads. Fig. 8 shows V – curves of Motor A. The simulations were verified by experiments. The stator voltage was changed from 230 V to 150 V which represents ratio U_{sph}/E_{PM} from 1.73 to 1.13 respectively. The motor was loaded from 1 kW to 2.5 kW. The results of simulation and experiments are in good coincidence. V – curves show the optimal operational point in their minimum. Also the minimal current I_{as} and corresponding ratio U_{sph}/E_{PM} at the given power can be seen.

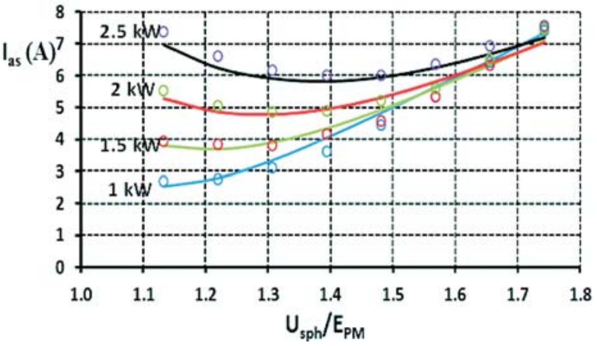


Fig. 8 V – curves of Motor A at frequency 36 Hz. Lines represent simulated data, circles measured data

Fig. 9 shows V – curves of Motor B. Motor B was loaded, in simulation only, by loads 0.5 kW, 1 kW, 1.5 kW, 2 kW and 2.5 kW. The stator voltage was changed from 230 V to 116 V which means ratio U_{sph}/E_{PM} from 6.8 to 3.4 respectively. As it is seen from Fig. 9 Motor B can provide 2 kW, but in comparison with other two motors mentioned above the stator current I_{as} is very high (around 10 A). The ratio U_{sph}/E_{PM} is very high and it doesn't meet the requirement for optimal ratio $U_{sph}/E_{PM} = 1.11$ [7]. A solution could be its operation at lower stator terminal voltage to achieve a ratio of U_{sph}/E_{PM} closer to 1.11. In that case for $E_{PM} = 34$ V the U_{sph} should be around 38 V. This would result in power around 50 W. It means that Motor B cannot operate satisfactorily in the required range of powers and stator voltage.

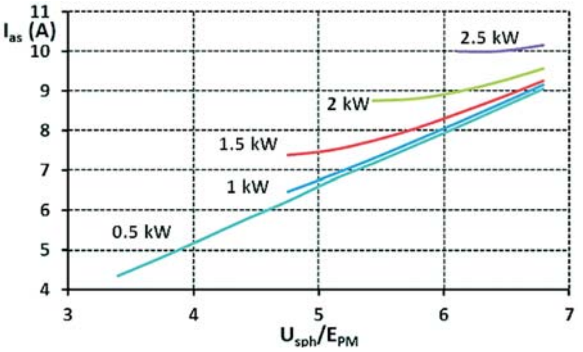


Fig. 9 V – curves of Motor B at frequency 36 Hz

The V – curves of Motor C are shown in Fig. 10. The stator voltage was changed from 266 V to 175 V, which represents the ratio U_{sph}/E_{PM} from 1.26 to 0.82, respectively. At 2 kW, the minimum of stator current 3.15 A is achieved at $U_{sph}/E_{PM} = 1.16$ which is very close to the recommended value 1.11.

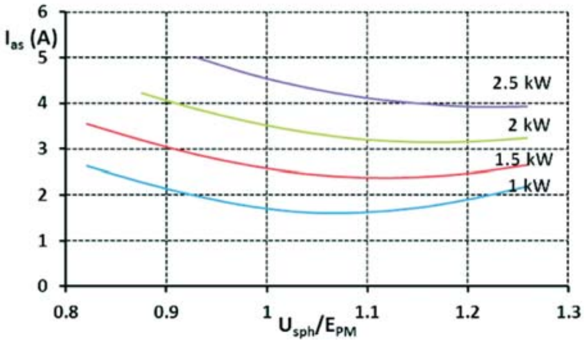


Fig. 10 V - curves of Motor C at frequency 36 Hz

C. Losses

Losses were investigated in all three motors. The next types of losses were taken into account:

- Joule losses ΔP_J
- Iron losses ΔP_{iron}
- PM losses ΔP_{PM}
- Mechanical losses ΔP_{mech}
- Additional losses ΔP_{ad}

The Joule losses are calculated by:

$$\Delta P_J = 3R_{sph} I_{as}^2 \tag{3}$$

where R_{sph} is the phase resistance at 75 °C.

Iron losses in n-th element (element means tooth, yoke, etc.) were calculated by the following formula [7]:

$$\Delta P_{iron,n} = p_{10}\left(\frac{B_{\max,n}}{1T}\right)^2 m_{iron,n}\left(\frac{f}{50}\right)^{1.3} \tag{4}$$

where $p_{10} = 3.1$ Wkg^{-1} is the iron loss per unit of mass at magnetic flux density $B = 1$ T, $B_{\max,n}$ is the amplitude of magnetic flux density in n-th element of the machine and $m_{iron,n}$ is its mass. The investigation of magnetic flux densities was done by means of FEM.

The amplitude of magnetic flux density $B_{\max,n}$ has two components, tangential and normal as can be seen in Fig. 11.

$$B_{\max,n} = \sqrt{B^2_{\max,n,\tan} + B^2_{\max,n,norm}} \tag{5}$$

The waveforms of tangential and normal components of the magnetic flux density versus electrical angle of the stator inner periphery of Motor A at the output power 2 kW and optimal $U_{sph}/E_{PM} = 1.27$ is shown in Figs. 12a,b. These values are taken from Fig. 1, point X1, in the stator yoke, and X2, in the stator tooth.

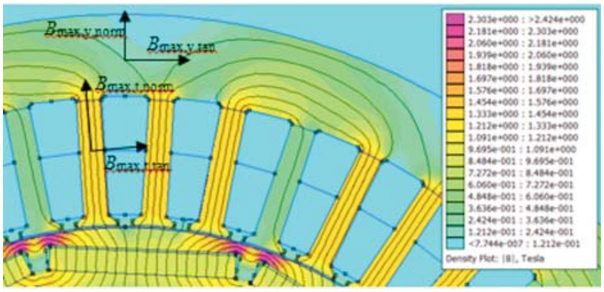
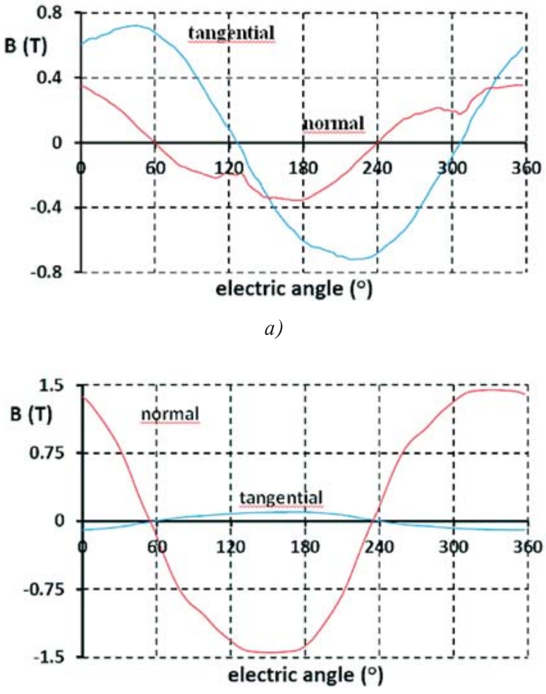


Fig. 11 Cross section PMSM with marked normal and tangential components of the magnetic flux density



a)

b)

Fig. 12 The tangencial and normal components of magnetic flux density in a) stator yoke, b) stator tooth of Motor A at rated load and optimal $U_{sph}/E_{PM} = 1.27$

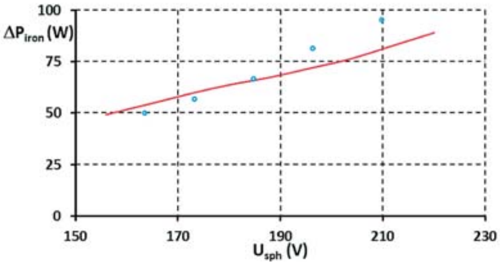


Fig. 13 The iron losses of Motor A versus stator voltage at 500 rpm. The line represents simulated data, circles experimental data

On the basis of (4) and (5) the iron losses were calculated and compared with the experimental data, see Fig. 13.

The experimental data were gained by measurements where the input phase stator voltage was gradually changed (see Fig. 13) and current, input power, output power and temperature were measured. The total input power is expressed as:

$$P_{in} = P_{out} + \Delta P_{js} + \Delta P_{iron} + \Delta P_{ad} + \Delta P_{PM} + \Delta P_{mech} =$$

$$= P_{out} + \Delta P \tag{6}$$

It is clear that to evaluate together the $\Delta P_{iron} + \Delta P_{ad}$ from the ΔP, it is necessary to subtract the components of ΔP_{js}, ΔP_{PM}, and ΔP_{mech}. Joule losses ΔP_{js} were corrected by temperature to involve the changing of the stator resistance. The mechanical losses $\Delta P_{mech} = 51$ W at 500 rpm were determined by special measurement. Losses in PM were finally neglected, see (7), (8). This approach was accepted in simulation as well as in measurement evaluation. The measurement was carried out at sinusoidal voltage and frequency of 50 Hz.

The simulated iron losses as a function of stator phase voltage of Motor B and C are shown in Fig. 14.

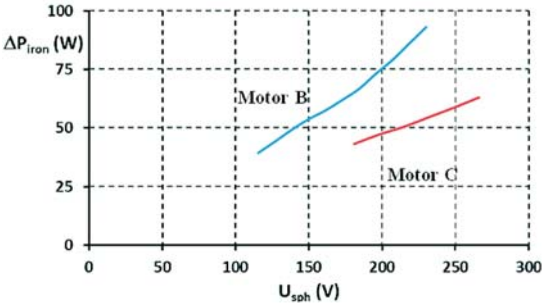


Fig. 14 The simulated iron losses of Motors B an C

The PM losses were calculated on the base of Ruoho at al. [8], where the next expression is given:

$$\Delta P_{PM} = \frac{V_{PM} f^2 w_{PM}^2 B_{PM}^2}{12\rho} \tag{7}$$

The w_{PM} is the PM width, $\rho = 150 \cdot 10^{-6}$ Ωcm is the resistivity of the NdFeB PM and B_{PM} is a magnitude of B pulsation in permanent magnet. If the PM volume V_{PM} is substituted by $w_{PM} l_{PM} h_{PM}$, it is seen that eddy current loss in PM depends on the 3rd power of the PM width. This model neglects the eddy current loss on the edge of the magnets. If this kind of the eddy current loss should be taken into account than the next formula must be used:

$$\Delta P_{PM} = \frac{f^2 B_{PM}^2}{16\rho} \frac{w_{PM}^3 l_{PM}^3 h_{PM}}{\left(w_{PM}^2 + l_{PM}^2\right)} \tag{8}$$

where l_{PM} is the PM length in the axial direction and h_{PM} is the PM height in the radial direction.

The PM loss was investigated by simulation by means of 2D FEM, [9]. The dashed red line on the PM surface represents the place where magnetic flux density was investigated, see Fig. 15.

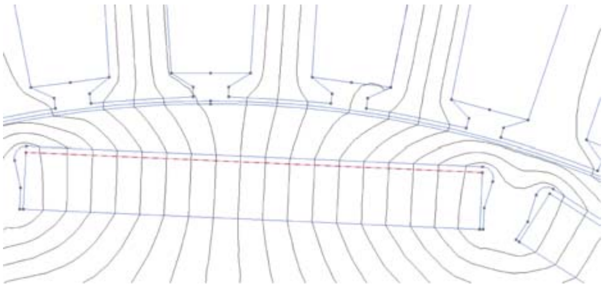


Fig. 15 FEM model for investigation of magnetic flux density on the PM surface (see dashed red line)

The 2D FEM model represented the no-load condition in a generating mode. In this mode the influence of the stator slots could be seen on the magnetic flux density taken on the PM surface. However, this phenomenon is clear only in the case of the surface mounted PM. In the investigated motor the PMs are immersed in the rotor body and no influence of the stator slots can be observed. This is clearly seen in Fig. 16 where the smooth waveform of the magnetic flux density on the PM surface along the PM width is presented.

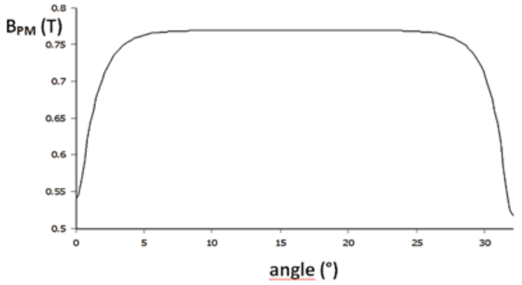
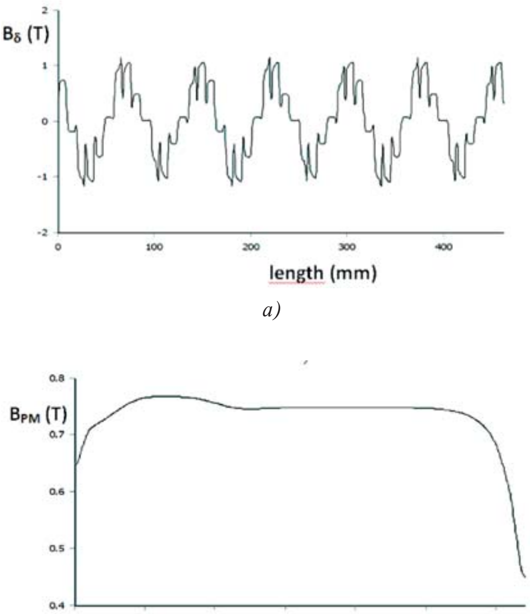


Fig. 16 The waveform of the magnetic flux density on the PM surface versus PM width in generating no-load condition (Motor A)

In the next step the simulation of the loaded motoring condition at rated 53 Nm was carried out. The profile of the air gap magnetic flux density along the rotor periphery, where the influence of the stator slots can be seen, is in Fig. 17a, the profile of the magnetic flux density investigated on the PM surface, if the PMSM is loaded, is in Fig. 17b. It is seen that the waveform is distorted by armature reaction in comparison with no load condition (Fig. 16).

The waveform of the B_{PM} from Fig. 17b was kept constant which means no variation in time and as a consequence no eddy currents in the PM. It means that the immerse PMs in the rotor



a)

b)

Fig. 17 Magnetic flux density versus air gap periphery: a) in the air gap of the PMSM, b) on the surface of the immersed PM if PMSM was loaded by the rated torque 53 Nm

body are shielded by a layer of iron over them and protected against the influence of the stator slotting. Therefore, no eddy currents in PMs can be observed in this case and the losses in the PMs can be neglected.

In motor C where ferrites are employed and they are mounted on the surface of the rotor, the PM losses can be also neglected but due to another reason.

The ferrites have very high resistivity, 10^6 Ωcm, see Table 1, which is much more higher than in NdFeB PMs. This fact considerably affects their operational performances in a positive way.

Mechanical losses were measured and calculated by empirical formulas by [7]. For Motors A and B the ΔP_{mech} is 36 W and for Motor C ΔP_{mech} is 40 W at 360 rpm.

D. *Efficiency*

The operation at the optimal stator current will result in lower winding losses and higher efficiency. The comparison of all three motor types can be done on the base of Figs. 18, 19 and 20.

Efficiency was calculated by a well known formula:

$$\eta = \frac{P_{out}}{P_{in}} = \frac{P_{out}}{P_{out} + \Delta P_{js} + \Delta P_{iron} + \Delta P_{ad} + \Delta P_{mech}} \quad (9).$$

Motor A was simulated and really measured, Motors B and C were investigated by simulations only. Fig. 18 shows the efficiency of Motor A at different loads. Into the simulations the Joule losses, iron losses simulated by 2D FEM models and the mechanical losses were introduced.

The maximal efficiency of Motor A $\eta = 86.5$ % at rated load was achieved at ratio $U_{sph}/E_{PM} = 1.33$, see Fig. 18.

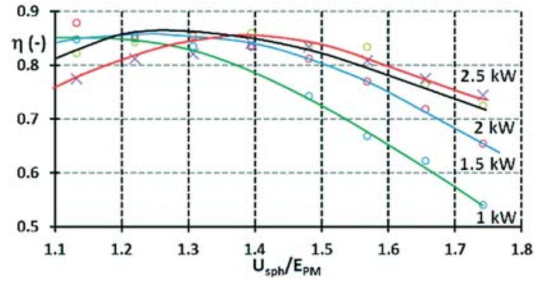


Fig. 18 The efficiency of Motor A, Lines represent simulated data, circles represent measurements

Fig. 19 shows the efficiency of Motor B at different loads, Fig. 20 the Motor C.

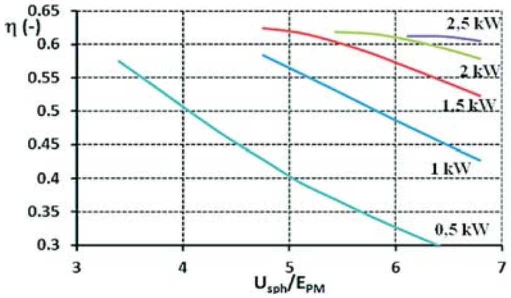


Fig. 19 Efficiency of Motor B

Motors A and C, see Figs. 18 and 20, provide good efficiency. The efficiency of motor B (Fig. 19) is very low, which doesn't match the requirements of high efficiency industrial applications. If the design of the PMSM with ferrites is changed appropriately as it was done in motor C, the efficiency is comparable with the PMSM with NdFeB magnets: the maximal efficiency of motor C is $\eta = 87.5$ % at ratio $U_{sph}/E_{PM} = 1.16$.

E. *Thermal behavior*

An investigation of thermal behavior of electric motors is very important because of PM and insulation of the stator winding. The stator winding insulation has a temperature limit of 155 °C which belongs to thermal class F. The NdFeB PM have maximal allowed working temperature around 120 °C. In order not to exceed it, the thermal measurement and thermal simulation was carried out by means of a thermal equivalent circuit and equation

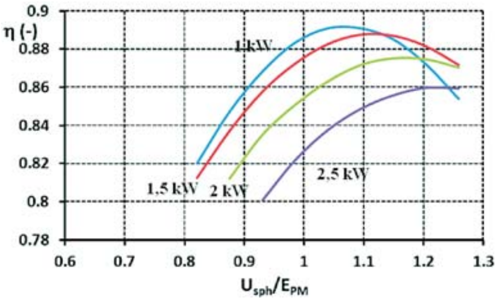


Fig. 20 Efficiency of Motor C

system by means of Matlab. In Motor A the simulations were confirmed by the experiments. The temperature curves are plotted in Fig. 21. The experiments and simulations were performed for the rated load 2 kW and for two values of the stator voltages. At the rated stator voltage 230 V the temperature exceeds the allowed values of 155 °C (class F) in the stator winding as well as 120 °C in PM. Then such stator voltage, at which the stator current is the smallest one (I_{as} = 4.9 A), see Fig. 8, was applied. The ratio is U_{sph}/E_{PM} = 1.27, which gives U_{sph} = 180 V. In this operating point the Joule losses should be the smallest and also temperature is the smallest but it is close to the temperature limit (see Figs. 21 c, d).

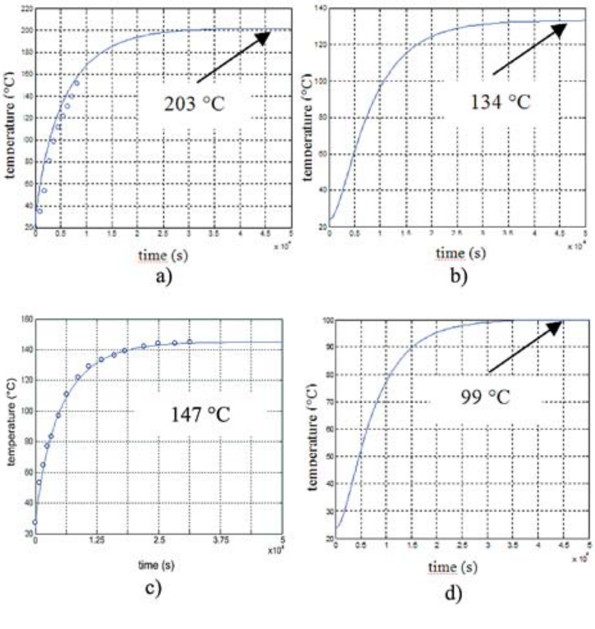


Fig. 21 The temperature curves of Motor A for rated load 2 kW: a) stator winding temperature, U_{sph} = 230 V, I_{as} = 7.3 A, b) PM temperature, U_{sph} = 230 V, I_{as} = 7.3 A, c) stator winding temperature, U_{sph} = 180 V, I_{as} = 4.9 A, d) PM temperature, U_{sph} = 180 V, I_{as} = 4.9 A. Lines represent simulated data, circles represent experimental data

The Motor C operates at the rated load 2 kW very close to the optimal smallest stator current. At this point the U_{sph}/E_{PM} is 1.09, which means U_{sph} = 230 V. The simulation was carried out in this operating point, see Fig. 22.

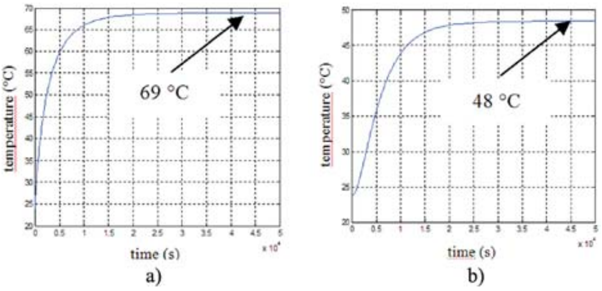


Fig. 22 The temperature curves of motor C for rated load 2 kW: a) stator winding temperature, U_{sph} = 230 V, I_{as} = 3.2 A, b) PM temperature, U_{sph} = 230 V, I_{as} = 3.2 A

The Fig. 22 shows that motor C works at rated stator voltage and rated load in safe temperature limits and there is no risk of its thermal damage.

4. Conclusion

The effect of different PM material on PMSM properties was investigated. The experiments and simulations showed how important it is to choose the right ratio of the U_{sph}/E_{PM} in PMSM operation. Very low and also very high ratios cause higher stator current which results in dangerous thermal conditions in PMSM. The simulations also showed that the replacement of the NdFeB by ferrites in the original configuration causes very poor properties of PMSM and a new design of PMSM is required. Although the ferrites do not have so excellent properties as NdFeB, it is possible to design PMSM with a very good efficiency if the design is proper. The progress in ferrite properties is expected in the future and then they could be employed in PMSM designs more often also due to their lower costs. The investigation confirms the fact that a PMSM design is a complex process. Many aspects have to be taken into account if the PMSM with excellent properties is required.

Acknowledgment

This work was supported by the R&D operational program Centre of excellence of power electronics systems and materials for their components II. No. OPVaV-2009/2.1/02-SORO, ITMS 26220120046 funded by European regional development fund (ERDF).

References

[1] http://www.hitachi-metals.co.jp/e/prod/prod03/p03_10.html

[2] http://www.magsy.cz

[3] LIANG FANG, LEE, B.H.; LEE, J.J.; KIM, H.J.; JUNG-PYO Hong: *Study on High-Efficiency Characteristics of Interior Permanent Magnet Synchronous Motor with Different Magnet Material, Electrical Machines and Systems,* ICEMS 2009

[4] RICHTER, E., NEUMANN, T.: Line Start Permanent Magnet Motors with Different Material, *IEEE Trans. Magnetics,* vol. 20, p. 1762–1764, ISSN: 0018-9464

[5] CHAUDHARI, B. N., FERNANDES, B. G.: *Synchronous Motor Using Ferrite Magnets for General Purpose Energy Efficient Drive,* TENCON 99, 1999, ISBN: 0-7803-5739-6

[6] SEKERAK, P., HRABOVCOVA, V., RAFAJDUS, P., KALAMEN, L.: *Interior Permanent Magnet Synchronous Motor Parameters Identification,* ISEM 2010, Prague, 2010, AFC, pp.: 107–116, 978-80-01-04621-0

[7] PYRHONEN, J., JOKINEN, T., HRABOVCOVA, V.: *Design of Rotating Electrical Machines,* Wiley, 2008, ISBN: 978-0-470-69516-6

[8] RUOHO, S., KOLEHMAINEN, J., IKAHEINEN, J., ARKKIO, A.: Interpedance of Demagnetization, Loading and Temperature Rise in a permanent-Magnet Synchronous Machine, *IEEE Trans Magnetics,* vol. 46, No. 3, March 2010

[9] RAFAJDUS, P., SEKERAK, P., KALAMEN, L., HRABOVCOVA, V., CHEN, H.: Static and Dynamic Analysis of Linear Switched Reluctance Machine, *Communications – Scientific Letters of the University of Zilina,* vol. 6, No. 4, 2011.

Miroslav Benco – Robert Hudec – Martina Zachariasova – Patrik Kamencay – Slavomir Matuska *

NOVEL APPROACH TO COLOR TEXTURE RETRIEVAL BASED ON GLCM

The color and texture are very important features in image analysis. In this paper, the combination of these features is presented. Several approaches for color texture features extraction are researched. The improvement of the one-dimensional version of GLCM (Gray Level Co-occurrence Matrix) for color textures was designed. We named this method as Color Level Co-occurrence Matrix (1D-CLCM). The experiments on database of 2600 color images are provided. Finally, the evaluation and comparison of image retrieval results for separate methods are presented.

Keywords: *Co-occurrence, color texture, GLCM, one-dimensional, texture extraction*

1. Introduction

Texture has been one of the most popular representations in image retrieval [1]. Forasmuch as grayscale textures provide enough information to solve many tasks, the color information was not utilized. But in the last years, many researchers have begun to take color information into consideration [2-7].

However, the human eye perceives the image as combination of primary parts (color, texture, shape), our previous research was oriented to combination of these primary parts of image. Exactly, the combination of color and texture was researched and new possibilities of GLCM were researched [8].

Tou et al. [9], introduce the reduction of computations for the GLCM method by reducing the matrix dimension from two to one dimension. Based on these two approaches, the one-dimensional GLCM method for color-texture feature extraction is introduced.

In this paper, new and extended experiments of GLCM are described and new methods are proposed. Experiments are realized on database of 2600 color textures and 110 queries are used for evaluation.

The outline of the paper is as follows. In the next section, an overview of basic principles of GLCM and 1D-GLCM are introduced. Next, the color feature extractions based on 1D-GLCM are introduced in section 3. In this section, the 1D-GLCM is applied on separated color channels of RGB color image and new 1D-CLCM method is proposed. Finally, the experiments are described in 4 and brief summary is discussed in section 5.

2. Gray level methods

2.1 GrayLevel Co-occurrence Matrix

The GLCM (Gray Level Co-occurrence Matrix) is a powerful method in statistical image analysis. This method is used to estimate image properties related to second-order statistics. GLCM considers the relation between two neighboring pixels in one offset, as the second order texture, where the first pixel is called reference and the second one the neighbor pixel. GLCM is the two dimensional matrix of joint probabilities between pairs of pixels, separated by a distance d in a given direction θ.

For the scale invariant of texture pattern, the GLCM is standardized by total pairs of pixels as follows:

$$p(i,j) = \frac{Pd,\theta(i,j)}{All\_pairs\_of\_pixel\_used}\ , \tag{1}$$

where $Pd,\theta(i,j)$, expresses joint probabilities between pairs in distance d and direction θ and i,j, are luminance intensities of those pixels.

Haralick [10] defined 14 statistical features from gray-level co-occurrence matrix for texture classification. However, these features are strongly correlated [11], we decide to use only one feature descriptor for methods comparison. The feature Inverse Difference Moment also called “homogeneity” is used. The homogeneity is defined as follows:

$$Homogeneity_{d,\theta} = \Sigma i \Sigma j = \frac{Pd,\theta(diff)}{1 + |i - j|^2}\,. \tag{2}$$

---

* **Miroslav Benco, Robert Hudec, Martina Zachariasova, Patrik Kamencay, Slavomir Matuska**
Department of Telecommunications, Faculty of Electrotechnical Engineering, University of Zilina, Slovakia, E-mail: miroslav.benco@fel.uniza.sk

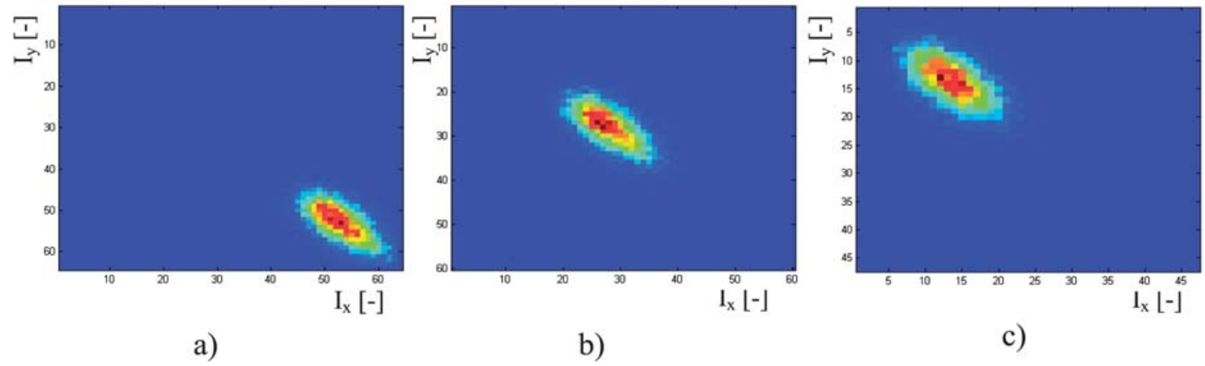


Fig. 1 Example of GLCM matrices plot for RGB color components (d = 1, θ = 0°), a) GLCM in R channel, b) GLCM in G channel, c) GLCM in B channel

2.2 One-dimensional GLCM

Tou et al. [9] reduce computations of feature extraction vector. They reduce GLCM dimension from two to one dimension by combination of certain values of the matrix. In conventional two-dimensional GLCM, the matrix in size G × G is used, where G represents the maximum gray level value in image. By focusing only on the differences of the gray level, a one-dimensional GLCM has a significantly smaller size, which is only 2 × G − 1. By reducing the dimension of the GLCM, the calculation of features is faster as fewer values are involved in the calculation.

Likewise in GLCM, the scale invariant of texture pattern is provided as follows [9]:

$$p(diff) = \frac{Pd,\theta(diff)}{All\_pairs\_of\_pixel\_used}, \tag{3}$$

where $Pd,\theta(diff)$ expressesdifference probabilities between pairs in distance d and direction θ and *diff* shows the differences of gray value between the two pixels of the pixel pairs.

The textural feature homogeneity for one-dimensional GLCM is defined as follows [9]:

$$Homogeneity_{d,\theta} = \Sigma i \Sigma j = \frac{Pd,\theta(diff)}{1+(diff)}. \tag{4}$$

3. Color level methods

However, the computations of texture features are reduced by one-dimensional GLCM, we used this advantage of computing time for next approach. Exactly, the input color image is not reduced to gray level for next methods. Thus, against previous methods, where only gray level value of pixels were considered, the color information is considered for texture feature extraction.

3.1 One-dimensional GLCMforseparatedcolorchannels

The easiest way for using gray-level method for color image is applying method to each 2D matrix of three-dimensional color image representation.

Subsequently, the color feature extraction can be defined as follows:

$$FV = [FV(C_1), FV(C_2), FV(C_3)], \tag{5}$$

where, FV is feature vector and C_1, C_2, C_3 are one-dimensional GLCM matrices of particular color channels. The example of difference of GLCM matrices for RGB color components is shown in Fig. 1. In our experiments, the color space RGB is used, thus, we named this method as RGB one-dimensional GLCM (RGB-1D-GLCM).

3.2 One-dimensionalColorLevel Co-occurrence Matrix

In our previous research [8], we used the 3D image interpretation directly for co-occurrence matrix computing. In this work, we applied this approach for one-dimensional GLCM and we called this method Color Level Co-occurrence Matrix (1D-CLCM).

The principle of CLCM method is shown in Fig. 2. For the distance and angles the cube of size 3×3×3 was created. In this case, three neighborhoods for every direction (#1–#12 in Fig. 2) was used. There are also neighborhoods on same position in image in different color components. Therefore, the direction #13 was also taken in consideration. The 13 directions neighborhood system was created. These 13 probability matrices express relations between component x_2 and its neighbors in all channels of color space. In order to get information of all channels relations, it is necessary to use this procedure in three iterations, where x_1, x_2, x_3 are changing the color space components (Table 1).

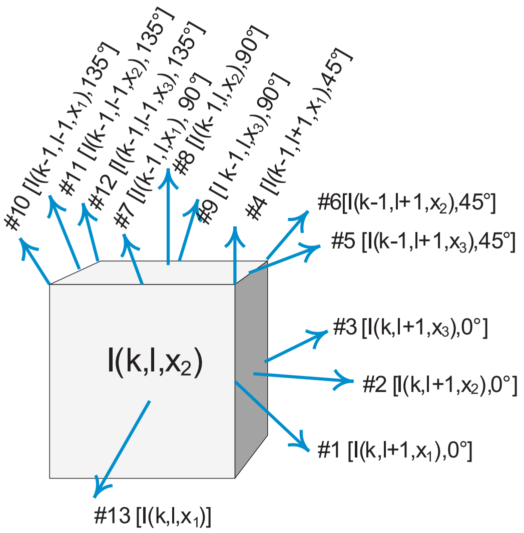


Fig. 2 Principle of 13 directions neighborhood system, x_1, x_2, x_3 are color components of color image I

The combination of color space components for GLCM Tab. 1

Colorcomponent	Iteration		
	1	2	3
x_1	C_1	C_2	C_3
x_2	C_2	C_3	C_1
x_3	C_3	C_1	C_2

For 1D-CLCM the differences for all neighbors by formula (3) are computed. Finally, feature vector consists of information of all three channels and their relations in 39coefficients (13x3).

4. Experiments

For our experiments, the Matlab environment was used. For the extraction of standard GLCM features, the library functions were used. For the others, the own functions were developed. However, we wanted to test algorithms on big database, a dataset from several smaller datasets (e.g. Vistex [12], Outex [13]) was created. This dataset contains 130 types of textures in 2600 color images (20 images for each texture) in resolution 128 × 128. The example images from texture database are shown in Fig. 3. For all experiments, the parameters distance $d = 1$ and angles $\theta = 0°$, 45°, 90°, 135° degrees are used. As was mentioned above, the only one feature “homogeneity” is used.

The scale invariant of texture pattern is provided by standardization of total pairs of pixels as was defined in (1) and (3).

For the rotation-invariant of texture pattern, the feature vector is rotated in stage of feature matching as follows:

$$VP_Q^{0^\circ} = [p(0°), p(45°), p(90°), p(135°)], \tag{6}$$

$$VP_Q^{45^\circ} = [p(45°), p(90°), p(135°), p(0°)], \tag{7}$$

$$VP_Q^{90^\circ} = [p(90°), p(135°), p(0°), p(45°)], \tag{8}$$

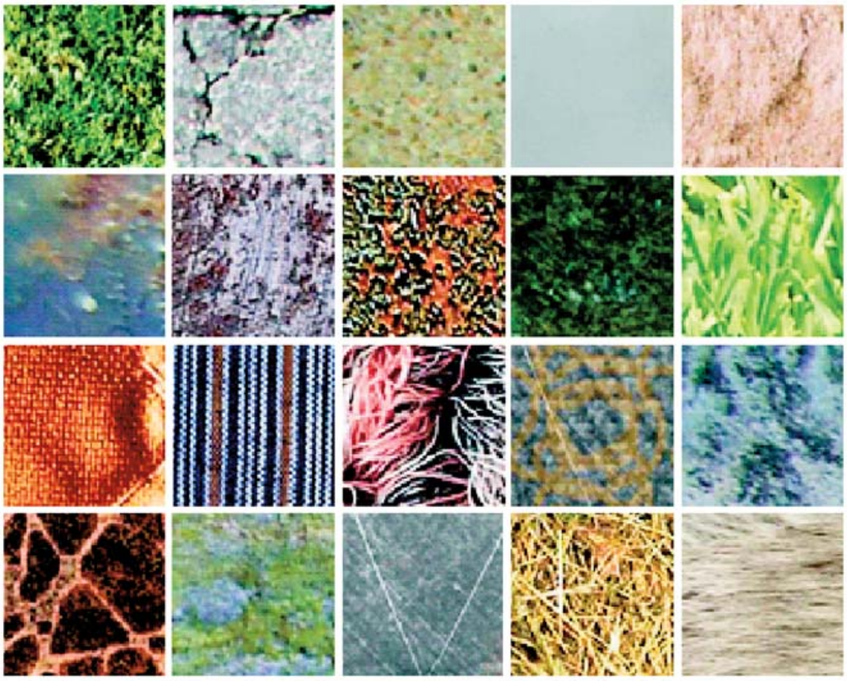


Fig. 3 Example images from texture database

$$VP_Q^{135°} = [p(135°), p(0°), p(45°), p(90°)], \tag{9}$$

where VP_Q is feature vector of retrieved texture and p is feature value in rotation 0°, 45°, 90° or 135°.

The equation for computing the minimal vector's distance d is defined as [14]:

$$d = \min[\mathrm{dist}(VP_Q^{(n-1)\theta_\Delta}) \mid VP_{DB})], \tag{10}$$

where $n = (1, 2, 3, 4)$ is iteration number for rotation of feature vector, and θ_Δ is rotation step in degrees. In our experiments, the rotation step 45° was used. The V and VP_{DB} are feature vectors of query image and database texture.

The percentage expression of image retrieval precision p is provided by a simple method defined as follows:

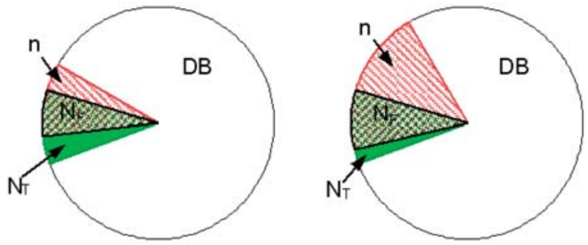
$$p = \frac{N_F - 1}{N_T - 1} * 100\%, \tag{11}$$

where N_F is number of truly found texture images and N is total number of relevant texture images. The subtraction of 1 is used, because there was a query image in database.

For a better view, the extra measure p^* was used for evaluationof image retrieval. The p^* expresses the precision of retrieval on first $n = 2{*}N_T$ positions. Thereby, it is possible to show if there are any images close to correct match. The principle of evaluation is shown in Fig. 4.

In our experiments where 20 images for each texture were used, $n = N_T = 20$ for p^* and $n = 2{*}\ N_F = 40$ for p^* .

Two simple applications for annotation and retrieval were created. The first application is used for creating the annotated database where the database of images is the input and extracted feature vectors, exported into XML file, are the output. The query image and the annotated image database are inputs for a second application. Output of this application is a sorted set of images by minimal distance (10).



a) b)

Fig. 4 Experiments evaluation, a) precision p, b) precision p* (DB is image database, N_F is number of truly found texture images, N_T is total number of relevant texture images, n expresses first n-positions of sorted set of images for evaluation)

4.1 Experimental results

The results of GLCM, 1D-GLCM, RGB-1D-GLCM and 1D-CLCM for 110 random query images are shown in Fig. 5. The average values of image retrieval precision are shown in Fig. 6. The conventional GLCM reached in our experiments precision $p = 45.7\%$ and $p' = 59.9\%$. The 1D-GLCM reached only $p = 38.0\%$ and $p' = 51.6\%$. After applying the 1D-GLCM on separated color channels, the better result were obtained. The RGB-1D-GLCM method reached $p = 56.5\%$ and $p' = 69.0\%$. Finally, the 1D-CLCM

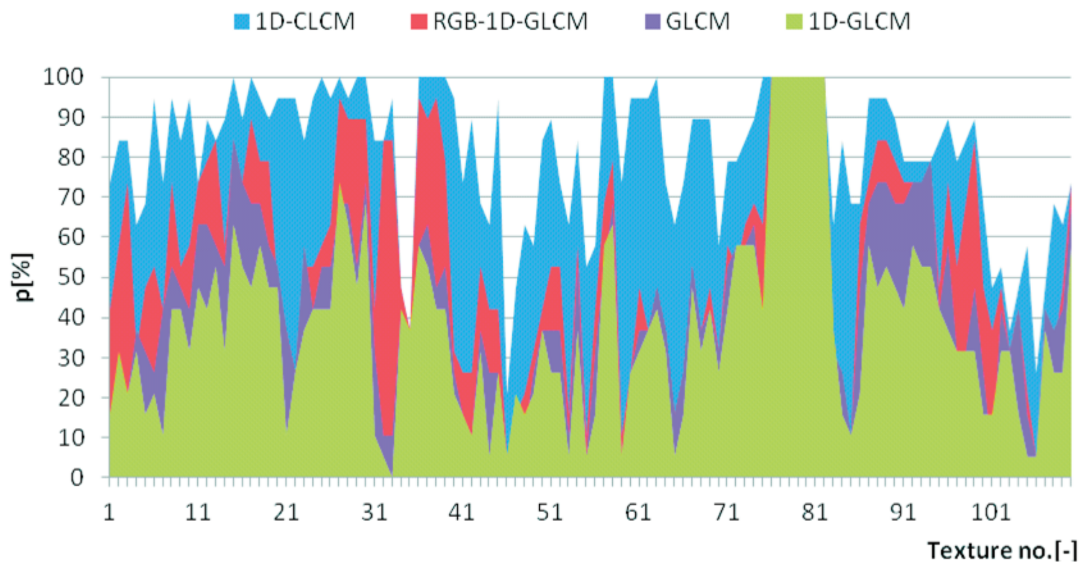


Fig. 5 The experimental results for 110 retrieved image textures

method reached $p = 81.1\%$ and $p' = 90.1\%$. Thus, we reached improvement of 43.1% over basic 1D-GLCM and 35.4% over GLCM, after color information wasapplied. These results make 1D-CLCM fast and strong descriptor for color textures.

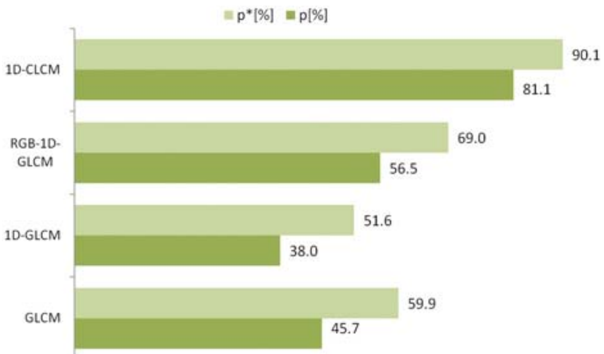


Fig. 6 The final experimental results for GLCM, 1D-GLCM, RGB-1D-GLCM and 1D-CLCM methods

5. Conclusion

In this paper, the methods for color texture feature extraction based on GLCM were proposed and tested. Exactly, we aimed our research to modification of one-dimensional GLCM and 1D-CLCM method was introduced. The experiments verified on the database of 2600 color images were provided. At first, the comparison of conventional GLCM, 1D-GLCM, applied on a separated channel of color image was presented. Next, we used our previous research on GLCM and modified this method for direct extraction from color image. The 1D-CLCM method was introduced and the precision of retrieval was rapidly increased. The designed 1D-CLCM method reached about 81% in texture retrieval with improvement of 43% over basic 1D-GLCM and 35% over GLCM were obtained.

Acknowledgements

This contribution/publication is the result of the project implementation: Centre of excellence for systems and services of intelligent transport, ITMS 26220120050 supported by the Research & Development Operational Programme funded by the ERDF.

References

[1] ASKOY, S., HARALIC, R., M.: Using Texture in Image Similarity and Retrieval, Texture Analysis in Machine Vision, M. Pietikainen, Ed., *World Scientific,* Singapore, vol. 20, pp. 129–149, 2000.

[2] PASCHOS, G.: Perceptually Uniform Color Spaces for Color Texture Analysis: an Empirical Evaluation, *Image Processing, IEEE Transactions,* vol. 10, No. 6, pp. 932–937, 2001, doi: 10.1109/83.923289.

[3] MEI, Y., ANDROUTSOS, D.: *Color Texture Retrieval Using Wavelet Decomposition on the Hue/Saturation Plane,* Multimedia and Expo, 2008 IEEE Intern. Conference, pp.877–880, 2008.

[4] ASSEFAA, D., MANSINHAB, L., TIAMPOB, K. F., RASMUSSENC, H., ABDELLAD, K.: Local Quaternion Fourier Transform and Color Image Texture Analysis, *Signal Processing,* vol. 90, No. 6, 2010, pp. 1825–1835, ISSN 0165-1684, DOI: 10.1016/j.sigpro 2009.11.031.

[5] MUTASEM, K., S., A., KHAIRUDDIN, B., O., SHAHRUL, A., N., ALMARASHDAH, I.: Fish Recognition Based on Robust Features Extraction from Color Texture Measurements Using Back-propagation Classifier, *J. of Theoretical and Applied Information Technology,* vol. 18. No. 1, 2010, Paper ID: 1401 -JATIT-2K10.

[6] PABAMALIE, L., A., I., PREMARATNE, H., L.: *A Grain Quality Classification System,* Inter. Conference on Information Society (i-Society), London, pp. 56–61, 2010, ISBN: 978-1-4577-1823-6.

[7] HOSSAIN, K., PAREKH, R.: *Extending GLCM to Include Color Information for Texture Recognition*, Intern. Conference on Modeling, Optimization and Computing, Book Series: AIP Conference Proc., vol. 1298, pp. 583–588, 2010, DOI: 10.1063/1.3516370, ISSN: 0094-243X, ISBN: 978-0-7354-0854-8.

[8] BENCO, M., HUDEC, R.: Novel Method for Color Texture Features Extraction Based on GLCM, *Radioengineering,* vol. 16, No. 4, pp. 64–67, ISSN 1210-2512, 2007.

[9] TOU, J., Y., TAY, Y., H., LAU, P., Y.: One-dimensional Grey-level Co-occurrence Matrices for Texture Classification, *Information Technology,* 2008, ITSim 2008, Intern. Symposium, vol. 3, pp.1–6, 2008.

[10] HARALICK, R., M.: *Statistical and Structural Approaches to Texture,* Proc. of the IEEE, 67, pp. 786–804, 1979.

[11] HARALICK, R. M.; SHANMUGAM, K.; Dinstein, Its''Hak, "Textural Features for Image Classification," Systems, Man and Cybernetics, *IEEE Transactions on,* vol. SMC-3, No. 6, pp.610, 621, Nov. 1973, MITMediaLab,Vision texture, VisTexdatabase, http://wwwhite. media.mit.edu/vismod/imagery/VisionTexture/vistex.html, 1995

[12] MITMediaLab,Vision texture, VisTexdatabase, http://wwwhite.media.mit.edu/vismod/ imagery/VisionTexture/vistex.html, (1995)

[13] OJALA T. at al.: *Outex-new Framework for Empirical Evaluation of Texture Analysis Algorithms,* Proc. of the 16th Intern. Conference on Pattern Recognition, vol. 1, Quebec, 2002, pp. 701–706, http://www.outex.oulu.

[14] ZACHARIASOVA, M., HUDEC, R., BENCO, M., KAMENCAY, P., LUKAC, P., MATUSKA, S.: The Effect of Metric Space on The Results of Graph Based Colour Image Segmentation, *Communications - Scientific Letters of the University of Zilina,* vol. 14, No. 3, 2012, ISSN 1335-4205.

Peter Dubravka – Pavol Rafajdus – Pavol Makys – Valeria Hrabovcova – Lorand Szabo *

ANALYSIS OF SWITCHED RELUCTANCE MOTOR BEHAVIOR UNDER ELECTRICAL FAULT CONDITIONS

This paper deals with analysis of Switched Reluctance Motor (SRM) behavior under fault conditions. The static parameters of the healthy SRM and SRM with defined faults of phases are investigated by Finite Element Method (FEM). These static parameters are used in dynamic simulation of the SRM. Transients are simulated for the healthy motor and for the motor under fault conditions. Results of the simulations are compared with experimental measurements.

Keywords: *Switched reluctance motor, fault operation, parameter analysis*

1. Introduction

The Switched Reluctance Motor (SRM) is one of the simplest and robust electrical machines. Torque of SRM is produced by tendency of rotor to move to a position where the inductance of the excited winding is maximized [1].

The SRM construction is very simple. It is a double salient electrical machine which has only stator winding and rotor is passive. Each phase comprises two coils wound on opposite stator poles and connected in series or parallel [2]–[4]. The cross-section area of the three phase 6/4 SRM is shown in Fig. 1a. In Fig. 1b, there is a photo of the investigated SRM.

In many applications, the drive reliability is a very important task from the point of view of fault operation. There are several types of fault states which could occur during the SRM operation. Mechanical, magnetic or electrical faults of the motor could occur. This paper is focused on electrical faults of SRM.

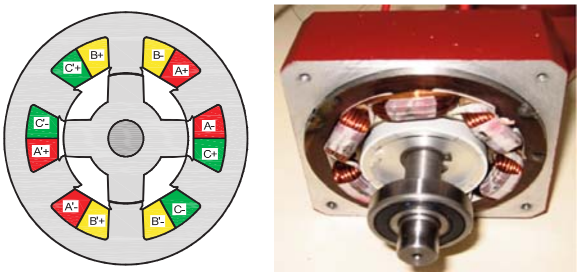
The electrical faults could be: short circuit in one coil of a phase (all turns or some turns), a whole coil is bridged by a short circuit, the whole phase is short circuited, open circuit in one coil of a phase, a short circuit between two different phases, a short circuit from one winding to ground [5], [6].

In [7] electrical faults detection and special fault tolerant design of the 12/14 SRM are studied, where the winding scheme is a six phased duplex type (each phase is doubled). When a fault of one channel occurs the second channel still contributes to the torque generation. Some simulations of the winding faults are presented.

In [8] a fault-tolerant SRM drive with adaptive fuzzy logic controller providing smooth torque with minimum ripple under normal and fault conditions is described.

Some fault states are analyzed in [9]. In this paper the static and dynamic faults of real 6/4 SRM are investigated. The static analysis is carried out by means of FEM and the phase inductance, flux linkage and air gap torque are analyzed under normal and fault conditions. These static parameters are used in the dynamic simulation of the SRM and transients are calculated also for health and fault operated motor. The simulated results are compared with experimental ones.

The nameplate of the investigated SRM is shown in Table 1.



a) *b)*

Fig. 1 The investigated SRM, a) cross-section area, a photo of open motor

The nameplate of the investigated 6/4 SRM Table 1.

Rated voltage	Rated current	Rated speed
3x 10V	28.5A	5000 rpm

---

* **Peter Dubravka[1], Pavol Rafajdus[1], Pavol Makys[1], Valeria Hrabovcova[1], Lorand Szabo[2]**
[1] Department of Power Electrical Systems, Faculty of Electrical Engineering, University of Zilina, Slovakia, E-mail: pavol.rafajdus@kves.uniza.sk
[2] Department of Electrical Machines and Drives, Technical University of Cluj-Napoca, Cluj-Napoca, Romania

2. FEM analysis of the SRM static parameters

There are several methods how to analyze the static parameters of the SRM [10] and [11]. Very useful and accurate is Finite Element Analysis (FEA) which is used in this paper. For the FE magnetostatic analysis the following input data are needed: geometrical dimensions of the machine, current density of one phase, material constants (winding conductivity and relative permeability, B-H curve of SRM ferromagnetic circuit material) and boundary conditions. The parametric model of SRM was created in LUA script in the FEMM 4.2 software for more convenient calculation with lower time consumption. The LUA script also enables the SRM model configurations (changes of the machine geometry or other parameters) during the execution of the script.

The static characteristics of the SRM were analyzed by means of FEM under four different winding fault conditions. Namely 0%, 20%, 50%, 70% of winding turns are in the short circuit that corresponds to 25, 20, 13, 8 winding turns respectively.

The accuracy of the result depends on the size of FEM mesh and accuracy of the input parameters. In this model, 11578 nodes were used. The calculation was carried out for each individual rotor position and current under static condition. The rotor position Θ was moved from unaligned Θ_u to aligned position Θ_a with step of 1° and in each position the current was changed within its working range from 1 to 28 A. In Fig. 2 the distribution of magnetic flux lines of healthy SRM for aligned and unaligned position can be seen. For comparison, in Fig. 3, there is a distribution of magnetic flux lines of fault SRM for aligned and unaligned position if 50% of coil "B" turns are in short circuit.

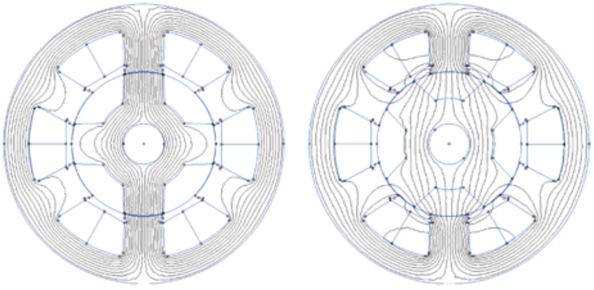

a) b)

Fig. 2 FEM analysis of the SRM, flux lines of the healthy SRM, a) aligned rotor position, b) unaligned rotor position

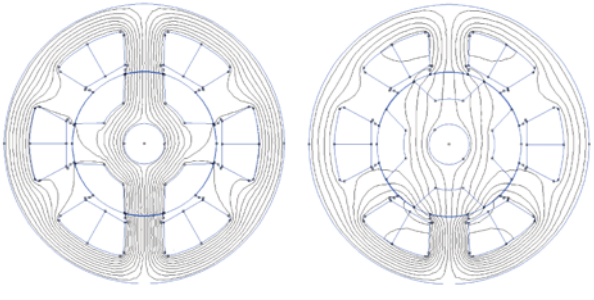

a) b)

Fig. 3 FEM analysis of the SRM, flux lines of the SRM, where 50 % of coil "B" turns are in short circuit; a) aligned rotor position, b) unaligned rotor position.

Magnetic flux linkage calculation

The first parameter to be analyzed is the flux linkage versus phase current for different rotor position $\psi = f(I,\Theta)$. The area bounded by maximal phase current and by both $\psi - I$ curves for aligned and unaligned positions is equal to mechanical energy, which is converted to electromagnetic force [1]. In Fig. 4 the flux linkage for various rotor positions and currents can be seen. In Fig. 5 the curves obtained by means of FEM for different rotor positions can be seen, if the phase current is kept constant 13 Amps (0° is equal to unaligned rotor position and 45° is equal to aligned rotor position). The investigated phase is B. As it can be seen from this Fig. 5, the flux linkage is decreased with increasing phase turns in short circuit.

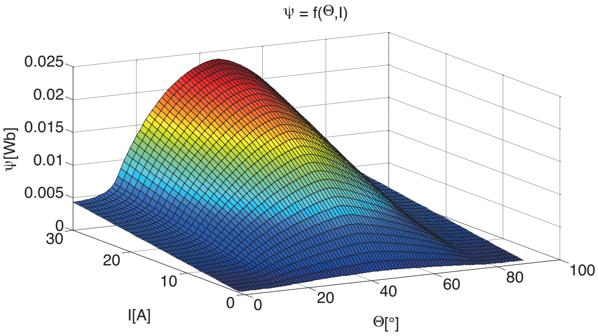


Fig. 4 FEM analysis of the SRM, flux linkage for various rotor positions and currents

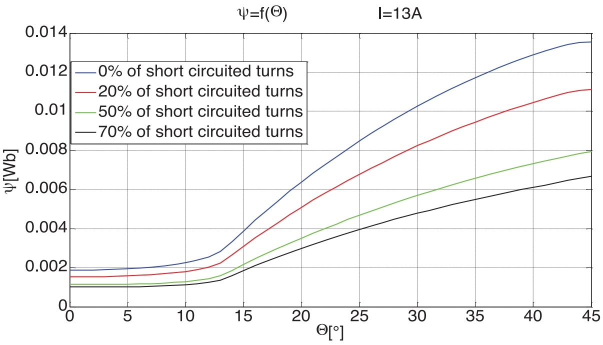


Fig. 5 FEM analysis of the SRM, flux linkage versus rotor position for constant phase current 13 A

A. Phase inductance calculation

The phase inductance $L = f(I,\Theta)$ versus rotor position for full current range is a static parameter which is needed in SRM mathematical model for dynamic simulations. The analysis was made for the whole working range. The phase inductance profiles are shown in Fig. 6. The results are obtained by means of FEM for SRM in accordance with the following equation:

$$L = \frac{\int A \cdot J dV}{I^2}. \qquad (1)$$

where A is magnetic vector potential, J is current density, V is volume and I is phase current.

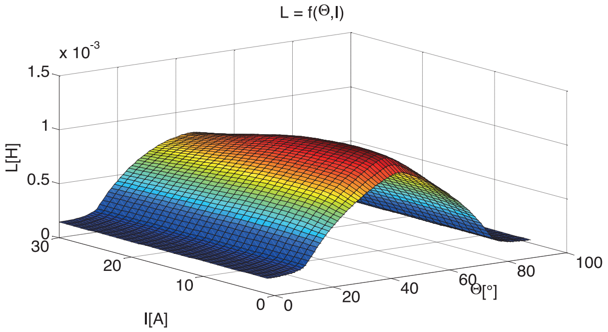


Fig. 6 FEM analysis of the SRM, the phase inductance for various rotor positions and various phase currents

The phase inductance for various rotor position and current 13A under different winding faults can be seen in Fig. 7. The phase inductance is decreased with increasing phase turns in short circuit.

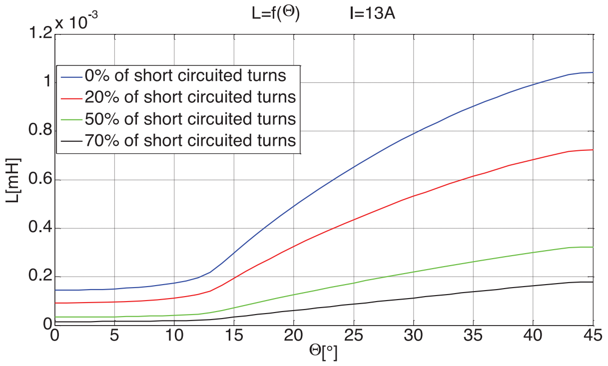


Fig. 7 FEM analysis of the SRM, phase inductance versus rotor position for constant phase current 13 A

B. Electromagnetic torque calculation

The electromagnetic torque was also calculated by means of FEM. The static torque characteristics were obtained for the whole working range. Maxwell's stress tensor prescribes the torque per unit area produced by the magnetic field in the air gap of the motor. The differential torque produced is:

$$dT = \frac{1}{2}\left(H(B \cdot n) + B(H \cdot n) - (H - B)n\right). \qquad (2)$$

where n denotes the direction normal to the surface at the point of interest, B is flux density and H is intensity of magnetic field.

The static electromagnetic torque of phase B in the air gap T_δ for various rotor positions and various phase currents is shown in Fig. 8.

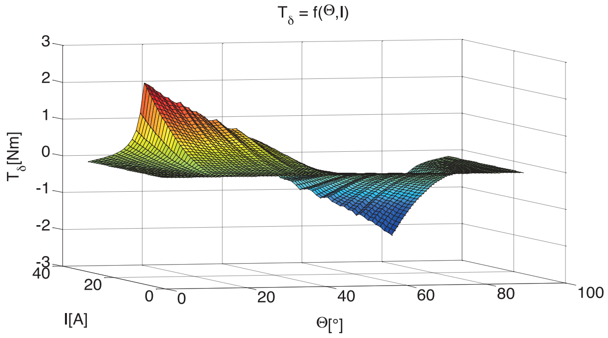


Fig. 8 FEM analysis of the SRM, the static electromagnetic torque of phase in the air gap T_δ for various rotor positions and various phase currents

The static electromagnetic torque of phase B in the air gap T_δ for various rotor positions and constant phase current 13A under different winding faults is shown in Fig. 9. The static electromagnetic torque in the air gap is decreased with increasing phase turns in short circuit.

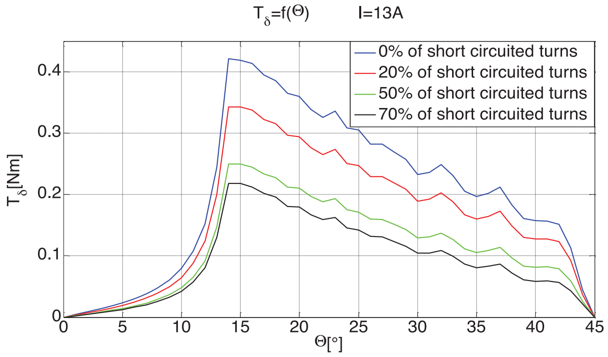


Fig. 9 FEM analysis of the SRM, the static electromagnetic torque of phase B in the air gap T_δ for various rotor positions and constant phase current 13 A

3. Mathematical model of SRM

As it is known the mathematical model of the SRM needs input static parameters to be able to solve transients. The static parameters: flux linkage, phase inductance, developed electromagnetic torque of SRM are obtained by means of FEM and they are used in this mathematical model to calculate following values: phase current, speed, voltage and dynamic torque of the SRM during normal operation and also during fault operation.

As it is known the electromagnetic torque of SRM can be calculated from:

$$T_e = \frac{\partial \int_0^i \Psi \, di}{\partial \Theta} \tag{3}$$

where i is phase current, Ψ is phase flux linkage and Θ is rotor position. To obtain dynamic total torque of SRM, the simulation of its mathematical model is needed.

The mathematical model of SRM consists of the following equations, if:

- leakage inductances between phases are neglected,
- iron losses are neglected,
- phase inductance depends on phase current and rotor position.

The voltage equation of one SRM phase is given as:

$$v = Ri + \frac{d\psi}{dt}. \tag{4}$$

where R is phase resistance, i is phase current and ψ is flux linkage. The flux linkage depends on both parameters: phase current and rotor position ($\psi = f(i, \Theta)$). Then

$$\frac{d\psi}{dt} = \frac{\partial \psi}{\partial i}\frac{di}{dt} + \frac{\partial \psi}{\partial \Theta}\frac{d\Theta}{dt}. \tag{5}$$

The phase current is calculated from combination of (4) and (5) as:

$$\frac{di}{dt} = \frac{v - \left(R + \dfrac{dL(i,\Theta)}{d\Theta}\omega\right)i}{L(i,\Theta)}. \tag{6}$$

The real angular speed is calculated from equation:

$$\frac{d\omega}{dt} = \frac{1}{J}\left(\sum_{j=1}^{m} T_j(\Theta, i) - T_{load}\right). \tag{7}$$

where J is the moment of inertia and T_{load} is load torque.

The SRM is controlled on the base of rotor position Θ, therefore it is as follows

$$\Theta = \int \omega dt. \tag{8}$$

4. Dynamic simulation of SRM

Simulation model was created under Matlab. The model of a single phase is shown in Fig. 10.

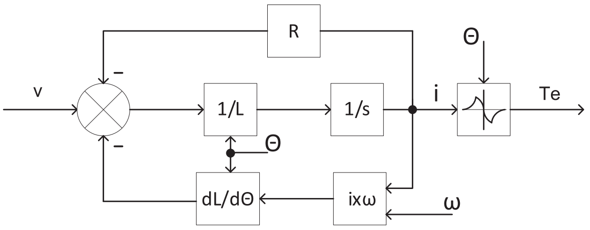


Fig. 10 Dynamic model of the SRM, one phase is shown only

To control rotor speed, PI controller is used. The PI controller is implemented in the discrete version, therefore, the control signal $u(k)$ is given as:

$$u(k) = K_p\left[e(k) + \frac{T}{T_i}\sum_{i=0}^{k-1} e(i)\right]. \tag{9}$$

where k is a discrete time instant and T is sampling time.

The hysteresis controller is used to control phase current. The phase voltage in (4) is given by DC voltage source, which supplied the converter of SRM. By this control the voltage can have three values, if the voltage drops on the transistors and diodes are neglected:

a) $v = +V_s$, if the phase current i is lower than low limit of the hysteresis region,
b) $v = 0$, if the phase current is higher than high limit of the hysteresis region,
c) $v = -V_s$, if the phase current is switched off.

The control structure which is used in order to simulate SRM's transient is shown in Fig. 11. The coefficients of PI controller are tuned by a trial and error method.

The SRM investigation was made for different speeds.

In dynamic simulation, the start up of the SRM was simulated for three required speeds. In the first case the start up of a healthy

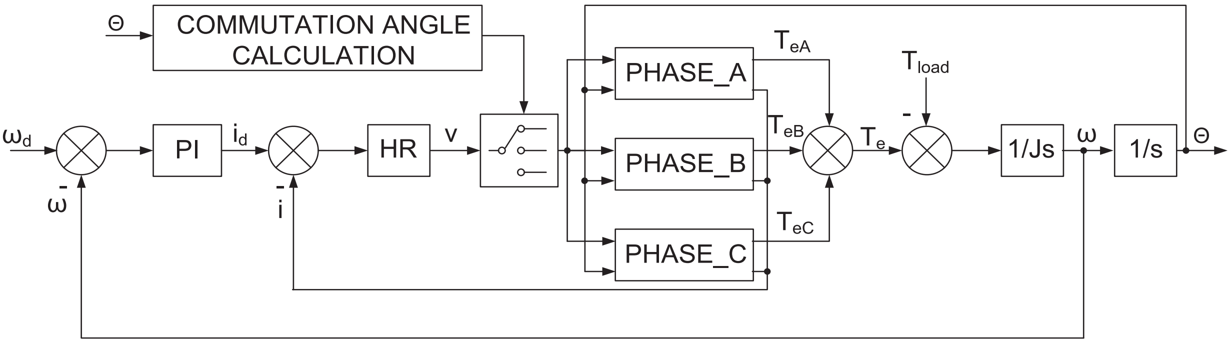


Fig. 11 The block diagram of SRM speed control with PI controller

motor was simulated up to time 0.11 (see Fig. 12) and then, the phase A fallout occurred. It means that the phase A is not active. The load torque was kept constant 0.1 Nm and required speed was 180 rpm. After the phase fallout occurs, the speed decreases under the required speed value and speed ripple occurred about 17% from the required speed. Currents in three phases up to time 0.1 can be seen in Fig. 13. After this time the phase A fallout occurred (the phase current is 0A). The phase current of phase B is higher than in phase C, which can be caused by the switch ON and switch OFF angles setup.

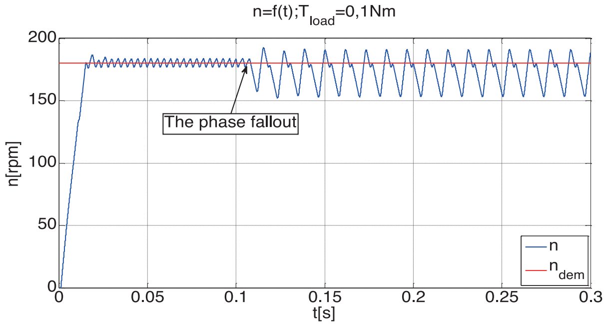


Fig. 12 Dynamic simulation of SRM speed with n_{dem} = 180 rpm, the start up of healthy motor and the phase A fallout occurred

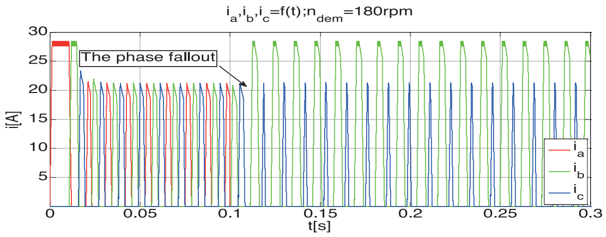


Fig. 13 Dynamic simulation of SRM currents with n_{dem} = 180 rpm, the start up of healthy motor and the phase A fallout occurred

In Fig. 14, there is start up of the motor up to 300 rpm with three phases and in time $t = 0.1$ s A phase fault occurred. After the phase fallout occurs, the speed decreases under the required speed value and speed ripple occurred about 5% from the required speed. Currents in three phases up to 0.1 s can be seen in Fig. 15. After this time the phase A fallout occurred (the phase current is 0A). The phase current of phase B is higher than in phase C, which can be caused by the switch ON and switch OFF angles setup.

In Fig. 16, there is start up of the motor up to 1000 rpm with three phases and in time $t = 0.12$ s one phase fault occurred. After the phase fallout occurs, the speed decreases under the required speed value and speed ripple occurred about 2% from the required speed. Currents in three phases can be seen in Fig. 17.

As it can be seen from these analyses, when the speed increases, the speed ripple decreases. It is given by the inertia of the system.

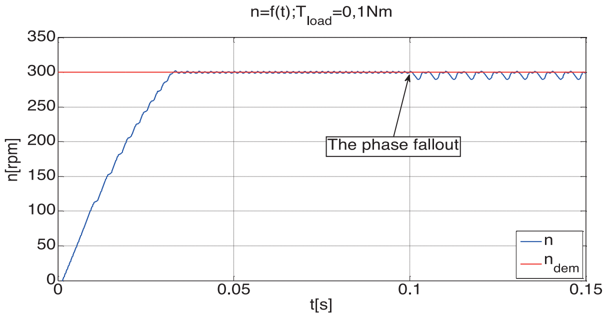


Fig. 14 Dynamic simulation of SRM speed with n_{dem} = 300 rpm, the start up of healthy motor and the phase A fallout occurred

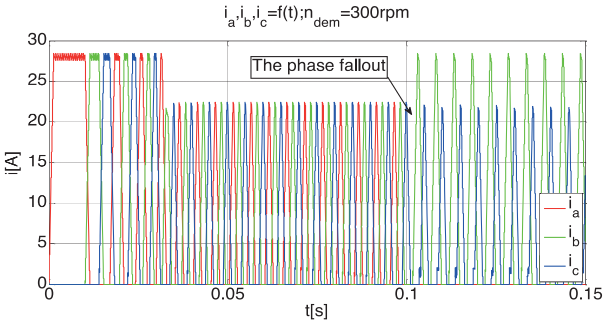


Fig. 15 Dynamic simulation of SRM currents with n_{dem} = 300 rpm, the start up of healthy motor and the phase A fallout occurred

This was simulated for torque 0.1 Nm and speed up to 1000 rpm. In the future it will be analyzed for rated speed and rated torque.

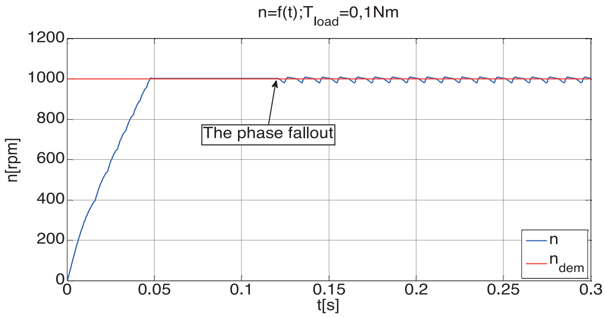


Fig. 16 Dynamic simulation of SRM speed with n_{dem} = 1000 rpm, the start up of healthy motor and the phase A fallout occurred

5. Experimental results

To verify the simulated dynamic result an experimental test rig with SRM coupled with BLDC load was carried out. The test rig is shown in Fig. 18. The control board equipped with a Digital Signal Controller by Freescale was used [12] and [13]. The control board is connected with low power and a low voltage converter which supplies the SRM. The control loop consists of a speed

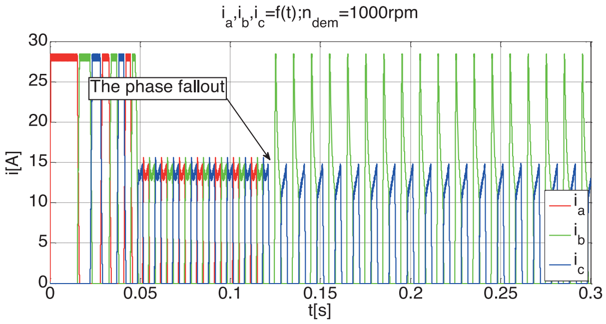


Fig. 17 Dynamic simulation of SRM currents with $n_{dem} = 1000$rpm, the start up of healthy motor and the phase A fallout occurred

loop with PI controller and a slave current loop made by hysteresis current controllers.

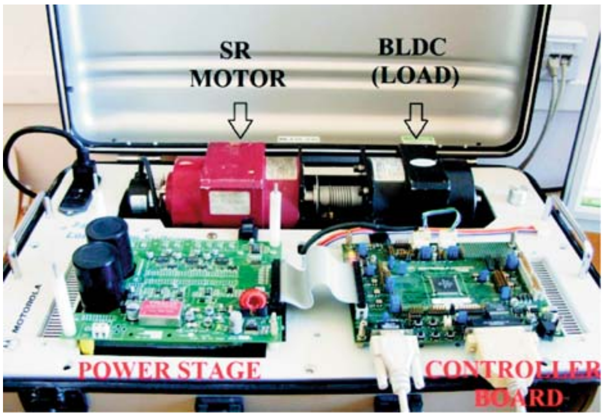


Fig. 18 Experimental test rig equipped with SRM and BLDC load

The measured speed start up of the healthy SRM is shown in Fig. 19. The load torque was kept constant 0.1 Nm.

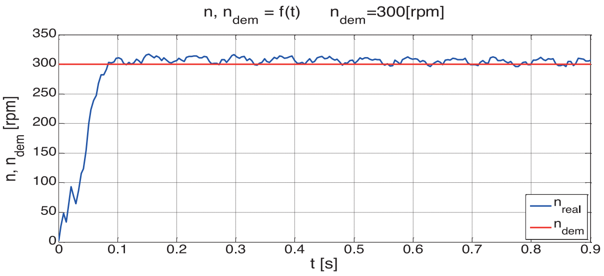


Fig. 19 The measured speed start up of the healthy SRM.

The measured phase currents for start up and steady state of the healthy SRM are presented in Fig. 20.

The measured start up of the healthy SRM where the required speed is 1000 rpm is shown in Fig. 21.

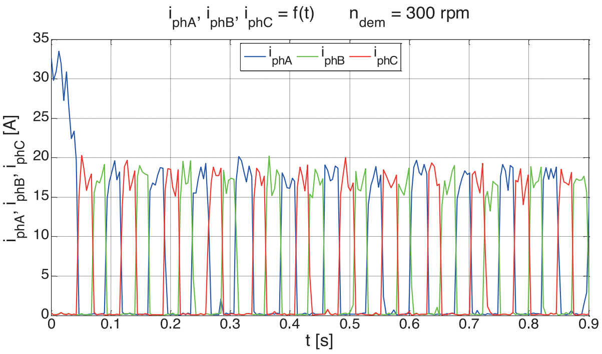


Fig. 20 The measured phase currents for start up of the healthy SRM

The measured phase currents for start up and steady state of the healthy SRM for the required speed 1000 rpm are shown in Fig. 22.

The start up of the healthy SRM and fallout of one phase in time 1.9 s, for the required speed 180 rpm are shown in Fig. 23. It means that one phase was un-plugged from the motor.

The measured phase currents for start up of the healthy SRM for the required speed 180 rpm and in time 1.9 s and fallout of one phase are shown in Fig. 24.

The speed and currents are filtered with an exponential filter to remove the noise.

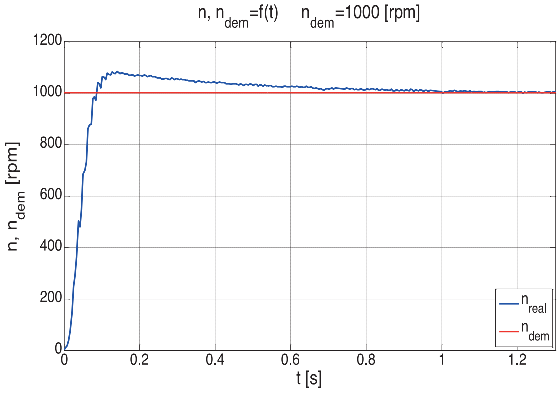


Fig. 21 The measured speed start up of the healthy SRM at 0.1 Nm

6. Conclusion

In this paper the static and dynamic analyses of the SRM fault operations are presented. The SRM static parameters were calculated by means of FEM for a healthy and short circuit phase. The mathematical model of the SRM was derived and used for dynamic simulation. The dynamic health and the phase fallout of the motor are analyzed by simulations and the results are compared with the measured ones. There are some differences caused by lower accuracy of the switch ON and switch OFF angles setup in simulation

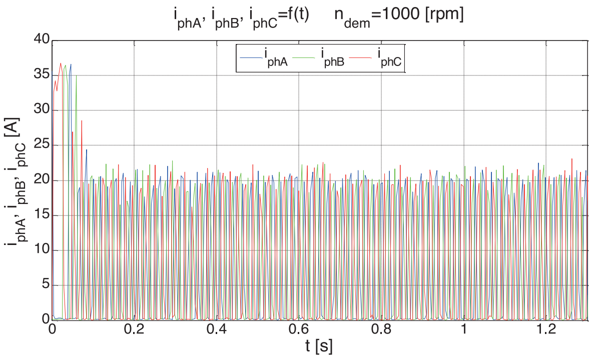


Fig. 22 The measured phase currents for start up of the healthy SRM

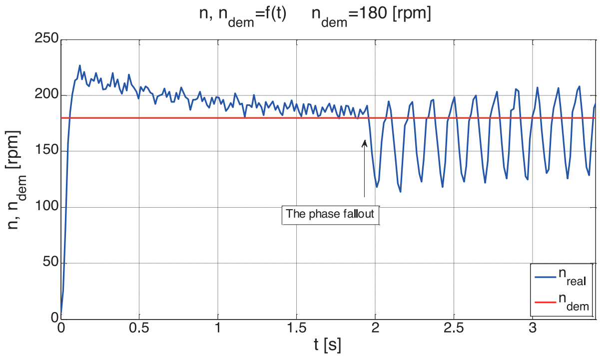


Fig. 23 The start up of the healthy SRM and fallout of one phase

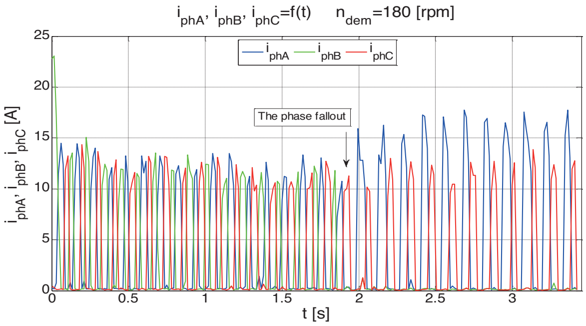


Fig. 24 The measured phase currents for start up of the healthy SRM and fallout of one phase

and in real SRM drive. This will be solved in the future also with higher speed and torque and for some other faults in the SRM.

On the base of this analysis verification, the simulation model can be used for other fault operation investigation and some recommendations can be done.

Acknowledgment

This work was supported by the Slovak Research and Development Agency under the Contract No. SK-RO-0016-10 and VEGA No. 1/0940/13 and R&D operational program Centre of excellence of power electronics systems and materials for their components II. No. OPVaV-2009/ 2.1/02-SORO, ITMS 26220120046 funded by European regional development fund (ERDF). Authors also like to thank the Freescale Semiconductors in Czech Republic for their support.

References

[1] MILLER, T. J. E.: *Electronic Control of Switched Reluctance Machines.* Oxford (U.K.) : Newnes, 2001.

[2] MILLER, T. J. E.: Switched Reluctance Motors and their Control, *Magna Physics,* 1992.

[3] PYRHONEN, J., JOKINEN, T., HRABOVCOVA, V.: *Design of Rotating Electrical Machines,* John Wiley & Sons, 2008.

[4] KRISHNAN, R.: *Switched Reluctance Motor Drives - Modeling, Simulation, Analysis, Design, and Applications,* CRC Press LLC, FLA, USA, 2000.

[5] SCHINNERL, B., GERLING, D.: *Analysis of Winding Failure of Switched Reluctance Motors,* Proc. of the IEEE Intern. Electric Machines and Drives Conference (IEMDC '09), Miami, pp. 738-743.

[6] TEREC, R., BENTIA, I., RUBA, M., SZABO, L., RAFAJDUS, P.: *On the Usefulness of Numeric Field Computations in the Study of the Switched Reluctance Motor's Winding Faults,* ISCIII 2011 - 5th Intern. Symposium on Computational Intelligence and Intelligent Informatics, art. no. 6069753, pp. 117-120.

[7] SZABO, L., RUBA, M.: *On Fault Tolerance Increase of Switched Reluctance Machines,* EUROCON 2009, pp. 18-23, May 2009

[8] MIR, S., ISLAM, M.S., SEBASTIAN, T., HUSAIN, I.: Fault-Tolerant Switched Reluctance Motor Drive Using Adaptive Fuzzy Logic Controller, *Power Electronics,* IEEE Transactions on, March 2004

[9] KACENKA, V., RAFAJDUS, P., MAKYS, P.,VAVRUS, V., SZABO, L.: *Static and Dynamic Fault Analysis of Switched Reluctance Motor,* Intern. conference ELEKTRO 2012, Rajecke Teplice, 2012, Slovakia, IEEE Xplore

[10] RAFAJDUS, P., HRABOVCOVA, V., HUDAK, P.: *Investigation of Losses and Efficiency in Switched Reluctance Motor,* EPE-PEMC 2006, Portoroz, Slovenia

[11] RAFAJDUS, P., SEKERAK, P., KALAMEN, L., HRABOVCOVA, V.; CHEN, H.: Static and Dynamic Analysis of Linear Switched Reluctance Machine, *Communications - Scientific Letters of the University of Zilina,* 4/2011

[12] DIRENZO, M. T.: *Switched Reluctance Motor Control - Basic Operation and Example Using the TMS320F240,* Application Report, Texas Instruments, 2000

[13] VISINKA, R., BALAZOVIC, P.: *3-Phase Switched Reluctance Motor Control with Encoder Using DSP56F80x*, MOTOROLA Inc., 2002.

Daniel Benedikovic – Jan Litvik – Michal Kuba – Milan Dado *

SPECTRAL TRANSMISSION CHARACTERISTICS OF ANDVANCED AMPLITUDE MODULATION FORMATS

In this paper we focus our attention on numerical investigation of high-order amplitude modulation format in single-channel transmission systems utilizing different kinds of optical fibers. The aim is to show the spectral behavior of advanced amplitude modulated signals in view of fundamental parameters of fiber-optic system. The investigation is realized by solving the nonlinear Schrodinger equation (NLSE) through the pseudospectral split-step Fourier method (SSFM). The obtained results show that the advanced amplitude modulation formats are sensitive to nonlinear phenomenon and they are more desirable for using at short-haul networks.

Keywords: *Amplitude modulation format, nonlinear Schrodinger equation, split-step Fourier method, optical communication system.*

1. Introduction

In general, the implementation of novel kinds of modulation formats follows the goal of increasing the system capacity and improves the spectral efficiency of currently used fiber optic transmission systems. The most popular transmission modulation technique has been on-off keying (OOK) over the years, in which the optical power is modulated according to incoming sequence of bits [1, 2]. The major drawback of OOK format is that it enables to transmit only one information bit per symbol. On the other hand, the advanced modulation techniques allow transmission of several information bits per symbol and so they fulfill the mentioned requirements for the future fiber-optic communications [2, 3].

From a wide range of novel types of high-order modulation formats [1–3], our investigation is oriented to high-order amplitude formats, usually denoted as M-ary amplitude shift keying (*M*-ASK). This modulation format represents the simplest modulation technique from a point of view of transmitter structure [3]. *M*-ASK modulation scheme provides good choice for capacity enhancement of short-haul networks [4, 5] or serves for conversion application between non-return-to-zero (NRZ) and return-to-zero (RZ) format [6].

2. Theory

In general, the high-order modulated signals at the output of an optical transmitter have complex nature and can be described as follows [1, 7]:

$$E_m(t) = \Re e\left\{A_m(t)\cdot e^{j\omega_0 t}\right\} \tag{1}$$

where $E_m(t)$ denotes a modulated optical field, $A_m(t)$ is a modulated complex envelope with Gaussian shape and ω_0 is optical carrier frequency. The *M*-ASK modulated complex envelope of optical field can be expressed by [1, 3]:

$$A_m(t) = A_0\cdot\sum_{i=0}^{N_s} S_k^{ASK}\cdot A(t - i\cdot T_s) \tag{2}$$

where N_s is the number of transmitted symbols, A_0 is the envelope amplitude, T_s is symbol interval proportional to the symbol rate R_s and S_k^{ASK} represents a complex amplitude symbol chosen from a discrete alphabet of elements corresponding to the number of modulation states *M*.

Due to the complex nature of high-order modulation formats, its representation is always illustrated in a complex plane as the superposition of in-phase (*I*) and quadrature (*Q*) component. The constellation diagrams for *M*-ASK modulation formats are depicted in Fig. 1.

The fundamental equation for studying the pulse evolution in fiber–optic transmission system is so-called nonlinear Schrodinger equation NLSE [8], which includes the impact of various fiber degradation mechanisms on transmitted optical signals. The NLSE can be expressed in the following form [7, 8]:

$$\begin{aligned}\frac{\partial A_m}{\partial z} &= j\gamma|A_m|^2 A_m - \frac{\alpha}{2}A_m - \beta_1\frac{\partial A_m}{\partial t} - \\ &-j\frac{\beta_2}{2}\frac{\partial^2 A_m}{\partial t^2} + \frac{\beta_3}{6}\frac{\partial^3 A_m}{\partial t^3},\end{aligned} \tag{3}$$

where β_1 is related to the group velocity, β_2 is the group velocity dispersion (GVD) parameter, β_3 is the third-order dispersion parameter, α is the fiber attenuation and γ is the nonlinear coefficient

---

* **Daniel Benedikovic, Jan Litvik, Michal Kuba, Milan Dado**
Department of Telecommunications and Multimedia, Faculty of Electrical Engineering, University of Zilina, Slovakia,
E-mail: daniel.benedikovic@fel.uniza.sk

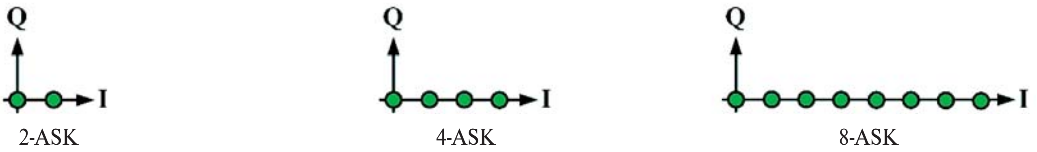


Fig. 1 Constellation diagrams of M-ASK modulations

of optical fiber. A_m is a modulated complex envelope carrying the encoded information in a given number of amplitude levels.

From a numerical point of view there are several options how to solve the NLSE, which only has the nontrivial solution. The most desirable numerical approach is so-called Beam Propagation Method (BPM) [7] in two widely used algorithmic expressions; the Finite Difference Method (FDM) [7, 9] and Split-Step Fourier method (SSFM) [3, 7, 8]. For our investigation, the SSFM technique was used. The SSFM is based on solving the NLSE in two separated parts according to the degradation effects. The linear and nonlinear impairments are solved in a small step in spectral and in time domain, respectively. The algorithmic implementation and operational principle of SSFM is shown in Fig. 2.

3. Results and discussion

The employing of advanced modulation formats into the fiber-optic system brings novel issues into the pulse propagation. After a modulation process, the optical pulses obtained new properties corresponding to the type of digital modulation and actual transmitted pulse. In context of *M*-ASK signals, this means that the transmitted pulses have different levels of amplitudes, so the propagation is quite different in comparison with the traditional power scheme OOK. The fact of various amplitude levels of transmitted pulses leads to the different sensitivity to the nonlinear phenomenon of self-phase modulation (SPM), whose impact on pulses is the limiting factor from the point of spectral broadening in a single-channel system.

In Fig. 3 are depicted the average values of spectral broadening depending on the fiber length of 2-ASK, 4-ASK and 8-ASK signals with input power $P_{in} = 1$ mW and bit rate $R_b = 10$ Gbps at $\lambda = 1550$ nm for two types of optical fibers; Standard single-mode fiber (SSMF) [10] and Dispersion shifted fiber (DSF) [11], respectively.

It is obvious that the value of spectral broadening is rapidly growing with a higher number of amplitude modulation levels. The worst results were obtained for 8-ASK format, for which the spectral broadening is significantly larger for both fiber types. For low-order amplitude levels (2-ASK and 4-ASK), the role of linear transmission properties (attenuation and chromatic dispersion) influence the impact of nonlinear SPM in sense of balancing the strength of signal degradation, especially after the effective fiber length, after which the impact of chromatic dispersion becomes the dominant factor of degradation and the growth of spectral broad-

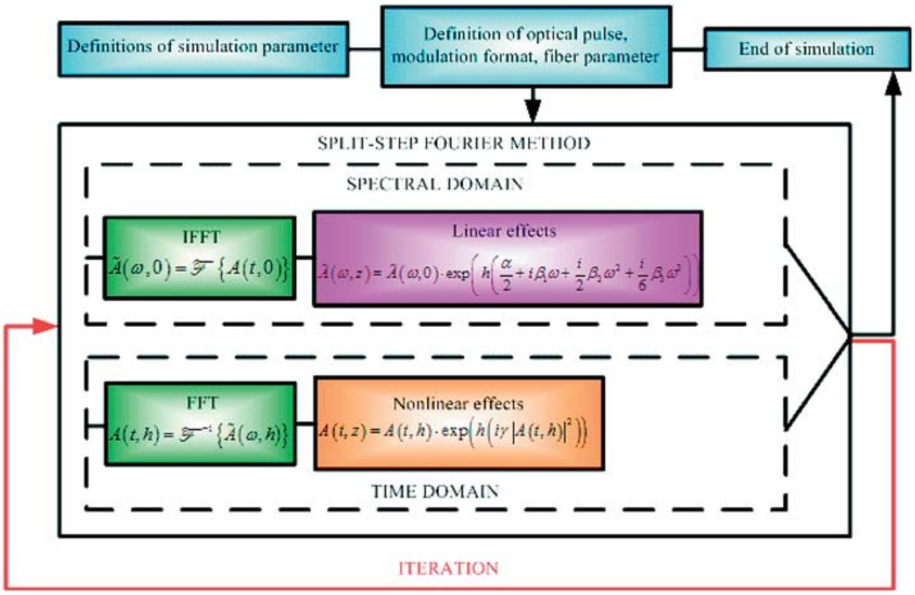
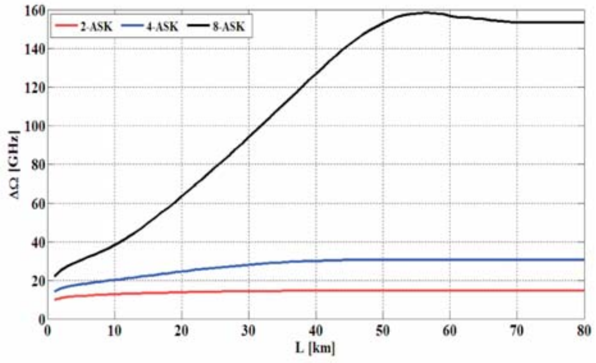
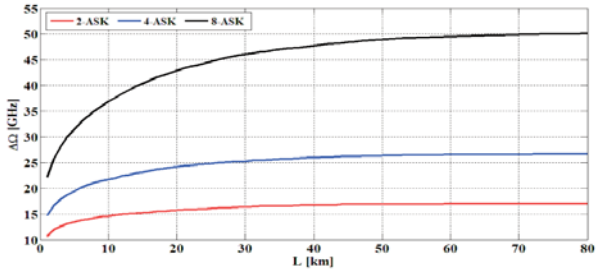


Fig. 2 Algorithm of split-step Fourier method

ening is significantly slower with a further increasing fiber length. From a transmission point of view, the DSF fibers exhibit considerably lower spectral broadening in comparison with SSMF fibers.



a) SSMF

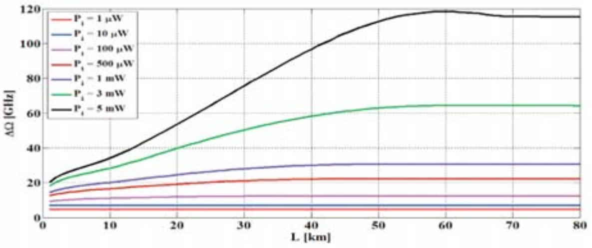


b) DSF

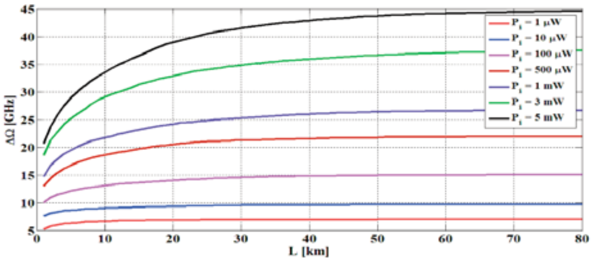
Fig. 3 Spectral broadening of different M-ASK signals

3.1 Effect of input power

In Fig. 4 can be seen the evolution of spectral broadening depending on the fiber length for various levels of input powers of 4-ASK signals with the same bit rate of system $R_b = 10$ Gbps at operating wavelength $\lambda = 1550$ nm for two types of optical fibers; SSMF 4a. and DSF 4b., respectively. The sequence of simulated symbol was generated with the same probability. It can be observed that for lower values of input powers, the spectral broadening is nearly constant or increases very slowly for both types of used optical fibers. The using of lower powers should be therefore recommended for implementation of *M*-ASK modulation in a fiber-optic system. The transmission symbols, which are encoded into the higher power levels represent larger and main contributions to the resulted effect of spectral broadening, when the higher value of power was launched into the fiber. We can also see that DSF fibers exhibit better robustness against nonlinear SPM degradation for a wider range of fiber lengths in comparison with SSMF fibers for the same input power levels. For this reason, the SSMF fibers should be used in a transmission system for shorter transmission distances with higher power levels. On the other hand, by using DSF fibers, it is possible to reach longer transmission distances with good spectral performance or the transmission system can operate with higher powers resulting in longer distance reach.

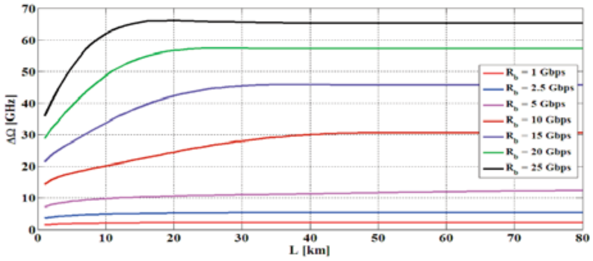


a) SSMF

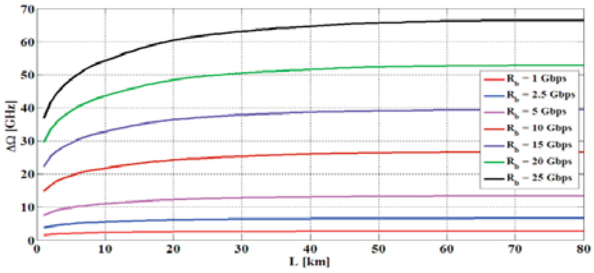


b) DSF

Fig. 4 Spectral broadening of 4-ASK modulation



a) SSMF



b) DSF

Fig. 5 Spectral broadening of 4-ASK modulation

3.2 Effect of bit rate

In Fig. 5 are illustrated the dependences of value of spectral broadening on the fiber length with various values of total bit rate Rb with the same moderate input power $P_{in} = 1$ mW for SSMF and DSF optical fibers.

From Fig. 5 an interesting fact can be observed. If the values of bit rates of an optical transmission system are higher, the spectral transmission performance of SSMF fibers is better. At lower bit rates (< 10 Gbps), the value of spectral broadening is nearly insignificant and with using longer fiber lengths, it increases in a slow way. On the other hand, at higher bit rates (> 10 Gbps), the resulted spectral broadening exhibits growing nature, but not as large as in the case of increasing power. The reason is that shorter optical pulses (higher bit rates) are more influenced by the effect of chromatic dispersion, which balances the impact of SPM. On the other hand, for DSF fibers this fact is not so obvious due to a very low (nearly zero) value of chromatic dispersion at an operating wavelength. From this point of view, the spectral transmission performances of SSMF and DSF are comparable.

For higher bit rates, the amplitude modulation formats provide suitable choice for implementation in a fiber-optic system, which is in good agreement with requirements for the next generation optical system.

4. Conclusion

In this paper we numerically investigated the M-ASK modulated signals. It was shown that the optical signals, which employ novel multilevel amplitude formats, are sensitive to the nonlinear effect of SPM in the sense of increasing the value of spectral broadening. The most promising and suitable amplitude format seems to be 4-ASK for SSMF and DSF fibers at moderate powers and bit rates. The power dependence of transmission pulses sets the significant limit for long-haul application, so the short-haul distance applications are preferred. On the other hand, by using the shorter duration of optical pulses, the desirable performance of a fiber-optic system in sense of balanced interaction between chromatic dispersion and SPM for both types of investigated fibers may be achieved.

Acknowledgement
This work has been supported by the Slovak Grant Agency under the project VEGA No. 1/1271/12 *("Studying the influence of degradation effects in physical layer on transmission mechanisms in high-speed optical networks").*

References

[1] PROAKIS, J. G.: *Digital Communications.* McGraw-Hill : New York 3rd ed., 2008
[2] WINZER, P. J., ESSIAMBRE, R.-J.: Advanced Modulation Formats for High-Capacity Optical Transport Networks. *IEEE/OSA J. of Lightwave Technology,* 24 (12), 2006
[3] SEIMETZ, M.: *High-order Modulation for Optical Fiber Transmission.* Springer Verlag Berlin : Heidelberg, 2009
[4] WALKIN, S., CONRADI, J.: Multilevel Signaling for Increasing the Reach of 10 Gb/s Lightwave Systems. *IEEE J. of Lightwave Technology,* 17 (11), 1999
[5] ZHAO, J, et al.: *Analytical Investigation of Optimization, Performance Bound, and Chromatic Dispersion Tolerance of 4-Amplitude-Shift-Keying Format.* Proc. of Optical Fiber Communication Conference (OFC), 2006
[6] YU, W., LU, D., WANG, D., LOU, C., HUO, L., PAN, S.: Proposal and Simulation Investigation of Optical Format Conversion between Quaternary Amplitude-Shift Keying Signals Based on Cascaded Modulators. *Optical Fiber Technology,* 2012
[7] OKAMOTO, K.: *Fundamentals of Optical Waveguides.* Academic Press San Diego 2nd ed., 2006
[8] AGRAWAL, G.P.: *Nonlinear Fiber Optics.* Academic Press New York 3rd ed., 2001
[9] JURISOVA, E., MULLEROVA, J.: Study of the Chalcogenide Nonlinear Fiber Bragg Gratings for All-Optical Switching, *Communications - Scientific Letters of the University of Zilina,* No. 2, 2012
[10] ITU-T RECOMMENDATION G. 652, *Characteristics of a single-mode optical fiber and cable*
[11] ITU-T RECOMMENDATION G. 653, *Characteristics of a dispersion-shifted single-mode optical fiber and cable.*

Oleg V. Chernoyarov – Martin Breznan – Alexander V. Terekhov *

RESTORATION OF DETERMINISTIC AND INTERFERENCE DISTORTED SIGNALS AND IMAGES WITH USE OF THE GENERALIZED SPECTRA BASED ON ORTHOGONAL POLYNOMIALS AND FUNCTIONS

Restoration algorithms of signals and images on the basis of their generalized spectra in bases of orthogonal polynomials and functions at absence and presence of random distortions have been examined. It is shown that in absence of hindrances the number of coefficients of the generalized spectrum of a restored signal (image) is determined by the desired approximation error at use of one or another metrics of functional space. If hindrances take place then there is an optimum number of coefficients of the generalized spectrum for signal (image) restoration. Working data of the proposed algorithms for various types of useful signals have been illustrated.

Keywords: *Signal and image restoration, generalized signal spectrum, orthogonal polynomials, Gauss quadrature formulas, approximation mean square error, maximum likelihood method, incomplete signal reception.*

1. Introduction

One of the methods of increasing throughput of multimedia data transmission, receiving and processing systems is optimization and improvement of coding (decoding) procedures of a signal source and redundancy reduction. At present two essentially different approaches for information compression of signals and images are used: 1) lossless compression; 2) controllable lossy compression [1, 2]. The first approach is based on principles of differential pulse code modulation, Huffman entropy coding, Lempel-Ziv-Welch method of repetitive chapter search on samples, as well as some others and it cannot provide essential compression of a data flow. More expedient is the coding procedure of a signal source with controlled losses. Within the scope of this approach various linear orthogonal transforms are applied: a) discrete cosine transform and its updatings; b) wavelet-transform; c) discontinuous piecewise constant function basis expansion (such as Walsh, Haar, S-transformation, etc.), however all of them are not flawless [2]. It causes need of new basic function search.

It is obvious that new algorithms of signal and image compression should be optimized not only in respect of computational burden reduction, complexity of hardware implementation but also in statistical sense taking into account probabilistic nature of hindrances, messages and performance measures. This problem is so severe under image transmission and processing as images have sufficiently great information capacity. In the present work it is shown that, as in [3, 4], the application of orthogonal polynomials or functions connected with them allows to receive effective, practically realizable procedures of signal and image restoration including presence of random distortions.

2. Signal and image restoration without noise influence

As systems of orthogonal polynomials (functions), the following polynomials (functions) can be chosen [5]

Legendre polynomials

$$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}(x^2-1)^n, \tag{1}$$

Chebyshev polynomials of the first kind

$$T_n(x) = \frac{(-2)^n n!}{(2n)!}\sqrt{1-x^2}\frac{d^n}{dx^n}\left(\sqrt{1-x^2}\right)^{2n-1}, \tag{2}$$

Hermite polynomials

$$H_n(x) = (-1)^n \exp(x^2)\frac{d^n}{dx^n}\exp(-x^2), \tag{3}$$

Hermite functions

$$\varphi_n(x) = \frac{1}{\sqrt{2^n n! \sqrt{\pi}}} H_n(x)\exp\left(-\frac{x^2}{2}\right), \tag{4}$$

etc. In Eqs. (1)-(4) n is nonnegative integer.

---

* **Oleg V. Chernoyarov[1], Martin Breznan[2], Alexander V. Terekhov[1]**
[1] National Research University "Moscow Power Engineering Institute", Moscow, Russia, E-mail: o_v_ch@mail.ru
[2] Department of Telecommunications and Multimedia, University of Zilina, Slovakia

The generalized spectra C_n^P, C_n^T, C_n^H, C'^H_n of the one-dimensional signals $s(x)$ specified on the interval $x \in [a,b]$ when using (1)-(4) can be accordingly found as

$$C_n^P = \frac{2n+1}{2}\int_{-1}^{1} s\left(\frac{b-a}{2}x + \frac{a+b}{2}\right)P_n(x)dx\,, \tag{5}$$

$$\begin{gathered} C_n^T = \frac{2}{\pi\alpha_n}\int_{-\pi/2}^{\pi/2} s\left[\frac{b-a}{2}\sin(x) + \frac{a+b}{2}\right]\cdot \\ \cdot T_n\left[\sin(x)\right]dx, \quad \alpha_n \begin{cases} 2. & n = 0; \\ 1, & n \neq 0; \end{cases} \end{gathered} \tag{6}$$

$$\begin{gathered} C_n^H = \frac{1}{2^n n!\sqrt{\pi}}\int_{-A}^{A} s\left(\frac{b-a}{2A}x + \frac{a+b}{2}\right)\cdot \\ \cdot H_n(x)\exp(-x^2dx), \\ C'^H_n = \int_{-A}^{A} s\left(\frac{b-a}{2A}x + \frac{a+b}{2}\right)\varphi_n(x)dx\,, \\ A >> 1 \end{gathered} \tag{7}$$

Then in absence of hindrances the signal restoration operation on the generalized spectrum will be written in the form

$$s(x) = \sum_{n=0}^{\infty} C_n \Psi_n\left(\gamma\frac{2x-a-b}{b-a}\right), \tag{8}$$

where $\{\Psi_n(x)\}$ is one of the orthogonal function systems (1)-(4), $\{C_n\}$ is generalized spectrum (5)-(7) corresponding to the basis $\{\Psi_n(x)\}$, $\gamma = A$ for Hermite polynomials (functions) (7) and $\gamma = 1$ for Legendre and Chebyshev polynomials. Generalization of formulas (5)-(8) on a two-dimensional case is obvious.

As the analysis showed, expansion coefficients (5)-(7) decrease rapidly with increasing n for both continuous (differentiable and nondifferentiable) and discontinuous signals. However for practical applications the calculation of the generalized spectra according to the formulas (5)-(7) should be made numerically in most cases. Thereupon it is of interest to use the expansion procedures which minimize computing expenses. One of the possible ways for creating such procedures is the use of Gauss quadrature formulas of the maximum degree of precision [6]:

$$\int s(x)\rho(x)dx = \sum_{k=0}^{N-1}\lambda_k s(x_k)\,. \tag{9}$$

Here x_k are zeros of a polynomial of N-th order, ρ_x is a weight function for polynomial orthogonality [7], λ_k are Christoffel numbers, and integration is conducted on an orthogonality interval of polynomials used in expancion. Special cases of Eq. (9) for polynomials (1)-(4) have the appearance:

1. Gauss-Legendre formula:

$$\int_{-1}^{1} s(x)dx = \sum_{k=0}^{N-1}\lambda_k s(x_k)\,, \tag{10}$$

where x_k are zeros of equation $P_N(x) = 0$, $\lambda_k = 2/(1 - x_k^2[P'_N(x_k]^2$.

2. Mehler (Gauss-Chebyshev) formula:

$$\int_{-1}^{1}\frac{s(x)}{\sqrt{1-x^2}}dx = \sum_{k=0}^{N-1}\lambda_k s(x_k)\,, \tag{11}$$

where $x_k = \cos[\pi(2k+1)/2N]$ are zeros of Chebyshev polynomial $T_N(x)$, $\lambda_k = \pi/N$.

3. Gauss-Hermite formula:

$$\int_{-\infty}^{\infty} s(x)\exp(-x^2)dx = \sum_{k=0}^{N-1}\lambda_k s(x_k)\,, \tag{12}$$

where x_k are zeros of equation $H_N(x) = 0$, $\lambda_k = 2^{N+1}N!\sqrt{\pi}/ /[H'_N(x_k)]^2$.

Using (10)-(12) the formulas (5)-(7) for the generalized signal spectra can be rewritten as

$$\begin{gathered} C_n^P = \sum_{k=0}^{N-1}\frac{2n+1}{(1-x_k^2)\left[P'_N(x_k)\right]^2}\cdot \\ \cdot s\left(\frac{b-a}{2}x_k + \frac{a+b}{2}\right)P_n(x_k), \end{gathered} \tag{13}$$

$$\begin{gathered} C_n^T = \frac{2}{N\alpha_n}\sum_{k=0}^{N-1} s\left[\frac{b-a}{2}\cos\left(\frac{\pi(2k+1)}{2N}\right) + \frac{a+b}{2}\right]\cdot \\ \cdot\cos\left[\frac{\pi n(2k+1)}{2N}\right], \end{gathered} \tag{14}$$

$$\begin{gathered} C_n^H = \frac{2^{N-n+1}N!}{n!}\sum_{k=0}^{N-1}\frac{1}{\left[H'_N(x_k)\right]^2}\cdot \\ \cdot s\left(\frac{b-a}{2}x_k + \frac{a+b}{2}\right)H_n(x_k), \end{gathered} \tag{15}$$

$$\begin{gathered} C'^H_n = N!\sqrt{\frac{2^{N-n+3}\sqrt{\pi}}{n!}}\sum_{k=0}^{N-1}\frac{1}{\left[H'_N(x_k)\right]^2}\cdot \\ \cdot s(x_k\sqrt{2})H_n(x_k\sqrt{2}). \end{gathered} \tag{16}$$

On the basis of Eqs. (13)-(16) and similar expressions the study of spectra for various signal models (differentiable, continuous nondifferentiable and discontinuous) was conducted. As an example in Table 1, coefficients C_n^P, C_n^T, C_n^H, C'^H_n, $0 \leq n \leq 20$ calculated according to formulas (13)-(16) for typical functions $s(x) = \exp(-x^2)$ (differentiable); $s(x) = 1 - |x|$, $-1 \leq x \leq 1$ (continuous nondifferentiable); $s(x) = \mathrm{sgn}(x)$, where $\mathrm{sgn}(x) = 1$ if $x \geq 0$ and $\mathrm{sgn}(x) = -1$ if $x < 0$ (discontinuous) are resulted. In that table the values of mean square error $\delta = \delta(N)$ [8]

$$\delta = \frac{1}{b-a}\int_a^b\left[s(x) - \sum_{n=0}^{N} C_n\Psi_n(x)\right]^2 dx \tag{17}$$

of signal approximation on the basis of the generalized spectra calculated at $a = -1$, $b = 1$ (orthogonality interval of polynomials (1), (2)) and various N are specified. Here C_n is one of the coefficients C_n^P, C_n^T, C_n^H, C'^H_n (13)-(16), and $\Psi_n(x)$ is corresponding polynomial (function) of n-th order (1)-(4).

From the Table 1 and conducted analysis including other orthogonal bases there follows that for differentiable and continuous nondifferentiable signals the construction of the generalized spectra on the basis of Hermite polynomials (3) or functions (4) appears the most effective. Really, coefficients C_n^H (15), C'^H_n (16) converge rapidly enough to zero with increasing n so mean square error (17) becomes relatively small at $N < 5$–6. The use of Legendre and Chebyshev polynomials can provide smaller extent of δ (17) in comparison with Hermite polynomials (functions). However sets $\{C_n^P\}$, $\{C_n^T\}$ can essentially exceed sets $\{C_n^H\}$, $\{C'^H_n\}$ by number of significant coefficients.

For discontinuous signals the Hermite polynomials (functions) also provide the least number of significant coefficients of the generalized spectrum. However approximation mean square error (17) received at their use reaches relatively large values. In this case the construction of the generalized spectrum on the basis of Chebyshev polynomials can appear to be expedient. Really, at loss by number of coefficients in 2 times approximately in comparison with Hermite series the accuracy gain of the signal approximation estimated according to (17) can reach 3–3.5 times. Also it should be noted that expansion coefficients calculated with the use of quadrature formulas of the maximum degree of precision which have an order less than 10^{-4} can be omitted at construction of the generalized spectrum of an analyzed signal. Besides, rate of convergence for series (8) on orthogonal polynomials can be more rapid generally than for basis of trigonometric functions and discrete cosine transformation [2].

3. Signal and image restoration in the presence of random distortions

As it has been established in the section 2 at restoration of the deterministic signals and images by their generalized spectra the number of restored signal (image) modes (coefficients of the generalized spectrum) is determined by the desired approximation error at use of one or another metrics of functional space. It will be shown below that the account of fluctuating noises actually existing together with a useful signal leads to occurrence of optimum number of modes which should be used for image restoration.

So, let the field

$$\xi(r) = s_M(r) + \eta(r), \quad r = (x,y)$$

be received to the optical-electron converter's input. Here s_M is observable useful signal for which the form and other characteristics can be apriori unknown, $\eta(r)$ is hindrance. Let us suppose that the useful signal $s_M(r)$ can be presented as series

$$s_M(r) = s_M(x,y) = \sum_{m=0}^{M}\sum_{n=0}^{M} c_{mn}\varphi_m(x)\varphi_n(y) \qquad (18)$$

in one or another orthonormal function system $\{\varphi_{m\,n}(x)\}$. Here $C_{m\,n} = \int_\Omega s_M(x,y)\varphi_m(x)\varphi_n(y)$ are expansion coefficients (generalized signal spectrum) and Ω is area of orthogonality of basis $\{\varphi_m(x)\varphi_m(y)\}$.

Models of hindrances at optical and optoelectronic processing of the information are in detail considered in [9, 10]. If a quantum character of weak optical signals is not taken into consideration and signal $s_M(r)$ is assumed as intensive enough then it is possible often to believe that a hindrance $\eta(r)$ is Gaussian random field. In addition, if processing algorithms do not consider thin difference of spatial spectra of a useful signal and hindrance but focus the attention on the analysis of generalized spectrum's modes of the field $s_M(r)$ then hindrance $\eta(r)$ may be approximated by Gaussian white noise with a correlation matrix of a kind $K_\eta(r_1, r_2) = (N_0/2)\delta(r_2 - r_1)$. Here $N_0/2$ is spectral density (intensity) of white noise. The model $\eta(r)$ as white noise is quite proved if hindrance sources are wideband processes in electronic devices [11] and will be used by us further.

On observable realization $\xi(r)$ and the available prior information it is necessary to restore an original signal $s_M(r)$ in an optimal way.

In the task of signal (image) restoration at the presence of random hindrances the vector $\vec{C} = (C_0, ..., C_M)$ is unknown and cannot be calculated directly as in (5)–(7). Therefore in the processing operations it is necessary to receive an estimate of this vector. At synthesis of estimation algorithm we will use a maximum likelihood method. Without loss of generality of obtained results for simplification of mathematical calculations we will believe a signal $s_M(r)$ as one-dimensional in the sequel, i.e. $s_M(r) = s_M(x)$. Then following [12] a logarithm of functional of likelihood ratio L as function of current values of all unknown coefficients C_n, $0 \le n \le M$ can be written down in the form:

$$L(\vec{C}) = \frac{1}{N_0}\left[2\int_\Omega \xi(x)s_M(x,\vec{C})dx - \int_\Omega s_M^2(x,\vec{C})dx\right]. \qquad (19)$$

Using the representation of a useful signal (18) in Eq. (19) we have

$$L(\vec{C}) = \frac{1}{N_0}\sum_{n=0}^{M}(2X_nC_n - C_n^2). \qquad (20)$$

Here $X_n = \int_\Omega \xi(x)\,\varphi_n(x)dx$. As maximum likelihood estimates (MLEs) $\tilde{\vec{C}} = (\tilde{C}_0, ..., \tilde{C}_M)$ of measured parameters the position of the greatest maximum of solving statistics (20) is taken:

$$\tilde{\vec{C}} = \arg\sup_{\vec{C}} L(\vec{C}).$$

Estimates $\tilde{C}_n$, $0 \le n \le M$ can be found from the decision of the likelihood equations

$$\partial L(\vec{C})/\partial C_0 = 0, \; \partial L(\vec{C})/\partial C_1 = 0, \; ...,$$

$$\partial L(\vec{C})/\partial C_M = 0. \qquad (21)$$

Substituting (20) in (21) and carrying out simple transformations we come to MLEs of a kind

$$\vec{C}_n = X_n = \int_\Omega \xi(x)\varphi_n(x)dx, \; 0 \le n \le M. \qquad (22)$$

Polynomial coefficients and approximation mean square error for various types of signals

Table 1

n	$\exp(-x^2)$				$1-\lvert x\rvert$				$\mathrm{sgn}(x)$			
	C_n^P	C_n^T	C_n^H	C'^H_n	C_n^P	C_n^T	C_n^H	C'^H_n	C_n^P	C_n^T	C_n^H	C'^H_n
0	0.747	0.645	0.707	0.816	0.499	0.363	0.479	0.482	$-5.9*10^{-5}$	0	$-2.8*10^{-8}$	$-3.9*10^{-8}$
1	$-6.4*10^{-5}$	0	$-4.2*10^{-8}$	$-3.7*10^{-8}$	$3.1*10^{-6}$	0	$-2.0*10^{-11}$	$-3.4*10^{-12}$	1.503	1.275	0.576	1.152
2	−0.446	−0.313	−0.088	−0.068	−0.623	−0.423	−0.085	−0.091	$-3.1*10^{-4}$	0	$6.6*10^{-8}$	$1.8*10^{-7}$
3	$-1.6*10^{-4}$	0	$8.1*10^{-9}$	$-4.5*10^{-9}$	$7.8*10^{-6}$	0	$4.0*10^{-12}$	$1.9*10^{-13}$	−0.885	−0.428	−0.050	0.088
4	0.074	0.039	$-5.5*10^{-3}$	$2.8*10^{-3}$	0.184	0.084	$7.3*10^{-3}$	$8.1*10^{-3}$	$-6.2*10^{-4}$	0	$-6.2*10^{-9}$	$1.1*10^{-8}$
5	$-2.9*10^{-4}$	0	$-9.9*10^{-11}$	$1.1*10^{-9}$	$2.0*10^{-5}$	0	$-3.9*10^{-13}$	$5.4*10^{-14}$	0.708	0.261	$3.9*10^{-3}$	0.017
6	$-7.3*10^{-3}$	$-3.2*10^{-3}$	$-2.3*10^{-4}$	$-7.7*10^{-5}$	−0.097	−0.035	$-4.1*10^{-4}$	$-4.2*10^{-4}$	$-1.1*10^{-3}$	0	$-1.3*10^{-9}$	$1.7*10^{-8}$
7	$-4.9*10^{-4}$	0	$6.3*10^{-11}$	$-3.9*10^{-11}$	$4.9*10^{-5}$	0	$2.3*10^{-14}$	$-1.0*10^{-14}$	−0.619	−0.191	$-2.4*10^{-4}$	$6.8*10^{-4}$
8	$-5.6*10^{-4}$	$2.0*10^{-4}$	$-7.2*10^{-6}$	$1.2*10^{-6}$	0.061	0.019	$1.7*10^{-5}$	$1.2*10^{-5}$	$-1.9*10^{-3}$	0	$2.0*10^{-10}$	$3.9*10^{-10}$
9	$-8.9*10^{-4}$	0	$6.3*10^{-12}$	$-2.1*10^{-12}$	$4.8*10^{-5}$	0	0	0	0.569	0.154	$1.2*10^{-5}$	$6.7*10^{-5}$
10	$-1.2*10^{-4}$	$-1.0*10^{-5}$	$-1.8*10^{-7}$	$3.2*10^{-8}$	−0.041	−0.011	$-5.4*10^{-7}$	7.7*10–8	$-3.3*10^{-3}$	0	$-1.1*10^{-11}$	$1.6*10^{-10}$
11	$-1.6*10^{-3}$	$-1.2*10^{-15}$	$-3.2*10^{-13}$	$2.3*10^{-13}$	$-1.0*10^{-4}$	0	0	0	−0.539	−0.132	$-5.3*10^{-7}$	$1.7*10^{-6}$
12	$-5.4*10^{-5}$	$4.2*10^{-7}$	$3.7*10^{-9}$	$-5.4*10^{-9}$	0.029	$7.1*10^{-3}$	$1.4*10^{-8}$	$-2.8*10^{-8}$	$-4.7*10^{-3}$	$-4.2*10^{-15}$	$1.4*10^{-13}$	$5.0*10^{-12}$
13	$-2.1*10^{-3}$	$1.7*10^{-15}$	$9.5*10^{-15}$	$-8.8*10^{-15}$	$-3.9*10^{-4}$	0	0	0	0.522	0.117	$2.0*10^{-8}$	$1.2*10^{-7}$
14	$-8.0*10^{-4}$	$-1.5*10^{-8}$	$-6.6*10^{-11}$	$4.2*10^{-10}$	−0.020	$-4.5*10^{-3}$	$-2.9*10^{-10}$	$1.8*10^{-9}$	$-4.2*10^{-3}$	$-1.5*10^{-14}$	$1.2*10^{-14}$	$-3.0*10^{-13}$
15	$-1.3*10^{-3}$	$-1.0*10^{-14}$	0	0	$-5.0*10^{-4}$	$-5.2*10^{-15}$	0	0	−0.522	−0.108	$-6.4*10^{-10}$	$2.2*10^{-9}$
16	$-9.9*10^{-4}$	$4.6*10^{-10}$	$9.9*10^{-13}$	$-2.5*10^{-11}$	0.012	$2.7*10^{-3}$	$4.9*10^{-12}$	$-7.9*10^{-11}$	$-4.2*10^{-4}$	$1.3*10^{-14}$	0	$1.0*10^{-14}$
17	$1.5*10^{-3}$	$6.7*10^{-14}$	0	0	$-1.2*10^{-4}$	$3.5*10^{-14}$	0	0	0.525	0.103	$1.8*10^{-11}$	$1.2*10^{-10}$
18	$5.3*10^{-4}$	$-5.1*10^{-12}$	$-1.1*10^{-14}$	$1.2*10^{-12}$	$-6.3*10^{-3}$	$-1.3*10^{-3}$	$-5.9*10^{-14}$	$2.8*10^{-12}$	$8.6*10^{-3}$	$-2.3*10^{-13}$	0	0
19	$5.1*10^{-3}$	$-3.0*10^{-13}$	0	0	$5.9*10^{-4}$	$-1.6*10^{-13}$	0	0	−0.539	−0.100	$-4.8*10^{-13}$	$1.6*10^{-12}$
20	$2.2*10^{-3}$	$-6.9*10^{-11}$	0	0	$6.7*10^{-4}$	$-1.0*10^{-11}$	0	0	0.016	$9.0*10^{-13}$	0	0
δ(5)	$4.2*10^{-6}$	$5.3*10^{-6}$	$2.0*10^{-3}$	$9.8*10^{-5}$	$1.3*10^{-3}$	$1.4*10^{-3}$	0.016	$6.5*10^{-3}$	0.098	0.100	0.186	0.148
δ(10)	7.1*10–8	$8.8*10^{-14}$	$1.2*10^{-5}$	$5.2*10^{-6}$	$1.5*10^{-4}$	$1.6*10^{-4}$	$2.4*10^{-3}$	$2.2*10^{-3}$	0.061	0.061	0.145	0.125
δ(15)	$4.2*10^{-7}$	0	$7.1*10^{-7}$	$8.9*10^{-5}$	$7.1*10^{-5}$	$7.6*10^{-5}$	$1.8*10^{-3}$	$4.4*10^{-3}$	0.040	0.040	0.114	0.124
δ(20)	$1.2*10^{-6}$	0	$8.0*10^{-8}$	$9.6*10^{-4}$	$5.1*10^{-5}$	$5.5*10^{-5}$	$1.7*10^{-3}$	0.014	0.034	0.033	0.108	0.120

The found estimates (22) obviously allow to restore a useful signal $s_M(x)$ according to Eq. (18).

Let us find characteristics of MLEs (22). As MLEs' characteristics we will use their conditional biases (systematic errors) $b(\tilde{C}_n | C_n) = \langle \tilde{C}_n \rangle - C_n$ and variances (error mean squares) $V(\tilde{C}_n | C_n) = \langle (\tilde{C}_n - C_n)^2 \rangle$. Here angle brackets $\langle\ \rangle$ designate averaging operation on all possible realizations $\xi(r)$. Carrying out direct averaging Eq. (22) we receive that the estimate $\tilde{C}_n$ is conditionally unbiased: $b(\tilde{C}_n | C_n) = 0$ and its conditional variance has the appearance $V(\tilde{C}_n | C_n) = N_0/2$. Besides, by virtue of orthogonality of basis $\{\varphi_n(x)\}$ the estimates $\tilde{C}_n$, $0 \le n \le M$ are uncorrelated: $\langle (\tilde{C}_k - C_k)(\tilde{C}_n - C_n) \rangle = (N_0/2)\int_\Omega \varphi_k(x)\, \varphi_n(x) dx = 0,\ k \ne n$.

Let us assume that for restoration of a useful signal $s_M(x)$ K modes of expansion (22) are used. Then the restored signal $\tilde{s}_K(x)$ can be presented as follows

$$\tilde{s}_K(x) = \sum_{n=0}^{K-1} \tilde{C}_n \varphi_n(x). \tag{23}$$

Define difference of the restored signal $\tilde{s}_K(x)$ (23) from original signal $s_M(x)$ (18) as [12, 13]

$$\langle\delta\rangle = \left\langle \int_\Omega \left[\tilde{s}_K(x) - s_M(x)\right]^2 dx \right\rangle = \sum_{n=0}^{K-1} V(\tilde{C}_n | C_n) + \sum_{n=K}^{M} C_n^2 = \frac{N_0}{2}\left(K + \sum_{n=K}^{M} z_n^2\right). \tag{24}$$

Here $z_n^2 = 2C_n^2/N_0$ is the signal-to-noise ratio (SNR) for a mode with an index n. According to Eq. (24) error $\langle\delta\rangle$ consists of two components. The first component is the fluctuating error increasing with K which is attributed to a deviation $\tilde{C}_n$ from C_n. The second component is caused by energy of unrestored modes. It is obvious that it decreases with K.

Let us denote $z^2_{\min} = \min(z^2_K, z^2_{K+1}, \dots, z^2_M)$ and $z^2_{\max} = \max(z^2_K, z^2_{K+1}, \dots, z^2_M)$ as minimum and maximum SNR in the residual mode signal expansion. Then from Eq. (24) following inequalities can be written down:

$$(M+1)z^2_{\min} + K(1 - z^2_{\min}) < N_0\langle\delta\rangle/2 <$$

$$< (M+1)z^2_{max} + K(1 - z^2_{max}).$$

Thus, the normalized error $\mu(K) = N_0\langle\delta\rangle/2$ as function of integer variable K lies between two majorizing functions

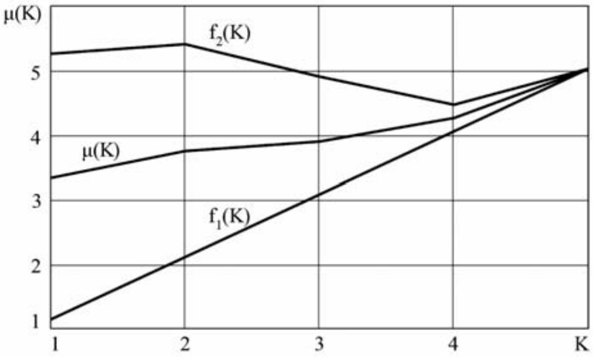
$$f_1(K) = (M+1)z^2_{\min} + K(1 - z^2_{\min}),$$

$$f_2(K) = (M+1)z^2_{\max} + K(1 - z^2_{\max}). \tag{25}$$

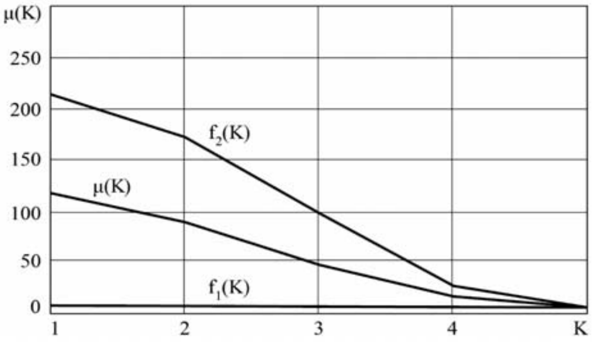
In Fig. 1 the typical behavior of normalized error $\mu(K)$ for three cases is shown: a) $z^2_{\min} < 1$, $z^2_{\max} < 1$ (Fig. 1a); b) $z^2_{\min} > 1$, $z^2_{\max} > 1$ (Fig. 1b); c) $z^2_{\min} < 1$, $z^2_{\max} > 1$ (Fig. 1c). Also here for comparison the dependences of functions $f_1(K)$ and $f_2(K)$ (25) are plotted. As test signal the signal $s_M(x) = \exp[-\alpha^2(x - \tau)^2]$ in basis of Hermite functions (4) was used. The generalized spectrum of such signal has the appearance [4]

$$C_n = \frac{\alpha^n}{2^n n! (\alpha^2 + 1)^{(n+1)/2}} \exp\left(-\frac{\alpha^2\tau^2}{\alpha^2 + 1}\right) H_n\left(\frac{\alpha\tau}{\sqrt{\alpha^2 + 1}}\right).$$

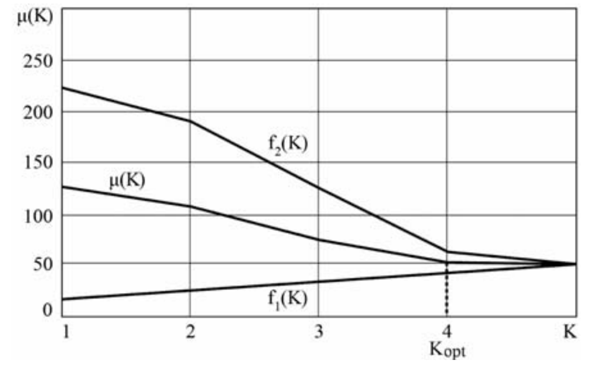
At construction of graphs it was taken $\alpha = 1$, $\tau = 2$ and first six modes of expansion ($M = 5$) were discounted. The noise spectral density for Figs. 1a-1c was accepted equal 1, 0.02 and 0.2 respectively.



a)



b)



c)

Fig. 1 Normalized restoration error of a useful signal at various intensities of acting hindrances

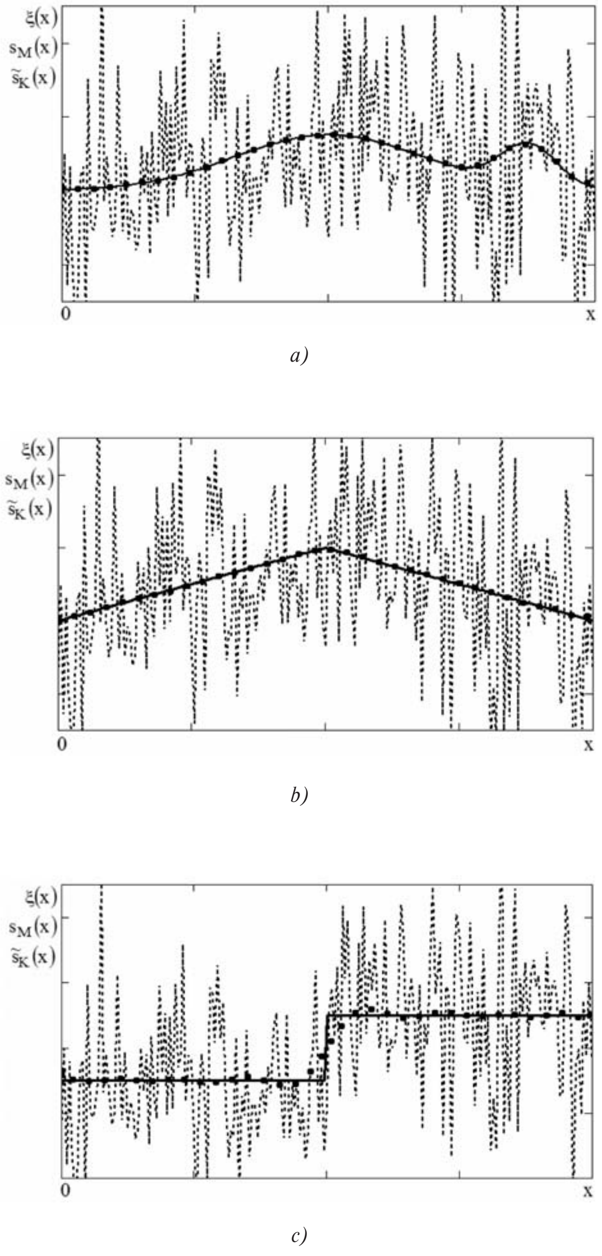
As follows from Fig. 1 under conditions $z_{\min}^2 < 1$, $z_{\max}^2 < 1$ the function $\mu(K)$ is monotonously increasing (i.e., with increasing number of processed modes the restoration error of a useful signal (24) also increases). If $z_{\min}^2 > 1$, $z_{\max}^2 > 1$ then the function $\mu(K)$ (and therefore the restoration error (24)) monotonously decreases with increasing number of processed modes. At last, in case of $z_{\min}^2 < 1$, $z_{\max}^2 > 1$ the behavior of function $\mu(K)$ becomes nonmonotonic. Really, having rewritten Eq. (24) as

$$\mu(K) = M + 1 + \sum_{n=K}^{M} \left(z_n^2 - 1\right) \tag{26}$$

it is possible to see that while condition $z_n^2 > 1$ holds summands under summation symbol (26) are positive and $\mu(K)$ decreases with increasing K. However, with the beginning of fulfillment of the condition $z_n^2 < 1$ summands in Eq. (26) change a sign on opposite and the error $\mu(K)$ starts to increase. The optimum number of restored modes K_{opt} corresponding to the minimum restoration error $\mu(K)$ of a useful signal $s_M(x)$ is determined by the relation $z_K^2 = 1$ (or more exactly $z_K^2 > 1$, $z_{K+1}^2 < 1$) for the considered example $K_{opt} = 4$ (Fig. 1c). Physically this is explained by the fact that under $z_n^2 < 1$ the fluctuating error connected with restoration of such modes is more than a deterministic component of the total error. Therefore, further expansion of signal components $\tilde{s}_K(x)$ becomes inexpedient.

Qualitative illustration of operation of the synthesized restoration algorithm (22), (23) at various types of the useful signal observed against white noise is shown in Fig. 2. Here the original useful signal $s_M(x)(M = \infty)$ is plotted by solid lines, the realization of the observed data $\xi(x)$ – by dashed lines and the restored signal $\tilde{s}_K(x)$ – by points. In Fig. 2a the useful signal is described by differentiated function with two local maxima, in Fig. 2b – continuous nondifferentiable function and in Fig. 2c – discontinuous function. As basic functions $\{\varphi_n(x)\}$ Legendre polynomials (1) were used. The number of coefficients of the generalized spectrum used at restoration was chosen based on the condition $\langle\delta\rangle \leq 0.05$ (24) for the "worst" type of signal. As "worst" the signal for which the coefficients of the generalized spectrum decrease more slowly with increasing sequence number is meant (as follows from Table 1 this type of signal will be discontinuous signal). Thereby the K value made $K = 15$ (it may be noted that generally speaking $K < K\leq_{opt}$ for the considered examples).

From Fig. 2 follows that the proposed algorithm (22), (23) allows to restore information signals against uncorrelated hindrances at input SNRs' values of the order or less -3 dB to a high accuracy. The conducted additional researches have shown that the mentioned algorithm can be also used for effective extraction of signals against stationary Gaussian and Laplace hindrances (including hindrances with high enough intensity).



a)

b)

c)

Fig. 2 Original, noisy and restored signals with various behavior

4. Conclusion

In this work the restoration algorithms of signals and images on the basis of their representation in the form of a set of coefficients of the generalized spectrum at absence and presence of random distortions were considered. It was shown that the number of coefficients of the generalized spectrum of a restored deterministic signal (image) is defined by the desired approximation error at use of one or another metrics of functional space. Herewith at processing of continuous signals the application of Hermite polynomials or functions appears the most effective. If the useful signal is discontinuous then the construction of the generalized spectrum in basis of Chebyshev polynomials is more expedient.

In the presence of stationary hindrances the maximum-likelihood approach in aggregate with the generalized spectral signal representation in bases of orthogonal polynomials (functions) allows to receive the restoration algorithm of signals (images) which is rather effective (including the action of hindrances of enough high intensity), requires the minimum amount of the prior information (in particular, the form and other characteristics of a useful signal can be unknown) and can be practically realized in both analogue and digital variants sufficiently simply. Besides, in the general case there is an optimum number of coefficients of the generalized spectrum (relatively small because of fast decrease of the generalized spectrum with increasing the number of coefficient) required for restoration of a useful signal against hindrances here.

The results obtained allow to recommend the proposed algorithms for restoration of information signals and images which are deterministic or distorted by stationary hindrances (including correlated and non-Gaussian) in various practical applications.

Acknowledgement

The reported study was supported by Russian Foundation for Basic Research, research projects No. 12-01-09406-mob_z, 13-08-00735a, and by The Ministry of education and science of Russian Federation, projects 14.B37.21.2102, 14.B37.21.2032, 14.B37.21.2015.

References

[1] VATOLIN, D., RATUSHNYAK, A., SMIRNOV, M., YUKIN, V.: Methods of Data Compression. *The Archivers' Structure, Image and Video Compression [in Russian], Dialog-MIFI,* Moscow, 2003.
[2] Digital Processing of Computer and Television Images, Edited by Zubarev, Y. B., Dvorkovich, V. P. [in Russian], *Moscow Centre of the Scientific and Technical Information,* Moscow, 1997.
[3] CHERNOYAROV, O.V., BREZNAN, M.: Optimal and Quasioptimal Algorithms of Distinction of the Compressed Images in Bases of Orthogonal Polynomials, *Communications - Scientific Letters of the University of Zilina,* vol. 14, No. 2, 2012, pp. 22–26.
[4] CHERNOYAROV, O. V., BREZNAN, M., GOLOBORODKO, P. A.: The Comparative Analysis of Optimal Estimations of Images' Shift Parameter in Synchronous and Asynchronous Orthogonal Function Bases [in Russian], *Herald of the Moscow Power Engineering Institute,* No. 3, 2012, pp. 103-109.
[5] SZEGO, G.: Orthogonal Polynomials, *American Mathematical Society,* 4 edition, 1975.
[6] HAZEWINKEL, MICHIEL: Encyclopedia of Mathematics, *Springer,* 2002.
[7] GULDAN, V., MARCOKOVA, M. Orthogonal Polynomials and Related Special Functions Applied in Geosciences and Engineering Computations, *Communications - Scientific Letters of the University of Zilina,* vol. 12, No. 1, 2010, pp. 12–15.
[8] RICHARD, L. BURDEN, J., DOUGLAS, FAIRES: Numerical Analysis, *Thomson Brooks/Cole,* 2005.
[9] Adaptation in Information Optical Systems, Edited by N. D. Ustinov [in Russian], *Radio i Svyaz',* Moscow, 1984.
[10] The Theory of Coherent Images, Edited by N. D. Ustinov [in Russian], *Radio i Svyaz',* Moscow, 1987.
[11] Systems of Machine Vision, Edited by Pisarevsky, A. N., Chernyahovsky, A. F. [in Russian], *Mashinostroienie,* Leningrad, 1988.
[12] TRIFONOV, A. P., SHINAKOV, Y. S.: Joint Discrimination of Signals and Estimation of their Parameters against Background [in Russian], *Radio i Svyaz',* Moscow, 1986.
[13] PRATT, W. K.: *Digital Image Processing, John Wiley & Sons, Inc.,* Hoboken : New Jersey, 2007.

Eliska Jurisova – Libor Ladanyi – Jarmila Mullerova *

INVESTIGATION OF THE THERMAL SENSITIVE SPECTRAL RESPONSE OF CHALCOGENIDE FIBER BRAGG GRATING

This contribution presents numerical studies of the nonlinear fiber Bragg grating (FBG). The proposed model corresponds well to chalcogenide FBG as a considerable tool for optical switching. The spectral response of this device is discussed theoretically. Simulations based on the nonlinear coupled mode theory are used for investigating the relationships between the spectral transmission of nonlinear chalcogenide FBG and the grating parameters. The influence of the temperature dependence of FBG is numerically investigated. Numerical results show that the ambient temperature has an influence on the spectral response of FBG. The results are indications for applications in optical switches for all-optical communication networks and this can be important for wavelength division multiplexing optical networks.

Keywords: *Chalcogenide glass, fiber Bragg gratings, optical switching, thermal sensitivity, spectral response.*

1. Introduction

Optical bistable devices have a large number of applications in all-optical communication systems. They are building blocks of optical logic, memories, switches, light power limiters etc. that are crucial parts of optical network. In this paper we focused on the optical switches devices. Ultrafast optical switches are necessary parts of optical network nodes in high bit rate all-optical communication networks including systems of wavelength division multiplexing (WDM) [1, 2]. The phenomenon of optical bistability (OB) is important to realize nonlinear switching in FBG at a desired wavelength. This optical phenomenon refers to the situation in which an optical device exhibits two stable different output intensities under one given incident intensity. OB arises from the energy exchange between light and nonlinear optical medium. It arises in feedback structures with some loss or nonlinearity [3, 4].

Nonlinear chalcogenide glasses as the nonlinear switching medium have advantages over other materials including ultrafast response time [5]. Chalcogenide glasses (CGs) are based on the chalcogen elements S, Se and Te with the addition of other elements such as Ge, As and Sb to form stable glasses. CGs are low-phonon energy materials and are generally transparent from near-infrared to mid-infrared. These glasses are well-known as optical materials with high nonlinear optical properties. High third-order Kerr nonlinearities up to 1000 times higher than those of silica glass make them distinguished candidates for applications such as nonlinear switching, Raman amplification, optical regeneration, parametric amplification and supercontinuum generation [5–7].

FBG is an important component of wavelength division multiplexed (WDM) optical communication systems [8, 9]. But there is still the problem of variable optical communication path characteristics and environmental fluctuations [10, 11].We investigate the nonlinear chalcogenide fiber Bragg gratings (FBG) as a promising device for all-optical switching. This effect can be modified depending on the incident light power which is assumed to be due to the optical Kerr effect. The aim of this contribution is to examine the bistable behavior and the spectral characteristics in a nonlinear chalcogenide FBG under the influence of the ambient temperature theoretically. The incident intensity thresholds of the bistable regimes are analyzed numerically to deduce possible switching properties of FBGs based on one chalcogenide glass sample taken from [12]. Nonlinear coupled mode equations (NCMEs) are used to model forward and backward light waves in nonlinear chalcogenide FBG. As the thermo-optic changes of the refractive index usually dominate over the changes due to the fiber elongation, only the thermo-optic effect on the refractive index is considered here.

2. Background

The nonlinear effect available in the third-order Kerr nonlinear media results in the intensity dependent refractive index change, which causes an intensity dependent phase modulation. When this phase modulation is converted into an amplitude modulation the switching can result [13].

A Fiber Bragg Grating

FBG (Fig. 1) is fabricated by exposing an optical photosensitive fiber to a pattern of UV light. This causes the changes of the refractive index along the FBG length L periodically. In Fig.1 I_{in} represents the incident light intensity, I_{ref} the reflected intensity and I_{tran} the transmitted intensity. The transmittance is given as

---

* **Eliska Jurisova, Libor Ladanyi, Jarmila Mullerova**
Institute of Aurel Stodola Liptovsky Mikulas, Faculty of Electrical Engineering, University of Zilina, Slovakia, E-mail: jurisova@lm.uniza.sk

the ratio between I_{tran} and I_{in} as $T = I_{tran}/I_{in}$. The key property of gratings is that they exhibit the Bragg reflection around a designed wavelength, the so-called Bragg wavelength λ_{Bragg}. This corresponds to the central rejection peak in the transmittance given by the Bragg condition [13, 14]

$$\lambda_{Bragg} = 2n_{eff}\Lambda \tag{1}$$

where n_{eff} is the effective index of the fiber core and Λ is the grating spatial period.

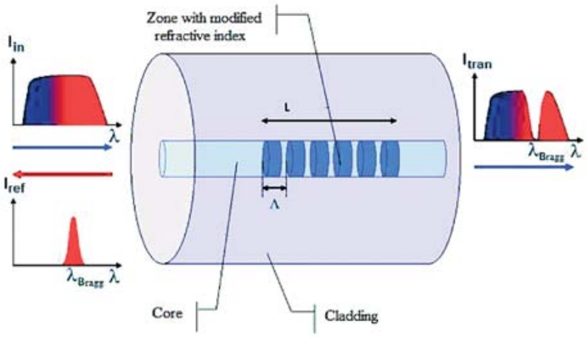


Fig.1 Design of propagating waves in the fiber Bragg grating [15]

FBG consists of a material, the optical parameters of which could be affected by heat. As the FBG temperature changes, a wavelength change defined by the shift of $\Delta\lambda_{Bragg}$ from the original wavelength λ_{Bragg} can be observed. This change is given by

$$\Delta\lambda_{Bragg} = \lambda_{BO}\,(\alpha_e + \alpha_o)\Delta T_{FBG} \tag{2}$$

where λ_{BO} is the FBG Bragg wavelength at a reference temperature of T_R, α_e is the thermal expansion coefficient of the fiber, α_o is the thermo-optic coefficient. $\Delta T_{FBG} = T_A - T_R$, T_A is the ambient temperature. The FBG thermal sensitivity S_{FBG} is defined as $S_{FBG} = \Delta\lambda_{Bragg}/\Delta T_{FBG}$. Hence, the thermally influenced Bragg wavelength is defined as [16]

$$\lambda_{Bragg} = \lambda_{BO} + S_{FBG}\Delta T_{FBG} \tag{3}$$

The refractive index inside the fiber core varies along the grating length periodically. The index variation can be expressed as [15, 17]

$$n = n_{eff} + V_n\nu\cos\left(\frac{2\pi z}{\Lambda}\right) + n_2|E|^2 \tag{4}$$

Here V_n is the depth of the modulation, ν the fringe visibility, n_2 is the nonlinear index of optical fiber, E is the electric field amplitude. The z coordinate denotes the grating direction. The electric field inside the grating can be written as

$$E = \left[Fe^{(ik_Bz)} + Be^{(-ik_Bz)}\right]e^{(-i\omega_0t)} \tag{5}$$

where $k_B = \pi/\Lambda n_{eff}$ is the wavenumber in the medium at the Bragg resonance, F and B are the amplitudes of the transverse core mode fields of the forward and backward (counter-propagating) waves, respectively. F and B exhibit additional slow variations as functions of position z and time t; they are thus the envelope functions. The carrier optical frequency at which the wave spectrum is initially centered is marked as ω_0. We assume that $|E|^2 \sim I$ represents the intensity of light.

B Nonlinear coupled mode theory

The forward and backward light waves exist simultaneously and satisfy NCMEs [17–19]:

$$i\frac{\partial F}{\partial z} + i\frac{1}{v_g}\frac{\partial F}{\partial t} + \kappa B + \delta F + + \gamma|F|^2F + 2\gamma|B|^2F = 0 \tag{6}$$

$$-i\frac{\partial B}{\partial z} + i\frac{1}{v_g}\frac{\partial B}{\partial t} + \kappa F + \delta B + + \gamma|B|^2B + 2\gamma|F|^2B = 0 \tag{7}$$

where the variable $v_g = c/n_{eff}$ represents the group velocity in the absence of the grating. The parameter κ is the coupling coefficient representing the grating strength per unit length. It determines the energy ratio between the forward and the backward fields and is defined as

$$\kappa = \frac{\pi V_n}{\lambda_{Bragg}} \tag{8}$$

The detuning parameter δ is the relation between the Bragg wavelength (λ_{Bragg}) and the carrier wavelength (λ) given as

$$\delta = 2\pi n_{eff}\left(\frac{1}{\lambda} - \frac{1}{\lambda_{Bragg}}\right) \tag{9}$$

The variable γ symbolizes the nonlinear (Kerr) coefficient (if $\gamma = 0$ the linear regime occurs)

$$\gamma = \frac{2\pi n_2}{\lambda_{Bragg}} \tag{10}$$

This parameter enters the term $\gamma|F|^2F$(or $\gamma|B|^2B$) in (6) (or 7) determining the strength of the self-phase modulation. The last term in (6) (or 7) belongs to the cross-phase modulation.

The boundary conditions in the spatial domain are usually chosen at the two opposite sides of the grating. The initial and boundary conditions take the well-known form [17-19]

$$F = (z = 0, t) = A(t);\ B = (z = L, t) = 0;$$

$$F(z, 0) = 0;\ B(z, 0) = 0 \tag{11}$$

The second of these equations indicates that the input wave is launched only from the $z = 0$ position. In this paper the finite difference method (FDM) was used to solve NCMEs (6), (7).

C Optical bistability

OB is a promising way to realize an optical switch, a key component of integrated photonic devices. The Kerr (dispersive) optical bistability is based on coupling the feedback mechanism in a nonlinear optical medium that exhibits a change in the refractive index as a function of incident light intensity. Fig. 2 shows the shape of the curve representing OB as the following change of the transmittance (T) as a function of the incident intensity (I_{in}).

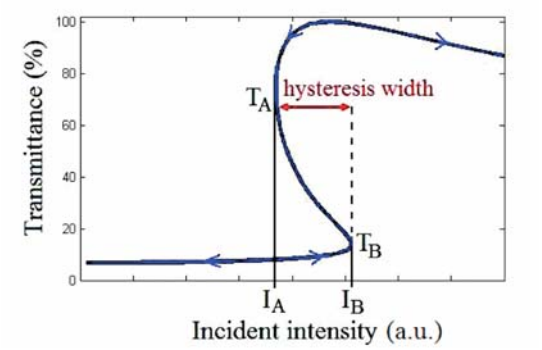


Fig. 2 The optical bistability curve

While I_{in} is pumped slightly the system behaves as linear until reaching the point I_B (the switch-on threshold). Here the output intensity makes a sudden jump to the upper branch of the hysteresis loop. When the optical intensity decreases from high values, it can return to the lower branch at the switch-off threshold (I_A, Fig. 2).The width of the hysteresis loop is defined as the difference between the switch-on (I_B) and -off (I_A) thresholds. Let us call the state of a linear increase of I_{in} as the "OFF" state and the state of a sharp rise in I_{in} for $I_{in} > I_A$ as the "ON" state. The transmission difference between "ON" and "OFF" states is the so-called on-off contrast. The width of the S-shaped curve (the so-called hysteresis width, w) is the difference between the switch-on (point I_B) and switch-off (point I_A) intensities. The middle branch of the S-curve is unstable against amplitude fluctuations due to the negative slope of the transmittance curve.

3. Results and numerical simulations

Here we present the simulation results and discuss the bistable behavior and the threshold of optical switching intensities in nonlinear chalcogenide FBGs. Then, the spectral responses of nonlinear FGBs at different incident intensities and at varied temperatures are numerically investigated. The experimental data necessary for numerical studies were taken from [12, 20]. The parameters of As_2Se_3 chalcogenide used in the simulations are in Table 1.

We investigated OB in the 3rd transmission spectral window of the optical communication WDM systems. The Bragg wavelength was positioned at the wavelengths of $\lambda_{Bragg1} = 1550.12$ nm and $\lambda_{Bragg2} = 1552.52$ nm representing telecommunication C-band channels of the 100 GHz and 50 GHz channel spacing grid according to G.694 ITU-T standard. The neighboring channel separation in the vicinity of 1550 nm for the 100 GHz (50 GHz) channel spacing is 0.8 nm (0.4 nm). In our investigation we assume 50 GHz channel spacing.We suppose the stationary behavior of the nonlinear system. Then NCMEs (6) and (7) have the form

Parameters of FBG and CG taken from literature [12, 20]   Table 1

Grating length	L	1 cm
Depth of the modulation	V_n	1×10^{-4}
Bragg wavelength	λ_{Bragg1}	1550.12 nm
Bragg wavelength	λ_{Bragg2}	1552.52 nm
Carrier wavelength	λ	($\lambda_{Bragg} \pm 0.2$) nm
Fringe visibility	v	1
Effective index As_2Se_3	n_{eff}	2.81
Nonlinear index As_2Se_3	n_2	14×10^{-14} cm^2/W
Reference temperature	T_R	10 [°C]
Ambient temperature	T_A	<8,15> [°C]
Thermal expansion coefficient	α_e	4.10^{-6} [K^{-1}]
Thermo-optic coefficient	α_o	21.10^{-6} [K^{-1}]

$$\frac{\partial F}{\partial z} = i\left[\kappa B + \delta F + \gamma\left(|F|^2 + 2|B|^2\right)F\right] \qquad (12)$$

$$\frac{\partial B}{\partial z} = -i\left[\kappa F + \delta B + \gamma\left(|B|^2 + 2|F|^2\right)B\right] \qquad (13)$$

A Transmittance versus incident intensity

Fig. 3 shows the nonlinear characteristics for As_2Se_3. The detuning parameter $\delta = 0$, that means $\lambda = \lambda_{Bragg}$. We observe the change in the transmittance: when the incident intensity increases from zero, the transmittance traces the lower branch of the line until the input intensity achieves the switching point, where the output makes a rapid jump to the upper branch. From this figure it is clear that the change in the Bragg wavelength does not cause a considerable change in the bistable curves. The S-shapes, the transmittances, switching intensities, the hysteresis widths and the on-off contrasts are very similar.

B Investigation of spectral responses

The possible distortions of spectral responses may occur in the OB incident intensity regions. Figs. 4 and 5 show the transmittance spectra of one samle of nonlinear chalcogenide FBG with the incident intensity as a parameter. The transmittance spectrum of linear FBGs is shown for the comparison (the dashed line). In the linear regime ($\gamma = 0$), the grating response does not depend on the incident intensity. As seen from the figures, the spectral curve is symmetrical on both sides of λ_{Bragg} at which a pronounced central transmittance minimum occurs. The simulations were carried out for changing the difference $\Delta\lambda = \lambda - \lambda_{Bragg}$ of the carrier wavelength λ and the Bragg wavelength λ_{Bragg} from $\Delta\lambda = -0.2$ nm ($\lambda =$ = 1552.32 nm) up to $\Delta\lambda = 0.2$ nm ($\lambda = 1552.72$ nm). Two cases for the same sample of chalcogenide nonlinear FBG were studied in more details: the incident intensity $I_{in} = 100$ MW/cm^2 (Fig. 4) and $I_{in} = 230$ MW/cm^2 (Fig. 5).

At the incident intensity of 100 MW/cm^2 the transmittance is very low and the system behaves almost as linear. However from Fig. 4 we can clearly see that nonlinear variation of the refractive index shifts the transmittance spectrum toward longer wavelengths and the change of the shape of the spectral curve (SC) occurs. We observe the SC narrowing for wavelengths below but near to λ_{Bragg} ($\Delta\lambda < 0$) and enlarging for wavelengths above λ_{Bragg} ($\Delta\lambda > 0$). The side lobes are shifted to higher $\Delta\lambda$.

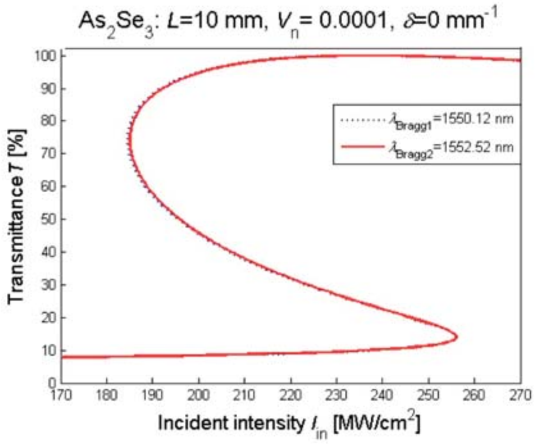


Fig. 3 FBG transmittance versus the incident intensity calculated as the time-independent solutions of NCMEs. The comparison of the optical bistability curves for chalcogenide FBG at the Bragg wavelengths of λ_{Bragg1} = 1550.12 and λ_{Bragg2} = 1552.52

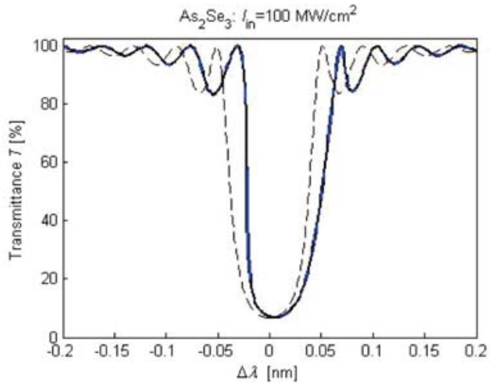
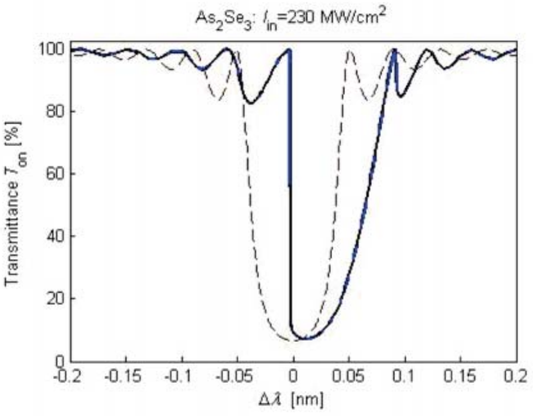


Fig. 4 Transmittance spectra of the As_2Se_3 nonlinear FBG at the incident intensity of 100 MW/cm^2 (solid line) and for the linear FBG of the same parameters (the dashed line)

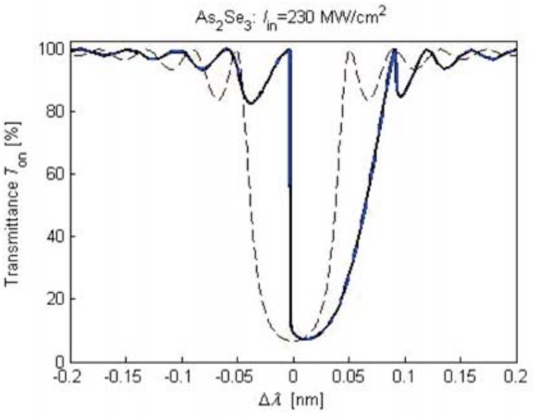
The incident intensity of 230 MW/cm^2 corresponds to the case when two transmittance states are possible for the same incident intensity. Fig. 5a) corresponds to the "ON" state, Fig. 5b) to the "OFF" state. It is obvious that the nonlinearity causes the distortion of the spectrum in the "ON" state and the distortion and the shift toward higher $\Delta\lambda$ in the "OFF" state. The spectral change between "ON" and "OFF" states was found to be $\Delta\lambda \sim 0.019$ nm that is smaller than the WDM channel spacing. Therefore, no problems with the undesirable channel switching between "ON" and "OFF" states are expected.

C Investigation of spectral responses by the temperature influence

Here we investigate the impact of ambient temperature on the nonlinear chalcogenide FBG spectrum. The same sample parameters as in Table 1 were used. Figure 6 shows the OB curve for five different values of the ambient temperature. We observe that the temperature change has a great influence on the curve OB and the width of the hysteresis of on – off contrast. In Fig. 7 the bistable behavior for the input intensity of 255 MW/cm^2 was simulated at "ON" and "OFF" state at varied ambient temperatures.

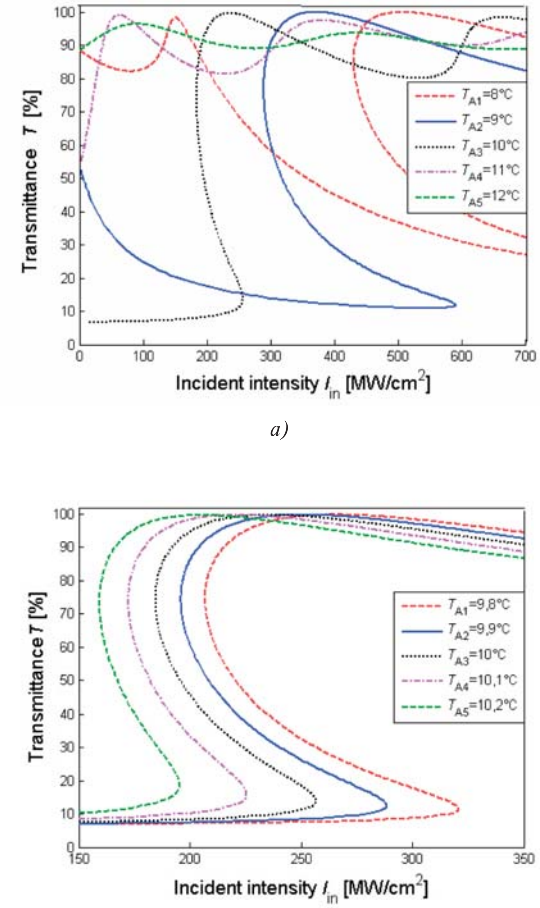


a) "ON" state



b)"OFF" state

Fig. 5 Transmittance spectra in the "ON" (a)) and "OFF"(b)) states at the incident intensity of 230 MW/cm^2 (solid line) and for the linear FBG of the same parameters (the dashed line)



a)

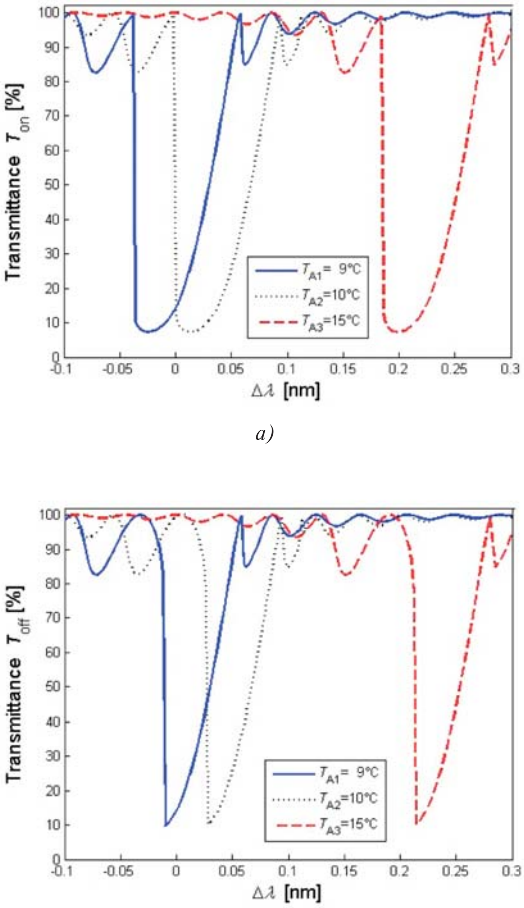
b)

Fig. 6 FBG spectrum at varied ambient temperature
a) T_A in the range from 8 °C to 12 °C and b) T_A in a narrow temperature range from 9.8 °C to 10.2 °C

From Fig. 7 we conclude that spectral curves do not change in shapes, but the shift occurs. If the temperature is increased by 5 °C, there is a shift of ~ 0.186 nm. The variation of approx. 12 °C shifts the spectrum by more than 0.4 nm. This value corresponds to the spacing of the 50 GHz WDM grid in the vicinity of the 1550 nm wavelength. Then the FBG temperature change may affect FBG spectrum and channel switching in WDM systems may occur.

4. Conclusion

In this article, optical switching based on optical bistability in nonlinear chalcogenide FBG is investigated. The sample of As_2Se_3 is numerically studied as a FBG medium. The nonlinear coupled mode equations are applied to analyze the optical bistable behavior. Numerical results are gathered using the finite difference method. The shape of the hysteresis loop is considered from the point of view of optical switching. The influence of the detuning of the carrier and the Bragg wavelength on the optical bistability behavior and the threshold of the incident switching intensity is showed. Spectral responses of the investigated FBGs are numerically studied and shown to be distorted at some investigated input intensities causing the nonlinearities. We found that the temperature change the linear shift in the spectrum of wavelengths that can affect switching channels in the WDM systems.



a)

b)

Fig. 7 FBG spectrum at varied temperatures, a) "ON" state and b) "OFF" state

Acknowledgment
This work was partly supported by the Slovak Grant Agency under the project No. 1/1271/12.

References

[1] DADO, M., KRAJCI, S., DUBOVAN, J., SABOL, D.: Numerical Investigation of Optical Burst Switching. *Communication – Scientific Letters of the University of Zilina,* No. 2, pp. 20–24, 2008

[2] HENKER, R. et al.: A Review of Slow- and Fast- Light Based on Stimulated Brillouin Scattering in Future Optical Communication Networks. *Communication – Scientific Letters of the University of Zilina,* pp. 45–52, No. 4, 2008

[3] WINFUL, H. G., MARBURGER, S. D., GARMIRE, E.: Theory of Bistability in Nonlinear Distributed Feedback Structures, *Appl. Phys. Lett.*, 35 (5), 1979

[4] YOSIA, Y., PING, S.: Optical Bistability in Periodic Media with Third-, Fifth-, and Seventh-Order Nonlinearities, *J. Lightwave Technol.,* vol. 25, pp. 3875–3882, 2007

[5] HARBOLD, J. M.,ILDAY, F. O., WISE, F. W. and AITKEN, B. G.: Highly Nonlinear Ge-As-Se and Ge-As-S-Se Glasses for All-Optical Switching, *IEEE Photonics Technology Letters,* vol. 14, No. 6, pp. 822–824, June 2002

[6] SANGHERA, G. J. S. et al.: Nonlinear Properties of Chalcogenide Glass Fibers, *Optoelectron. Adv. M,* vol. 8, No. 6, pp. 2148–2155, 2006

[7] ZAKERY, A., ELLIOT, S. R.: Optical Nonlinearities in Chalcogenide Glasses and their Applications, *Springer-Verlag Berlin Heidelberg,* 2007

[8] ERDOGAN, T.: Fiber Grating Spectra, *J. of Lightwave Technology,* vol. 15, No. 8, 1997

[9] LI, W., Y.W. LI, X.D. HAN, G.Q. YU: *The Study of Enhancing Temperature Sensitivity for FBG Temperature Sensor,* 8th Intern. Conference on Machine Learning and Cybernetics, Baoding, 2007

[10] MAHMOUD, M. and GHASSEMLOOY, Z.: Tunable Fiber Bragg Gratings Modeling and Simulation, *IEEE, Computer Society,* 2003

[11] KAMIKAWACHI, R.C., KALINOWSKI, H.J., MULLER, M., FABRIS, J.L.: Thermal Behaviour of an Etched FBG Immersed in Different Surroundings. *J. of Microwaves, Optoelectronics and Electromagnetic Applications,* vol. 6, No. 2., 2007

[12] BOUDEBS, G., et. al.: Nonlinear Optical Properties of Chalcogenide Glasses at Telecommunication Wavelength Using Nonlinear Imaging Technique, *Transparent Optical Networks,* vol. 2, pp. 145–150, 2004

[13] IMAI, M., SATO, S.: Optical Switching Devices Using Nonlinear Fiber-Optic Grating Coupler, *Photonics Based on Wavelength Integration and manipulation, IPAP Books 2*, pp. 293–302, 2005

[14] CARVALHO, J.C.C. et al.: A New Acceleration Technique for the Design of Fiber Gratings, *Optical Society of America, Optic Express,* vol. 14, No. 22, pp.10715-10725, 2006

[15] MISSINNE, J., HOE, B.: Artificial Skin Based on Flexible Optical Tactile Sensors, *SPIE Newsroom,* DOI: 0.1117/ 2.1201001.002582, 2010

[16] MAHMOUD, M., GHASSEMLOOY, Z: Tunable Fiber Bragg Gratings Modeling and Simulation. *IEEE, Computer Society,* 2003

[17] PENG, Y., QIU, K.,WU, B., JI, S.: Study on Characteristics of Optical Bistable Devices Based on Fiber Bragg Grating, *Passive Components and Fiber-based Devices VI, SPIE-OSA-IEEE, Proc. SPIE,* vol. 7630, 2009

[18] LEE, H., AGRAWAL, G. P: Nonlinear Switching of Optical Pulses in Fiber Bragg Gratings, *IEEE J. of Quantum Electronics,* vol. 39, No. 3, 2003

[19] STERKE, C.M., JACKSON, K.R., ROBERT, B.D.: Nonlinear Coupled Mode Equations on a Finite Interval: a Numerical Procedure, *J. Opt. Soc. Am B,* vol. 8, No. 2, 403–412, 1991

[20] GUILLEVIC, E. et al.: Optimatization of Chalcogenide Glass in the As-Se-S System for Automative Applications. *OPT MATER,* pp. 1688–1692, vol. 31, No. 11., 2009.

Daniel Kacik – Peter Tatar *

PHOTONIC CRYSTAL FIBER MODAL INTERFEROMETER FOR REFRACTIVE INDEX SENSING

We present a modification of core-cladding intermodal interferometer for refractive index sensing. Its structure consists of double cladding photonic crystal fiber (PCF) as a sensing element. First cladding allows fundamental mode propagation only. Second cladding consists of ring of holes with the average refractive index higher than the basic material of the fiber and it is situated close to interface fiber-external environments. So the cladding mode is confined in the ring of holes and its evanescent field extends to external environment. Advantages of this arrangement against core – cladding interferometers are higher sensitivity and, by suitable choice of refractive index of second cladding holes, the possibility to investigate the external refractive index in the region where phase constants of interfering modes shows an extreme (equalization wavelength) in the spectral region.

Keywords: *Intermodal interference, optical fiber sensor, photonic crystal fiber, refractive index sensing.*

1. Introduction

There has been great interest in optical sensors in recent days. A large attention is given to sensors based on Bragg gratings and periodical structures made inside the fiber [1], or other optical materials [2] but also the fiber sensors based on intermodal interference [3] has the potential. For construction of these sensors optical fibers with different geometrical and propagation parameters can be used [4, 5]. Especially interesting fibers for sensor construction based on intermodal interference are those with photonic structure [6] due to the unique guiding mechanisms. In photonic crystal fibers (PCFs) can be observed the interference of two core modes [7], or the interference of two modes spreading in different areas of fiber. First approach is not very suitable for sensor construction due to critical launching and polarization conditions. In contrast, sensors based on second approach enjoy greater popularity in the sensor community. Sensors based on interference of modes spreading in different cores (two cores and more cores fibers) [8] have low temperature dependence due to the same material and are suitable for measurement of parameters as a curvature, a strain and a pressure. But usually, these structures are not suitable for measurement of external refractive index. In contrast, the interference of core and cladding mode is suitable, except the above-mentioned parameters, to measure the refractive index [9].

Here, we propose a modification of core-cladding intermodal interferometer for refractive index sensing based on double cladding photonic crystal fiber as a sensing element. First cladding allows fundamental mode propagation only (endlessly single mode propagation). Second cladding consists of ring of holes located close to the cladding-surroundings interface. Since the average refractive index of the ring is higher than basic material of the fiber the "cladding" mode propagates through the holes. This makes the transmission of the sensing element dependent on the refractive index of the external environment. By suitable choice of refractive index of second cladding holes, it is possible to investigate the external refractive index in the region where phase constants of interfering modes show an extreme [10] – center of interference (equalization wavelength) in spectral region. The sensing element can operate at different wavelengths as well as coated with thick films made of variable index materials.k

2. Intermodal interference

Light in an optical fiber propagates by means of modes of electromagnetic fields. A particular mode propagates with phase constant β which depends on parameters of an optical fiber surroundings and can be different for particular modes. The phase constants difference allows observing intermodal interference under special conditions.

The signal s generated by a quadratic detector at the end of the fiber (length of z) can be expressed as

$$s(z) = \int_S c(x,y) \cdot \sum_i \psi_i(x,y,z) \cdot \sum_i \psi_i^*(x,y,z)\,dxdy \qquad (1)$$

where $c(x,y)$ is the detector sensitivity, $\psi_i(x,y,z)$ are the functions describing the propagating modes and are equal to $\psi_{i\ i,0}(x,y) \cdot \exp(j\beta_i z)$, where $\psi_{i,0}$ are the modal functions, β_i are the phase constants of particular modes, x and y are the coordinates perpendicular to the direction of the propagation, S is area on which the modal function is nonzero and * denotes complex conjugation.

---

* **Daniel Kacik, Peter Tatar**
Department of Physics, University of Zilina, Slovakia, E-mail: daniel.kacik@fel.uniza.sk

Expression (1) can be rewritten as

$$s(z) = \int_S c(x,y)\cdot\sum_k \psi_{k,0}(x,y)\cdot\psi^*_{k,0}(x,y)\cdot dxdy + \\ + \int_S c(x,y)\cdot\sum_{l\neq k}\psi_{l,0}(x,y)\cdot\psi^*_{k,0}(x,y)\cdot\exp(j\beta_{lk}z)\cdot dxdy \qquad (2)$$

$\Delta\beta_{lk} = \beta_l - \beta_k$ is phase constants difference. The first integral of Eq. 2 represents the sum of particular mode intensities. The second integral of Eq. 2 is the interference term. It is clear that while the sensitivity c does not depend on coordinates its value is zero, because

$$\int_S \psi_{l,0}(x,y)\cdot\psi^*_{k,0}(x,y)dxdy = 0 \quad \text{for } l\neq k \qquad (3)$$

It means that the intermodal interference can be observed only when the detector sensitivity is not uniform or when the eigenfunctions are transformed into non-orthogonal functions.

We let $\phi = \Delta\beta z$ be the total phase difference accumulated along the intermodal interferometer.

While an external perturbation is imposed on the sensing element phase changes by the amount of $d\phi$. The physical quantities (temperature, strain, pressure and refractive index n of surrounding medium) can lead to phase changes.

The resulting phase change depending on the refractive index of surrounding medium can be written as [11]

$$d\phi = \frac{2\pi}{T_n}dn \qquad (4)$$

where T_n is the quantity of refractive index that introduces a 2π phase change. So, the smaller the T_n, the greater the sensitivity.

3. Methodology

The proposed structure of double cladding photonic crystal fiber is shown in Fig. 1.

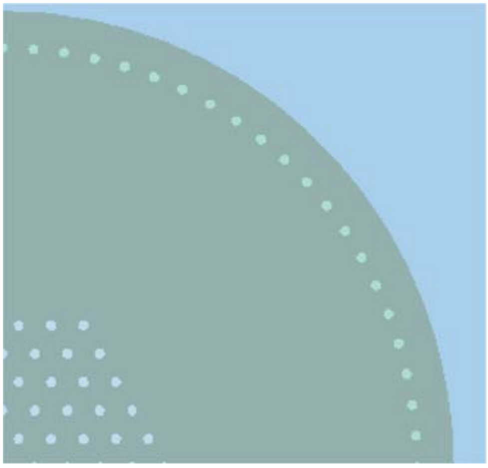


Fig. 1. The double cladding structure of PCF in quadrant (0,π/2)

Diameter of core is 9 μm, the first holey region diameter is 65 μm and the second cladding, consisting of ring of holes, is located close to the cladding-surroundings interface. First cladding holes are filled by air with refractive index 1. Only one mode is propagated by the core of this structure. Refractive index of second cladding holes is 1.4602. So the average refractive index of the ring is higher than basic material and light (or cladding mode) is propagated through the ring (Fig. 3). Since the ring is close to cladding-surrounding interface the change of refractive index of surrounding will change the phase constant of mode propagation through the ring and thus it changes the character of modal interference.

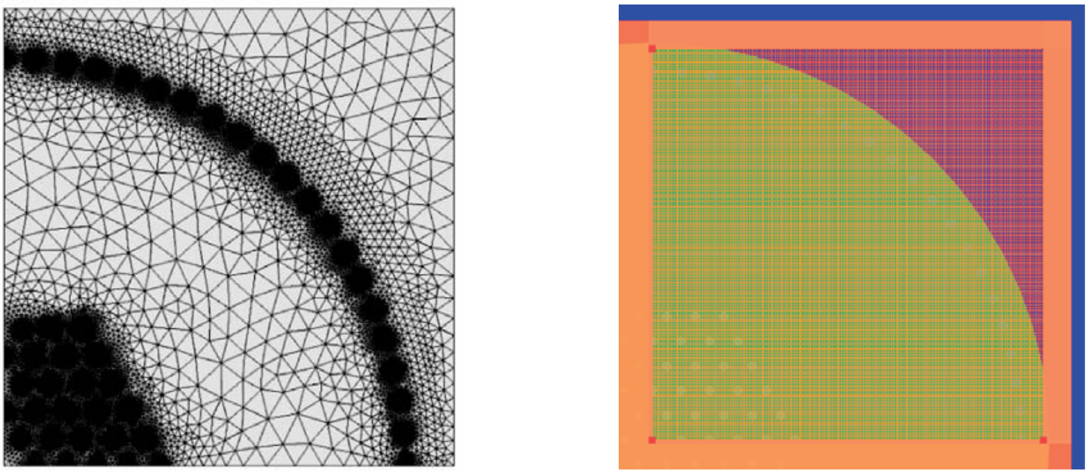
To calculate modes in PCFs one needs to use numerical methods because of the complexity of the structure in PCFs. There have been several numerical methods invented throughout the time period of PCFs and even before. Some of them solve the wave equation and others solve directly the Maxwell equations without any assumptions for dielectrics. Although these methods differ they have one feature in common – they all use spatial mesh which discretizes simulated structure or simulation area into a finite set of points. Naturally, methods for discretization also differ and each is suitable for particular problem. The most known is the Finite Element Method (FEM) which discretizes the structure based on its complexity (Fig. 2a). From the user's view of point it is possible to setup the number of calculated points per domain. Second most used method for creating mesh is regular mesh with equal differences between neighboring points. This type of mesh does not depend on the shape of structure and usually has option of override mesh over complex regions (Fig. 2b).

One of the approaches of solving electromagnetic problems is Finite Difference Time Domain (FDTD), based on Yee algorithm (5). The mathematical relationship of the electromagnetic fields radiated by time-dependent current or charge densities is governed by Maxwell's equations. These are discretized using central difference approximations to the space and time partial derivates. The resulting finite-difference equations are solved. So we simplify the vectorial equation taking just one component and rewrite it to discrete form taking the central difference approximations for both temporal and spatial derivatives [11]

$$\frac{E_x^{n+1/2}(k) - E_x^{n-1/2}(k)}{\Delta t} = -\frac{1}{\varepsilon_0}\frac{H_y^n(k+1/2) - H_y^n(k-1/2)}{\Delta x} \\ \frac{H_y^{n+1}(k+1/2) - H_y^n(k-1/2)}{\Delta t} = \frac{1}{\mu_0}\frac{E_x^{n+1/2}(k) - E_x^{n+1/2}(k)}{\Delta x} \qquad (5)$$

We used FDTD method based commercially available software MODE solutions (Lumerical Inc.) to perform the simulation.

The whole process of simulation in this software enviroment consists of two main parts. First is the layout part for drawing the structure, setting the material properties (Sellmeir coefficients and experimental data in our case), setting mesh and definition of boundary condition (PML). The other part is the analysis, where the modal properties of the structure are calculated. In purpose of intermodal



a)

b)

Fig. 2. Finite Element Method mesh (COMSOL 4.2.0), Regular mesh (MODE solutions 3.5a)

interference investigation it is necessary to choose suitable modes. In case of core mode it's quite simple, but cladding mode can be chosen only based on some physical parameter like polarization, overlap value etc. After that we are able to sweep the mode properties through the frequency or wavelength spectrum.

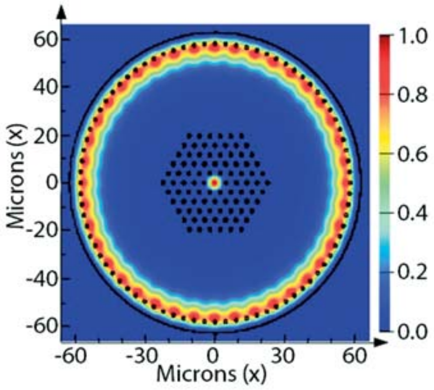


Fig. 3. Power flow distribution in linear scale of core (fundamental) and selected cladding mode for wavelength 1600 nm

4. Results and discussion

Since the mode in the second cladding is guided in the waveguide that is formed by glass – air interface the change of surrounding refractive index will also change its phase constant. This causes a shift in the interference fringes. Immersing double cladding PCF into the water ($n = 1.317$), we created glass – water waveguide and we simulated the spectral response. Similarly, we replace the water for higher refractive index substance ($n = 1.367$). Comparison of spectral responses is shown in Fig. 4.

In the case of glass–air interface the extreme of phase constants difference appears at 952.12 nm (blue curve), for glass–water interface is the extreme at 941.06 nm (red) and for glass – substance interface at 935.64 nm (green). If we consider the length (1.8 cm) of double cladding PCF, the interference between these two modes could be interpreted in the form of interference terms (Fig. 5).

Equalization wavelength dependence on external refractive index changes is shown in Fig. 6. There is also drawn a prediction of the dependence.

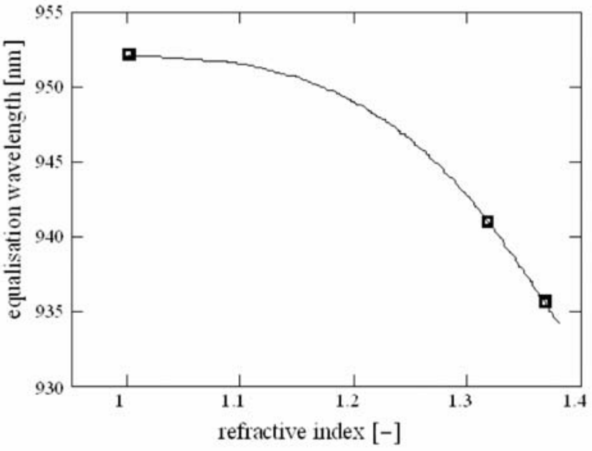


Fig. 6 Equalization wavelength dependence on external refractive index changes (points) and cubic spline interpolation

Similar to core-cladding interferometers the structure is insensitive (weakly sensitive) to changes of the refractive index of air [9]. But the structure is significantly more sensitive to changes of

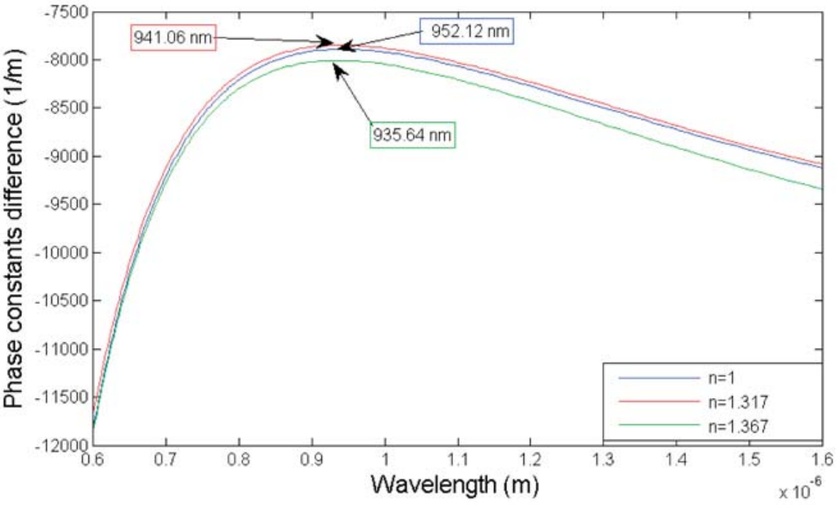


Fig. 4. Spectral Dependence of phase constants difference for different external refractive index

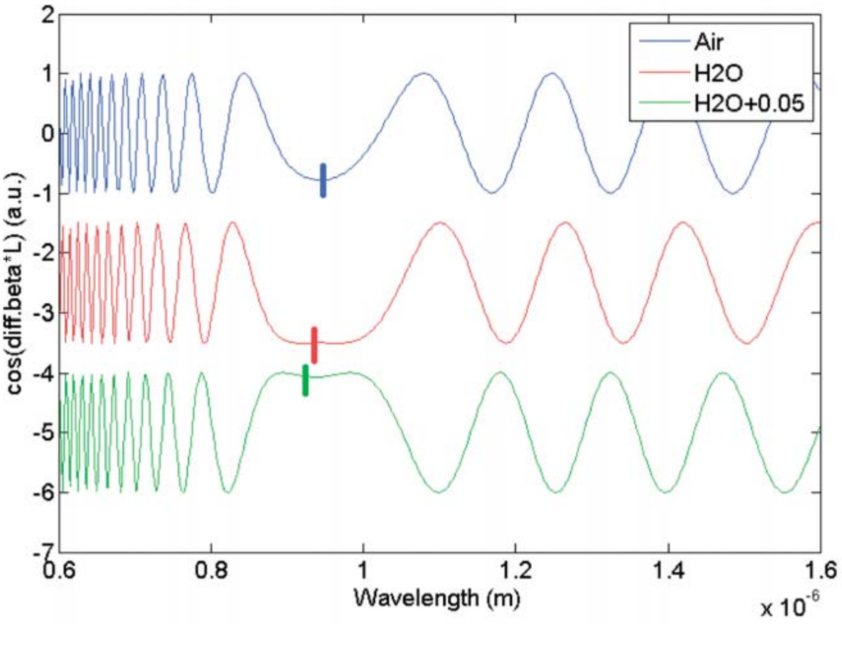


Fig. 5 Interference terms for glass-air(a), glass-water(b), glass-water+0.05(c) interface; curves are vertically shifted

refractive index of water (substances based on water). The equalization wavelength can be determined with accuracy better than 1 nm [12] so it is possible to determine the changes at 0.0095 refractive index units.

5. Conclusion

We propose a modification of core-cladding intermodal interferometer based on double cladding photonic crystal fiber as a sensing element. Main advantage of this structure is the possibility to investigate the external refractive index in the region where phase constants of interfering modes show an extreme. Moreover, data obtained from the simulation shows that the sensor element can also be used to measure the refractive index of water-based substances where conventional core–cladding interferometers are insensitive.

Acknowledgements

This work was partly supported by Slovak National Grant Agency No. VEGA 1/1058/11 and No. VEGA 1/0528/12. The authors wish to thank for the support to the R&D operational programs Centre of excellence of power electronics systems and materials for their components I. and II. No. OPVaV-2008/2.1/01-SORO, ITMS26220120003, No. OPVaV-2009/2.1/02-SORO, ITMS 26220120046 funded by European regional development fund (ERDF).

References

[1] KERSEY, A. D., DAVIS, M. A., PATRICK, H. J., LEBLANC, M., KOO, K. P., ASKINS, C. G., PUTNAM, M. A., FRIEBELE, E. J.: Fiber Grating Sensors, *J. of Lightwave Technology*, 15 (8), pp. 1442–1462, 1997.

[2] TARJANYI, N.: Real-time Imaging of Grating Formation in LiNbO3:Fe using Mach–Zehnder Interferometer, *Opt. Eng.*, vol. 49, No. 8, 085602, 2010.

[3] STATKIEWICZ-BARABACH, G., CARVALHO, J. P., FRAZAO, O., OLSZEWSKI, J., MERGO, P., SANTOS, J. L., URBANCZYK, W.: Intermodal Interferometer for Strain and Temperature Sensing Fabricated in Birefringent Boron Doped Microstructured Fiber, *Applied Optics,* vol. 50, No. 21, pp. 3742–3749, 2011.

[4] TRIPATHI, S. M., KUMAR, A., VARSHNEY, R. K., KUMAR, Y. B. P., MARIN, E., MEUNIER, J. P.: Strain and Temperature Sensing Characteristics of Single-mode-multimode-single-mode Structures, *J. of Lightwave Technology,* vol. 27, No. 13, pp. 2348 –2355, 2009.

[5] IVANOV, O. V.: Fibre-optic Interferometer Formed by a Section of Small-core Fibre Spliced Between Standard Fibres, *Optics Communications* 282, pp. 3895–3898, 2009.

[6] KACIK, D., MARTINCEK, I., PUDIS, D., TARJANYI, N., TUREK, I.: Photonic Crystals – Optical Structures for Advanced Technology, *Communications – Scientific Letters of the University of Zilina,* No. 2, 2008, pp. 25–29, 2008.

[7] KACIK, D., TUREK, I., MARTINCEK, I., CANNING, J., ISSA, N. A., LYYTIKAINEN, K.: Intermodal Interference in Photonic Crystal Fibre, *Opt. Express* 12, 3465-3470, 2004.

[8] MACPHERSON, W.N., GANDER, M. J., MCBRIDE, R., JONES, J. D. C., BLANCHARD, P. M., BURNETT, J. G., GREENAWAY, A. H., MANGAN, B., BIRKS, T. A., KNIGHT, J. C., RUSSELL, P. ST. J.: Remotely Addressed Optical Fibre Curvature Sensor using Multicore Photonic Crystal Fiber, *Optics Communications,* vol. 193, No. 1–6, pp. 97–104, 2001.

[9] KACIK, D., TVAROZEK, P., MARTINCEK, I., SCHUSTER, K.: Refractive Index Measurement based on Core-cladding Mode Interferometry in Endlessly Single Mode Fiber, *Optik* 123, 1746–1749, 2012.

[10] MARTINCEK, I., PUDIS, D.: A Theoretical Study of the Temperature Sensor Based on the LP01-LP02 Intermodal Interference in Optical Fiber with a Liquid Core, *Communications – Scientific Letters of the University of Zilina,* No. 2, 2010, pp. 10–13, 2010.

[11] BOCK, W. J., EFTIMOV, T. A., MIKULIC, P., CHEN, J.: An Line Core-cladding Intermodal Interferometer using a Photonic Crystal Fiber, *J. of Lightwave Technology* 27, pp. 3933–3939, 2009.

[12] MARTINCEK, I., PUDIS, D., KACIK, D., SCHUSTER, K.: Investigation of Intermodal Interference of LP01 and LP11 Modes in the Liquid-core Optical Fiber for Temperature Measurements, *Optik – J. for Light and Electron Optics,* vol. 122, No. 8, pp. 707–710, 2011.

Pavel Pribyl – Vit Fabera – Vladimir Faltus – Lukas Tyfa *

SEMANTIC WEB RELATED TO ITS SYSTEMS

Abstract: This paper describes a way to better navigate in an ever increasing number of standards in the field of ITS. Short overview about fundamental ITS models is presented, the principles of related knowledge system are described. Ontology approach is used, especially as domain-oriented ITS ontology, where a new hybrid method is explained.

Keywords: *ITS, standardization, ontology, knowledge system, knowledge unit.*

1. Problem declaration

The elaboration of ITS (Intelligent Transport Systems) standards strongly supported by common European policy is relatively successful. It is possible to identify about 180 active standards within European standardization committee CEN/TC278 at present. Together with ISO/TC204, there are more than 310 standards. Some of them are being elaborated, some are under revision and a significant part of standards has been finished. They are focused on the different market applications, starting from CEN/TC278/WG1 "Electronic Fee Collection" (EFC) up to the ISO/TC204/WG16 "Wide Area Communications/Protocols and Interfaces" (CALM) communication standards as examples.

There are two basic problems limiting successful utilization of ITS standards in real praxis. They are very complex (in general, a standard often has almost 100 pages). There are more or less 20 000 pages of ITS standards, which is a huge quantity for comprehensive reading. The standards are frequently written in the Universal Modelling Language (UML) or eXtensible Markup Language (XML) conventions and they need an advanced reader. Even for experts in ITS standardization it is often difficult to follow such a complex set of standards. Generally said, standards are not readable at all, it is a problem especially for investors and decision makers.

This article gives short overview about fundamental models describing complex ITS systems. Briefly describes principles of knowledge system developed in the frame of ZNALSYS project supported by Ministry of Transport of the Czech Republic. Main focus is on ontology; especially domain-oriented ITS ontology. New hybrid method is explained. It uses not only terms creation based on different ITS dictionaries but ontology is completed by terms coming from object-oriented models of standards describing processes within them.

2. User-oriented web applications

The fundamental goal of presented research is linked to end user who needs to find out appropriate information within set of standards. The ontology enables to create semantic meaning of the term which is under his/her consideration. Web searching mechanism looks for these qualified questions. Their quality is done by the quality of ontology.

The aim of the user-oriented web application is to satisfy users and to ensure their positive experience with the application. Reference [1] shows that the users never spend more than 50 % of their on-line time with any web application. Any user is more satisfied if he is carried out through given subject by fixed rules. Domain-oriented ontology provides a fixed taxonomy with well-defined semantic meaning of terms as shared framework of standards. Web-oriented search system was tested in practice in the frame of ZNALSYS research project supported by Czech Ministry of Transport. Brief look at the semantically oriented web-searching system is presented in chap. VII.

3. Model of ITS systems

ITS systems (IT systems generally) are modelled by data models. These models identify and describe a part of real word formally (with help of mathematics, formal text, graphs ...) using terms like entities, relationships etc. Data models are created on some abstraction levels:

- *Conceptual model* (which is also called the semantic one) is situated at the highest level. The model is not dependent on the specific programming language or software/hardware platform. Crucial entities and relationships are identified in this model. The typical representative is just E-R model (Entity-Relationship model).

---

* **Pavel Pribyl[1], Vit Fabera[2], Vladimir Faltus[2], Lukas Tyfa[1]**
[1] Department of Transporting Systems, Faculty of Transportation Sciences, Czech Technical University, Prague, Czech Republic,
E-mail: pribyl@fd.cvut.cz
[2] Department of Transport Telematics, Faculty of Transportation Sciences, Czech Technical University, Prague, Czech Republic

- *Logical model* is expanded conceptual model and it is created by adding specifications of entities, attributes and relationships. Ranges of values and parameters of attributes are specified. The model is still independent on the technological platform.
- *Physical model* is adapted to the realization. Structures suitable for implementation are created and technical details are added to realize IT.

Three well-known data models are:
- *E-R model (Entity-Relationship model)* is a conceptual model used in the database area. They are described by E-R diagrams. E-R diagrams contain entities, relationships and attributes. Top-down method is usually applied when E-R model is created (modelled subject is decomposed into smaller parts until indivisible entities are identified).
- *Relation model* is derived from E-R model. All entities and relations are implemented by (database) relations. Relational algebra defines operations over relations.
- *Object-oriented data models* are used to describe real world systems when objects are searched to be similar to objects in real world. Object is a real or abstract entity containing data (attributes) and descriptions how to manipulate with them (methods). The object is defined by its state, behaviour and its identity.

Object-oriented models are notated with UML. UML uses structure diagrams (Class diagrams, Component diagrams ...) and behavioural diagrams (Activity diagrams, Use case diagrams ...).

In connection with other considerations, we will stream to create a semantic domain-oriented model of ITS. To do this exactly we will use object-oriented models to be more precise.

4. Role of ontologies

The ability to search and fuse information from heterogeneous standards significantly contributes to the discovery of added value knowledge that is unreachable using classical searching methods. In order to provide an efficient information mining from standards a knowledge-based model is an optimal solution. Using an ontology approach, a coherent, consistent and non redundant knowledge model could be designed.

In information science, ontology is defined as a formal representation of knowledge as a set of individuals (instances, terms), classes (concepts), attributes, and relations between those concepts. Instance depicts the basic object and it is indivisible. In the sense of the paper an instance unit of text in an extract could be represented as a semantically uniform term. It is also possible to speak about hierarchical categorization which describes the same model as the ontology is.

In theory, an ontology is a “formal, explicit specification of a shared conceptualization”, [2], expressed as a shared vocabulary, which can be used to model a certain area. There are universal ontologies, as for example SUMO (Suggested Upper Merged Ontology). SUMO ontology creates high-level ontologies for different domains – finance, computers, geography, transport etc. and it is too generic to help us mining appropriate information from standards.

It is possible to divide ontologies into following groups:
- *Terminological ontologies* – they are used especially in librarianship and in next regions oriented to text sources. Terms play crucial role, relations carry taxonomies, i.e. specify relationships of general or special terms.
- *Information ontologies* – they are superstructures over relational-database sources; they ensure conceptual abstraction level necessary to conceptual querying.
- *Knowledge representation (KR) ontologies* – classes and instances are systematically defined using formal language.

According to the subject of formalization it is possible to use next classes:
- *Domain ontologies* – are the most used ones, they are focused on the specified area;
- *Generic ontologies* – they describe general concepts and relations across several areas;
- *Task ontologies* – ontologies that (unlike the other ontologies covering the state) are focused on deduction processes, diagnosis etc.;
- *Application ontologies* – ontologies that are focused on the specific application, they involve both domain and task ontologies.

In the following consideration we will present methodology how to create domain-oriented ontology focused on ITS systems.

A. ITS standards and ontologies

The design of methodology should start under assumption that it is possible to define the ITS systems, including their boundaries over which it is realistic to create ontology. Selected assumptions that influenced the development of methodology are summarized and commented in the following text:
- Analyses show that no complex domain ontological model has been created for the present to conceptualize terms over all ITS area. It is not realistic to finish it in next years due to complexity and dynamics of ITS field.
- Next, generic ontology is too general to be used practically.
- Domain and application ontologies can be used in practice. Ontologies are known to be used in the area of location-based services, [3] or to analyze research documents, [4].
- Analyses created within ZNALSYS project show that it is better to focus on domains which are described by generally accepted form. ITS standards have this form.
- The advantage of ITS standards is their similar syntactic form (structure) – the documents are more or less formalized and have therefore formally similar structure.

Created ontology only describes a fraction of hierarchy of concept of ITS terms in ZNALSYS project. Tree structure of this hierarchy captures subordination of concepts but no additional relations and no additional assertion about concepts and terms. Just relations and terms in the first-order logic are basics of knowledge.

5. Domain-oriented ontology – new approach

Authors describe several methodologies how to construct ontologies, [5]. Presented methodologies both depend on areas to cover by ontologies and on the way how the reality is informally described. The METHONTOLOGY, [6], is a suitable technique for our purposes. This methodology describes the ontology creation in eight tasks (steps). The first four tasks represent coarse steps of the ontology creation process, the tasks number 5 – 8 determine how to define relations, attributes etc. in details.

The first possibility is to build a glossary of terms to be included in the future ontology. This task has been already done for ITS because the Czech and English explanatory dictionary of ITS terminology had been published, [7]. It will also be able to extract Concept taxonomies (task 2) and ad-hoc binary relations (task 3) from this dictionary.

ITS Standards are specific and well-structured documents. Our idea is to find sets of typical (and analogous) statements in ITS standard and then to collect a set of rules how to transform these typical statements into first order logic sentences.

There is a wide discussion within literature whether is better to create ontology based only on "terminological" procedure (as described above) or is it better to use object-oriented models. The newly proposed procedure is not based on the approach "either/or", but it is complementary approach based on hybrid model. It uses both components but in the three stages:

- (A) *The basic ontology* is created by using terms which are usually defined at the beginning of the standard. They can be corrected and supplemented according to technical dictionaries. The output from this stage is hierarchically organized tree of terms.
- (B) The standards are described as *object-oriented models*, i.e. for a specific process described in a standard the model is created using object-oriented diagrams. Expert or knowledge engineer decides whether to use object-oriented models and which kind of model will be used (structure diagrams and/or behavioural diagrams, e.g. class diagrams, activity diagrams, sequence diagrams, etc.). These models reveal number of new terms.
- (C) Basic ontology (A) is completed by new terms generating by object models (B) and the semantic meaning of all terms is unified.

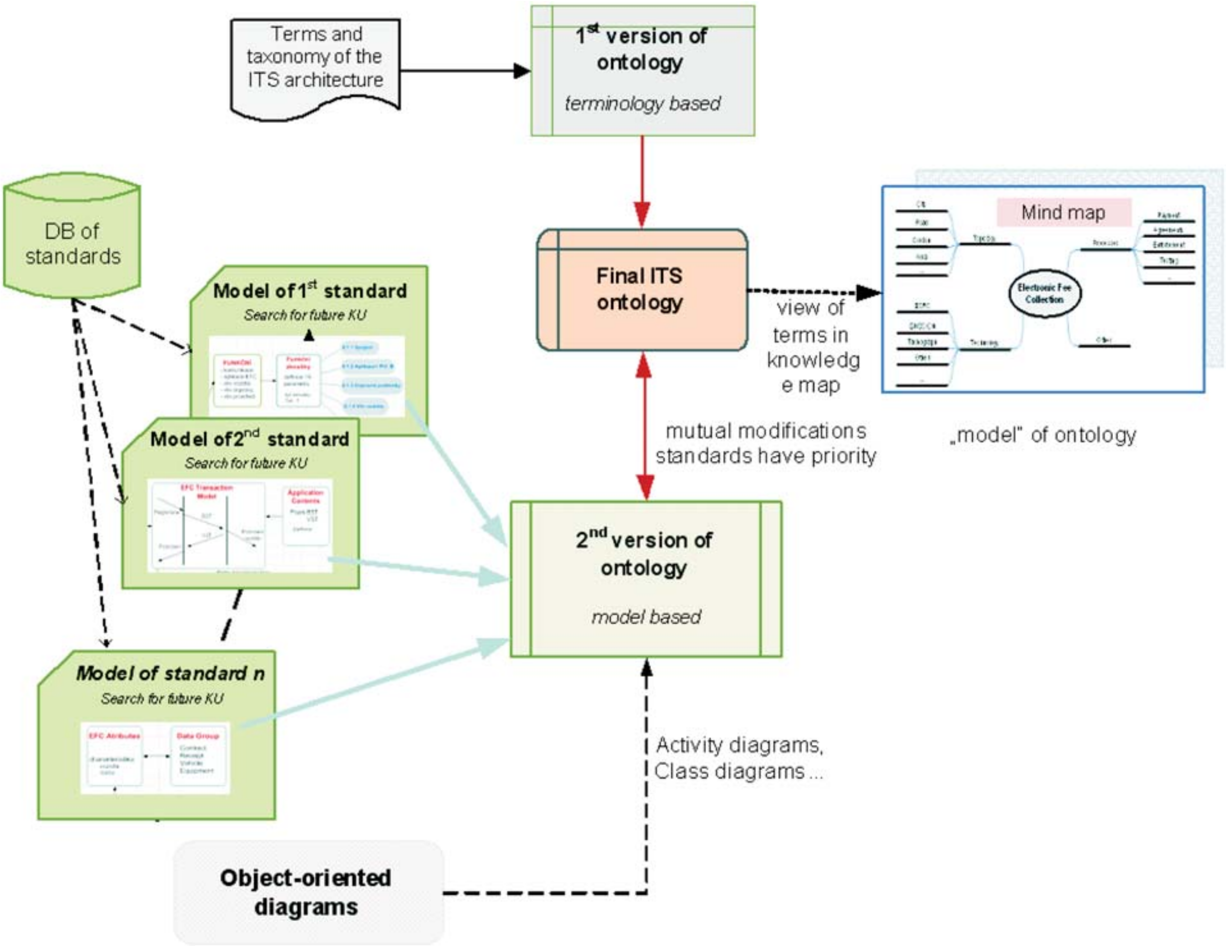


Fig. 1 Hybrid model of creation of domain ontology for ITS

The model created according to this algorithm is valid generally, is well-comprehensible and verifiable. It covers ITS domain more complex by unmasking terms which are not present in primary model (created ad-hoc). Newly found terms are subsequently included in knowledge map. Knowledge map creation is not necessary but its graphical form makes possible back verification of hierarchy and the completeness of the semantic model.

The creation of the basic ontology is depicted in Fig. 1. First stage uses terms defined in standards or in dictionaries. There are minimally two disadvantages: terms coming from standards or dictionaries did not cover the whole area and in addition they have not any hierarchical structure. The structure has to be created more or less intuitively. The advantage is to acquire many new terms. Using Mind Map graphical tools helps to prepare best possible hierarchy and better understand a structure of a system.

Second version of ontology is elaborated on the basis of models of standards. Models are usually created with help of CASE tools, for example Enterprise Architect. It is possible to use static description like Class diagrams or dynamic description (Activity diagrams, State diagrams). The deployment of knowledge expert is necessary to build up suitable diagrams.

On the other hand, these diagrams can detect many disadvantages in created ontology so we consider adding, verify or correct terms on the basis of object-oriented model (OOM) in these hybrid models. If it is very difficult to create OOM due to complexity of concrete standard it is recommended to use different kinds of simplification. This approach is possible to see in Fig. 1. As a result of this stage is that the missing terms are added into ontology.

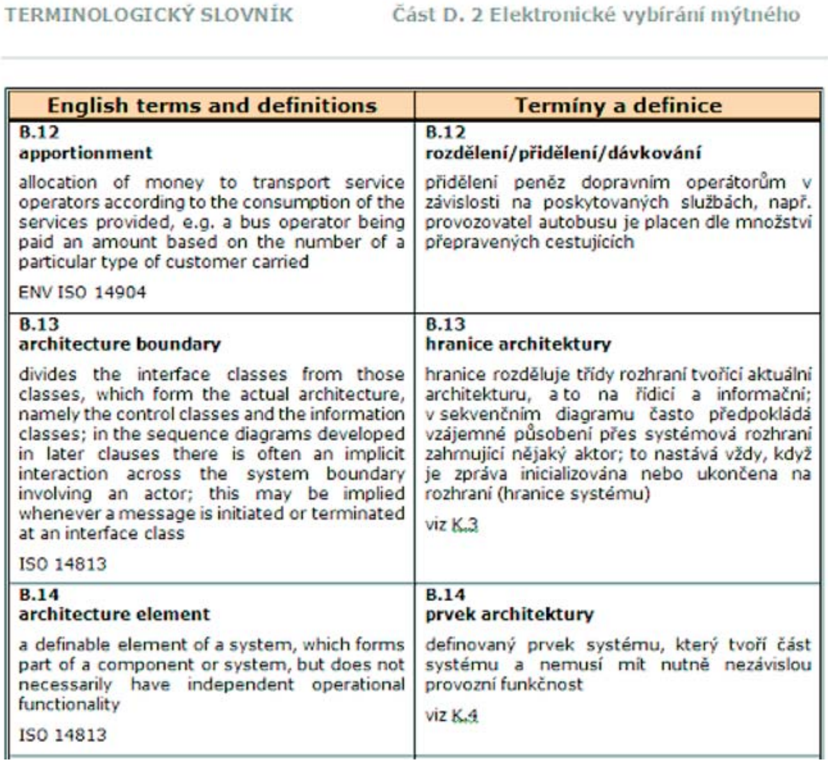
6. Test of proposed method

In the verification stage of proposed method, which should take in account both the above approaches (see A and B) a few standards will be analyzed, i.e. the standards of group CEN/WG1 "Electronic Fee Collection": 14906 "Application interface for DSRC", 14907-1, 2 "Test procedures"; 17574 "Safety frames" and 17575 "GNSS interface". Only first two will be explained more in detail.

A. Domain ontology - terminology based approach

As primary source of terms, design of basic ontology uses standards which cover domain of our interest. Keywords are listed at the beginning of standards. Very good source is also technical report ISO/DTR 14812 "ITS glossary of standard technologies for the transport information control sector". This rather older document (2002) could give good overview on 350 pages.

The basis for Czech ITS domain-oriented ontology is "Terminological dictionary" prepared by SDT (Traffic Telematics Association) which include several hundred (450 pages) of terms organized in the clusters according to WG in the CEN committee. Fig. 2 depicts small part of the chapter "Electronic fee collection".



TERMINOLOGICKÝ SLOVNÍK   Část D. 2 Elektronické vybírání mýtného

English terms and definitions	Termíny a definice
B.12 **apportionment** allocation of money to transport service operators according to the consumption of the services provided, e.g. a bus operator being paid an amount based on the number of a particular type of customer carried ENV ISO 14904	**B.12** **rozdělení/přidělení/dávkování** přiděleni peněz dopravnim operátorům v závislosti na poskytovaných službách, např. provozovatel autobusu je placen dle množstvi přepravených cestujicich
B.13 **architecture boundary** divides the interface classes from those classes, which form the actual architecture, namely the control classes and the information classes; in the sequence diagrams developed in later clauses there is often an implicit interaction across the system boundary involving an actor; this may be implied whenever a message is initiated or terminated at an interface class ISO 14813	**B.13** **hranice architektury** hranice rozděluje třídy rozhraní tvořici aktuálni architekturu, a to na řidici a informačni; v sekvenčním diagramu často předpokládá vzájemné působeni přes systémová rozhrani zahrnující nějaký aktor; to nastává vždy, když je zpráva inicializována nebo ukončena na rozhraní (hranice systému) viz K.3
B.14 **architecture element** a definable element of a system, which forms part of a component or system, but does not necessarily have independent operational functionality ISO 14813	**B.14** **prvek architektury** definovaný prvek systému, který tvoří část systému a nemusí mít nutně nezávislou provozní funkčnost viz K.4

Fig. 2 Terms and their definition (semantic meaning) also in Czech language

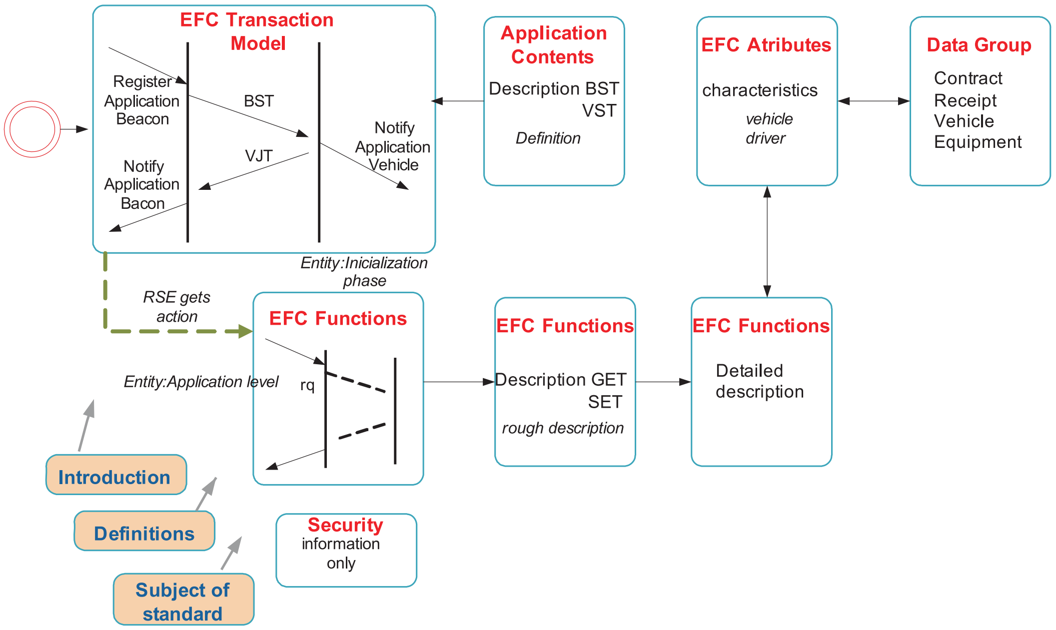


Fig. 3 Structure and logical relation as model of 14906 (product-oriented standard)

Practical experience shows that terms are usually (significantly) different from the following aspects:

- level of detail (while one term describes insignificant detail, other may describe system properties);
- frequency in the documents (importance);
- semantic meaning (that is often different according to the context).

B. Domain ontology – object-oriented model of standards

This chapter clarify topic B – "The standards are described as object-oriented models". ISO DIS 14 906 and 14705-1 will demonstrate the process of simple object-oriented model. The aim is to process simplified model of standard, in the first stage as model of classes as it is known in UML.

In order to distinguish the differences of subjects among multiple standards, the standard 14906 is first analyzed, which almost exclusively addresses the manner and form of communication between on-board unit (OBU) and the infrastructure facilities (RSE). Standard 14 705-1 is focused only on testing of OBU and RSE.

A closer look at Fig. 3, it is evident that the connection is first initiated by "EFC Transaction Model" block whose parameters are transmitted in the "Application Content" block. After establishing and confirming the connection, the data exchange itself follows, while a rough diagram of the transaction is in the first "EFC Functions" block, followed by two blocks of the same name with more and more detailed description of the transmitted data.

"EFC Attributes" block describes parameters that are transmitted in data files and "Data Group" block clusters it into blocks. Relatively small and unspecific mention is about security ("Security" block).

Discussion to "model" of standard 14906:

- This simplified model (more activity diagram than class diagram) gives a good idea about the possible creation of knowledge-based segments linked to appropriate terms.
- There is a problem with a link to ontology – in terms of creating entities in a real standard there is necessary to logically identify potential knowledge unit (KU) and a KU with unambiguous interpretation of the meaning ⇒ it can be inferred that the "blind" creating of model of standard without accordance with pre-built ontology or real needs would not lead to the goal of developing the most competent ontology.
- In this configuration and at this level of resolution it is necessary to still think how competency questions may be formulated. Respective terms need to be targeted to help to prepare competitive question.

Further examined standard, see Fig. 4, describes the test procedures of dedicated short-range communication. The structure is clear, there are three parallel testing procedures consisting of practically the same blocks.

Discussion to "model" of standard 14705-1:

- Model is similar to Classes diagram, but it is not the same. Terms linked to testing can be formulated very well at the highest level.
- If it is necessary to expand a list of terms and to go more in detail, more detailed descriptions of each test will be needed and detailed model shall be formed as Fig. 4 shows.
- Next group of terms can be directed to the requirements for certification and documentation requirements for DSRC testing.

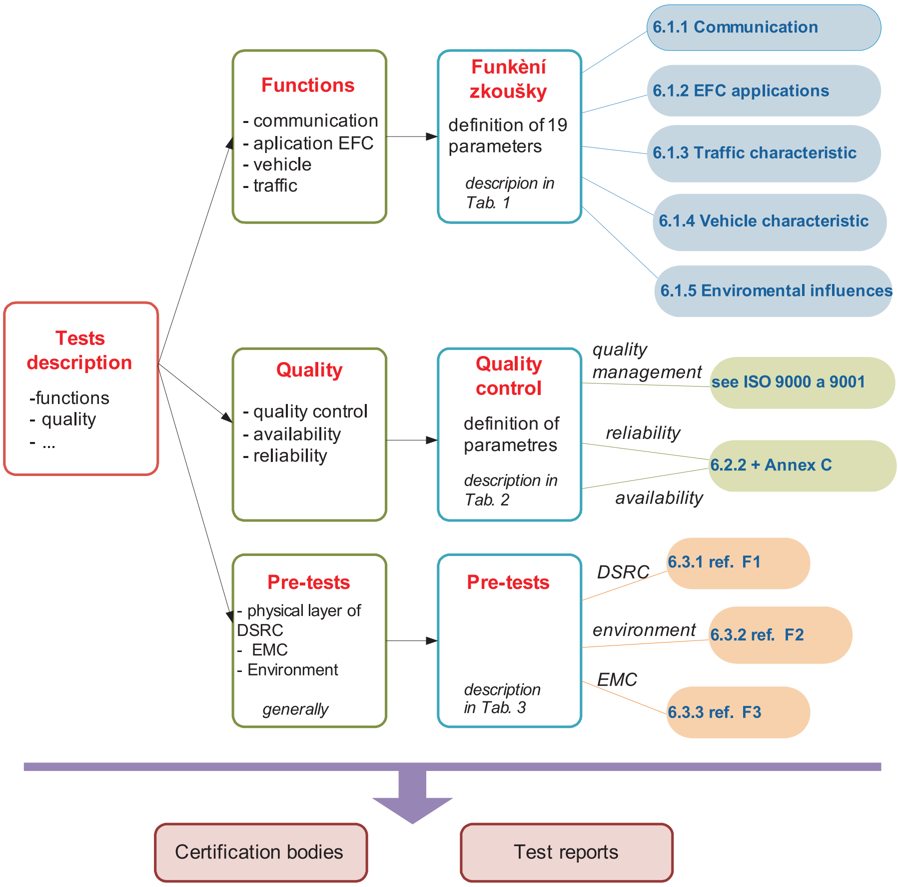
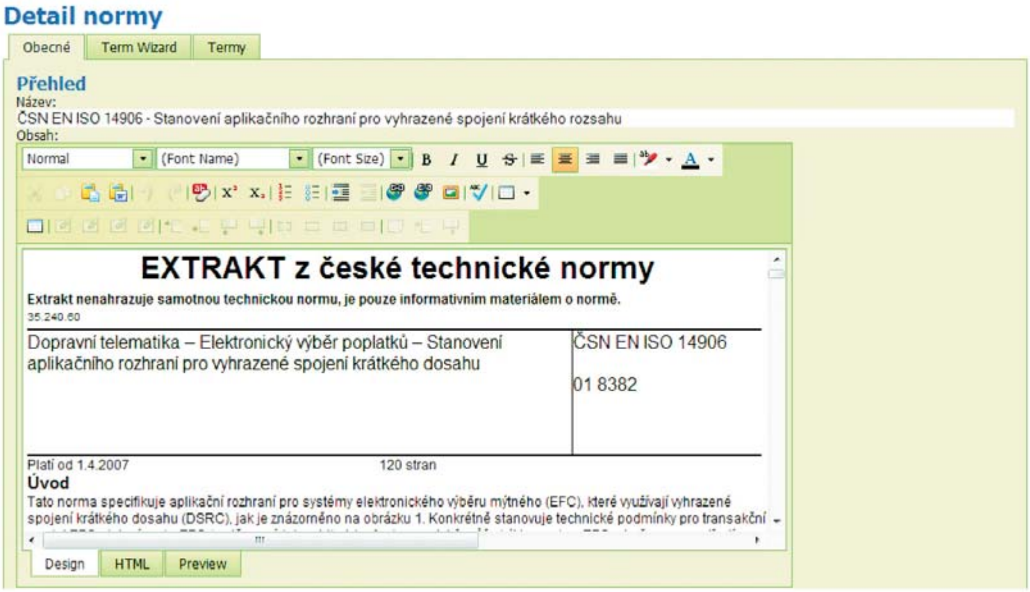


Fig. 4 Structure model of 14 705-1 (standard for testing)



Detail normy

Přehled

Název:
ČSN EN ISO 14906 - Stanovení aplikačního rozhraní pro vyhrazené spojení krátkého rozsahu

Obsah:

EXTRAKT z české technické normy

Extrakt nenahrazuje samotnou technickou normu, je pouze informativním materiálem o normě.

35.240.60

Dopravní telematika – Elektronický výběr poplatků – Stanovení aplikačního rozhraní pro vyhrazené spojení krátkého dosahu	ČSN EN ISO 14906 01 8382

Platí od 1.4.2007 120 stran

Úvod

Tato norma specifikuje aplikační rozhraní pro systémy elektronického výběru mýtného (EFC), které využívají vyhrazené spojení krátkého dosahu (DSRC), jak je znázorněno na obrázku 1. Konkrétně stanovuje technické podmínky pro transakční

Fig. 5 Example of knowledge unit provided by Czech web application

- The disadvantage of this simple model is that the created ontology is rather general. Nevertheless to go more in a detail within different tests described in standard will significantly extend levels of ontology.

7. Pilot application of semantic web

The basic aim of the pilot application was to present knowledge units (well defined short texts of standard linked to ontology) for selected categories of users. The web application promoted the use of knowledge units (KUs) in ITS standards, see Fig. 5, so the ontology model and its units (terms) have offered not only the definitions, but also the links to the knowledge units through competitive questions. The end user found out a set of KUs according to his questions and KUs gave him overview about a problem which is under solution.

The web-oriented search engine was worked out to mind out much focused knowledge contained in the standards. The result was in a form as an on-line application with the possibility to start with a term or uncertain need and to end with a set of knowledge units coming from the source standards.

The web application promoted the use of ITS standards on a simple way available for non-experts too.

8. Conclusion and next works

The article presents the hybrid model of creation of domain-oriented ontology of ITS systems. The basic idea is built on ITS standards. These standards cover significant part of ITS applications. Standards are also very proper due to their context form which is very similar also for very different applications.

In theory, an ontology is a “formal, explicit specification of a shared conceptualization”, expressed as a shared vocabulary, which can be used to model a certain area. ITS systems are usually connected through data interfaces and it works quite well. On the other hand there is not commonly used ontology covering this area. It means that there is no interface on semantic level.

Presented hybrid model uses two basic steps. The first one, more traditional, creates fundamental ontology coming up from different technical dictionaries containing appropriate terms of discussed domain. This basis is completed and verified by new terms derived from object-oriented models within step two. In terms of finding out terms which could complete basic ontology, the simplified models of standards are developed by experts.

The idea of this approach is under discussion with Japan. The ministry MLIT provides some financial support to bi-lateral project oriented to formulate ontology of ITS systems. Czech experts will discuss next development of proposed methodology in Tokyo in April this year.

References

[1] KREMENOVA, I., GALOVIC, M.: Proposal of Key Factors for User-oriented Web Application, *Communications - Scientific Letter of the University of Zilina,* ISSN 1335-4205, 2, 2010, pp. 72–75.

[2] GRUBER, T.: *A Translation Approach to Portable Ontology Specification,* Knowledge System Laboratory – Stanford University, CA, Technical Report KSL 92–71, 1992, pp. 192.

[3] PFOSER, D., TRYFONA, N.: *The Use of Ontologies in Location-based Services: The Space and Time Ontology in Protege,* Research Academic Computer Technology Institute, Research Unit 3, Athens, Greece, 2002, pp. 15.

[4] WENG, S., CHANG, H.: Using Ontology Network Analysis for Research Document Recommendation, Science Direct, *Expert System with Applications 34,* 2008, pp. 1857–1869.

[5] PEREZ, A., LOPEZ, M., CORCHO, O.: *Ontological Engineering,* Springer, London, ISBN 1-85-233-551-3, 2004, pp. 403.

[6] LOPEZ, M., PEREZ, A., JURISTO, N.: *Methodology: From Ontological Art Towards Ontological Engineering,* Spring Symposium on Ontological Engineering of AAAI, Stanford University, California, 1997, pp. 33–40.

[7] PRIBYL, P. et al.: *Terminological Dictionary of Traffic Telematics,* SDT, Praha, 2006, ISBN 80-239-7780-6, 2006, pp. 480 (Czech-English).

Peter Vestenicky – Martin Vestenicky – Juraj Palecek *

CALCULATION AND MEASUREMENT OF RFID TAG CRITICAL FREQUENCY

This paper deals with undesirable phenomenon in inductively coupled RFID system, i. e. by fading out the amplitude modulation of carrier signal in RFID reader. It is caused by inaccurate tuning of RFID transponder resonant circuit. The phenomenon is described by mathematical model of the inductively coupled RFID system. The numerical and graphical results of mathematical modelling is then compared by measurement of resonant frequency of real (tuneable) RFID circuit on which the amplitude modulation fades out.

Keywords: *RFID, tag, transponder, quality factor, critical frequency, mathematical model.*

1. Introduction

Inductively coupled RFID (Radio Frequency Identification) systems [1] are being widely used for marking of goods and animals, in access control system, bus tickets and data acquisition systems etc. In these applications the reading range of RFID transponder is not extremely long and the quality factor Q of the transponder resonant circuit is relatively small. Mostly $Q = 15$ [2] for identification cards in accordance with the standard ISO 14443. The requirements for accurate tuning of the transponder resonant frequency are not generally critical.

If the required reading range of the RFID transponder must be essentially longer, for example, if underground engineering networks are marked by the RFID transponders [3], the reading range can be increased by increasing of transponder coil area together with increasing of transponder resonant circuit quality factor [4]. Then inaccurate tuning of the transponder resonant frequency can make the data transfer from the transponder to the RFID reader impossible because the amplitude modulation of carrier signal fades out on the critical resonant frequency of the transponder.

2. Mathematical model of inductively coupled RFID system

The mathematical model was created to calculate system parameters in design process of RFID system, to analyse the influence of inaccurate tuning of reader and transponder resonant frequencies and to estimate maximum reachable reading range. The model goes out from the general schematic diagrams depicted in Figs. 1 and 2. Simplification of the model is given by the neglecting of RFID transponder chip nonlinearity. These chips usually include a voltage limiter in the form of two anti-serially connected Zener diodes in parallel to the transponder resonant LC circuit which limit the voltage across the chip to values about 14 – 15 V peak – to – peak. Therefore, this mathematical model is suitable to calculate RFID features when the limiter is not in operation, i. e. when the distance between RFID reader and transponder is relatively long.

2.1 Dynamic model

Let the transponder have two-state (0 and 1) modulator and the modulation frequency is 1/64 of carrier frequency, i. e. $f_{Mod} = f/64$. We denote $R(t) = R_P$ if the modulator is in state 0 and $R(t) = R_M$ if the modulator is in state 1. Let the modulator be controlled by the $Data(t)$ function according to equation (1). The resistance of RFID chip is then given by function (2). Let the excitation signal of the RFID reader be given by function (3).

$$Data(t) = \frac{Sign\left(\sin\left(2\pi \frac{f}{64} i\right)\right) + 1}{2} \tag{1}$$

$$R(t) = Data(t)R_P + (1 - Data(t))R_M \tag{2}$$

$$u(t) = U \sin(2\pi f t) \tag{3}$$

The model in Fig. 1 can be described by a system of 2nd order linear differential equations (4) which has not constant coefficients [5].

---

* **Peter Vestenicky[1], Martin Vestenicky[2], Juraj Palecek[3]**
[1] Department of Control and Information Systems, Faculty of Electrical Engineering, University of Zilina, Slovakia,
E-mail: peter.vestenicky@fel.uniza.sk
[2] Department of Telecommunications and Multimedia, Faculty of Electrical Engineering, University of Zilina, Slovakia
[3] Amicus SK s.r.o, Skalica, Slovakia

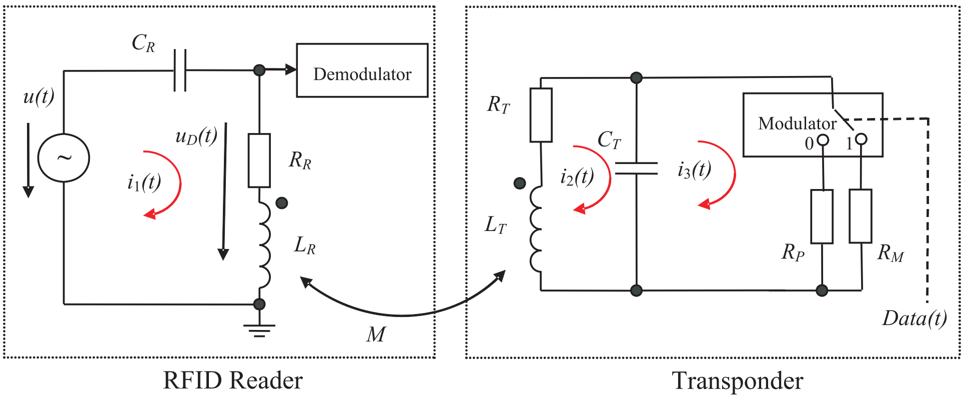


Fig. 1 Dynamic model of RFID system

$$L_R \frac{d^2 i_1(t)}{dt^2} + R_R \frac{di_1(t)}{dt} + \frac{1}{C_R} i_1(t) - M\frac{d^2 i_2(t)}{dt^2} = \frac{du(t)}{dt}$$
$$-M\frac{d^2 i_1(t)}{dt^2} + L_T \frac{d^2 i_2(t)}{dt^2} + R_T \frac{di_2(t)}{dt} + \qquad (4)$$
$$+\frac{1}{C_T} i_2(t) - \frac{1}{C_T} i_3(t) = 0$$
$$\frac{1}{C_T} i_2(t) = R(t)\frac{di_3(t)}{dt} - \frac{1}{C_T} i_3(t) = 0$$

2.2 Static model

This model considers only steady states of the RFID system modulation process. This is symbolically represented by selection of resistance value R (see Fig. 2) from two elements $\{R_P, R_M\}$ i. e. $R = R_P$ when modulation switch is in the logical zero state and $R = R_M$ when modulation switch is in the logical one state. The resistance R_P represents energy consumption of the RFID chip and the resistance R_M represents the additional damping of resonant circuit when the actual data bit is in logical one state.

Going out from Fig. 2 we can create the next system of equations

$$\left(R_P + j\omega L_R + \frac{1}{j\omega C_R}\right) I_1 - j\omega M I_2 = U$$
$$-j\omega M I_1 + \left(R_T + j\omega L_T + \frac{R}{1 + j\omega R C_T}\right) I_2 = 0 \qquad (5)$$

which can be solved by application of the Cramer's rule. Then the loop currents I_1, I_2 are given by (6) and (7).

$$I_1 = \frac{U\left(R_T + j\omega L_T + \frac{R}{1 + j\omega R C_T}\right)}{\left(R_R + j\omega L_R + \frac{1}{j\omega C_R}\right)\left(R_T + j\omega L_T + \frac{R}{1 + j\omega R C_T}\right) + \omega^2 M^2} \qquad (6)$$

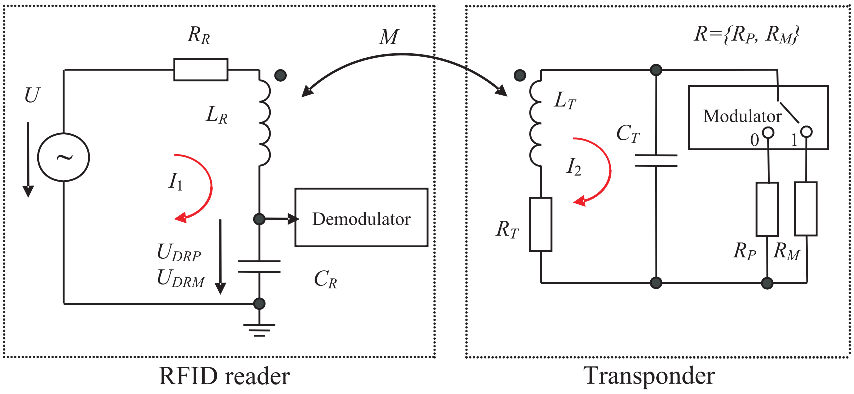


Fig. 2 Static model of RFID system

$$I_2 = \frac{j\omega MU}{\left(R_R + j\omega L_R + \dfrac{1}{j\omega C_R}\right)\left(R_T + j\omega L_T + \dfrac{R}{1 + j\omega RC_T}\right) + \omega^2 M^2} \tag{7}$$

$$U_{DRP} = \frac{U\left(R_T + j\omega L_T + \dfrac{R_P}{1 + j\omega R_P C_T}\right)}{j\omega C_R\left(\left(R_R + j\omega L_R + \dfrac{1}{j\omega C_R}\right)\left(R_T + j\omega L_T + \dfrac{R_P}{1 + j\omega R_P C_T}\right) + \omega^2 k^2 L_R L_T\right)} \tag{8}$$

$$U_{DRM} = \frac{U\left(R_T + j\omega L_T + \dfrac{R_M}{1 + j\omega R_M C_T}\right)}{j\omega C_R\left(\left(R_R + j\omega L_R + \dfrac{1}{j\omega C_R}\right)\left(R_T + j\omega L_T + \dfrac{R_M}{1 + j\omega R_M C_T}\right) + \omega^2 k^2 L_R L_T\right)} \tag{9}$$

The voltage at the demodulator input in Fig. 2 is given by (8) for modulator switched to logical zero and by (9) for modulator switched to logical one. Note that M is the mutual inductance of reader and transponder coils L_R and L_T, $M = k\sqrt{L_R L_T}$, k is coupling factor and $\omega = 2\pi f$ is angular frequency.

Going out from equations (8) and (9) we can calculate the modulation depth of amplitude modulated signal at the input of RFID reader demodulator:

$$m_{AM} = \frac{|U_{DRP}| - |U_{DRM}|}{|U_{DRP}| + |U_{DRM}|} \tag{10}$$

3. Measurement of the RFID chip parameters

To determine the resistances R_P and R_M a measurement of RFID transponder chip EM4100 (manufactured by EM Microelectronic – Marin SA, Switzerland) was performed according to Fig. 3.

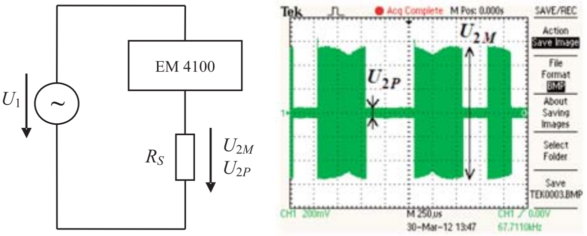


Fig. 3 Measurement of EM 4100 parameters

The equivalent resistances R_P and R_M are calculated from the measured voltages U_{2M} and U_{2P} (see Table 1) for $R_S = 110\ \Omega$ according to equations (11) and (12).

$$R_M = R_S\left(\frac{U_1}{U_{2M}} - 1\right) \tag{11}$$

$$R_P = R_S\left(\frac{U_1}{U_{2P}} - 1\right) \tag{12}$$

For simplicity, the RFID chip EM4100 will be approximately modelled by the resistances $R_P = 16$ kΩ and $R_M = 1.6$ kΩ (Figs. 1 and 2).

Measurement of the equivalent EM 4100 resistances Table 1

U_1 [Vpp]	U_{2P} [Vpp]	U_{2M} [Vpp]	R_P [kΩ]	R_M [kΩ]
5.0	0.032	0.045	17.08	12.11
7.6	0.054	0.250	15.37	3.23
10.0	0.068	0.640	16.07	1.61
12.4	0.080	1.050	16.94	1.19
14.0	0.094	1.440	16.27	0.96
15.0	0.328	1.680	4.92	0.87

4. Graphical results

The calculations according to equations (4) and (10) were performed for these parameters of RFID system:

- Voltage of signal generator $U = 10$ V, frequency $f = 125$ kHz
- $C_R = 1.621$ nF, $L_R = 1$ mH, $R_R = 15.7\ \Omega$, i. e. reader resonant frequency is 125 kHz, quality factor of $L_R C_R$ tuned circuit is $Q_R = 50$
- $C_T = 1.621$ nF (or variable if the resonant frequency of transponder f_T is variable), $L_T = 1$ mH, $R_T = 7.85\ \Omega$
- $R_P = 16$ kΩ, $R_M = 1.6$ kΩ (chapter 3)
- $k = 0.01$ or variable

4.1 Dynamic model

By numerical solving the system of differential equations (4) for parameters given above the time responses of signals at the input of reader demodulator were obtained. These time responses are shown in Fig. 4 for nominal resonant frequency of transponder (f_T = 125 kHz) and in Fig. 5 for detuned transponder (f_T = 136 kHz). In case of detuned transponder the amplitude modulation of signal fades out and the transponder becomes unreadable.

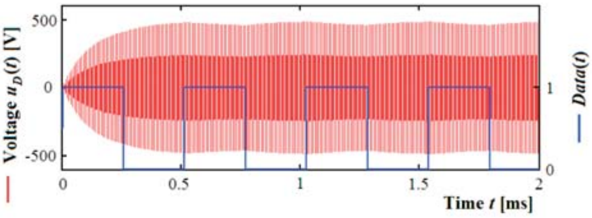


Fig. 4 Voltage at demodulator input as a function of time, transponder resonant frequency f_T = 125 kHz

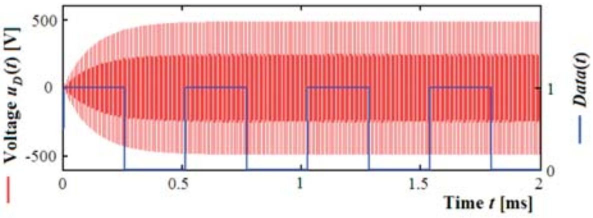


Fig. 5 Voltage at demodulator input as a function of time, transponder resonant frequency f_T = 136 kHz

4.2 Static model

The fade of amplitude modulation in RFID reader under condition of transponder detuning to the critical resonant frequency is evident from the three dimensional graph of function (10) which is shown in Fig. 6.

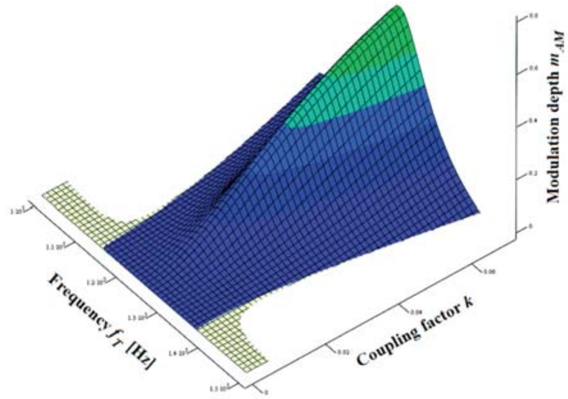


Fig. 6 Modulation depth m_{AM} as a function of coupling factor k and transponder resonant frequency f_T

By substituting (16) into (8) and (9) we can calculate the dependence of AM modulation depth m_{AM} (10) on distance x between the RFID transponder and the RFID reader and on the transponder resonant frequency f_T. The 3D graph is shown in Fig. 7.

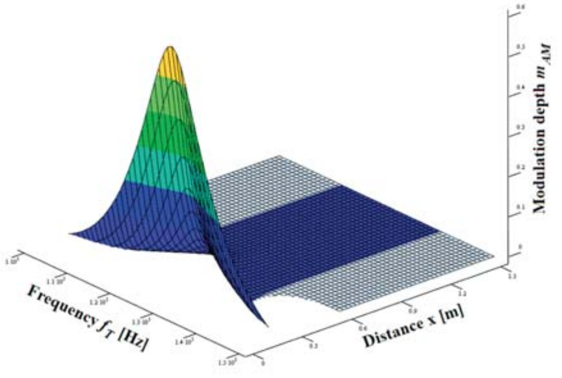


Fig. 7 Modulation depth m_{AM} as a function of distance x and transponder resonant frequency f_T

Similar, if we substitute (13) into (8), (9) and then (10) we can calculate the dependence of modulation depth on the resonant frequency f_R of the reader LC circuit and coupling factor k or distance x respectively, see Fig. 8 and Fig. 9.

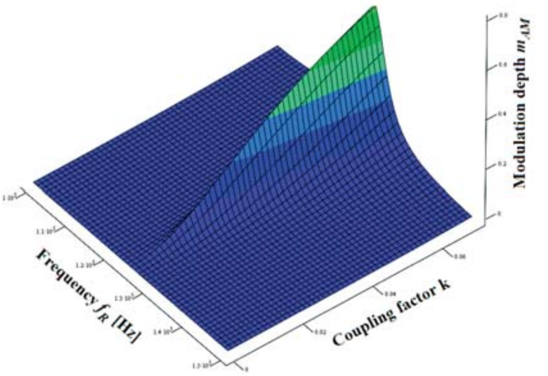


Fig. 8 Modulation depth m_{AM} as a function of coupling factor k and resonant frequency f_R of reader antenna

From the series of 3D graphs it is evident that detuning of transponder causes its unreadability due to fall of amplitude modulation, i. e. certain critical frequency exists which is dependent on distance (or coupling factor) and on other parameters of RFID system especially on quality factors of tuned LC circuits. The detuning of the reader antenna LC circuit is not so critical because the surfaces of 3D graphs in Figs. 8 and 9 do not cross zero plane of modulation depth.

The critical frequency f_C can be calculated from equation (14) substituting (15) into (8) and (9). The equation (14) has two solutions f_{C_HIGH}, f_{C_LOW} whose dependency on coupling factor k is

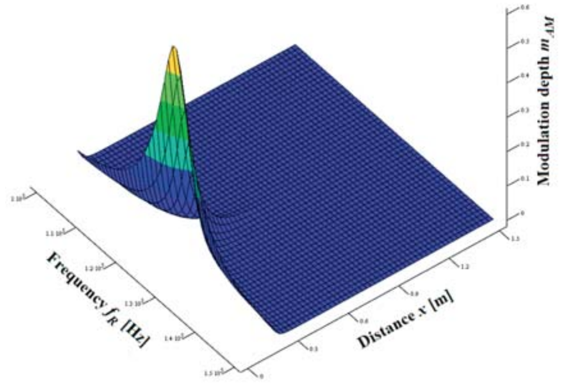


Fig. 9 Modulation depth m_{AM} as a function of distance x and resonant frequency f_R of reader antenna

shown in Fig. 10. After substituting (16) into (8) and (9) we can obtain dependency of critical frequencies on distance x between reader and transponder coils. It is shown in Fig. 11.

The equation (16) describes coupling factor of two circular coils as a function of their radiuses r_R, r_T, angle θ, and distance x [6]. The calculations of graphs in Figs. 7, 9 and 11 were performed for $r_R = 0.1$ m, $r_R = 0.1$ m, and $\theta = 0$ °, i. e. both coils are parallel.

$$C_R = \frac{1}{(2\pi f_R)^2 L_R} \tag{13}$$

$$|U_{DRM}| - |U_{DRP}| = 0 \tag{14}$$

$$C_T = \frac{1}{(2\pi f_C)^2 L_T} \tag{15}$$

$$k = \frac{r_R^2 \cdot r_T^2 \cdot \cos\theta}{\sqrt{r_R \cdot r_T} \cdot (r_R^2 + x^2)^{\frac{3}{2}}} \tag{16}$$

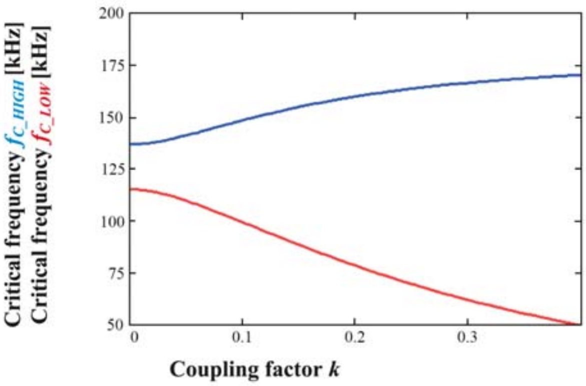


Fig. 10 Transponder critical frequencies as a function of coupling factor k

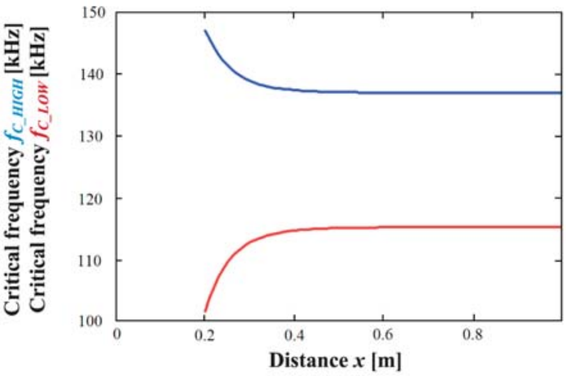


Fig. 11 Transponder critical frequencies as a function of distance x between coils L_R and L_T

5. Comparison of calculated and measured critical frequencies

To compare the results of mathematical model described in chapters 2 and 4 with real properties of inductively coupled RFID system the measurement of critical frequencies and their dependency on distance between transponder and reader was performed. The measured circuits were arranged according to Fig. 1 where transponder capacitor C_T was replaced by parallel connection of fixed and variable capacitor to enable the tuning of transponder resonant frequency. The critical frequencies were found when the amplitude modulation became extinct. The modulator and its resistances R_P and R_M were replaced by RFID chip EM 4100 (chapter 3).

Measured and calculated critical frequencies Table 2

Distance x [m]	Measured critical frequency		Calculated critical frequency	
	f_{C_LOW} [kHz]	f_{C_HIGH} [kHz]	fC_LOW [kHz]	fC_HIGH [kHz]
0.20	102.12	199.22	94.19	139.15
0.25	103.84	154.39	102.47	133.35
0.30	107.16	138.63	106.01	130.61
0.35	109.51	130.39	107.50	129.41
0.40	111.76	126.08	108.15	128.88
0.45	113.25	123.53	108.44	128.63
0.50	113.90	121.96	108.59	128.51

The results of calculations according to equation (14) and measurements are listed in Table 2 and displayed in Fig. 12. The nominal working frequency used in experiment and calculations has nonstandard value 117.7 kHz which is given by used inductor L_R and capacitor C_R. Other parameters were set as follows:

- Voltage of signal generator $U = 5$ V, frequency $f = 117.7$ kHz
- $C_R = 2.4$ nF, $L_R = 760$ μH, i. e. reader resonant frequency is $f_R = 117.7$ kHz, measured quality factor of $L_R C_R$ tuned circuit is $Q_R = 62$, then equivalent resistance $R_R = 9.08\ \Omega$
- C_T is variable i. e. transponder resonant frequency f_T is variable), $L_T = 1.04$ mH, $R_T = 8.37\ \Omega$
- $R_P = 16$ kΩ, $R_M = 1.6$ kΩ (chapter 3)

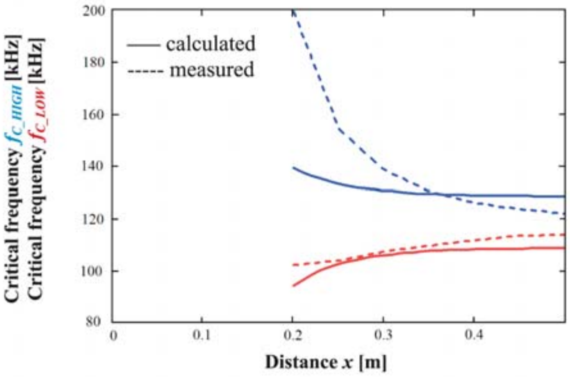


Fig. 12 Comparison of calculated and measured transponder critical frequencies

6. Conclusion

The method of RFID transponder critical frequency calculation can be useful at the design of transponder LC circuit. Especially, it can be useful to determine the acceptable manufacturing tolerances of transponder resonant frequency. Moreover, the required bandwidth of modulated signal from transponder must be considered. Note that the required bandwidth is about 4 kHz for the used RFID chip EM 4100 (with Manchester coding and with data transfer rate 1/64 of carrier frequency 125 kHz).

The used mathematical model is simplified by neglecting the nonlinearity of transponder (RFID chip) modulation circuits, which is evident from Table 1. Therefore, the results of critical frequency calculations in Figs. 10 and 11 must be considered as informative and their validity must be limited to the lower values of the coupling factor or higher values of the distance between the transponder and reader. This limitation of linear mathematical model is evident by comparison of calculated and measured results in Fig. 12. Moreover, high difference between measured and calculated critical frequencies f_{C_HIGH} in Fig. 12 (upper curves) is probably caused by external overvoltage suppressors connected to the used RFID chip, because capacity of the suppressor is not constant and is dependent on the applied voltage. At higher resonant frequencies the capacity of overvoltage suppressors becomes dominant compared to external capacitor C_T. This limitation will be subject to further improvement of the mathematical model.

Acknowledgment

This publication is the result of the project implementation: "Centre of excellence for systems and services of intelligent transport II", ITMS 26220120050 supported by the Research & Development Operational Programme funded by the ERDF. "Podporujeme vyskumne aktivity na Slovensku / Projekt je spolufinancovany zo zdrojov EU"

Európska únia
Európsky fond regionálneho rozvoja
Operačný program VÝSKUM a VÝVOJ
Agentúra
Ministerstva školstva, vedy, výskumu a športu SR
pre štrukturálne fondy EÚ

References

[1] FINKENZELLER, K.: RFID *Handbook. Fundamentals and Applications in Contactless Smart Cards and Identification.* 2nd edition. John Wiley and Sons, Chichester, 2003. ISBN 978-0-470-84402

[2] ROMERO, H. P., REMLEY, K. A., WILLIAMS, D. F., CHIH-MING, W., BROWN, T. X.: Identifying RF Identification Cards From Measurements of Resonance and Carrier Harmonics. *IEEE Transactions on Microwave Theory and Techniques*, vol. 58, No. 7, July 2010, pp. 1758-1765. ISSN 0018-9480

[3] VESTENICKY, P.: Solution of Some Technical Problems in Marker and Marker Locator Development. *Communications - Scientific Letters of the University of Zilina,* vol. 6, No. 4, 2004, pp. 103-106. ISSN 1335-4205

[4] VESTENICKY, P.: Increasing of RFID System Read Range. *Advances in Electrical and Electronic Engineering,* vol. 5, No. 1-2, 2006, pp. 171-173. ISSN 1336-1376

[5] RUZICKOVA, M.: Discrete and Differential Equations in Applied Mathematics. *Communications - Scientific Letters of the University of Zilina,* vol. 10, No. 2, 2008, pp. 72-78. ISSN 1335-4205

[6] GEHRING, U., ROZ, T.: *RFID Made Easy.* EM Microelectronic-Marin SA, Marin, 2000.

Ladislav Janousek *

INFLUENCE OF SELECTED PARAMETERS ON EDDY CURRENTS ATTENUATION IN NON-DESTRUCTIVE INSPECTION

Eddy currents attenuation in conductive materials under non-destructive inspection is analyzed in the paper. Influence of selected parameters of the inspection on the attenuation is studied by numerical means. A circular driving coil is employed to induce eddy currents in conductive plates under harmonic excitation. Outer diameter of the coil is varied together with the clearance between the coil and a plate surface for two plates with different thickness. Attenuation of the eddy current density in plates is studied and the results are analyzed in absolute values and the normalized ones as well. It is demonstrated that the investigated parameters of the inspection have significant impact on the attenuation.

Keywords: *Non-destructive evaluation, eddy currents, attenuation, penetration depth.*

1. Introduction

Eddy current testing (ECT) is one of the most common electromagnetic methods utilized in non-destructive evaluation of conductive materials. The principle of ECT underlies in the interaction of induced eddy currents with a structure of an examined body [1], [2]. A primary alternating exciting electromagnetic field is generated in the vicinity of a coil driven by a time-varying current according to the Ampere's law. Electromotive force is induced in a conductive object which is in proximity of the coil according to the Faraday's law. Eddy-currents flow in the conductive object according to the Ohm's law and their vector lines must be closed. A secondary electromagnetic field generated by the eddy-currents counterworks to the primary exciting electromagnetic field according to the Lenz's theorem. The induction coupling therefore exists between the coil and the conductive object. It can be simply considered as an interaction between the primary and the secondary electromagnetic fields. ECT can thus be employed for detection of surface and subsurface non-homogeneities, measurements of material thickness as well as of the electromagnetic parameters and so on [1]. However, the most wide spread area of its application in present is the detection of discontinuities [2].

ECT probes are one of the most important elements in the non-destructive testing, because they transfer information between an ECT instrument and a conductive object through the induction coupling. Optimal ECT probe should assure high sensitivity to expected defects, high probability of detection of expected defects and classification possibility of expected defects (location, dimensions, etc.) [3]. Exciting coil(s) of ECT probe should thus induce eddy currents with high density and such distribution that eddy current lines are significantly perturbed when a defect is in presence. The detection circuit should assure that maximum of the perturbation field is sensed. Many ECT probes have been developed over past decades reflecting special demands of particular applications. Probe design and development is still of high interest because the area of ECT utilization is gradually wide-spreading [4]–[7].

Eddy current distribution in an inspected material significantly determines basic features of the probe such as sensitivity, penetration and resolution. To be able to detect expected flaws with high probability in a material under given test conditions it is very indispensable to have a knowledge about the eddy current distribution. Many probes have been developed and designed mainly based on practical experiences. Maturity of the numerical means make it possible to build up systematic approach in this sense.

According to the electromagnetic field theory, the distribution of eddy currents along material depth depends mainly on a testing frequency and the electromagnetic parameters of a material as it is given by the well known equation for the standard depth of penetration:

$$\delta = \frac{1}{\sqrt{\pi f \mu \sigma}}, \qquad (1)$$

where δ [m] is the standard depth of penetration, f [Hz] is the inspection frequency, μ [H·m^{-1}] and σ [S·m^{-1}] are the magnetic permeability and the electric conductivity of the inspected material, respectively. However, under real conditions there are much more parameters with substantial influence on this distribution such as shape of coils, their dimensions, configuration of inspection, material thickness, lift-off, etc [8].

---

* **Ladislav Janousek**
Department of Electromagnetic and Biomedical Engineering, Faculty of Electrical Engineering, University of Zilina, Zilina, Slovak Republic,
E-mail: ladislav.janousek@fel.uniza.sk

The paper numerically analyzes influences of selected parameters of ECT inspection on attenuation of eddy currents in a conductive material.

2. Simulation Model

Numerical simulations using the finite element method are carried out to investigate influences of certain inspection parameters on the eddy currents attenuation.

A plate conductive specimen, shown in Fig. 1, is modeled in this study to explore attenuation of eddy currents along its depth. The plate thickness is adjusted to two values $t = 10$ and 30 mm to explore its influence on the attenuation. The electromagnetic parameters of the plate are adjusted to the following values $\sigma = 1.35 \cdot 10^6$ S·m^{-1} $\mu_r = 1$ that correspond to a material SUS316L frequently used for structural components in petrochemical and nuclear industries.

Fig. 1 Layout of model

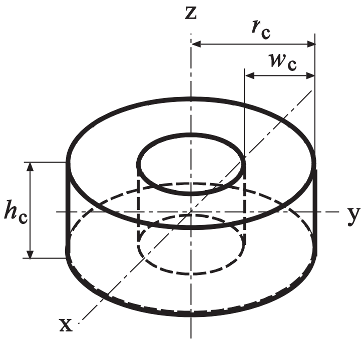


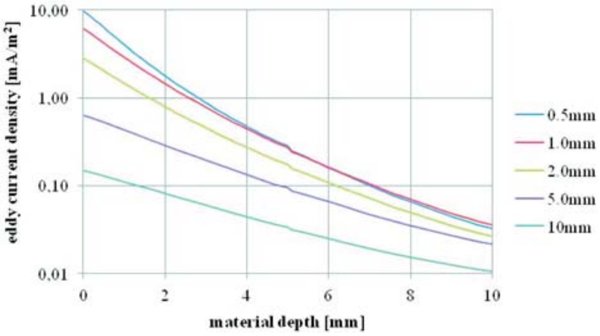
Fig. 2 Layout of circular exciting coil

A circular exciting coil, displayed in Fig. 2, is employed to drive the eddy currents in the plates. Axis of the coil is oriented normally regarding the plate s' surface. This type of exciting coil in the given orientation is commonly utilized in variety of applications because an ECT probe with such exciting coil does not account for the directional properties. Cross-section of the coil winding is kept constant for all the calculations, in concrete $w_c = 1$ mm, $h_c = 1$ mm (see Fig. 2), because these dimensions have only minor impact on the eddy current attenuation along the material depth [9]. The coil radius rc is varied to investigate its influence on the attenuation. It is gradually adjusted to the following values $r_c = 2, 5, 10, 15, 20$ mm. Clearance between the coil and the plate surface, hereinafter referred to as the lift-off, is sequentially tuned to $l_f = 0.5, 1, 2, 5, 10$ mm. The coil is driven with the harmonic current while the current density is kept constant at a value of 1 A·mm^{-2} and the frequency is adjusted to 10 kHz. The standard depth of penetration for the given parameters equals to $\delta = 4.33$ mm. Another three values of the exciting frequency 1, 5 and 20 kHz are also used for the investigations for the plate with a thickness of 10 mm.

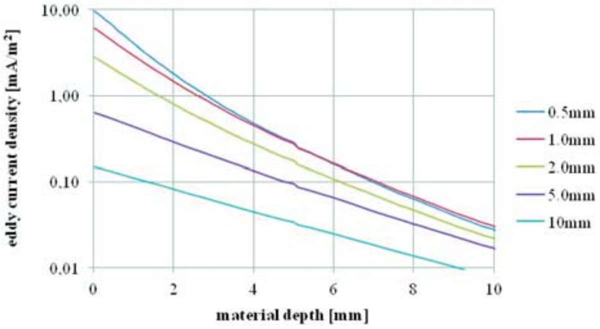
Very fine model of the system is built up based on the finite tetrahedral elements. Only ¼ of the system, as shown in Fig. 1, is considered employing appropriate boundary conditions. The mesh consists of approximately 4 million elements. Results of numerical simulations are summarized in the following section.

3. Results and discussions

Three parameters of the inspection, i.e. the coil radius r_c, the lift-off l_f and the plate thickness t are varied to investigate influences of these parameters on the eddy current attenuation along the material depth. Important findings are reported and discussed here.



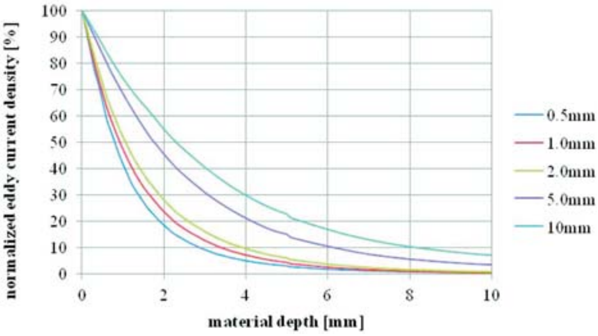
a) plate thickness $t = 10$ mm



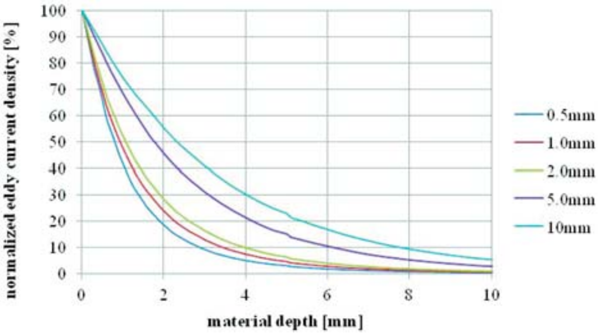
b) plate thickness $t = 30$ mm

Fig. 3 Dependences of eddy current density in absolute values on the material depth for different adjustments of lift-off, $r_c = 2$ mm

The distribution of eddy current density vector in the material is stored after successful execution of each simulation. The results are then processed in such a way that only one dependence of the eddy current density absolute value along the material depth is taken for the evaluation. The dependence is taken along the plate thickness direction (see Fig. 1) under a certain surface position where the eddy current density has the maximum value.



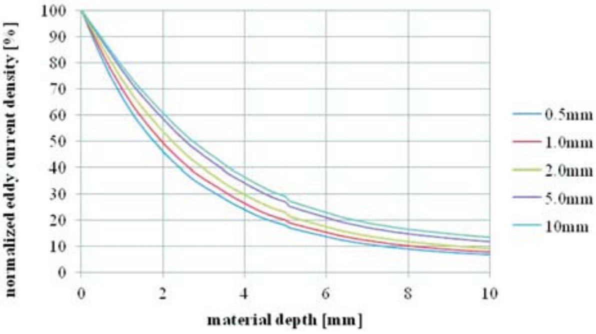
a) plate thickness t = 10 mm



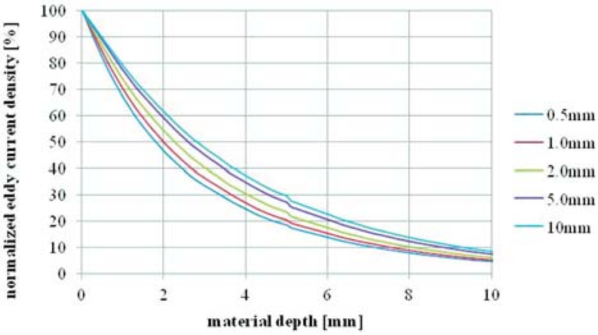
b) plate thickness t = 30 mm

Fig. 4 Dependences of eddy current density in normalized values on the material depth for different adjustments of lift-off, $r_c = 2$ mm

The dependences of the eddy current density absolute value on the material depth for the coil with a radius of $r_c = 2$ mm and for several values of the lift-off are shown in Fig. 3. The exciting frequency is adjusted to 10 kHz. The dependences are shown for the two considered values of the plate thickness. Note that the vertical axis is in logarithmic scale. The upper limit of x-axes for both the plate thicknesses are set to 10 mm so one can simply compare the results. Attenuation of eddy currents along the material depth has the exponential character. However, it can be observed that the results gained for different values of the lift-off show diverse behaviour. Moreover, it can be seen that the plate thickness almost does not have impact on the attenuation for the given standard depth of penetration. The same dependences but in normalized values are shown in Fig. 4 in order to highlight contrast between the curves. Each of dependences of the eddy current density absolute value on the material depth is normalized by its maximum value. It can be seen that the eddy current attenuation strongly depends not only on the standard depth of penetration, i.e. the frequency, the conductivity and the permeability, but also on the lift-off, however it does not depend on the material thickness. Larger clearance between the coil and the plate surface provides eddy current distribution with less attenuation along the material depth.



a) plate thickness t = 10 mm



b) plate thickness t = 30 mm

Fig. 5 Dependences of eddy current density in normalized values on the material depth for different adjustments of lift-off, $r_c = 20$ mm

Similar dependences in normalized values for the coil with a radius of $r_c = 20$ mm are shown in Fig. 5. The presented results clearly show that the impact of the lift-off on the eddy current attenuation along the material depth decreases for the coil with larger diameters. It can be also noticed that the eddy currents are slightly less attenuated in the 10 mm thick plate comparing to the one with a thickness of 30 mm; however, this impact is not significant.

A specific value of the distance from material surface along its depth, denoted hereinafter as the penetration depth, is calculated for all the evaluated distributions of the eddy current density. It is the value where the eddy current density falls to 50% of its surface value. The penetration depth is then plotted as a dependence on the coil radius rc and the lift-off lf as well. Fig. 6 shows the calculated dependence only for the plate with a thickness of *t* = 10 mm as the results presented above showed that the plate thickness does not have significant impact on the attenuation under the frequency of *f* = 10 kHz. As it can be seen, by proper adjustment of the exciting system parameters, the penetration depth can be increased several times comparing to a case of only a small coil placed with a small lift-off over a material.

The eddy current attenuation for the plate with a thickness of *t* = 10 mm is investigated also under frequencies *f* = 1.5 and 20 kHz. Table 1 reports values of the standard penetration depth δ for the given frequencies calculated according to (1) together with a ratio of the plate thickness (*t* = 10 mm in this case) and the standard depth of penetration *t*/*δ*.

The penetration depth, as defined above, as dependence on the coils radius rc and on the lift-off for another considered values of the frequency are shown in Figs. 7-9. It can be observed that the parameters have strong impact on the penetration depth especially when the value of material thickness is comparable to the standard depth of penetration. Importance of the parameters decreases with increasing the ratio of the material thickness and the standard depth of penetration.

Standard depth of penetration and its relation to material thickness t = 10 mm Table 1

f [kHz]	1	5	10	20
δ [mm]	13.70	6.13	4.33	3.06
t/*δ* [−]	0.73	1.63	2.31	3.27

The exciting coil radius and the lift-off have significant influence on the eddy current attenuation along a material depth when the material thickness is at maximum double comparing to the standard penetration depth. It should be noted that in such a case larger coils and/or larger lift-off provide deeper penetration of eddy currents inside a material and thus it would provide better resolution. However, sensitivity can be decreased. Sophisticated design of the exciting system requires therefore appropriate compromise to reach essential performance of an ECT probe.

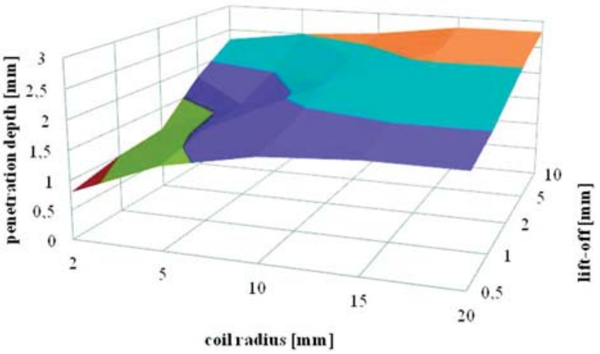


Fig. 6 Penetration depth as function of coil radius and lift-off, t = 10 mm, f = 10 kHz

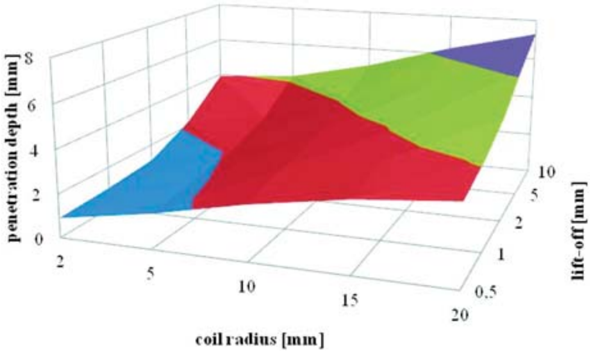


Fig. 7 Penetration depth as function of coil radius and lift-off, t = 10 mm, f = 1 kHz

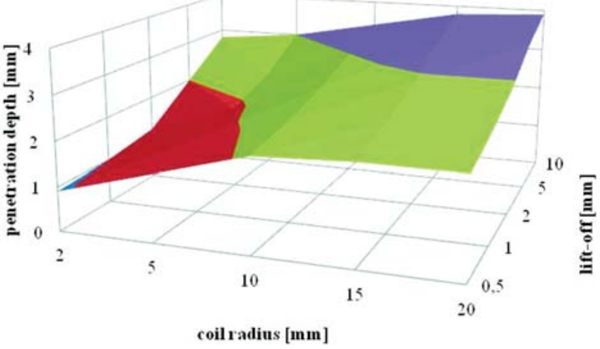


Fig. 8 Penetration depth as function of coil radius and lift-off, t = 10 mm, f = 5 kHz

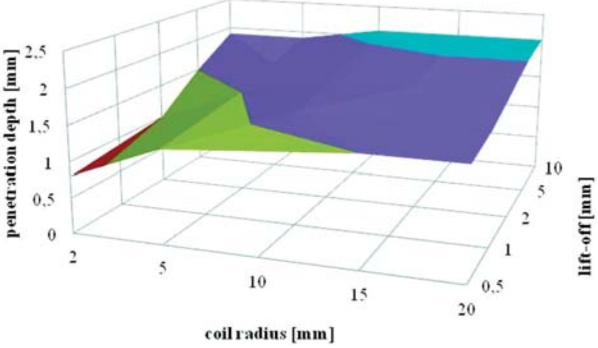


Fig. 9 Penetration depth as function of coil radius and lift-off, t = 10 mm, f = 20 kHz

4. Conclusions

The paper concerned on attenuation of eddy currents along material depth in non-destructive inspection. Influences of selected parameters of eddy current exciting system and of material thickness on the attenuation were studied by numerical means. A conductive plate specimen having the electromagnetic parameters of SUS316L was used for the study. A circular exciting coil positioned normally regarding the plate surface drove eddy currents. Three parameters of the exciting system, specifically the coil radius, the lift-off and the inspection frequency were altered together with the plate thickness. Distribution of eddy current density was evaluated for various adjustments of the parameters. The results clearly demonstrated that the coil radius and the lift-off as well have strong impact on the eddy current attenuation along material depth when a material thickness is comparable to the standard penetration depth.

Acknowledgment

This work was supported by the Slovak Research and Development Agency under the contracts No. APVV-0349-10 and APVV-0194-07. This work was also supported by grant of the Slovak Grant Agency VEGA, project No. 1/0765/11.

References

[1] RAO, B.P.C.: P*ractical eddy current testing.* Alpha Science International Ltd.: Oxford, 2007.

[2] JANOUSEK, L., T. MAREK, T., D. GOMBARSKA, D.: Eddy Current Non-destructive Evaluation of Conductive Materials. *Communications - Scientific Letters of the University of Zilina,* 1, pp. 29-33, 2006.

[3] CHEN, Z., MIYA, K.: A New Approach for Optimal Design of Eddy Current Testing Probes. *J. of Nondestructive Evaluation,* vol. 17, pp. 105-116, 1998.

[4] STRAPACOVA, T., CAPOVA, K., JANOUSEK, L.: Artificial Heart Valve Inspection Using Eddy Current Techniques. *Communications - Scientific Letters of the University of Zilina,* vol. 13, No. 2A, pp. 83-86, 2011.

[5] JANOUSEK, L., SMETANA, M.: Uncertainty in Depth Evaluation of Partially Conductive Cracks from Eddy Current Testing Signals. *Communications - Scientific Letters of the University of Zilina,* vol. 13, No. 2A, pp. 55-60, 2011.

[6] SMETANA, M., STRAPACOVA, T.: Pulsed Eddy Currents as a New Approach in Non-destructive Evaluation of Conductive Materials. *Proc. of TRANSCOM 2009,* Zilina, pp. 63-66, 2009.

[7] GOMBARSKA, D., SMETANA, M.: Wavelet Based Signal Analysis of Pulsed Eddy Current Signals. *Electrical review,* vol. 87, pp. 37-39, 2011.

[8] MOTTL, Z.: The Quantitative Relations between True and Standard Depth of Penetration for Air-cored Probe Coils in Eddy Current Testing. *NDT International,* vol. 23, pp. 11-18, 1990.

[9] ALMAN, M.: Excitation Coil Parameters in View of Eddy Current Distribution in ECT Applications. *Proc. of Elektro 2010,* Zilina, pp. 221-224, 2010.

Matej Pacha – Jiri Stepanek *

PERFORMANCE AND FUEL CONSUMPTION OPTIMIZATIONS OF SHUNTING HYBRID LOCOMOTIVES

In recent years hybrid electric drive has become very popular in public transportation and automobiles. This article deals with hybrid drive technology and its impact on performance and energy consumption optimization of shunting locomotives in comparison with conventional diesel-electric locomotives. Based on classic longitudinal train dynamics theory of train movement and with support of simulation results, a comparison of electrochemical accumulators and super-capacitors impact is also introduced.

Keywords: *Hybrid drive, shunting locomotive, energy savings, accumulator, supercapacitor.*

1. Introduction

Hybrid electric drive (HEV) became a very popular alternative of conventional drives last years. We can see many hybrid vehicles for personal use but its popularity is not motivated by energy or fuel savings. The configuration of these vehicles is not optimal for general use in combined city and highway operation [1].

With new types of accumulators and with positive costs development of power converters the hybrid technology is being utilized in public transport vehicles as well. Unfortunately, in scope of railway transport there are only several light trains and locomotive prototypes equipped with the hybrid drive.

In this paper we would like to introduce positive impact of vehicle hybridization on fuel consumption in special operational conditions of yard-switching and shunting locomotives.

2. Energy analysis of train movement

Energy analysis of a train movement helps us to focus on reasonable train operation types for hybrid drive applications. The key parameter is kinetic energy, being developed with every acceleration and lost when braking [2] and [3].

Based on the analysis published in [4], the train movement model can be described by a longitudinal train dynamics theory based on 3-axes multipoint movement to a single point linear movement simplification with concentrated parameters (multipoint analysis is not necessary for energy computations). Tractive effort covers resistive forces (vehicle resistance F_v, elevation and curve resistance F_e and F_c, dynamic/inertia force F_a – usually expressed by their specific values p_v, p_e, p_c, p_a). From energetics point of view of the longitudinal analysis, we can recognize two main parts of the overall work A_r – a kinetic energy E_k and a tractive work A_t.

$$A_r = E_k + A \tag{1}$$

Kinetic energy E_k can be written as

$$E_k = \frac{1}{2} \cdot m \cdot \xi \cdot v_d^2 \tag{2}$$

where m is a mass of the train, ξ is an addition of rotational parts of the vehicle(s) and v_d is a speed difference (or final speed when accelerating from zero speed).

Work part of the energy is represented simply by

$$A_t = F_r \cdot l = (F_v + F_e + F_c) \cdot l \tag{3}$$

Where l is a distance passed by the train and F_r is a sum of resistive forces.

The same conditions are valid for the braking period respectively. The only difference to be mentioned is the sign of the tractive work has to be negative, because the kinetic energy is used to cover resistive forces.

$$A_b = E_k - A \tag{4}$$

To simplify the energy investigation, a train mission cycle can be divided into four periods investigated separately:

- Acceleration period
- Constant speed period
- Coasting (similar to the constant speed period)

---

* **Matej Pacha**[1], **Jiri Stepanek**[2]
[1] Faculty of Electrical Engineering, University of Zilina, Slovakia, E-mail: matej.pacha@fel.uniza.sk
[2] CZ LOKO, a.s., Ceska Trebova, Czech Republic

- Braking (deceleration) period

Overall energy analysis is then a simple summary of periods' results mentioned above.

Energy savings during a train movement are dependent on transport requirements such as speed, a number of stops, locomotive power transmission efficiency and type of the rolling stock, or a coasting to constant speed ratio [4]. All these parameters are based on real world conditions and components and they are discussed separately in many publications. Nevertheless, considering equations above we can focus on energy that can be saved without any impact on transport requirements.

In braking period the whole kinetic energy is partially consumed by resistive forces and the rest is usually irreversibly spent in braking system of conventional vehicles. Focusing on this period of train movement, we can define a relative braking work β_b [5] and [6]:

$$\beta_b = \frac{A_b}{A_t + A_b} \tag{5}$$

Using formulas from [4] and [5] we can obtain a parametric equation showing most of dependencies

$$\beta_b = \frac{\frac{1}{2}\cdot m\cdot\xi\cdot v_b^2 - m\cdot g\cdot\left(p_{v(AV)} + p_{e,c}\right)\cdot l_b}{m\cdot g\cdot\left(p_{v(\ )} + p_{e,c}\right)\cdot l_b + \frac{1}{2}\cdot m\cdot\xi\cdot v_d^2 + m\cdot g\cdot\left(p_{v(AV)} + p_{e,c}\right)\cdot l_d} \tag{6}$$

Where A_b is computed for the braking period only, the A_t represents a constant speed period and the A_r represents the accelerating period of the train movement. It can be seen that (6) is a rolling stock dependent and nonlinear equation. Fig. 1 shows (6) in a chart $\beta_b = f(l, v)$ valid for a mixed wagons (passenger and freight) train with average acceleration 0.6 m/s^2. It should be noticed that the relative braking energy is higher for short distance and for higher speed trains. Most of the relative braking energy can be saved in an accumulator (flywheel, ultracapacitor, etc). To get the real energy savings or energy recuperated to the accumulator the efficiency should be considered.

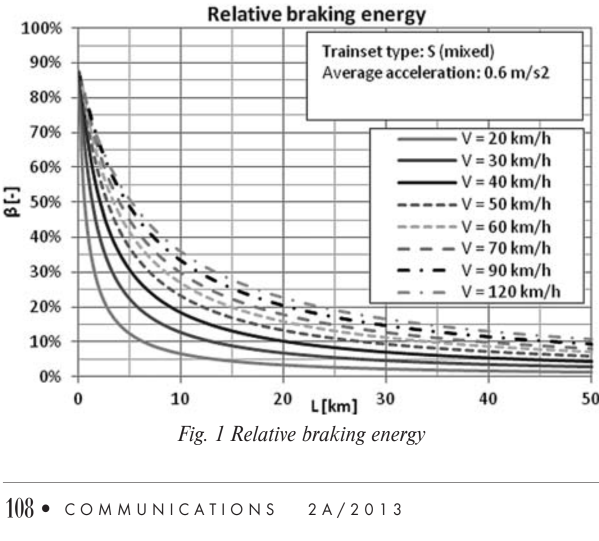


Fig. 1 Relative braking energy

Based on results from (6) and Fig. 1, a conclusion on braking energy can be made. Energy recuperation in general is more effective for high speed short distance trains like city shuttles or public transport. Another field of possible applications is yard switching and shunting where the speed is low but the average distance is approximately 1–2 km.

In public transport several hybrids have been introduced (city buses, trams). Their power mission is well known and scheduled, thus the system design is very easy. Hybrid driven yard switchers and station shunting locomotives are forgotten except some prototypes in the Czech Republic (CKD 718.501, 1984, [6]), USA (GE Ecomagination, 2008 [7]), Canada (RailPower Green Goat 2009 [8]) and France (ALSTOM V200, 2010 [9]). The focus on hybrid shunting locomotives is motivated by another benefit of hybridization – diesel engine operation optimization.

3. Power mission profile

To optimize any operation and energy savings, power mission profile is needed to be well known. Dealing with shunting locomotives, their mission is not a regular or a repeating pattern of any kind. Based on current needs the locomotive operates very randomly.

There is a common denominator in such missions operated by diesel locomotives: very low diesel power utilization. A long idling time stands side by side the short time full power operation intervals what causes higher specific fuel consumption and emissions [6]. The power demand is changing very fast and the system operates in steady state for a very short time, which only causes other losses.

A. Shunting modes

Heavy and light shunting operation can be recognized. Light shunting operation is a typical for stations and some industrial railways. The mission is to pick or drop wagons at sidings for loading and unloading. The acceleration and the speeds are usually higher due to operational reasons and allowed by the low weight of the trains.

Heavy shunting locomotives operate on yards and usually move a high load train sets between yards for or after the sorting, or sorting trains directly. Fig. 2 shows a typical heavy shunting operation in comparison to the light shunting operation. The speed uses to be lower (from 10 to 20 km/h) due to high load of the train.

To fully understand the mission a set of measurements and simulations were made and the results are introduced.

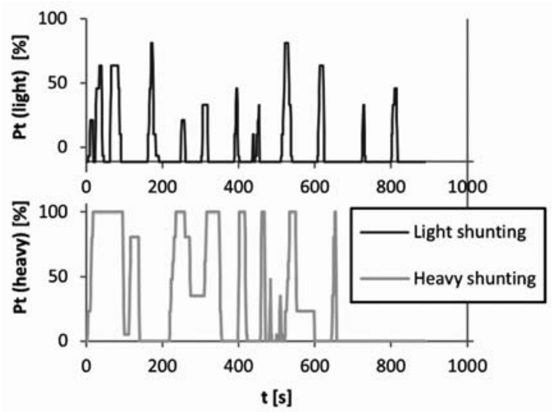


Fig. 2 Light and heavy shunting power mission comparison

B. On-board Measurements

On-board measurement is the most accurate way to get the real mission profile. The real world measuring well considers waiting periods, traffic interactions, weather conditions, coupling issues, etc. Therefore results are valuable for an analysis of power missions. A several days a week at least should be measured to capture irregularities.

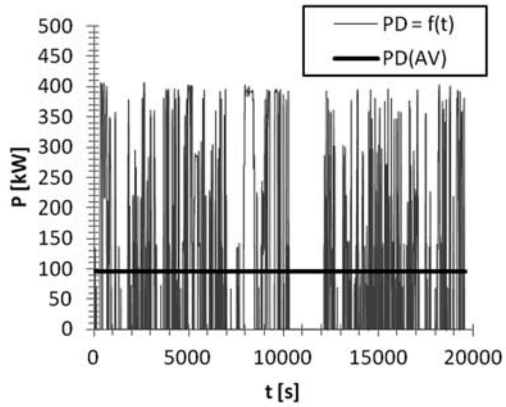
For this purpose, several measurements were made in 2008 at OKD Transportation (today AWT Corporation) industrial railway in Kladno, Czech Republic, operated by 740.3 class locomotives [4]. The main mission consists of several shunting tasks nearby local mounting halls and freight train cars distribution from the central station to several companies placed within the industrial railway in Kladno. It can be characterized as a light shunting operation.

Parameters and results are summarized in Table 1. Data collection was realized with support of on-board control unit NES RV08 via diagnostic serial link. An example of one day power mission is shown in Fig. 3a. It should be noticed that the mission power average is 23.5% of the diesel engine rated power. Typical values of shunting locomotive power average are less than 15% in most cases, depending on power installed and the mission as well.

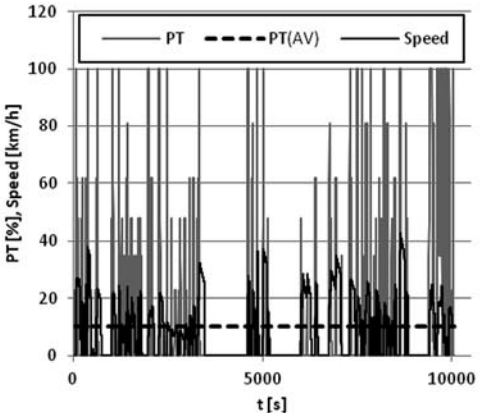
On-Board Measurements summary Table 1

Parameter	Value
Axle arrangement	Bo'Bo'
Diesel engine	CAT C15
Diesel power	403 kW
Power transition	AC-DC
Weight	64 t
Gearbox ratio	78:15
Max. tractive effort	230 kN
Max. velocity	60 km/h
Cont. tractive effort	114 kN
Cont. speed	8 km/h
Entire mission time (per day)	6 hours
Average speed	4 km/h
Average diesel power	93 kW
Load range	100 - 1100 tons
Operation	Shunting, Short freight run

There is another issue with such a mission operated by diesel locomotive. Due to low speeds and short runs the power demand is changing very fast and the system cannot reach its steady state. Thus the diesel engine is not allowed to operate in optimal mode. All these indicators lead to a solution with diesel engine operating in a constant power mode described below supported with some kind of non-conventional power source.



a)



b)

Fig. 3 a) One-day power mission of 740.3 class; b) Power mission example of standard shunting

Such a mission represents an industrial light shunting operation with low average-to-maximum power ratio and high frequencies of power demand, which leads to higher specific consumption and higher emissions of the diesel engine.

To compare the industrial shunting with standard shunting mission, Figs. 3a and 3b are introduced. The mission was recorded on 731 class with 600 kW diesel engine using a special data acquisition system on a standard light shunting mission in Kolín station. These two missions are nearly comparable but the power average. The speed shown in Fig. 3b fits the high relative braking energy presumption from statement II.

Lower rated power of a diesel engine leads to its better average utilization. On the other hand, such a locomotive is not able to serve higher train loads.

The common sign of both records and the most of diesel locomotive operations is a very low power average represented by $PD(AV)$ in Fig. 3a and $PT(AV)$ in Fig. 3b.

Depending on a power mission, the average-to-maximum ratio is usually 40 – 70% for track missions and 10 – 20% for shunters. Similar results are presented in papers over the world.

Shunting missions cannot be operated by a low powered locomotive. The maximal diesel power is needed for acceleration and high load train shunting. Low power is requested when locomotive runs the train at constant speed. The rest is idling time needed when coasting or to keep the engine on optimal temperatures.

Dealing with low power average and high power peaks a conventional diesel driven locomotive cannot be optimized for shunting operations. Electrification of yards or industrial railways is very expensive and therefore out of discussion. With a hybrid drive technology a shunting locomotive operation can be optimized.

4. Options of performance optimizations

Based on power mission profile the optimization process can be divided into two separated aims. The first is to save energy losses when braking; the second aim is to optimize diesel engine operation and to lower fuel consumption and emissions. Both aims call for some kind of energy storage and additional power source. Therefore, a diesel-electric power transition combined with electric accumulator is discussed.

A. Braking energy savings

To save the braking energy expressed by (6), designed energy storage should be able to store it. The main problem is not in the amount of energy but in the short time the energy is produced – the braking power. The braking power PB depends on demanded deceleration a, train speed v and its mass m (7)

$$P_E(t) = F_B(t) \cdot v(t) = m \cdot \xi \cdot a(t) \,.\, v(t) \qquad (7)$$

Where F_B is a braking force and ξ is a coefficient of rotational parts. There are three main limitations of electro-dynamic brake – maximal braking power, maximal braking force and the brake fade-out in low speed. The final braking force is generated by electro-dynamic brake in cooperation with conventional pneumatic brake.

Typical braking power values can be computed from accumulators' data shown in Table 2. Based on these values and typical accumulator parameters (Table 3), we can select the right accumulator type for the energy storage. Charging power allowed can be computed simply as

$$PCh = ChargingRatio \cdot C \qquad (8)$$

Where *ChargingRatio* is a parameter from Table 2 and *C* is capacity of the accumulator in kWh. To utilize the most of the braking power, the accumulator should be sized according to (8). For example, NiCd accumulator can be charged with 2C [4], which means 500 kWh for 1000 kW of braking power. Such an accumulator should weigh approximately 11.5 tons with 3.5 m^3 of volume, which might be unsuitable for some locomotives. Considering the charging current or charging power, conventional accumulators are out of discussion. Modern LiFePo accumulators became very popular thanks to their specific parameters. In other words, LiFePo battery for 1000 kW and 500 kWh weighs about 5 tons with 1.5 m^3 of volume. The main disadvantage of LiFePo accumulators is a high price [4].

Typical accumulator data comparison [5] Table 2

Parameter	Note	NiCd	NiMH	LiFePo	NiNaCl	SC
Cell Voltage [V]	*Nom.*	1.2	1.2	3.2	2.58	2.5
	Min	0.85	0.85	2.5		0
	Max	1.85	1.85	4.25	2.7	2.5
Specific Energy [Wh/kg]		45	70	100	100	5
En. Density [Wh/l]		150	230	330	172	0.07
Spec. Power [W/kg]		60	150	720	170	2000
Efficiency	*Amper-hour*	0.8	0.99	0.99	0.99	0.99
	Watt-hour	0.6	0.96	0.96	0.98	0.98
Discharg. Ratio[h^{-1}] / [A]	*Cont.*	1 C	2 C	3 C	1 C	150A
	Pulse	3 C	20 C	20 C	4 C	750A
Charging Ratio[h^{-1}] / [A]	*Cont.*	0.5 C	0.5 C	0.5 C	-	150A
	Fast	2 C	1 C	2 C	2 C	750A
Temp. [°C]	*Min*	−40	−20	−45	−40	−40
	Max	45	50	85	50	65
Life time [cycles]	*At 80% DOD*	3500	4500	5000	2000	1E6
Approx. Price [USD/Wh]		1.5	4.3	6.5	150	80

To reduce the accumulator size, a supercapacitor can be used. In this case, charging power is limited by the maximal current only and the supercapacitor is sized according to the kinetic energy of the typical train load. Stored energy helps with the next acceleration, thus the diesel can be undersized. The main disadvantage is low energy density and pricing. Due to dimension limitations, the capacity is limited to approximately 10 kWh for most of locomotives [4]. This amount of energy can help with acceleration but it cannot fit high loaded train moves between yards where the full power is needed for minutes. Thus supercapacitors are suitable for very specific applications with well-known periodic start-stop cycle, e.g. city buses, trams and light trains. Nevertheless, there is a project coming that will introduce a hybrid shunting locomotive with a supercapacitor of 6 kWh of energy [10]. The results will be presented in future publications.

B. Diesel engine performance optimization

On shunting locomotives, diesel engine operates at idle most of time and at low power with short time power peaks, as mentioned above. This leads to higher specific consumption and low efficiency [4].

Hybrid technology can help to solve this issue. Accumulator serves as a peak source with sufficient capacity to cover usual missions. Then, diesel engine can be undersized to overall power average of the mission. Such a design brings additional costs of the accumulator but the undersized diesel engine can partially compensate these costs and bring many advantages such as lower fuel consumption. There are other advantages of the hybrid drive, such as recuperative braking, silent or zero emission operation, etc. The most common hybrid concept for railway applications is Full Hybrid with series configuration and it can be found e.g. in RailPower project Green Goat [8]. In moder hybrid concept the AC traction drive should be used to improve overall efficiency [11].

To demonstrate these advantages, a hybrid vehicle simulation model was designed and analyzed and presented in [5]. The analysis is based on power mission from Kladno industrial railway and serves as an example of predesign of hybrid vehicles. The main results and a very brief description of the design method are introduced.

5. Hybrid vehicle design

There are several methods for hybrid vehicle design based on different attitudes. The very basic method, Power average method, introduced e.g. in [6] is based on the average of the mission power. Therefore the mission power time record should be known. The diesel engine is than undersized but with a sufficient power to move the train in case of the accumulator failure. Results of such a design have been introduced in 1980's with TA436 class CKD hybrid locomotive design [6].

Considering the time-power mission record is a time signal, it can be transformed by FFT (fast Fourier transform) and used for the hybrid drive design. The constant part of the signal is the power demand for the diesel engine, the rest of the power demand signal is divided based on the low-pass filter between the accumulator and supercapacitor. This method was published in [12] and [13].

The hybrid drive design is a multi-point optimization problem, thus evolutionary algorithms can be used. The use of particle swarm optimization [14] for the hybrid vehicle design was introduced in [5]. The cost function (correctness of the solution) is not a classic function, but it is a simulation model running with the real-world data. Overall analysis of the simulation is used to evaluate each step of the optimization process and the results are for the next step. Parameters of the final design are shown in Table 3.

The system uses the accumulator as the main power source, up to 576 kW. In the meantime the diesel operates on the average power of 50 kW. If the power demand is higher, the diesel engine power is set to full (up to 224 kW). If the battery state-of-charge is low, the diesel power rises slowly to charge the battery and to replace the battery in the role of the traction power source.

This concept can significantly lower the fuel consumption: from 9 to 16 kg per day [4]. Operating on the optimal power, the diesel engine emissions are lower. With recuperation, energy saving is up to 70% for every acceleration.

6. Return on investment

The supercapacitor increases input costs as mentioned above, thus the usage is optional and depends on the target application. In the following balance the supercapacitor is not considered.

Balancing the costs, fuel pricing is one of the most important factors. Nevertheless, the fuel costs are hardly predictable but the trend can be estimated based on the long term statistics [16].

Input costs can be compared with a conventional vehicle of similar performance and expressed as the additional costs of hybrid

Hybrid Locomotive Components

Table 3

Component	Parameters	Component	Parameters
Diesel engine	CAT C7 (157 – 224 kW) [15]	Supercapacitor (optional)	Type: Maxwell BMOD0094 35 F / 600 V Charging: 600 kW Discharging: 600 kW Mass: 600 kg No. of units: 24
Accumulator	Type: Saft NHP 10-100 NiMH, 100 Ah / 576 V Charging: 70 kW Discharging: 576 kW Mass: 893 kg No. of units : 48		

drive components. Considering the life-time of the accumulator (usually 5 years) and the fuel savings computed in the previous part.

Balance of the hybridization in 30 years is introduced in [4]. In the diesel performance optimization the costs can be returned in 24 years. Considering the energy savings due to the recuperation of the braking energy the investments can be paid back in 19 years. Such a long term return on investments is not ideal at all. But the prognosis of the fuel price gives the hybrid drive technology good position for the future. With the accumulator and supercapacitor technology development the costs of hybrid drive should lower the initial investments [5].

7. Conclusion

Hybrid technology is a great possibility to eliminate braking energy losses and to optimize internal combustion engine operation.

Based on the longitudinal analysis, the braking energy savings can be up to 70% of the overall energy during one start-stop cycle. Unfortunately, present electro-chemical accumulators are not suitable for the energy storage of such energy due to its low charging power limitation. Therefore, supercapacitors should be used.

Many diesel locomotives operating yards, station shunting and industrial railways missions must deal with high power peaks and long-term idling periods. This leads to non-optimal power utilization, higher fuel consumption and higher emissions. To optimize diesel engine operation, the accumulator should be used as a power peak source. The diesel engine can operate on its optimal power with the lowest specific fuel consumption. Instead of idling, the accumulator is pre-charged.

Hybrid locomotive design can be supported by computer technology and modern methods such as Particle swarm method. Based on simulations, performance parameters can be optimized off-line before the real locomotive is finished.

Although the costs of modern accumulators are still high, the hybrid drive technology should be considered as the future of city transport and some specialized operations such as yard switching and shunting. For the present implementation the return on investments should be measured in two tens of years.

Acknowledgment
We would like to thank to CZ LOKO Company for the support and help with the design. Measurements were made thanks to AWT Kladno support.

References

[1] BARRERO, R., MIERLO, J. V., TACKOEN, X.: Energy Savings in Public Transport, *IEEE Vehicular Technology Magazine,* vol. 3, pp. 26–36, Sep. 2008. ISSN 1556-6072
[2] IWNICKI, S.: *Handbook of Railway Dynamic,* CRC Press, Taylor & Francis. New York, 2006. ISBN 978-0-8493-3321-7
[3] CENE, B., RADOSAVLJEVIC, A., GAVRILOVIC, B.: Driving Energy Consumption of Electric Traction Vehicles. *Communications – Scientific Letters of the University of Zilina,* No. 3, pp. 50–56, 2008. ISSN 1335-4205
[4] PACHA, M.: *Hybrid Propulsion of Railway Vehicles,* PhD. Thesis, University of Zilina, Slovakia, 2010.
[5] PACHA, M., STEPANEK J.: *Energy Savings and Performance Optimizations of Hybrid Shunting Locomotives.* Proc. of 2012 Elektro, University of Zilina, 2012. ISBN 978-1-4673-1178-6
[6] POHL, J.: *Hybridní pohon posunovací lokomotivy / Hybrid Driven Shunting Locomotives,* CKD Report, 1982. unpublished.
[7] GE transportation – North America. [Online]. http://www.getransportation.com, 2009, Feb. [cit. 2012-12-05]
[8] COUSINEAU, R.: Development of a Hybrid Switcher Locomotive, *IEEE Instrumentation & Measurement Magazine,* pp. 25–29, Feb. 2006. ISSN 1094-6969
[9] Alstom. Hybrid Loco, A New Environmental-friendly Concept for Shunting. [online] http://www.transport.alstom.com, 2010, May. [cit. 2012-12-06]
[10] I-mash, Трансмашхолдинг работает над созданием гибридного тепловоза с асинхронным тяговым приводом /Transmash Holding is Working on Hybrid Locomotive with AC Traction Drive, [online] http://www.i-mash.ru, 2011, Feb. [cit. 2012-12-17]
[11] KUCHTA, J., FULIER, M.: Electrical Transmission of Modernized Locomotives of Independent Traction Power, *Communications – Scientific Letters of the University of Zilina,* No. 2–3, pp. 76–88, 2001. ISSN 1335-4205
[12] AKLI, C. R., ROBOAM, X., JEUNESSE, A.: *Energy Management and Sizing of a Hybrid Locomotive,* Power Electronics and Applications, 2007 European Conference, pp. 1–10, 2007, [Online]. http://www.ieee.org, 2007. [cit. 2010-05-20], E-ISBN 978-92-75815-10-8
[13] JAAFAR, A., AKLI, C. R., SARENI, B., ROBOAM, X., JEUNESSE, A.: Sizing and Energy Management of a Hybrid Locomotive Based on Flywheel and Accumulators, *Transactions on Vehicular Technology,* vol. 58, No. 8, pp. 3947–3957, Oct. 2009. ISSN 0018-9545
[14] ZELINKA, I., OPLATKOVA, Z., SEDA, M., OSMERA, P., VCELAR, F.: *Evolucni vypocetni techniky – principy a aplikace / Evolutionary Computation – Principles and Applications (in Czech),* Prague : BEN, 2009. ISBN 80-7300-218-3
[15] Caterpillar. [Online]. http://www.cat.com. 2009, Feb. [cit. 2012-12-18]
[16] EIA. Energy Information Administration. [Online]. http://tonto.eia.doe.gov. 2010, May. [cit. 2010-05-20]

Georgy Pavlov *

STUDY OF THE ENERGY PARAMETERS AND POSSIBILITIES FOR MODERNIZATION OF ELECTRIC VEHICLES FOR AC POWER

The main problem with thyristor locomotives operated in the country is the low value of the power factor. The paper presents results of a survey of the quality of power consumption in this type of locomotives. Based on the experimental data analyzed, we have determined the main parameters characterizing the traction power consumption. Indicated are some general guidelines for optimization.

Keywords: *Electric locomotives with thyristor rectifier, Power factor, Reactive Power Compensator.*

1. Introduction

The main requirement for AC-driven electric vehicles (EV) such as locomotives and electric multiple units (EMU) in operation in the railway system is to have the capability of smooth regulation of traction and braking forces by means of suitable control gear (CG) allowing the vehicles to maintain their energy parameters (i.e. power coefficient k_M) at high levels across the entire operational range. The power coefficient at non-sinusoidal current and voltage is calculated using the following formula:

$$k_M = \frac{P_{(1)}}{S} = \frac{U_{(1)} \cdot I_{(1)} \cdot \cos\varphi_{(1)}}{U_{eff} \cdot I_{eff}}, \text{ with } U_{(1)} \approx U_{eff},$$

$$k_M = \frac{U_{eff} \cdot I_{(1)} \cdot \cos\varphi_{(1)}}{U_{eff} \cdot I_{eff}} = \nu_1 \cdot \cos\varphi_{(1)}, \qquad (1)$$

where:

$U_{(1)}$, $I_{(1)}$, $\cos\varphi_{(1)}$ and $P_{(1)}$, are respectively the effective values of voltage and current of the 1st harmonic, the phase angle between them, and the active power carried by the main harmonic, W;

U_{eff}, I_{eff} – the actual effective (True rms) voltage and current;

ν_I – the current non-sinusoidality coefficient, $\nu_I = I_{(1)}/I_{eff}$;

S – the apparent power, VA.

A large percentage of the electric vehicles used in Bulgaria's railway system have diode and thyristor control gear (current rectifiers) and direct-current motors. Only the SIEMENS DESIRO Classic electric railcars by Siemens are equipped with four-quadrant power converters with IGBT, which allow for regulation of the power factor and propulsion by induction motors.

The main problem which all alternating current EVs with DC drives face, irrespective of their particular solution for regulating the mode of operation, are the deteriorated energy parameters. The combination of increased consumption of reactive power and the deformation of the propulsion current curve leads to a lower power coefficient. All of this is further compounded by the additional losses in the traction power supply system thereby decreasing its efficiency.

It is a well known fact that Bulgaria's rolling stock is worn out and technologically obsolete. It needs urgent renovation by replacing part of the EV fleet and upgrading certain vehicles with IGBT [1] converters. Of course, for a vehicle to be eligible for upgrade, it must have a sufficient residual operational life. This should be determined by a detailed preliminary assessment of the technical status of the available rolling stock.

For the EVs with diode converters used in the Bulgarian railway system (locomotives of the 44 and 46 series) increasing k_M can be achieved by remodeling their power circuitry to retrofit power inverters but leaving most of the propulsion installation intact. For the locomotives of the 44-001 and 46-200 series previously upgraded with thyristor rectifiers this approach would require for the rectifiers to be removed, which would not be economically justifiable. A better solution for them would be to outfit them with active compensators for reactive power connected to the secondary windings of the locomotive's propulsion transformer.

This paper is based on data from experimental measurements of the energy parameters of locomotives with thyristor rectifiers conducted in real-life operational environment. It proposes technical solutions for upgrading this type of EVs to decrease their consumption of reactive power and improve their power consumption profile.

---

* **Georgy Pavlov**
Transportation Power Supply and Power Equipment Department of the Todor Kableshkov University of Transport, Bulgaria, Sofia,
E-mail: g_pavlov61@abv.bg

2. Experimental study of the energy parameters of an upgraded, series 46-200 locomotive

Since 1998 the KONČAR – ELEKTRIČNE LOKOMOTIVE d.d. factory in Zagreb, Croatia has upgraded 13 locomotives from the BDZ fleet. These include 2 locomotives from the 44 series and 11 from the 46 series (their codes have been amended as follows: 44-001 and 44-002, and Series 46-200). The upgrade consisted in retrofitting thyristor rectifiers with natural commutation and two-zone voltage adjustment capability. In terms of energy performance, this solution is not optimal and leads to deterioration in the power consumption indicators.

There are a wide range of technical solutions for improving the energy parameters of EVs with thyristor current rectifiers (single-mode thyristors). However, only two of them seem to be favored in the industry. The first relies on using multi-zone adjustment (up to four adjustment zones) and yields relatively good results in terms of k_M. Increasing the number of the adjustment zones provides for better energy parameters (increased k_M) but makes the construction of the propulsion transformer increasingly more complicated and is therefore avoided. The second solution for increasing k_M relies on the use of artificial commutation in the thyristor rectifiers. It offers a significant improvement of the EVs energy parameters [2].

To study and analyze the energy parameters of the upgraded EVs from the BDZ fleet, we conducted experimental measurements on trains drawn by 46-200 Series locomotives across the territory of pre-selected traction power substations and feeders. The measurements were performed using microprocessor measurement devices (MC 750 and UMG 604E grid analyzers). The devices performed real-time measurement of voltage and current. The data was then used to calculate the effective electric power (active, reactive, and apparent) and the power factor. A harmonic analysis of the propulsion current and voltage curves in the overhead lines is performed [3]. The measurements were taken with an integration period of 1 minute. All specified values were stored in the devices' onboard memory chips. The analyzers were connected to the traction power supply system via 800/5 A current transformers and 35000/100 V voltage transformers.

We presented the results from measurements conducted in the area of Iliantsi Propulsion Power Substation on a train drawn by an upgraded, Series 46-200 locomotive with rated traction power of 5 400 kW. No other trains were operating in the area during the measurements. Based on the data gathered during the measurements, we were able to determine the main energy parameters for the studied type of locomotive.

Figure 1 shows the change in propulsion current I_{el}, A of the locomotive between 12:00 PM and 12:50 PM. This time of day was selected because it featured all main movement patterns of the train such as acceleration according to a selected algorhythm, cruising at a predetermined speed, and dynamic resistor braking.

Figure 2 shows the changes in the active (P, W), reactive (Q, VAr) and apparent (S, VA) power of the train. We witnessed a significant consumption of reactive power with inductive nature, which was similar in its value to the consumed active power. Within the time interval subject of our study the train had performed several run sequences for each of which the train engineer had chosen different algorhythms. During each of those, as we observed, the ratio between the reactive and active power changed and between 12:16 PM and 12:30 PM the reactive power had higher values than the active. From the charts in Figs. 2 and 3 we can conclude that to a large degree these values for the active and reactive power depend on the pattern of movement of the EV.

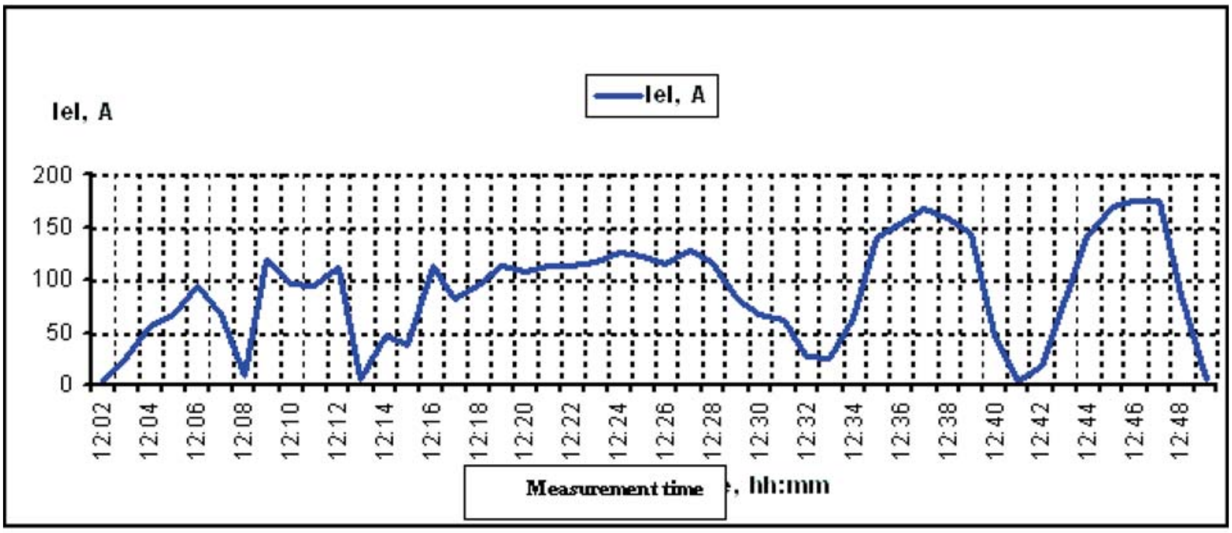


Fig. 1 Locomotive Propulsion Current Dynamics

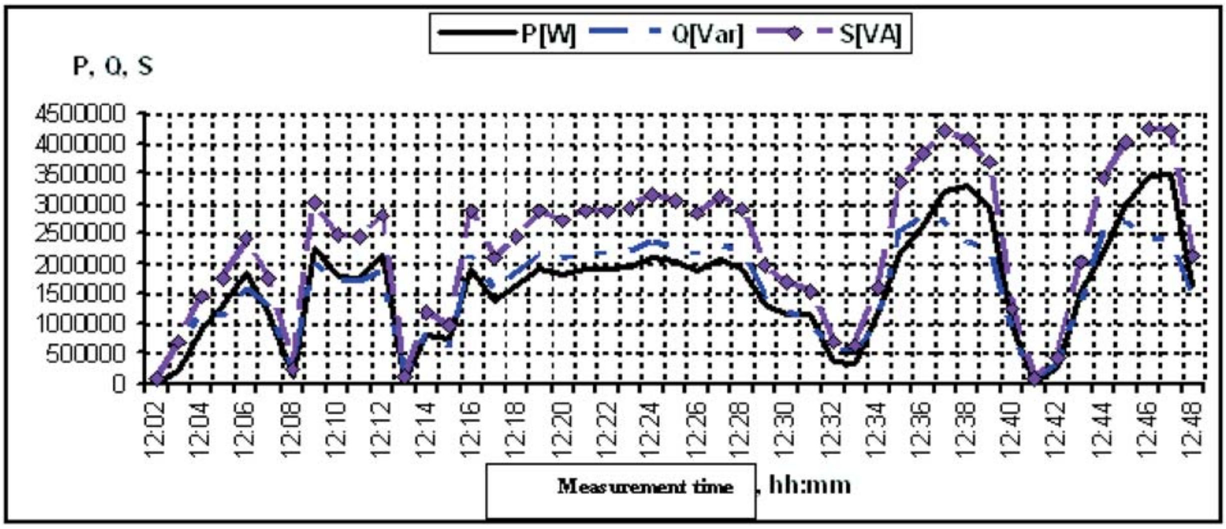


Fig. 2 Locomotive Active, Reactive, and Apparent Power Dynamics

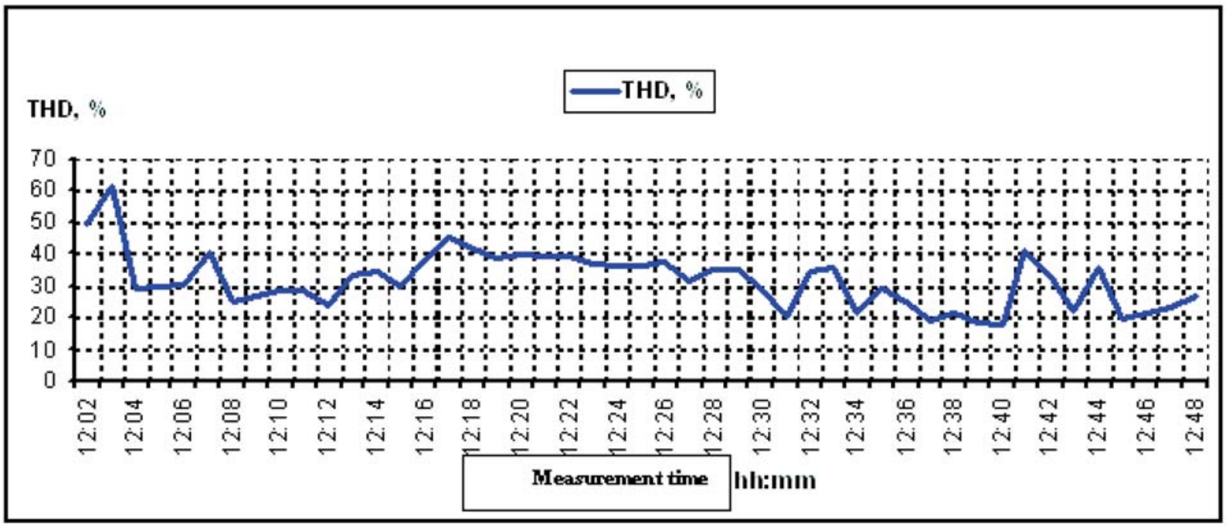


Fig. 3. Propulsion Current Harmonic Distortion Coefficient Dynamics

Figure 3 provides a graphical representation of the dynamics of the THD_I, % indicator (total harmonic distortion) of the propulsion current as measured at the 27.5 kV terminal of the traction power supply substation. This indicator characterizes the distortion of the current curve and the presence of high-order harmonics as a percentage of the effective value of the current of the first harmonic $I_{(1)}$.

Over the specified time interval the value of THDI, % varies widely (from 20 % to 60 %) depending on the movement pattern of the train and the current value of the power. By using the calculation device [4] and formula (2) below, we managed to determine the change in the non-sinusolidality coefficient of the propulsion current (νI), which itself varied between 0.858 and 0.98.

$$\nu_I = \frac{1}{\sqrt{1+\left(\frac{THD_I\%}{100}\right)}} = \frac{1}{\sqrt{1+\left(\frac{\sqrt{\sum_{i=2}^{\infty} I_{(1)}^2}}{I(1)}\right)^2}} = \frac{1}{\sqrt{\left(\frac{I_{eff}^2}{I_{(1)}^2}\right)}} = \frac{I_{(1)}}{I_{eff}} \qquad (2)$$

In addition to the above, we studied the changes in power factor PF_{el} and the results are presented in Fig. 4. Due to the low content of higher-order harmonics in the voltage supplied ($THD_U\% <$ $< 10\%$), the measured PF_{el} (in accordance with the method described in [4]) almost coincided with the power coefficient k_M,

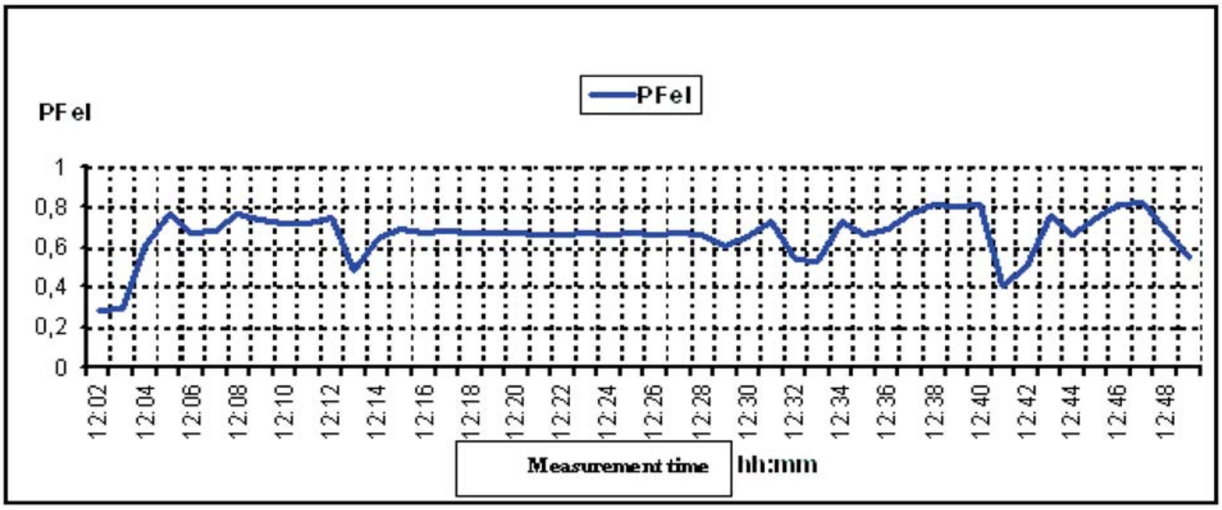


Fig. 4 Locomotive Power Factor Dynamics

i.e. $PF_{el} \approx k_M$. The chart shows that for the most part of the measurement interval it remained under 0.8. Between 12:19 and 12:22, when the train had a relatively steady rate of power consumption, the average value of PF_{el} remained around 0.67. By a more detailed analysis of the measurement data we determined cos $\varphi_{(1)}$ to be equal to 0.72 and the nonsinusoidality coefficient of the propulsion current (v_I) to be equal to 0.93. On the basis of these results we inferred that the low k_M value is to a significant degree the result of the consumption of reactive power. We must also note that under this movement pattern, the locomotive's current power was about 36 % of its rated power. During the intervals 12:37 PM to 12:39 PM and 12:45 PM to 12:47 PM when the locomotive was operating at over 60 % of its rated power, these parameters had the following values: cos $\varphi_{(1)} \approx 0.81$; v_I, ≈ 0.98; and PF_{el}, respectively k_M was equal to 0.80–0.81. This indicates that increasing the power leads to a significant improvement of the energy parameters of the EV. However, in a real-life operational environment this type of movement pattern is difficult to achieve.

During our studies we conducted numerous measurements on locomotives from the 44-001 and 46-200 Series. Different measurements on different trains under identical patterns of movement showed no significant differences of the power coefficients in terms of shape or form. The results gathered were completely consistent with those from similar measurements conducted within other studies [5].

Under the existing scheme, after the upgrade of the studied locomotives, the best way to improve their energy parameters is to retrofit them with active compensators for reactive power (ACRP)

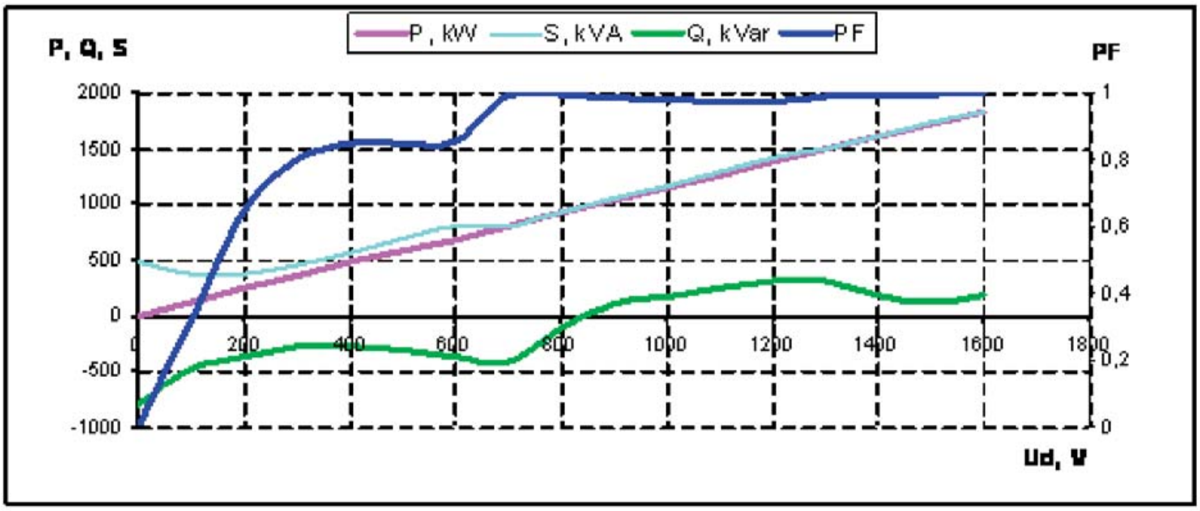


Fig. 5 Power measurement and Energy Parameters in Relation to the Average Value of the Rectified Voltage

by incorporating these devices in the power circuitry of the EVs. There have been numerous trials of locomotives with thyristor rectifiers retrofitted with ACRP and the results from them have been described in detail in various sources. They show that k_M for propulsion rises from 0.8 to 0.96 and the consumption of Q sees a significant decrease [6, 7].

To further study the process, a k_M study was conducted using non-adjustable Reactive Power Compensators (RPC) set at $Q = 650$ kVAr connected to a group of secondary windings in the propulsion transformer.

Figure 5 shows the power measurements and the energy parameters in propulsion mode with 1180 A of current supplied to the engine and the RPC turned on. They indicate that during zone-2 operation and, for the most part, during zone-1 operation, k_M is over 0.9.

It must be noted that decreasing the propulsion current of the EV leads to an expansion of the overcompensation area into a larger zone. As with undercompensation, overcompensation of reactive power leads to additional increase in current loads in the traction power supply system and therefore has to be avoided. Furthermore, overcompensation increases the voltage in the receiver and can cause emergency shutdown of the power supply. For these reasons onboard compensators must be of the active (adjustable) type to allow their operation to be set in accordance with the particular requirements of the EV.

3. Conclusion

The trial measurements of the power consumption qualities of the upgraded electric locomotives of the 44-001 and 46-200 Series performed across numerous substations show unsatisfactory results in terms of their energy parameters. In real-life operational environments, the power factor PF_{el} varies between 0.6 and 0.8, which is substantially lower than the European norms. Figure 3 indicates that across the different patterns of movement the reactive power consumed by the locomotives is significant and that translates directly into a poor power coefficient.

Upgrading the thyristor locomotives operated within the Bulgarian railway system is unavoidable in order for the system to comply with the European norms and requirements regarding the quality of consumed power. This leaves us with two options:

A) To remove the thyristor rectifiers and retrofit four-quadrant converters with IGBT. The advantage of this option is that it will ensure optimal energy parameters across the entire adjustment range and allow recuperation. Its disadvantage is that it would not be economically justified.

B) To design and retrofit the EVs with active compensators for reactive power (ACRP). This would require a considerably smaller investment and would be more problem-free in comparison. Studies show that in terms of its energy footprint, this solution would guarantee compliance with the requirements of the European norms.

With mathematical models and trial measurements of the performance of the upgraded trains we can optimize the process of designing of active solutions to compensate the reactive power consumed by the vehicles. To improve the energy parameters and the quality of the power consumed by the vehicles, we must develop an innovative technical solution to upgrade them, namely an active compensator for reactive power (ACRP) to be connected to the secondary propulsion windings of the locomotive transformer. This, to a large extent, will eliminate the main problem related to the operation of these EVs, which is related to the consumption of considerable amounts of reactive power from the traction power supply system, and would improve the energy efficiency of the specialized railway power supply system.

References

[1] PAVLOV., G.: *Increasing the Energetic Parameters of Electric Vehicles Using Input Pulse Rectifiers,* Proc. of Seminar Electrical Transport Engineering (in Bulgarian), pp. 23-28, National Association of Electrical Engineers in the Transport in the Republic of Bulgaria, 2002.

[2] Dissertation thesis: *Investigation of Artificial Switching Converters in Electric Rolling Stock Power-line,* Technical University of Sofia, 2001

[3] DIMITROV G., PAVLOV, G.: Specialized Modular Measurement Stand for Research on Power Consumption and Energy Efficiency in Transport, International Scientific Conference Transport 2011, Sofia, *Mechanics, Transport, Communications AJ,* No. 3, article No. 0597, 2011, ISSN 1312-3823.

[4] MACK GRADY, W., GILLESKIE, R. J.: *Harmonics and How They Relate to Power Factor,* Proc. Of the EPRI Power Quality Issues & Opportunities Conference (PQA '93), San Diego, November 1993.

[5] MATOV P., VETSKOV, A.: *Efficiency of Alternating Current Traction Network in a Uniflow Track Section,* XVIII Intern. Scientific Conference Transport 2008, Sofia, Mechanics, Transport, Communications AJ, ISSN 1312-3823, No. 3, article No. 0293, 2008.

[6] DONSKOJ, D. A.: *Improving Energy Performance ac Electric Locomotives.* Proc. of Scientific and Technical Conference Science-Transport-2006, Moscow Engineering, 2006

[7] DONSKOJ, D. A.: Improving the Efficiency of Electric Traction on AC. *J. of Electronics and Electrical Equipment in the Transport.* No. 5, 2006.

Sona Pavlikova – Dusan Maga – Boris Simak *

INVERSED MATRIXES IN ELECTRIC CIRCUITS PARAMETERS ESTIMATION BY ORIENTED GRAPHS

The base of electric circuit analysis is often set up on numerical methods. When taking a look into the principle of any numerical method, a matrix calculus always rises out. The larger the problem is, the larger is the size of the matrix, representing the logical structure of the solved problem [1]. The problem (especially from the mathematical point of view) is in acquiring the inverse matrix, which can be solved with some of a large scale of available numerical (or analytical; based on the size of matrixes) methods. This paper shows another possible approach to obtaining the inverse matrix. The graph theory has already been successfully used to invert the matrixes; although, the idea of evaluating the alternating paths, presented by authors, is completely new and had not been used before. The utilization of the designed procedure is demonstrated on a part of an analogue signal processing filter. This technology gives a new tool to software designers and could speed up the numerical solutions; or, in case of impossible analytical and/or diverge numerical technique can lead to a successful and reasonable solution.

Keywords: *Invert of matrixes, mathematical representation of electric circuits, graph theory.*

1. Introduction

Graphs are historically and generally considered as geometrical objects. In spite of this, they can be represented in many different ways: lists of edges, adjacency matrixes or by a lot of other different structures [2]. The reality represented by graphs can be anything between chemical formulae, mechanical system, map structure and/or computer processes. Graphs are visually impressive and can easily introduce the observer into the core of the problem.

The paper will introduce the necessary base of graph theory to the readers and describe the inverse matrix acquisition based on graph theory. This means that neither numerical nor analytical method is used [3, 4]. This new approach can both spare computational time (for certain types of problems) and lead to solution where numerical or analytical methods cannot be used. In addition, the presented technology can be implemented into the cores of numerical (or analytical) tools to obtain the required solution of the matrix formulated problem. The paper is based on latest research of authors presented in [5] and completed with description of background theory and presenting new and completed examples.

The paper describes the principles of topological description of electric circuits based on graphs theory. The electric circuit, described by its topology, could easily be represented by graphs. The equivalent circuit of an asynchronous rotary machine can be presented as an important and well-known example (Fig. 1). The parameters from Fig. 1 are generally known and it is not necessary to describe them in this paper. Important is that the edges in Fig. 1 are easy to identify. The result of this assignment can be seen in Fig. 2. These steps might be successfully used for almost every rotary or linear [6] machinery properties analysis.

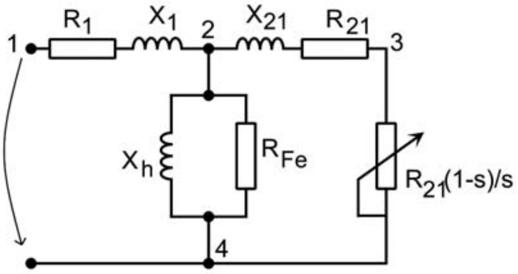


Fig. 1 Equivalent circuit of an asynchronous machine

The utilization of different graph types in electric circuits' analysis is not a new idea at all. There are technologies (based either on oriented graphs or directly describing the topological structure of the circuit) used parallel to numerical or analytical techniques [7]. One of the most referred are signal graphs [5]. Signal Graphs are mapping the flow of electrical signals through the given circuit. They are able to interpret the graphical representation of Kirchhoff's Laws; nevertheless, they are not able to follow the topology of the circuit. The signal graph, following the given rules, can be simplified together with a change of individual legs

---

* **Sona Pavlikova1, Dusan Maga2, Boris Simak2**
[1] Slovak University of Technology, Bratislava, Slovakia, E-mail: sona.pavlikova@stuba.sk
[2] Czech Technical University in Prague, Prague, Czech Republic, dusan.maga@fel.cvut.cz

transformation. Finally, the algebraic expressions are obtained from the graph structure. The investigated parameters can easily be evaluated according to a numerical (or analytical – where possible) solution of equation sets representing Kirchhoff's laws.

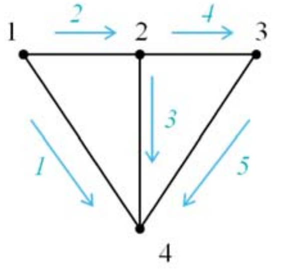


Fig. 2 Edges of the circuit

2. Topological graphs

In contrast to oriented signal graphs, this type of graphical structure represents directly the topological structure of the circuit. The vertices are the nodes and the edges are the arms (Fig. 1). The topological graph can be used for analytical representation of a circuit's topological structure in the form of matrix. The matrixes obtained from these are known as *Incidence topological matrixes*.

To be able to build the topological graph of the circuit, it is necessary to replace the dipole components of the circuit with their representative – edge of the graph. The most important term in graph theory is the graph tree. This definition covers such a subgraph, which contains all of the vertices, but no closed cycle. The edges of the graph forming the tree are called arms, while the complementary edges are called chords.

The graphical connection can be transformed into an algebraic representation with several types of topological matrixes.

3. Incidence matrix

The adjacency matrix of oriented graph with vertices $v_1, ..., v_n$ is a square matrix $A(G) = (a_{ij})$ of n-th degree, where a_{ij} is equal to number of edges between vertices v_i and v_j. The incidence matrix $M = (m_{ij})$ is a matrix with size of $p \times q$, where p represents the number of vertices and j number of edges. The value of m_{ij} is equal either to 1 or 0 – based on incidence of vertex *i* and edge *j*. It is also important to define the diagonal matrix $D\ m_{ij}\ (d_{ij})$, where value of d_{ij} is equal to degree of vertex *i*. The relation between the adjacency matrix and incidence matrix is as presented:

$$A = D - MM^T \tag{1}$$

The concrete construction of the incidence matrix can be illustrated at an example presented in Fig. 1 and its graph representation (Fig. 2). As can be seen in Fig. 2, there is either one or none connection (edge) between the vertices. The rows of the matrix are based on vertices and the columns on edges of the graph. As described above, each row i of this matrix express which edges are connected to vertex *j*. For example, edge 1 leaves from vertex 1 so m_{11} is equal to -1. The same goes for edge 2 and vertex 1: m_{12} is equal to -1 again. Edges 3, 4 and 5 have no connection (incidence) to vertex No. 1, so their values are equal to 0. The edge 2 arrives to vertex 2: the value of m_{22} is equal to 1. Based on these steps the following matrix is obtained:

$$\begin{pmatrix} -1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 1 & 0 & 1 & 0 & 1 \end{pmatrix} \tag{2}$$

It is generally known that any row in the incidence matrix is a linear combination of the rest of the rows. This means that one of the rows contains redundant information and it is possible to leave it out of the matrix. It is typical to quit the row corresponding to reference (grounded) node, e.g. the fourth row from the above presented incidence matrix. The obtained type of matrix is called *node matrix* and is labeled as R. For the case presented in Fig. 2 it is equal to:

$$R = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{pmatrix} \tag{3}$$

When assigning every n^{th} graph vertex to potential φ_n (relative to reference node), each single component a_{nk} of matrix *R* can be considered as an expression of potential φ_n influence on current flowing through edge *k*. When the voltages on the edges are represented in the form of column vector *u* and the potentials of vertices in the form of column vector φ, then the voltage of edges can be calculated as:

$$u = R^T \varphi \tag{4}$$

The matrix *R* and the column vector $\underline{u}$ can be separated into parts R_v, R_t and u_v, u_t. These components correspond to tree arms and chords. The dependence between the potential of each vertex and the voltage of each edge can be obtained from (5).

$$\begin{pmatrix} R_v^T \\ R_t^T \end{pmatrix} \varphi = \begin{pmatrix} u_v \\ u_t \end{pmatrix} \tag{5}$$

More detailed description of connection to electric circuits presented by the authors can be seen, e.g. in [5].

4. Inverse matrix by graph construction

There are numerous methods used to obtain the inverse matrix. One of them is based on the graph construction. The method is described in a simple example with alternating paths (the term alternative path covers those paths where the edges of 1-factor and edges out of 1-factor are alternating, and it begins and ends with edge out of 1-factor [4]). The matrix *A* is adjacent to graph *G*. There also is an inverse graph G^{-1} with its adjacent matrix $A(G^{-1})$. The matrix obtained this way is an inverse matrix to the original one. Graph G is understood as a finite non-oriented graph without loops with the possibility of multiple edges [8]. The labeling of graph G is represented as $l{:}E(G) \rightarrow R$, where R is the set of real numbers. When labeling l on graph G is available, the definition of adjacent matrix $A = (a_{ij})$ to graph G is a square matrix of order n, where:

$$a_{ij} = \begin{cases} l(v_i, v_j) & if\ (v_i, v_j) \in E(G) \\ 0 \end{cases} \tag{6}$$

Graph G is called bipartite graph (or simply bigraph) when it does not contain a circle of odd length. The non-empty sets $R \subset \subset V(G)$ constitute the bipartition to bigraph $G = (V(G), E(G))$ when $R \cap C = \varnothing$ and $R \cup C = V(G)$ and none of two vertices from R (or C) are connected with an edge. When the bipartition (R, C) of bigraph G is given, the $R \times C$ of matrix $B_{RC}(G) = b_{ij}$ can be defined ($|R|$ means cardinality of set R) as follows:

$$b_{ij} = \begin{cases} l(v_i, v_j)\ if\ (v_i, v_j) \in E(G),\ where\ v_i \in R,\ v_j \in C \\ 0 \end{cases} \tag{7}$$

This matrix is called bipartite matrix of graph G. It is obvious that:

$$A(G) = \begin{pmatrix} 0 & B^T_{RC}(G) \\ B^T_{RC}(G) & 0 \end{pmatrix} \tag{8}$$

Additionally, if $B_{RC}(G)$ is a regular square matrix then also $A(G)$ is regular. Graph G contains 1-factor (perfect matching) if a factor of the graph with each vertex with a degree of 1 exists. Symbol $M(G)$ or (M) will be assigned to 1-factor of graph G (if included in) and symbol G/M to the graph obtained from G by contraction on edges of 1-factor [9]. The path with n vertices is labeled as P_n. The path P in G of length k is a subgraph of G isomorphic with P_{k+1}. Let G contain 1-factor $M(G)$. The path P is in G alternating in respect of $M(G)$ if one of the following criteria is fulfilled:

- P is of length 1 (i. e. edge),
- from each pair of adjacent edges in P there is just one belonging to $M(G)$ and P begins and ends with an edge not belonging to $M(G)$.

The P is even (odd) alternating path when P contains even (odd) number of edges out of $M(G)$. When P is only an edge, it will be considered even (odd) alternating path if it belongs (does not belong) to $M(G)$.

Let G be a bigraph without multiple edges with unique 1-factor $M(G)$. Nodes u and v are arbitrary vertices in G. Then the number of even (odd) alternating paths connecting the nodes u and v can be specified as ${}_{p+}(u,v)$ $({}_{p-}(u,v))$.

Let G be a bigraph with unique 1-factor $M(G)$. Then the labeled bigraph G^{-1} with label $l : E(G) \rightarrow Z$ can be obtained as follows [1, 10–12]:

- $V(G^{-1}) = V(G)$
- let vertices $u,v \in V(G^{-1})$; then the edge $(u,v) \in V(G^{-1})$ if and only if ${}_{p+}(u,v) \neq {}_{p-}(u,v)$ and $l(u,v) = {}_{p+}(u,v) - {}_{p-}(u,v)$.

5. Examples

A simple general example to illustrate the described theory has been presented in [5]. A bipartite matrix $B(G)$ is defined to simplify the transcription.

It is obvious that the matrix $B(G)$ is square and it is regular as well. The concrete example can be transformed, according to the rules described above, to the graph structure presented in Fig. 4. The description of vertices corresponds to 1-factor. Since the 1-factor and the alternating paths are very important in the process, it is recommended to mark these components. So, the edges of 1-factor are displayed bold in Figs. 4 and 5. The bipartite matrix of graph G is:

$$B(G) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \end{pmatrix} \tag{9}$$

The practical realization of defined rules to follow is very simple. Point No. 1 from Fig. 4 (represents the input to the system) is connected only to the resistor R_1 (Fig. 3) at the beginning part of the scheme (point 1′ in Fig. 4). Thus the value B_{11} is equal to

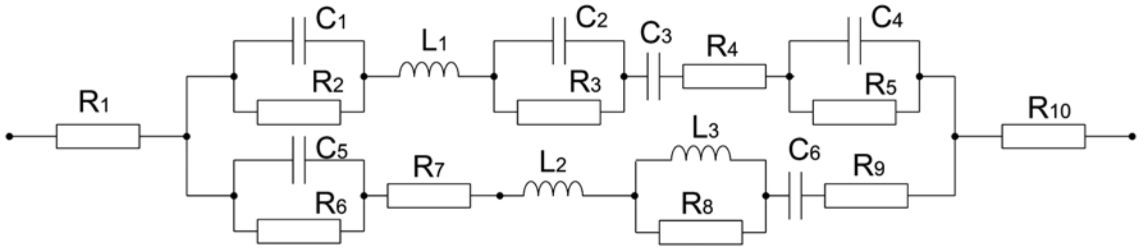


Fig. 3 Example of the electric circuit scheme

1, the other values in the first row are equal to 0. On the other hand, the parallel connection of resistor R_2 and capacitor C_1 (represented by point No. 2 in Fig. 4), has an association to points No. 1′ and 2. So the relevant values in $B(G)$ are set to 1 while the others in the second row are set to 0. This process leads to construction of complete matrix $B(G)$, finished by line No. 8, where R_{10} (point No. 8 in Fig. 4) has an association to points No. 4′, 7′ and 8′.

When constituting the graph G^{-1} (with bipartite matrix $B^{-1}(G)$ – an inverse matrix to $B(G)$) the below described steps have to be followed:

1, the edges of 1-factor are put into the graph G^{-1} and a value of 1 is assigned to them (Fig. 4).
2, the other edges are included into the graph structure following the algorithm:
   - if between two vertices u' and v an alternative path(s) exists, then the number of even $_{p+}(u',v)$ and odd $_{p-}(u',v)$ occurrences have to be counted. A new edge will be added in case when $_{p+}(u',v) \neq {}_{p-}(u',v)$ and its value (label) is $l(u',v) = {}_{p+}(u',v) - {}_{p-}(u',v)$,
   - if not, the edge is not added.

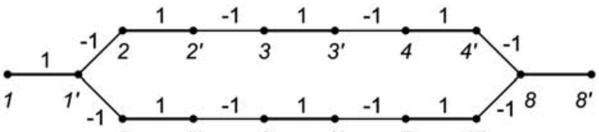


Fig. 4 Structure of graph G

The even alternative path has an even number of edges out of 1-factor, while odd has an odd number of those edges. The graph constructed following these rules can be seen in Fig. 5. The corresponding $B(G^{-1})$ is:

$$B(G^{-1}) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & -1 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 1 & -1 & 1 & 0 \\ 2 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \end{pmatrix} \tag{10}$$

The practical realization of the above described steps is again very easy. For instance, when analyzing the existence of alternative paths to point No. 4, only two of them exist (Fig. 5):
- from 2′ – this includes 3 occurrences, the final value assigned to this path is: 1, and
- from 1′ – this includes 5 occurrences, the final value assigned to this path is: −1.

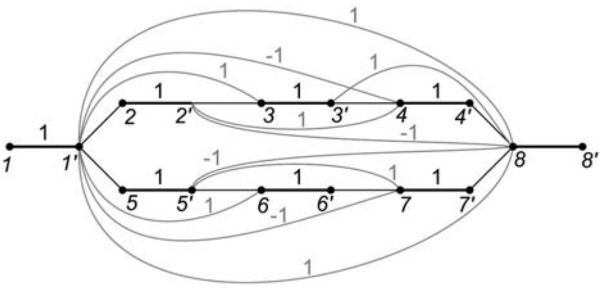


Fig. 5 Paths in G

Together with paths:
- from 3′ – assigned as −1, and
- from 4′ – assigned as 1

the fourth row of $B(G^{-1})$ is set to (−1, 1, −1, 1, 0, 0, 0, 0).

The alternative paths illustrated in Fig. 5 represent all available alternatives. The second half of the paths can be easily constructed based on problem symmetry; although this cannot be generalized. Nevertheless, one important remark has to be made in this point: the alternative paths coming to point No. 8 are as illustrated in Fig 5, but because of the mentioned symmetry case there are two alternative paths from 1′ available. Thus the value of B_{81} is equal to +2!

Finally, it can easily be proven that the bipartite matrix of graph G^{-1} is an inverse matrix to $B(G)$:

$$B^{-1}(G) = B(G^{-1}) \tag{11}$$

The procedure for adjacent matrix and its inverse matrix establishment is presented in [5].

At the conclusion of this chapter it has to be written that a fully new principle for obtaining the adjacent matrix has been successfully tested. The bipartite matrix (as a base of the adjacent matrix) is a fundamental component for (almost) any following both numerical and analytical analysis of (but not only) electric circuits. Even though this, it also has to be mentioned that there are certain issues where this technology cannot be applied or where the computational complexity of this kind of solution does not allow its successful implementation.

6. Conclusion

This paper is based on the authors' results presented in [5] completed with the latest knowledge and authors' experience. According to these, the base of any numerical method used in technical sciences to analyze and simulate the technical processes is constituted on matrix calculus [1]. This is especially clear for electric circuits used in electrical engineering and power electronics [13-16], where the physical reality and description of the circuit is transformed into a matrix form. The numerical methods in these cases are based on inverse matrix calculus. This can be done either in an algebraic or numerical way.

The presented paper discovers a purely new technique to obtain the inverse matrix necessary to construct the sets of equations. This technique is based on graph theory and represents a unique and new approach to numerical solutions of electrical problems. Generally, according to the presented technology, only the number of components is important for building the adjacent matrix. The concrete value of components is not important for this step of the solution and may vary from element to element.

The method itself had been introduced in the paper. The possible way of electric circuit representation by graph structures had been presented, as well as the matrix construction based on the obtained graph. The mathematical background of inverse matrix by graph construction had been presented, including the alternating path influence on matrix system construction.

Compared to the authors' results presented in [5] the background theory was elaborated in a more precise way. The practical example (asynchronous machine) were chosen to demonstrate the theory and possible application of graphs. More complex example is presented here by the authors and a simple solution based on bipartite matrix is offered. This example (Fig. 3) can be considered as an analogue filter used as a basic building block in signal processing. The topology of the filter was chosen to demonstrate the mathematical concept of the designed procedure. To be able to follow the procedures a detailed description of the most important steps is also presented.

The advantages of the presented unique method can be summarized as the following:

- the inverse matrix can be obtained without using either algebraic or numerical technique; the necessity of algebraic description of the task requiring the computation of inverse matrix is completely eliminated,
- graphical (clear) visibility,
- compared to [17] an alternating paths are taken into account,
- a simplified solution based on bipartite matrix is suggested.

The presented method has also several disadvantages. These might sometimes also lead to unsuccessful application of the presented methods:

- necessity of graph scanning and finding the alternating path(s), including the finding the 1-factor,
- algorithm is valid for bipartite graph structures only.

On the other hand, there is a computational complexity of the method. As has been mentioned above, the method based on alternating paths is a completely new method. The searching of 1-factor is sufficiently described in the literature, so it is not presented in the paper. The authors describe the principles and offer an example – the question of computational complexity is outside the scope of this paper. Nevertheless, it is clear that the presented method can be successfully used mostly with "sparse" graphs with limited number of edges.

Acknowledgement

This paper has been supported by the following projects:

Theory of Numbers and Its Application in Fuzzy Modeling, Grant No. Vega 1/0569/12, STU Bratislava, Slovakia.

Developing Skills of Scientific a Research Employees at the CTU in Prague, Grant No. OPPA CZ.2.17/1.1.00/32205, CTU Prague, Czech Republic.

References

[1] FABO, P., PAVLIKOVA, S.: *GSIM – Software for Simulation in Electronics.* 14th Intern. Power Electronics and Motion Control Conference, EPE-PEMC 2010, New York: IEEE, 2010, pp. S414–S417, ISBN 978-142447854-5.

[2] ARCHDEACON, D.: *Topological Graph Theory – A Survey*. [p. l.]: [Citeseer], 1996, 67 p.

[3] KONAR, R., MICIAN, M.: Numerical Simulation of Residual Stresses and Distortions in Butt Weld in Simulation Programme Sysweld, *Communications – Scientific Letters of the University of Zilina,* vol. 14, No. 3, pp. 49–54, 2012, ISSN 1335-4205.

[4] DADO, D. et al.: Numerical Investigation of Optical Burst Switching. *Communications – Scientific Letters of the University of Zilina*, vol. 10, No. 2, pp. 20–24, 2008, ISSN 1335-4205.

[5] PAVLIKOVA, S., MAGA, D., SIMAK, B.: *Graph Based Inversed Matrix in Electric Circuits Solutions,* Proc. of 9th Intern. Conference, Elektro, 2012. pp. 443–447. ISBN 978-146731179-3

[6] RAFAJDUS, P. et al.: Static and Dynamic Analysis of Linear Switched Reluctance Machine. *Communications – Scientific Letters of the University of Zilina,* vol. 13, No. 4, pp. 6–11, 2011, ISSN 1335-4205.

[7] PAVLIKOVA, S.: *Extending the Class of Matrices Invertible Using Graphs. 6th Conference on Mathematics and Physics at Technical Universities,* Brno: UO, 2009, pp. 203–210, ISBN 978-80-7231-667-0.

[8] CVETKOVIC, D., DOOB, M., GUTMAN, I., TORGASEV, A.: *Recent Results in the Theory of Graph Spectra.* Annals of Discrete Mathematics, Amsterdam : North-Holland, 1988, ISBN 0-444-70361-6.

[9] Godsil, C. D.: Inverses of Trees, *Combinatorica,* 5, pp. 33–39, 1985.

[10] PAVLIKOVA, S., KRC-JEDINY, J.: On the Inverse and the Dual Index of a Tree. *Linear and Multilinear Algebra,* vol. 28, pp. 93–109, 1990, ISSN 0308-1087.

[11] PAVLIKOVA, S.: Application of Eigenvalue Linear Mapping with Utilization of Graph Representation of Linear Mappings. *Mechatronika 2009,* Trencin : TnUAD, 2009, pp. 99–102, ISBN 978-80-8075-392-4.

[12] PAVLIKOVA, S.: Applications Graphs Spectral Properties in Technical Sciences. *Cybernetic Letters,* 2011. ISSN 1802-3525

[13] MAGA, D., HARTANSKY, R.: *Numerical Solutions.* University of Defence, Brno, 2006, 174 p, ISBN 80-7231-130-1.

[14] BAUER, P., VAN DUIJSEN, P.: Challenges and Advances in Simulation. *36th Annual IEEE Power Electronic Specialists.* New York: IEEE, pp. 1030–1036, 2005, ISSN: 0275-9306. ISBN 0780390334; 978-078039033-1.

[15] KRUPKA, Z., STEFEK, A.: Method Cooperation in Optimization Process. *Mechatronika 2010.* Trencin: TnUAD, pp. 21–23, 2010, ISBN 978-808075461-7.

[16] SVOBODA, J., SIMAK, B., ZEMAN, T., HRAD, J.: *Educational Activities of the CTU Prague Team in EMC and Support of the EMC Development Problems in the Czech Republic.* 13th Intern. Wroclaw Symposium and Exhibition on Electromagnetic Compatibility. Wroclaw : Institute of Telecommunications, pp. 208–211, 1996, ISBN 83-901999-4-7.

[17] FIELDLER, M.: On Inverting Partitioned Matrices. *Math. J.* 13 (88), pp. 574–586, 1963.

Matus Pleva – Jozef Juhar *

BUILDING OF BROADCAST NEWS DATABASE FOR EVALUATION OF THE AUTOMATED SUBTITLING SERVICE

This paper describes the process of recording, annotation, correction and evaluation of the new Broadcast News (BN) speech database named KEMT-BN2, as an extension for our older KEMT-BN1 and COST-278 databases used for automatic Slovak continuous speech recognition development. The database utilisation and statistics are presented. This database was prepared for evaluation of the automated BN transcription system, developed in our laboratory, which is mainly used for subtitle generation for recorded BN shows. The speech database is the key part of the acoustic models training for specific domains and also for speaker and anchor adapted models creation.

Keywords: *Broadcast news, segmentation, speech recognition, Transcriber.*

1. Introduction

The development of continuous speech recognition (CSR) systems in Slovak language expects a large amount of different language resources to be collected [1 and 2]. First of all, the speech database needs to be built, which is also the most expensive and demanding task [3]. The building of the textual database of Slovak texts for language modelling is also challenging [4 and 5] and could be done using modern crawling technologies and post-processing, morphological analysis etc. [6].

The KEMT-BN2 database campaign was carried out between 2009 and 2011. It consists of broadcast news (BN) shows from the first Slovak public service broadcaster television (STV1 – Jednotka). The transcription task was mainly realized by brigadiers, and then trained and evaluated by transcription specialist. The database is a follower of the KEMT-BN1 [7] and the Slovak part of the COST-278 [8] database realized in our laboratory [9] (recorder from TA3 news television).

The purpose of the specialized BN databases is to build and evaluate the automatic transcription system for BN shows [10]. This system should have special BN acoustic models for different types of speech in BN shows (F-conditions) [11], special acoustic models for anchors or speakers with high occurrence in the news (politicians, sportsmen, artists, etc.) [12, 13 and 14] and also a special language model from the BN domain [15 and 16]. To use these models in the special detection system for speakers, different types of speech etc. should be provided [17 and 18].

This paper describes the process of recording and collecting the audio materials of the BN shows. Next, the transcription process and the evaluation of the transcriptions are presented. Finally the statistics of the collected annotated data in the database is depicted and discussed in conclusions and future work.

2. Recording the shows

The broadcast news shows were recorded from DVB-T channel multiplex streaming the PS (program stream) data to the disk using Technisat Airstar PCI card [19] from testing broadcast on channel 25 in Kosice region before an official digitisation process. The MPEG2 Program Stream was captured with time reserve, but it was not truncated because a jingle detection algorithm based on Euclidian distance or DTW is planned to be developed and evaluated on this data later.

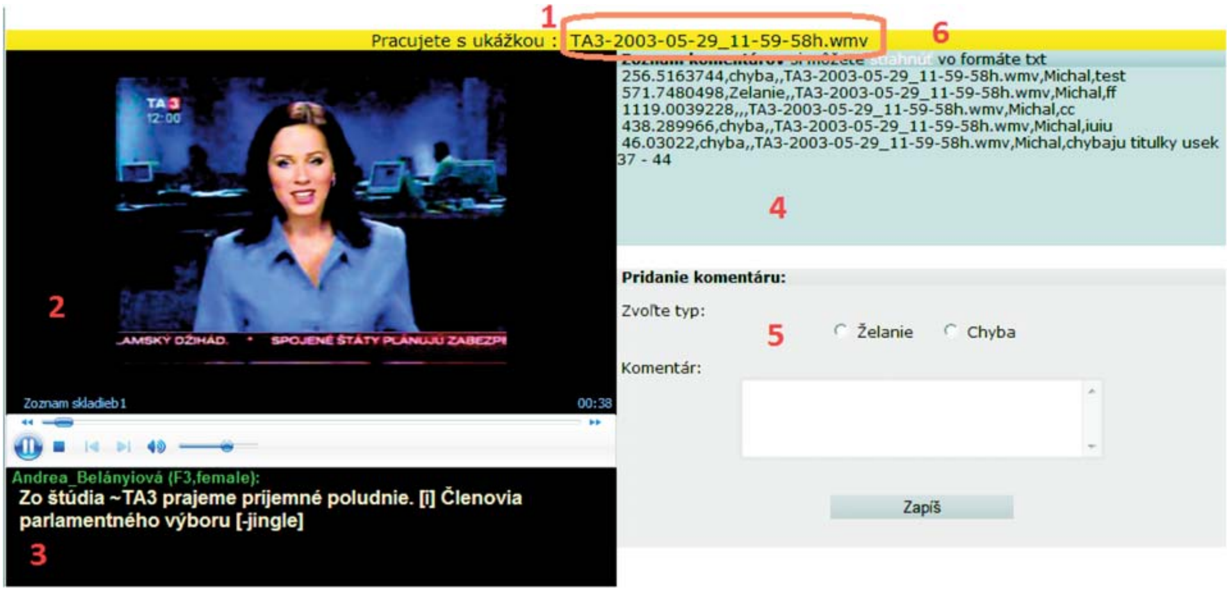
The audio subchannel is de-multiplexed from the stream using DGMPGDec DGIndex [20] GPL (GNU Public License) licensed software resulting in *.mp2* file (*48 kHz* stereo *128 kbits* CBR constant bitrate quality MPEG-1 Audio Layer 2 codec).

Next, the audio file needs to be converted to a format compatible with transcription software and delivered to annotators (the file size is also important). The used *Transcriber* software [21] mentioned in the next chapter has several bugs when using mp3 format (the time was not correlating with wav or video) so the *mp2* files (not compliant) were recompressed to *ogg* format (Ogg Vorbis *160 kbps* q5.0 mono) using the freeware foobar2000 tool [22] with *Vorbis* plugin. After that also a mono PCM *16 kHz wav* file was decompressed for database utilisation purposes.

The complete video recordings is planned to be converted also to a well supported video streaming format for using in web application for presenting the database with captions for the public

---

* **Matus Pleva, Jozef Juhar**
Department of Electronics and Multimedia Communications, Faculty of Electrical Engineering and Informatics, Technical University of Kosice, Slovakia, E-mail: Matus.Pleva@tuke.sk



Pracujete s ukážkou : TA3-2003-05-29_11-59-58h.wmv

Zoznam komentárov si môžete stiahnuť vo formáte txt

```
256.5163744,chyba,,TA3-2003-05-29_11-59-58h.wmv,Michal,test
571.7480498,Zelanie,,TA3-2003-05-29_11-59-58h.wmv,Michal,ff
1119.0039228,,,TA3-2003-05-29_11-59-58h.wmv,Michal,cc
438.289966,chyba,,TA3-2003-05-29_11-59-58h.wmv,Michal,iuiu
46.03022,chyba,,TA3-2003-05-29_11-59-58h.wmv,Michal,chybaju titulky usek
37 - 44
```

Pridanie komentáru:

Zvoľte typ: ○ Želanie ○ Chyba

Komentár:

Zapíš

Andrea_Belányiová (F3,female):
Zo štúdia ~TA3 prajeme príjemné poludnie. [i] Členovia parlamentného výboru [-jingle]

Fig. 1 Broadcast news shows (1/2) transcriptions (3) presentation on web interface (6) with ability to send an error report (5) to the administrator with automatic timestamp (4) of the paused video

(Fig. 1 – the COST278 TA3 part of the database on the web) [23]. The video recording is important when transcribing the speaker names (from captions in the video) and topics descriptions too.

3. Transcription of the speech and non-speech audio events

Transcription process consists of manual orthographic transcriptions of the whole audio recording using *Transcriber 1.5.1* tool – a free software under GPL license (Fig. 2) [21]. The annotation process follows the LDC (Linguistic Data Consortium) transcription conventions for HUB4 [24] (DARPA-sponsored Hub4 continuous speech recognition evaluation) extended using new rules for Slovak language and future use for lexical and language modelling. The native *xml* file format file is *.trs* file.

A) STM export

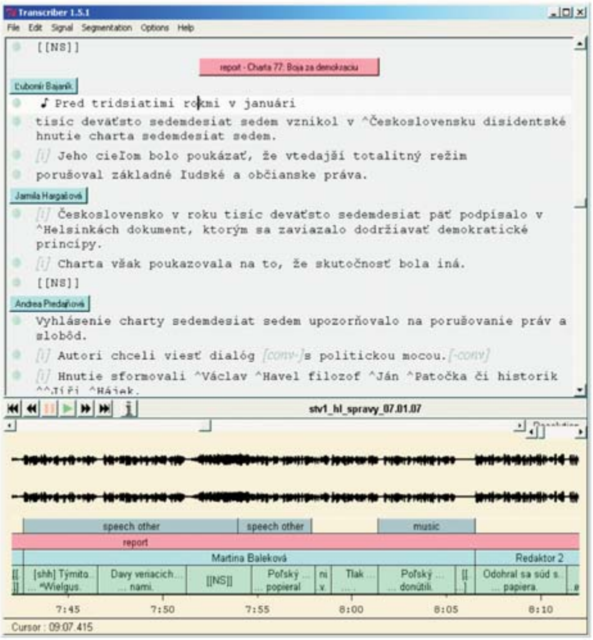
After completing the transcriptions the .stm (the NIST Scoring toolkit Sclite [25] – a more simple text file format exported from Transcriber) file is generated. The .stm file is the source format for next processing of the recordings, as segmentation and conversion to other speech database and online subtitles standards [26] which are suitable for using reference speech recognition training procedure described in [27]. We developed a special set of Perl scripts for conversion from *wav* and *stm* file pairs to the standardized SpeechDat database format for this purpose [28].

B) Transcriber modifications

The Transcriber toolkit was slightly modified for these purposes. The description of noise markers was translated and extended (the annotators have to enter the noise marker/tags only using menu – to avoid frequent typos in non-speech tags).

Next, also the conversion script for *stm* format export was modified to include all tags in resulted *stm* file (some of them were filtered).

Finally, the Slovak spellchecking feature was realized using free GPL licensed Aspell (http://aspell.net/) dictionary and modifying the corresponding spelling TCL/Tk script (which should send only the words to the dictionary – not tags) which was not



```
[[NS]]
♩ Pred tridsiatimi rokmi v januári
tisíc deväťsto sedemdesiat sedem vznikol v ^Československu disidentské
hnutie charta sedemdesiat sedem.
[i] Jeho cieľom bolo poukázať, že vtedajší totalitný režim
porušoval základné ľudské a občianske práva.
[i] Československo v roku tisíc deväťsto sedemdesiat päť podpísalo v
^Helsinkách dokument, ktorým sa zaviazalo dodržiavať demokratické
princípy.
[i] Charta však poukazovala na to, že skutočnosť bola iná.
[[NS]]
Vyhlásenie charty sedemdesiat sedem upozorňovalo na porušovanie práv a
slobôd.
[i] Autori chceli viesť dialóg [conv-]s politickou mocou.[-conv]
[i] Hnutie sformovali ^Václav ^Havel filozof ^Ján ^Patočka či historik
^^Jiří ^Hájek.
```

Fig. 2 Transcriber window with audio segments, visualized waveforms, transcribed text, speaker and topic tags able to work with Slovak symbols (CP1250 or UTF-8 encoded characters) in the text.

C) Segmentation and foreign languages

The speech utterances in the database have not been too long and every speaker inspiration event (breathing – tint) should be regarded as a potential breakpoint.

According to segmentation, the silence inside a speaker turns shorter than 0.5 seconds was not marked at all. Breakpoint in the middle was inserted when the pause in the speech utterance is between 0.5 and 1.5 seconds. When the pause was longer than 1.5 seconds, a special silence segment was inserted [7].

Foreign language utterances were marked with language event tags and should not be transcribed at all.

4. Database corrections

The database includes many typos, mistakes, misspelling and strange characters also after second annotator review of every transcription. The correction process is important because every wrong annotation could decrease the quality of the resulting acoustic or language models.

The process used in our laboratory for acoustic models training is very sensitive to every discontinuity in the database (refrec – Reference Recognizer from COST249) [27]. This process is also affected using the conversion scripts from file pairs (*wav + stm*) to SpeechDat format, including the mapping of the noise markers/tags and generating the phonetic lexicon.

There are more crucial points in the acoustic models (AM) training procedure described in [29]:

- Generating the *word level phonetic transcription* of all segments which will be used in the training procedure. Usually the script finished with error that *some word was not found in the phonetic lexicon*, generated during the conversion of the database.
  - Mainly it is a problem that some tags or non-word units passed to the training because the *tags mapping* (from huge set of noise tags to simple SpeechDat [sil] [spk] [sta] [int] tags [26]) *missed some new tag* (the developer needs to decide how to map it). Also typing errors (typos) are discovered during this stage like: missing character or mistyped character.

- Generating *initial monophone models*. Sometimes there is a problem that *for specific segment a proper label file is missing*.
  - This error is caused by *inconsistency of the two* filtering and index file generating *scripts*, when one script filters out a segment as not suitable for training (and do not include its labels in the master label file) but the script for generating the file-list of training segments decide that this segment could be used for training. The architecture of the used training procedure should be changed in the future to use only one set of filtering rules in both scripts and including them from specialized configuration file.

- Generating the *phone prototypes*. During this stage the developer sometimes discovers that an *unknown phoneme (or unwanted) is in the training* or some phoneme or noise *model* is missing.
  - This error is usually caused by a *non-Slovak word* (should not be annotated or should be marked with a special tag) or *filtering script error*, when some tag was filtered (during training set generation) like not important for AM training (lexical tags), but then we found out that also another important noise tag was filtered.

- Generating *tied triphone models* sometimes crash because some *phoneme is missing in the decision tree* or phonemes class definitions.
  - Sometimes we want to try a new phoneme set (reduced or more specific) for testing the impact of the precise phoneme definitions on the resulting system. During this stage sometimes the phoneme class definition should be changed or the *phoneme mapping is not properly defined* and should be corrected.

- Automatic *forced alignment* errors. When the *forced alignment procedure could not find a suitable automatic alignment* for the segment and its corresponding annotation the segment will be included to the outliers list and will be discarded from the training procedure.
  - Checking this outliers list and *reviewing the original file* and the corresponding *annotation*, the annotation should be corrected because there is usually some error in annotation (sometimes missing word or another word with similar meaning instead the right one – it is complicated to find this type of errors for the annotators because the brain is doing some automatic correction sometimes during monotonous work).

5. Database statistics

The database consists of 291 TV shows in 210 hours of material (including time reserve before and after). The total transcribed database includes *141 hours of annotated audio material* (1’169’832 words in 131’884 speech utterances) and the distribution of Focus conditions is depicted in Table 1 below.

The dictionary generated from this database consists of 95’376 Slovak words and 19’425 foreign words/names, noises, not correctly spelled words, partial/misspelled words or abbreviations.

The phonetic transcription (pronunciation lexicon) was generated using our developed Perl tool, and it is a very important element of the database. The phoneme description is based on SAMPA format [30] standard. Recently we found out that the phonetic transcription based on words for SpeechDat databases is not suitable for sentences, so we decide to change the training script to accept also whole sentences phonetic transcription for better inter-word phonetics.

Speech utterances distribution in the KEMT-BN2 database Table 1

Focus conditions of the utterances	
F0 – prepared speech in studio	73.46 h
F1 – spontaneous speech in studio	23.13 h
F2 – prepared telephone speech (reduced-bandwidth)	1.20 h
F3 – speech with music in background (SNR<10dB)	13.80 h
F4 – speech under degraded acoustical conditions	35.03 h
F5 – speech performed by non-native speaker	0.36 h
FX – combined conditions of types mentioned above	18.89 h

6. Conclusions

The collection of speech databases is the crucial problem when developing an automatic speech recognition engines for different domains and conditions. The broadcast news task is a very popular issue nowadays, because the government regulation specifies the minimal amount of shows with hidden subtitles for hearing impaired spectators.

The new KEMT-BN2 database brings a very important contribution to broadcast news processing. Not only for speech recognition but also for jingle detection, speech detection, speaker/anchor detection (anchor – hosting character in broadcast programs), segmentation, speaker clustering and different specialized noise modelling for domain specific tasks.

The KEMT-BN2 database has 3 times more data in every important parameter than the previous KEMT-BN1 and the Slovak part of COST-278 database together [7].

In the next period, we plan to use this database for building new acoustic models for broadcast news automatic continuous speech recognition, evaluate these models with previous versions (built on KEMT-BN1 and other databases) on new BN domain specific test set. Together with our colleagues we already prepared the new language model (LM) for BN task (adapted from huge universal LM used in our previous projects [4]). We plan also to implement the sentence level phonetic transcription process in the training script.

Acknowledgement

The research presented in this paper was supported by the Research & Development Operational Program funded by the ERDF (ITMS 26220220155) 50% and by 7th Framework Programme EU ICT project INDECT (FP7 - 218086) 50%.

References

[1] JUHAR, J., CIZMAR, A., RUSKO, M., TRNKA, M., ROZINAJ, G., JARINA, R.: Voice Operated Information System in Slovak, *Computing and Informatics,* vol. 26 (6), pp. 577–603, 2007.

[2] NOUZA, J., SILOVSKY, J., ZDANSKY, J., CERVA, P., KROUL, M., CHALOUPKA, J.: Czech-to-Slovak Adapted Broadcast News Transcription System, *Proc. of INTERSPEECH 2008*, pp. 2683–2686, 2008.

[3] PLEVA, M., JUHAR, J., CIZMAR, A.: About Development and Evaluation of Multilingual Database for Automatic Broadcast News Transcription Systems, *Acta Electrotechnica et Informatica (AeI),* vol. 4 (2), pp. 56–59, 2004.

[4] STAS, J., HLADEK, D., PLEVA, M., JUHAR, J.: Slovak Language Model from Internet Text Data, *Lecture Notes in Computer Science*, Vol. 6456 LNCS, pp. 340–346, 2011.

[5] PROCHAZKA, V., POLLAK, P., ZDANSKY, J., NOUZA, J.: Performance of Czech Speech Recognition with Language Models Created from Public Resources, *Radioengineering,* Vol. 20 (4), pp. 1002–1008, 2011.

[6] HLADEK, D., STAS, J.: Text mining and processing for corpora creation in Slovak language, *Journal of Computer Science and Control Systems,* vol. 3 (1), pp. 65–68, 2010.

[7] PLEVA, M., JUHAR, J., CIZMAR, A.: Slovak Broadcast News Speech Corpus for Automatic Speech Recognition, *Proceedings of RTT 2007 conference,* Zilina, p. 4, 2007.

[8] VANDECATSEYE, A. et al.: The COST278 pan-European Broadcast News Database, *Proc. of LREC 2004*, vol. 6, May 2004, Lisbon, pp. 873–876, 2004.

[9] PLEVA, M.: Building European Broadcast News Database, *Proc. of 4. Doktorandska konferencia a SVOS TU v Kosiciach – SCYR 2004*, Kosice, pp. 85–86, 2004.

[10] PLEVA, M., CIZMAR, A., JUHAR, J., ONDAS, S., MIRILOVIC, M.: Towards Slovak Broadcast News Automatic Recording and Transcribing Service, *Lecture Notes in Computer Science: Verbal and Nonverbal Features of Human-Human and Human-Machine Interaction,* vol. 5042 LNCS, p. 158–168, 2008.

[11] JARINA, R., KUBA, M.: Speech Recognition Using Hidden Markov Model with Low Redundancy in the Observation Space, *Komunikacie (Communications)*, Vol. 6 (4), pp. 17–21, 2004.

[12] NOUZA, J. et al.: Making Czech Historical Radio Archive Accessible and Searchable for Wide Public, *Journal of Multimedia*, Vol. 7 (2), pp. 159–169, 2012.

[13] HRIC, M., CHMULIK, M., JARINA, R.: Comparison of Selected Classification Methods in Automatic Speaker Identification, *Komunikacie (Communications)*, Vol. 13 (4), pp. 20–24, 2011.

[14] CERVA, P., NOUZA, J., SILOVSKY, J.: Study on Cross-lingual Adaptation of a Czech LVCSR System Towards Slovak, *Lecture Notes in Computer Science.* Vol. 6800 LNCS, pp. 81–87, 2011.

[15] NOUZA, J., SILOVSKY, J.: Adapting Lexical and Language Models for Transcription of Highly Spontaneous Spoken Czech, *Lecture Notes in Computer Science,* vol. 6231 LNAI, pp. 377–384, 2010.

[16] JUHAR, J., STAS, J., HLADEK, D.: Recent Progress in Development of Language Model for Slovak Large Vocabulary Continuous Speech Recognition, *New Technologies: Trends, Innovations and Research,* Rijeka: InTech, pp. 261–276, 2012.

[17] PLEVA, M., JUHAR, J., CIZMAR, A.: Speech Detection in the Broadcast News Processing, *Proc. of DSP-MCOM 2005.* Kosice, pp. 84–85, 2005.

[18] VAVREK, J.: Audio Content Classification using SVM Binary Decision Trees, *Proc. of SCYR 2012: 12th Scientific Conference of Young Researchers,* May, Herlany, pp. 80–83, 2012.

[19] http://www.technisat.com does not present the end-of-life product, see spec.: http://www.digitalnow.com.au/product_pages/airstar2.html

[20] http://neuron2.net/dgmpgdec/dgmpgdec.html - developer site

[21] http://trans.sourceforge.net Transcriber 1.5.1 developer site

[22] http://www.foobar2000.org/ - developer site

[23] PLEVA, M., JUHAR, J., CIZMAR, A.: Multimedia Database Management for Annotators of the Metadata Content, *Proc. of RTT 2009*, Praha : CVUT, p. 3, 2009.

[24] http://www.ldc.upenn.edu/Projects/Corpus_Cookbook/transcription/broadcast_speech/english/index.html – LDC recommendation

[25] NIST SCLITE scoring toolkit: http://www.itl.nist.gov/iad/mig/tools/

[26] POLLAK, P. et al.: SpeechDat(E)-Eastern Speech Databases, *Proc. of LREC 2000, XLDB satellite workshop,* Athens, Greece, pp. 20–25, 2000.

[27] ZGANK, A. et al.: The COST 278 Initiative – Crosslingual Speech Recognition with Large Telephone Database, *Proc. of LREC 2004*, Lisbon, May 2004, pp. 2107–2110.

[28] PLEVA, M.: Automatic Processing of Speech Data in Multimedia Databases (Automaticke Spracovanie Recovych Dat v Multimedialnych Databazach), *PhD Thesis (in Slovak)*, FEI TU of Kosice, p. 93, 2009.

[29] LINDBERG, B. et al.: A Noise Robust Multilingual Reference Recogniser Based on Speechdat (II), *Proc. of INTERSPEECH 2000*, Beijing, China, October 16–20, 2000, pp. 370–373, 2000.

[30] IVANECKY, J., NABELKOVA, M.: Phonetic transcription SAMPA and Slovak language (in Slovak), *Jazykovedny casopis,* vol. 53, pp. 81–95, 2002.

Jiri Vodrazka – Petr Jares – Boris Simak *

MORE ACCURATE MODELLING OF DIGITAL SUBSCRIBER LINES WITH INHOMOGENEITY

The digital subscriber lines are still base of access network to provide Internet connectivity to the public. The transmission capacity is continually increased and performance of digital subscriber lines must be estimated for network provider use. This paper describes two ways for correction of digital subscriber line modelling. The resulting bitrates are calculated with respect to the length of the line and the number of disturbance sources. The article brings new ideas for simulations and calculations of transmission function for line with inhomogeneity and calculations new value of code gain.

Keywords: *Access network, model of twisted pair, inhomogeneity, digital subscriber lines.*

1. Introduction

In May 2011, installation of VDSL2 (Very-high-speed digital subscriber line) was started by Telefonica Czech Republic. The spectral profile 998ADE17 was chosen for VDSL2 lines operating in the access network owned in the Czech Republic [1]. In a year of operation, practical experience with the operation of VDSL is already available and we proceeded to correct our simulation model for digital subscriber lines [2]. The VDSL lines are widely used for the transmission video-streams of IPTV (Internet Protocol Television). The required bandwidth and quality of video stream on higher layers of communication model is researched in many works, for example [3]. This article is focused to transmission directly on the physical medium and the physical layer. Two major problems were identified with precision modelling of digital subscriber lines on the physical layer.

Firstly, transmission function of the line must be accurate. A real telecommunication line can be simplified as a homogeneous line with evenly distributed electrical parameters. As the frequency grows, the inhomogeneities along the line have more influence.

Secondly, code gain of the modem must be updated. The code gain is depended on a vendor and its error correction and equalization algorithms implementation.

Both problems are described and analysed in this article and practices are recommended to address them. Another big problem of metallic lines is crosstalk, which was addressed in many articles, such as [4] and [5].

2. Types inhomogeneities

The telecommunications practice uses the following classification for specific types of inhomogeneities that can appear in subscriber lines in local access networks; they have originated from the analogy technology:

- Inhomogeneities of type 1 – they include transitions between different core diameters, transitions between different materials (copper–aluminium), and transitions between different cable arrangements (pairs–quads).
- Inhomogeneities of type 2 – representing “broken” quads, pairs, pair with bad insulation, interruptions.
- Inhomogeneities of type 3 – they originate from disobedience of rules for connecting groups (bundles) and layers (positions).

This classification is, however, insufficient for digital subscriber lines that occupy the frequency bands up to tens of MHz, and therefore it is desirable to introduce the classification according to Table 1. The authors have further divided the individual types into subtypes that are identified by lower-case letters.

It is particularly important to introduce the inhomogeneities of Type 0 that are characterized by inhomogeneity of the cable section itself resulting from imperfect manufacturing and installing, and also from unmatched characteristic impedances of terminal devices and the line section in the operating frequency range (compromise termination) due to extremely wide band.

Type 1 inhomogeneities are caused by the branched access network topology and they must be taken into account in real oper-

---

* **Jiri Vodrazka, Petr Jares, Boris Simak**
Department of Telecommunication Engineering, Faculty of Electrical Engineering, Czech Technical University in Prague, Czech Republic,
E-mail: vodrazka@feld.cvut.cz

ation. Type 2 inhomogeneities, on the other hand, have the nature of impedance faults, and therefore they should be eliminated by manufacturing and installing; if they occur, the affected cable elements cannot be used for transmission. Type 3 inhomogeneities do not introduce more serious problem than type 0a and their consequences include mainly crosstalk. It is not possible to state that these consequences are just negative, since a change of layer causes changes in disturbance nature along the line, which may influence the even crosstalk conditions inside the cable tree in a positive way. Type 4 has been also introduced that includes parasite open-ended (4a) as well as functional (4b) taps on a line. The open-ended taps should be eliminated (if possible), while the functional taps should be correctly terminated. Type 5 is linked to additional elements in the access network, such as patch panels, distribution frames, splitters, etc.

Types of inhomogeneities Table 1

Type	Description	Consequences
0	Cable Inhomogeneities	
0a	Manufacturing and installation imperfections	Partial reflections, irregular ripple of the characteristics
0b	Compromise line termination	Partial reflections, regular ripple of the characteristics
1	Line Inhomogeneities	
1a	Cable type change	Partial reflections, regular ripple of the characteristics
1b	Cable construction change	
2	Impedance defects	
2a	Mismatching of cable elements	Major damage of line function with rapid growth of attenuation
2b	Insulation leakage	
2c	Conductor interruption	
3	Position change	Partial reflections, irregular ripple of the characteristics
4	Taps	
4a	Open-ended taps	Strong reflections and local maximums of attenuation
4b	Functional taps	Reflections and local maximums of attenuation
5	Added components	
5a	Patch cords and cables on MDF	Add attenuation and partial reflections
5b	Distribution frames and cross-connects	Add attenuation and partial reflections
5c	Splitters	Impact on parameters for low frequency

3. Model of line with inhomogeneities

The transmission function of a line is influenced by many factors, and the first option is to model it according to [6] as an environment with multipath propagation:

$$H(f) = \sum_{i=1}^{N} g_i \cdot e^{-\alpha(f) \cdot l_i} \cdot e^{-j2\pi \cdot f \cdot T_i} \tag{1}$$

where g_i is the path weight, T_i is the delay of a path with length l_i, N is number of paths and α is specific attenuation.

Another modelling option is to use the telegraph equations, in this case not with constant coefficients, but with variable ones, i.e. primary parameters RLCG depending on the position within the line. Simplified model can be used for practical applications, which is based on real conditions in access networks – the inhomogeneities are not simulated continuously, but discretely, as a limited number of impedance mismatches on the line. In such a case we can model the line as a cascade of individual sections with different parameters and lengths.

If we want to determine the transmission function of a line composed of several sections with different parameters and containing also taps or other elements, we can express it from the product of matrices describing the individual cascaded sections. We should use the cascade parameters of a two-port network ABCD [7]. The resulting matrix of the cascade will be given by the product of the sectional matrices:

$$A = A_1 \cdot A_2 \dots A_n = \begin{bmatrix} a(f) & b(f) \\ c(f) & d(f) \end{bmatrix} \tag{2}$$

where A_1 to A_n are matrices describing the cascaded elements and $a(f)$ to $d(f)$ are the resulting parameters of the entire cascade.

Similar procedure can be applied also to modelling of lines with inhomogeneity. The model will be formed by matrices describing homogeneous sections of various lengths. Between each two homogeneous sections there will be inserted inhomogeneities of different nature. The entire cascade will be terminated by a homogeneous section. There are two possibilities of defining the inhomogeneities: either using the impedance defects as combinations of serial and parallel impedances (suitable for less serious defects), or using electrically short sections with different secondary parameters; also, a combination of both methods can be used (suitable for strong inhomogeneities).

Then we can describe a section of a homogeneous line with length l_i, characteristic impedance Z_{ci} and propagation coefficient γ_i using the following matrix:

$$A_{Li} = \begin{bmatrix} \left(\cosh\left(\gamma_i(f) \cdot l_i\right)\right) & \left(Z_{Ci}(f) \cdot \sinh\left(\gamma_i(f) \cdot l_i\right)\right) \\ \left(\dfrac{\sinh\left(\gamma_i(f) \cdot l_i\right)}{Z_{Ci}(f)}\right) & \left(\cosh\left(\gamma_i(f) \cdot l_i\right)\right) \end{bmatrix} \tag{3}$$

Impedance Z_S that is in series between the input and output contacts will be expressed using the matrix (4); it can be used if there is an inhomogeneity caused by increased specific resistance or inductance (e.g. contact resistance in cross-connect and distribution frames).

$$A_S = \begin{bmatrix} 1 & Z_S \\ 0 & 1 \end{bmatrix} \qquad (4)$$

Impedance Z_p that is in parallel with the common input and output contacts will be expressed using the matrix (5); it can be used if there is an inhomogeneity caused by increased specific leakage or capacitance (e.g. capacitance of cross-connect and distribution frames).

$$A_p = \begin{bmatrix} 1 & 0 \\ \left(\dfrac{1}{Z_p}\right) & 1 \end{bmatrix} \qquad (5)$$

The resulting matrix will be calculated as a product of odd number of partial matrices $(1+2n)$ as follows:

$$A = A_{L1} \cdot \prod_{i=2}^{n} \left(A_{Di} \cdot A_{Li} \right) \qquad (6)$$

where A_{Li} are matrices describing homogeneous sections (Lines) and A_{Di} are matrices modelling the inhomogeneities (Defects), but with different secondary parameters Z_{ci} and γ_i than in the matrix A_{Li} (3). The number of matrices depends on the required nature of inhomogeneities. For practical use, modelling of inhomogeneity by electrically short sections of lines with different secondary parameters can be recommended. The formula (6) can then be rewritten as a product of matrices, as follows:

$$A = A(\gamma_1, Z_{c1}, l_{L1}) \cdot \prod_{i=2}^{n} \left[A(\gamma_{Di}, Z_{cDi}, l_{Di}) \cdot A(\gamma_{Li}, Z_{cLi}, l_{Li}) \right] \qquad (7)$$

where l_{Li} are the homogeneous sections lengths and l_{Di} are the lengths of sections modelling the inhomogeneity.

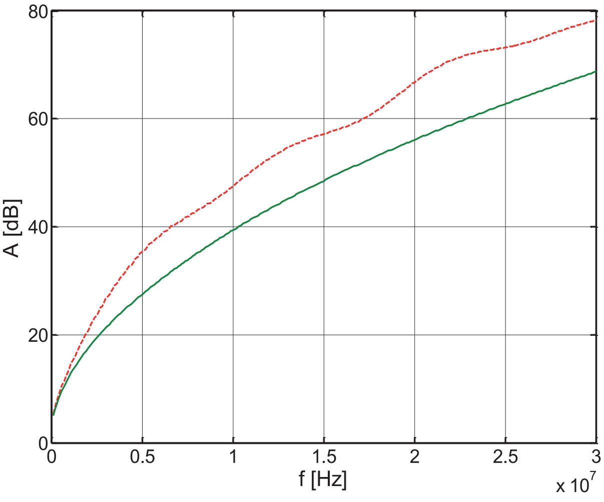


Fig. 1 Attenuation of a homogeneous line (green) and of a line with inhomogeneity and added components (red)

In order to approach the pseudo-random nature of inhomogeneity originating by cable manufacturing and installing, it is advisable to choose the lengths so that they are not integer multiples of each other. An example of modelling a line with inhomogeneity and added components (e.g. capacitance of cross-connect and distribution frames, patch cords, connection from main distribution frame to DSLAM) is shown in Figs. 1 and 2 where the value of a 500m section is compared to that of a homogeneous line in a TCEPKPFLE 75×4×0.4 local cable.

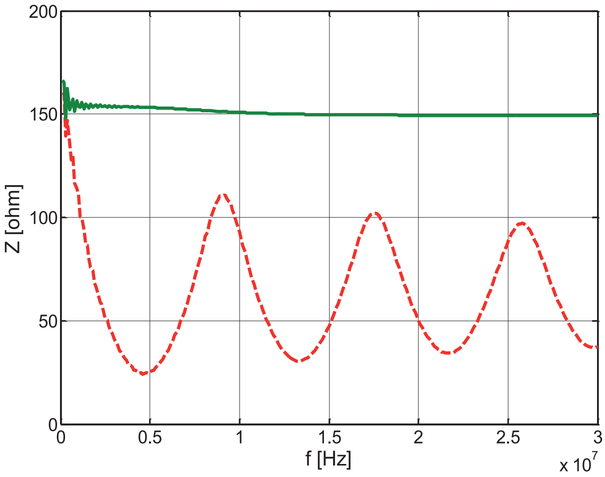


Fig. 2 Module of input impedance for a homogeneous line (green) and for a line with inhomogeneity and added components (red)

4. Improvement code gain calculation

Mentioned more precise calculations for line transmission parameters were implemented into Simulator xDSL program [2]. The simulation program is designed to perform calculations within the spectral compatibility sphere in a metallic access network and calculations concerning a transmission performance for several types of digital subscriber lines.

The simulation program allows performing simulation of all lines xDSL. For each transmission technology it is possible to set up an additional parameter. For example, for VDSL2 it is necessary to set up frequency band and PSD mask (Power Spectral Density). The transmission environment is modelled with the use of the noise profiles A, B, C, D, defined by ITU-T for each technology. However, it is also possible to use one's own combination of different transmission technologies.

For more precise calculations of transmission performance in Simulator xDSL program, there was a gain code correction applied. The code gain (CG) of modem depends on how the subscriber's data are protected during the transmission. What is more, a value of code gain affects bits allocation in each sub-channel of DMT

(Discrete Multi-tone) modulation. Number of allocated bits for i tone (channel) is determined by (8) [5]:

$$b_i = \log_2\left(1 + \frac{S_i}{N_i \cdot k_b}\right) \qquad (8)$$

where, b_i is the number of allocated bits, S_i is a signal power, N_i is a noise power. Constant k_b is given by (9):

$$k_b = \frac{SG \cdot NM}{CG} \qquad (9)$$

where SG is Shannon Gap 9.55 for BER $= 10^{-7}$, NM is Noise Margin (usually for xDSL NM $=$ 6 dB). Figure 3 shows a dependence of signal to noise ratio (SNR) on a number of channel. Blue curve is again measured, while red comes from theoretical simulation.

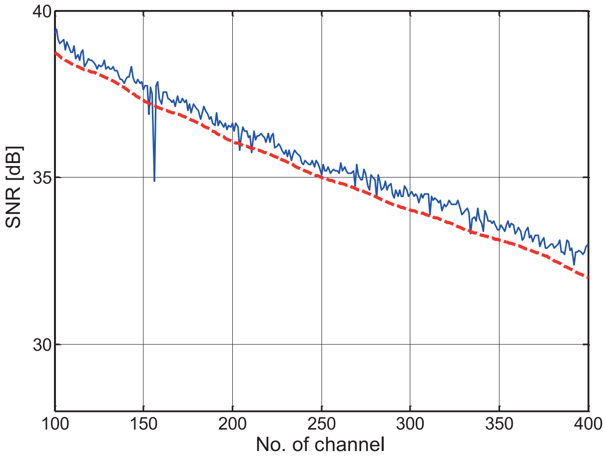


Fig. 3 Dependence of SNR on order of DMT tones

An example of code gain calculated from measured values of noise margin is shown in Fig. 4. The noise margin value varies because of the user modem that is trying to provide a maximum bitrate, which depends on the actual SNR value for a given tone. The noise margin is not really constant, but varies between the minimum and maximum values around the desired value of 6 dB. The calculated code gain from noise margin and SNR values is around 6 dB.

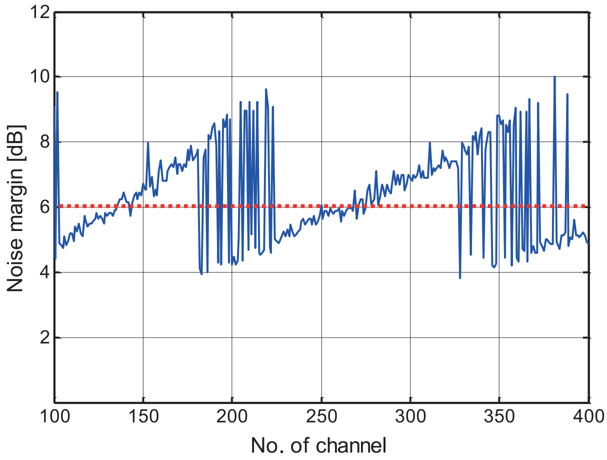


Fig. 4 Dependence of noise margin (blue) and code gain (red) on order of DMT tones

5. Conclusion

Two major problems were identified with precision modelling of digital subscriber lines: inhomogeneity along the line and code gain vendor dependence. The updated classification and inhomogeneity was used to refine model lines. In result, code gain value differs, depending on a vendor and algorithms implemented into the end user modem. For that reason, improvements can be applied only to a specific modem. Despite these conditions, Simulator xDSL allows to perform calculations according to specific conditions and obtain exact values for transmission performance of xDSL system. Table 2 summarizes the results before correction and after correction of the downstream bitrate modelled values for both the intended effects. Parameters of local cable TCEPKPFLE 75×4×0.4, spectral profile 998ADE17 to 17 MHz and noise model B (ETSI) were used for all cases. The noise model B is calculate from typical crosstalk from other 15 lines with VDSL2 modems and mix of SHDSL and ADSL modems in reference cable with 50 pairs.

Bitrates of VDSL2 profile 998ADE17 comparison for 1 km length of line Table 2

Description	Downstream bitrate
Effect of inhomogeneity:	
Without inhomogeneity	16 Mbps
With inhomogeneity	12 Mbps
Effect of code gain correction:	
Original value of code gain = 3 dB	12 Mbps
New value of code gain = 6 dB	16 Mbps

Acknowledgement

This work was supported by the Grant of the Technology Agency of the Czech Republic, No. TA02011015, "Research and development of a new communication system with multi-channel approach and multi-layer co-operation for industrial applications", and was researched in cooperation with CERTICON.

References

[1] VODRAZKA, J., JARES, P., HRAD, J.: *Optimal Position of External Node for Very-high-bitrate Digital Subscriber Line,* 34th Intern. Conference on Telecommunications and Signal Processing (TSP). Budapest, AUG 18-20, 2011. Pages: 175–177. ISBN: 978-1-4577-1411-5

[2] JARES, P., VODRAZKA, J.: *xDSL Simulator*, on-line http://matlab.fel.cvut.cz/en

[3] SMIESKO, J., URAMOVA, J.: Access Node Dimensioning for IPTV Traffic Using Effective Bandwidth, *Communications - Scientific Letters of the University of Zilina,* No. 2, 2012. ISSN 1335-4205.

[4] LAFATA, P., VODRAZKA, J.: *Simulations and Statistical Evaluations of FEXT Crosstalk in xDSL Systems Using Metallic Cable Constructional Arrangement, TSP - 31st Intern.* Conference Telecommunications and Signal Processing [CD-ROM]. Budapest: Asszisztencia Szervezo Kft., 2008, ISBN 978-963-06-5487-6.

[5] STARR, T., SORBARA, M., CIOFFI, J. M., SILVERMAN, P.: *DSL Advances.* Upper Saddle River, USA: Prentice Hall, 576 p., 2002. ISBN 0-13-093810-6.

[6] HRASNICA, H.: *Broadband Powerline Communications Networks,* John Wiley & Sons : Chichester, 2004. ISBN 0-470-85741-2.

[7] VODRAZKA, J., HRAD, J.: *Modeling of a Subscriber Line with Inhomogeneity,* 32nd Intern. Conference on Telecommunications and Signal Processing (TSP 2009), Dunakiliti, pp. 79–83, 2009. ISBN 978-963-06-7716-5.

Jan Papaj – Lubomir Dobos – Anton Cizmar *

FUNCTIONALITY VALIDATION OF THE NEW QOS AND SECURITY INTEGRATION MODEL FOR MANET

In this article, the simulation testbed of the new designed model for cooperation between QoS and security is introduced. The field of research about QoS and security provide dynamically developed and important areas of the research and up to now no protocol or mechanisms have been designed to this process. The cooperation is based on using a Cross Layer Design model (CLD) and modified Security Service Vector (SSV). A main idea of testbed is the indication how could security and QoS be integrated as a parameter to the MANET via this model. We also show how this model can affect the functionalities of the whole MANET.

***Keywords:** MANET, QoS, Security, Security Service Vector.*

1. Introduction

In this article the simulation testbed of the new QoS and security integration model for mobile ad-hoc network (MANET) is introduced. The MANET is characterized as a set of mobile nodes and devices connected to each other through wireless. A basic feature of abstinence is very solid infrastructure nodes which are able to create and maintain a connection between them. Research in the field of MANET is oriented to following areas: QoS, Security and Cross Layer Design.

Quality of Service (QoS) is a dynamically evolving field of research dealing with different areas. In literature, research is oriented to *QoS models*, *QoS resource reservation signalling, QoS routing and QoS Medium Access Control (MAC)* [1]. There are many definitions of the term QoS [2]. From network point of view, the QoS is a guarantee provided by the network to satisfy a set of predetermined service performance constraints for the user in terms of the end-to-end delay statistics, available bandwidth, probability of packet loss, and so on [2] and [3].

Security is also a dynamically evolving field. The major role of security mechanism is the ensuring of services: Confidentiality, Authentication, Availability, Integrity and Non-repudiation. Research communities in MANET's solve problems of *Secure Routing, Key Management and Intrusion Detection System* [4].

The main idea of Cross Layer Design (CLD) is to increase the performance and adaptability of MANET. Cross-layering tries to share information among different layers, which can be used as inputs for algorithms, for decision processes, for computations, and adaptations [5]. Based on research the three main CLD architectures have been designed: *Direct communication between layers, Shared database architectures, Heap architectures or Completely novel approaches.*

Optimized Network Engineering Tools (OPNET) Modeler [6] is a useful tool for simulation and design of different type networks and related technologies. OPNET modeler provides a wide range of different types of simulation models of devices, communication technologies, wired and mobile network protocols and technologies. OPNET Modeler combines the C/C++ programming language and PROTO-C language. PROTO-C is characterized as a finite state transition diagrams (FSM) to implement different simulation models and supporting technologies. OPNET modeler provides fairly straight-forward tools to simulation study using standard models. Developing new models or modifying existing ones is a challenging and often frustrating undertaking [6].

2. Main motivation for the validation of the new model in OPNET modeler

The main ideas of the integration model are to provide QoS and security mechanisms at the same time, and that user or services have the possibilities to interact with system via CLD. Integration provides also the proper functionalities of QoS and security mechanisms. Our model enables cooperation between QoS and security mechanism and also between users and system by modified security service vector. The goal of the simulation testbed in OPNET modeler is to show that the proposed model does not affect the activity of the network and greatly contributes to the integration between QoS and security.

---

* **Jan Papaj, Lubomir Dobos, Anton Cizmar**
Department of Electronics and Multimedia Communications, Faculty of Electrical Engineering and Informatics, Technical University in Kosice, Slovakia, E-mail: jan.papaj@tuke.sk

3. New QoS and security integration model for MANET

In this section the new QoS and Security integration model for MANET is introduced.

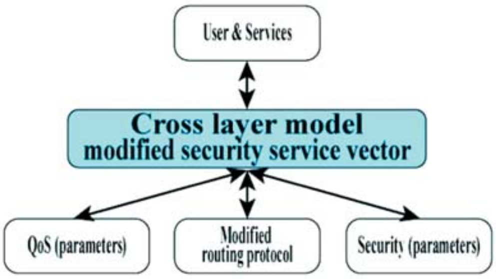


Fig. 1 The new QoS and security integration model with cross layer interface and modified security service vector

The model provides the ability of mutual cooperation between QoS and security related systems and algorithms by a new cross layer interface (model) with modified security service vector. Our new model is displayed in Fig. 1. The model includes all components for interactions between the user and system to integrate security as one parameter. ***A block cross layer model and modified security service*** vector is a block where the Cross layer model (CLD) is used to create interactive environment between users and the system and also provides support interactions between the routing protocol and modified security service vector (SSV). ***The block QoS (parameters)*** represents a mechanism for delivering of QoS in MANET network environments. It defines and specifies the QoS parameters necessary to provide the required services. ***The block Security (parameters)*** represents a mechanism to provide security-related services and also defines the necessary parameters used to process services providing. ***The block User&Service*** enables the interaction between the user and the system. ***The block Modified routing protocol*** represents the routing protocol with implemented modified SSV algorithm for selecting the optimal way based on user defined requirements (QoS and security).

The main idea of the model is based in the current use of QoS and security mechanisms as well as the interaction between users and systems in order to provide the type of service. Integration itself is necessary for proper functioning of both mechanisms in terms of QoS and security. The model also provides users with the possibilities to specify requirements for new services in MANET.

The main part of the model is modified SSV and cross layer model or interface (CLD) [7]. The modified SSV is based on the security service vector designed especially for wired IP networks [8]. A main idea of the modified SSV is shown in Fig. 2. Modification of the SSV is defined into two parts: user and system. The user part deals with a process of collecting the relevant data about requested services. In our case, these data are created by QoS and security parameters. Parameters can represent different QoS and security parameters or mechanisms for providing QoS and security processes [9].

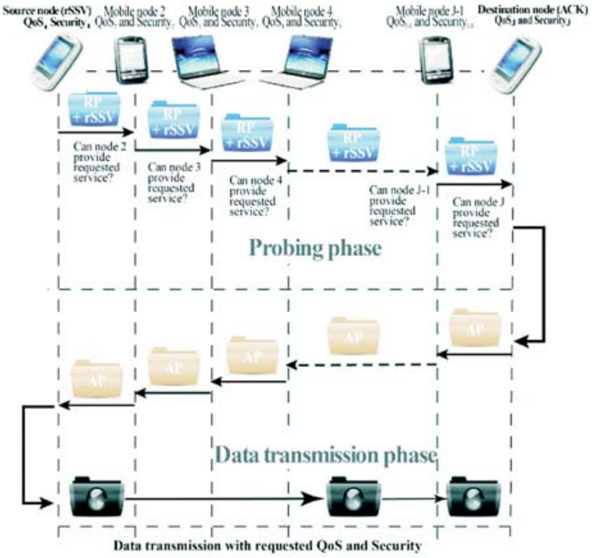


Fig. 2 The modified SSV model in MANET

In this model, users can specify the required parameters and the using of this approach can actively affect the system (routing) processes. The system part of our modification represents the new method of processing collected data and also deals with routing processes of the routing protocol. Each MANET node includes the algorithm to process the routing packet (RP). Algorithms analyze the routing information stored in RP and analyze the information about requested parameters, QoS and security (rSSV) [7], [9]. CLD is used to process bidirectional collection of relevant data from application or network layer by the modified SSV. These data are used for routing which uses the DSR routing protocol. This model enables cooperation between QoS and security mechanisms by the new designed cross layer model and modified SSV.

The main idea of CLD is depicted in Fig. 3. In the case of a source node, the user defines the SSV attributes via CLI inter-

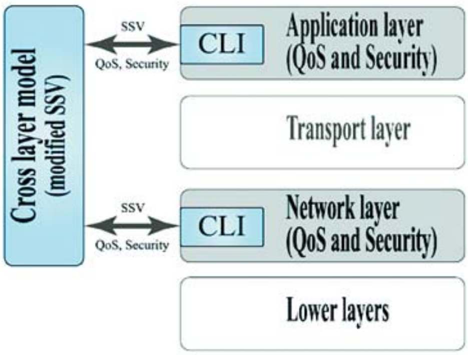


Fig. 3 The new cross layer interface for MANET

face located on the application layer and CLD interface sends these data to the network layer where they are stored to the modified route cache [9]. In the case of a routing node, the CLI analyses the incoming packet and reads information about the SSV stored in the packet. If the modified route caches do not include the information about security and QoS from the application layer, the CLD interface activates recollecting process of these data from the application layer. In the case of a destination node, the CLD collects data about requested QoS and security from the routing packet and from the modified route cache.

4. Simulation setup and experiments

The main ideas of the testbeds were to verify possibilities of implementing and testing a new designed model in MANET terminals. All behaviour of the proposed model was simulated in OPNET modeler 16.0. The three simulation scenarios were used to evaluate effectiveness of integrating a new model with CLD and modified SSV (Table 1):

- ***The model where the nodes used the routing protocol DSR without modified SSV and CLD (DSR)*** – data are transmitted by each layer without CLD and modified SSV.
- ***The model where the nodes used the modified routing protocol with the implemented modified SSV (DSR+SSV)*** – data are transmitted by each layer without CLD with implemented modified SSV.
- ***The model where the nodes used the modified routing protocol with the implemented modified SSV and CLD (DSR+SSV_CLD)*** – data are transmitted by new CLD interface and modified SSV is implemented.

The three parameters were used to check functionality of the proposed model with the integrated modified SSV and CLD, namely:

- ***Time to processing*** – means the process time necessary to process all operations of data on nodes. Time is measured from the time of creation, from the application layer or from arrival on the physical layer.
- ***Delay of MANET*** – represents the value of the average end-to-end delay measured from the network layer on the source node, where the MANET packet is created, to the delivery of the packet to the destination node. This parameter also includes the processing time which is necessary for all SSV processes during source-destination transport.
- ***Total packet processing delay*** – this parameter represents the average delay in MANET networks from sending a packet to the adoption of the packet on the IP layer of the target node. The parameter does not reflect the time needed to processing information SSV.

OPNET modeler simulation parameters Table 1

Parameters	Values
Number of nodes	10 – 100
Simulation areas	500×500 m^2 (for 10–50 nodes) 1000×1000m^2 (for 60–100 nodes)
Simulation period	1000 s
Number of simulations	100
Number of collected values per simulation	1000
Speed of nodes	0–2 m.s^{-1}
Mobility model	Random waypoint model
Space model	Free space

In the first experiment, the processing time on source, routing and destination nodes was analyzed. This parameter represents the time required for processing and creating the modified packet. The term processing, in the case of the source node, means the time since the creation of the requirements to transmit data at the application layer to the time of the packet departure from the physical layer. It is the time of data arrival at the physical layer and of return to the physical layer in the case of a routing node. It is also the time that is necessary to perform required activities in the case of a destination node.

In the second experiment, the parameter delay of MANET was simulated and analyzed. This parameter gives the information how long it takes to deliver a MANET packet from the source to destination nodes. It provides good information about the time necessary to deliver a MANET packet from the source to destination nodes and includes the time for CLD and for carrying out the activities of the modified SSV.

The third experiment shows how the processing of modified SSV and CLD affect the total packet processing delay that represents the time necessary to transmission of the packet from the source to a destination node through the MANET network. This transmission depends on dynamic source routing protocol (DSR) with the implemented modified SSV and CLD algorithms.

5. Simulation Results

The results of monitoring the processing time depending on the type of node (first experiment) are shown in Fig. 4. The processing time parameter is monitored on the source, routing and destination nodes. Based on the collected results, we can conclude that in all cases (source, routing, destination nodes), the better results were obtained for model DSR+SSV_CLD in comparison with DSR and DSR+SSV. For example, in the case of the source node, the implementation of the DSR+SSV increased the value of the processing time by 11.70 % as compared with DSR. When DSR + SSV_CLD were applied, the values of the processing time were reduced by 19.09 % as against the DSR and by 24.89 % when compared with DSR+SSV.

When the DSR+SSV was implemented into the destination node the processing time increased by 12.37 % as compared with DSR under the same conditions. In the case of the destination

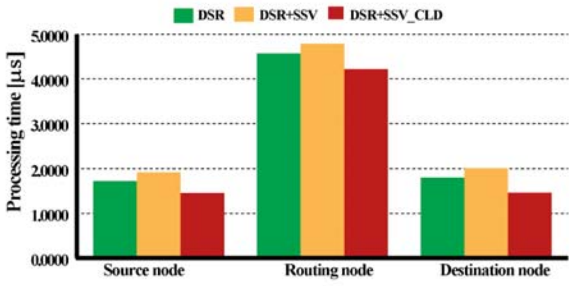


Fig. 4 Processing time for source, routing and destination nodes

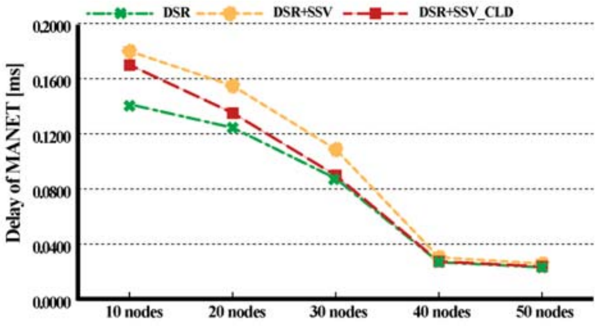


Fig. 5 Delay of MANET [ms] analysis depending on the number of nodes for 10–50 nodes

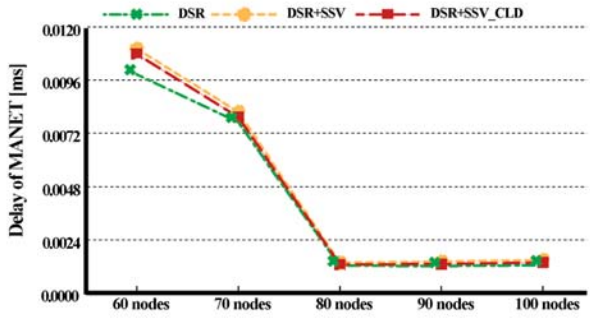


Fig. 6 Delay of MANET [ms] analysis depending on the number of nodes for 60 -100 nodes

node, implementation of the DSR+SSV_SSV represents a decrease of processing time when compared with DSR (by 16.80 %) and DSR_SSV (by 19.15 %). The implementation of CLD into the MANET model (DSR+SSV_CLD) provides reduction of the processing time compared with DSR model and model DSR+SSV.

The final results from second experiment, in which the delay of the MANET network was analysed and studied, are displayed in Fig. 5 and 6. The results showed that the delays of the MANET were increased after implementation of DSR+SSV_CLD by 20.21 % as compared with the standard DSR and by 27.24 % when using DSR+SSV. However, applying DSR + SSV_CLD reduced the delay by 5.52 % as compared with DSR+SSV. The lowest increase of the delay value was achieved for 50 nodes – the average delay after applying DSR + SSV_CLD increased only by 2.41 % and by 11.36 % using DSR + SSV as compared with the standard DSR protocol and applying DSR + SSV increased by 11.36 %.

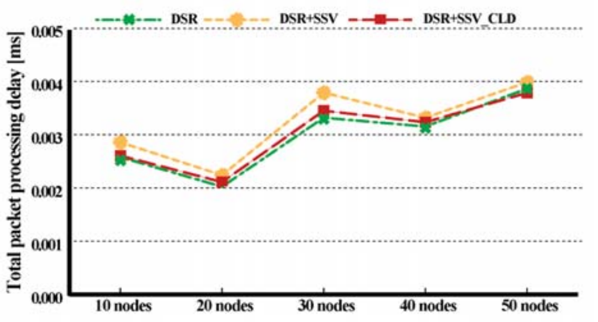


Fig. 7 Total packet processing delay of MANET [ms] analysis depending on the number of nodes for 10-50 nodes

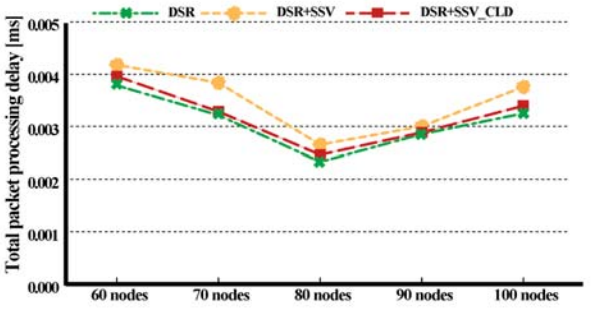


Fig. 8 Total packet processing delay of MANET [ms] analysis depending on the number of nodes for 60-100 nodes

In the last experiment the total packet processing delay was analyzed. All the obtained results are displayed in Fig. 7 and 8. For example, when DSR+SSV_CLD were applied in MANET consisting of 50 nodes, the total packet processing delay was reduced by 3.13 % against the standard DSR protocol and the application of the SSV + DSR meant achieving an increase (about 3.16 %).

6. Conclusions

The new model has been designed for cooperation between QoS and security mechanisms in MANET. In order to test the functionalities of the new model, which includes the modified SSV and CLD, the simulation testbed analysis of three MANET models were presented. The models DSR, DSR+SSV, DSR+SSV_CLD were analysed. All simulations were simulated in OPNET modeler. Based on the collected results (see section Simulation Results) of processing time, delay and total packet processing delay, we can conclude that our designed integration model with CLD and modified SSV for MANET represents a insignificant increase of these parameters in MANET. This model is also useful for the processing of QoS and security integration for MANET.

Acknowledgments

This work was supported by the EU ICT Project INDECT (FP7-218086) (30 %) and by the Ministry of Education of Slovak Republic under research VEGA 1/0386/12 (30 %) and by Research & Development Operational Program funded by the ERDF under the ITMS project code 26220220155 (40 %).

References

[1] KARIMI, M.: *Quality of Service (QoS) Provisioning in Mobile Ad-Hoc Networks (MANETs),* Mobile Ad-Hoc Networks, Protocol Design, InTech, ISBN: 978-953-307-402-3, 2011..

[2] MONFORT, J-Y., ADLER, K.P.F., GIERLICH, H. W., POMY, J.: Multimedia Quality as Perceived by the User, *Communication – Scientific Letters of the University of Zilina,* vol. 4, 2008, ISSN 1335-4205, pp. 5–10.

[3] BRIDA, P., BENIKOVSKY: Radio MaF framework for GSM Positioning, *Communication – Scientific Letters of the University of Zilina,* vol. 04, 2009, ISSN 1335-4205, pp. 24–27.

[4] DJENOURI D., KHELLADI L., BADACHE, A.N.: A Survey of Security Issues in Mobile Ad Hoc and Sensor Networks, *Communications Surveys & Tutorials,* IEEE, 2005, vol. 7, No. 4, pp. 2–28.

[5] SRIVSTAVA V., MOTANI M.: *The Road Ahead for Cross-layer Design,* Proc. of 2005 2nd Intern. Conference on Broadband Networks, IEEE, 2005, pp. 551–556.

[6] OPNET Modeler Simulation Software, http://www.opnet.com

[7] PAPAJ, J., DOBOS, L., CIZMAR, A.: *New Cross Layer Model to Integration QoS and Security as a one Parameter in Mobile ad hoc Network,* MCSS 2010: Multimedia Communications, IEEE intern. conference Services and Security, Krakow, ISBN 978-83-88309-92-2, 2010, 6-7, May, pp.1–6.

[8] SAKARINDR, P., ANSARI, N., ROJAS-CESSA, R., PAPAVASSILIOU, S.: *Security-enhanced Quality of Service (SQoS) Networks,* IEEE Sarnoff Symposium on Advanced in Wired and Wireless Communications, 2005, April, pp. 129–132.

[9] PAPAJ, J., CIZMAR, A., DOBOS, L.: Implementation of the New Integration Model of Security and QoS for MANET to the OPNET. *Communications in Computer and Information Science,* 149 CCIS, ISSN: 1865-0929, 2011, pp. 310–316.

Tomas Macha – Radko Krkos – Vit Novotny *

PROPOSAL OF LOAD AWARE ROUTING FOR OSPF ROUTING PROTOCOL

This paper deals with route cost assignment using one of the most commonly used routing protocols - Open Shortest Path First (OSPF). The limitation of this protocol is that the algorithm does not take actual link load into consideration. This paper proposes using an alternate costing strategy to avoid this problem. The alternate strategy uses an integer value representing the load of a link as the additive component of metric. Since the link load is a dynamic variable, Exponential Weighted Moving Average is used to counter quick changes. This paper proposes a novel approach to calculate OSPF metrics, aimed to better distribute the load in network.

Keywords: *OSPF, routing, metric, load.*

1. Introduction

The Internet or internetworking communication systems are based on universal network-level interconnections. These interconnections enable delivery of data units from their source to their desired destination. Internetworking also deals with the complexity of various underlying communication technologies that today's heterogeneous networks comprise of. The IP (Internet Protocol) plays a central role in the Internet architecture, as it is responsible for appropriate routing of packets in the Internet [1].

Routing is one of the main properties of the Internet and enables establishment of robust and efficient networks. The data unit is sent directly to the destination, if the destination is on the same subnetwork as the source is. Otherwise, the data unit is sequentially being sent to the routing devices in the network that make decisions about directions so the data unit reaches desired destination [1]. Such devices, which interconnect subnet works and pass packets from one to the other, are called routers. The behaviour of routers depends on a routing protocol. A routing protocol defines a set of rules used by a router for communication with neighbouring routers. Routers between a source node and a destination node must choose and maintain paths to be used. The router must create routing tables and compute the best route to use [2].

Since OSPF (Open Shortest Path First) currently represents the most widely used routing protocol, any valuable improvement to keep pace with the rapidly changing Internet environment would be greatly appreciated. The limitation of this protocol is that its link cost calculation algorithm does not take actual link load into consideration. If there are two links with the same bandwidth and the link utilization of the first is very low and for the second link very high, OSPF assigns both links the same metric [2]. An explanatory example is in Fig. 1. This network consists of six routers; most of the routers are connected with 100 Mbps Ethernet links and their cost corresponds to a metric of 10, see Table 1. Technology of 10 Mbps Ethernet is used to connect router E with routers B and I. Metric assigned to links between these routers is 10. In this network, User 1 and User 3 communicate with the server. Based on path metrics calculated as a sum of metrics of all links on the path from the source to the destination the entire data traffic passes the link between routers D and E. With increasing users' demands on the server this link may become congested. Situation becomes even worse when User 2 also starts to communicate with the server. Congestions cause service performance degradation and in extreme cases even service disruption.

This paper brings a novel mechanism that also includes link load into link cost computation to improve routing in OSPF networks. This mechanism considers link load as an additive parameter for a final metric, which solves the problem of absence of traffic awareness and is aimed to balance the load among network links and routers, which in general decreases latency for transported data and overall improves network performance. This paper is a summary of on-going research presented in [3] and [4].

New Cost Adaptive OSPF (CA-OSPF) is proposed in [5]. However, this solution is not applicable for heavily loaded networks where the CA-OSPF cannot improve the network performance [5]. The recent works related to routing algorithms improvements [6], [7], [8] and [9] can be used to choose a path with specific bandwidth requirements. An implementation of QoS (Quality of Service) routing extensions for OSPF is proposed in [10]. Also many balance heuristic techniques were proposed to avoid con-

---

* **Tomas Macha, Radko Krkos, Vit Novotny**
Department of Telecommunications, Faculty of Electrical Engineering and Communication, Brno University of Technology, Czech Republic,
E-mail: tomas.macha@phd.feec.vutbr.cz

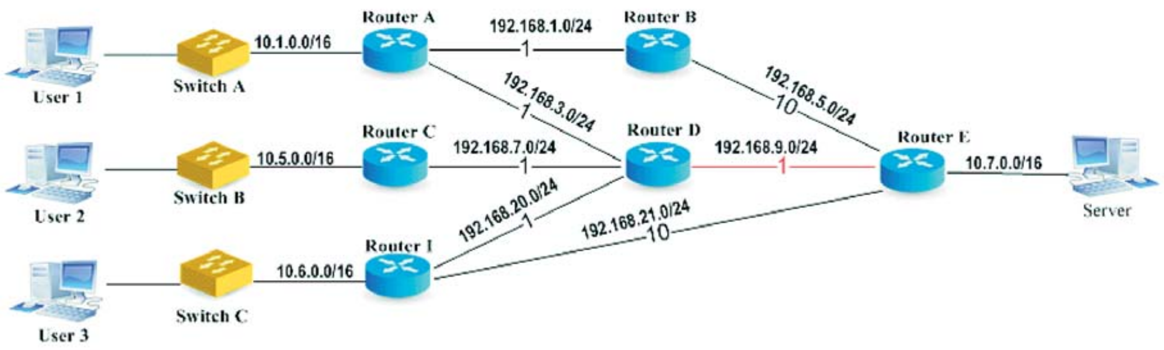


Fig. 1 OSPF network example

gestion and utilize links with low load [11], [12]. However all of these solutions focus only on the case of full link congestion and there is no rerouting of traffic until then. This load balancing gets to a stabilized state after a number of reached link congestions and so the convergence time is very long. The equilibrium is again broken when the volume of traffic changes and convergence process has to start afresh.

The main goal of this paper is to propose a novel and more efficient mechanism to improve routing in data networks using OSPF protocol. To achieve this purpose, a couple of extensions of current routing methods will be examined.

A) OSPF metric

Dijkstra's algorithm is used by OSPF for calculating the best path through complex networks. The shortest-path calculation is executed with a full knowledge of the topology.

RFC document [13] does not specify what the link metric should be. Most of the routers calculate the metric based on the configured bandwidth. The metric is then inversely proportional to the bandwidth; higher bandwidth means lower metric.

Metric of a path is the sum of the individual link metrics on the path to the destination. The formula used to calculate the metric of one interface is:

$$metric = \frac{reference\ bandwidth}{bandwidth} \tag{1}$$

Different manufacturers use different reference bandwidth, however, common reference bandwidth is 100 Mbps (100,000,000 bps). The reference bandwidth has to be equal for all routers in the same routing area. Only one metric can be assigned per interface. According to formula (1), the metric for Ethernet technology is 10 (10^8/10 000 000 bps). OSPF metrics for frequently used link types and based on formula (1) are summarized in Table 1. The interface metric is an unsigned 16-bit number; the total metric of a path is a 24-bit unsigned integer, as can be seen from the format of LSA-3, LSA-4, LSA-5, and LSA-7. Metrics are rounded down to the nearest integer. If its total metric overflows the 24-bit value, the path becomes unusable.

If the link bandwidth is 100 Mbps or higher, then based on formula (1) and commonly used reference bandwidth, the link metric will be 1, which is not optimal as there is no distinction between 100 Mbps and faster links. To solve this problem, it is possible to set new reference bandwidth for all routers in the entire network, using a router configuration command.

Default OSPF metrics Table 1

Technology	10^8/bps	Metric
Gigabit Ethernet	10^8/1 000 000 000 bps	1
Fast Ethernet	10^8/100 000 000 bps	1
Token Ring (16 Mbps)	10^8/16 000 000 bps	6
Ethernet (100 Mbps)	10^8/10 000 000 bps	10
E1	10^8/2 048 000 bps	48
T1	10^8/1 544 000 bps	64
64 kbps link	10^8/64 000 bps	1562
56 kbps link	10^8/56 000 bps	1785
9.6 kbps link	10^8/9 600 bps	10 416

2. Proposed method

Dijkstra's algorithm solves the shortest path problem by finding a path of minimal total cost between the nodes based on link metrics. Link metrics do not include the link load. Solution is that an integer value is added to the load of a link as an additive component of bandwidth based metric. This method builds on common metric calculations used by OSPF protocol. The computation of a novel metric for the novel method derives from formula (1):

$$novel\ metric = \frac{reference\ bandwidth}{bandwidth} + \\ + load - \left(metric_{next\ technology} - metric\right) \tag{2}$$

The component metric is a link metric computed according to the formula (1), metricnext technology is calculated using the same formula, but bandwidth of the next slower technology according to Table 1 is used. This means that if the link load approaches full utilization, the next slower link is used as an alternative path because the metric of the faster link is degraded linearly with load to match metric of 0% utilized link of next slower technology when the link is fully loaded. The situation is described in Fig. 2, where the metric for 100 Mbps link grows with link load to finally match the metric of Token Ring link.

Along with the method, we propose several improvements. The first one includes a moving average because of its averaging characteristics and the ability to smooth fluctuations. The second one considers real maximum link load utilization. The third one calculates link metric using highest link bandwidth present in OSPF area as reference bandwidth. And the fourth improvement consists of using next slower technology present in current topology instead of next slower technology according to Table 1, in link load impact calculation.

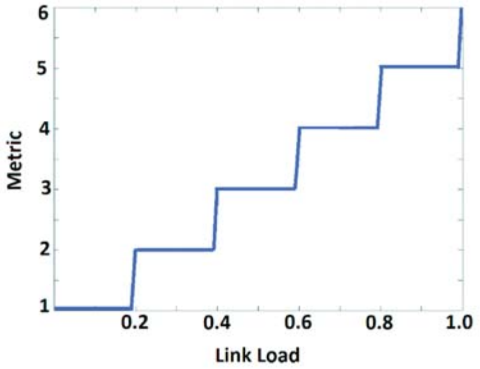


Fig. 2 Metric degradation of 100 Mbps link based on load

A change in the network topology is represented as a change in one or more of the LSAs. LSAs flooding process can noticeably degrade performance. The change causes new LSA flooding and calculation of new routing tables. However, even this amount of necessary traffic is undesirable.

A) Exponential weighted moving average application

Data collected in the time order can be averaged over several samples. Moving averages are often used in time series data analysis. Moving averages can be useful for measuring the changes in a trend, smoothing fluctuations and as a forecast for the next period. In this paper, the used raw data are link load observations of a router. Link load values represent data from which the moving average will be computed.

Exponential Weighted Moving Average (EWMA) [13] has become popular process-monitoring tool in a process-control field. Due to EWMA's robustness and ability to monitor a dynamic process with memory and drift, this approach is adopted for proposed method.

EWMA, sometimes also called Exponential Moving Average (EMA), provides a higher level of accuracy when compared with simple or weighted moving average. EWMA uses weight factors that decrease exponentially, which brings much more importance to the recent observation, but all the previous observations are included. Let c_n be an observation at time n, the explicit formulation of EWMA is:

$$A(n) = (1 - \lambda)^k A(n - k) + \lambda \sum_{i=0}^{k-1} (1 - \lambda)^i c_{k-i}\,, \tag{3}$$

where $A(n)$ is the current EWMA average, c_i is the value of link load at time i, k is the number of previous values used in calculation or the averaging depth and $A(n - k)$ is the EWMA average of k samples before. The starting value c_0 equals zero or is generally being set to the mean of former observations. The effect of the starting constant c0 decreases over time. The formula relies on an effective period for the exponential moving average called smoothing factor λ, $0 < \lambda \leq 1$, $\lambda = 2/(n+1)$, where n is the period of the moving average. The parameter λ determines the rate at which previous observations influence the calculation of EWMA. A value of $\lambda = 1$ implies that only the most recent observation influences the EWMA, thus higher value of λ gives more weight to recent data. The weight of each older observation is then decreased by a factor of $(1 - \lambda)$ [13]. EWMA can be computed recursively as [14]:

$$A(n) = \lambda c_n + (1 - \lambda) A(n-1)\,. \tag{4}$$

The recursive form of the EWMA calculation simplifies the formula (3) and decreases computing complexity. The recursive fashion (4) requires only two pieces of information to be processed.

Application of the EWMA technique is demonstrated using an example depicted in Fig. 3. The blue line stands for raw link load data, representative of current bandwidth usage. The red line stands for EWMA of the raw data. Formula (4) was used for EWMA computation. It can be seen that the series is smoothed,

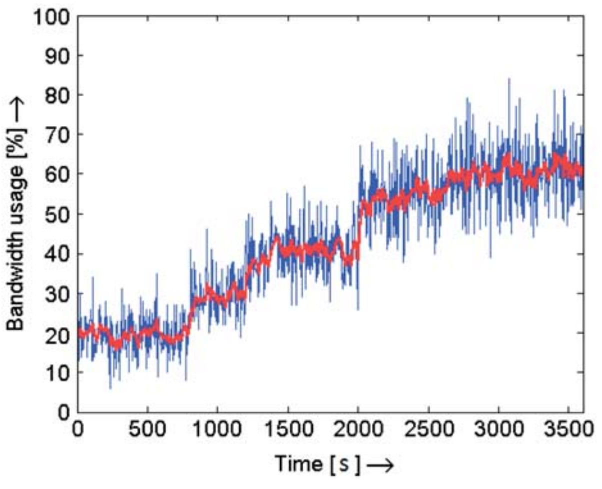


Fig. 3 Exponential weighted moving average example

with much lower variance. The decrease of the weights is an exponential function of the weighting factor λ. If the value of λ is close to zero, the moving average at any time n is significantly influenced by older values.

As mentioned before, the first improvement of the novel method is the application of EWMA technique onto the novel metric (3). The resulting formula can be expressed as:

$$novel\ metric_1 = metric + \mathrm{EWMA}(load) \cdot \cdot \left(metric_{next\ technology} - metric\right) \quad (5)$$

Figure 4 describes the impact of sampling frequency on EWMA convergence time. There are four similar graphs where the blue line stands for link load and the red line stands for EWMA of load. If the sampling is done every 30 minutes, it can be seen from the first graph that the convergence time is about 2 hours, which is not feasible. On the other hand, if the sampling is done every minute, the EWMA curve copies the load curve. So in this case, parameter λ should be increased.

B) Maximum link load utilization

This improvement considers the real maximum load of the links. This method is advantageous when used in heterogeneous networks or in networks supported by routers of different performance where it provides better scalability and network performance.

The best way how to explain this method is to show an example. For assembling the topology in Fig. 1, Linksys routers WRT54GL were used. The manufacturer claims 100 Mbps link bandwidth. However, according to the real measurement, the maximum bandwidth is only 32 Mbps. If 32 Mbps is the real maximum utilization instead 100 Mbps, the routers are not able to reach link utilization of 100% and the network performance is correspondingly lower. This problem prevails especially with less powerful routers. Therefore, as a solution, a maximum relative load parameter is applied. The resulting formula can be expressed as:

$$novel\ metric_2 = metric + load \cdot \cdot \left(\frac{metric_{next\ technology} - metric}{max\ relative\ load}\right). \quad (6)$$

When the network infrastructure is homogenous, the maximum relative load is approximately equal for all the links. But this approach still enables better scaling of metrics with link load. Maximum relative load is most useful when it is quite different from 100%, so the load effect on metric is spread evenly in the range of real bandwidth, which the link is capable of.

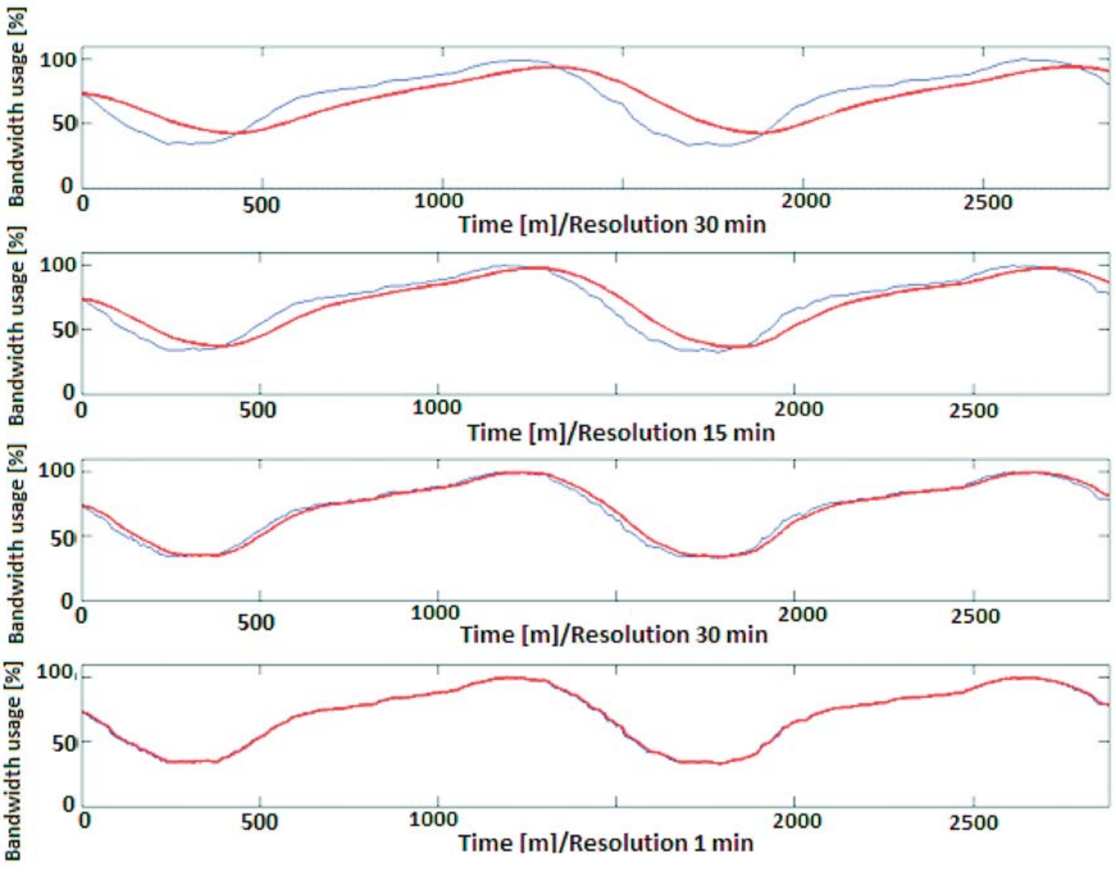


Fig. 4 Impact of sampling frequency on EWMA convergence time

C) Reference bandwidth value optimization

Original OSPF metric is calculated using reference bandwidth set to 108 [15]. This value is too small for today's link bandwidths and it causes that 100 Mbps, 1 Gbps and faster links have equal metric, see Table 1. Reference bandwidth value can be changed by network administrator, but when altered, it has to be manually set to the same value on all routers in order to shortest path selection to work properly. The solution would be an automatic fastest link seeking. This can be done in two ways. First possibility is to specify and distribute information about fastest link before the first topology creation. This is easier and often should be enough, because link bandwidths rarely change during operation. But it requires new LSA type or alteration of existing one to spread the highest encountered link bandwidth value. Second solution is to resolve it during topology creation. This is more complex, because the metrics have to be calculated before the maximum link bandwidth in network is known. Routers can calculate metrics using largest value already known to them, but also have to transmit used reference bandwidth, so other routers can correct their values if used reference values are mismatched. The advantage is possibly lower convergence time, especially for larger topologies.

Extension to this technique is an automatic detection of all present link bandwidths. This enables, in cooperation with improvement in section D), more accurate spreading of link load effect on metric because non present link bandwidths are left out and only the existing ones are considered when calculating next line speed metric, as shown by equation (6).

D) Matching technologies present in topology

Based on the formula (3), link load impact is multiplied by metric difference of current technology and the next slower technology according to Table 1. This is often sufficient, as multiple technologies used in one network mostly differ by one level. In some situations, when this is not the case, e.g. 100 Mbps lines are used for headquarters networking, but remote sites are interconnected by a leased line E1 link, this approach would yield no benefit.

The problem is, that in Table 1, there is another technology with metric value between the two mentioned. So in full load of 100 Mbps link, the metric for this link would be degraded to match the next slower technology, 10 Mbps link. The metric of an E1 link would still be much higher and therefore the link would not be used as an alternative. Solution is using the next technology present in current topology when calculating $metric_{next\ technology}$. This requires modifications to the OSPF topology discovery procedure to ensure that all routers know all technologies present in topology and therefore can change the link metrics based on load in a consistent and efficient manner. Such procedure is a question for further research beyond this paper.

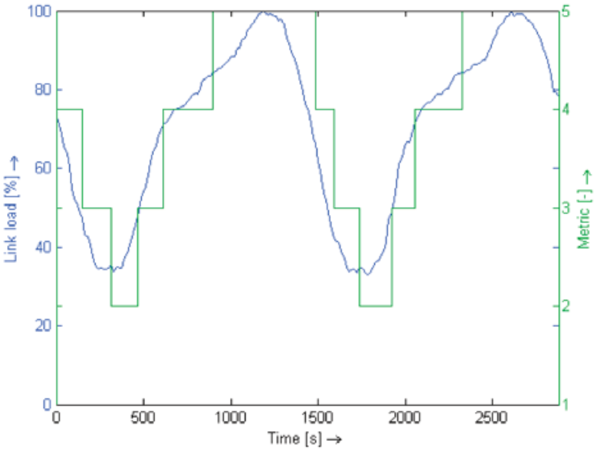


Fig. 5 Metric evolution based on link load for 100 Mbps link

3. Conclusion and future work

Open Shortest Path First is the most commonly used Internet routing protocol. The limitation of this protocol is that the algorithm does not take link load into consideration. This paper suggests using an alternate costing strategy to counter this problem. A novel method of OSPF metric calculation was proposed. An additive metric that considers current load of the link is added to the default metric. Also four methods to improve the OSPF metric calculation were proposed. All the proposals are summarized in the paper. Fig. 5 shows the metric evolution based on link load for 100 Mbps link. It can be seen that the novel method is functional and adapts the metric according to the link load. Proposed method may cause unwished characteristic, as frequent routing table changes can degrade the service and create downtimes. Also as IP protocol routing is dynamic, two-stage oscillation effect can occur. In this case, the traffic switches between two paths back and forth. Further work will be focused on solutions to these problems. As discussed in [3], further simulations will be done in ns-3 with direct code execution support, running altered Quagga OSPF daemon on the simulated nodes. Next, this altered OSPF implementation will be deployed on real devices to confirm the simulated results and carry out performance measurements.

Acknowledgment

The described research was performed in laboratories supported by the SIX project; the registration number CZ.1.05/2.1.00/03.0072, the operational program Research and Development for Innovation.

References

[1] CHAO, JONATHAN H., BIN LIU: *High Performance Switches and Routers.* Hoboken, New Jersey: John Wiley & Sons, 2007, 613 p., ISBN 978-0-470-05367-6.

[2] SPORTACK, Mark A.: *IP Routing Fundamentals.* Indianapolis: Cisco Systems, 1999, xvii, 510 p. ISBN 15-787-0071-X.

[3] MACHA, T., KRKOS, R., NOVOTNY, V.: OSPF Alternate Costing Strategy. In 2012 Intern. Conference on Telecommunication Systems, Modeling and Analysis (ICTSM2012), Prague: Czech Technical University in Prague, pp. 136-140, 2012, ISBN: 978-0-9820958-6-7.

[4] MACHA, T., KRKOS, R.: *A Novel Approach to OSPF Metric Calculation.* Proc. of Research in Telecommunication Technologies, 14th Intern. Conference, pp. 85-91, 2012, ISBN: 978-80-554-0570- 4.

[5] ZHOU, H., PAN, J., SHEN, P.: *Cost Adaptive OSPF.* Proc. of the Fifth Intern. Conference on Computational Intelligence and Multimedia Applications. 2003.

[6] GUERIN, R., ORDA, A., WILLIAMS, D.: *QoS Routing Mechanisms and OSPF Extensions.* Proc. of IEEE GLOBECOM, Phoenix. AZ. pp. 1903-1908, November 1997.

[7] WANG, Z., CROWCROFT, J.: Quality-of-service Routing for Supporting Multimedia Applications. IEEE *J. on Selected Areas of Communication,* vol. 14, pp. 1288-1234, September 1996.

[8] MA, Q., STEENKISTE, P., ZHANG, H.: *Routing High-bandwidth Traffic in Max-min Fair Share Network.* Proc. of ACM SIGCOMM. Stanford. CA. pp. 206-217. August 1996.

[9] WANG, J., NAHRSTEDT, K.: *Hop-by-hop Routing Algorithms for Premium Class Traffic in DiffServ Networks.* Proc. of IEEE INFOCOM, pp. 705-714, June 2002.

[10] APOSTOLOPOULOS, G., KAMAT, S. GUERIN, R.: *Implementation and Performance Measurements of QoS Routing Extensions to OSPF.* Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies, pp. 680-688, March 1999.

[11] FORTZ, B., THORUP. M.: Increasing Internet Capacity Using Local Search. *Computational Optimization and Applications,* pp. 13-48, 2004.

[12] FORTZ, B., REXFORD, J., THORUP. M.: Traffic Engineering with Traditional IP Routing Protocols. *IEEE Communication Magazine,* vol. 40, pp. 118-124, October 2002.

[13] HUNTER, J. S.: The Exponential Weighted Moving Average. *J. of Quality Technology*.

[14] ROBERTS, S. W.: *Control Chart Tests Based on Geometric Moving Averages.* Technometrics.

[15] MOY, J.: OSPF Version 2. RFC 2328. *Internet Engineering Task Force.* April 1998.

Petr Cika – Martin Zukal *

NOVEL WATERMARKING METHODS BASED ON FREQUENCY DOMAIN AND SINGULAR VALUE DECOMPOSITION

Nowadays, the copyright laws protecting the digital content are getting quite common. One of the important things is to protect the original or compressed images against illegal use. *Frequency domain watermarking methods are used in most cases in today's image- and video-watermarking algorithms. Especially Discrete Wavelet Transform (DWT) and Discrete Cosine Transform (DCT) are used in watermarking algorithms. This paper deals with development of two novel digital image watermarking methods that use DWT with Singular Value Decomposition (SVD) and DCT with SVD. The paper includes step by step process of watermark embedding and extraction and testing of the watermark robustness. The robustness was tested with the Checkmark, one of the watermark benchmarking tools. Experimental results show that both schemes are robust against a lot of attacks and have very high data hiding capacity.*

Keywords: *Singular value decomposition, frequency domain, robustness, watermark.*

1. Introduction

Nowadays, most of digital images are stored in the digital form for future utilization. The digital data protection against illegal use becomes still more important. In the field of digital images, the protection against copying or modifications is most important. One way how to protect digital images is to use digital watermarking techniques. There exist two general ideas: the visible watermark and the invisible watermark. Visible watermarks are mainly known for example as logos of television stations. The invisible watermarks are used especially for author rights authentication to the digital content (image, video). Generally speaking, invisible watermarks can be inserted in the spatial domain or in the transform domain [1]. There exist lots of methods which use the spatial domain for watermark insertion [2] and [3]. The most famous and firstly published method uses the least significant bits of the image data for watermark insertion. However, most of these methods are not robust enough.

This paper focuses on the invisible watermarks and techniques that use the transform domain of digital images for watermark insertion and are useful, for example, to secure the images in biomedical systems [4].

Discrete Cosine Transform (DCT), Discrete Fourier Transform (DFT), Discrete Wavelet Transform (DWT) and Walsh-Hadamard Transformation (WHT) are used to transform from spatial to frequency domain. These transforms are mostly used in many frequency domain watermarking methods [5], [6], [7] and [8].

This paper deals with watermarking methods that use DWT, DCT and SVD together. We developed two novel techniques that are robust to various types of attacks. The first technique is based on 2D-DCT and SVD transforms, the second one is based on 2D-DWT and SVD. Experimental results described in last sections show that both of these methods are suitable for image data protection against illegal use.

2. Common techniques used in watermarking systems

In most cases the frequency domain representation of the digital image is acquired using two dimensional DCT (2D-DCT) or two dimensional DWT (2D-DWT). The 2D-DCT is usually applicable on the square image data (matrix) of dimension 8×8 pixels. Similarly to other transforms, the 2D-DCT attempts to decorrelate the image data. Low frequencies are concentrated in the top left corner of the transformed matrix, high frequencies are concentrated in the bottom right corner. This transform is used in the well-known JPEG and MPEG compressions schemes. The second frequently used transform in image and video compressions schemes is the 2D-DWT. The 2D-DWT decomposes the original image into four bands, the watermark can be embedded in all frequencies [7].

SVD as a general linear algebra technique is used in a variety of applications, for example in watermarking. Modifying the singular value decomposition of the image is one of the most prevalent techniques in transform domain watermarking. SVD is described by the equation

---

* **Petr Cika, Martin Zukal**
Department of Telecommunications, Faculty of Electrical Engineering and Communication, Brno University of Technology, Czech Republic
E-mail: cika@feec.vutbr.cz

$$A = USV^T, \tag{1}$$

where the diagonal entries of S matrix are the singular values, the U matrix contains the left singular vectors and the V matrix contains the right singular vectors. Each of the new matrices has the same dimension as the original matrix A. The important fact is that each singular value s from the equation

$$S = \begin{pmatrix} s_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & s_i \end{pmatrix} \tag{2}$$

specifies the luminance value of image layers while the corresponding pair of singular vectors specifies the geometry of the image layer [9].

3. Quality parameters

The Peak Signal to Noise Ratio (*PSNR*) was used for testing the quality of the image with embedded watermark. The final *PSNR* value is expressed as

$$PSNR = 10\log_{10}\left(\frac{255^2}{MSE}\right) \tag{3}$$

In case of gray-scale images. *MSE* is Mean Square Error defined by the equation

$$MSE = \frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\left[x(m,n) - x'(m,n)\right]^2 \tag{4}$$

where M, N define the image size, x is the pixel value of the original image, and x' is the pixel value of the watermarked image.

The Normalized Cross Correlation (NCC) function was used for quality evaluation of the extracted watermark. NCC is defined by the equation

$$NCC = \frac{\sum_{i=0}^{I-1}\sum_{j=0}^{J-1} W_{ij} W'_{ij}}{\sum_{i=0}^{I-1}\sum_{j=0}^{J-1} \left[W_{ij}\right]^2} \tag{5}$$

where I, J define the size of the embedded watermark and W, W' define the original and extracted watermark bits.

4. Proposed watermarking methods

New watermarking schemes based on a method described in [5] will be presented in the following paragraphs. The first proposed method uses 2D-DWT and SVD, the second one uses 2D-DCT and SVD. Both methods divide the original image I into four sub-images before embedding.

The following steps describe the embedding process of the proposed watermarking methods (Fig. 1):

- The first step is to divide the gray-scale image I into 4 sub-images $I_1, ..., I_4$,
- 2D-DWT or 2D-DCT is applied on each sub-image $I_1, ..., I_4$ (only the first decomposition is used in case of 2D-DWT),
- in the next step the SVD is used:
  - *2D-DWT:* SVD is used to CA sub-band.
  - *2D-DCT:* All coefficients are zig-zag scanned at first, the SVD is used only for the first quarter of DCT coefficients.

The SVD is computed by the equation

$$A_k = U_k S_k V_{k,}^T, \quad k = 1, 2, 3, 4 \tag{6}$$

where k specifies the sub-image.

- The 2D-DWT/2D-DCT is applied on the whole watermark image W. Dimensions of the watermark image must be equal to quarter of the size of the original image I.
- Next the SVD is applied. The CA sub-band of the watermark W is processed in case of 2D-DWT. The first quarter of zig-zag scanned DCT coefficients is used in case of 2D-DCT:

$$A_W = U_W S_W V_W^T \tag{7}$$

- Singular values of each sub-band are modified by the singular values of the DWT transformed visual watermark:

$$s_i^{*k} = s_i^k + \alpha s_{wi}, \quad i = 1, ..., n, \tag{8}$$

where s_i^k are singular values of S^k, s_{wi} are singular values of S_w and α is the scale factor.

- Finally 4 sets of modified DWT/DCT coefficients are obtained by the equation

$$A^{*k} = U_A^k S_A^{*k} V_A^{kT}, \tag{9}$$

where $k = 1, 2, 3, 4$

and the watermarked image can be obtained using inverse 2D-DWT / 2-DCT.

Watermark extraction process (Fig. 2) can be summarized in the following steps:

- The watermarked image is divided into 4 sub-images $I_{w1}, ..., I_{w4}$.
- The 2D-DWT (only the first decomposition) or 2D-DCT is applied on each sub-image $I_{w1}, ..., I_{w4}$.
- The SVD is applied on each of the CA sub-bands in the case of 2D-DWT or on the first quarter of the zig-zag scanned coefficients in case of DCT.

$$A = USV^T \tag{10}$$

- The 2D-DWT/2D-DCT is applied on the entire visual watermark W.
- The singular values are extracted from each quadrant A_k, where $k = 1, 2, 3, 4$

$$s_{wei}^k = (s_i^{*k} - s_i^k)/\alpha, \quad i = 1,..., n, \tag{11}$$

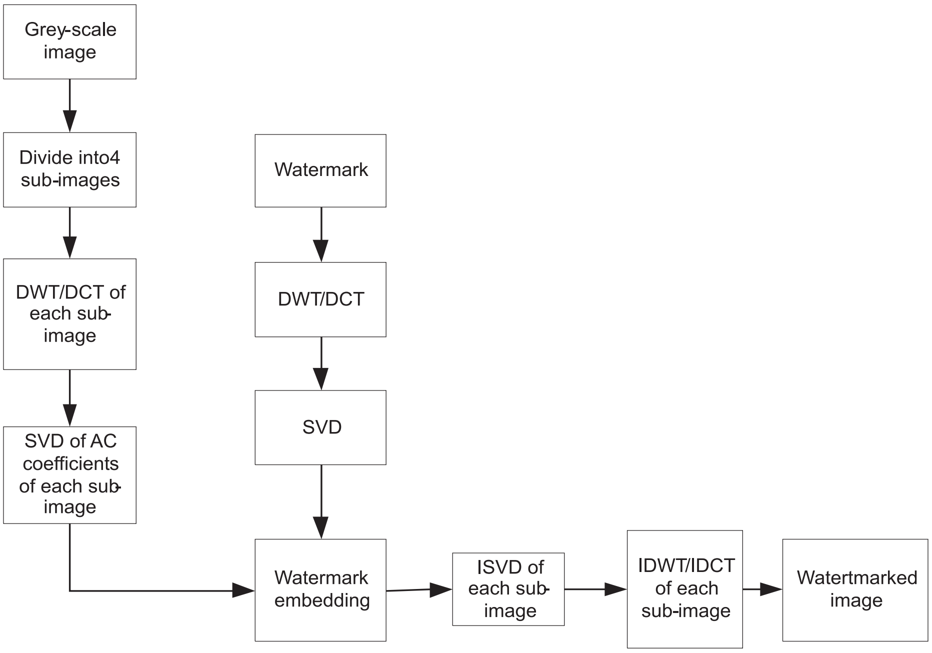


Fig. 1 Watermark embedding process

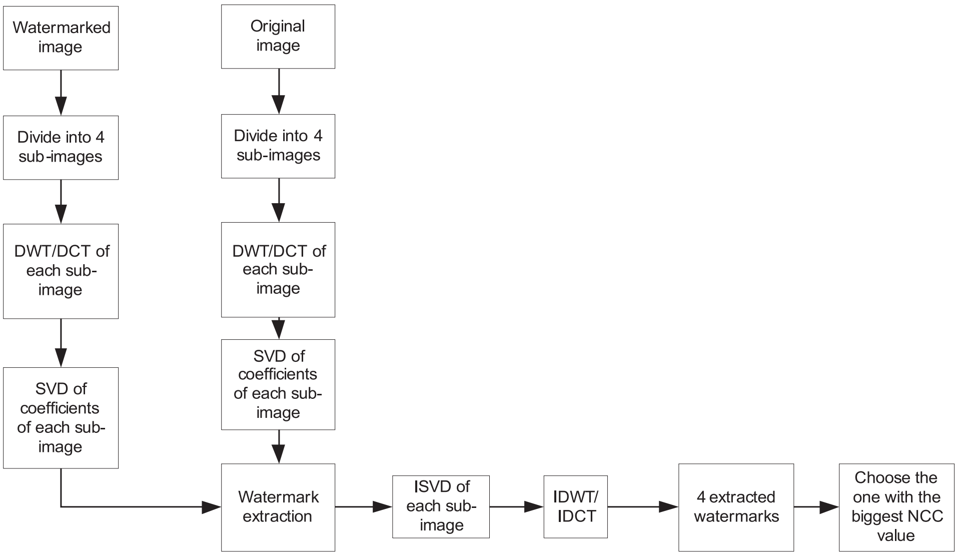


Fig. 2 Watermark extraction

where s^k_{wei} are singular values of extracted watermark s^k_{we} in k quadrant.

- The inverse 2D-DWT or 2D-DCT is applied on each set to construct four extracted watermark images.
- The best extracted watermark is chosen according to the computed NCC.

5. Evaluation of the proposed methods

We use the Checkmark benchmark tool to evaluate the proposed methods. The Checkmark easily carries out attacks on watermarked image and returns the results of watermark extraction success.

There exist 4 groups of attacks in the Checkmark: removal attacks, geometric attacks, cryptographic attacks and protocol

a) b) c)

Fig. 3 a) Lenna original, b) Lenna - watermarked image, c) Watermark

attacks. The image Lenna (Fig. 3a) as the original image I and the image of face (Fig. 3c) as the watermark image were used to test the proposed methods. The scale factor 0.16 was used as a tradeoff between watermark invisibility and robustness. The differences between the original and the watermarked image are shown in Figs. 3a) and 3b). The watermarked image (Fig. 3b) has notably different brightness than the original image (Fig. 3a).

We compared the novel method with a method that uses only DCT domain [10]. This method uses the luminance part of the image and DCT to embed the watermark in the image. Table 1 shows the results of watermark extraction after the selected attacks. 384 were carried out in total. However, Table 1 shows only the most common and most frequently used attacks.

6. Conclusion

The paper describes newly proposed image watermarking methods based on the wavelet transform and disrete cosine transform and singular value decomposition and. Our methods were compared with method described in [10]. The results introduced at the end of this paper show that our methods are highly robust to a lot of attacks.. The advantage of the proposed methods lies in high percentage of correctly extracted watermarks. The disadvantage of both methods is high computational complexity. We would like to improve these methods and add error correction codes to improve the reliability in the future.

Acknowledgment

This work was financially supported by the SIX Centre, OPVVI fund number CZ.1.05/2.1.00/03.0072 sponsored by the Ministry of Education, Youth and Sports of the Czech Republic and European Union.

Results of the test Table 1

Attacks	NCC - DCT	NCC - DWT	NCC-
Cropping 10%	0.9627	0.9178	
Cropping 20%	0.8863	0.8728	
Cropping 50%	0.7245	0.7357	
Cropping 75%	0.6301	0.6046	
Dithering	0.7840	0.6966	
Gaussian blur 3×3	0.9872	0.9899	0.8264
Gaussian blur 5×5	0.9870	0.9897	0.8264
JPEG q = 10	0.9974	0.9911	>0.7000
JPEG q = 40	0.9995	0.9982	>0.7000
JPEG q = 90	0.9997	0.9999	0.8667
Median filtering (3×3)	0.9813	0.9878	
Median filtering (4×4)	0.9930	0.9575	
Ratio change (0.8:1)	0.9890	0.9899	
Ratio change (1:0.8)	0.9847	0.9945	
Ratio change (1:1.2)	0.9903	0.9973	
Rotation 15°	0.8174	0.7234	
Rotation −1°	0.7232	0.7570	
Rotation 45°	0.9880	0.9600	
Scale (0.50x)	0.8169	0.9102	
Scale (0.75x)	0.9656	0.9806	
Scale (0.90x)	0.9722	0.9794	
Scale (1.10x)	0.9865	0.9906	
Scale (1.50x)	0.9969	0.9968	
Scale (2.00x)	0.9950	0.9958	
Sharpening	0.8316	0.8700	
Wiener filter (3×3)	0.9718	0.9773	
Jpeg 2000 (0.1 bit/pix)	0.9148	0.9110	
Jpeg 2000 (3.5 bit/pix)	0.9998	0.9997	
Jpeg 2000 (8.0 bit/pix)	1.0000	0.9997	

References

[1] COX, J., MILLER, M. L., BLOOM, J. A.: *Digital Watermarking,* Ed. London: Academic Press, 2002, ISBN 1-55860-714-5

[2] DARMSTAEDTER, J. DELAIGLE, J. QUISQUATER, M. B.: Low-cost Spatial Watermarking. *Computer&Graphics.* vol. 33, No. 4, 1998, p. 417–424, ISSN 0097-8493

[3] CIKA, P.: *New Watermarking Scheme for Colour Image. 12th Intern. Conference on Personal Wireless Communications (PWC 2007)*, Prague, , vol. 245, 2007

[4] HARGAS, L.: Application of Communication Systems in Biomedical Engineering, 2006, *Communications – Scientific Letters of the University of Zilina,* No. 1, p. 43–47, 2006, ISSN 1335-4205

[5] SVERDLOV, X., DEXTER, S., ESKICIOGLU, A. M.: Robust DCT/SVD Domain Image Watermarking for Copyright Protection: Embedding Data in all Frequencies, *EUSIPCO*, Antalya, 2005

[6] CIKA, P., ZUKAL, M.: *Efficiency of BCH Codes in Digital Image Watermarking,* 35th Intern. Conference on Telecommunications and Signal Processing (TSP), Prague, 2012

[7] CIKA, P., KOVAR, P., KOHOUTEK, M.: *Improved Digital Image Watermarking Method Based on DWT and SVD,* 31st International Conference on Telecommunications and Signal Processing (TSP 2008), Paradfurdo, 2008

[8] CIKA, P.: *Watermarking Scheme Based on Discrete Wavelet Transform and Error – Correction Codes,* 16th Intern. Workshop on Systems, Signals and Image Processing (IWSSIP 2009), Chalkida, pp. 250–253, 2009

[9] VOLOSHYNOVSKIY, P. S., MADUENO, M., MARCHAND-MAILLET, S., PUN, T.: *Second Generation Benchmarking and Application Oriented Evaluation, Information Hiding Workshop,* Pittsburgh, April 2001.

[10] MENG Z, YU P, YU G. *Copyright Protection for Digital Image Based on Joint DWT-DCT Transformation,* Wavelet Analysis and Pattern Recognition (ICWAPR), 2012 pp.11–14, 15–17 July 2012 doi: 10.1109/ICWAPR.2012.6294746.

Miroslav Voznak – Jiri Slachta – Filip Rezac – Jan Rozhon *

EXPLORATORY ANALYSIS OF VIRTUALIZATION TOOLS IN DELAY-SENSITIVE ENVIRONMENT

This article deals with an exploratory analysis of virtualized real-time applications. We compare selected virtualizations tools and present the achieved results in boxplots based on the performed exploratory analysis. The paper analyzes the impact of memory size and the number of processor cores on the delay itself and its variance in selected set of virtualization tools.

Keywords: *ANOVA, VMWare, KVM, VirtualBox, Real-time applications.*

1. Introduction

Nowadays virtualization presents a solution which can enhance reliability, security, portability and ease the maintenance of computer clusters [1] and [2]. In order to answer the question, how the virtualization techniques influence overall delay of transmission; we chose several free full virtualization tools and compared each other. We found out that the selected virtualizations have different characteristics and their median values do not match. There are many advantages of virtualization but besides them we also need to take into account the drawbacks of the technology, especially now when its use has become so prevalent. The main disadvantage is an overhead generated by the virtualization tool. The overhead causes virtual machines to be less efficient than the physical devices with similar attributes and reduces their performance. Further, the overhead can have a negative impact especially on real-time applications since it can cause long delays and increase the variance of delay between the individual packets [3], [4] and [5]. The aim of this article is to find out what impact different implementations of the virtualization technology have on the real-time traffic represented here by IP telephony as it is one of the most widely spread real-time technologies. The influence of the number of processor cores and memory size is also to be analyzed.

2. Virtualization tools

This part presents the three most common virtualization tools: VMware Player, Kernel-based Virtual Machine (KVM) and VirtualBox. Regarding KVM, high performance requirements on the instruction translation in the binary form resulted into a combination of experience gained in the different virtualization models. When the hardware-assisted virtualization emerged, a new kernel-module-based hypervisor started to be developed for the GNU/Linux platform. This hypervisor combines both high performance and versatile usability. By extending the Linux kernel with the KVM hypervisor, the advantages of the model which allows for maintaining each single virtual machine as a standard Linux process [6] can be exploited. The second VirtualBox is a multiplatform virtualization tool designated to run under OS Windows, Mac OS X, GNU/Linux or Solaris on platforms using most common HW architectures. The tool allows performing full virtualization with a hosted hypervisor which means that an already installed operating system is required to run this tool. The vast versatility of this tool enables an easy transition between different hosts with different operating systems [7]. The most favoured and well-known producer in the field of virtualization is indisputably VMware. Its products belong among the most used solutions and are mainly designated for the x86 architecture and its descendant x86-64. The company offers products that implement a so called “bare-metal” hypervisor or a hosted hypervisor allowing the company to cover a larger part of the market spanning from end-users with low requirements to servers and data centres in which high efficiency, performance and scalability is a must.

3. Measuring platform and methodology

The methodology used in this paper relies on and uses the free full virtualization tools. Their virtual machines will be run on high performing hardware with a hardware assisted virtualization support.

3.1. Measuring platform preparation

As the KVM needs a hardware-assisted virtualization support, it is necessary to use a computer equipped with a processor supporting the Intel-VTx or AMD-V technology, two incarnations of the mentioned hardware-assisted virtualization technology from

---

* **Miroslav Voznak, Jiri Slachta, Filip Rezac, Jan Rozhon**
Department of Telecommunications, Faculty of Electrical Engineering and Computer Science, VSB-TU Ostrava, Czech Republic,
E-mail: voznak@ieee.org

both largest x86 processor producers. Main hardware and software parameters of the used computer are summarized in further points: processor Intel(R) Core(TM) i7, 8 GB RAM, two 1Gbps NIC and 64-bit operating system Debian Squeeze. The tested topology consists of one computer with SW Asterisk PBX on virtualized platform, one 1Gbps switch and traffic generator Optixia XM2 with IxLoad control SW [8], the situation is depicted in Fig. 1.

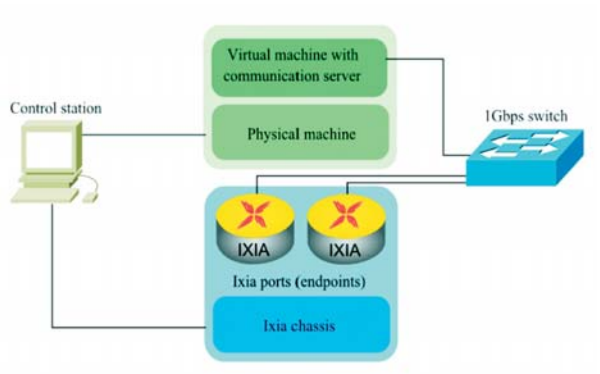


Fig. 1 Test topology

3.2. Measured parameters

Real-time network applications including IP telephony depend on network parameters that influence the transmission quality [9] and [10]. Using the Optixia XM2 we are able to measure Interarrival Jitter, Delay Variation Jitter, One Way Delay, Post Dial Delay, Media Delay and Post Pickup Delay [8]. We define a variation of delay as a jitter. It is the difference between the expected and real time of the packet reception. This appears during the packet transport through the IP network when the time shift between packets occurs because of the queue ordering in routers [10], [11] and [12]. In this article, the following parameters were chosen to carry out measurements on generator and analyzer: Interarrival Jitter, Delay Variation Jitter and Post Dial Delay. Due to the low virtual machine utilization and low number of the UDP sockets, finally, we decided to implement a codec translation which, of course, increases the utilization and after that we could observe a difference in performance of real-time applications on various virtualization platforms. The configuration can be split into three parts – the first part with global parameters, the second with network parameters, and the third describes the selected test activity. Therefore, the test scenario consists of a fixed part which is the same for all the tests, and the variable part which is determined for the selected activity. Activities can be combined, enabling measuring multiple parameters during single test iteration. The test starts with the MakeRegistration procedure. Once both sides of the communication are registered, the SIP MakeCallAuthentication and SIP ReceiveCallAuthentication procedures are executed. These are followed by the authentication RTP session. Once it is over, the call is ended.

3.3. Test methodology

The test scenario remains the same for all the tests though several parameters of the virtual machines, including RAM capacity, number of processor cores and used virtualization tool, changed. Due to the above mentioned limitations, the end-to-end delay variation can only be measured between the UA that generates the call and the communication server. Under our scenario, the traffic has a linearly increasing trend but the utilization and delay increases are not linear at all. Asterisk PBX responds to an increasing load with a notch increase once a certain load threshold is exceeded. Once the hardware limit has been reached, Asterisk begins to refuse registrations and first unsuccessful calls appear.

4. Results

The data files were analyzed using the exploratory analysis applied to each individual parameter. The ANOVA test was applied to verify data independence and other required properties. Every result category consists of charts describing how the three most important parameters are influenced by the current environmental setting. These parameters are Post Dial Delay, Delay Variation Jitter and Interarrival Jitter.

4.1. Classification according to virtual tools performance

Fig. 2 depicts three variables (KVM, VirtualBox, VMware) and their effect on the Post Dial Delay and Delay Variation Jitter. The first variable, KVM, has a very limited range of measured data especially when compared to the other variables (VirtualBox and VMware). However, due to this limited range it is impossible to determine how these values are distributed.

VirtualBox has its median value lower than the average meaning that the most values were observed mainly under the average which is affected by several high values. This can be said about VMware as well, since the data distribution is similar to VirtualBox except the narrower data range. In the second parameter, Delay Variation Jitter, we can see the similar behaviour.

4.2 Classification according to memory size

The classification according to a memory size describes data properties for four variables which represent individual memory sizes. We had in use following four different values successively of the RAM size 512 MB, 1GB, 2GB and 4GB. Our exploratory analysis showed a low variety of results and a presence of outliers, as is depicted in Fig. 3. According to the results of measurements classified by a memory size, we can assume that virtual machines are memory independent when speaking about the reasonable amounts of memory.

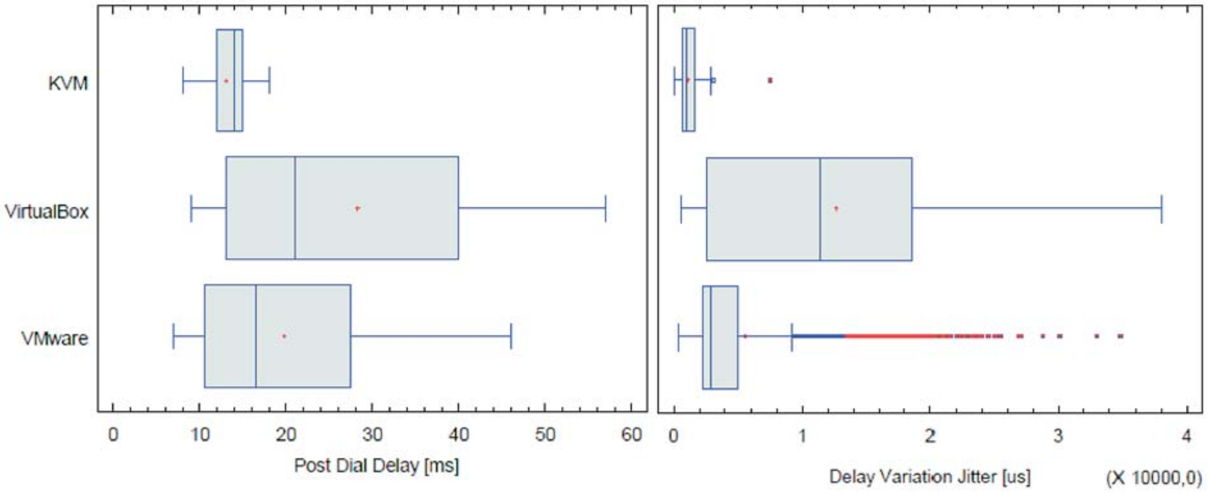


Fig. 2 Boxplot of Post Dial Delay and Delay Variation Jitter for all three virtualization tools

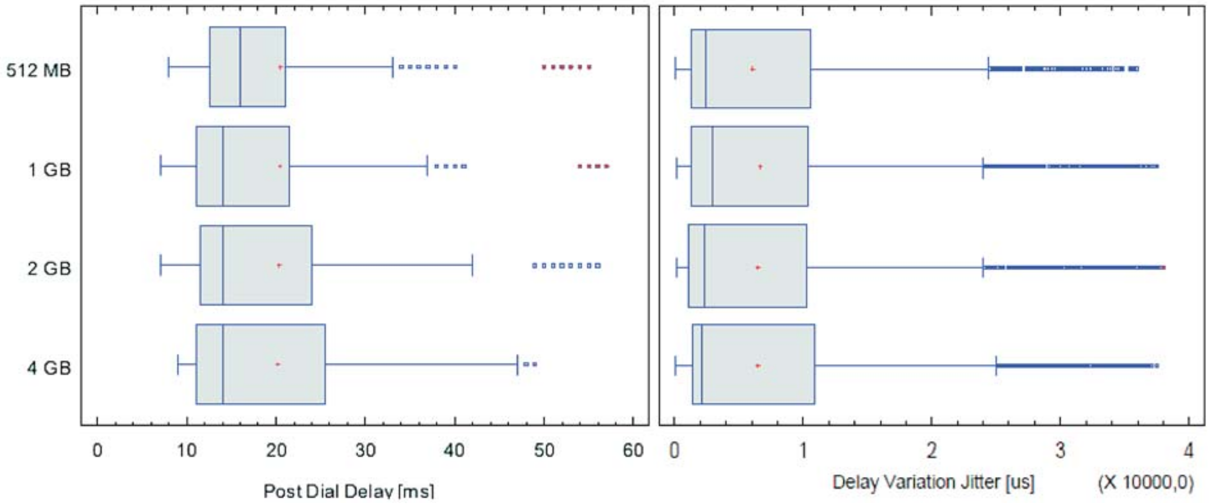


Fig. 3 Boxplot of Post Dial Delay and of Delay Variation Jitter for test runs with 512 MB to 4 GB RAM size

4.3 Classification according to processor cores

The classification according to a number of processor cores describes data properties for four variables which represent individual number of processor cores. We had the following sequence of processor cores: one core, two cores, three cores and four cores. Figure 4 shows the results for Delay Variation Jitter. The number of core processors significantly affects the evaluated results in all the measured parameters. Exploratory analysis was again performed from data set across all the tested platforms.

4.4 Variance analysis

All the tables that are presented in this section contain values of Delay Variation Jitter in relation to the type of the used virtu-

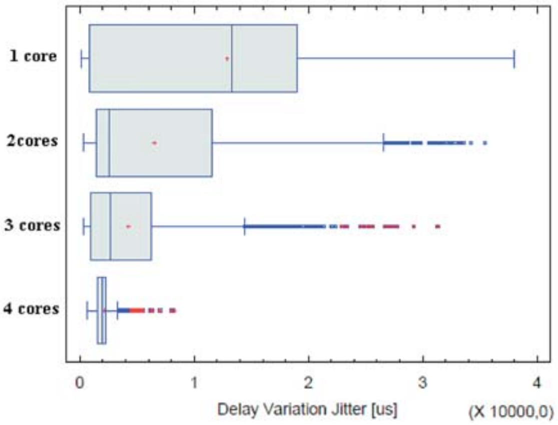


Fig. 4 Boxplot of Delay Variation Jitter for various numbers of cores

alization tool. First, we need to find out whether the data set is compliant with the Normal distribution; the results of chi-squared test are in Table 1.

Chi-Squared Test of Delay Variation Jitter Data Set Normality — Table 1

	Test	Chi-Squared
KVM	P-value	< 0.01
VirtualBox	P-value	< 0.01
VMware	P-value	< 0.01

A chi-squared test or $\chi 2$ test, is any statistical hypothesis test in which the sampling distribution of the test statistic is a chi-squared distribution. Since P-value for all the virtualization tools is equal to zero, we can state that data distribution in all cases is compliant with the normal distribution. Then we continue with the homoscedasticity test to find out if the variances of the data sets are equal or not. Data variables, $N(\mu_i, \sum_i)$ are homoscedastic if they share a common covariance (or correlation) matrix $\sum_i = \sum_j, \forall i,j$. Because all the data variables do not come from the normal distribution, to confirm their homoscedasticity we use the Levene's test instead of Bartlett's test. Levene's test is used to test the null hypothesis that all k population variances are equal against the alternative that at least two are different. Let $Z_{ij} = |X_{ij} - \overline{X}_i|$, n be the total number of samples and n_i be the number of samples in the i-th group, then denote

$$\bar{Z}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} Z_{ij}\,,\ \bar{Z} = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{n_i} Z_{ij} \tag{1}$$

$$SS_{ZB} = \sum_{i=1}^{k} n_i\left(\bar{Z}_i - Z\right)^2,\ SS_z = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(\bar{Z}_{ij} - Z\right)^2 \tag{2}$$

If the null hypothesis is valid then test statistic returns approximately Fisher-Snedecor distribution with $k-l$ degrees of freedom in the nominator and $n-k$ degrees of freedom in the denominator. Levene's test statistic is expressed in relation (3).

$$\frac{\dfrac{SS_{ZB}}{k-1}}{\dfrac{SS_Z}{n-k}} \tag{3}$$

Since the null hypothesis assumes that the variances of individual data sets are equal, we can now state that according to the P-value obtained from the Levene's test in Table 2 this hypothesis can be rejected. This means for us that we can take the data sets as different and, therefore, continue with the Kruskal-Wallis test. For k independent observations ranked as $X_{11}, X_{12}, \dots, X_{1n_1} \dots X_{k1}, X_{k2}, \dots, X_{kn_k}$ we denote n as the total number of observations across all the groups. Then we determine R_{ij} as the rank (among all the observations) of observation j from group i, thereby T_i is expressed as their mean value (4) and the test statistic is given by (5).

$$T_i = \frac{1}{n}\sum_{j=1}^{n_i} R_{ij} \tag{4}$$

$$Q = -3(n+1) + \frac{12}{n(n+1)}\sum_{i=1}^{k}\frac{T_i^2}{n_i} - 3(n+1) \tag{5}$$

Levene's Test of Delay Variation Jitter Data Set Homoscedasticity Confirmation — Table 2

	P-value
Levene's test	< 0.01

Using this test can point us in direction of further data analysis and more importantly provide us with the information whether median values of individual data sets are equal. Since the P-value obtained from this test is 0 we can reject the null hypothesis and we now know that the median values differ, which can be confirmed from the presented boxplots. With Kruskal–Wallis test done, we can now proceed to post-hoc analysis of this test using the so called Dunn's test. Dunn's method is used in cases of rejecting the zero hypotheses in the Kruskal-Wallis test. It is used for multiple median comparisons and can say whether two chosen data sets differ greatly in their distribution, mainly median. The results of Dunn's method are presented in Table 3.

Results of Delay Variation Jitter Analysis using the Dunn's Method — Table 3

	$t_i - t_j$	Critical Value
KVM-VirtualBox	4189.91	92.244
KVM-VMware	4130.29	92.244
VirtualBox-VMware	1812.44	92.244

Using the critical value from the table above, we can learn that all three pairs differ significantly in their median values. This way confirmed the properties of all the data sets in this article, but for the sake of the reasonable size of the paper we do not publish them for all the measured parameters.

5. Conclusion

Although the properties of the test did not allow measuring data traffic between the sender's and receiver's user agents due to the codec translation, it was possible to compare the obtained values of communication between the user agent and the server providing the IP telephony services. Looking at the results of the exploratory analysis, we can conclude that the pre-test assumptions regarding the virtualization tool performance were correct. The lowest range of values of the Post Dial Delay, Delay Variation Jitter and Interarrival jitter is achieved using the KVM. Although the VMware has a background of large and prosperous company, it did not perform well enough to beat its competitor KVM especially as regards the stability of results. From the real-time application point of view, VirtualBox can be considered as the least efficient and advantageous solution as the values of all three measured parameters obtained while measuring with this virtualization tool were the worst in every aspect. Looking at the results, we can also assume that the virtual machines are not memory dependent. Their dependence on the number of processor cores, on the other hand, is rather obvious. Using other statistical techniques we have confirmed that the data for different categories (KVM, VMware, VirtualBox; CPU core categories) have different characteristics and their median values do not match. This and other possible interpretation of the results can be read from the presented boxplots.

Acknowledgement

This work was supported by the European Regional Development Fund in the IT4Innovations Centre of Excellence Project (CZ.1.05/1.1.00/02.0070) and by the Development of human resources in research and development of latest soft computing methods and their application in practice Project (CZ.1.07/2.3.00/20.0072) funded by the Operational Programme Education for Competitiveness, co-financed by ESF and state budget of the Czech Republic.

References

[1] BIRKOS, K., CHRYSIKOS, T., KOTSOPOULOS, S.: QoS Issues in Complex-based High Platform Communication Systems, *Communications - Scientific Letters of the University of Zilina*, vol. 9, No. 4, pp. 5–9, 2007.

[2] VOZNAK, M., SLACHTA, J., ROZHON, J.: Performance Analysis of Virtualized Real-Time Applications, *Intern. J. of Mathematical Models and Methods in Applied Sciences,* vol. 6, No. 2, pp. 305–313, 2012.

[3] KLIMO, M., KOVACIKOVA, M., SEGEC, P.: Selected Issues of IP telephony, *Communications - Scientific Letters of the University of Zilina,* vol. 6, No 4, pp. 63–70, 2004.

[4] VOZNAK, M., ROZHON, J.: Methodology for SIP Infrastructure Performance Testing, *WSEAS Transactions on Computers,* vol. 9, No. 9, pp. 1012–1021, 2010.

[5] MONFORT, J., ADLER, K., GIERLICH, H., POMY, J.: Multimedia Quality as Perceived by the User, *Communications - Scientific Letters of the University of Zilina,* vol. 10, No. 4, pp. 5–10, 2008.

[6] Red Hat. Red Hat Enterprise Linux 6 Virtualization Guide: Guide to Virtualization on Red Hat Enterprise Linux 6, 2010. URL https://access.redhat.com/knowledge/.

[7] Oracle Corporation, VM Virtual Box, p.287, 2011. URL http://download.virtualbox.org.

[8] Ixia, Hardware and Reference Manual Release, 2010. URL http://ixiacom.com.

[9] BLASKOVA, L., HOLUB, J.: How do Non-native Listeners Perceive Quality of Transmitted Voice?, *Communications - Scientific Letters of the University of Zilina,* vol. 10, No. 4, pp. 11–14, 2008.

[10] KOVAC, A., HALAS, M., ORGON, M., VOZNAK, M.: E-model MOS Estimate Improvement through Jitter Buffer Packet Loss Modelling, *Advances in Electrical and Electronic Engineering,* vol. 9, No. 5, pp. 233–242, 2011.

[11] VOZNAK, M., TOMES, M., VACLAVIKOVA, Z., HALAS, M.: E-model Improvement for Speech Quality Evaluation Including Codecs Tandeming, *Intern. Conference on Data Networks, Communications, Computers, Faro,* pp. 119–124, 2010.

[12] VOZNAK, M.: E-model Modification for Case of Cascade Codecs Arrangement, Intern. *J. of Mathematical Models and Methods in Applied Sciences,* vol. 5, No. 8, pp. 1439–1447, 2011.

Petr Munster – Radim Sifta – Vladimir Tejkal – Miloslav Filka *

THE INFLUENCE OF BINARY MODULATIONS IN OTDM

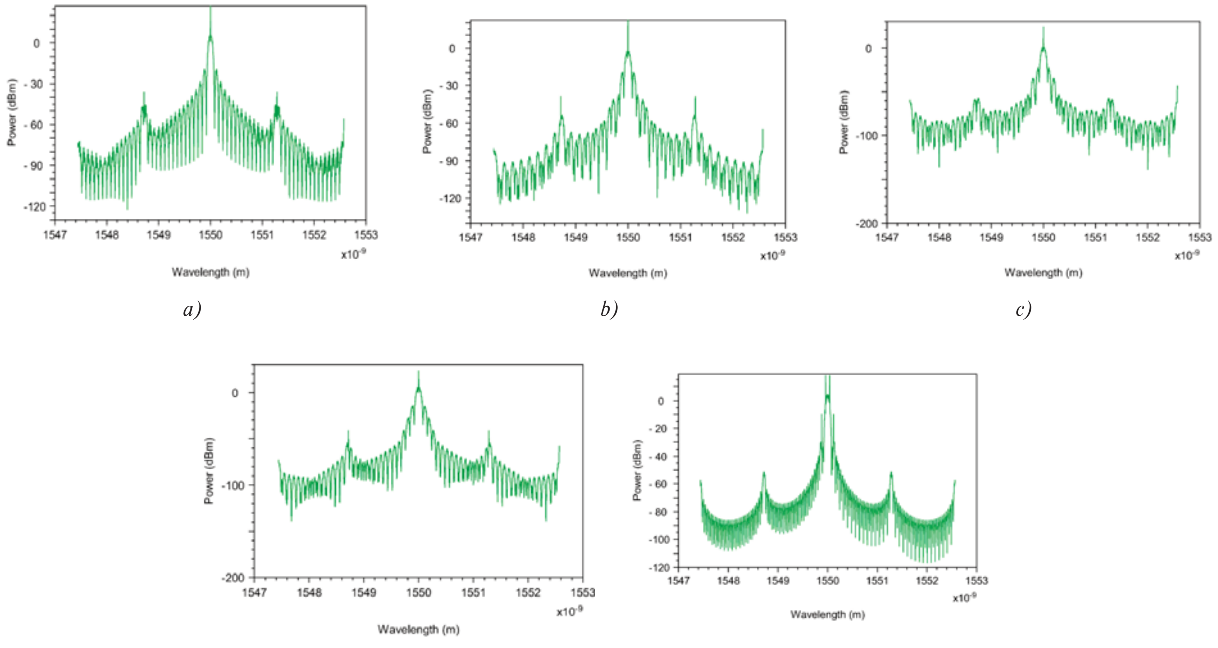
In this paper we propose a simulation test using different binary modulation formats in optical network based on the Optical Time Division Multiplexing (OTDM). OTDM in optical networks is used for high speed data generation in point–to–point networks. Dispersion and nonlinearities in optical fibers can cause transmission disturbances, which can influence the throughput of the transmission system and maximal achievable distance. Basic modulation formats RZ and NRZ with new CRZ and CSRZ formats were tested in OTDM simulation model.

***Keywords:** OTDM, modulation, RZ, NRZ, CSRZ, CRZ, BER, Q-factor.*

1. Introduction

In addition to increasing user demands and new internet services like Triple Play, the requirement for greater bandwidth grows around 50–70% per year [1]. Copper wirings are close to their limits – both achievable distance and bandwidth. Optical fiber seems to be the best appropriate solution. Gradually, optical fibers spread from core networks to access networks.

Data transport in optical networks can be presented by the Time Division Multiplexing (TDM) and the Wavelength Division Multiplexing (WDM), where TDM can be realized electrically (ETDM) or optically (OTDM). OTDM has no electronics limitations and so can be used for high-speed data generation (actually the research groups achieved over 10.2 Tbps per one channel). In the comparison with limitation of ETDM networks, OTDM appears to be a solution for networks of next generations.



a) *b)* *c)*

d) *e)*

Fig. 1 Frequency spectrum of transmitted 160 Gbps OTDM signals at the output of multiplexer for modulation formats: a) NRZ, b) RZ 33%, c) RZ 50%, d) CRZ, e) CSRZ

---

* **Petr Munster, Radim Sifta, Vladimir Tejkal, Miloslav Filka**
Brno University of Technology, Czech Republic, E-mail: munster@feec.vutbr.cz

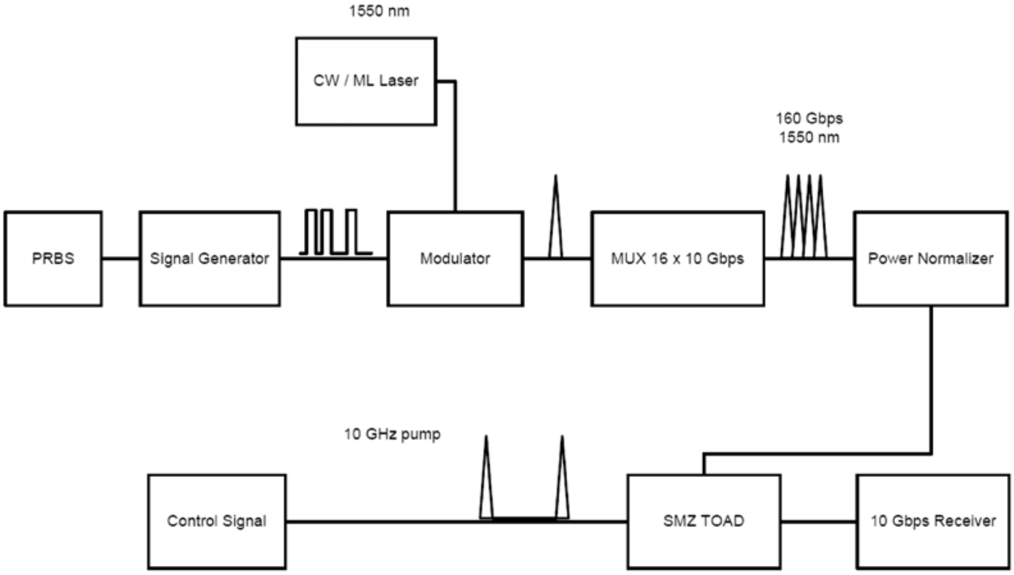


Fig. 2 Block scheme of the 160 Gbps OTDM network. Except the block modulator, the scheme was the same for all modulation formats

Throughput of every transmission system can be affected by nonlinearities and dispersions. Modulation formats have a big importance for transmission systems and could have a positive influence on maximal achievable distance and bit-rate. Except basic modulation formats the Non-Return-to-Zero (NRZ) and the Return-to-Zero (RZ), new types of modulation formats for optical networks were tested – the Chirped Return-to-Zero (CRZ) and the Carrier-Suppressed Return-to-Zero (CSRZ).

A) NRZ

NRZ is a main modulation format for transmission of data signal over optical fiber. Level of logical 1 takes the whole bit interval and in this duration the level of logical signal does not return to zero. The return to zero is caused by the change of the signal. Phase is 0 for the level of logical zero and π for the level of logical 1. In the comparison with other modulation formats, NRZ has a narrower spectrum of the central lobe and signal peaks are at multiples of the bit-rate [2].

B) RZ

The RZ modulation is based on return of the signal to logic zero. Pulses wide are shorter than bit interval wide. RZ is based on the NRZ modulation which is further modulated by a sinusoidal function. Commonly used formats are: RZ with 50% duty cycle, RZ with 33% duty cycle, RZ with 67% duty cycle (CSRZ). Pulses for all RZ signals have the same shape independent of the neighboring bit value. In the ideal modulation the pulse phases are identical. The central lobe of the optical spectrum of these RZ signals is wider than central lobe of the NRZ signal due to the narrower width of the RZ pulses [3].

C) CRZ

In CRZ a chirp is added to RZ by applying a phase modulation where the sinusoidal control voltage has twice the frequency compared 50% RZ and oscillates twice the amplitude. Compared to classical RZ pulses phase is changed by π every following bit period. Carrier frequency is destructively affected and no peak is apparent on it, which improves the signal modulation [3].

D) CSRZ

In case of the CSRZ modulation the RZ optical signal after Mach-Zehnder modulator goes through a phase modulator driven by an analog sine wave generator at a frequency equal to half of the bit rate. That will introduce a π phase shift between any two adjacent bits and the spectrum will be modified so that the central peak at the carrier frequency is suppressed [4].

2. OTDM Simulation model

In OTDM narrow optical pulse lasers are used, hence it is possible to use extremely narrow time slots with a correspondingly high bandwidth. The higher bit-rate than for ETDM in one wavelength channel is achieved by multiplexing lower bitrates in an optical domain [5].

This section demonstrates simulation of 160 Gbps transmission and demultiplexing in OTDM network for different modulation formats. Simulations were done with OptSim software from RSoft Design Group [6]. The block scheme of simulation model in Fig. 2 is divided into 4 main parts – transmitter, control signal, the Symmetrical Mach-Zehnder Interferometer (SMZ) and receiver.

The transmitter part is based on 16×10 Gbps data generation in sixteen channels on the same wavelength 1550 nm and with pseudorandom bit sequence. In this paper two almost same simulation models with same parameters were tested. Just the CW laser was used in the first model and the ML laser was used in the second model (the default setups were used for both laser models). Pseudorandom bit sequence is modulated and then split in 8 directions. Every direction is delayed by 1/8 time window and multiplexed. The signal is multiplexed again by 1/2 time window. Spectra of the transmitted signals are shown in Fig. 1.

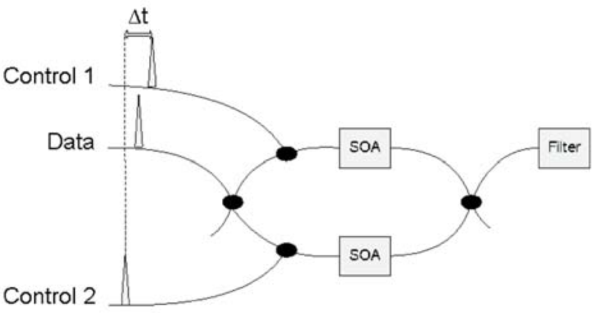


Fig. 3 TOAD based on Symmetrical Mach-Zehnder Interferometer

The important part of the Symmetrical Mach–Zehnder Interferometer (SMZ in Fig. 2) is 10 GHz control signal. The control signal is used for switching when it saturates the Semiconductor optical amplifier (SOA) in the loop and changes the index of refraction. Between two counterpropagating data pulses the differential phase shift is achieved for data pulses switching to the output port. SOA offset from the center position then provides switching window duration. The principle of the Terahertz Optical Asymmetric Demultiplexer (TOAD) based on SMZ is shown in Fig. 3.

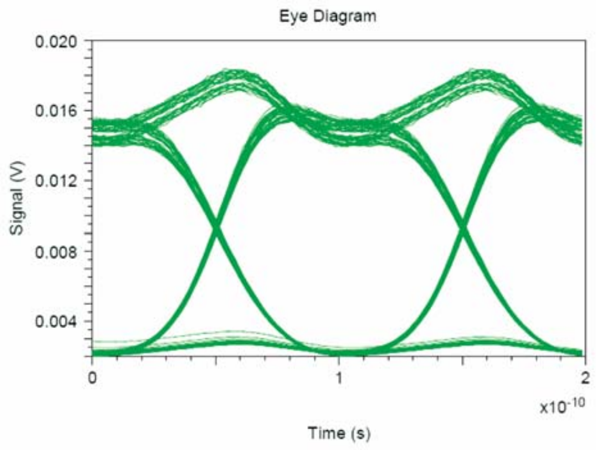


Fig. 4 Eye diagram for RZ 33% modulation with BER 10^{-10}

The output signal at the switching port carries data information of demultiplexed channel and suppressed interference from other channels. Fig. 4 shows corresponding Eye diagram of the switching port at the receiver part for the RZ modulation.

3. Description of the parameters

The main parameters for assessing the quality are the Bit Error Rate (BER) which shows the signal to noise ratio, Eye diagram and Q-factor.

The Q-factor is a performance estimator. It shows the quality of the signal with regard to the signal to noise ratio (SNR). This includes all physical signal disturbances. These disturbances reduce the quality of the signal and cause bit errors [4]. Consequently, the higher Q-factor ratio means a better signal to noise ratio and thus lower *BER*. The *BER* is approximately [4]:

$$BER \cong 0.5 erfc\left(\frac{Q}{\sqrt{2}}\right) \tag{1}$$

The *BER* expresses the frequency of bit errors given by the ratio of the incorrectly transmitted elements in a digital signal to the total number of transferred elements [4].

In telecommunication systems an eye pattern (Eye diagram) is an oscilloscope display where the vertical input repetitively sampled a digital data signal from the receiver and data rate is used to trigger the horizontal sweep. From the Eye diagram is possible to offer some other parameters like jitter, Eye Amplitude, Eye Delay or Intersymbol Interference (ISI), which shows the interleaving of the modulation pulses.

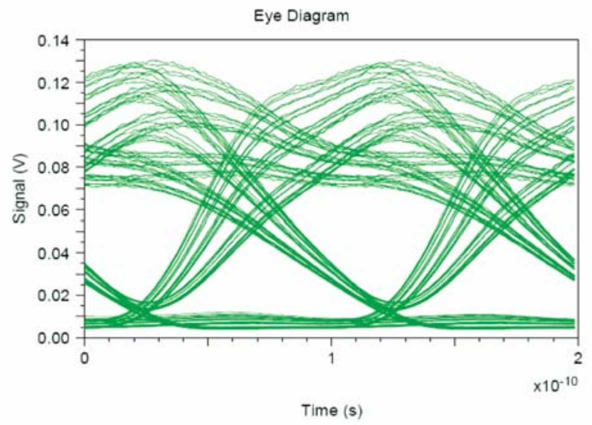


Fig. 5 Eye diagram for NRZ modulation with BER 10^{-2}

Corresponding Eye diagrams with good BER and with low BER are shown in Fig. 4 or in Fig. 5.

4. Results

The spectra of transmitted signals are shown in Fig. 1. The central peaks of all transmitted signals are set to 1550 nm wavelength. Both NRZ and CSRZ have a similar narrow power spectrum peak, but the CSRZ modulation does not have a chirp in the middle of the carrier frequency. Absence of this chirp in the central of the carrier frequency reduces the intersymbol interference. The CRZ modulation has a wider spectrum than other modulation formats and is less affected by nonlinearities (but a wider spectrum is worse for WDM systems).

Results of BER and Q-factor for all modulation formats for both lasers Table 1

CW Laser

	BER [-]	Q_{dB} [dB]
NRZ	2.16E-02	6.11E+00
RZ 33 %	1.16E-10	1.60E+01
RZ 50 %	2.35E-13	1.72E+01
CRZ	1.14E-13	1.73E+01
CSRZ	5.53E-18	1.87E+01

ML Laser

	BER [-]	Q_{dB} [dB]
NRZ	9.77E-15	1.77E+01
RZ 33 %	5.40E-13	1.71E+01
RZ 50 %	3.56E-15	1.78E+01
CRZ	3.26E-17	1.84E+01
CSRZ	1.33E-05	1.25E+01

Note that the widely accepted definition is used to express the factor in decibel units.

As shown in Fig. 6 the best BER was achieved with the CSRZ modulation and CW laser in transmitter. The worst BER (2.16 E-2) was achieved with the CW laser and NRZ modulation. This value is low and the system could not correctly recognize the signal. Results of the Q-factor and BER for all modulation formats are shown in Table 1.

Graphs of the Q-factor and BER (Figs. 6 and 7) show all modulation formats with their values. As you can see with a higher Q-factor also BER grows. The minimal value of BER for the correctly recognized signal should be about 10^{-5} and for the Q-factor about 4 dB.

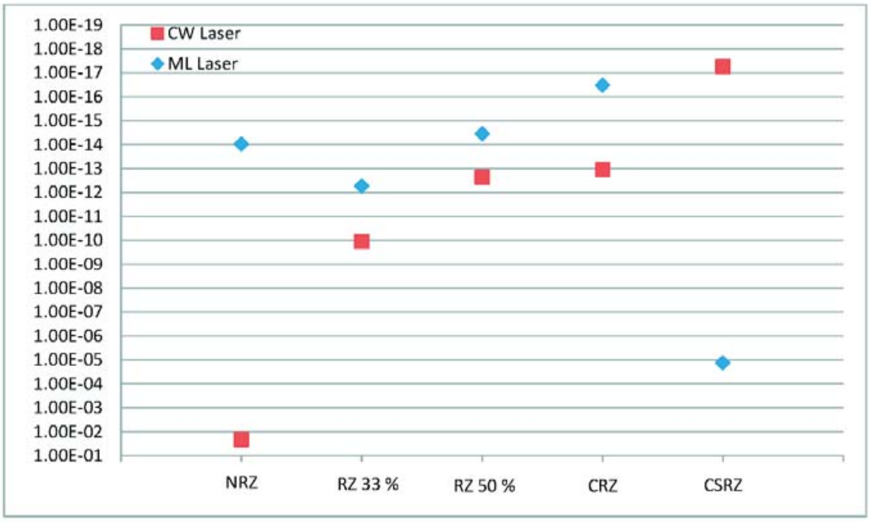


Fig. 6 Graph of comparison BER parameter for all modulation formats for both lasers

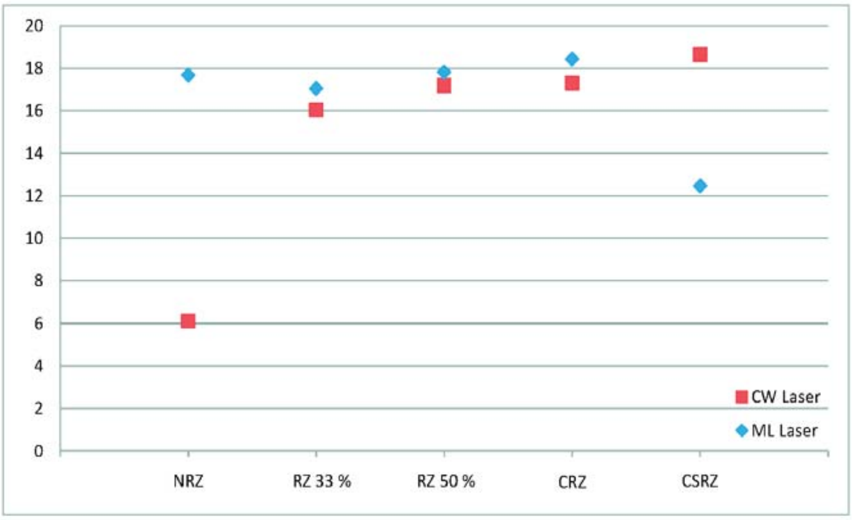


Fig. 7 Graph of comparison Q-factor parameter for all modulation formats for both lasers

5. Conclusion

In this paper new modulation formats for OTDM were tested. At present the NRZ modulation format is the most used format for optical communications because of its simple implementation and good transmit parameters. From the results you can see that other modulation formats (RZ, CSRZ and CRZ) had better parameters in the tested OTDM models and could be used in optical communications. The CSRZ modulation due to its good tested transmit parameters should be used for longer distances. The CRZ modulation then should be used in optical networks with big influence of nonlinearities and dispersions. Laser also has big influence on the transmitted signal. With mode lock lasers the short pulses can be generated and, therefore, these lasers are more suitable for high-speed data generation.

Acknowledgement

This research work is funded by projects SIX CZ.1.05/2.1.00/03.0072, EU ECOP EE.2.3.20.0094, CZ.1.07/2.2.00/28.0062, and MPO FR-TI4/696, Czech Republic.

References

[1] SWANSON, B., GILDER, G.: *Estimating the Exaflood – the Impact of Video and Rich Media on the Internet*, Tech. Rep., Discovery Institute Seattle: Washington, 2008.

[2] TEJKAL, V., FILKA, M., REICHERT, P., SPORIK, J.: Binary Modulation Formats in Optical Access Networks (in Czech), *Optics and Optoelectronic,* vol. 8, pp. 96–101, 2010.

[3] TEJKAL, V., FILKA, M., SPORIK, J., REICHERT, P., MUNSTER, P.: *The Influence of Binary Modulations in Passive Optical Network Based on Wdm,* 34th Intern. Conference on Telecommunications and Signal Processing (TSP), vol. 34, pp. 141–144, 2011.

[4] BOSCO, G., CARENA, A., CURRI, V., GAUDINO, R., POGGIOLINI, P.: On the Use of NRZ, RZ, and CSRZ Modulation at 40 gb/s with Narrow Dwdm Channel Spacing, *J. of Lightwave Technology,* vol. 9, pp. 1694–1704, 2002.

[5] MUNSTER, P.: *Otdm Based Passive Optical Network*, EEICT, Proc. of the 18th conference, vol. 3, pp. 24–28, 2012.

[6] RSOFT DESIGN GROUP, New York, OptSim, 2010.

David Kurc – Jiri Schimmel *

SIMULATION OF LINE ARRAY SOUND SOURCE

This paper is focused on the simulation of line array source that consists of eight separate loudspeaker enclosures of the same type having some directional characteristics influenced by the position within the line array. The influence on these characteristics caused by the different positions of enclosures in the array was tested. The characteristics of the line array source were measured in the anechoic chamber. The comparison between the results of the simulation of ideal and real sources is shown.

Keywords: *Sound source, acoustic signal, acoustic simulation, directional characteristics.*

1. Introduction

Sound systems are nowadays increasingly confronted with the requirement of high fidelity sound reproduction. This is closely related to the spectral composition of reproduced acoustic signal. Thus, the reproduced signal should contain the same spectral components as the original and have the least possible distortion (linear or nonlinear). However, high fidelity also increasingly depends on the transfer of information about the original sound source location. It means the direction and distance of the original source related to the listener. For example, this information is very important in the area of teleconferencing technology, but also in the area of arts or entertainment. Among other methods of multichannel sound reproduction systems there is wave field synthesis (WFS) [1]. WFS is a modern method of sound reproduction allowing the true reproduction of sound field including the spatial information. This method utilizing almost exclusively horizontally installed line array sources for the reproduction of acoustic signal.

It is very useful to simulate the acoustic field generated by single sound source or even by more sources. The final image can display the interaction of the sources in areas where they are summed together according to the phase and amplitude relations between them. Therefore, we can get the picture of sound pressure level in the area in front of the loudspeaker array.

The method of ideal sound source simulation is presented in [2]. In our work, we take measured directional characteristics of real sources into account, including diffraction.

The line array sound source described in this paper was designed as the sound source for the use with the WFS. The system consists of eight loudspeaker enclosures, each of the same construction and with 3-inch speaker inside.

The paper presents the process of directional characteristics measurement and subsequent application of the measured data to simulate directional behavior of the real sources. All the simulations were done in Matlab.

2. Ideal sound source simulation

The essential assumption for all of the presented simulations is that all sources are placed in the free field. Therefore, reflections from the walls or other barriers are not considered [3]. Moreover, the simulations and measurement were done only in the horizontal plane. One of the possible ways of sound field description is utilizing the velocity potential Φ [3] and [4]. It is a scalar function and acoustic (particle) velocity $v = \text{grad}\,\Phi$ can be calculated as [5]. An advantage of description with velocity potential is the scalar addition of individual sources contributions as described

$$\Phi = \sum_{i=1}^{n} \Phi_i \,. \tag{1}$$

It means that the total velocity potential in every point of the space is the sum of potentials induced of all the sources. Then acoustic pressure p of harmonic wave with angular frequency ω can be written as

$$p = -j\omega\rho_0\Phi \tag{2}$$

where ρ_0 is medium density (air density in our case).

Basic source used in acoustic simulations is a point source [6]. Its main property is frequency-independent omnidirectional polar pattern. Velocity potential of the *i*-th point source at the point X in distance of r_i from the source can be calculated as [3] and [4]

$$\Phi_i = A_i \frac{e^{-jkr_i}}{r_i} \tag{3}$$

---

* **David Kurc, Jiri Schimmel**
Department of Telecommunications, FEEC, Brno University of Technology, Czech Republic, E-mail: david.kurc@phd.feec.vutbr.cz

where A_i is an amplitude of the source and k is the wave number ($k = \omega/c_0$, where c_0 is speed of sound). However, the real sources (loudspeakers) can be considered as the point sources only in the situations when the wavelength of emitted sound wave is much greater than physical dimensions of the source [6]. Nevertheless, only very low frequencies satisfy this requirement – more precisely, the frequencies meeting condition $kR < 1$, where R is the radius of speaker diaphragm. In other cases it is more accurate to utilize vibrating circular piston sources in the simulations [4]. It is considered to be located on the wall surface, which is perfectly rigid, flat and infinitely large. Due to this assumption, there is no need to consider the phenomena associated with diffraction on the edges [4]. However, the properties of this kind of source are more similar to the real loudspeakers with circular diaphragm, especially in directional characteristics. Velocity potential of the i-th vibrating circular piston source at the point X in distance of r_i from the source can be calculated as [4]

$$\Phi_i = -A_i \frac{e^{-jkr_i}}{r_i}\pi R_i^2 \frac{2J_1(kR_i\sin\vartheta_i)}{kR_i\sin\vartheta_i} \tag{4}$$

where R_i is the radius of i-th circular piston source, is the Bessel function of the first kind and ϑ_i is the angle between the acoustic axis of the loudspeaker and the line connecting the speaker reference point and calculated point X. For example, the loudspeakers used for our simulations have the diaphragm radius of 3.81 cm. Thus, eq. (3) can be used for the simulations of frequencies below 1.42 kHz, while eq. (4) have to be used for higher frequencies.

The comparison example of these kinds of sources with the frequency of 4 kHz can be seen in Fig. 1. As can be seen in the

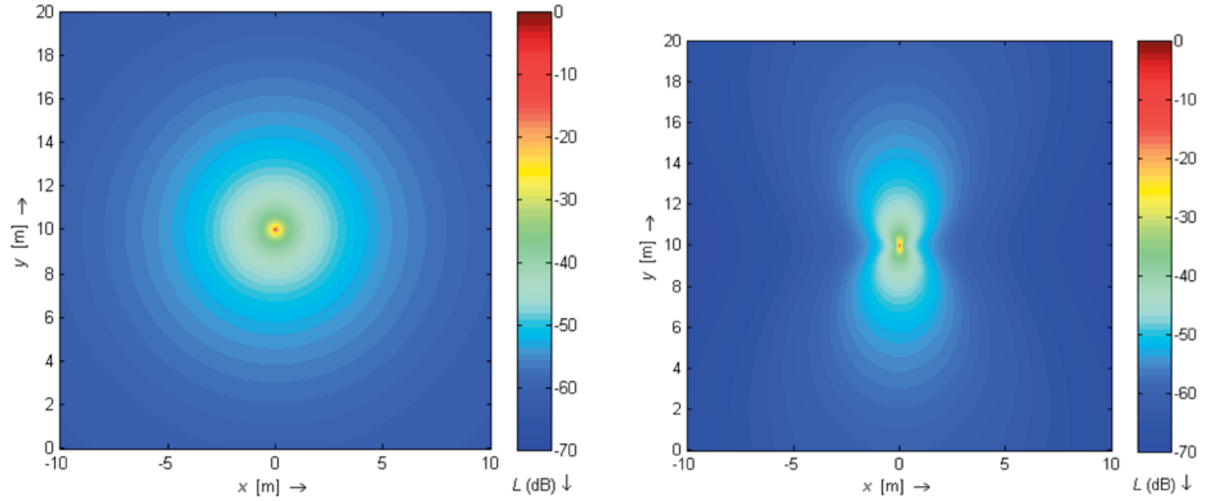


Fig. 1 The comparison of point source (left) and circular piston source (right) at the frequency of 4 kHz. Diaphragm radius is 3.81 cm

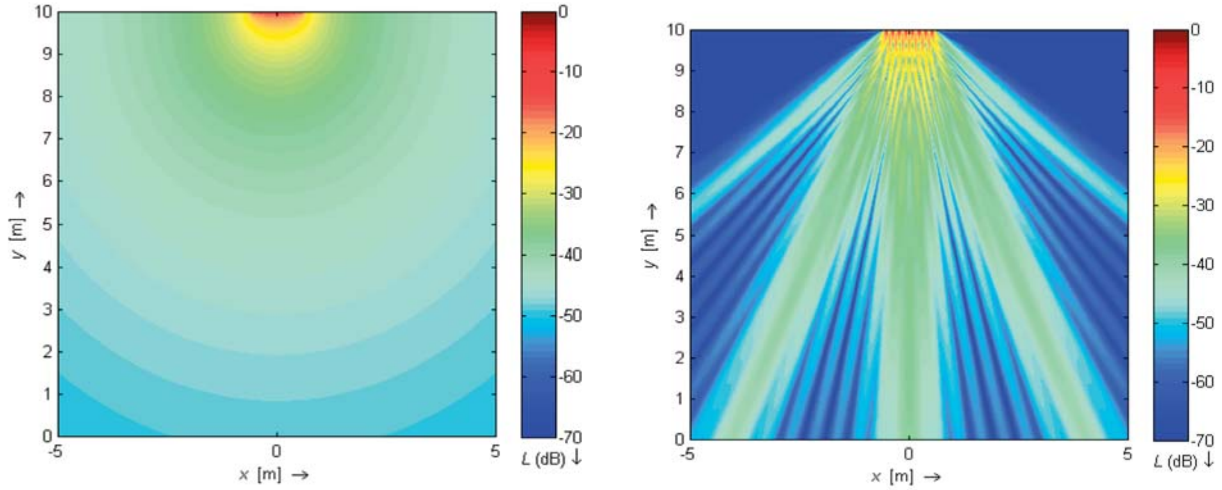


Fig. 2 The line array source simulation with eight individual sources spaced 15 cm at the frequency of 100 Hz (left) and at the frequency of 6 kHz (right). Diaphragm radius is 3.81 cm

figures, the main difference between these sources is in their directivity. Polar pattern of a point source is the same at all frequencies while the polar pattern of a circular piston source gets narrower with increasing frequency.

An important reason for simulation of sound pressure distribution is the visualization of interactions between two and more sources. Two major factors are crucial in the process of source summing. They are the difference in amplitude and the difference in phase of the sources [6] and [7]. The resulting picture of the interaction of the line array source containing eight loudspeaker enclosures at low frequency and high frequency is shown in Fig. 2. The point source model was used for low frequencies and the circular piston source model was used for high frequencies.

Different behaviour at low and high frequencies is apparent from the simulation. At frequency of 100 Hz the wavelength is quite long (3.43 m at the temperature 20 °C), thus more than three times longer than the overall length of the simulated line array source (1.05 m). Therefore, the directional pattern is almost omnidirectional. Nevertheless, at the frequency of 6 kHz the wavelength is only 0.057 m long, which is much less than the line source length. Due to that, there are dark blue areas in the simulation, where the individual sources are summing with similar or even the same amplitude, however, with the phase difference higher than 120° and lower than 240° [7]. The result is that the energy of the sources in this area (at this frequency) is cancelled instead summed. On the other hand, the bright colours denote the areas where the sources are in phase.

Anyway, the important feature of the simulated line array source is the front-back symmetry. Simulations are shown only for front directions in the figures presented in this paper, thus only the directional response in the front of the line source is displayed. That is because the measurements of the real directional characteristics were done only for this interval of angles.

3. Directional characteristics measurement process

Directional response data measured on real sources can provide more accurate information about some phenomena (such as diffraction) not considered in the simulations with ideal sources. A standard procedure of directional characteristic measurement is measuring the acoustic pressure effective value of loudspeaker fed with harmonic signal at given frequency [8]. A special turntable is employed to rotate the loudspeaker around its axis. However, this process is not very flexible when higher number of frequencies or the individual loudspeakers of composite sound source are needed to be measured. It is the reason why we designed and utilized the method named directional-frequency analysis.

3.1 Directional-Frequency analysis

The method is based on the use of a turntable with the constant angular speed of rotation and the broadband signal with constant power spectral density (e.g. white noise) feeding the measured loudspeaker. The output signal of a measurement microphone is recorded during rotation of the loudspeaker. The next step is time-frequency analysis of the recorded signal. For instance a discreet short-time Fourier transform (STFT) can be utilized for this purpose. It can be written as

$$s_{STFT}\left(e^{j\omega_i}, m\right) = \sum_{n=-\infty}^{\infty} s[n] w[n-m] e^{j\omega_i n} \tag{5}$$

where $s[n]$ is the sequence of recorded signal samples, $w[n]$ is the weighting function (Hamming window was used in our case), ω_i are frequency points, and m are time points. Nonetheless, several other transformations could be used as well. Due to the known constant radial frequency of turntable rotation, the time points can be substituted by the corresponding angle points. The angle α between the acoustic axis of the loudspeaker and the line connecting the point of measurement and the loudspeaker reference point can be written as

$$\alpha = v_r t \ \ [\mathrm{rad}] \tag{6}$$

where t is the time from the beginning of the record and v_r is the radial frequency in rad·s^{-1}. Following relation holds for time point m and angle α

$$m = \alpha \frac{f_s}{v} \tag{7}$$

where f_s is the sample rate of recorded signal. Therefore, results of time-frequency analysis can be interpreted as the directional-frequency analysis results. Set of transformed coefficients provide frequency responses for the particular angles of rotation as well as directional characteristics for particular frequencies. Frequency characteristics are obtained for the points ω_l with $l = 0, 1, ..., N-1$ (N is the length of weighting window) and directional characteristics are obtained for the points $m = 0, 1, ..., 2\pi f_s/v$. Both the frequency resolution $\Delta\omega$ and the angular resolution $\Delta\alpha$ depend on the length of weighting window N according to equations

$$\Delta\omega = \frac{f_s}{N} \tag{8}$$

$$\Delta\alpha = \frac{N}{f_s} v \tag{9}$$

3.2. Process of measurement

The measurement was done in the semi-anechoic chamber to simulate the free field. The critical frequency of the chamber is 200 Hz. Because of the dimensions of the line array source it was easier to rotate the measurement microphone around the source instead of the rotation of the line array source itself. The turntable was utilized to rotate the microphone fixed on a 1.25 m long beam. The reference point of the microphone was set to the same height as the reference point of the measured speaker. The turntable allows setting of the lowest constant speed of rotation (convenient for the highest precision of analysis) to $\pi/360$ rad·s^{-1}. The directional

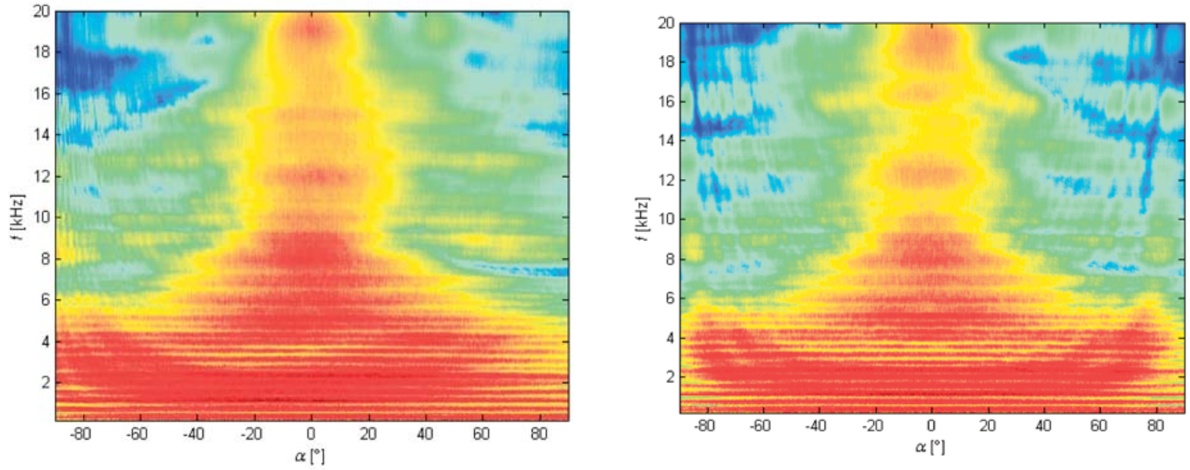


Fig. 3 The directional-frequency analysis of the first (left) and the fourth (right) individual source (from the front view and the left side)

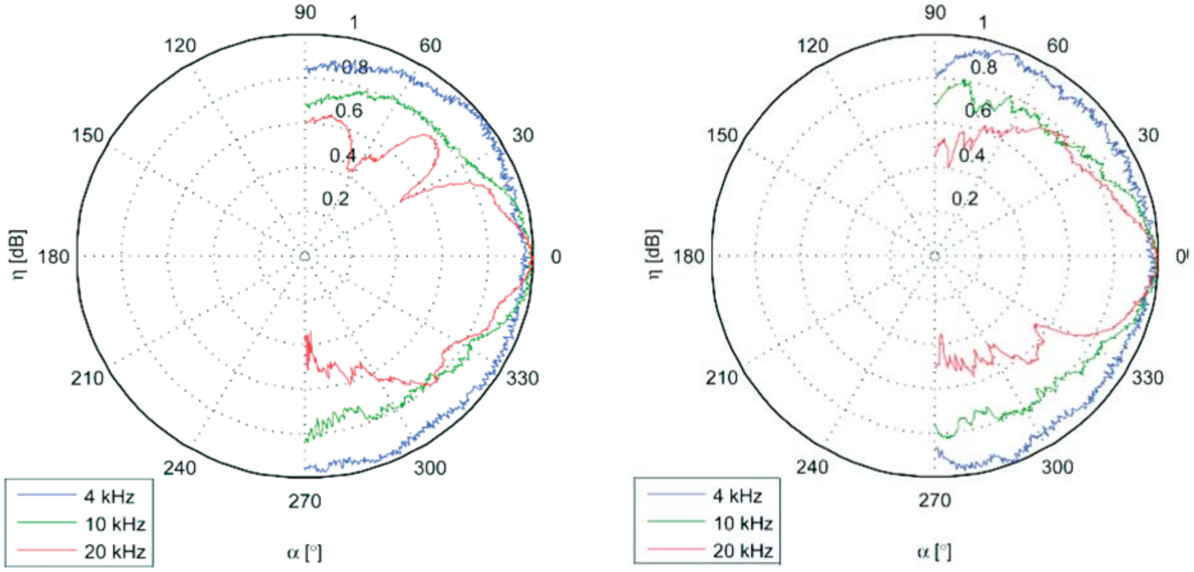


Fig. 4 Directional characteristics of the first (left) and the fourth (right) individual source (from the front view and the left side)

characteristics were measured only for the angles from the interval (−90°; 90°). White noise with a crest factor of 1.5 was used as a measurement signal. Sample rate was set to 48 kHz.

3.3. Measurement results

The results of the measurement and analysis can be seen in Figs. 3 and 4. Floating average filter with the width of 1/6 octave was used to smooth the resulting characteristics [9]. There is a comparison of the outer and the middle segment of the line array source in Fig. 3.

The differences are probably the consequence of different position within the line array source, thus the different diffraction on the edges of the line array source enclosures. Another feature of the sources visible in the analysis is that their directional pattern gets narrower at the frequency about 6 kHz. Different behaviour of the line array source segments can also be seen in Fig. 4. The differences are the most apparent at the highest frequencies.

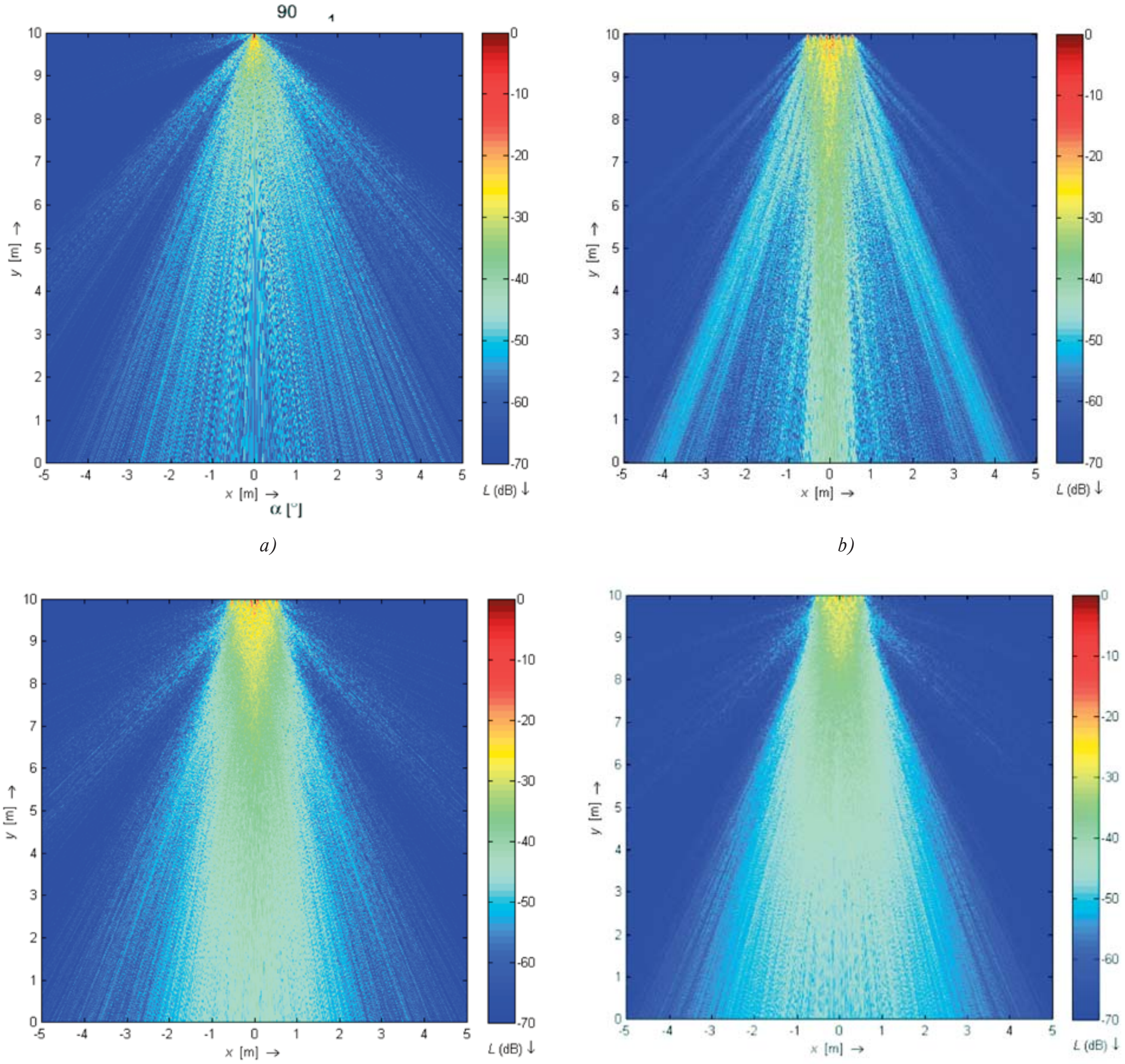
4. Real sound source simulation

Because of the different directional behaviour of the real sources beside the ideal ones, the resulting coverage pattern of the area in front of the line array source is also different. The phase relations between the sources (at the certain point) remain unchanged, though the amplitude ratios between the sources are modified by the individual directional characteristics. Therefore, the measured directional data should be implemented into the simulations with ideal sound sources.

At the certain point of the space it can be done by multiplying the ideal source amplitude by the normalized value of the directional characteristic at given frequency. For the *i*-th ideal point source with applied real source directionality the velocity potential at the certain point X of the space can be calculated using (3) as

$$\Phi_i = \frac{A_i d_i \left(f, \alpha\right) e^{-jkr_i}}{r_i} \qquad (10)$$

where d_i ($d_i \in (0,1)$) is the normalized directional characteristic coefficient corresponding to the given simulated frequency f of the source and the angle α between the acoustic axis of the loudspeaker and the line connecting the speaker reference point and calculated point X. The accuracy of the correct coefficient determination depends on the mentioned resolution of the directional-frequency analysis.



a) *b)* *c)* *d)*

Fig. 5 The simulation of the single source with applied directional characteristic at frequency of 4 kHz (a); The simulation of the line array source with applied directional characteristics at frequency of 6 kHz (b), 1 kHz (c) and 500 Hz (d)

An example of the single source simulation with applied directional characteristic is in Fig. 5 (a). Compared with the ideal circular piston source (see Fig. 1) the source in Fig. 5 (a) has a narrower directional pattern (at the same frequency). Therefore, the directivity of the overall line array source will be higher as can be seen in Fig. 5 (b). However, the directional patterns of the individual sources are not as wide as the ideal ones, thus the ratios of the source amplitudes are higher at the places of sources cancellation (caused by the big phase difference), which means the cancelation is not such dramatic. And the acoustic pressure variance (at this frequency) is also lower. The width of the main beam (the area of the highest acoustic energy) depends on frequency, or on wavelength (higher frequency means narrower beam) and the overall length of the line array source (longer source means narrower beam) [10]. Figure 5 (c) shows the line array source at the frequency of 1 kHz. Its directional pattern is wider than in Fig. 5 (b) due to lower frequency. Fig. 5 (d) demonstrates the further directional pattern widening at the frequency of 500 Hz.

Although, the main result of the simulation is the figure of predicted sound field, we can also compare the results of the simulation of the ideal models of sound sources, the simulation of sound sources with imported directional characteristics and the real measured line array source. Accordingly, the comparison of the directional characteristics in 1 m distance on axis and frequency 1 kHz, measured or simulated with line array sources containing eight segments with 0.15m spacing is shown in Fig. 6.

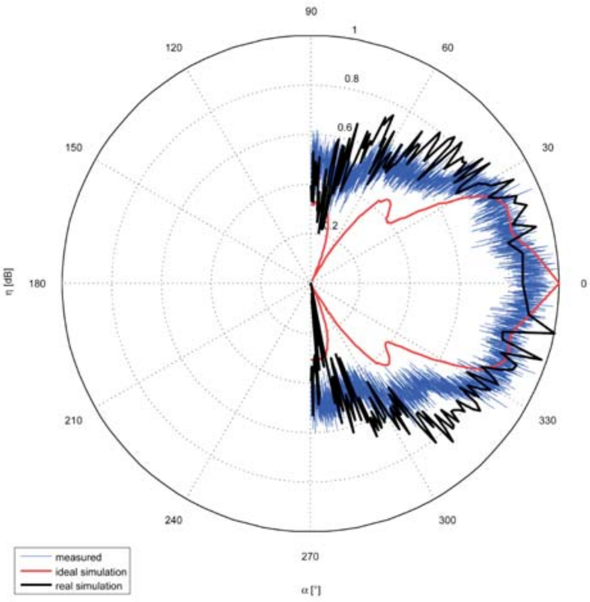


Fig. 6 The comparison of the measured and simulated directional characteristics of line array source containing eight speakers with spacing 0.15 m on 1 kHz

The measured directional characteristic was smoothed with the floating average filter with the width of 1/3. It can be seen that the trace of directional characteristic simulated with the sources with imported real characteristics (black line) is more similar to the measured characteristic (blue line) than the one simulated with the ideal models of sources (red line). However, the simulated characteristics were made with the resolution of 0.01 m, which is not very sufficient especially around 0°, where the number of values is quite small. The higher degree of precision can be reached with even higher resolution, but the computational cost will substantially increase too.

5. Conclusion

The acoustic line array source simulation method was presented utilizing the directional-frequency analysis to implement the directional characteristics of the real measured acoustic sources. Presented method of loudspeaker measurement and directional-frequency analysis allows obtaining the data very quickly as compared with the standard measurement process. Directional characteristics can be obtained as the frequency slices of the directional-frequency analysis while frequency responses can be obtained as the time (angle) slices. Furthermore, various kinds of analysis can be applied on the measured signals and the frequency or directional resolution can be adjusted as necessary. The directional characteristics were implemented into the simulations and the differences between the ideal case and the real sources simulations were shown.

Resulting simulations are useful for the visualization of the acoustic pressure coverage of the area. The function created in Matlab is able to simulate individual sources as well as line array sources or the interaction of any other spatial configuration of multiple sources.

The only drawback is the big computational complexity when simulating higher number of sources with high resolution. The overall time of the simulation of eight sources and the resolution of 0.01 m on Intel Core i7 2800GHz CPU and the RAM memory of 4 GB was approx. 14 minutes, while with ten times higher resolution (0.001 m) the computational time increased to approx. 99 minutes.

Acknowledgement

The described research was performed in laboratories supported by the SIX project; the registration number CZ.1.05/2.1.00/03.0072, the operational program Research and Development for Innovation.

References

[1] THEILE, G., WITTEK, H.: *Wave Field Synthesis – A Promising Spatial Audio Rendering Concept,* to be published in J. of Image Information and Television Engineers, 2007, pp. 77. Online: http://www.irt.de/fileadmin/media/downloads/Produktion/ WaveFieldSynthesis.pdf

[2] RUND, F.: *Simulace zvukoveho pole vice zdroju [Simulation of the Sound Field of Multiple Sources].* Proc. of Technical Computing Prague, Praha : Humusoft, 2009, pp. 5. ISBN 978-80-7080-733-0

[3] BRUNEAU, M.: *Fundamentals of Acoustics.* London : ISTE Ltd, 2006. ISBN-13: 978-1-905209-25-5

[4] SKVOR, Z.: *Akustika a elektroakustika [Acoustics and Electroacoustics].* Praha, Academia, 2001, pp. 527. ISBN 80-200-0464-0

[5] RUZICKOVA, M.: Discrete and Differential Equations in Applied Mathematics. Published in *Communications – Scientific Letters of the University of Zilina,* No. 2, 2012. ISSN 1335-4205.

[6] RAICHEL, D. R.: *The Science and Application of Acoustics.* New York: Springer, 2006. ISBN-13: 978-0387-26062-4

[7] OLSON, H. F.: *Elements of Acoustical Engineering.* New York: D. Van Nostrand Company, Inc., 1940, pp. 368.

[8] MERHAUT, J. et al: *Prirucka elektroakustiky [The Electroacoustic Handbook].* Praha, Statni nakladatelstvi technicke literatury, 1964, pp. 610.

[9] MALACKA, Z.: Pursuit Curves and Ordinary Differential Equations. Published in *Communications – Scientific Letters of the University of Zilina,* No. 1, 2012, ISSN 1335-4205.

[10] ZIOMEK, L. J.: *Foundamentals of Acoustic Field Theory and Space-Time Signal Processing.* Boca Raton : CRC Press, 1995. ISBN: 0-8493-9455-4.

Jakub Sedy – Pavel Silhavy – Ondrej Krajsa – Ondrej Hrouza *

PERFORMANCE ANALYSIS OF TURBO CODES

The article focuses on the performance analyses of Turbo Codes. These codes belong to the group of error correction codes. By their use, it is possible to achieve high system performance. The performance analysis is based on simulations for different code parameters that affect the code gain, bit error rate and computational complexity. Furthermore, it presents the basic structure of the encoder and decoder and the principle of encoding and decoding, which uses the Viterbi algorithm with soft input and soft output.

Keywords: *Turbo encoder, turbo decoder, iterative decoding, SOVA, BER.*

1. Introduction

The basic elements of turbo codes are convolutional codes [1, 2] and decoding algorithms that use soft input and soft output [3, 4, 5, 6]. The input bit sequence is encoded by two encoders, between which is stored interleaver to ensure that the encoded sequences are mutually independent. RSC (Recursive Systematic Convolutional) encoders [1, 2] are often used where each RSC encoder produces a systematic output, which is equivalent to input information, and produces parity bits. Both parity sequences can be punctured before they are transferred with systematic bits to the decoder. Via puncturing it is possible to reduce the number of parity bits to one half and thereby also increase the information rate to 1/2.

For decoding, special algorithms must be used which use soft input and soft output [4, 5, 6]. These soft inputs and outputs do not determine only whether the decoded bit has the logical value 0 or 1 but the likelihood ratio which determines the probability of whether the bit was correctly decoded. The turbo decoder operates iteratively. The first iteration of the first decoder gives an estimate of the original data sequence, based on the soft output channel. It also provides an extrinsic output. Extrinsic output for a given bit is not dependent on the value of the transmission channel for this bit but on the information for the surrounding bits. This extrinsic output from the first decoder is used as a-priori information for the second decoder together with input information from the channel.

The second decoder will give us again extrinsic information and soft output. In the second iteration the extrinsic information from the second decoder in the first iteration is used as a-priori information for the first decoder. Thus the decoder achieves a more accurate estimate of the decoded bits than was the case in the first iteration. This cycle is continuously repeated. In each iteration of the two decoders soft output and extrinsic information are calculated based on the input sequence and a-priori information obtained from extrinsic information of the previous decoder. After each iteration, the BER (Bit Error Rate) decreases.

2. Turbo encoder

The block diagram of turbo encoder [5, 7] is shown in Fig. 1. Two identical encoders are used here, usually RSC, which are separated by interleaver. It is possible to use a structure with more than two encoders, but in this chapter we will deal with the classical structure of two RSC encoders [1, 2].

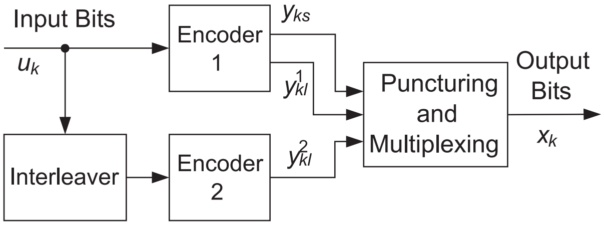


Fig. 1 Turbo encoder block diagram

The input bit sequence is fed to the input of the first encoder where it is encoded. The output of the first encoder is formed by systematic and parity bits. The input bits for the second encoder are interleaved and encoded in the second encoder. The input bits for the second encoder thus become independent of the input bits of the first encoder. Typically, pseudo-random or block interleaver is used. The second encoder produces parity bits only. The output from the two encoders is punctured and then multiplexed. Usually both RSC encoders have an information rate of 1/2 and give one systematic and one parity bit for each input bit. This means that the turbo encoder output sequence contains for each input bit one systematic and two parity bits, i.e. $y_{1s'}, y^1_{1l'}, y^2_{1l'}, y_{2s'}, y^1_{2l'}, y^2_{2l'}, \dots,$

---

* **Jakub Sedy, Pavel Silhavy, Ondrej Krajsa, Ondrej Hrouza**
Faculty of Electrical Engineering and Communication, Brno University of Technology, Czech Republic, E-mail: jakub.sedy@phd.feec.vutbr.cz

$y_{ks'}$, $y^1_{kl'}$, $y^2_{kl'}$. For this output sequence the turbo encoder has an information rate of 1/3. For the total information rate to be 1/2, the output bits from the turbo encoder must be punctured. The output sequence is punctured so that all the systematic bits are preserved and only the parity bits are punctured. Puncturing of the systematic bits will degrade the code performance. After puncturing and multiplexing the turbo encoder output sequence x_{kl} would be $y_{1s'}$, $y^1_{1l'}$, $y^2_{1l'}$, $y_{2s'}$, $y^1_{2l'}$, $y^2_{2l'}$, ..., $y_{ks'}$, $y^1_{k+1s'}$, $y^2_{k+1l'}$.

3. Turbo decoder

A. Soft output Viterbi algorithm

The turbo decoder uses the Viterbi algorithm which is referred to as the SOVA (Soft-Output Viterbi Algorithm [3, 4, 5]. For decoding turbo codes, this algorithm has two modifications. The first modification adapts the path metric so that it takes into account a-priori information when selecting the maximum likelihood paths in the trellis diagram. The second modification of the algorithm consists in soft output in the form of a-posteriori LLR (Log Likelihood Ratio) $L(u_k|\underline{y})$ for each decoded bit.

The first modification considers the state sequence $\underline{s}_k^s$ which gives the states along the surviving paths at the state $S_k = s$ at stage k in the trellis diagram. The metric should be easy to compute via the recursive way where we go from stage $k - 1$ to the kth stage in the trellis diagram. A suitable metric for the path $\underline{s}_k^s$ is defined as [4, 5]:

$$M(\underline{s}_k^s) = M(\underline{s}_{k-1}^{\acute{s}}) + \frac{1}{2}u_k L(u_k) + \frac{L_c}{2}\sum_{l=1}^{n} y_{kl}x_{kl'} \qquad (1)$$

where $M(\underline{\acute{s}}_k^s)$ is the metric of surviving path through the state S_{k-1} at stage $k-1$ in the trellis diagram, u_k is the encoder input bit, x_{kl} is the transmitted channel sequence (output from encoder) associated with a given transition, and y_{kl} is the received sequence from the transmission channel for that transition. Using the transmission channel with BPSK (Binary Phase Shift Keying) modulation and AWGN (Additive White Gaussian Noise), the channel reliability L_c is defined as follows [5]:

$$L_c = 4\alpha\frac{E_b}{2\sigma^2}, \qquad (2)$$

where E_b is the transmitted energy per bit, α is the fading amplitude, and σ is the noise variance.

Now we will discuss the second modification of the algorithm which is the soft output. In a binary trellis diagram there will be two paths reaching the state $S_k = s$ at the stage k. The modified Viterbi algorithm takes a-priori information, calculates the metric of these two paths according to Equation (1) and discards the path with a lower metric. When both paths $\underline{s}_k^s$ and $\underline{\hat{s}}_k^s$ reaching state S_k have the metric $M(\underline{s}_k^s)$ and $M(\underline{\hat{s}}_k^s)$, respectively and the path with the higher metric $\underline{s}_k^s$ is selected as surviving, we define the difference metric Δ_k^s of these paths as [4, 5]:

$$\Delta_k^s = M(\underline{s}_k^s) - M(\underline{\hat{s}}_k^s) \geq 0, \qquad (3)$$

where $M(\underline{s}_k^s)$ is the metric for the surviving path, and $M(\underline{\hat{s}}_k^s)$ is the metric for the discarded path.

When we reach the end of the trellis diagram and find the ML (Maximum Likelihood) path, it is necessary to find the LLR. This determines the reliability of deciding on the bits around the ML path. The Viterbi algorithm shows that all the surviving paths at the stage in the trellis diagram come from the same path a few steps before this stage. This previous stage may attain δ transitions before the stage k, where δ is usually set to five times the constraint length of the convolutional code. Therefore, the bit value u_k associated with the transition from the state $S_{k-1} = \acute{s}$ to the state $S_k = s$ on the ML path may be different when the Viterbi algorithm selects one path merged with the ML path instead of the ML path after the δ transitions, i.e. $k + \delta$ stage in the trellis diagram. If the algorithm selects one of the paths merged with the ML path, it will not affect the value u_k, because this path will differ from the ML path from the transition $S_{k-1} = \acute{s}$ to $S_k = s$. When we calculate the LLR for the bit u_k, SOVA has to take into account the probability of paths merging with the ML path at the stage k to stage $k + \delta$. By comparing the differences in the metric $\Delta_i^{s_i}$ for all states s_i along the ML path from the state $i = k$ to $i = k + \delta$. This LLR is defined as [4, 5]:

$$L(u_k|\underline{y}) \approx u_k \min_{i=k\cdots k+\delta\; u_k\neq u_k^i} \Delta_i^{s_i}, \qquad (4)$$

where u_k is the bit value of the ML path, and u_k^i is the value of the bit of the path that merged with the ML path and was discarded in the state i. The minimization in Equation (4) is only used for paths merging with the ML path which gives a different value for the bit u_k when this path is selected as the surviving path. The paths that gave the same value u_k as the ML path do not affect the decision.

B. Implementation of the SOVA

SOVA is implemented as follows. In every state at every stage in the trellis diagram the metric $M(\underline{s}_k^s)$ is calculated for the two paths merging into the state using Equation (1). The path with the higher metric is chosen as the surviving for this state and the metric indicator stored as the Viterbi algorithm does it. However, in order to provide reliable decoded bits, it also stores the value of $L(u_k|\underline{y})$ calculated by using Equation (4). The metric differences between the surviving and the discarded path are stored together with the binary vector of $\delta + 1$ bits in length, which indicates the sequence of discarded path bits u_k from k back to $k - \delta$ to compare the differences with the surviving path. This series of bits is called the update sequence and is given by output modulo 2 between the previous $\delta + 1$ and the decoded bit along the surviving and the discarded paths. When SOVA identifies the ML path, the update sequences and metric differences along the path are stored and used to calculate the value of $L(u_k|\underline{y})$.

C. Iterative decoding

Now we will describe how iterative decoding works. Fig. 2 shows the schematic of turbo decoder, and it describes the inputs and outputs of individual blocks.

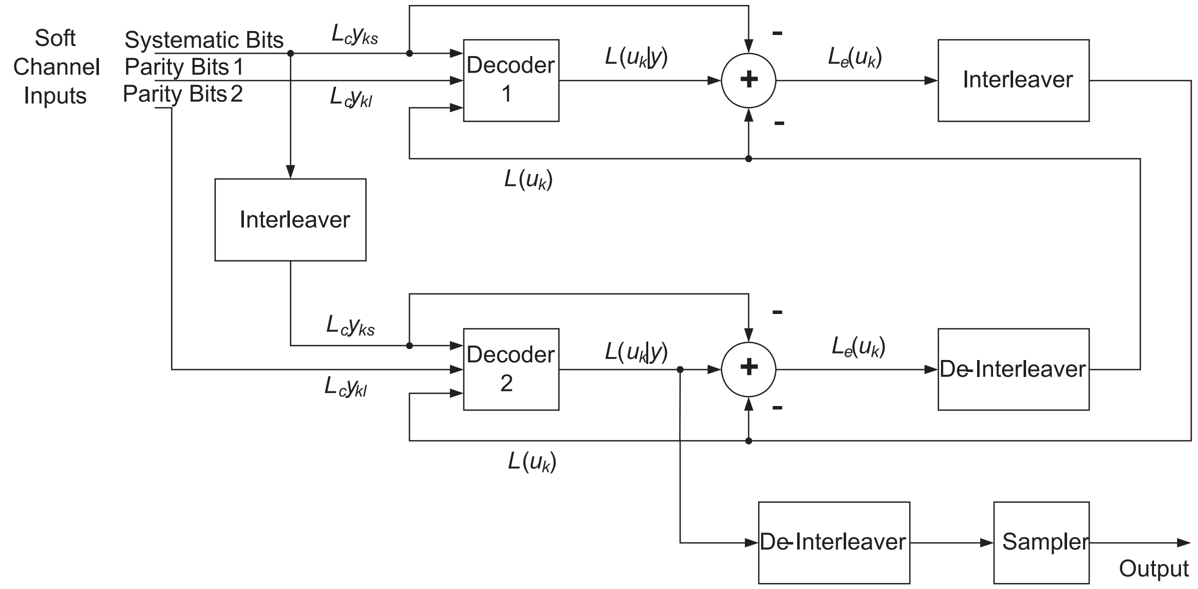


Fig. 2 Turbo decoder schematic

The first decoder in the first iteration receives a sequence $L_c \underline{y}^{(1)}$ from the transmission channel, which includes systematic bits $L_c y_{ks}$ and parity bits $L_c y_{kl}$ from the first encoder. Usually only half of the parity bits are received because these bits have been punctured in the transmitter. The decoder inserts zeros on the punctured places in the soft channel output $L_c y_{ks}$. The first decoder begins processing the soft input from the channel. The output of the first decoder is conditional LLR $L_{11}(u_k \mid \underline{y})$ of data bits u_k, where $k = 1, 2, \ldots N$. The subscript of symbol $L_{11}(u_k \mid \underline{y})$ denotes a-posteriori LLR in the first iteration from the first decoder. In the first iteration the first decoder has no a-priori information about bits, therefore the value of $L(u_k) = 0$, which corresponds to an a-priori probability of 0.5. Now the second decoder begins to operate. It receives the sequence $L_c \underline{y}^{(2)}$ which contains systematic bits for the first decoder which passes through the interleaver and the parity bits from the second encoder. Furthermore, it receives a-priori LLRs $L(u_k)$ which is generated from the conditional LLR $L_{11}(u_k \mid \underline{y})$ from the first decoder. As can be seen from the figure, the extrinsic information $L_s(u_k)$ from the first decoder is adjusted by the interleaver to match with the sequence of input bits entering the second decoder. The second decoder uses this information and the received interleaved sequence $L_c \underline{y}^{(2)}$ to calculate the a-posteriori LLR $L_{12}(u_k \mid \underline{y})$. Now by the equation [4, 5]:

$$L_s(u_k) = L(u_k|\underline{y}) - L(u_k) - L_c y_{ks} \qquad (5)$$

the systematic soft input $L_c y_{ks}$ and a-priori information $L(u_k)$ from the previous decoder are subtracted from the decoder output $L(u_k \mid \underline{y})$. The calculated value is the extrinsic information $L_s(u_k)$ and it is used as a-priori information for the first decoder in the second iteration. This ends the first iteration for both decoders.

In the second iteration the first decoder processes the received sequence $L_c \underline{y}^{(1)}$ again, but now it has available a-priori information which is de-interleaved extrinsic information $L_s(u_k)$ calculated by the second decoder in first iteration from the a-posteriori $L_{12}(u_k \mid \underline{y})$. Now, the first decoder can calculate a more accurate a-posteriori LLR $L_{21}(u_k \mid \underline{y})$. The second iteration continues in the second decoder. It uses the more accurate a-posteriori LLR $L_{21}(u_k \mid \underline{y})$ from the first decoder which calculated more accurate a-priori information $L(u_k)$ by using Equation (5). This information is used together with the received sequence $L_c \underline{y}^{(2)}$ to calculate $L_{22}(u_k \mid \underline{y})$ from which $L_s(u_k)$ is then calculated for the following (first) decoder.

When the series of iterations is completed, the turbo decoder output is given by de-interleaving the a-posteriori LLR $L_{12}(u_k \mid \underline{y})$ of the second decoder where i is the number of iterations used. The signs in a-posteriori sequences give the hard decision output, that is $+1$ or -1.

4. Performance analysis of turbo codes

In this chapter we will present simulations based on the effect of parameters on the performance of turbo codes. The parameters that were used in the simulation are shown in Table 1. Turbo encoder uses two parallel concatenated encoders. Selected as the code was the RSC with generator polynomials $G_0 = 37$, $G_1 = 21$ (octal) and constraint length of code $K = 5$. The interleaver chosen was the pseudo-random interleaver with length $L = 2048$ bits. Unless specified otherwise, puncturing the parity bits to one half will always be used, which will increase the information rate to $R = 1/2$. The decoder uses the SOVA algorithm; usually 8 iterations were used for decoding. The AWGN transmission channel with BPSK modulation is used in the simulation.

Parameters of turbo encoder and decoder Table 1

Channel	AWGN
Modulation	BPSK
Encoders	Two identical RSC
RSC Parameters	$n = 2$, $k = 1$, $K = 5$, $G_0 = 37$, $G_1 = 21$
Puncturing	Half parity bits from each encoder, information rate $R = 1/2$
Decoder	SOVA
Iterations	8
Interleaver	2048-bit pseudo-random interleaver

The performance of turbo codes can be influenced by many parameters. Some of these parameters are:

- The number of decoding iterations used.
- The use of puncturing in encoding.
- The generator polynomials of the codes.
- The frame lengths of input data.

A. Effect of the number of iterations

Fig. 3 shows the performance of turbo codes depending on the number of decoder iterations. Uncoded BER is shown for comparison. The performance after the first iteration of the turbo decoder should be theoretically comparable with the performance of the convolutional code [5]. As the number of iterations increases, the performance of the decoder increases too. For example, the improvement of the performance between the first and second iterations is about 1.2 dB at BER 10^{-4}. This performance increase continues up to the eighth iteration. Code gain between the eighth and fourteenth iteration is only 0.1 dB at BER 10^{-4}. From the figure it is possible to conclude that the increasing number of iterations increases not only the performance of the code but also the computational complexity in decoding; therefore it is recommended to use between 4 and 14 decoder iterations. For this reason, only 8 decoder iterations are used in the following simulations.

B. Effect of puncturing

As already described, the turbo encoder uses two or more encoders which produce parity bits. In these simulations the RSC encoders are used. This is the most common solution which is able to achieve an information rate of below 1/3. In order to achieve an information rate of 1/2, every second parity bit from each encoder must be punctured. It is also possible to use the code without puncturing and thus keep the information rate at 1/3. The performance of unpunctured code is shown in Fig. 4. Encoders use the same parameters as in the previous simulation, Fig. 3. The turbo encoder for unpunctured code has at BER 10^{-4} of a code gain which is 0.5 dB better than the turbo encoder which used puncturing. Very similar gains may also be achieved for different

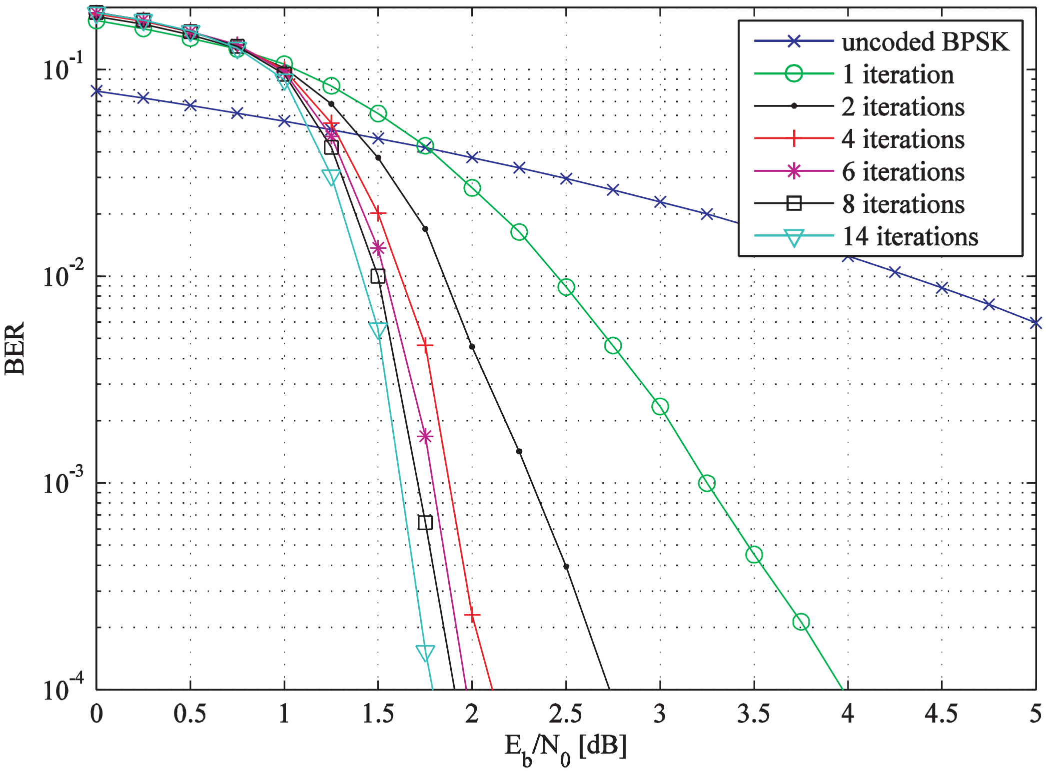


Fig. 3 Turbo coding BER performance using different numbers of iterations

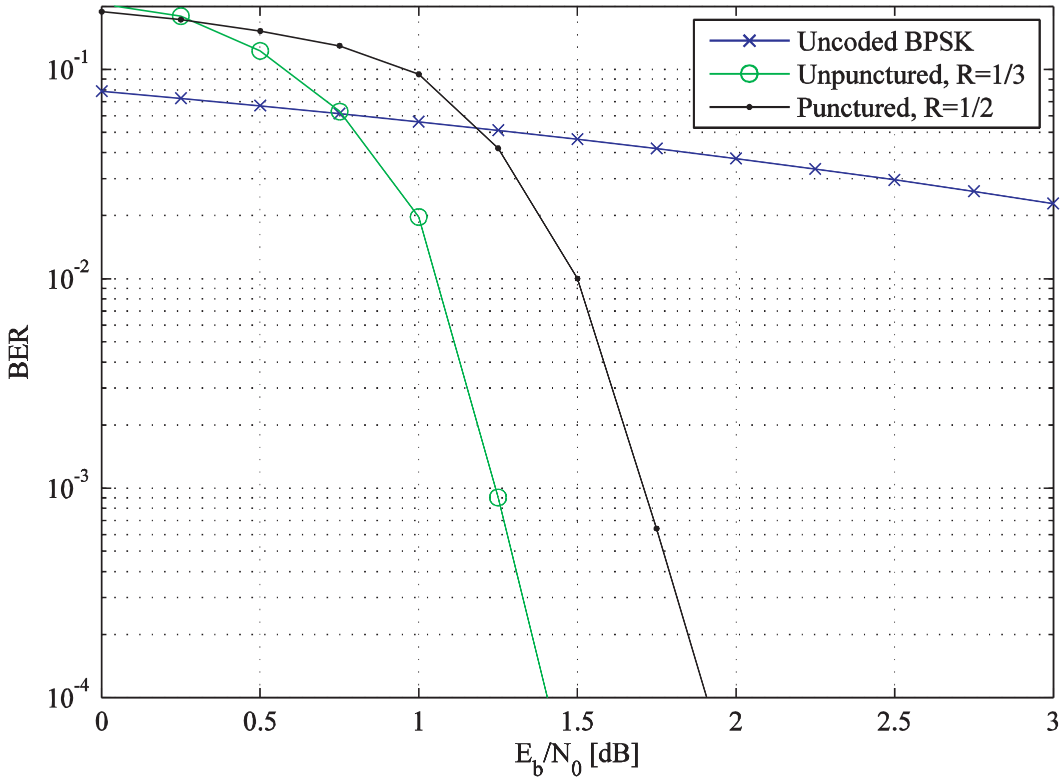


Fig. 4 Comparison of BER performance between punctured and unpunctured turbo codes

generator polynomials. From the figure it is possible to conclude that better results are obtained if puncturing is not used. But this is an improvement in the order of tenths of a decibel. Taking into account the lower number of transmitted bits in the case of puncturing, in some cases it may be considered preferable to use puncturing. Puncturing does not affect the computational complexity of encoding and decoding. It only reduces the number of bits transmitted by a transmission channel and thus reduces the transmission bandwidth.

C. Effect of generator polynomial

Fig. 5 shows the dependence of the performance of turbo convolutional code on the generator polynomial. The first code selected was the RSC code with generator polynomials $G_0 = 7$, $G_1 = 5$ and constraint length $K = 3$. The second code selected was $K = 4$, $G_0 = 17$, $G_1 = 15$. This code achieves performance that is about higher than that achieved by the code with constraint length $K = 3$ at BER of 10^{-4}. The third selected code, which was used for all simulations, has a constraint length $K = 5$ and generator polynomials $G_0 = 37$, $G_1 = 21$. Compared with the first code ($K = 3$), it reaches a performance that is about 0.3 dB higher at a BER of 10^{-4}; in comparison with the code $K = 4$, its performance increases by about 0.125 dB. With increasing constraint length of the code and with greater generator polynomials the performance of turbo codes increases, but what also increases is the size of trellis diagram and thus the computational complexity of decoding.

D. Effect of frame length

Fig. 6 shows the performance of turbo codes depending on the frame length. For many applications, such as applications using real-time transmission, a large frame length is absolutely unacceptable. Frames with a length of 256 bits are useful for voice transmission and 1024 to 2048 bits for video transmission. Systems with larger frame lengths can be used to transfer data and for applications that do not require real-time transmission. The best result in the simulation was reached by a turbo code with a frame length of 65536 bits. The turbo code with a frame length of 65536 bits has a code gain of 0.35 dB compared to turbo codes with a frame length of 2048 bits and 0.6 dB to turbo codes with a frame length of 1024 for BER of 10^{-4}. With growing frame length the performance of turbo convolutional codes increases but the delay gets affected and for shorter frames reaches lower values.

5. Conclusion

This article deals with the problem of turbo codes. It describes a basic structure of turbo encode using two identical RSC codes and turbo decoder which uses Viterbi algorithm. Furthermore, it also presents basic mathematical equations for the SOVA decoding algorithm and describes iterative decoding. Simulations were performed for different parameters of turbo codes. Based on these simulations, it is possible to conclude that the performance of

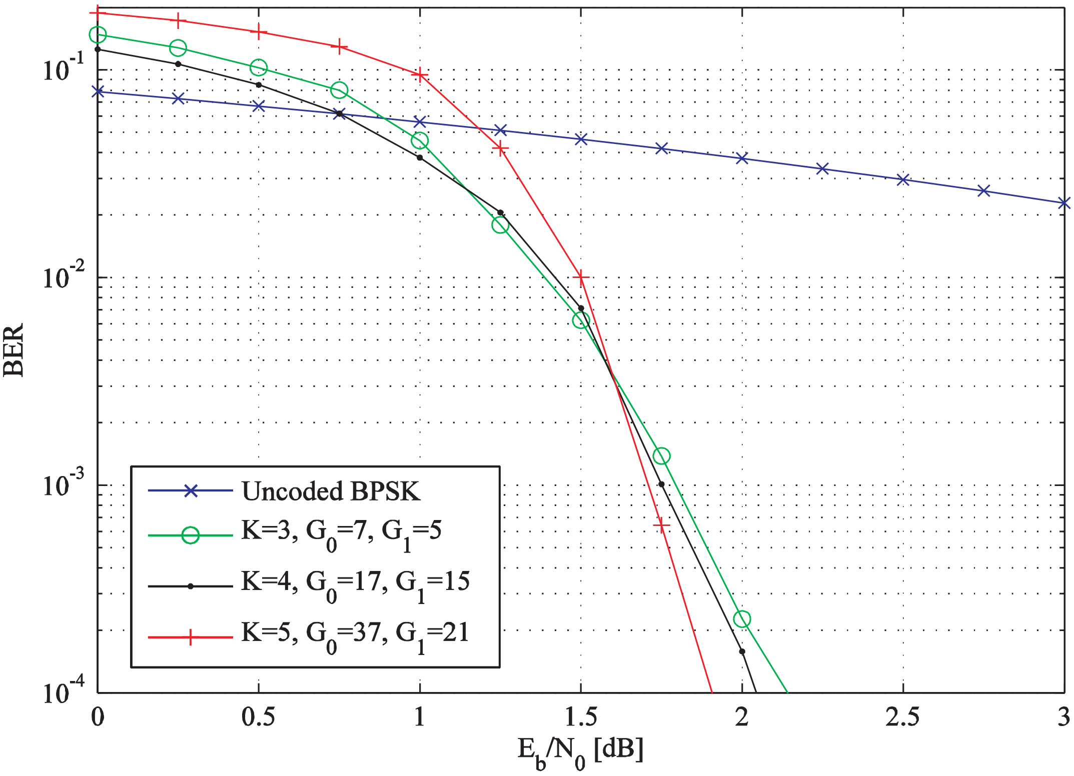


Fig. 5 Effect of constraint length and generator polynomial on the BER performance of turbo coding

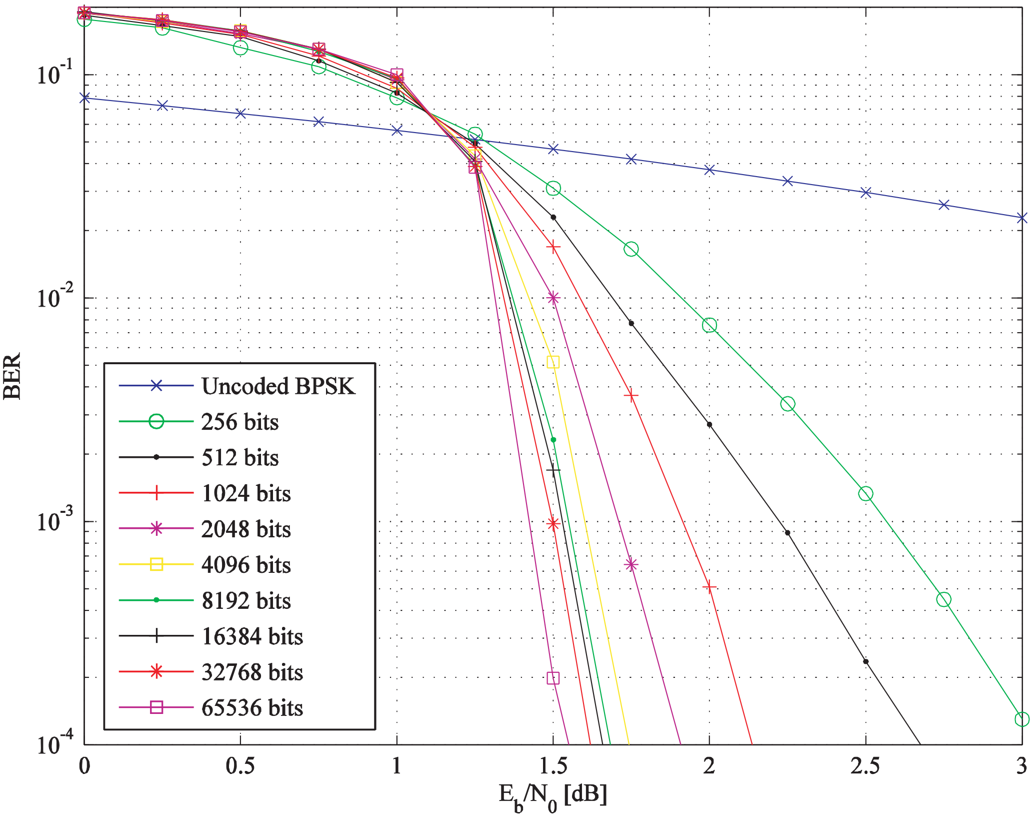


Fig. 6 Effect of frame length on BER performance of turbo coding

turbo codes decreases when puncturing is used. On the contrary, the performance of turbo codes increases with increasing number of the decoding iterations performed by an appropriate choice of the code (generator polynomial) or by changing the frame length. It is possible to implement a high performance codec.

Acknowledgement

The described research was performed in laboratories supported by the SIX project; the registration number CZ.1.05/2.1.00/03.0072, the operational program Research and Development for Innovation and the Technology Agency of Czech Republic project No. TA02020856.

References

[1] LEE, L. H. C.: *Convolutional Coding – Fundamentals and Applications.* Artec House, 1997, ISBN 0-89006-914-X.
[2] LIN, S., COSTELLO, D. J.: *Error Control Coding: Fundamentals and Applications, second edition.* Prentice Hall: Englewood Cliffs, NJ, 2005, ISBN: 0-13-042672-5.
[3] FRANEKOVA, M., NAGY, P.: Ciphering Systems Based on The Error Correcting Coding Techniques, *Communications – Scientific Letters of the University of Zilina,* No. 3, 1999, ISSN 1335-4205
[4] GLAVIEUX, A.: *Channel Coding in Communication Networks: From Theory to Turbo Codes.* Wiley-ISTE, 2007, ISBN: 978-1-90520-924-8.
[5] HANZO, L., LIEW, T. H., YEAP, B. L.: *Turbo Coding, Turbo Equalisation and Space-Time Coding for Transmission over Fading Channels.* John Wiley, 2002, ISBN: 0470847263.
[6] MOON, T. K.: *Error Correction Coding: Mathematical Methods and Algorithms.* Wiley-Interscience, 2005, ISBN-13: 978-0070010697.
[7] FARELL, P. G., MOREIRA, J. C.: *Essentials of Error-Control Coding.* John Wiley, 2006, ISBN-13 978-0-470-02920-6.

Jan Spirik – Jan Zatyik *

IMAGE EXTRAPOLATION USING SPARSE METHODS

Image extrapolation is the specific application in image processing. You have to extrapolate the image for example when you want to process the given image piecewise. When the border patches are incompleted you must extrapolate them to the given size. Nowadays,some basic extrapolations, e.g. linear, polynomial etc. are used. The advanced methods are presented in this paper. We are using the algorithms that are based on finding the sparse solutions in underdetermined systems of linear equations. Three algorithms are presented for image extrapolation. First one is the K-SVD algorithm. K-SVD is the algorithm that trains a dictionary which allows the optimal sparse representation. Second one is Morphological Component Analysis (MCA) which is based on Independent Component Analysis (ICA). The last is the Expectation Maximization (EM) algorithm. This algorithm is statistics-based. These three algorithms for image extrapolation are compared on the real images.

Keywords: *Image extrapolation, sparse, K-SVD, MCA, EM.*

1. Introduction

Image processing using underdetermined systems of linear equations is a very promising approach in different applications. In the previous years methods for image denoising based on image splitting were used [1]. Today, we could denoise the image via underdetermined systems of linear equations [2]. Except image denoising we use these systems for image compression [3], deblurring [4] or image inpainting [5] and for many other applications.

We will use the algorithms originally designed for image inpainting to image extrapolation. Image inpainting is the framework for filling in the known holes in the input image. Classical inpainting methods assume filling the holes from different directions. If we want to use these algorithms for image extrapolation we must redefine or modify some parameters or parts of the original algorithms.

If we are using underdetermined systems of linear equations we want to find the sparse solution. The sparse solution is the one which contains only a few nonzero coefficients. The basics of sparse signal representations are introduced in [3]. All methods for finding sparse solutions start from the basic problem (P_0) which is defined as follows:

$$(P_0)\colon \min_x \|\boldsymbol{x}\|_0 \quad \text{subject to } \boldsymbol{y} = \boldsymbol{Dx} \tag{1}$$

where y is the known signal (e.g. image) we want to reconstruct, $\boldsymbol{D}$ is the dictionary which consists of n atoms ("elementary signals") in columns and m rows which denote the length of atoms. An unknown sparse solution represents amounts of each atom in the original signal. We define the norm of a vector:

$$\|\boldsymbol{x}\|_p = \left(\sum_{i=0}^{N} |\boldsymbol{x}_i|^p\right)^{1/p} \tag{2}$$

and $0 < p < \infty$. If $0 < p < 1$, it is actually not the norm, it is the quasinorm. The quasinorm is similar to the standard norm, but it does not satisfy the triangle inequality.

The optimization problem (1) is defined in ℓ_0. When $\|\boldsymbol{x}\|_0 \ll n$ for $\boldsymbol{x} \in \mathbb{R}^n$, we could say that $\boldsymbol{x}$ is sparse. Searching for the sparse solution without any algorithm in ℓ_0 norm is NP-hard problem. The problem can be redefined to:

$$(P_0)\colon \min_x \|\boldsymbol{x}\|_0 \quad \text{subject to } \|\boldsymbol{Dx} - \boldsymbol{y}\|_2 \le \varepsilon \tag{3}$$

where ε is the error of solution. There are various algorithms dealing with this problem. The overview of these algorithms is presented in [6].

2. Applied algorithms

A) K-SVD algorithm

In most applications via underdetermined systems, we assume fixed dictionary $\boldsymbol{D}$. K-SVD algorithm is the algorithm for adaptive learning the dictionary that allows sparse representation of the input signal. It is called K-SVD because of the SVD (Singular Value Decomposition) algorithm which is always performed K times where K is the count of columns in $\boldsymbol{D}$. K-SVD algorithm is the generalization of the K-means algorithm [7]. K-SVD is the iterative algorithm containing two basic steps. The first one is finding the

---

* **Jan Spirik, Jan Zatyik**
Department of Telecommunications, Brno University of Technology, Brno, Czech Republic,
E-mail: jan.spirik@phd.feec.vutbr.cz, zatyik.jan@phd.feec.vutbr.cz

sparse solution **x** for the actual dictionary. The second step is the dictionary update. Before application of these steps we must initialize the dictionary. The initialized dictionary must have all atoms (all columns) ℓ_2 normalized. We could use, for example, DCT dictionary or otherwise a random matrix for the initial dictionary.

We could formulate the K-SVD as:

$$\min_{\boldsymbol{D},\boldsymbol{X}} \left\{ \left\| \boldsymbol{Y} - \boldsymbol{D}\boldsymbol{X} \right\|_F^2 \right\} \quad \text{subject to } \forall i, \left\| \boldsymbol{x}_i \right\|_0 \leq T_0 \,, \tag{4}$$

where $\boldsymbol{Y}$ is a matrix that contains training samples $\{\boldsymbol{y}_i\}_{i=1}^N$ in columns, $\boldsymbol{X}$ is a matrix that contains the corresponding coefficients, T_0 is the error of the representation and $_F$ denotes the Frobenius norm. The Frobenius norm is defined as:

$$\left\| \boldsymbol{D} \right\|_F = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} \left| d_{i,j} \right|^2} \,, \tag{5}$$

In the first stage we assume $\boldsymbol{D}$ fixed. We can express the penalty term as:

$$\left\| \boldsymbol{Y} - \boldsymbol{D}\boldsymbol{X} \right\|_F^2 = \sum_{i=1}^{N} \left\| \boldsymbol{y}_i - \boldsymbol{D}\boldsymbol{x} \right\|_2^2 \,, \tag{6}$$

Then we could divide (6) into problems:

$$i = 1, 2, ..., N \quad \min_{s_i} \left\{ \left\| \boldsymbol{y}_i - \boldsymbol{D}\boldsymbol{x}_i \right\|_2^2 \right\}$$

$$\text{subject to } \left\| \boldsymbol{x}_i \right\|_0 \leq T_0 \,, \tag{7}$$

These problems could be solved by the known algorithms for finding sparse solutions in underdetermined systems of linear equations.

In the second stage the dictionary $\boldsymbol{D}$ is updated. We assume $\boldsymbol{D}$ and $\boldsymbol{X}$ fixed. In each step only one atom (column in $\boldsymbol{D}$) $\boldsymbol{d}_k$ and corresponding row $\boldsymbol{x}_T^k$ in $\boldsymbol{X}$ will be updated. Based on previous statement we define sets ω_k that consists of indexes of vectors $\{\boldsymbol{y}_i\}$ which use the atom $\boldsymbol{d}_k$, in fact where $\boldsymbol{x}_T^k$ is nonzero:

$$\omega_k = \left\{ i \middle| 1 \leq i \leq N, \boldsymbol{x}_T^k(i) \neq 0 \right\}. \tag{8}$$

Then we define matrices of errors $\boldsymbol{E}_k$ that express error for all N samples with missing k-atom:

$$\boldsymbol{E}_k = \boldsymbol{Y} - \sum_{j \neq k} \boldsymbol{d}_j \boldsymbol{x}_T^j \,. \tag{9}$$

With these conditions we could rewrite (4) as:

$$\begin{aligned} \left\| \boldsymbol{Y} - \boldsymbol{D}\boldsymbol{X} \right\|_F^2 &= \left\| \boldsymbol{Y} - \sum_{j \neq k} \boldsymbol{d}_j \boldsymbol{x}_T^j \right\|_F^2 = \\ &= \left\| \left(\boldsymbol{Y} - \sum_{j \neq k} \boldsymbol{d}_j \boldsymbol{x}_T^j \right) - \boldsymbol{d}_j \boldsymbol{x}_T^k \right\|_F^2 = \left\| \boldsymbol{E}_k - \boldsymbol{d}_j \boldsymbol{x}_T^k \right\|_F^2 \end{aligned} \tag{10}$$

For simplification we apply sets ω_k to the matrices $\boldsymbol{E}_k$ in that manner we choose only columns that correspond with ω_k and we get $\boldsymbol{E}_{k\cdot}^R$. Using the algorithm SVD we divide $\boldsymbol{E}_{k\cdot}^R$ into:

$$\boldsymbol{E}_{k\cdot}^R = \boldsymbol{U}\Delta\boldsymbol{V}^T \tag{11}$$

The last step in the dictionary update stage is only replacing the current atom (column) of the dictionary $\boldsymbol{d}_k$ with the first column of the matrix $\boldsymbol{U}$ and the corresponding row of coefficients $\boldsymbol{x}_T^k$ with the first column of matrix $\boldsymbol{V}$ multiplied by .

Extrapolation via K-SVD is realized as training the dictionary on the image patches with the chosen size (for example 8 × 8, 16 × 16 pixels). The count of atoms in the dictionary is chosen by user too. For the image extrapolation we add random pixels in the place where we want to extrapolate the image. We also define the binary matrix $\boldsymbol{M}$ which extends the original image. Ones in the binary mask represent the known pixels in the image and zeros indicate the missing ones. We can reformulate problem (P_0) as:

$$\min_x \left\| x \right\|_0 \quad \text{subject to } \boldsymbol{MDx} = \boldsymbol{y}. \tag{12}$$

If we perform K-SVD we apply $\boldsymbol{M}$ to the dictionary $\boldsymbol{D}$, but in the extrapolated part of the image we use the whole trained dictionary. For better results we completely overlap all the patches from the image [3].

One of the biggest disadvantages of K-SVD algorithm is the dependency on the length of extrapolation with the patch size. We must train the dictionary for the patches of minimum $(e + 1) \times$ $\times$ $(e + 1)$ pixels, where e is the length of extrapolation in pixels. It means that the computational efficiency is dependent quadratically on the length of extrapolation. This patch size is chosen because of the overlapping patches. If we do not fulfill this limit, the extrapolation will be unsuccessful, the end of the extrapolated part will be black.

On the other hand, the big advantage of K-SVD extrapolation is that the training part for one image is done only once. You can extrapolate the image to different directions, but the trained dictionary remains the same. Only the reconstruction will be done. For example, if you train the dictionary for patches 8 × 8 pixels, you should perform extrapolation to any directions with the extrapolation length of maximum 7 pixels.

B) Morphological Component Analysis

The basics of Morphological Component Analysis (MCA) are introduced in [8]. We assume that the input picture is a linear combination of two independent parts: cartoon and texture. This idea comes from Independent Component Analysis (ICA). We use two incoherent dictionaries. The dictionary $\boldsymbol{D}_t$ allows sparse decomposition of the part of the image $\boldsymbol{y}_t$ (we assume the picture as 1-D vector) that contains only texture. On the opposite side, the dictionary $\boldsymbol{D}_c$ allows sparse decomposition of the part of the image $\boldsymbol{y}_c$ that contains only cartoon. We could formulate the problem as:

$$\left\{ \boldsymbol{x}_t^{opt}, \boldsymbol{x}_c^{opt} \right\} = \arg \min_{\{x_t, x_c\}} \left\| \boldsymbol{x}_t \right\|_0 + \left\| \boldsymbol{x}_c \right\|_0 \quad \text{subject to}$$

$$\boldsymbol{y} = \boldsymbol{D}_t \boldsymbol{x}_t + \boldsymbol{D}_c \boldsymbol{x}_c \,. \tag{13}$$

Because of the problem for finding the numerical solution we must reformulate equation (13) as:

$$\{\boldsymbol{x}_t^{opt},\boldsymbol{x}_c^{opt}\} = \arg\min_{\{x_t,\,x_c\}}\|\boldsymbol{x}_t\|_1 + \|\boldsymbol{x}_c\|_1 + + \lambda\|\boldsymbol{y} - \boldsymbol{D}_t\boldsymbol{x}_t - \boldsymbol{D}_c\boldsymbol{x}_c\|_2^2 + \gamma TV\{\boldsymbol{D}_c\boldsymbol{x}_c\}, \quad (14)$$

where the third term of the equation reflects the reconstruction error and TV is the abbreviation for total variation penalty. We use the TV for recovering piecewise smooth objects with pronounced edges, when applied to the cartoon layer [9]. The total variation is essentially the ℓ_1 norm of the gradient.

For the purpose of image extrapolation, we must apply the binary mask $\boldsymbol{M}$ the same way as for K-SVD to the image:

$$\{\boldsymbol{x}_t^{opt},\boldsymbol{x}_c^{opt}\} = \arg\min_{\{x_t,\,x_c\}}\|\boldsymbol{x}_t\|_1 + \|\boldsymbol{x}_c\|_1 + + \lambda\|\boldsymbol{M}(\boldsymbol{y} - \boldsymbol{D}_t\boldsymbol{x}_t - \boldsymbol{D}_c\boldsymbol{x}_c)\|_2^2 + \gamma TV\{\boldsymbol{D}_c\boldsymbol{x}_c\}. \quad (15)$$

Ones in the binary mask express the known pixels in the image and zeros indicate the missing ones.

If we assume that $\boldsymbol{y}_t = \boldsymbol{D}_t\boldsymbol{x}_t$ and we know the texture part of the image $\boldsymbol{y}_t$, we can calcute the sparse vector $\boldsymbol{x}_t$ as $\boldsymbol{x}_t = \boldsymbol{D}_t^+\boldsymbol{x}_t + \boldsymbol{r}_t$, where $\boldsymbol{r}_t$ is a residual vector in the null-space of $\boldsymbol{D}_t$ and $^+$ denotes Moore-Penrose pseudoinverse. We apply the same properties to the cartoon part. For simplification we assume $\boldsymbol{r}_t = \boldsymbol{r}_c = 0$. Then we could minimize the problem as:

$$\min_{\{y_t,y_c\}}\|\boldsymbol{D}_t^+\boldsymbol{y}_t\|_1 + \|\boldsymbol{D}_c^+\boldsymbol{y}_c\|_1 + + \lambda\|\boldsymbol{M}(\boldsymbol{y} - \boldsymbol{y}_t - \boldsymbol{y}_c)\|_2^2 + \gamma TV\{\boldsymbol{y}_c\}. \quad (16)$$

When we implement the algorithm we first choose the threshold factor, number of iterations and the parameters λ and γ. The last two parameters should be constant during the iterations or they could be descended. Then we initialize the cartoon part of the image as $\boldsymbol{y}_c = \boldsymbol{y}$ and the texture part as $\boldsymbol{y}_t = 0$. After that we iterate the algorithm to the stopping rule (threshold or number of iterations). In the iteration part of the algorithm we first fix the texture part of the image $\boldsymbol{y}_t$ and we update the cartoon part $\boldsymbol{y}_c$ and then vice-versa. After these two stages we apply the TV penalization.

The important step in the algorithm is the choice of the dictionaries. For texture part we should use local DCT, Gabor or wavelet packets transforms and for the cartoon part wavelet, curvelet, ridgelet, contourlet and many other transforms [9]. MCA extrapolation significantly depends on the choice of dictionaries. For different types of pictures (real-life pictures, cartoons, computer images) a different combination of dictionaries is more vital. The extrapolation is performed for every single direction separately. You cannot do any temporary calculations. You only perform the reconstructions phase, there is no training phase as for K-SVD. All the conditions for successful extrapolation are included in the algorithm.

C) Expectation-Maximization algorithm

Expectation-Maximization algorithm (EM) is based on mathematical statistics. It is iterative method for estimation of maximum-likelihood. In this case we use this method for finding sparse solutions based on penalized maximum-likelihood estimator estimation [5]. In terms of statistics inpainting is a problem of estimation from incomplete data sets. EM algorithm is used for computation of sparse vector x from the previous iteration. We reconstruct the whole image, not only the missing parts. As well as MCA the efficiency of EM depends on the choice of the dictionary. One of the biggest advantages is that we could make a dictionary from several different transformations, i.e. utilize underdetermination of the system.

At first we must define the penalized maximum-likelihood estimator [5]:

$$\hat{\boldsymbol{x}} = \arg\min_{x}\frac{1}{2\sigma^2}\|\boldsymbol{Y} - \boldsymbol{D}\boldsymbol{x}\|_2^2 + \lambda\Psi(\boldsymbol{x}) \quad (17)$$

where σ^2 is a variance of zero-mean additive white Gaussian noise, $\lambda > 0$ is regularization parameter and Ψ is a penalty function. The penalty function must be non-negative, continuous, even-symmetric and non-decreasing function. But it must not be necessarily convex in $\mathbb{R}^+$. We often use ℓ_1-norm penalty for Ψ. We divide the input image into two sets: $\boldsymbol{y}_0$ which contains the known pixels from the image and $\boldsymbol{y}_m$ which contains the unknown pixels from the image. The incomplete observation makes impossible calculate (17) at once. We use EM algorithm for iterative reconstruction of missing data so that we use (17) for computation of new estimations until the convergence is achieved.

As already mentioned above the EM algorithm is iterative process. It consists of two steps: Expectation (E) step and Maximization (M) step. In the first step the conditional expectation of the penalized log-likelihood of complete data y_0 and actual parameters is calculated. The particular expectations can be expressed as the conditional expected squared values of the missing data:

$$\boldsymbol{y}_i^t = \exists\left[y_i \middle| \boldsymbol{D}, \boldsymbol{y}_0, \boldsymbol{x}^t, \sigma^{2t}\right] = = \begin{cases} \boldsymbol{y}_i, & for\ observed\ data, i \in \boldsymbol{I}_0 \\ (\boldsymbol{D}\boldsymbol{x}^t)_i & for\ mis\sin g\ data, i \in \boldsymbol{I}_m \end{cases} \quad (18)$$

$$\exists\left[\boldsymbol{y}_t^2 \middle| \boldsymbol{D}, \boldsymbol{y}_0, \boldsymbol{x}^t, \sigma^{2t}\right] = = \begin{cases} \boldsymbol{y}_i^2, & for\ observed\ data, i \in \boldsymbol{I}_0 \\ (\boldsymbol{D}\boldsymbol{x}^t)_i^2 + \sigma^{2t} & for\ mis\sin g\ data, i \in \boldsymbol{I}_m \end{cases} \quad (19)$$

We could express the estimation at t iteration as:

$$\boldsymbol{y}^t = \boldsymbol{y}_{act} + (\boldsymbol{I} - \boldsymbol{M})\boldsymbol{D}\boldsymbol{x}^t = \boldsymbol{D}\boldsymbol{x}^t + (\boldsymbol{y}_{act} - \boldsymbol{M}\boldsymbol{D}\boldsymbol{x}^t) \quad (20)$$

where $\boldsymbol{y}_{act} = \boldsymbol{M}\boldsymbol{y}$ and $\boldsymbol{M}$ is the binary mask with the same properties as above.

In $\boldsymbol{M}$ step we maximize the penalized function with the missing observations using the estimates from the $\boldsymbol{E}$ step at t iteration:

$$\sigma^{2t+1} = \frac{1}{n}\left[\sum_{i\in I_0}\left(\boldsymbol{y}_i - \left(\boldsymbol{D}(\boldsymbol{x})^t\right)_i\right)^2 + (n - n_0\sigma^{2t})\right], \quad (21)$$

where n_0 is the number of observed pixels.

These two steps are repeated *t* times until the convergence is reached:

$$\hat{\sigma}^2 \rightarrow \frac{1}{n_0}\sum_{i \in I_0}\left(\boldsymbol{y}_i - \left(\boldsymbol{D}\left(\hat{\boldsymbol{x}}\right)^t\right)_i\right)^2, \tag{22}$$

that is the maximum-likelihood estimate of the noise variance inside the mask with the observed pixels. The properties of convergence depend on the structure of the dictionary. The algorithm behaves differently when the dictionary is the basis, tight frame or union of several incoherent orthogonal dictionaries [5].

The EM extrapolation is in properties very similar to MCA extrapolation. There is no training phase. The result depends on the dictionary if it is basis or union of basis or etc.

The biggest disadvantage of EM algorithm is that the reconstruction is performed on the whole image, not only on planned parts. It means the extrapolated image is blurred. The blur effect depends on the length of extrapolation. You can improve the EM reconstruction result by some technique presented in [4].

3. Experimental results

We choose the excerpt of the Barbara image (Fig. 1) to compare the efficiency of the presented algorithms. The algorithms were compared for different length and type of extrapolation. For the purposes of experimental measurements, the binary mask ***M*** to the known part of the image was applied. It is because of possibility to

Fig. 1 Used part of Barbara image 256 × 256 pixels

(a) (b) (c) (d)

Fig. 2 Image extrapolation of 7 pixels: (a) original image with applied mask, (b) reconstruction via K-SVD: PSNR 36.3 dB, SSIM 0.9981, (c) reconstruction via MCA: PSNR 29.0 dB, SSIM 0.9946, (d) reconstruction via EM: PSNR 27.7 dB, SSIM 0.9445

(a) (b) (c) (d)

Fig. 3 Image extrapolation of 15 pixels: (a) original image with applied mask, (b) reconstruction via K-SVD: PSNR 29.5 dB, SSIM 0.9846, (c) reconstruction via MCA: PSNR 25.3 dB, SSIM 0.9765, (d) reconstruction via EM: PSNR 24.4 dB, SSIM 0.9264

Fig. 4 Image extrapolation of 25 pixels: (a) original image with applied mask, (b) reconstruction via K-SVD: PSNR 25.6 dB, SSIM 0.9612, (c) reconstruction via MCA: PSNR 21.9 dB, SSIM 0.9428, (d) reconstruction via EM: PSNR 21.9 dB, SSIM 0.8984

Fig. 5 Image extrapolation of 25 pixels into 2 directions: (a) original image with applied mask, (b) reconstruction via K-SVD: PSNR 22.1 dB, SSIM 0.9045, (c) reconstruction via MCA: PSNR 20.9 dB, SSIM 0.8981, (d) reconstruction via EM: PSNR 21.4 dB, SSIM 0.8541

measure the efficiency of the algorithms. The binary mask will be always rectangular. We choose the length e of extrapolation and then which direction we want to extrapolate. We could make extrapolation for more directions at once. We fill in e columns or rows (it depends on the directions of extrapolation) with zeros in $\boldsymbol{M}$.

The standard PSNR and SSIM are applied for measurement reconstruction quality. We are introducing some examples of extrapolation in the following figures (Figs. 2 – 5). For all extrapolations using MCA and EM algorithms two incoherent dictionaries were used: curvelets [10] and undecimated discrete wavelet transform [11].

4. Conclusion

The principles of three different algorithms for image extrapolation that are using sparse solutions have been presented. The results for different lengths and directions of extrapolation were shown in the figures. Our subjective perception of the quality of the reconstructed images corresponded with the objective measurement of quality: PSNR and SSIM. The best algorithm from the presented algorithms was K-SVD algorithm. The quality was especially in good reconstruction of the texture parts in the image. There is also only a small blur effect. Some modifications can be made in the original K-SVD, for example, non-square patches and so on. These and more modifications will be presented in the future work.

Acknowledgement

The described research was performed in laboratories supported by the SIX project; the registration number CZ.1.05/2.1.00/03.0072, the operational program Research and Development for Innovation.

References

[1] SIMAK, B., KARPF, M.: Salt & Pepper Noise Impact on Image Coding Based on Image Splitting, *Communications - Scientific Letters of the University of Zilina*, vol. 2, pp. 16-20, 2000.

[2] ELAD, M., AHARON, M.: Image Denoising Via Sparse and Redundant Representations Over Learned Dictionaries, *Image Processing, IEEE Transactions on,* vol. 15, no. 12, pp. 3736-3745, dec. 2006.

[3] ELAD, M.: *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing, Springer,* 2010.

[4] ZHANG, H., ZHANG, Y.: Sparse Representation Based Iterative Incremental Image Deblurring, in *Proceedings of the 16th IEEE intern. conference on Image processing,* Piscataway, 2009, ICIP'09, pp. 1285-1288, IEEE Press.

[5] FADILI, M. J., STARCK, J. L., MURTAGH, F.: Inpainting and Zooming Using Sparse Representations, *The Computer Journal,* 2007.

[6] SPIRIK, J.: Algorithms for Computing Sparse Solutions, *in Proc. of the 17th conference STUDENT EEICT vol. 3,* 2011, pp. 123-127.

[7] AHARON, M., ELAD, M., BRUCKSTEIN, A. M.: K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representations, *IEEE Transactions on Signal Processing,* vol. 54, pp. 4311-4322, 2006.

[8] STARCK, J. L., ELAD, M., DONOHO, D. L.: Redundant Multiscale Transforms and Their Application for Morphological Component Analysis, *Advances in Imaging and Electron Physics,* vol. 132, 2004.

[9] ELAD, M., STARCK, J. L., QUERRE P., DONOHO, D. L.: Simultaneous Cartoon and Texture Image Inpainting Using Morphological Component Analysis (MCA), *Applied and Computational Harmonic Analysis,* vol. 19, no. 3, pp. 340-358, 2005.

[10] CANDES, E., DEMANET, L., DONOHO, D. L., YING, L.: Fast Discrete Curvelet Transforms, *Multiscla Modeling & Simulation,* vol. 5, no. 3, pp. 861-899, 2006.

[11] GYAOUROVA, A., KAMATH, CH., FODOR, I. K.: Undecimated Wavelet Transforms for Image De-Noising, *Tech. Rep.*, 2002.

Filip Rezac – Jakub Safarik – Miroslav Voznak – Jan Rozhon – Karel Tomala – Jiri Vychodil *

BRUTEFORCE ATTACKS BLOCKING SOLUTION ON EMBEDDED SIP COMMUNICATION SERVER

This article deals with embedded SIP communication server with an easy integration into the computer network based on open source solutions and its effective defense against the most frequent attack in the present – Denial of Service. The article contains brief introduction into the Bright Embedded Solution for IP Telephony – BESIP and describes the most common types of DoS attacks, which are applied on SIP elements of the VoIP infrastructure including the results of defensive mechanism that has been designed.

Keywords: *BESIP, SIP, IP Telephony, DoS, IPS, Security, Embedded Solution.*

1. Introduction

Many large institutions operate small offices with tens or hundreds of employees. A common requirement is the full integration of these departments in the organization's environment (examples are libraries and branch offices). With our proposed solution BESIP (Bright Embedded Solution for IP telephony), the integration can be achieved easily with the use of IP telephony and supporting network infrastructure. The device is designed as a price acceptable solution that supports SIP (Session Initiation Protocol) IP telephony and also services such as ENUM (E164 NUmber Mapping) [1], secure communication using SRTP (Secure Realtime Transfer Protocol) and TLS (Transport Layer Security) [2], monitoring of call quality, tools for attacks detection, billing and clear configuration via a web interface. The whole solution is deeply described in the paper "Embedded multiplatform SIP server solution" [3].

The system consists of software PBX Asterisk [4] and a part of BESIP is also a module responsible for the safety. Today one of the most common attacks against these types of network elements is Denial of Service – DoS. It is because of high efficiency and relatively simple feasibility. It was therefore necessary to develop methods for security which can be used not only as part of our system, but also as a general solution for Asterisk.

The following chapters refer to the scheme of the system in more details, the vulnerability of Asterisk SIP proxy servers to DoS attacks and methods for server protection. For each attack, this paper describes their impact on a SIP server, evaluation of the threat and the way in which they are executed.

2. BESIP System Schemes and Modules

As mentioned above, the BESIP system is a modular solution where each element consists of several applications which are supported by core. (Fig.1). Modules are divided according to the function which they perform at Core, Security Module, Monitoring Module, PBX Module, and Module of Services.

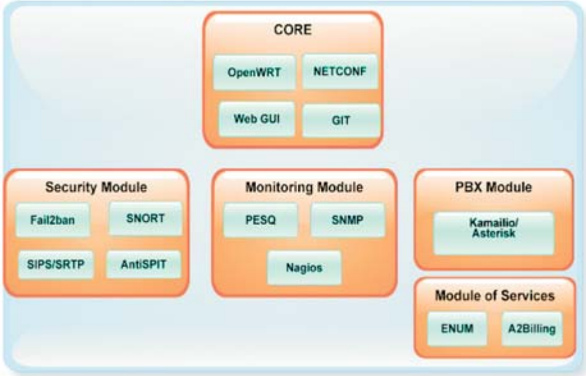


Fig. 1 BESIP Divided into Modules

2.1. Core

The core of the system consists of the Linux distribution OpenWRT [5] which is directly designed for embedded devices and has very low demands on computing power.

---

* **Filip Rezac, Jakub Safarik, Miroslav Voznak, Jan Rozhon, Karel Tomala, Jiri Vychodil**
Department of Telecommunications, Faculty of Electrical Engineering and Computer Science, VSB-TU Ostrava, Czech Republic,
E-mail: fillip.rezac@vsb.cz

To manage updates and revisions, Git [6] application is used. Another tool that is part of the core is the NETCONF protocol [7]. NETCONF provides mechanisms to install, manipulate, and delete the configuration of network devices.

The last part of the Core Module is Web GUI support. This is done using Lighttpd [8] which has a small memory demands and is therefore suitable for embedded devices.

2.2. Security Module

Security module is responsible for protecting the system itself against attacks from external subjects, as well as the analysis of these threats.The module is responsible also for signaling and media encryption.

The protection of system against threats is provided by Fail2ban [9] application. It is a tool which is able to block IP addresses in the firewall based on the logs scan. Another way to protect the system against attacks directed at an IP telephony service is the implementation of Snort and IPS (Intrusion Protection System) [10]. Snort and its sub-applications are the main defense used to protect against DoS, as described below. Protection against Spam over Internet Telephony – SPIT attacks is solved by the AntiSPIT [10] tool which was also developed as an original solution by the authors of article. The last element of the security module is the ability to encrypt calls using SRTP and TLS protocol. This security is ensured directly by communication server, in our case, the SIP PBX Asterisk in version 1.8.4.4 [10].

2.3. Monitoring Module

This module is able to monitor the speech quality for individual IP calls, as well as provide other monitoring of network devices in the network using SNMP (Simple Network Management Protocol) and Nagios tools [11].

2.4. PBX Module

This module is one of the most important of the entire system, as it contains the actual communication server for IP telephony calls. All other modules are interconnected with Asterisk because instead of sending or retrieving data, Asterisk is also responsible for encrypting and comparing call quality using algorithm which is based on PESQ (Perceptual Evaluation Speech Quality) method [9].

2.5. Module of Services

Module of services contains tools for providing additional services, such as billing or ENUM. End user can also define additional services that he/she needs, but above mentioned are part of the system by default.

2.6. Hardware

Since the beginning of the development, the BESIP was planned as the most mobile, portable and especially low cost device. Outside of these conditions it also had to offer sufficient computing power for smooth operation of all modules, applications and participants.

After a series of tests and analyses a standard desktop PC with Atom processor was chosen. It consists of the Intel Packton D410PT setwith the following configuration: CPU: x86 Intel Atom D410 - 1,66 GHz, chipset: Intel NM10 Express, NIC: Realtek 10/100 Mbps, USB 2.0: 8 ports, max. RAM: 4 GB, memory: Kingston 1GB 667MHz CL 5, HDD: Kingston 16GB SSD, interface: SATA 3 Gb/s 2,5" , case: Eurocase Mini ITX Wi-05, Size: 265 × 90 × 270 mm, power Supply: 200 W, number of 2,5" positions: 1.

3. Classification of the DoS Attacks

Denial of service can be achieved in several ways – flooding a server with malformed, damaged or useless packets as a result of which the server runs out of its resource capacity. The affected server is then unable to communicate with its regular users or process regular requests. DoS attacks can be divided into three general classes [12, 13] - Flooding Attacks which are targeting on server resources (CPU, memory or link capacity), Misuse Attacks specified by the hacker uses of a modified SIP message to cancel or redirect calls or misuses of the service and Unintentional Attacks where the attacker targets the supporting services (DNS, call billing, etc.) in order to distort or restrict the service. These attacks typically affect a small group of users only.

The impact of a DoS attack depends on the target. Targeting a particular client can lead to denying the service to this user only but when a SIP server such as BESIP is the target, no user can use its services.

4. BESIP Security Technology Used

Attacks against the embedded systems are more dangerous due to their relatively lower performance which makes the attacks more efficient. That is because we tried to use an effective secure solution in BESIP system. We chose an IPS system, consisting of three applications.

4.1. Snort

The core of the entire IPS solution is IDS (Intrusion Detection System) system Snort which detects malicious activity in the network. The detection is based on signatures or detection of anomalies. The whole IDS system is modular. The most important components are Packet Decoder that captures packets from network interfaces, prepares them for pre-processing. Pre-processor is responsible for processing or modification of the packets before processing (packet Defragmentation, URI decoding, reassembling TCP streams, etc.). Other modules are also important. Detection Engine is responsible for attack detection, Logging and Alerting System is linked to the Detection Engine and is used to log the activity or generate an alert. Plenty of plugins and extensions that enhance its features are also available for Snort.

4.2. SnortSam

This application operates on the client-server model. It allows Snort to dynamically intervene into iptables rules. To ensure its proper operation, we need to first upgrade our Snort installation with a SnortSam plugin. The user communicates with the Snort's sensor, sends commands to the server (where incident has been detected). The server listens on port 898, applying information from clients to iptables rules (see Fig. 2). Iptables is an open-source firewall for Linux-based operation systems. It is used to block malicious traffic on a server.

SnortSam messages are transferred as encrypted. A whitelist of non blockable IP addresses is also available. The detected traffic is then blocked for some time. Once the attack is over and timed out, the blocked IP is allowed to communicate again. Thus, only malicious traffic that poses a threat to our server is blocked.

5. Results

We created a testing topology to measure DoS and security solution effectiveness. It contained a BESIP system, hacker's PC and some endpoint devices registered on BESIP.

The malicious tools applied were as follows: Sipp, Inviteflood, Udpflood, Flood2, Juno.

5.1. Attacks on BESIP's CPU Using Sipp

The Sipp programme is primarily used to simulate calls and to carry out SIP proxy stress tests. But with a simple upgrade of the call scenarios, it can make malicious calls on SIP proxy. These calls are intended to overload system's CPU. Figure 3 shows the impact of these attacks on the BESIP. The attack scenario applied was the same for each attack. Sending malicious packets started in 10 s and continued for 60 s. Other 30 s shows the time for which the system is still inhibited by the attack.

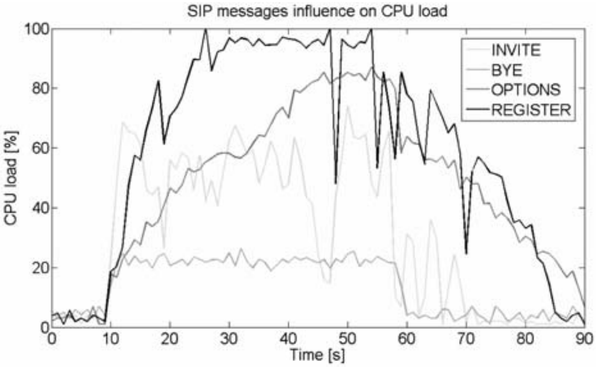


Fig. 3 The impact of different attack SIP message types on a BESIP's CPU load

To enable the comparison of the efficiency of individual malicious SIP messages, the messages had been sent to the SIP server with the same rate (250 messages per second). Clearly, the most effective SIP messages to attack a SIP server are REGISTER and OPTIONS. In the first case, the endpoint could not register or make calls, though running calls was not affected (the RTP stream only between endpoints). OPTIONS flood caused merely a delay in request processing, yet the situation deteriorated as the attack continued. In the end, not a single endpoint was able to register or make calls. The relatively long time necessary for the BESIP to recover (in both cases) was rather surprising.

The delay in connection was evident in the attack performed by means of INVITE messages. Some calls failed to be connected at all. The attack was performed by a non-existing source user.

Attacks performed by means of BYE, CANCEL and ACK messages returned almost the same results (the figure illustrates only the attack by means of the BYE message). During the attack, no call or registration was affected. BYE and CANCEL were not sent to end a particular call.

Security precautions against all these attacks include Snort rules tracking the number of messages sent to the SIP server from a particular source address. Where the limit for messages was exceeded, the blocking rule was activated on the firewall. The CPU load with the activated IPS system was about 9% during these attacks (Fig. 4.).

```
alert udp $EXTERNAL_NET any -> $SIP_PROXY $SIP_PORT (msg:"SIP DoS
attempt(registerflood)"; content:"REGISTER sip";detection_filter:track by_src,
count 50, seconds 5; classtype:misc-attack; sid:1000001; rev:1; fwsam:src, 5min;)

alert tcp $EXTERNAL_NET any -> $SIP_PROXY $SIP_PORT (msg:"SIP DoS
attempt(registerflood tcp)"; content:"REGISTER sip"; detection_filter:track by_src,
count 50, seconds 5; classtype:misc-attack; sid:1000007; rev:1; fwsam:src, 5min;)
```

Fig. 2 The example of the Snort rules - tracking the number of SIP REGISTER messages from one source

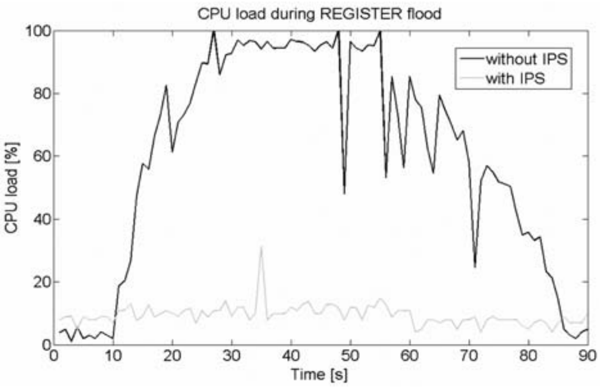


Fig. 4 The impact of an attack with (SSI) and without the protection

The attacker could be sending all the above mentioned malicious messages at a higher rate. In this way each malicious message can consume up to 100% of the BESIP's CPU. Just to compare, the INVITE messages need 10 times higher rate than the REGISTER messages to consume a similar load of the affected machines CPU. The INVITE messages can also send the inviteflood application and create a situation very similar to the flooding with UDP packets (the same is true for any attack with a high rate of packets sent).

5.2. Link Flooding Attacks

Unlike the above mentioned attacks, udpflood only floods the target destination with useless UDP packets. These packets contain a sequence from 1 to 9, followed by zeros. The packet size is 1400 bytes, and the tool can spoof the source address.

The CPU load is very low during the attack but all communication with the BESIP is blocked due to a high volume of traffic. Blocking the traffic on BESIP's interface is useless as the link would still be flooded. There is no efficient protection to be applied on the system, it is only possible to eliminate the impact of such an attack.

5.3. TCP SYN Flood Attacks

The last type of attack against BESIP tested was to flood it with TCP SYN flag set packets. We used flood2 and Juno applications. The Juno tool is especially dangerous as it can be easily upgraded to spoof the source address and ports. When the attack was launched, the connection with the system was lost almost instantly. Detecting this attack is simple but surprisingly useless. Even with an active firewall rule, Snort still analyzes the malicious traffic and the system's CPU load approaches 100%.

5.4. Assessment of Results

The performed tests clearly indicate that SIP proxy is rather vulnerable to DoS attacks. As the BESIP runs on a limited physical machine, only very basic protection mechanisms against certain DoS attacks can be implemented. This system consists of the following applications: Snort, SnortSam and Iptables. The tests proved that the analysis of the BESIP's traffic does not significantly affect system's performance (except for TCP SYN flood attack).

The most dangerous attacks include flooding with REGISTER, INVITE and OPTIONS messages, link bandwidth depletion using udpflood and TCP SYN flood attack. The attacks using malicious ACK, BYE or CANCEL messages are harmless at lower rates, with the same impact as udpflood at higher rates. No effective protection to be applied directly on the BESIP exists against certain attacks. In this case, a more secure network topology is the only solution (Fig. 5).

The main change in this topology is the inclusion of a demilitarized zone – DMZ. It is located between two firewalls (inner and outer). The purpose of this zone is to separate the safe inner part from the rather dangerous outer part of the network. Both firewalls run SnortSam agents so rules can be dynamically applied on both machines.

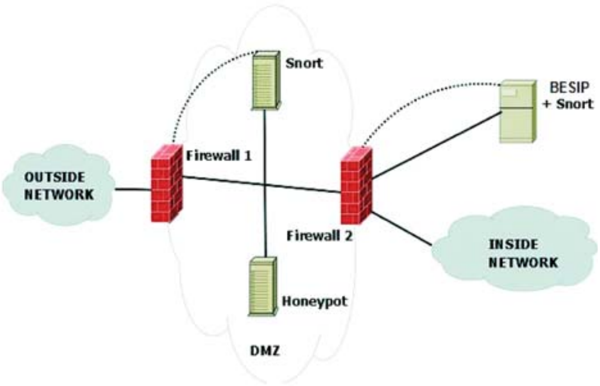


Fig. 5 The proposal of a safer topology

The inner firewall (marked as Firewall 2) serves to protect the BESIP system against the attacks from inside the network. All traffic to the BESIP has to pass through at least one firewall. The safe inner network should be implemented as a matter of course. The potential attack from inside the network would affect many users. Using encryption, VoIP VLANs and methods such as ARP inspection and DHCP snooping should provide an adequate response to possible security breaches. The implementation of a QoS mechanism should further reinforce the protection.

A honeypot located in the DMZ is an inspiration for further security precaution to be implemented.

6. Conclusion

We have developed and implemented a system with the working title BESIP, which allows easy integration of SIP IP telephony infrastructure to branch offices of large companies. On this solution were also implemented security solutions, which should reduce or completely eliminate the attacks, mostly DoS threats. We tested their efficiency in practice and documented the results. This article maps the most frequently used DoS attacks of today and evaluates the risk inherent to each of them. On the other hand the solution proposed in this article should ensure only a basic level of protection suitable for small and middle-size offices or detached workplaces for which the BESIP is intended. The contribution of this paper includes the performed comparison of the DoS attacks' efficiency. It was tested both without any protection and then with implemented Snort and SnortSam applications as proposed in our solution.

Acknowledgement

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement No. 218086.

References

[1] BRADNER, S., CONTROY, L., FUJIWARA, K.: *The E.164 to Uniform Resource Identifiers (URI) Dynamic Delegation Discovery System (DDDS) Application (ENUM)*, RFC 6116, URL: http://tools.ietf.org/html/rfc6116, 2011.

[2] DUHA, J., DADO, M., JARINA, R.: Communication Technologies and Services, *Communications - Scientific Letters of the University of Zilina,* vol. 5, No. 3, pp. 33-35, 2003.

[3] MACURA, L., VOZNAK, M., TOMALA, K., SLACHTA, J.: *Embedded Multiplatform SIP Server Solution,* 35th Intern. Conference on Telecommunications and Signal Processing, TSP 2012, Prague, pp. 263-266, 2012.

[4] KLIMO, M., KOVACIKOVA, M., SEGEC, P.: Selected Issues of IP Telephony, *Communications - Scientific Letters of the University of Zilina,* vol. 6, No. 4, pp. 63-70, 2004.

[5] SURHONE, L. M., TENOE, M. T.: *OpenWrt,* Betascript Publishing, ISBN-13: 978-6135271591, 2011.

[6] LOELLIGER, J.: *Version Control with Git: Powerful Tools and Techniques for Collaborative Software Development,* O'Reilly Media; 1 edition, ISBN-13: 978-0596520120, 2009.

[7] ENNS, R.: *NETCONF Configuration Protocol,* RFC 474, URL: http://tools.ietf.org/html/rfc4741, 2006.

[8] BOGUS, B.: *Lighttpd,* Packt Publishing, ISBN-13: 978-1847192103, 2008.

[9] VOZNAK, M., HALAS, M., REZAC, F., KAPICAK, L.: *Delay Variation Model for RTP Flows in Network with Priority Queueing,* 10th WSEAS Intern. Conference on EHAC'11 and ISPRA-11, Cambridge, pp. 344-349, 2011.

[10] REZAC, F., VOZNAK, M., TOMALA, K., ROZHON, J., VYCHODIL, J.: *Security Analysis System to Detect Threats on a SIP VoIP Infrastructure Elements,* Advances in Electrical and Electronic Engineering, vol. 9, No. 5, pp. 225-232, 2011.

[11] BARTH, W.: *Nagios: System and Network Monitoring,* No Starch Press; Second Edition, ISBN-13: 978-1593271794, 2008.

[12] ENDLER, D., COLLIER, M.: *Hacking Exposed VoIP,* McGraw-Hill Osborne Media, p. 321-363, 2009.

[13] SISALEM D., KUTHAN J., ELHERT T. S., FRAUNHOFER, F.: *Denial of Service Attacks Targeting SIP VoIP Infrastructure: Attack Scenarios and Prevention Mechanisms,* IEEE Network, 2006.

Patrik Kamencay – Martina Zachariasova– Robert Hudec – Miroslav Benco – Jan Hlubik – Slavomir Matuska *

IMAGE SEGMENTATION AND FEATURE EXTRACTION USING SIFT-SAD ALGORITHM FOR DISPARITY MAP GENERATION

In this paper, a stereo matching algorithm based on image segments and disparity measurement using stereo images is presented. We propose the hybrid segmentation algorithm that is based on a combination of the Belief Propagation and Mean Shift algorithms with aim to refine the final disparity map by using a stereo pair of images. Firstly, a color based segmentation method is applied for segmenting the left image of the input stereo pair (reference image) into regions. The aim of the segmentation is to simplify representation of the image into the form that is easier to analyze and is able to locate objects in images. Secondly, results of the segmentation are used as an input of the SIFT-SAD matching method to determine the disparity estimate of each image pixel. This matching algorithm is proposed by combining Scale Invariant Feature Transform (SIFT) with the Sum of Absolute Difference (SAD). Finally, the comparisons between the three robust feature detection methods SIFT, Affine SIFT (ASIFT) and Speeded Up Robust Features (SURF) are presented. The obtained experimental results demonstrate that the proposed method has a positive effect on overall estimation of disparity map and outperforms other examined methods.

Keywords: *Belief Propagation, Mean Shift, SIFT, ASIFT, SURF, disparity map.*

1. Introduction

This paper describes a set of algorithms for structure, motion automatic recovery and visualization of a 3D image from a sequence of 2D images. The important step to perform this goal is matching of corresponding pixels in the different views to estimate the depth map. The depth of an image pixel is the distance of the corresponding world point from the camera center. Detecting objects, estimating their pose, geometric properties and recovering 3D shape information are a critical problem in many vision and stereo computer vision application domains such as robotics applications, high level visual scene understanding, activity recognition, and object modeling [1]. The structure and motion recovery system follows a natural progression, comprising the following phases:

- feature matching using SIFT descriptor,
- image segmentation,
- feature detection using SIFT-SAD algorithm,
- disparity and depth map generation.

A classical problem of stereo computer vision is the extraction of 3D information from stereo views of a scene. To solve this problem, knowledge of view properties and feature point between views is needed. However, finding these points is notoriously hard to do for natural scenes. The fundamental idea behind stereo computer vision is the difference in position of a unique 3D point in two different images. As the object moves closer to the cameras, the relative position of object will change, and the positions in each image will move away from each other. In this way, it is possible to calculate the distance of an object, by calculating its relative positioning in the two images. This distance between the same objects in two images is known as disparity [1]. Disparity map computation is one of the key problems in 3D computer vision.

This paper employed a new feature projection approach based on SIFT-SAD method using hybrid segmentation algorithm. A comparison between these two different approaches for the image segmentation (Mean Shift and Belief Propagation) is described in [2]–[4].

The outline of the paper is as follows. The section 2 gives brief overview of the state-of-the-art in stereo matching and stereo correspondence. The proposed method of disparity map estimation from corresponding points using SIFT-SAD algorithm is described in section 3. Finally the experiment results and architecture of reconstruction algorithm are introduced in Section 4 and brief summary is discussed in Section 5.

2. Related work

In this section, we review related stereo. We refer the reader to a detailed and updated taxonomy of dense, two-frame stereo correspondence algorithms by Scharstein and Szeliski [5]. It also provides a tested for quantitative evaluation of stereo algorithms.

---

* **Patrik Kamencay, Martina Zachariasova, Robert Hudec, Miroslav Benco, Jan Hlubik, Slavomir Matuska**
Department of Telecommunications, University of Zilina, Slovakia, E-mail: patrik.kamencay@fel.uniza.sk

A stereo algorithm is called a global method if there is a global objective function to be optimized. Otherwise it is called a local method. The central problem of local or window-based stereo matching methods is to determine the optimal support window for each pixel. An ideal support region should be bigger in texture less regions and should be suspended at depth discontinuities. The fixed window is obviously invalid at depth discontinuities. Some improved window-based methods, such as adaptive windows [6], shift-able windows [7] and compact windows [7] try to avoid the windows that span depth discontinuities.

In stereo correspondence, two previous comparative papers have focused on the performance of sparse feature matchers [8] and developed new criteria for evaluating the performance of dense stereo matchers for image-based rendering application [9]. Our work is a continuation of the investigations begun by Szeliski and Zabih [10], which compared the performance of several popular algorithms.

3. Proposed method

In this section we present our method for fusing two approaches to disparity map estimation from input stereo images: hybrid segmentation algorithm and SIFT-SAD representation. This proposed algorithm based on the combination of the hybrid segmentation algorithm with SIFT descriptor and SAD stereo matching algorithm is faster, since a small portion of whole left and right images pixels are used for matching. The proposed method shown in Fig. 1 is implemented in MATLAB environment and improves the performance of disparity map calculation.

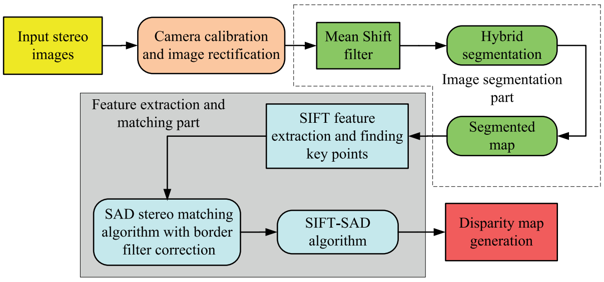


Fig. 1 Architecture for disparity map computation.

First, step is image rectification. It is transformation which makes pairs of conjugate epipolar lines become collinear and parallel to the horizontal axis (baseline). For the epipolar rectified images pair, each point in the left image lies on the same horizontal scan line as in the right image. This approach is used to reduce a search space for disparity map estimation algorithm. Next, we apply image filtering by Mean Shift filter. This step is very useful for noise removing, smoothing and image segmentation [11]. After filtration, the filtered image is split into segments using hybrid segmentation algorithm. Image segmentation (automatically partitioning an image into regions) is an important stage of our proposed algorithm for disparity map estimation. The combination of Mean Shift and Belief Propagation segmentation algorithms are deployed in order to improve precision of the key points search using SIFT and overall complexity. Finally, matching is performed using SAD algorithm, where a disparity map is obtained. The accuracy of SIFT-SAD algorithm depends on the correctness and quality of hybrid image segmentation.

3.1 Hybrid segmentation algorithm

This hybrid approach delivers accurately localized and closed object contours and brings together the advantages of both segmentation algorithms. Mean Shift is quick and Belief Propagation is very accurate segmentation. Initially, the noise corrupting the image is reduced by a noise reduction technique that preserves edges remarkably well, while reducing the noise quite effectively. At the second stage, this noise suppression allows a more accurate calculation and reduction of the number of the detected false edges.

First, we apply image filtering by Mean Shift algorithm. This step is very useful for noise removing, smoothing and image segmentation. For each pixel of an image, the set of neighboring pixels is determined. For each pixel of an image, the set of neighboring pixels is determined. Let X_i be the input and Y_i filtered image, where $i = 1, 2, ..., n$. The filtering algorithm comprises of the following steps

- Compute through the Mean Shift the mode where the pixel converges.
- Store the component of the gray level of the calculated value $Z_i = (x_{i,s}, y_{i,c})$, where $x_{i,s}$ is the spatial component and $y_{i,c}$ is the range component.

Secondly, the image is split into segments using Mean Shift algorithm. In the third step, means of segments are retrieved by applying mean shift theory. Fourth, the small segments are merged together to the most similar adjacent segments by the Belief Propagation method. Finally, we have integrated our proposed hybrid segmentation algorithm with the SIFT descriptor and Sum-of-Absolute-Differences (SAD) stereo matching algorithm. This proposed combination is able to produce highly accurate disparity map [12].

Parameters used in hybrid algorithm Tab.1

Parameter	Set value
p	5
S	50
Min_sh	1
Max_sh	40

The set up parameters of the used hybrid segmentation algorithm are shown in Table 1. The spatial resolution parameter p

affects smoothing and connectivity of segments. Moreover, parameter S is a size of the smallest segment, *Min_sh* is minimum shift and *Max_sh* is maximum shift of the pixels.

Advantages of the proposed hybrid algorithm:

- Efficient edge preserving smoothing guided by Mean Shift.
- Ability to change the image topology by using a simple merging mechanism, thus reducing over-segmentation.
- Relatively low sensitiveness to noise.
- Execution time directly proportional to the image size.

3.2 SIFT-SAD algorithm

The performance of stereo matching algorithms depends on the choice of matching cost. In our experiment we proposed SIFT-SAD matching method as matching cost. Scale Invariant Feature Transform (SIFT) is a local descriptor of image features insensitive to illuminant and other variants that is usually used as sparse feature representation. SIFT features are features extracted from images to help in reliable matching between different views of the same object [13]. Basically, in SIFT descriptors the neighborhood of the interest point is described as a set of orientation histograms computed from the gradient image. SIFT descriptors are invariant to scale, rotation, lighting and viewpoint change (in a narrow range). The most common implementation uses 16 histograms of 8 bins (8 orientations), which gives a 128 dimensional descriptor [13]. The SAD algorithm is based on accumulating absolute differences of the left image and right image pixels within a given window. It works by taking the absolute value of the difference between each pixel in the original block and the corresponding pixel in the block being used for comparison. The more similar the pixels are the less the SAD value becomes. These differences are summed over the block to create a simple metric of block similarity, the $L1$ norm of the difference image [14]. SIFT descriptor delivers most of local gradient information and SAD provides local intensity information. SIFT-SAD consists of two parts. Firstly, we get the $L1$ distance of SIFT between pixel p in the left image and $p + d_p$ in the right image.

$$D_{SIFT}(d_p) = \left\| x_L(p) - x_R(p + d_p) \right\|, \tag{1}$$

where d_p is the disparity of pixel p, $\|x_L(\mathrm{p}) - x_R(p + d_p)\|$ is the $L1$ distance. Next, we define SAD matching cost as:

$$D_{SAD}(d_p) = \exp\left(-SAD(p, p + d_p)\right), \tag{2}$$

where SAD $(p, p + d_p)$ is the SAD score in a square neighborhood searching window. Our algorithm computes the disparity for all pixels with a window size dimension at a square of 9×9 pixels. The minimum difference value over the frame indicates the best matching pixel, and position of the minimum defines the disparity of the actual pixel [15]. Then, a linear combination of SIFT-SAD algorithm is proposed as

$$D(d_p) = D_{SIFT}(d_p) + \lambda D_{SAD}(d_p), \tag{3}$$

where λ is a weighting factor that controls the contribution of SIFT part and SAD part. We set $\lambda = 1$ in all the experiments. Finally, we use one dimensional Gaussian weight with a scale factor s to get the matching cost. The underlying assumption is that if a minimum corresponds to the true surface, the neighboring pixels should have near values at a similar depth [15].

Quality of final 3D disparity map depends on square window size, because a bigger window size corresponds to a greater probability of correct pixel disparity calculated from matched points, although the calculation gets slower [15].

4 . Experimental results

In this section, some of the obtained experimental results will be presented. All the experiments were implemented in Matlab. We conducted experiments on Middlebury image database [16] using an Intel(R) Core2 Quad CPU with 2.40 GHz processor. The input to the proposed algorithm is a stereoscopic pair images. A SIFT-SAD matching algorithm lies in the heart of the 3D reconstruction procedure.

First, the proposed hybrid algorithm with three segmentation algorithms were compared using automatic algorithm evaluating the precision of segmentation, as is shown in Table 2. This plays important role for two reasons:

- it can be placed into a feedback loop to enforce another run of segmentation algorithm that may include more sophisticated steps for high precision segmentation,
- outcome of this evaluation can be treated as a quality factor and thus can be used to design a quality driven adaptive recognition system.

The definition of precision (P), recall (R) and $F1$ is given by

$$P = \frac{C}{C + F} * 100, \tag{4}$$

$$R = \frac{C}{C + M} * 100, \tag{5}$$

where C is the number of correct detected pixels that belonging to the boundary, F is the number of false detected pixels and M is the number of not detected pixels. Parameter $F1$ is a combined measure from precision and recall [15]. The definition of $F1$ is given by

$$F1 = \frac{2PR}{P + R} * 100, \tag{6}$$

Best results of image segmentation algorithms Tab. 2

Segmentation algorithm	P [%]	R [%]	F1 [%]	Computing time [s]
Belief Propagation	55.34	19.47	21.03	34.19
K-Means	43.27	15.13	17.56	34.21
Mean Shift	53.09	21.35	23.12	37.02
Hybrid segmentation	61.49	25.09	27.52	54.33

Next, combination of SIFT-SAD algorithm was tested (see Table 3 and Table 4). The result of this stereo matching process is a disparity map that indicates the disparity for every pixel with corresponding intensity. Quality of disparity map is represented as percentage of pixels with disparity errors (bad matching pixels (see Table 3) [5]:

$$P = \frac{1}{X * Y} \sum_{i=1}^{X} \sum_{j=1}^{Y} \left(\left| d_C(i,j) - d_T(i,j) \right| \right), \tag{7}$$

where $X * Y$ represent the size of the image, d_C is the computed disparity map of the test image and d_T is the truth disparity map.

$$d_T = \frac{fBI_{RES}}{D_T h}, \tag{8}$$

where D_T is ground truth depth map, h is height from the ground plane, $D_T * h$ is ground truth distance, B is baseline between the cameras, I_{RES} is image resolution and f is focal length.

Quantitative evaluation of the proposed method in terms percentage of error rates. Tab.3

	Aloe	Dolls	Reindeer
Graph Cut	4.27	3.59	5.03
Graph Cut + Occlusion	3.83	4.15	4.72
Dynamic Programming	4.35	3.78	4.89
Proposed method (SIFT-SAD)	2.12	1.97	2.75

Figure 2 indicates the computed disparity maps for the three scenes together with the used ground-truth maps. For quantitative evaluation, we examine the performance of several algorithms by their error rates. Table 3 shows the error rate for three test stereo image pairs (“Aloe”, “Dolls” and “Reindeer”).We see that the proposed disparity estimation procedure has boosted up the perfor-

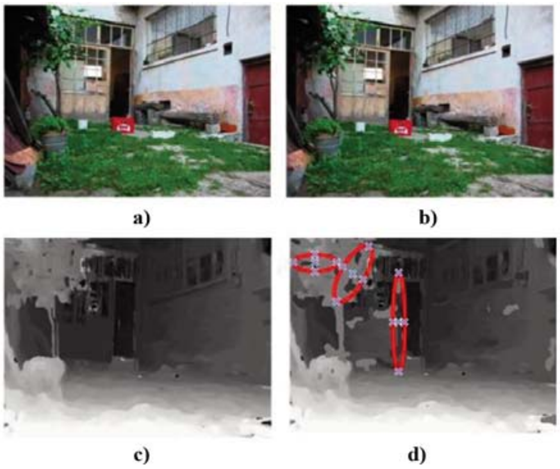


Fig. 3 Results for test “House” stereo image pair: a) left image, b) right image, c) the disparity results of the proposed method, d) the disparity result of [19]

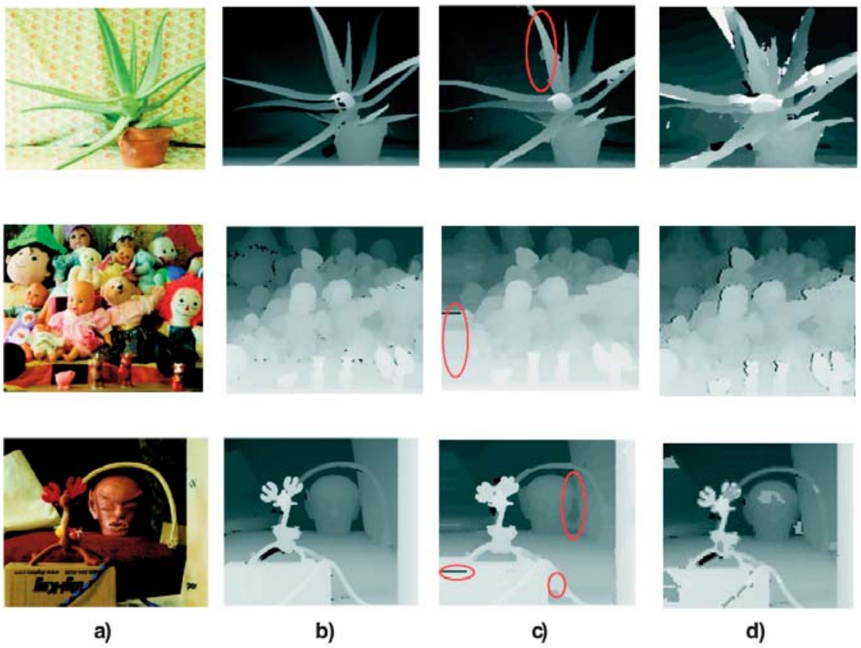


Fig.2 Aloe, Dolls and Reindeer, a) original images, left view; b) ground-truth referring to the left view with black labeled occlusions; c) computed disparity maps using hybrid segmentation algorithm and SIFT-SAD, computed disparity maps using only SAD method

mance of our algorithm. The ground truth disparity map is the inverse of the ground truth distance scale by the image resolution and the focal length. Equation (8) shows how to calculate the ground truth disparity map from the depth map [18]. The depth map is a 16 bit map with values ranging from 0 to 1 where the ground plane was at $D = 1$ and the cameras were at $D = 0$. D is distance of object from the camera. Time is counted for the complete processing which includes feature detecting and matching. Table 3 shows that SURF is the fastest one and SIFT-SAD is the slowest, but it finds most matches.

Although the proposed approach (see Fig. 1) improves the quality of the final disparity map and handles the occlusion, they cannot estimate correct disparity values when the background area behind an object is textureless. This wrong estimation as false matching are called. To give an example, the "Aloe" stereo image pair has many inaccurate disparity regions as illustrated by red circles in Fig. 2, which is obtained by the proposed algorithm. We can see the false matching clearly inside red circles.

In addition, we show another result in Fig. 3 which has the biggest search range of our test images. The size of each left and right images at Fig. 3 (a) and (b) is 800 by 600. The similarity measure was computed with a minimum quadratic 3×3 correlation window because of the high amount of texture in every scene.

The percentage of disparity found correctly, disparity error and the detected occlusion that are correct. Tab. 4

	SIFT	ASIFT	SURF	SIFT-SAD
Disparity correct [%]	86.69	82.07	89.78	92.35
Disparity error [%]	13.31	17.93	10.22	7.65
Occlusion correct [%]	67.45	65.32	72.76	72.03
Total matches	125	135	89	312
Total time [s]	5.52	5.07	2.78	4.95

Table 4 shows summary of overall performance. We compared performances obtained by the proposed method SIFT-SAD with those obtained by three common algorithms (SIFT, ASIFT and SURF). Approximately 92 percent of the disparity values were found correctly for our proposed algorithm. The final disparity map is labeled as correct if it is within one pixel of the correct disparity. Our result in Fig. 3 (c) successfully detects false matching areas and assigns more accurate disparity vector in occlusion regions than the result generated by [19] at Fig. 3 (d) (red circles). For example, our algorithm removes the false matching area around the door and tree which are most complexes regions of the stereo images.

5. Conclusion

The method for reconstructing a 3D scene and proposed algorithm for disparity map measurement from two input images was presented. This algorithm uses image segmentation and SIFT-SAD feature point detection method which extracts more key-points than other feature extraction methods such as SIFT, SURF or ASIFT. The proposed system is based on 3D reconstruction solution using stereo images. This system works with common cameras. The applications of these methods of 3D picture processing are very useful in sphere of medicine, for example detection and identification of tumor in brain and also in other branches as physics, biology or astronomy. In the future we could speed up computation time, improve precision of the hybrid algorithm and apply these methods in real situations.

Acknowledgements

This work was supported by the Slovak Science Project Grant Agency, Project No. 1/0705/13 "Image elements classification for semantic image description" and by project "Competence Center for research and development in the field of diagnostics and therapy of oncological diseases", ITMS: 26220220153, co-funded from EU sources and European Regional Development Fund.

Európska únia
Európsky fond regionálneho rozvoja

Operačný program VÝSKUM a VÝVOJ

Agentúra
Ministerstva školstva, vedy, výskumu a športu SR
pre štrukturálne fondy EÚ

„We support research activities in Slovakia/This project is being co-financed by the European Union"

References

[1] SUN, J., ZHANG, N. N., SHUM, H. Y.: *Stereo Matching Using Belief Propagation,* IEEE Transactions on Pattern Analysis and Machine Intelligence, 25(7):787-800, 2003.

[2] YEDIDA, J. S., FREEMAN, W. T., WEISS, Y.: *Understanding Belief Propagation and it is Generalizations,* Exploring Artificial Intelligence in the New Millennium, Chap. 8, pp. 239-236, January 2003.

[3] DONG, L., OGUNBONA, P., LI, W., YU, G., FAN, L., ZHENG, G: *A Fast Algorithm for Color Image Segmentation,* First Intern. Conference on Innovative Computing, Information and Control, 2006.

[4] YIN, Z., COLLINS, R.: *Belief Propagation in a 3D Spatio - Temporal MRF for Moving Object Detection,* Department of Computer Science and Engineering, The Pennsylvania State University, University Park, PA, 2007.
[5] SCHARSTEIN, D., SZELISKI, R.: A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms, *Intern. J. of Computer Vision,* 47(1), pp. 7-42, 2002.
[6] KANADE, T., OKUTOMI, M.: *A Stereo Matching Algorithm with an Adaptive Window: Theory and Experiment,* PAMI, 16(9), pp. 920-932, 1995.
[7] VEKSLER, O.: *Stereo Matching by Compact Windows via Minimum Ratio Cycle,* ICCV, 2002.
[8] HSIEH, Y. C., McKEOWN, D., PERLANT, F. P.: *Performance evaluation of scene registration and stereo matching for cartographic feature extraction,* IEEE TPAMI, 14(2), pp. 214-238, 1995.
[9] MULLIGAN, J., ISLER, V., DANIILIDIS, K.: Performance Evaluation of Stereo for Tele-Presence, *In ICCV*, vol. 2, pp. 558-565, 2002.
[10] SZELISKI, R., ZABIH, R.: An Experimental Comparison of Stereo Algorithms, *In International Workshop on Vision Algorithms,* Springer, pp. 1-19, 2000.
[11] KUHL, A.: *A Comparison of Stereo Matching Algorithm for Mobile Robots,* Centre for Intelligent Information Processing System, Western Australia, pp. 4-24, 2005.
[12] KAMENCAY, P., BREZNAN, M., JARINA, R., LUKAC, P., ZACHARIASOVA, M.: Improved Depth Map Estimation From Stereo Images Based On Hybrid Method, *The Radioengineering J.,* vol. 21, No. 1, April 2012, ISSN 1210-2512.
[13] LOWE, D. G.: Distinctive Image Feature from Scale Invariant Key Points, *Intern. J. of Computer Vision (IJCV),* vol. 2, No. 60, 2004, pp. 91-110.
[14] JUAN, L., GWUN, O.: A Comparison of SIFT, PCA-SIFT and SURF, *Intern. J. of Image Processing (IJIP),* vol. 3, No. 4, pp. 143-152.
[15] KE, Y., SUKTHANKAR, R.: *PCA-SIFT: A More Distinctive Representation for Local Image Descriptors,* Proc. Conf. Computer Vision and Pattern Recognition, pp. 511-517, 2003.
[16] Middlebury Stereo image pairs dataset. [online 10.10.2012] Available on the internet. http://vision.middlebury.edu/stereo/data/.
[17] ZACHARIASOVA, M., HUDEC, R., BENCO, M., KAMENCAY, P., LUKAC, P., MATUSKA, S.: The Effect of Metric Space on The Results of Graph Based Colour Image Segmentation, *Communications - Scientific Letters of the University of Zilina,* vol. 14, No. 3, 2012, ISSN 1335-4205.
[18] BOLECEK, L., RICNY, V., SLANINA, M.: *Fast Method for Reconstruction of 3D Coordinates,* 35th Intern. Conference on Telecommunications and Signal Processing (TSP 2012), July 2012, ISBN 978-1-4673-1115-1.
[19] KOLMOGOROV, V., ZABIH, R.: *Computing Visual Correspondence with Occlusions using Graph Cuts,* Proc. IEEE Int. Conference Computer Vision 2, pp. 508-515, 2002.

Jakub Safarik – Miroslav Voznak – Filip Rezac – Lukas Macura *

IP TELEPHONY SERVER EMULATION FOR MONITORING AND ANALYSIS OF MALICIOUS ACTIVITY IN VOIP NETWORK

The paper aims at gathering information about attacks from real internet infrastructure and their analysis. For this purpose, we prepared a set of honeypots monitoring various aspects of VoIP infrastructure including SIP endpoint and SSH terminal emulation. SIP endpoints are registered with real SIP registrar and the incoming calls are routed to a honeypot according the rules in dialplan. The honeypot gathers valuable data about hacker's activity with no threat to production systems. Analysis of the honeypot data is crucial for further improvement of existing security mechanisms in VoIP networks. The paper describes the honeypot's behaviour and brings an analysis of a detected malicious activity as well.

Keywords: *Artemisa, Dionaea, Kippo, VoIP attacks, VoIP honeypot.*

1. Introduction

The paper describes the use of honeypots in a VoIP infrastructure. These systems become increasingly necessary as the number of IP-based telephony solutions rises. Nearly all large companies today rely on some kind of IP telephony in their internal communication. This situation only induces greater hacker interest in these services. Nowadays, many companies have experienced abuse such as social engineering or spit calls [1].

The way to protect this infrastructure is to keep up with hackers and constantly improve security mechanisms. But achieving this simple goal is not easy at all. The basic rule is to keep all systems and their versions up to date, with at least access policies properly set and encryption of all crucial data. But this is not always possible in VoIP systems. The question is how do we find the system's bottleneck? Each of us needs to turn into a hacker in his own system to find system weaknesses. Although this gives a better understanding of the whole infrastructure, the number of security holes found depends on the skills of security auditor. Even if the auditor fixes all existing weaknesses there can still be security holes which can be exploited by either a foreign or inner attacker.

Another option is to create honeypots which lure hackers. Using these honeypots, we can obtain real data about hacker activities and map actual attacks in the network, which is otherwise not possible [2, 3].

The main purpose of a honeypot is to simulate the real system and interact with anyone in the same way as the production system would. It watches the behaviour of anyone who interacts with it [4].This article provides a close look on honeypots referred to as Artemisa and Kippo.

2. Honeypot features

An Artemisa honeypot can be deployed in any VoIP infrastructure which uses a SIP protocol. In this infrastructure it plays a role of a regular SIP phone (see Fig. 1)

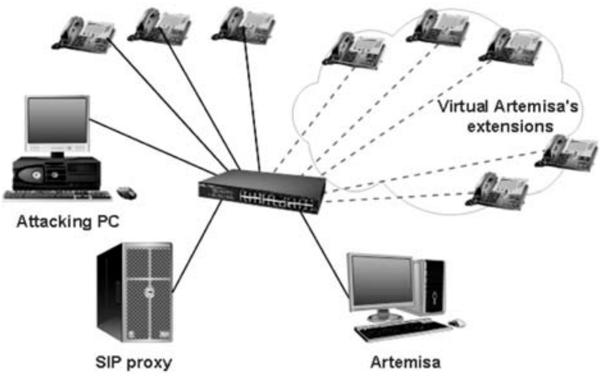


Fig. 1 A VoIP topology with a honeypot

The honeypot itself can run on a physical or virtual machine. The program connects to SIP proxy with the extensions defined in

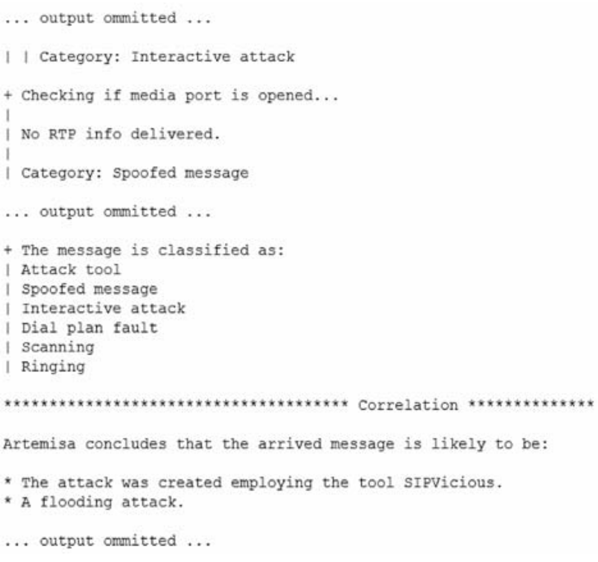
---

* **Jakub Safarik, Miroslav Voznak, Filip Rezac, Lukas Macura**
Department of Telecommunications, Faculty of Electrical Engineering and Computer Science, VSB-TU Ostrava, Ostrava-Poruba, Czech Republic,
E-mail: jakub.safarik@vsb.cz

a configuration file. The extensions should be within the range which is typically used for real accounts. The main purpose is to establish a better masking against the potential attackers [5]. Artemisa itself does not simulate PBX but rather an active end point device.

Once the call is established on one of Artemisa extensions, the honeypot simply answers the call. At the same time, it starts to examine the incoming SIP message. Artemisa then classifies the call and saves the result for a further review by the security administrator (Fig. 2).

The message is classified in the following steps. First of all, Artemisa looks for fingerprints of well-known attack tools. If the attacker uses some popular hacking tool, the fingerprint of this tool can easily reveal the malicious intentions. Then it checks domain names and SIP ports on the attacker side (provided they are really opened). There is also a similar check for media ports. Requested URI are also checked, as well as the ACK message received from the user. Finally, Artemisa checks the received RTP stream – provided a RTP stream was established (the audio trail of the received call can be stored in a WAV format).



```
... output ommitted ...

| | Category: Interactive attack

+ Checking if media port is opened...
|
| No RTP info delivered.
|
| Category: Spoofed message

... output ommitted ...

+ The message is classified as:
| Attack tool
| Spoofed message
| Interactive attack
| Dial plan fault
| Scanning
| Ringing

************************************** Correlation **************

Artemisa concludes that the arrived message is likely to be:

* The attack was created employing the tool SIPVicious.
* A flooding attack.

... output ommitted ...
```

Fig. 2 An example of the output file

This sequence of procedures helps Artemisa to classify the call. The result is then shown in a console. The results can be saved into a pre-defined folder or they can be sent as a notification by e-mail. Once the call has been examined, a series of bash scripts is executed. These scripts are executed with pre-defined arguments. Artemisa can launch some countermeasures against the incoming attacks [6].

A. Kippo features

The second tested honeypot is based on different foundations. It is not VoIP oriented as Artemisa. It simulates a SSH server. When someone tries to connect to a server with a honeypot running on it, the twistd application redirects this user to the honeypot. This happens where the user IP address is not included in the list of permitted IP address.

Once the connection with the honeypot is established, the attacker must enter correct username and password. These are set to the most used username *root* and password is the second most common combination of numbers *123456* (Table 1 lists Top 10 most frequently used passwords). Other combinations for the root access can be added to `data/pass.db` file.

Kippo logs every login attempt. Where the entered combination is valid, the intruder is granted access to a fake filesystem. Every command entered into the honeypot is logged and behaviour typical for a particular command is emulated (for the most common commands only). If the user tries to download something from the Internet, Kippo saves this file into a secure folder for further examination.

All logs made by *Kippo* are saved in a *MySQL* database which facilitates the subsequent analysis.

B. Dionaea features

All previously mentioned honeypots were single service oriented ones. Dionaea belongs to a multi-service oriented honeypot which can simulate many services at a time. Typically is information from these multiple services only general but dionaea serves only a small number of them like SMB (Microsoft's printers, files, serial ports sharing protocol), HTTP, FTP, TFTP, MSSQL (Microsoft SQL server), SIP protocols. Attackers abuse these protocols in most cases. Dionaea has also ability to save malicious content needed by hackers securely, but contrary to Kippo, it can also emulate code from these files.

Describing features of all these protocols is beyond the scope of this article and further features focus only on the SIP protocol. Dionaea works in a different way as Artemisa. There is no need for connecting to an external (or production) VoIP server. It simply waits for any SIP message and tries to answer it. It supports all SIP requests from RFC 3261 (REGISTER, INVITE, ACK, CANCEL, BYE, OPTIONS). Dionaea supports multiple SIP sessions and RTP audio streams (data from stream can be recorded). For better simulation of a real IP telephony system, it is possible to configure different user agent phone mimics with custom username, password combinations. There is functionality for a different pickup time on simulated phones via pickup delay feature. All traffic is monitored, and logs are saved in plain-text files and in sqlite database.

3. Monitoring traffic using artemisa

As mentioned before, Artemisa investigates all traffic which is routed to its extensions. That's not the whole truth. Artemisa can run in three different modes depending on the settings in the `behavior.conf` file. These modes are called passive, active and aggressive.

In the passive mode, Artemisa only takes the incoming calls and answers them. Using the active mode, we can achieve the same functionality as in the passive mode. In addition, Artemisa starts to examine the incoming SIP messages as described above. The last mode – the aggressive mode – attacks the intruder with its own bash script (the scripts are located under the `/scripts` directory). Typically, we run Artemisa in the active mode so that all SIP messages routed to the honeypot are analysed.

To test the usefulness of an *Artemisa* honeypot, we prepared some tests in our testing topology. We built a simple VoIP network with asterisk running as PBX and some end-point hardware and software phones. The honeypot was installed on a machine inside our network with five extensions. These extensions were running in the active mode. Since we are developing our own IPS system, we have chosen to test the honeypot under test scenarios similar to that IPS system.

4. Results

A. Artemisa's usability tests

First of all, we should start scanning the whole network from the point of view of a typical intruder. Many applications can be used for this purpose. We used two such applications – *nmap and SIPVicious* [7].

Both these applications yielded useful information. Yet neither *nmap* nor *svmap* was detected by the honeypot. In case of *svmap* there was information about the incoming SIP message, but this message was not analysed. No results were created after the network was scanned. This behaviour was quite surprising as typically, each attack starts by scanning the network. *Artemisa* should take account such situations.

With *svmap* we know about the running user-agent at honeypot's IP address (Fig. 3)

```
158.196.244.241:5060   |   Twinkle/1.4.2   |
T-Com  Speedport  W500V  /  Firmware  v1.37
MxSF/v3.2.6.26
```

Fig. 3 Svmap application output

Using this information we began looking for extensions running in the testing network. Direct scanning of the SIP proxy server was not detected by the honeypot, but when we use the *svwar* tool directly against the IP address on which a honeypot is running, we get information about all active extensions. This scan was recognized by the honeypot and an adequate result file was created. *Artemisa* correctly concludes that messages received came from a *SIPVicious* scanner. On the other hand we know from the *SIPVicious* output that these extensions do not behave as normal clients. This can stir up more caution on the side of the intruder.

The aim of other attacks was to flood the client's device with various types of SIP messages. Using some of these attacks, the intruder can achieve a DoS attack on a closed group of end-point devices [8]. For this kind of attack we used a number of tools including *udpflood, rtpflood, inviteflood* and *sipp* [9].

Each of these applications can launch a simple DoS attack. As *Artemisa* is a mere VoIP honeypot, it only detects attacks using the SIP protocol. Accordingly, only flood attacks from *inviteflood* and *sipp* were detected. In case of *inviteflood*, the application was successfully recognized thanks to its well-known fingerprint.

The *Sipp* application was not designed for hacking or penetration testing but this functionality can be achieved easily. We used specific call scenarios with a similar impact as the above mentioned flooding tools. Using sipp we can generate a high number of SIP messages which was immediately detected as a flood attack by the honeypot. In this situation, the whole honeypot stopped responding shortly and no result was recorded for the attack at all. Mere 250 SIP messages per second caused this situation. If we use lower sending rates, the attack was recognized well and the output file was successfully created.

Identifying a spit call is one of the most important features of the honeypot. We used the application called Spitfile for simulating these calls. Spitfile is an open-source SIP penetration tool. Using this tool, we can easily generate arbitrary calls. All of these calls were successfully detected by Artemisa and the appropriate output files were generated.

At last we tried to make a call to the honeypot extensions with hardware and software phone. Calls were marked as a scanning and ringing attack in both cases. So it seems that Artemisa evaluated almost every SIP message aiming at its extensions as some kind of attack.

The results from the detected attacks are stored in the `results`/directory inside the Artemisa folder. The output is in a simple text format and in html format, both of them containing the same data.

All functionalities mentioned before concern honeypots running in the active mode. The aggressive mode looks more interesting with its ability to counterattack the intruder. Artemisa contains three bash scripts to stop malicious activity. However, a close look at these scripts was surprising.

Let's start with the last script `on_spit.sh`. Inside this script, there is only one comment. This comment may activate a firewall rule, but the command is not included. This script is totally useless unless we rewrite it. Even the remaining scripts do not contain anything but comments inside. A simple condition (commented) is included in the `on_scanning.sh` script, which runs a python script to crash the scanning by the *SIPVicious* application (but only this particular application). The `on_flood.sh` script has a commented command inside to apply an iptables rule on the IP address and port. These are given by a parameter. This solution is not bad but if we want to block some traffic, there is a chance that a false positive attack will be blocked, so some automatic recovery mechanism should also be included. This can be easily solved by adding another script. This script will remove the rule after a certain interval. The main issue in blocking traffic using iptables is that the command applies the rule on a local machine. However, it only blocks the consequences on the honeypot, not on the main firewall which protects the whole infrastructure. This feature makes the aggressive mode useless against the attack of any intruder.

Using the honeypot in the passive mode is worthless because *Artemisa* only answers the call with no further analysis and without results being saved to a file.

B. Kippo data analysis

We use a *Kippo* honeypot to analyse SSH traffic in a real network with seven active monitoring sensors. The honeypot has been active and gathering data for a month. During this period, 873342 connection attempts were observed.

Only a small part of these connections was successful. Table 1 lists ten most frequently used password combinations enabling connections.

10 Most frequently udes password combinations Table 1

Password	count
	28146
123456	17625
password	6325
1234	5663
12345	5501
123	5342
1qa2ws3ed	5278
a	5121
test	4743
qwerty	4601

As we can see from the table, password *123456* was used 17625 times. The correct username root was used only in 2551 cases. These 2551 cases led to a connection to a fake filesystem. An intruder then typically uploads some kind of a script which should be used for a DoS attack on an outside server [10]. The in-depth analysis of the attacker's input is beyond the scope of this article.

Another interesting fact acquired from the honeypot is the approximate location of the origin of the attacker's connection (see Fig. 4).

The figure only shows first ten positions. Attacks from Germany account for 25.21% of the total connection attempts. China came second with 20.33% and Mongolia third with 15.52%. Almost 75% of connection attempts were made from one of the first four countries.

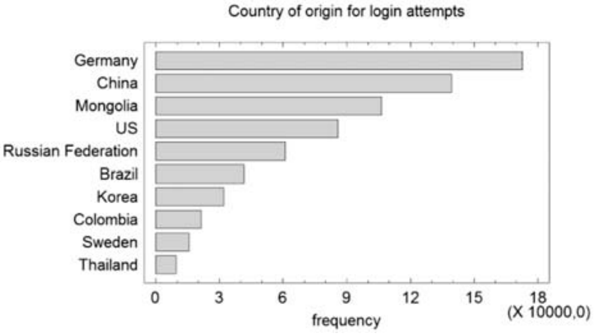


Fig. 4 Country of origin for logging attempts

As mentioned above, we have seven different sensors. With nearly a million attempts, each sensor was tested every 23 seconds.

C. Dionaea data analysis

Dionaea was monitoring malicious traffic for 18 days. The number of attacks is not as high as in Kippo case but still high enough. Fig. 5 shows the distribution of attacking IP addresses.

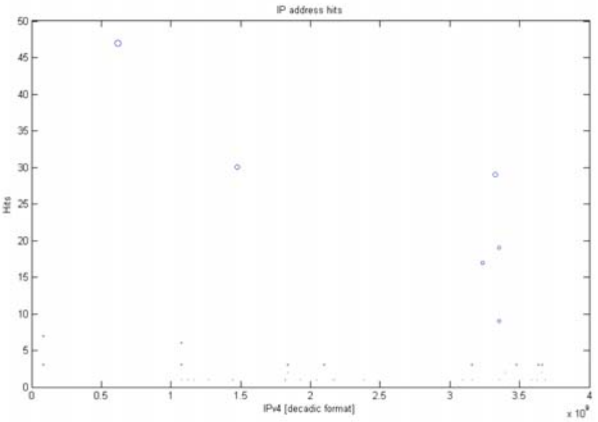


Fig. 5 Hits on attacker IP address

Most attacking attempts come from Israel (first peak in Fig. 5) with IP address 37.8.54.135. The second most used IP was originated in Germany and third in Russia.

Dionaea provides also additional information about attacks like used SIP messages, SIP header information, SDP and RTP statistics. Interesting is typical attack behaviour based on sent SIP messages. All attacks occur in typical sequences. Table 2 shows gathered SIP message data. In column groups is the number of SIP message's grouped by different connections. One connection is a single session with emulated honeypot's SIP proxy. Ratio simply illustrates the average number of messages in each connection group.

SIP message type analysis Table 2

SIP message	Groups	Count	Ratio
ACK	40	303	7,575
BYE	4	4	1,000
CANCEL	1	11	11
INVITE	18	85	4,722
OPTIONS	76	76	1,000
REGISTER	28	1745	62,321

Most of attacks can be divided into 2 groups. First represents various types of a PBX scanning & probing. Attacker send OPTION message and wait for an answer or simply try to place a call with immediate cancelation (it means INVITE message followed by CANCEL message).

Other group represents flood attacks. Our previous tests confirmed an extraordinary vulnerability of SIP proxy against OPTIONS floods, but attackers still use only REGISTER message flooding.

At last there are few attacks which cannot be simply placed in groups or generally categorized [11, 12].

5. Conclusion

All the tests which we carried out on a VoIP honeypot gave us a solid look on its features. The main goal of this article is to test the honeypot functionality and analyse gathered data. Another goal is to consider its deploying in our real IP telephony infrastructure.

The test revealed that Artemisa is not the silver bullet solution for discovering all security threats. Its main disadvantage is that it does not recognize a scanning attempt into the infrastructure. There is also a performance issue while analysing a flooding attack. It was a result of the flooding at higher rates. Accordingly, Artemisa cannot handle such a big mass of SIP messages.

But the biggest disappointment was the behaviour of *Artemisa* in the aggressive mode. It is a good idea to implement some mechanisms that can proactively block malicious activity but in case of this honeypot it was done weakly. The passive mode can only be used for testing purposes. We have not seen any option for using it in a real network. The active mode is the only applicable mode, and it should be a default setting for our honeypot.

On the other hand, Artemisa does exactly what we want it to do. An Artemisa honeypot is not a simulation of PBX, but rather of an endpoint device. Despite the fact that it freezes at high flooding rates, its principal utility is to detect suspicious call activity and spit attacks. This is by no means an easy task for many other tools.

SSH honeypot Kippo gives us a good idea about connection attempts on a standard SSH port. The information acquired was used to gain a better understanding of the attacker's behaviour. It is also a good source of malicious IP addresses. Another gathered information contains combinations of username and password used for connection (by attacker), command line activity history or samples of downloaded malicious software. After some time, the rate of connection attempts decreases. This happens once the honeypot has been discovered. It is necessary to change the honeypot configuration each time it is used.

Dionaea honeypot is great in simulating a SIP proxy. The data gathered show real attacks and gave us valuable feedback for improving existing security mechanisms. Quite surprising were flooding attacks using only REGISTER messages. We found no connection between IP addresses used for attacks on both kippo and dionaea.

Our previous research focused on creating an open-source IPS to protect VoIP PBX. Using honeypots to gather information about hacker's actions was extremely helpful. Now we orient our further research on creation of a honeypot network which should gather information on various locations. All this information would be stored in one datastore for the following analysis and other improvement of security mechanisms.

Acknowledgement
The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement No. 218086.

References

[1] SPITZNER, L.: *Honeypots: Tracking Hackers,* Addison-Wesley Professional, 2002.
[2] KLIMO, M., KOVACIKOVA, M., SEGEC, P.: Selected Issues of IP Telephony, *Communications - Scientific Letters of the University of Zilina,* vol. 6, No. 4, pp. 63-70, 2004.
[3] DUHA, J., DADO, M., JARINA, R.: Communication Technologies and Services, *Communications - Scientific Letters of the University of Zilina,* vol. 5, No. 3, pp. 33-35, 2003.
[4] SISALEM, D., KUTHAN, J., ELHERT, T. S., FRAUNHOFER, F.: *Denial of Service Attacks Targeting SIP VoIP Infrastructure: Attack Scenarios and Prevention Mechanisms.* IEEE Network, 2006.
[5] PROVOS, N., HOLZ, T.: *Virtual Honeypots,* Addison-Wesley Professional, 2007.

[6] REZAC, F., VOZNAK, M., TOMALA, K., ROZHON, J., VYCHODIL, J.: Security Analysis System to Detect Threats on a SIP VoIP Infrastructure Elements, *Advances in Electrical and Electronic Engineering,* vol. 9, No. 5, pp. 225–23, 2011.

[7] VOZNAK, M., REZAC, F.: Web-based IP Telephony Penetration System Evaluating Level of Protection from Attacks and Threats, *WSEAS Transactions on Communications,* vol. 10, No. 2, pp. 66–76, February 2011.

[8] JOSHI, R. C., SARDANA, A.: *Honeypots: A New Paradigm to Information Security,* Science Publishers, 2011.

[9] ENDLER, D., COLLIER, M.: *Hacking Exposed VoIP,* McGraw-Hill Osborne Media, 2009.

[10] SAFARIK, J., VOZNAK, M., REZAC, F., MACURA, L.: *Malicious Traffic Monitoring and its Evaluation in VoIP Infrastructure,* 35th Intern. Conference on Telecommunications and Signal Processing, TSP, Prague, pp. 259-262, 2012.

[11] VOZNAK, M., SAFARIK, J.: DoS Attacks Targeting SIP Server and Improvements of Robustness, *Intern. J. of Mathematics and Computers in Simulation,* vol. 6, No. 1, pp. 177–184, 2012.

[12] VOZNAK, M., REZAC, F.: Threats to Voice over IP Communications Systems, *WSEAS Transactions on Computers,* vol. 9, No. 11, pp. 1348–1358, November 2010.

COMMUNICATIONS – Scientific Letters of the University of Zilina
Writer's Guidelines

1. Submitted papers must be unpublished and must not be currently under review for any other publication.
2. Submitted manuscripts should not exceed 8 pages including figures and graphs (in Microsoft WORD – format A4, Times Roman size 12, page margins 2.5 cm).
3. Manuscripts written in good English must include abstract and keywords also written in English. The abstract should not exceed 10 lines.
4. Submission should be sent: By e-mail – as an attachment – to one of the following addresses: komunikacie@uniza.sk or holesa@uniza.sk (or on CD to the following address: Zilinska univerzita, OVaV – Komunikacie, Univerzitna 1, SK-10 26 Zilina, Slovakia).
5. Uncommon abbreviations must be defined the first time they are used in the text.
6. Figures, graphs and diagrams, if not processed in Microsoft WORD, must be sent in electronic form (as JPG, GIF, TIF, TTF or BMP files) or drawn in high contrast on white paper. Photographs for publication must be either contrastive or on a slide.
7. The numbered reference citation within text should be enclosed in square brackets. The reference list should appear at the end of the article (in compliance with ISO 690).
8. The numbered references (in square brackets), figures, tables and graphs must be also included in text – in numerical order.
9. The author's exact mailing address, full names, E-mail address, telephone or fax number, the name and address of the organization and workplace (also written in English) must be enclosed.
10. The editorial board will assess the submitted paper in its following session. If the manuscript is accepted for publication, it will be sent to peer review and language correction. After reviewing and incorporating the editor's comments, the final draft (before printing) will be sent to authors for final review and minor adjustments
11. Submission deadlines are: September 30, December 31, March 31 and June 30.

COMMUNICATIONS

SCIENTIFIC LETTERS OF THE UNIVERSITY OF ZILINA
VOLUME 15

Editor-in-chief:
Prof. Ing. Otakar Bokuvka, PhD.

Address of the editorial office:
Zilinská univerzita
Office for Science and Research
(OVaV)
Univerzitna 1
SK 010 26 Zilina
Slovakia

E-mail: komunikacie@uniza.sk

Each paper was reviewed by two reviewers.

Journal is excerpted in Compendex and Scopus.

It is published by the University of Zilina in
EDIS – Publishing Institution of Zilina University
Registered No: EV 3672/09
ISSN 1335-4205

Published quarterly

Single issues of the journal can be found on:
http://www.uniza.sk/komunikacie