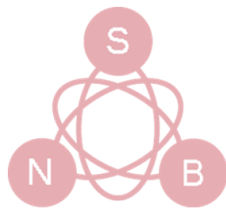# L INTERNATIONAL SCIENTIFIC CONFERENCE ON INFORMATION, COMMUNICATION AND ENERGY SYSTEMS AND TECHNOLOGIES

ICEST 2015
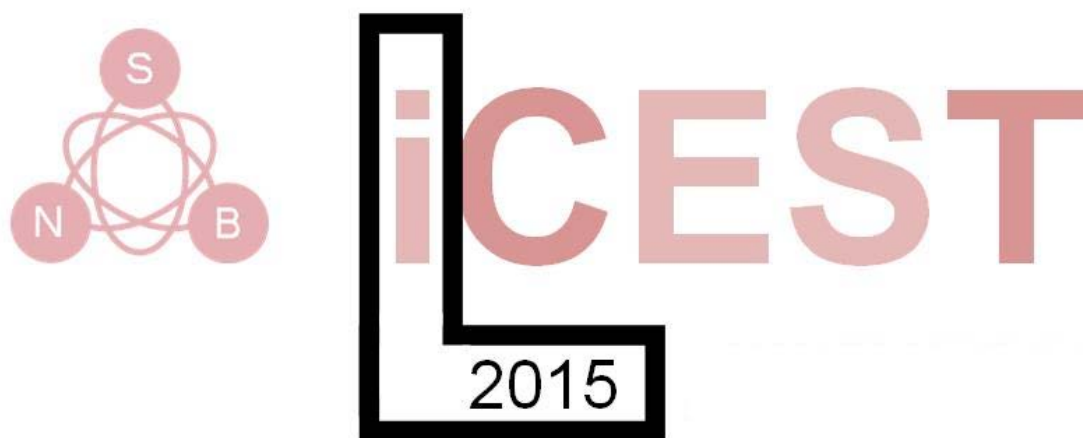

24 - 26 June 2015, Sofia, Bulgaria

Proceedings of papers

L INTERNATIONAL SCIENTIFIC CONFERENCE ON INFORMATION, COMMUNICATION AND ENERGY SYSTEMS AND TECHNOLOGIES

PROCEEDINGS OF PAPERS

SOFIA, 2015

ICEST 2015 Proceedings of the L International Scientific Conference on Information, Communication and Energy Systems and Technologies
Organized by the Faculty of Telecommunications, Technical University of Sofia, June 24-26, 2015, Sofia, Bulgaria

Proceedings of Papers
Editor: *Assoc. Prof. Kalin Dimitrov, PhD*
Published by: *Faculty of Telecommunications*
Printed by: *Publishing Company, TU-Sofia*

WELCOME TO ICEST 2015

Dear Colleagues,

Welcome to the L International Scientific Conference on Information Communication and Energy Systems and Technologies – ICEST 2015. The conference is going to be held from June 24 to 26 2015 at the Faculty of Telecommunications, Technical University of Sofia. The Conference is for fourteenth time jointly organized by the Faculty of Telecommunications, Sofia, Bulgaria, Faculty of Electronic Engineering, Niš, Serbia and by the Faculty of Technical Sciences, Bitola, Macedonia.

As to the earlier ICEST conferences many authors from all over the world submitted their papers. This year 96 papers have been accepted for oral or poster presentation.

After the conference opening one plenary invited paper "Future Cellular Networks for a Society in Motion" will be given by Univ.Prof. Dipl.-Ing. Dr.techn. Markus Rupp from TU Wien. Furthermore the Conference will also include workshop "Reasoning-based Intelligent Systems" with chairman: Prof. Dr. Kazumi Nakamatsu from University of Hyogo, Japan and meeting with IEEE Bulgaria section committee.

I hope that all participants will take opportunities not only to exchange their knowledge, experiences and ideas but also to make contacts and establish further collaboration. A social program, rich in events will provide more relaxing atmosphere for meeting the colleagues.

On the behalf of technical program committee I wish you successful presentations and pleasant stay in Sofia.

Assoc. Prof. PhD Kalin Dimitrov,
ICEST 2015 Conference Chairman
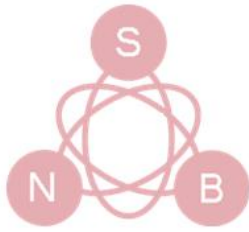
iCEST
2015

L INTERNATIONAL SCIENTIFIC CONFERENCE ON INFORMATION, COMMUNICATION AND ENERGY SYSTEMS AND TECHNOLOGIES ICEST 2015

organized by

Faculty of Telecommunications,
Technical University of Sofia, Bulgaria

Faculty of Technical Sciences,
University "St. Kl. Ohridski", Bitola, Macedonia

Faculty of Electronic Engineering,
University of Niš, Serbia

in co-operation with
- IEEE Bulgaria Section
- IEEE Macedonia Section
- IEEE Serbia Section

TECHNICAL PROGRAM COMMITTEE

Chairman:

K. Dimitrov   Technical University of Sofia, Bulgaria

Honorary Chairman:

R. Arnaudov   Technical University of Sofia, Bulgaria

Vice Chairmen:

B. Milovanović   University of Niš, Serbia
C. Mitrovski   St. Kliment Ohridski University, Bitola, Macedonia

Members:

V. Poulkov	Technical University of Sofia, Bulgaria, Dean of Faculty of Telecommunications
N. Acevski	University "St. Kliment Ohridski", Bitola, Macedonia
I. Atanasov	Technical University of Sofia, Bulgaria
M. Atanasovski	University "St. Kliment Ohridski", Bitola, Macedonia
M. Bakmaz	University of Belgrade, Serbia
A. Bekiarski	Technical University of Sofia, Bulgaria
W. Bock	University of Ottawa, Canada
O. Boumbarov	Technical University of Sofia, Bulgaria
V. Češelkoska	University "St. Kliment Ohridski", Bitola, Macedonia
D. Denić	University of Niš, Serbia
V. Demirev	Technical University of Sofia, Bulgaria
R. Dinov	Technical University of Sofia, Bulgaria
D. Dimitrov	Technical University of Sofia, Bulgaria
T. Dimovski	University "St. Kliment Ohridski", Bitola, Macedonia
D. Dobrev	Technical University of Sofia, Bulgaria
I. Dochev	Technical University of Sofia, Bulgaria
B. Dokić	University of Banja Luka, Bosnia and Herzegovina
N. Dončov	University of Niš, Serbia
D. Drača	University of Niš, Serbia
T. Eftimov	Plovdiv University "Paisii Hilendarski", Bulgaria
N. Gospić	University of Belgrade, Serbia
V. Georgieva	Technical University of Sofia, Bulgaria
G. Iliev	Technical University of Sofia, Bulgaria
I. Iliev	Technical University of Sofia, Bulgaria
M. Ivković	University of Novi Sad, Serbia
Z. Jakšić	IHTM Institute, Belgrade, Serbia
D. Jankov	ić University of Niš, Serbia
N. Janković	University of Niš, Serbia
M. Jevtić	University of Niš, Serbia
B. Jokanović	Institut of Physics, Belgrade, Serbia
I. Jolevski	University "St. Kliment Ohridski" Bitola, Macedonia
L. Jordanova	Technical University of Sofia, Bulgaria
Z. Jovanović	University of Niš, Serbia
V. Katić	University of Novi Sad, Serbia
B. Kolundžija	University of Belgrade, Serbia
Z. Konjović	University of Novi Sad, Serbia
M. Kostov	University "St. Kliment Ohridski" Bitola, Macedonia
R. Kountchev	Technical University of Sofia, Bulgaria
M. Lutovac	Singidunum University, Serbia
J. Makal	Tech. University of Byalistok, Poland
N. Maleš-Ilić	University of Niš, Serbia

D. Mančić	University of Niš, Serbia
G. Marinova	Technical University of Sofia, Bulgaria
V. Marković	University of Niš, Serbia
A. Markovski	University "St. Kl. Ohridski", Bitola, Macedonia
M. Milanova	University of Arkansas at Little Rock, USA
R. Miletiev	Technical University of Sofia, Bulgaria
S. Mirchev	Technical University of Sofia, Bulgaria
T. Mitsev	Technical University of Sofia, Bulgaria
P. Mitrevski	University "St. Kliment Ohridski", Bitola, Macedonia
K. Nakamatsu	University of Hyogo, Japan
I. Nedelkovski	University "St. Kl. Ohridski", Bitola, Macedonia
A. Nešić	IMTEL Institute, Belgrade, Serbia
T. Nikolov	Technical University of Sofia, Bulgaria
B. Nikolova	Technical University of Sofia, Bulgaria
Z. Nikolova	Technical University of Sofia, Bulgaria
D. Pantić	University of Niš, Serbia
E. Pencheva	Technical University of Sofia, Bulgaria
Z. Perić	University of Niš, Serbia
M. Petkovski	University "St. Kliment Ohridski", Bitola, Macedonia
P. Petković	University of Niš, Serbia
P. Petrović	IRITEL Institute, Belgrade, Serbia
S. Pleshkova	Technical University of Sofia, Bulgaria
P. Popnikolovska-Radevska	University "St. Kl. Ohridski", Bitola, Macedonia
O. Pronić-Rančić	University of Niš, Serbia
B. Radenković	University of Belgrade, Serbia
J. Radunović	University of Belgrade, Serbia
L. Raikovska	Technical University of Sofia, Bulgaria
B. Reljin	University of Belgrade, Serbia
I. Reljin	University of Belgrade, Serbia
R. Stanković	University of Niš, Serbia
Z. Stanković	University of Niš, Serbia
B. Stefanovski	University "St. Kliment Ohridski", Bitola, Macedonia
D. Stojanović	University of Niš, Serbia
M. Stojčev	University of Niš, Serbia
L. Stoimenov	University of Niš, Serbia
G. Stoyanov	Technical University of Sofia, Bulgaria
D. Tasić	University of Niš, Serbia
G. Todorov	St. Cyril and St. Methodius University of Veliko Tarnovo, Bulgaria
M. Todorova	St. Cyril and St. Methodius University of Veliko Tarnovo, Bulgaria
M. Trajanović	University of Niš, Serbia
Lj. Trpezanovski	University "St. Kliment Ohridski", Bitola, Macedonia
B. Tsankov	Technical University of Sofia, Bulgaria
A. Tsenov	Technical University of Sofia, Bulgaria
I. Uzunov	Technical University of Sofia, Bulgaria
S. Valtchev	NOVA-University, Lisbon, Portugal
L. Zieleznik	Brookes University of Oxford

CONFERENCE ORGANIZING COMMITTEE

Chairman:

K. Dimitrov   Technical University of Sofia, Bulgaria

International Coordinators:

B. Milovanović   University of Niš, Serbia
C. Mitrovski   Kliment Ohridski University, Bitola, Macedonia

Conference general technical editor:

I. Draganov   Technical University of Sofia, Bulgaria

Members:

T. Asenov   University of Niš, Serbia
A. Atanasković   University of Niš, Serbia
B. Bonev   Technical University of Sofia, Bulgaria
D. Denić   University of Niš, Serbia
T. Dimitrijević   University of Niš, Serbia
N. Dončov   University of Niš, Serbia
V. Đorđević   University of Niš, Serbia
A. Đorić   University of Niš, Serbia
R. Goleva   Technical University of Sofia, Bulgaria
D. Janković   University of Niš, Serbia
J. Joković   University of Niš, Serbia
V. Jović   University of Niš, Serbia
K. Kassev   Technical University of Sofia, Bulgaria
B. Kehayov   Technical University of Sofia, Bulgaria
D. Kireva   Technical University of Sofia, Bulgaria
N. Maleš-Ilić   University of Niš, Serbia
Z. Marinković   University of Niš, Serbia
V. Marković   University of Niš, Serbia
M. Milijić   University of Niš, Serbia
Ts. Mitsev   Technical University of Sofia, Bulgaria
N. Mojsovska   University "St. Kliment Ohridski", Bitola, Macedonia
M. Nenova   Technical University of Sofia, Bulgaria
N. Neshov   Technical University of Sofia, Bulgaria
K. Nikolova   Technical University of Sofia, Bulgaria
A. Panajotović   University of Niš, Serbia
J. Pargovski   University "St. Kliment Ohridski", Bitola, Macedonia
Z. Perić   University of Niš, Serbia
P. Petkov   Technical University of Sofia, Bulgaria
O. Pronić   Rančić University of Niš, Serbia
M. Spirovski   University "St. Kliment Ohridski", Bitola, Macedonia
Z. Stanković   University of Niš, Serbia
M. Stoilković   University of Niš, Serbia
B. Stošić   University of Niš, Serbia
K. Stoyanova   Technical University of Sofia, Bulgaria
M. Stoyanova   Technical University of Sofia, Bulgaria
D. Tasić   University of Niš, Serbia
M. Vasileva   Technical University of Sofia, Bulgaria

CONFERENCE SECRETARIAT

K. Stoyanova   Technical University of Sofia, Bulgaria

Address:

ICEST 2015 Conference
Technical University of Sofia
Faculty of Telecommunications
8, Kl. Ohridski, Blvd, 1000 Sofia, Bulgaria
1 bldg, 2 fl, rooms 1254, 1257
phone: +359 2 965 3145, +359 2 965 2272
E-mail: icest@tu-sofia.bg
Web: http://www.icestconf.org

LIST OF ICEST 2015 REVIEWERS

Prof. Dimiter Dimitrov, PhD
Technical University of Sofia, Bulgaria

Prof. Evelina Pencheva, DSc
Technical University of Sofia, Bulgaria

Prof. Dragan Denic, PhD
University of Niš, Serbia

Prof. Dragan Jankovic, PhD
University of Niš, Serbia

Prof. Miroslav Lutovac, PhD
Singidunum University, Serbia

Prof. Dragan Pantic, PhD
University of Niš, Serbia

Prof. Petar Spalevic, PhD
University of Priština, Kosovska Mitrovica, Serbia

Prof. Leonid Stoimenov, PhD
University of Niš, Serbia

Prof. Mile Stojcev, PhD
University of Niš, Serbia

Prof. Branko Dokic, PhD
University of Banja Luka, Bosnia and Herzegovina

Assoc. Prof. Angel Colov, PhD
Technical University of Sofia, Bulgaria

Assoc. Prof. Kalin Dimitrov, PhD
Technical University of Sofia, Bulgaria

Assoc. Prof. Ivo Dochev, PhD
Technical University of Sofia, Bulgaria

Assoc. Prof. Liljana Docheva, PhD
Technical University of Sofia, Bulgaria

Assoc. Prof. Anelia Doseva, PhD
Technical University of Sofia, Bulgaria

Assoc. Prof. Ivo Draganov, PhD
Technical University of Sofia, Bulgaria

Assoc. Prof. Veska Georgieva, PhD
Technical University of Sofia, Bulgaria

Assoc. Prof. Emil Manolov, PhD
Technical University of Sofia, Bulgaria

Assoc. Prof. Agata Manolova, PhD
Technical University of Sofia, Bulgaria

Assoc. Prof. Galia Marinova, PhD
Technical University of Sofia, Bulgaria

Assoc. Prof. Tsvetan Mitsev, PhD
Technical University of Sofia, Bulgaria

Assoc. Prof. Seferin Mirtchev, PhD
Technical University of Sofia, Bulgaria

Assoc. Prof. Marin Nedelchev, PhD
Technical University of Sofia, Bulgaria

Assoc. Prof. Kamelia Nikolova, PhD
Technical University of Sofia, Bulgaria

Assoc. Prof. Oleg Panagiev, PhD
Technical University of Sofia, Bulgaria

Assoc. Prof. Georgi Tsenov, PhD
Technical University of Sofia, Bulgaria

Assoc. Prof. Zlatka Valkova, PhD
Technical University of Sofia, Bulgaria

Assoc. Prof. Nebojsa Doncov, PhD
University of Niš, Serbia

Assoc. Prof. Dejan Milic, PhD
University of Niš, Serbia

Assoc. Prof. Nenad Milosevic, PhD
University of Niš, Serbia

Assist. Prof. Tihomir Brusev, PhD
Technical University of Sofia, Bulgaria

Assist. Prof. Stoyan Kirilov, PhD
Technical University of Sofia, Bulgaria

Assist. Prof. Rumen Mironov, PhD
Technical University of Sofia, Bulgaria

Assist. Prof. Peter Petkov, PhD
Technical University of Sofia, Bulgaria

Assist. Prof. Vencislav Trifonov, PhD
Technical University of Sofia, Bulgaria

Assist. Prof. Simeon Tsvetanov, PhD
Technical University of Sofia, Bulgaria

Assist. Prof. Julian Velchev, PhD
Technical University of Sofia, Bulgaria

Assist. Prof. Hristomir Yordanov, PhD
Technical University of Sofia, Bulgaria

Assist. Prof. Dejan Ciric, PhD
University of Niš, Serbia

Assist. Prof. Goran Djordjevic, PhD
University of Niš, Serbia

Assist. Prof. Zlatica Marinkovic, PhD
University of Niš, Serbia

Assist. Prof. Zoran Stankovic, PhD
*University of Niš, Serbia*Assist.

Assist. Prof. Dragan Stojanovic, PhD
University of Niš, Serbia

Assist. Prof. Biljana Stosic
University of Niš, Serbia

PLENARY SESSION
Chairman: Prof. Dr. V. Poulkov, Technical University of Sofia, Bulgaria

Future Cellular Networks for a Society in Motion
Univ.Prof. Dipl.-Ing. Dr.techn. Markus Rupp, TU Wien

Abstract: It is expected that by 2050 about 86% of the world's population shall be urbanized. Even though most people will concentrate around city centers, they will have to commute substantially. Not just going to work can require several hours a day but also every errand, every event people want to participate it. These hours of travelling can be unproductive but once wireless connections to the internet are available, lots of activities can be fulfilled. No matter if public transportation by bus or train, or individual automobile transportation, lots of wireless connections are required. But not only commuting people are causing high data traffic, also the busses, cars and trains themselves will participate in data traffic generation and finally even dominate due to many desirable features in security, safety and fleet control. The wireless internet today is based on a nomadic use for its participants.
In this talk the challenges for future wireless networks to deal with such mass of high mobile users are being set and first ideas to overcome such hurdles with existing and future networks are being proposed. The limitations of the current physical layer are explained, potential solutions such as Hetnets, distributed antennas and eMBMS services are explained in such context. Only few of the recent ideas proposed for 5G networks appear as feasible solutions.

WORKSHOP - REASONING-BASED INTELLIGENT SYSTEMS
Chairman: Prof. Dr. Kazumi Nakamatsu, University of Hyogo, Japan

Big Data Approach in an ICT Agriculture Application
Prof. Alireza AHRARY, PhD, Sojo University, Japan

Abstract: In the case of Agriculture, what kind of technologies can be applied in agriculture? Specifically, ICT technologies can be represented by any device, tool or application that allows the end user to exchange, share or collect any desire information. The term ICT encloses every technology from satellite connections to mobile devices or sensor networks. An new affordable devices enable developing countries to access this technology, allowing farmers to take advantage of the different characteristics. In this presentation we will introduce the ICT in agriculture in Japan and talk about the most important difficulties in this area.

From Search Engine to Text Mining
Prof. Dr. Sachio HIROKAWA, Kyushu University, Japan

Abstract: Search Engine is a "must" in today's life, not only for ordinary people but also business person or researcher. Use of search engine varies according to the purpose of search. When we want to know a nice restaurant, we can check the list of gourmet blog found as the search result. It won't take a few seconds. However, when we - researchers - make a survey or search on related work, it takes a few weeks or months. We have to find related articles, researchers and keywords. Moreover, we have to continue further search on those researchers and those keywords we found in the first stage search. In other words, we have to keep searching. In the present talk, I'd like to introduce my advanced search engines, "cross table" search engine and "mind map" search engine, which performs text mining.

Paraconsistent Logics and Applications
Prof. Dr. Jair Minoro ABE, Paulista University, Brazil

Abstract: In this work we summarize some of the applications of so-called Paraconsistent logics, mainly one class of them, the paraconsistent annotated logics. Roughly speaking such systems allow inconsistencies in a non-trivial manner in its interior; so it is suitable to handle themes in which inconsistencies become a central issue, like pattern recognition, non-monotonic reasoning, defeasible reasoning, deontic reasoning, multi-agent systems including distributed systems, collective computation, among a variety of themes.

TABLE OF CONTENTS

ORAL SESSIONS

COMPUTER SYSTEMS AND INTERNET TECHNOLOGIES I

COMPUTER SYSTEMS AND INTERNET TECHNOLOGIES II

I

DIGITAL IMAGE PROCESSING

ENERGY SYSTEMS AND EFFICIENCY

ENGINEERING EDUCATION

INFORMATICS AND COMPUTER SCIENCE

MEASUREMENT SCIENCE AND TECHNOLOGY

RADIO COMMUNICATIONS, MICROWAVES, ANTENNAS

SIGNAL PROCESSING

TELECOMMUNICATION SYSTEMS AND TECHNOLOGY

POSTER SESSIONS

POSTER SESSION 1 - COMPUTER SYSTEMS AND INTERNET TECHNOLOGIES

POSTER SESSION 2 - CONTROL SYSTEMS

POSTER SESSION 3 – DIGITAL IMAGE PROCESSING

POSTER SESSION 4 – ELECTRONICS

POSTER SESSION 5 – ENERGY SYSTEMS AND EFFICIENCY

POSTER SESSION 6 – ENGINEERING EDUCATION

POSTER SESSION 7 – INFORMATICS AND COMPUTER SCIENCE

POSTER SESSION 8 – MEASUREMENT SCIENCE AND TECHNOLOGY

Higher School of Transport "T. Kableshkov", Sofia, Bulgaria

M. Anđelić,V. Tasić, M. Pavlov-Kagadejev, D. Brodić*, I. Stojković**
Institute of Mining and Metallurgy, Bor, Serbia
*Technical Faculty in Bor, University of Belgrade, Bor, Serbia
**Faculty of Electronic Engineering, Niš, Serbia

POSTER SESSION 9 – RADIO COMMUNICATIONS, MICROWAVES, ANTENNAS

H. Gochev, B. Bonev, P. Petkov
Technical University of Sofia, Bulgaria

P. Petkov
Technical University of Sofia, Bulgaria

J. Shikoski, R. Arnaudov
Technical University of Sofia, Bulgaria

O. Panagiev
Technical University of Sofia, Bulgaria

ORAL SESSION
COMPUTER SYSTEMS AND INTERNET TECHNOLOGIES I

Access to Mobility Information in M2M Communications

Anastas Nikolov[1], Evelina Pencheva[1], Ivaylo Atanasov[1]

*Abstract –***Service capabilities in Machine-to-Machine (M2M) communications represent generic functionality accessible for different applications through application programming interfaces. For M2M communications Representational State Transfer (REST) style is adopted. This paper presents an approach to design of RESTful web services for M2M mobility. An abstraction of mobility information is synthesized and represented as resources accessible through four standard operations: create, retrieve, update and delete. Implementation aspects are considered.**

Keywords – **Machine-to-Machine Communications, Service capabilities, Resource structure, Web Services**

I. INTRODUCTION

Machine-to-Machine (M2M) communications allow intelligent objects which are uniquely identified and gather data from their environment to exchange information with network applications. In order to deploy the M2M solutions in a wide scale it is necessary to design capabilities that can be reused across several applications [1], [2]. Best software practices in information technology and communications (ICT) adopt separation of applications from service capabilities and network capabilities.

Service capabilities are software modules that are exposed to M2M applications through the use of application programming interfaces (APIs). Several API sets have been designed in the ICT area. The Web services model has been used to define communication functions such as call control, messaging, charging and access to mobility information.

The Web services model is based on the assumption that communications are like invocation of remote service whose nature can be ignored. But this model is not suitable for M2M applications, as M2M devices are constrained resources with tangible states that can be manipulated. REST (REpresentational State Transfer) is adopted as a method for M2M modeling. In REST, each physical or logical entity is represented as a resource which has particular state. The resources can be addressed through HTTP Uniform Resource Identifier (URI) and its states can be retrieved and updated. Resources can be created and destroyed respectively.

In this paper, we present an approach to design Mobility service capability that may be used to build horizontal service platform. Mobility as an M2M service capability provides information about device location. The access to the location of a device is through a request for the device location, a notification of a change in the device location and notifications of device location on a periodic basis. The related research concerning access to location information

[1]The authors are with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mails: nikolov.anastas@gmail.com; enp@tu-sofia.bg; iia@tu-sofia.bg

addresses specific solutions, but does not study the necessary generic functionality [3], [4], [5], [6], [7]. The related implementations deal with positioning methods and the usage of location data, but do not concern programmability issues. We apply the REST architectural style to identify the generic context functions providing access to location information and to design valuable resource structure with operations that can manipulate it.

Our approach, presented in the next sections, consists of annotation of semantic information for location services and synthesis of resource structure as specified in ETSI TS 102 690. The procedures used to manipulate the location information are described as create, retrieve, update and delete operations. Some implementation aspects of the designed RESTful services are considered.

II. ANNOTATION OF LOCATION SERVICE INFORMATION

The location information depends on the application. For example it is required for healthcare and fleet tracking applications but not for smart metering. Additionally, location information may be reported periodically or on demand, and notifications of distance changes between monitored devices may also be available to applications.

The location information is presented by latitude, longitude, altitude, accuracy and time stamp. Latitude, longitude and altitude values are expressed as floating point numbers. The accuracy values express the desire of the application for the location information to be provided. The choice of values may influence the price that the service provider charges. In triggered notifications, a tracking accuracy is defined. Two accuracy values (requested and accepted) may be used. For example, a taxi tracking service that locates the nearest taxi to the client requires fine grained accuracy while coarse grained accuracy may be appropriate for a truck nearing the vicinity of a warehouse. The accuracy of location provided in meters is expressed as an integer number. In some applications, the maximum age of location information may be useful, e.g. the location information may be cached rather than directly accessed. The maximum acceptance age, in seconds, is expressed in integers.

Fig.1 shows the resource structure for device location. Following the ETSI resource structure defined in TS 102690, the deviceLocation resource contains standard attributes common for all resources such as expirationTime, accessRightID, creationTime, and lastModifiedTime (not shown in Fig.1) and specific attributes such as deviceLatitude, deviceLongitute, deviceAltitude, accuracy and timestamp. The deviceLatitude, deviceLongitude and deviceAltitude attributes represent the measured device location and the accuracy attribute represents the measurement accuracy. The timestamp attribute represents the date and time that location

was collected. The <subscriptions> sub-resource of the deviceLocation resource contains a collection of 0..n <subscription> resources which represent active subscriptions to location information. Each <subscription> resource in addition to specified in ETSI TS 102 690 mandatory attributes has also requestedAccuracy and acceptedAccuracy attributes. The requestedAccuracy express the range in which the subscribed application wants to receive location information. The acceptedAccuracy expresses the range that the subscribed application considers to be feasable. If the location cannot be provided within this range, the application prefers not to receive the information. The <deviceDistances>, <devicePeriodicReporting> and <deviceChangeReporting> are sub-resources of the <deviceLocation> resource.
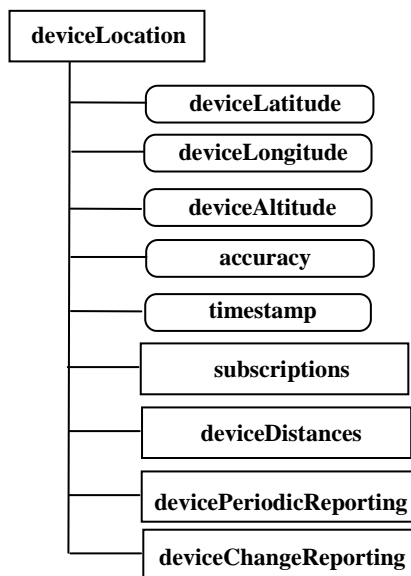
Fig.1 Structure of deviceLocation resource

Some applications may be interested in device distance from a location. The <deviceDistances> collection resource represents the collection of <deviceDistanceFrom> resources. The <deviceDistanceFrom> resource structure is shown in Fig.2, where the remoteLatitude and remoteLongitude attributes represent the latitude and longitude of the location to measure from, respectively. The distance attribute represents the distance from device to the location specified in meters. The subscriptions sub-resource of the <deviceDistances> resource contains a collection of 0..n <subscription> resources which represent active subscriptions to device distance from the specified location.
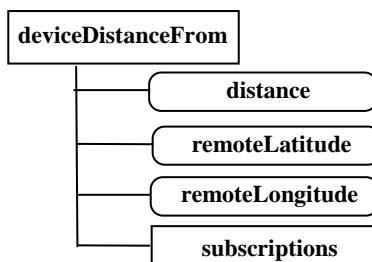
Fig.2 Structure of deviceDistanceFrom resource

Notifications of device location may be provided on a periodic basis. The periodic notifications provide location information at an application defined interval. The <devicePeriodicReproting> collection resource represents the collection of <devicePeriodicLocation> resources. Fig.3 shows the <devicePeriodicLocation> resource structure.
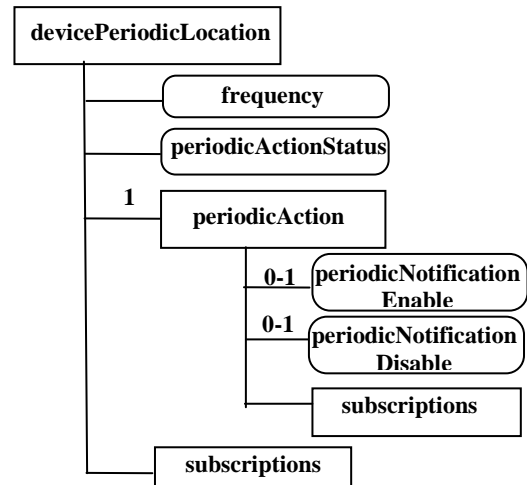
Fig.3 Structure of devicePeriodicLocation resource

The frequency attribute represents the minimum time between notifications (maximum frequency of notifications can also be considered). The periodicActionStatus attribute indicates the status of the action. The periodicAction resource represents the action. The periodicNotificationEnable attribute represents the action that enables the periodic notification. The periodicNotificationDisable attribute represents the action that disables the periodic notification.

An application can be notified of a device entering or leaving a geographical area. When a matching event occurs, a notification message will be sent to the application. An application may define a target area and notification criteria e.g. entering the target area or leaving the target area. The <deviceChangeReporting> collection resource represents the collection of <deviceLocationChange> resource. The <deviceLocationChange> resource structure for triggered location change notifications is shown in Fig.4. The locationChangeCriteria attribute is of enumerated type (entering or leaving an area). The areaLatitude, areatLongitude attributes represent the latitude and longitude of the center point, and radius attribute represents the radius of the circle around the center point in meters. The triggeredActionStatus attribute indicates the status of the action. The <triggeredAction> resource represents the action.

III. PROCEDURES FOR LOCATION SERVICES

Each procedure related to manipulation of device location information consists of Create, Retrieve, Update and Delete. For collection resource such as subscriptions and deviceDistances collection resources only Retrieve and Update procedures are defined for managing the retrieval and update information associated with the specified collection

resources. Such type of resource (like any other collection) cannot be created or deleted by means of request.
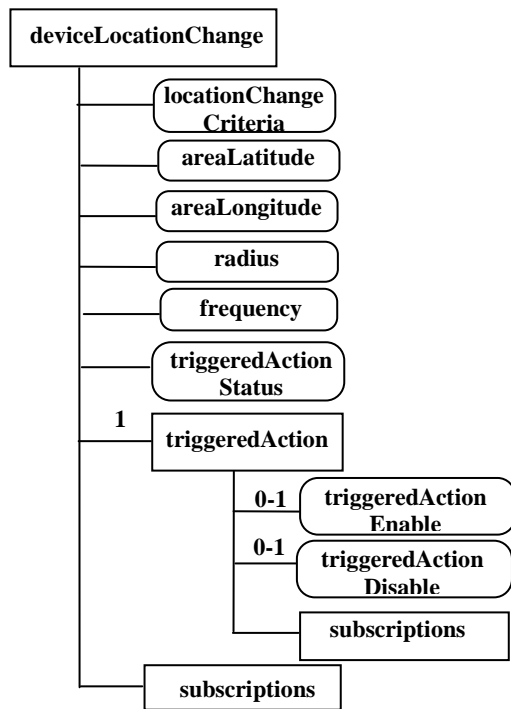
Fig.4 Structure of deviceLocationChange resource

Fig.5 illustrates the procedure used to retrieve the representation of deviceDistances collection resource. The representation includes the values of all the attributes and the references to the deviceDistanceFrom child resources for which the issuer is authorized to discover.
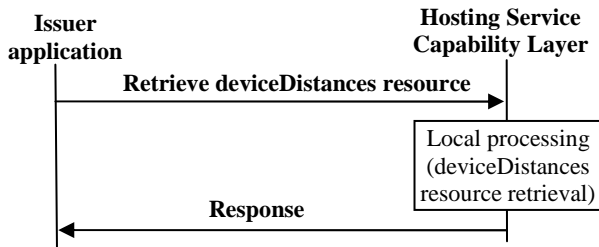
Fig.5 Retrieval of deviceDistances resource

Create procedure is to create a resource as a child of a collection resource. Fig.6 illustrates the procedure for creation of <deviceDistanceFrom> resource as a sub-resource of the <deviceDistances> collection resource.
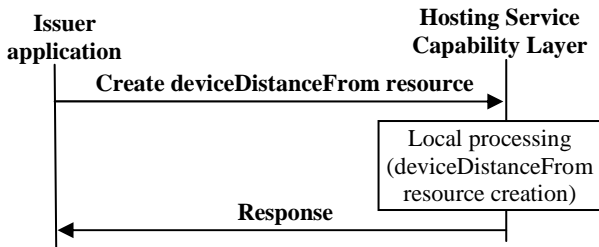
Fig.6 Creation of deviceDistanceFrom resource

Delete procedure are used to remove resources. For example, Delete <subscription> resource procedure may be used to delete an active subscription which means that the subscriber unsubscribes from notifications.

Update procedures are used to update/modify the content of existing resources. Fig.7 illustrates the procedure used to modify the attributes of the <devicePeriodicLocation> resource (e.g. frequency of notifications).
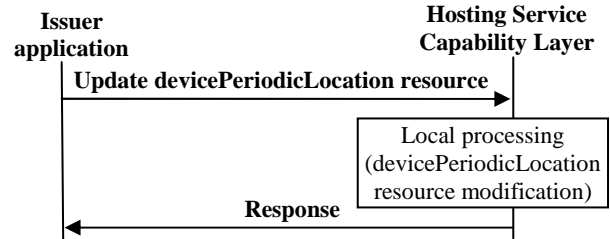
Fig.7 Update of devicePeriodicLocation resource

Table 1 lists the specific operations related to location service and the corresponding REST procedures used for these operations.

Table 1 Supported location service operations

Specific location service operations	REST procedures
getDeviceLocation	Retrieve <deviceLocation> resource
getDeviceDistance	Retrieve <deviceDistance> resource
startPeriodicNotification	Update <priodicAction> resource to enable
stopPeriodicNotification	Update <priodicAction> resource to disable
startTriggeredNotification	Update <triggeredAction> resource to enable
stopTriggeredNotification	Update <triggeredAction> resource to disable
deviceLocationNotification	Update <deviceLocation> resource
deviceLocationChanged	Update <deviceLocation> resource

The getDeviceLocation operation retrieves the location of a device. The getDeviceDistance operation determines the distance of a device from a location. The startPeriodic-Notification operation makes available notifications of a device location periodically, while the stopPeriodic-Notification operation disables periodic notifications about a device location. Similarly, the startTriggeredNotification operation makes available notifications of a device location change, while the stopTriggeredNotification operation disables notifications about a device location change. The deviceLocationNotification operation notifies periodicaly the location of a monitored device. The deviceLocationChanged operation notifies about location changes of a monitored device. These operations are implemented by retrieving or updating the content of the resources.

For example, in order to start device location change reporting, the <triggeredAction> resource has to be modified to triggeredActionEnable. Notify procedure is used to indicate the operation for reporting a notification about a change of a resource as a consequence of a subscription. Notify procedure is mapped to an Update procedure as the asynchronous mechanisms are used. For example, Fig.8 illustrates the notification procedures about device location as a result of periodic location reporting.
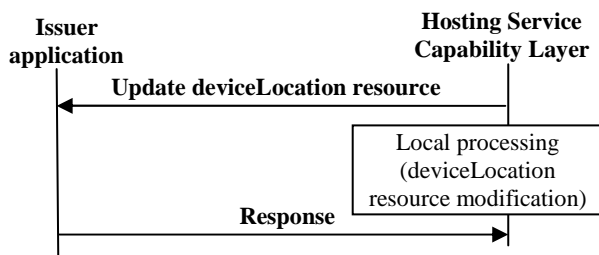
Fig.8 Periodic reporting of device location

Typical exceptions that may arise include the following: accuracy is out of limit, invalid remote location, service error, and invalid input value.

IV. IMPLEMENTATION ASPECTS

Location information may be used by different M2M applications for healthcare and transportation. Asset tracking functionality is useful in tracking of high-value assets such as intravenous pumps, wheel chairs, and stretchers within hospitals, and marking medications with RFIDs to eliminate errors in hospitals when administering medications, as well as in for tracking objects within trucks or other forms of transportation. Location information is valuable in fleet management to ascertain location of vehicles and sending dispatch notifications, in insurance for remotely monitor location and theft prevention, for emergency support, navigation, toll ticketing, remote patient monitoring, etc.

The support of the designed resource structure is up to device capabilities. Depending on its computing capabilities, the device may provide the full range of the designed operations or just some of them.

Each resource, in REST terms, can be addressed using HTTP URI. The resource is located inside a web server using the host, port, and path parts of the URI. The action to be taken is determined by the HTTP request. The HTTP request GET is used to retrieve the resource state, the HTTP request PUT is used to update the resource state. The HTTP request POST and DELETE can be used to create (to add) and destroy a resource.

Let us assume that the device runs a mobility application named deviceMobility that will generate location data. The device runs a device service capability layer (DSCL) that is configured to have the <sclBase>: HTTP://device5555.example.com. The network application that uses location services is already registered with the network service capability layer (NSCL) agreed to have <sclBase>: HTTP://mobility.com. The network application subscribes for registering of devices. When a device registers to DSCL, the DSCL requests registration through the creation of a resource under HTTP://example.com/scl/ with the resource identifier device5555, which will be the unique device identifier. The hosting NSCL responds positively to the request. After the registration of the device application deviceMobility to the DSCL and announcing to the NSCL, the device application requests the creation of a deviceLocation resource under the HTTP://example.com/scls/device5555/applications/deviceMobility/ collection with the identifier deviceLocation. The device location will be addressable through the link HTTP://example.com/scls/device5555/applications/deviceMobility/deviceLocation using HTTP GET request.

On request for device location, an HTTP response is returned where the location information is provided in the response body e.g. in JSON format: {"latitude":42.6795551, "longitude":23.2916345,"altitude":602.2, "accuracy": 20, "timestamp": 1388678400} where "timestamp" is the number of milliseconds elapsed since 1 January 1970 00:00:00 UTC (UNIX TimeStamp), the value is equal to 'Wed, 01 Jan 2014 23:00:00 GMT', and the location is Sofia, Bulgaria.

Following the same procedures the device application may request creation of child resource of deviceLocation resource.

Another protocol that may be used in REST-based architectures is CoAP (Constrained Application Protocol), which defines some primitives that allow REST on the top of TCP and UDP.

V. CONCLUSION

Location services may be used in different M2M application area. Design of M2M service capabilities for mobility provides access to location information that can be shared by different applications through reusable software modules. The paper presents an approach to design REST-based service capabilities for mobility. The semantic information related to mobility is synthesized by identification of basic use cases. The mobility semantic information is presented in a resource tree structure, where resources may be manipulated through their states. REST-based location services are designed by definition of operation performed on the resources. As each of the resources is uniquely addressable, it can be accessed using standard HTTP methods or CoAP primitives.

REFERENCES

[1] C. Pereira, A. Aguiar, "Towards Efficient Mobile M2M Communications: Survey and Open Challenges", *Sensors* vol.14, pp.19582-19608; DOI: 10.3390/S141019582, 2014.

[2] C. Im, C. Jeong, "ISOMP: Instant Service Orchestration Mobile M2M Platform", *International Journal of Distributed Sensor Networks,* Hindawi, Article ID 298251, 2015.

[3] S. K. Datta, C. Bonnet, "Smart M2M Gateway Based Architecture for M2M Device and Endpoint Management", Available at: http://www.Eurecom.Fr/Fr/Publication/4318/ Download/Cm-Publi-4318.Pdf, 2014.

[4] J. Kim, J. Lee, J. Kim, J. Yun, "M2M Service Platforms: Survey, Issues, and Enabling Technologies", *IEEE Communications Surveys & Tutorials*, vol.16, no.1, pp. 61-76, pp.2014.

[5] N. A. Surobhi, A. Jamalipour, "M2M-Based Service Coverage For Mobile Users In Post-Emergency Environments," *IEEE Transactions On Vehicular Technology*, vol. 63, no. 7, pp. 3294-3303, 2014.

[6] S. Wahle, T. Magedanz, F. Schulze, "Demonstration of OpenMTC – M2M Solutions for Smart Cities and the Internet of Things", Available at: http://Www.Ieeelcn.Org/Prior/ LCN37/Lcn37demos/Lcndemos12_Wahle.Pdf, 2013.

[7] F. Bai, K. S. Munasinghe, A. Jamalipour, "A Novel Information Acquisition Technique For Mobile-Assisted Wireless Sensor Networks," *IEEE Transactions on Vehicular Technology*, vol. 61, no. 4, pp.1752-1761, 2012.

Ontology Based Data Model for Power Control in Smart Homes

Ivaylo Atanasov[1], Anastas Nikolov[1], Evelina Pencheva[1]

Abstract – **One of the key enablers of service capabilities in Machine-to-Machine (M2M) communications which can be reused by different applications are data models. Data models explicitly determine the structure of data exchanged between M2M applications. The paper presents an ontology based model of home appliance power control. The model presents the domain knowledge as a set of concepts and the relationships between concepts. The model description is formalized using Ontology Web Language (OWL).**

Keywords – **Machine-to-Machine Communications, Semantic information, Formal representation**

I. INTRODUCTION

In recent years, the concept of creating more smart homes gained popularity. Smart Home has a certain degree of "intelligence" which serves people living in the home. The main objective is to increase the comfort of residents, while minimizing energy consumption and reducing costs. This idea already embedded into another technological concept, namely Machine-to-Machine (M2M), acquires clearly defined and realizable projection in many areas of our daily lives [1], [2], [3], [4].

The potential of M2M system to provide applications with capabilities to discover, interpret and use data from smart objects (M2M data) is essential for the creation of high-level M2M services and for the development of an open M2M data market. To enable data transmission in an M2M system, it is necessary to synthesize semantic information. Through such semantic information, applications can discover data without prior knowledge about their existence [5].

In this paper, we present a structural approach to semantic annotation for power control of appliances in a smart home. Survey on research works in the area shows that adding semantic information to a system can be done in various forms [6], [7], [8], [9], [10]. Our approach is to focus on what is sensed and acted upon and it does not limit the M2M system to considerations of a sensor network. Devices are modeled by a generic approximation with minimal information and the system should be able to integrate them using this information. This semantic information describes the devices as things which may be considered as device shadows in a M2M system.

The paper is structured as follows. First we present the ETSI approach to structuring semantic data of domain-specific knowledge. Next we describe the ontology for the power control in a smart home. Then we model the data

[1]The authors are with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mails: iia@tu-sofia.bg; nikolov.anastas@gmail.com; enp@tu-sofia.bg

following the ETSI approach. Parts of formal model description with the OWL language are presented. The conclusion summarizes our contribution.

II. THE ETSI APPROACH TO MODELING M2M SEMANTIC DATA

In (ETSI TR 101 584), it is outlined the approach to modelling M2M semantic data. The main guiding principles are separation of the abstract information model from its representation in the ETSI M2M resource structure and separation of domain-specific knowledge from instance-specific data. The synthesis of data abstraction needs to be independent of ETSI M2M resource architecture defined in ETSI TS 102 60 and which may be updated in future releases. With separation of instance and domain knowledge, then the instance-specific data can be understood without any further a-priori knowledge, and different applications can interact on the bases of common domain model.

Fig.1 shows the structure of ETSI domain specific knowledge.
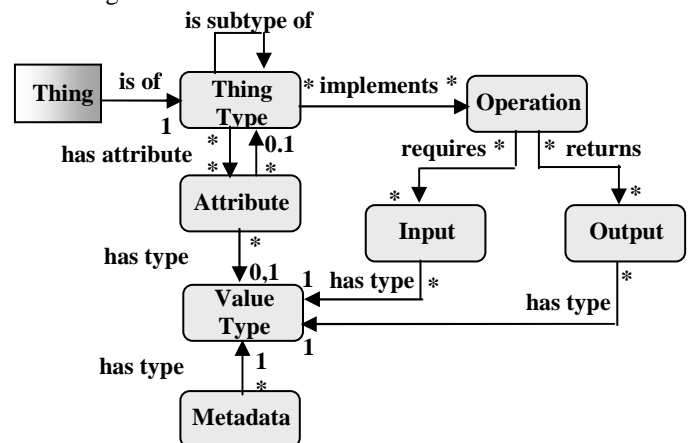
Fig.1 Structure of domain specific knowledge (ETSI TR 584)

Ontology is used to express the information represented by an application domain model. The model identifies Thing types, which have attributes, data about data (metadata) and implements operations which require inputs and return outputs of value types. Attributes, metadata, inputs and outputs have certain value types.

III. ONTOLOGY FOR POWER CONTROL IN A SMART HOME

Let us consider a home energy saving system (HESS) whose general purpose is energy saving through power

control on electrical appliances at home. For example, HESS may control the heating, cooling and ventilation home systems and the hot water tanks. For heating and cooling the energy saving is achieved by small changes in the referenced (preferred by the home owner) temperature during predefined time periods when nobody is at home or everybody sleeps [11]. The ontology web that represents the domain knowledge is shown in Fig.2.
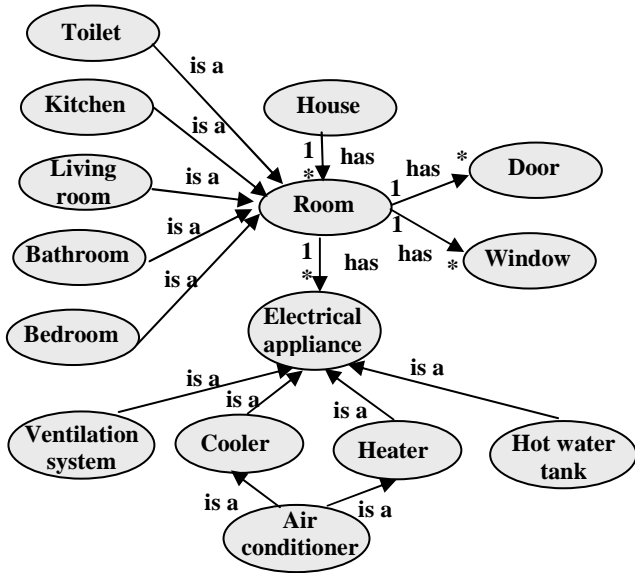
Fig.2 Ontology web for power control in a smart home

The basic concepts in the domain of power control in a smart home are as follows. A house has rooms, a room has windows and doors, a room has electrical appliances that may be remotely controlled. The room may be a kitchen, a living room, a bathroom, a toilet and a bedroom. The home owner may define power shedding schedule for heating/cooling control and a policy stating that if a room door or a room window is open, the heater/cooler has to be switched off.

IV. DOMAIN-SPECIFIC MODEL FOR POWER CONTROL IN A SMART HOME

The domain specific model defines Thing types, Thing type attributes and operations. Fig.3 shows the domain-specific model of power control of electrical appliances in a room.

The *ElectricalAppliance* type has the following attributes. The *ElectricalOperation* attribute may have on-off or on-off-standby values. The *UsageMode* attribute may have autonomous or manual (user directed) values. The *Usage* attribute may have periodic, permanent or semi-random values. Only electrical appliances that have autonomous usage mode and on-off-standby electrical operation may be controlled remotely. The *ElectricalAppliance* type implements the following operation. The *powerOn* operation is used to power on the appliance, while the *powerOff* operation is used to power off the appliance. A controllable appliance such as a cooler, heater, air conditioner, ventilation system and hot water tank, when is powered on, goes in a standby state.
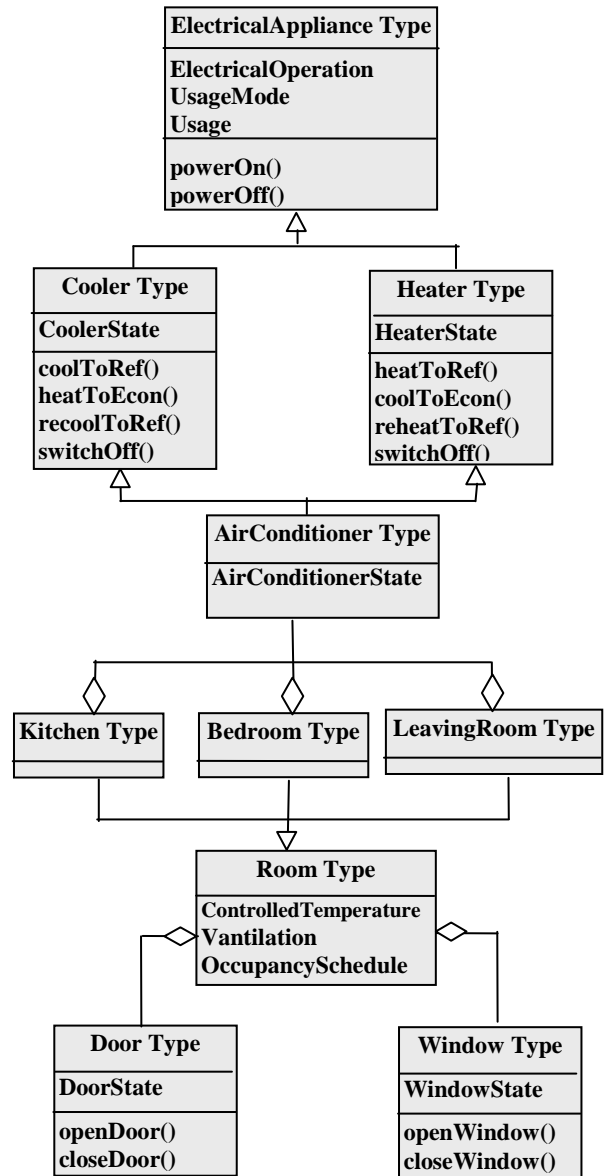
Fig.3 Domain information model of controllable appliances

The *CoolerState* attribute of the *Cooler* type indicates the cooler state (e.g. powered off, standby, etc.).

The *Cooler* type implements the following operations. The *coolToRef* operation instructs the cooler to cool to the referenced temperature T_{ref} (the temperature preferred by the home owner) and it requires T_{ref} as an input. The *heatToEcon* operation instructs the cooler to switch off while the room temperature increases by a small amount (Δ) to realize energy economy and it requires Δ as an input and $T_{ref-\Delta}$ as an output. The *recoolToRef* operation instructs the cooler to re-cool to the referenced temperature and it requires T_{ref} as an input. Similarly, the *Heater* type implements the following operations. The *heatToRef* operation instructs the heater to heat to the referenced temperature T_{ref} and it requires T_{ref} as an input. The *coolToEcon* operation instructs the heater to switch off while the room temperature decreases by a small amount (Δ) to realize energy economy and it requires Δ as an input and $T_{ref-\Delta}$ as an output. The *reheatToRef* operation

instructs the heater to reheat to the referenced temperature and it requires T_{ref} as an input. The *HeaterState* attribute of the *Heater* type indicated the heater state. Both *Cooler* type and *Heater* type implement the *swithOff* operation which instructs the appliance to switch off. More information about heating/cooling control may be found in [11].

The *Room* type has the following attributes. The *ControledTemperature* attribute indicates whether the room is heated or not, e.g. the toilet is an unheated room, while the bedroom is a heated one. The *Ventilation* attribute indicates whether the room has a fan that pumps the air or not. The *OccupancySchedule* attribute indicates occupancy schedule of the room. The *Door* type implements *openDoor* operation and *closeDoor* operation and it has a *DoorState* attribute. The *Window* type implements *openWindow* operation and *closeWindow* operation and it has a *WindowState* attribute.

V. FORMAL DESCRIPTION OF THE DOMAIN-SPECIFIC MODEL

This section describes parts of the formal description of the proposed semantic information model of air conditioner device using OWL 2.

The description of the domain-specific model for power control of electrical appliances in a smart home is formalized by the OWL (Ontology Web Language) which is used to represent ontologies in the web. OWL is a language based on computational logic, so that the knowledge expressed in OWL can be analyzed by a computer program in order to verify compliance of this knowledge and turn it from implicit to explicit [12]. OWL ontologies can be published on the World Wide Web and can refer to other ontologies or directed by them.

The software product Protégé is used for the creation and verification of the ontology of the domain-specific model for power control in a smart home. The product fully complies with the standardized specification of the WWW Consortium to create OWL 2 and RDF descriptions [13].

Every "thing" in the OWL world is a member of the class owl:Thing. So, every class is implicitly a subclass of owl:Thing. Class definition has two parts: a name and a list of restrictions. Instances of the class belong to the intersection of these restrictions. Element subClassOf is a down of possible restrictions. Instances are members of classes.

The properties allow describing the general relations for class members and specific facts about instances. There are two types of properties:

- properties of the types of data that describe the relations between instances of classes and RDF literals and XML Schema data types;
- properties of objects that describe the relations between instances of two classes.

When defining the properties, e.g. relation, there are different ways to impose restrictions. As the relations in mathematics, properties are defined with domain (domain) and the range of values (range).

Statements in Figure 4 declare a fragment of named individuals, classes, and object properties.

Figure 5 describes the part of of the relationships between declared individuals, classes and attributes.

Figure 6 shows the data type declarations and definitions.

VI. CONCLUSION

Data models representing semantic information of the domain specific knowledge allow information sharing between different applications. We propose a structural approach to modeling semantic information for power control of electrical appliances in a smart home. A target physical system with distributed monitoring sensors and controlling actuators is represented in the digital world as a set of things which might be controlled or monitored by M2M applications. Things are identified by approximation to a generic model and the M2M system integrates them on the basis of semantic information. Using domain-specific ontology, we define things types, which have attributes and implement operations with input and output parameters. The data model is formally described by OWL and verified by the use of the Protégé tool.

REFERENCES

[1] C. Hung, Y. Bai, P. Chen, J.-M. Hsu, "Remote control of home lighting devices with RFID identification and current detection of outlets", IEEE 17th International Symposium on Consumer Electronics ISCE '13, Taiwan, Conference Proceedings, pp.7-8, 2013.

[2] K. Hwang, S. Yoon, "A Multifunctional RF Remote Control for Ultralow Standby Power Home Appliances", *International Journal of Distributed Sensor Networks,* Available at: http://www.hindawi.com/journals/ijdsn/2014/381430/, 2014.

[3] V. Kumar, A. Fensel, G. Lazendic, U. Lehner. „Semantic Policy-Based Data Management for Energy Efficient Smart Buildings". OTM'2012 Workshops, *Lecture Notes in Computer Science*, vol.7567, pp.272-281, 2012.

[4] G. Wu, X. Yu, "Remote control system for energy efficient home", IEEE Conference Energytech'2013, Ohio, USA, Conference Proceedings, pp.1-5, 2013.

[5] A. Fensel, S. Tomic, V. Kumar, M. Stefanovic, S. Aleshin, D. Novikov, "SESAME-S: Semantic Smart Home System for Energy Efficiency", *Informatik Spektrum*, vol.36 issue 1, pp.46-57, 2013.

[6] Z. Qunzhi, S. Natarajan, Y. Simmhan, V.Prasanna, "Semantic Information Modeling for Emerging Applications in Smart Grid", International Conference on Information Technology: New Generations ITNG'2012, Las Vegas, USA, Conference proceedings pp. 775-782, 2012.

[7] S.Rohjan, M.Uslar, M. Specht, A. Niesse, M. Troschel, "Towards Semantic Service Integration for Automation in Smart grids" *International Journal of Distributed Energy Sources*, vol.8 number 2, pp.119-146, 2012.

[8] S. Rohjans, K. Piech, W. Mahnke, "Standardized Smart Grid Semantics using OPC UA for Communications", *IBIS – Interoperability in Business Information Systems*, vol.6, issue 10, pp.21-32, 2012/

[9] Y. Simmhan, Q. Zhou, V. Prasanna, "Semantic Information Integration for Smart Grid Applications," Chapter 19 in edited book *Green IT: Technologies and Applications*, pp.361-380, 2012.

[10] Y. Tang, I. Ciuciu, "Semantic Decision Support Models for Energy Efficiency in Smart-Metered Homes", IEEE

International Conference on Trust, Security and Privacy in Computing and Communications TrustCom'2012, Liverpool, UK, Conference Proceedings. pp. 1777-1784, 2012.

[11] I., Atanasov, M. Ivanov, A. Nikolov, R. Dimova, E. Pencheva. "Towards web services for power control", International Journal on Information Technologies & Security, IJITS, 2014, no.4, pp.3-16.

[12] W3C OWL 2 Web Ontology Language Primer, Second edition, W3C Recommendation, 2012.

[13] Protégé, Stanford University, Available at: http://protege.stanford.edu/, Accessed 2015.

```
Declaration ( NamedIndividual ( :TheAirConditioner ) )
Declaration ( Class ( :AirConditioner ) )
Declaration ( Class ( :Appliance ) )
AnnotationAssertion ( rdfs: comment :Appliance "Represents the set of all appliances." )
Declaration ( ObjectProperty ( :hasUsage ) )
Declaration ( ObjectProperty ( :hasElectricalOperation ) )
Declaration ( ObjectProperty ( :hasUsageMode ) )
Declaration ( ObjectProperty ( :hasAirConditionerID ) )
Declaration ( ObjectProperty ( :implementsCooling ) )
Declaration ( ObjectProperty ( :implementsReheating ) )
Declaration ( ObjectProperty ( :implementsHeating ) )
Declaration ( ObjectProperty ( :implementsReCooling ) )
Declaration ( ObjectProperty ( :requiresTref ) )
Declaration ( ObjectProperty ( :requiresDelta ) )
Declaration ( ObjectProperty ( :requiresTrefminus ) )
Declaration ( ObjectProperty ( :requiresTrefplus ) )
Declaration ( ObjectProperty ( :requiresTrefplus ) )
Declaration ( ObjectProperty ( :requiresTrefminus) )
Declaration ( ObjectProperty ( :returnsTrefminus) )
Declaration ( ObjectProperty ( :returnsTrefplus ) )
```

Fig.4. Declarations related to OWL description of air conditioner information model

```
SubClassOf ( :AirConditioner :Appliance)
ClassAssertion ( :AirConditioner :TheAirConditioner)
SubObjectPropertyOf ( :requiresTref  :implementsCooling)
SubObjectPropertyOf ( :requiresDelta :implementsCooling)
SubObjectPropertyOf ( :returnsTref-  : implementsCooling)
SubObjectPropertyOf ( :requiresTref- :implementsReheating)
SubObjectPropertyOf ( :requiresDelta  :implementsReheating)
SubObjectPropertyOf ( :returnsTref  : implementsReheating)
```

Fig.5 OWL relationships description in the an air conditioner information model

```
Declaration ( Datatype ( :UsageMetadata )
Declaration ( Datatype ( :ElectricalOperationMetadata )
Declaration ( Datatype ( :UsageModeMetadata )
DatatypeDefinition (
        :UsageMetadata
        DataOneOf ( "1"^^xsd:integer "2"^^xsd:integer "3"^^xsd:integer"
        Annotation (rdfs:comment "1 means permanent, 2 means semi-random, 3 means periodic" )
        )
)
DataPropertyRande  (:requiresDelta xsd:nonNegativeFloat)
DataPropertyRande ( :hasAirConditionerID xsd:nonNegativeInteger)
```

Fig.6. OWL data type declarations and definitions in the air conditioner information model

BikeEge - Bicycle Sharing System for Ege University Students

Nese BAZ [1], Bahadir AKTAS [1], Nail AKCURA[2], Radosveta SOKULLU,[1] Erol UYAR[3]

Abstract – **In the last decade Bike Sharing systems have become very popular all over the world. There are numerous examples bike sharing systems designed for big metropolitan cities which help reduce both traffic and CO2 emissions. More recently similar systems have become popular for universities campuses. In this paper we present the architecture and details of the free bicycle sharing system designed for Ege University students.**

Keywords – **bike sharing systems, system architecture, user inferface**

I. INTRODUCTION

A Bicycle Sharing System (BSS) is an alternative type of transportation system that allows bicycles, belonging to a certain organization to be used in a shared manner by a group of designated users, mainly for short-term short-distance transportation. Until the end of 2014 there were over 200 cities in the world that had registered bike sharing systems, with the number of bicycles varying from 50 (Dumfries, UK) to 65000 (Hangzhou, China) and stations from 10-15 to several thousand. Supported by a sophisticated software and network database, including an hourly based payment subsystem such BSS relieve the users from transporting and keeping a bike, while at the same time allow them to benefit from all the advantages that biking in cities can provide. In the last decade similar systems have been started in a number of universities all over the world. Most interesting examples in this respect are ZotWheels. (University of California, Irvine automated bike share program, inaugurated in 2009), viaCycle@GT, (Georgia Institute of Technology, Atlanta, Georgia), Reggie Ride (Illinois State University, Illinois), Crimson Bikes (Harvard University, Cambridge, MA), NYU Bike Share (New York University, New York) and Bike Emory, (Emory University, Atlanta, Georgia) to mention just a few. Generally these are not-for-profit enterprises, and city programs are usually sponsored by large banks or city municipalities, while university bike sharing is usually installed, operated and backed up by university funding.

[1]Nese Baz, Bahadir Aktas and Radosveta Sokullu are with the Department of EEE, Faculty of Engineering, Ege University, Izmir, Turkey: corresponding author email: radosveta.sokull@ege.edu.tr

[2]Nail Akcura is with the Dept. of EEE, at Katip Celebi University, Izmir Turkey

[3] Erol Uyar is with the Mechatronic Department, Institute of Natural Sciences, Ege University, Izmir, Turkey

In this paper we will present the details of such a prototype systems that is currently being developed in Ege University, Izmir, Turkey. The project, led by Prof. Dr. Fazil Apaydin, includes participants from different departments from the Faculty of Engineering, the EU Student Biking Society and it is also supported by Accell, the largest bike manufacturing company in Turkey.

II. BIKE SHARING SYSTEM ARCHITECTURE

A. Operational principles of a BSS

A bike sharing system in general consists of a number of stations, where bicycles are located, one main terminal at each station, which is used for carrying out the renting and returning operations and a main server located somewhere else. Docking stations are usually fixed mounted and bicycles are locked/unlocked to them. The terminal (kiosk) consists of a user interface part, storage, operating unit, power unit and a LAN or WAN network connection, which provides connection to a central server. A sample view of a bike
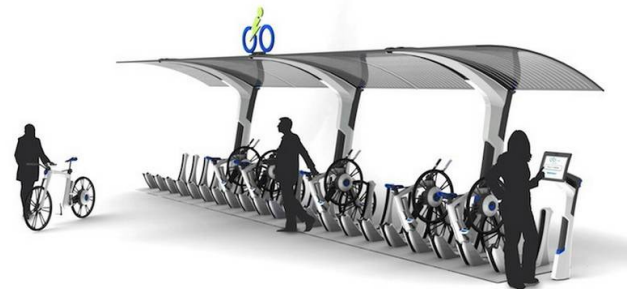
Fig.1 An example of a bike sharing station

sharing system is given in Fig. 1

The process of renting and using a bike has the following 4 major steps:

a) User Registration
b) Renting
c) Usage
d) Return

In most cases there is a small fee for renting the bike, which is

based on the duration and either pre-paid cards or credit cards can be used. Most of the systems require user registration in

terms of name, address and contact information. Once a bicycle is rented from the terminal, the lock on the bicycle stand is opened and the user can take the bicycle. Payment is usually per hour and is charged on the return. Bicycles can be returned at any station which has available bicycle stands.

B. University BSS

Different from the city BSS, university systems usually do
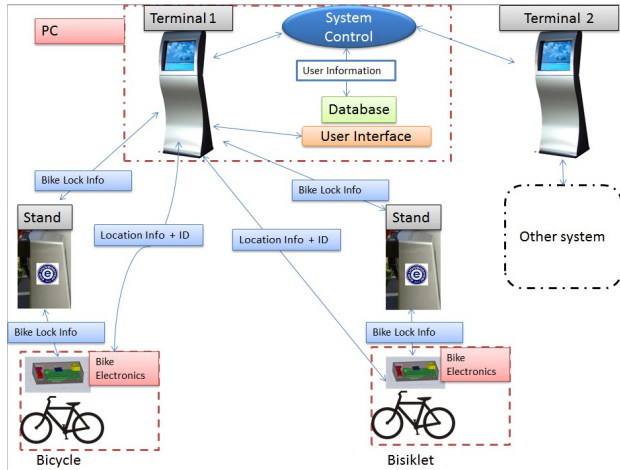
Fig.2 BikeEge Architecture

not require credit card payment but instead need initial user registration over the internet and a valid student ID card. Some of the systems provide also SMS notifications, location information as well as additional user related information. The architecture of the proposed system is given in Fig.2.

The proposed BikeEge system has the following main features:
a) It is free of charge for university students
b) Works directly with student ID cards
c) Integrated GPS service for bicycle location tracking
d) Location mapping of rented and free bicycles

III. MAIN COMPONENTS OF BIKEEGE

As can be seen in Fig. 2 the proposed system comprises the following modules: a Bike Module, a Station Module and System Control Module. The bike module consists of the electronics on the bike itself as well as the lock and the stand. The station module includes the terminal and the related user interfaces and databases. The system control module is connected to all terminals and contains the system management interface and all the user related databases.

List of the components used in the Bike Module is provided in Fig.3 below:
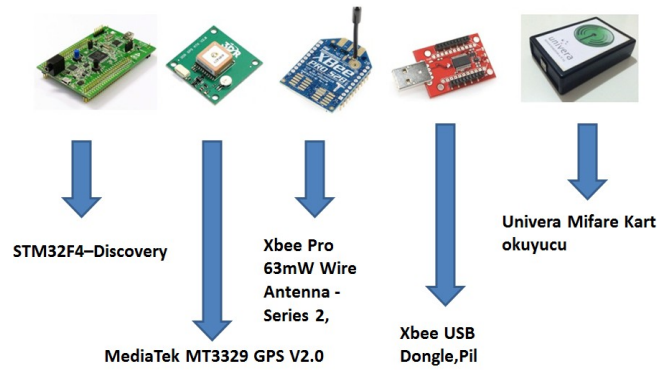
Fig.3 Hardware for the Bike Module

The Bike Module contains an ARM based STM32f4 processor, [13] [14], MT3329 GPS unit [9] and XBee 63mW wireless module which provides the connection between the docking stations and the terminal. [7], [9], [10], [13], [14]. Location information from the GPS and bike ID data is processed and transferred wirelessly to the terminal. The Station Module is responsible for connecting to the main System Control Module and database for verifying the user information. Through the designed user interface the operations of registration, password change, renting and return can be carried out at the terminal. System registration of a user can either be realized remotely (over the internet by using the student ID number) or directly at the terminal. Once the user is registered it is prompt for a password, which will be used later in all renting and return operations. The Station Module includes software and hardware for collecting the user information – in our case this is a MiFare card reader, compatible with the student card system used in Ege University [11] as well as processing unit and storage. All the data is stored and controlled by a database, which at this stage is a limited copy of the university one and is not linked to it in real time. The system control module is running on a server, which physically holds the whole user database, the bicycles, stands and terminals inventory and is implemented using SQL SERVER 2012. [12]
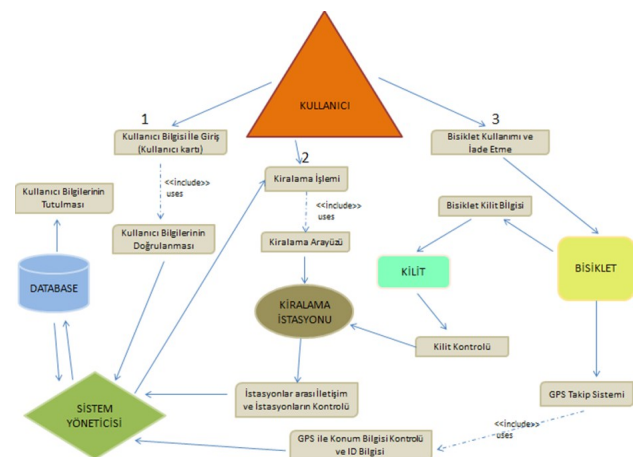
Fig.3 BikeEge Algorithm Flowchart

IV. BikeEge Interfaces and Operation

The designed system provides and easy to use interface. The interface is backed by the software system given in Fig.3. The communication is user-centered. After the initial welcoming screen (see Fig.4), the user action selection screen
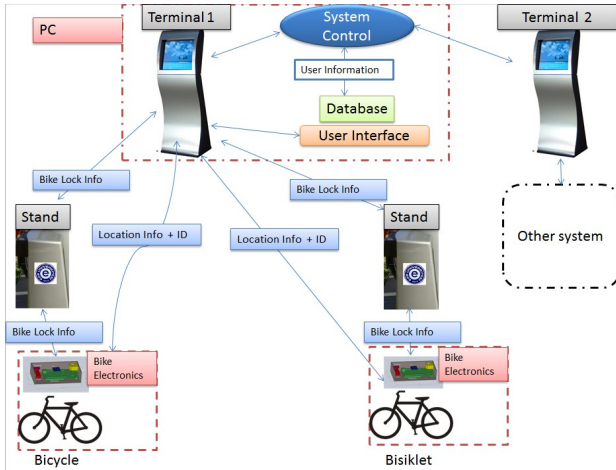
Fig.2 BikeEge Architecture

Fig.4 Welcome screen

Fig.5 User Action Selection screen

pops up that gives information about the available bicycles in terms of stations and number of bicycles available at each station. The user is requested to select an option to proceed. In order to rent a bike the user is prompted for his ID card and is directed to the renting window shown in Fig. 7. which requires user password.

Fig.6 User ID card request

At this stage the designed system accommodates both initial registration over the internet, based on the student email and registration at the terminal, when again the user is prompted for student ID number and email address. The information is checked against the existing EU student database.
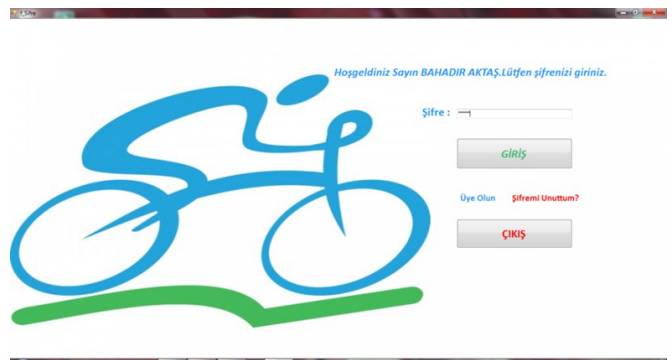
Fig.7 User registration screen

Once the registration process is completed, the user selects a bike number from the available pool. The numbers of the bikes are paired with the locks and the check is used for releasing the bicycle. An additional feature which is provided is a Google map, which shows the number of available bicycles at each station in the university campus. The terminals and docking stations are only within the limits of the university campus (which however is the largest in Izmir).

Since the service is provided free of charge to the students, there is no time limit to the renting period. However, the operator is notified if a bike is not returned within 24 hours. In that case the student should be contacted based on the information available in the database. In the future we plan to provide additional SMS service for return reminders, out of campus notifications and delayed returns notifications.

The return process is very similar to the renting process. It starts with the same Action Selection screen (given in Fig. 5) and upon selecting the return option, the student is prompt for his identity information and the ID number of the bicycle he wants to return. The data is used to match up the lock at the stand and it is also recorded in the management database together with a time stamp.

As part of the System Control Module a separate management interface developed for the administrator. It is shown in Fig. 8
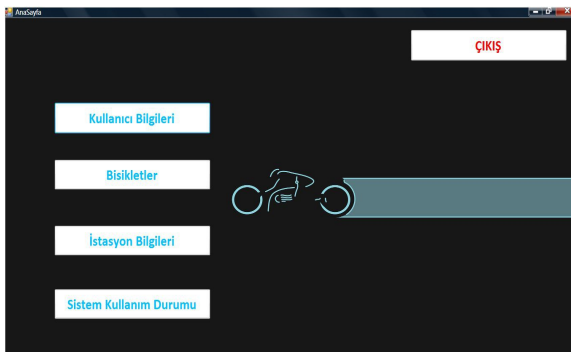
Fig.8 Administrator Screen

The administrator screen allows an authorized user to display and check the information in the different databases: user information, bike information, station information and system usage information. At any time the administrator can audit which users are renting which bikes and for how long they have been rented; he can view the number and ID of the bikes found at each station as well as the information about the occupied and free stands at each station. It also provides a notification if there is a damaged bicycle/lock, which can be used for timely intervention of the maintenance team.

V. The BikeEge Project

The work described in this paper is part of a much large project carried out currently in Ege University. Bike Ege is a large organizational project which requires the cooperation of very different units both in the university and outside the university. Besides the Department of Electrical and Electronics Engineering, which is responsible for the electronic part of the system and its supporting software there are several other departments participating. The Department of Machine Engineering is responsible for the docking systems (the stands, the locks and the terminals design); the Civil engineering Department is responsible for designing, establishing and maintain the bicycle tracks that will run through the campus and connect the two major parts (the Hospital Part and the Main Campus Part); the International Computer Institute will provide the final overall system software. As for the bikes themselves, for the first prototypes Accell has provided a limited number of all-purpose bikes. However, once the final bike and docking station design is determined, the manufacturing will be carried out by Accell. Finally, to have the system up and running, the active participation of the EU student Biking Society as well as the Faculty of Business will be required.

VI. Conclusion

In this paper we have presented details about an ongoing project for providing alternative "green" transportation to the

nearly 45 00 students in Ege University – Bike Ege. The scope of this paper is the design of the electronic sub-system that will be located on the bikes and will allow renting, tracking and return of the bikes to their docking stations. The operational principle, the hardware, software and the designed user interfaces are described in detail. The paper is concluded with future work goals and a brief overview of the aims of the project as a whole

References

[1] Alta Planning + Design: Bike Sharing/Public Bikes: An Overview of Programs, Vendors and Technologies. Portland, 2009
[2] Schneider, R. Krizek, K. J., and E. W. Ston, TCRP Synthesis of Practice 62: Integration of Bicycles and Transit, Transportation Research Board of the National Academies, Washington, D.C., 2005
[3] DeMaio, P., Smart bikes: Public transportation for the 21st century, Transportation Quarterly 57(1): 9–11, 2003
[4] New York City Department of City Planning, Bike-Share Opportunities in New York City, NY, 2009
[5] ViaCycle@GT, Georgia Tech, http://www.viacycle.com/
[6] Bixi Bike, Toronto, https://toronto.bixi.com/ride-with-bixi/station)
[7] http://en.wikipedia.org/wiki/Zigbee (accessed 03.2015)
[8] http://www.jennic.com/elearning/zigbee/index.htm (accessed 05.2015)
[9] MediaTek, MT3329 DataSheet, Rev, A03, 2010/04/30, 2010
[10] http://en.wikipedia.org/wiki/Global_Positioning_System (accessed 03.2015)
[11] http://www.sonmicro.com/trx/downloads/Mifare/ds_SM130 pdf
[12] Database Fundamentals, First Edition, Neeraj Sharma, Liviu Perniu.
[13] http://stm32f4-discovery.com/stm32f4-links/ (accessed 03.2015)
[14] http://www.st.com/st-web-ui/static/active/en/resource/technical/document/reference_manual/DM00031020.pdf (accessed 03.2015)

Quality of Service Considerations for two DiffServ scenarios in IP networks

Vesna Radonjić Đogatović [1], Saša Stojanović [2], Aleksandra Kostić-Ljubisavljević [3] and Branka Mikavica [4]

Abstract – **Differentiated Services (DiffServ) architecture model provides the most extensive and appealing solution for Quality of Service (QoS) support in current Internet Protocol (IP) networks. The main considerations in DiffServ model are scalability and traffic classification in order to handle large number of data efficiently. Multi Protocol Label Switching (MPLS) is a network protocol technology that can be useful for improving scalability and routing flexibility in IP networks. In this paper we compare DiffServ scenario with the scenario that represents integration of DiffServ and MPLS Traffic Engineering (TE). The results show that the implementation of MPLS TE mechanism to support DiffServ architecture in the backbone network can improve network performances.**

Keywords – **QoS, DiffServ, MPLS TE, IP.**

I. INTRODUCTION

Quality of service in IP networks is used to denote different concepts - from customer perception and evaluation of a certain service to a set of technical parameters that are required for the implementation of specific telecommunication services with required quality [1], [2], [3]. QoS can be quantitatively expressed by different parameters such as bandwidth, packet loss, delay, jitter, echo, availability etc.

Since QoS plays a crucial role in ensuring proper support for many types of applications with different quality requirements in IP networks, special attention should be paid to mechanisms which provide QoS in these networks. The most popular architecture for providing QoS in IP networks is DiffServ. Main drawbacks of DiffServ are the lack of QoS guarantees for individual IP traffic flows and inability to provide end-to-end QoS guarantees. Therefore, service providers often use MPLS to improve QoS in order to meet specific Service Level Agreements (SLAs) on the significant performance measures [4], [5], [6]. In addition, MPLS TE can improve scalability, network efficiency and service guarantees.

In this paper we explain and compare two QoS scenarios: DiffServ itself and DiffServ supported by MPLS TE. The paper is organized as follows. After the introduction, in Section 2 a brief overview of DiffServ and MPLS TE are given. In Section 3 simulation results are discussed and concluding remarks are given in Section 4.

II. QOS ISSUES IN IP NETWORKS

In terms of IP technology, QoS is characterized by diverse parameters depending on the particular service and traffic carried over the IP network. QoS is mainly technical issue but it can cover the complete end-to-end system effects including user perception and evaluation of a service. It describes the ability of network to provide a service with an assured level and is related to service performances that can be measured and controlled at the users' access point [7]. QoS is defined by the bilateral contract or agreement i.e. SLA between two interconnected parties. SLA regulates guarantees involved in providing and utilization of a service as well as the responsibilities of all contracting parties.

Due to the rapidly increasing deployment of interactive and multimedia applications in IP networks, QoS becomes an integral part of various protocols, mechanisms and services in enabling computing and communication systems [8].

A. DiffServ concept

DiffServ operates at Layer 3 only and is not engaged with lower layers of the OSI model. It relies on the traffic classification in order to provide different QoS level on a per-hop basis. Traffic can be classified according to different criteria which include Type of Service (ToS) value in an IP header.

Before clustering of traffic into an aggregation, the packets belonging to the aggregation must be identified. With aim to protect the service guarantees of each aggregation, limitations in terms of the amount of traffic that any user can inject must be conducted. These aspects are addressed by DiffServ functions named classifying and policing. Routers within DiffServ architecture perform functions of measuring, shaping and dropping packets in a flow [9]. DiffServ architecture also defines the relationship among multiple administrative domains, which are specified in a SLA.

DiffServ concept implies the division of network into DiffServ domains (DS domains), each of which corresponds to an Internet service provider domain or a network that is centrally managed from a single point. DS domain nodes use the common pre-defined rules for the network resources utilization and methods of packet processing depending on the IP traffic class. A DS domain consists of DS ingress nodes, DS interior nodes (core nodes), and DS egress nodes. An

All authors are with the Faculty of Transport and Traffic Engineering at University of Belgrade, Vojvode Stepe 305, Belgrade, Serbia.

[1] Vesna Radonjić Đogatović, e-mail: v.radonjic@sf.bg.ac.rs
[2] Saša Stojanović, e-mail: stojanovicsasa991@yahoo.com
[3] Aleksandra Kostić-Ljubisavljević, e-mail: a.kostic@sf.bg.ac.rs
[4] Branka Mikavica, e-mail: b.mikavica@sf.bg.ac.rs

ingress or egress node might be denoted as a DS boundary node, connecting two DS domains together.

Typically, DS boundary nodes perform traffic conditioning. They perform traffic classification based on the input values of several fields from the IP packet header, so that the individual streams fit into a finite and limited set of aggregate flows - a class of traffic. Each IP packet is marked by certain bits in the header, which defines the method of processing packets at each DS domain node (PHB, Per-Hop Behavior). Existing octet TOS in the IPv4 packet header or the Traffic Class octet in the IPv6 packet header is replaced by the Differentiated Services (DS) field. Six bits of the DS field are used as code PHBs (Differentiated Services Code Point, DSCP), while the remaining two bits remain unused and they are reserved for a future use. In this fashion it is theoretically possible to differentiate up to 64 different classes of traffic. Core routers perform packet classification based only on the value of DSCP field. Incoming packets are classified into pre-defined aggregates, metered in order to determine compliance to traffic parameters (whether the packet fits the profile or not), marked appropriately by writing/re-writing the DSCP and shaped (buffered to achieve a target flow rate) or being droped in case of congestion.

PHBs are applied on each network element and their assignment is to determine processing priority, allocated bandwidth, delay-bound, jitter-bound, packet drop probability etc. This combination of packet marking and well-defined PHBs results in a scalable QoS solution for any given packet and any application.

One of the drawbacks of DiffServ architecture refers to the lack of precise requirements specification and target values for the QoS parameters and consequently QoS guarantees for individual IP traffic flows, which derives from the very concept of providing QoS for the aggregate traffic flows.

Another drawback is related to the extent of providing QoS, which is limited to the DS domain without providing end-to-end QoS guarantees. DiffServ architecture does not provide explicit mechanisms for communication between applications and network in terms of dynamic QoS negotiation. Applications are not able to dynamically adapt IP traffic profiles according to the network resources availability. In case of such a change of state in DiffServ network that it is no longer able to handle accumulated traffic load, neither the input routers nor the applications will be informed about it. Furthermore, in terms of QoS negotiations, interaction between transmitting and receiving parties is not possible.

B. Multi Protocol Label Switching with Traffic Engineering

In order to ensure explicit QoS support, MPLS can be integrated with the DiffServ model. MPLS is one of the solutions which enable providing QoS guarantees and traffic prioritization for a variety of applications [4], [5]. MPLS is not a substitute for DiffServ but can be applied to support differentiated services architecture, primarily due to differences in the position within the OSI model, as can be seen in Figure 1. Contrary to the DiffServ, MPLS specifies modes that Layer 3 traffic can be mapped to connection-oriented Layer 2. It can be interpreted as an integration of

Layer 2 and Layer 3 technologies. MPLS sits between these traditional layers, providing additional features for the transport of data across the network and often is summarized as a layer 2.5 networking protocol.
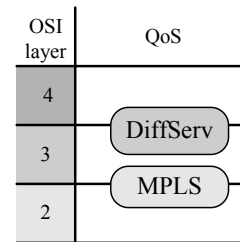
Figure 1. DiffServ and MPLS in OSI model

MPLS implies generating a special packet header - label, independent of the address that is inserted between the layers 2 and 3 headers within the OSI model. By making traditional Layer 2 features available to Layer 3, MPLS enables traffic engineering. The traffic engineered tunnels allow mapping traffic streams onto available network resources thus preventing the excessive use of subsets of network resources while other subsets are under-utilized.

MPLS TE employ label switching to improve traffic performance while network resources are efficiently utilized [10]. With MPLS, traffic engineering capabilities are integrated into Layer 3, which optimizes the routing of IP traffic, given the constraints imposed by backbone capacity and topology [11]. Labels are assigned and distributed between routers using the Label Distribution Protocol (LDP) or the Tag Distribution Protocol (TDP). After the Label Edge Router (LER) i.e. ingress router assign labels to the packet, it is forwarded across the network using label switching based solely on the label and not on the IP header information. A Label Switching Path (LSP) exists between all routers within the MPLS domain. Label Switching Router (LSR) i.e. transit router is responsible only for MPLS switching in the middle of an LSP. At the end of an LSP, egress router removes the label and the packet is again forwarded as an IP packet. MPLS TE automatically establishes and maintains LSPs across the network backbone, using Resource Reservation Protocol with Traffic Engineering (RSVP-TE) which improves scalability.

III. SIMULATION RESULTS ANALYSIS

In this section we present and analyze simulations that are performed for two QoS scenarios: DiffServ itself and DiffServ supported by MPLS TE. For the purpose of comparison these two scenarios in terms of providing the required QoS, OPNET Modeler was used, focusing on the backbone network.

The first scenario involves the application of DiffServ with a local setup of edge routers for QoS routing and prioritization of VoIP communications. The second scenario implies the application of MPLS TE with the previous setting of DiffServ approach as in the first scenario. It includes a global setting of static LSP in order to perform traffic engineering and to reduce the load from the main transmission route, defined by Open Shortest Path Protocol (OSPF) in the first scenario.

Session Initiation Protocol (SIP) signaling is not the goal of this research and SIP proxy server is not set up in the access network, as this analysis shows only the differences in the achieved QoS with focus on the backbone.

A. DiffServ scenario

The first scenario (DiffServ scenario) is designed in a manner that different PHB classes are differently transmitted through the network. Six sources of a video conferencing at the same time establish a video conferencing session and one source of VoIP communication establishes a VoIP session (in 100s of the simulation). After that each source adds session by session on the current one (in 420s of the simulation) and finally, three simultaneous sessions are established by each source (in 720s of the simulation). 2.048Mbps links are placed in the backbone network between two edge routers and 100BaseT Ethernet is set from the Ethernet switch to the edge router. 10BaseT Ethernet link is set from the source to the switch. This network is shown in Figure 2.

The total throughput after the establishment of all three sessions is approximately 6.91Mbps. Applying the OSPF routing the shortest selected path of 2.048Mbps (assuming the same rate of each link) will not be able to support the transmission of all sessions with minimal losses and therefore Class Based Weighted Fair Queuing (CBWFQ) is implemented. Dynamic queue management is performed by Weighted Random Early Detection (WRED). CBWFQ and WRED profiles for each PHB traffic class are consolidated through the traffic policy profile. Also, the extended access control list is created in the edge routers in order to identify future video and audio streams for traffic classification.

DSCP values in the header and the assigned address are used for packet identification. Each source is connected to the corresponding destination, thus the communication between two workstations can be accurately performed.
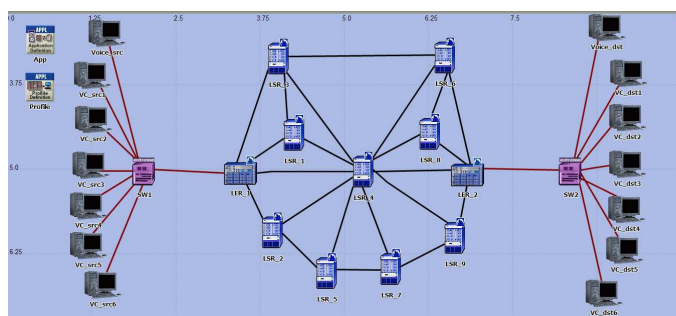
Figure 2. DiffServ scenario

B. DiffServ with MPLS TE scenario

MPLS enables the distribution of sessions to statically or dynamically determined paths. In the simulation static LSPs are used. For the purpose of classification and clustering of packets, Seven Forwarding Equivalence Classes (FECs) and seven traffic trunk profiles are defined. Based on classes and trunk profiles, flows are mapped to the defined static routes (by configuring the edge router). The layout of the network with defined paths for the second scenario is shown in Figure 3.
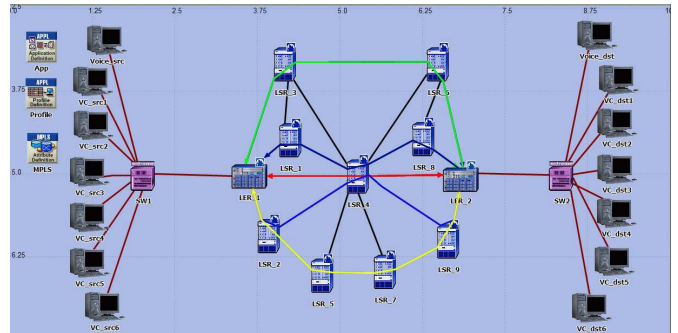
Figure 3. DiffServ / MPLS TE scenario

C. Simulation results comparisons

We performed simulations for different QoS parameters. In this paper the results for average voice end-to-end delay, average video conferencing end-to-end delay and voice traffic received are presented and discussed.

In Figure 4 it can be noticed that end-to-end delay for voice is below 68ms when DiffServ scenario is applied and below 66ms in case of DiffServ with MPLS TE (delay should be kept below 150ms). Therefore, for voice communication end-to-end delay shows acceptable level in both scenarios. In terms of delay second scenario is slightly better than the first one.
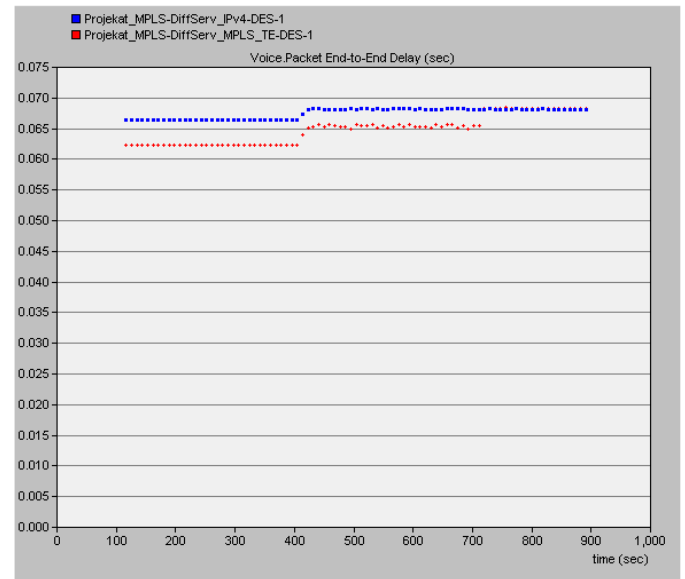
Figure 4. Average voice end- to-end delay

In Figure 5 the average end-to-end delay for video conferencing streams are shown. Applying CBWFQ algorithm has led to considerable delay due to prevention of video conference flows to take the most of the output link router LER_1 capacity. Average end-to-end delay is on the satisfactory level in the first scenario until a second session is established while excellent results in the second scenario are deteriorated with the establishment of the third session although these results are also acceptable (acceptable is up to 400ms).
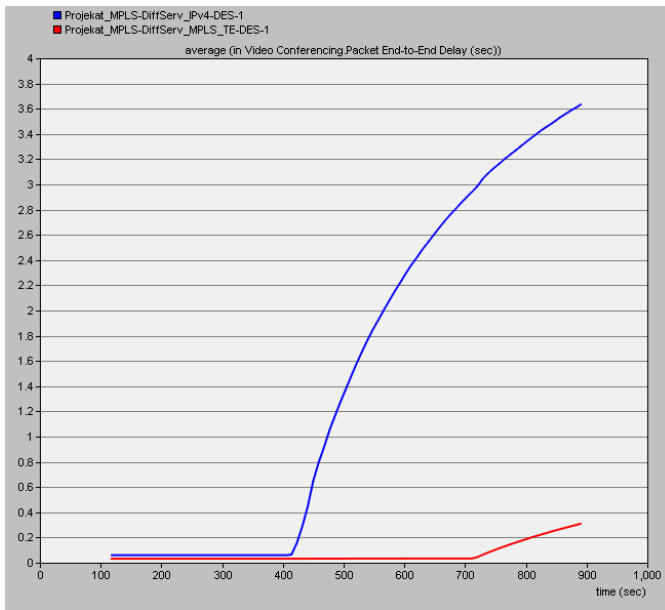
Figure 5. Average video conferencing end-to-end delay

Of all simulation results voice traffic received is the best indicator of differences between these two scenarios. In Figure 6 it is shown that the establishment of the second flow leads to a significant packets loss in the first scenario, primarily due to the overload of link LER_1 - LSR_4 - LER_2 established by OSPF protocol. On the other hand, there are no significant packet loss of voice sessions in the second scenario.
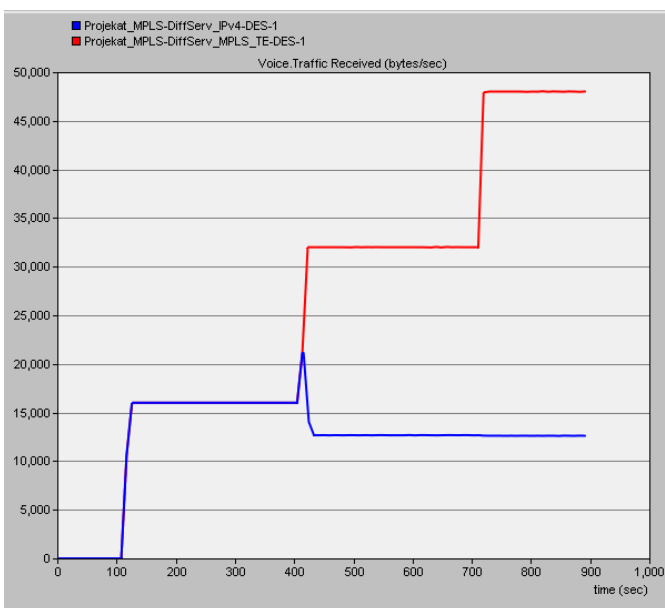
Fig 6. Voice traffic received

IV. CONCLUSION

DiffServ and MPLS are two separate standards dealing with QoS issues in IP networks. DiffServ itself is not capable to control the traffic which has been taken to end-to-end path with a number of congested links. On the contrary,

MPLS TE has ability to control the traffic and to set up end-to-end routing path before data has been forwarded. The evolution of QoS solutions in IP networks has led to the integration of DiffServ and MPLS TE which can be suitable even for the applications with high requirements in terms of QoS, such as real-time applications.

In this paper we have focused on two scenarios: DiffServ without implementing MPLS TE and DiffServ supported by MPLS TE. Some of the simulation results for voice communication and video conferencing are presented and analyzed. The results show improvement of network performances for the scenario which includes implementation of MPLS TE mechanism to support DiffServ architecture in the backbone network.

ACKNOWLEDGEMENT

REFERENCES

[1] J. Gozdecki, R. Stankiewicz, A. Jajszczyk, "Quality of Service Terminology in IP Networks," IEEE Communication Magazine, vol.41, no.3, pp. 153-59, 2003.

[2] ITU-T Recommendation E.800, "Definitions of terms related to quality of services", Geneva, 2008.

[3] ITU-T Recommendation P.10, "Vocabulary for performance and quality of service", Geneva, 2006.

[4] M. Romdzi, A. Rahimi, H. Hashim, R. Ab Rahman, "Implementation of Quality of Service (QoS) in Multi Protocol Label Switching (MPLS) Networks", Proceedings of the 5th International Colloquium on Signal Processing & Its Applications (CSPA), Kuala Lumpur, Malaysia, pp. 98-103, 2009.

[5] A. Jamali, N. Naja, D. EI Ouadghiri, R. Benaini, "Improving Quality of Service (QoS) in Multi-Protocol Label Switching Module", Mediterranean Microwave Symposium (MMS), Tangiers, Morocco, pp. 1-4, 2009.

[6] Md. Tariq Aziz, Mohammad Saiful Islam, Md. Nazmul Islam khan, Prof. Adrian Popescu "Effect of packet delay variation on video/voice over DiffServ-MPLS in IPv4/IPv6 networks", International Journal of Distributed and Parallel Systems IJDPS, vol. 3, no.1, pp. 27-47, 2012.

[7] V. Radonjic, A. Kostic-Ljubisavljevic, M. Stojanovic, V. Simic, B. Dimitrijevic, "Different Aspects of Quality in NGN", Proceedings of the 6th International Quality Conference, Kragujevac, Serbia, pp. 817-822, 2012.

[8] G. Rovithakis, A. G. Malamos, T. Varvarigou, M. A. Christodoulou, "Quality of Service Mechanisms in Multimedia Telecommunication Services", Reliability, Survivability and Quality of Large Scale Telecommunication Systems, edited by P. Stavroulakis, John Wiley & Sons Ltd, 2003.

[9] W. Weiss, "QoS with Differentiated Services", Bell Labs Technical Journal, vol.3, no.4, pp. 48-62, 1998.

[10] S. Alvarez, QoS for IP/MPLS Networks, Cisco Press, 2006.

[11] S. Naveed, S. V. Kumar, "MPLS Traffic Engineering – Fast Reroute", International Journal of Science and Research (IJSR), vol.3, no.5, pp. 1796-1801, 2014.

Development of Hardware Control on Dot Matrix Liquid Crystal Display

Goran Goranov [1] Pavel Hubenov [2]

Abstract – **This paper proposes a new method of display control based on CPLD, using logical gate functions. It allows easy creating an external block with desired characters of the alphanumerics, upper or lower case, symbols. Assignment of symbols is set in binary code, which has to be written in memory of the display's controller.**

Keywords – **Logic gate functions, Dot matrix, Control, LCD, CPLD, Xilinx.**

I. INTRODUCTION

LCD is a display device used to display information. Its applications starts from simple clocks and reaches to machine's control units or railway's monitor boards and so on. A special controller converts instructions into signals, which switch on or off lights in the matrix.

There are many LCD dot matrix modules containing the corresponding LCD driver and controller, e.g. "KS0066U" (40SEG DRIVER & CONTROLLER FOR DOT MATRIX LCD) designed by Samsung or "HD44780U" (Dot Matrix Liquid Crystal Display Controller/Driver) designed by Hitachi. These both are compatible. [1, 2]

The display consists of a dot matrix of lights, which are organized by lines and columns. The controller defines which single dot to turn on or off. Every single dot has coordinates C_n (Common 1-16) and S_n (Segment 1-40). The combinations of dots, which formed figures, letters, and symbols, are saved in library called character generator.

Every single character can be summoned by a binary code according to manufacturer's arranged table. The internal memory contains three character generators:

1. CGROM (Character Generator ROM): 10,080 bits (204 characters 5x8 dots) & (32 characters 5x11 dots);

2. CGRAM (Character Generator RAM): 64x8 bits (8 characters 5x8 dots);

3. DDRAM (Display Data RAM): 80x8 bits (80 characters max.);

The Special character pattern is directly programmable by the Character Generator RAM and owing to a mask option, customer can program character pattern in his own.

The principle of managing a display like this is the same for different manufacturers. The strategy is based on few

[1]Goran Goranov is Assistant Professor – Doctor at Technical University of Gabrovo, 5300 Gabrovo, ul. Hadji Dimitar 4, Bulgaria, E-mail: g_goranov@bitex.bg.

[2]Pavel Hubenov is PhD student at Technical University of Gabrovo, 5300 Gabrovo, ul. Hadji Dimitar 4, Bulgaria, E-mail: pavel_hubenov@yahoo.com.

commands of initialization, followed by instructions to choose mode and data for visualization, Table 1, a. All codes are accompanies by a strobing frequency which is the clock frequency for data bus transfer. [1]

TABLE I
INSTRUCTIONS FOR INITIALIZATION

a

Power on

Wait for more than 30ms after VDD raises to 4,5V

Function Set

RS	R/W	DB7	DB6	DB5	DB4	DB3	DB2	DB1	DB0
0	0	0	0	1	1	*N*	*F*	X	X

Wait for more than 39µs.

Display ON/OFF Control

RS	R/W	DB7	DB6	DB5	DB4	DB3	DB2	DB1	DB0
0	0	0	0	0	0	1	*D*	*C*	*B*

Wait for more than 39µs.

Display Clear

RS	R/W	DB7	DB6	DB5	DB4	DB3	DB2	DB1	DB0
0	0	0	0	0	0	0	0	0	1

Wait for more than 1,53ms.

Entry Mode Set

RS	R/W	DB7	DB6	DB5	DB4	DB3	DB2	DB1	DB0
0	0	0	0	0	0	0	1	*I/D*	*S/H*

Initialization End

b

LEGEND OF CHARACTERS

N	0	1-line mode	*F*	0	Display off
	1	2-line mode		1	Display on
D	0	Display off	*C*	0	Cursor off
	1	Display on		1	Cursor on
B	0	Blink off	*I/D*	0	Decrement mode
	1	Blink on		1	Increment mode
S/H	0	Entire shift off			
	1	Entire shift on			

When the power is turned on, KS0066U is initialized automatically by power on reset circuit. During the initialization, the following instructions are executed, and BF (Busy Flag) is kept "High" (busy state) to the end of initialization. The description of parameters for instructions of initialization are shown in Table 1, b.

II. BLOCK AND PRINCIPLE SCHEMES

A. Block scheme

In order to make usage of the display faster and easier could be created a scheme with inputs labeled as alphanumeric and outputs corresponding to pins of LCD module. For example, the block named "Converter" on figure 1 is logic schematic, function "OR" gate dedicated.

The input frequency "CLK1" is the clock frequency of the "Shift Register", with serial input and parallel outputs. It defines the duration of time when an output is active. The active one enables internal logic combination of "Converter", based on "OR" gate array. [6, 7]

Outputs "D0" to "D2" are responsible for the initialization of the display, step 2, 3, 4 in Table 2. They are followed by sequence of symbols which forming the word on the display. The usage of lower frequency for "CLK1" to the input of the register allows displaying the word as slow as is necessary. The higher value of it has to be pursuant by the manufacturer's timings, Table 1, a.
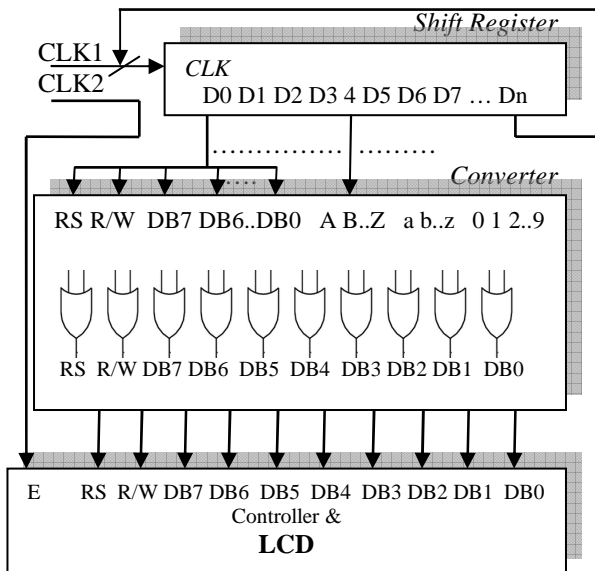
Fig. 1. Block schematic to manage LCD

Block "Converter" owns outputs, which are managing the LCD module. Each of them is the output of a single "OR" gate with a number of inputs extended as it is necessary. The same rule applies for inputs of the block. Whether there is a symbol used more than once, his Converter's input has to be extended to the number of usage. Character "A" in the example is used twice. In this case an auxiliary logic scheme, based on "OR" gate, between "Shift Register" and "Converter" is needful.

The second frequency "CLK2" is the strobing frequency. It defines the speed of data transfer into the controller.

When the desired information is displayed on the LCD, both frequencies, "CLK1" and "CLK2", have to be stopped. Otherwise, the word will disappear. To achieve that, last output of the decoder has to be involved in a logic scheme which will block them.

B. Principle schemes

The frequency "CLK1" proceeds to the input of block "Shift Register". It consist of a 4-bits counter and 4-bits decoder, figure 2. Every pulse turns the output state of the counter by one, and the binary code changes the output of the decoder.
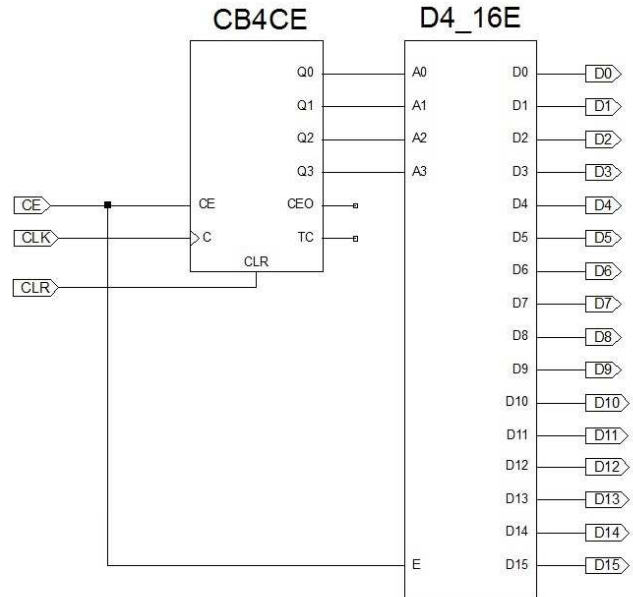
Fig. 2. Shift register scheme

First of decoder's outputs is "D0", which starts the
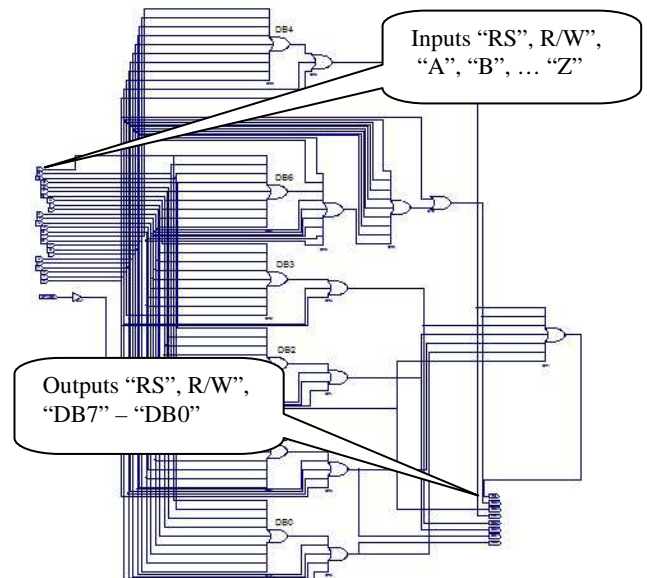
Fig. 3. Converter scheme "OR" gates array
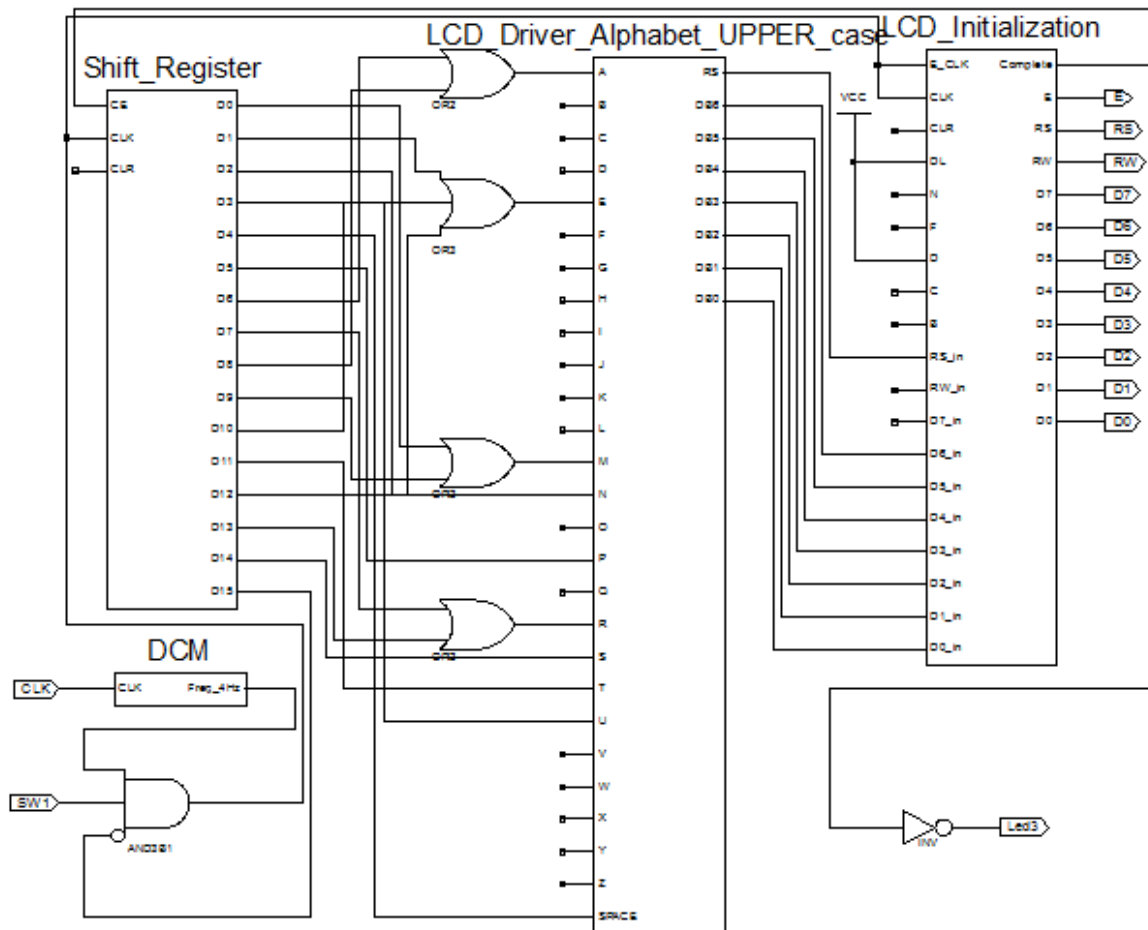
Fig. 4. Principle scheme of external programmable driver for LCD 2x16 dot matrix

sequence of initialization regarding Table 1, a. After the process of initialization, "D0-D2", commands for desired word are run to the converter block. Every command is coherent character. This block contains logic gates, function "OR" dedicated, which spread a single signal from inputs ("A", "B", … "Z") to a combination of outputs "DB0"-"DB7" corresponding to the binary code, figure 3. The addition of "RS" and "R/W" busses in this block, is useful for the final scheme of the project. As a matter of fact, every data code for character has to be accompanied by a strobing pulse, and adjacent pulse to "RS".

III.   EXPERIMENTS

In order to make usage of the display faster and easier could be created a scheme with inputs labelled as alphanumeric and outputs corresponding to pins of LCD module. For example, the scheme of "Converter" is assembled to "LCD_Driver_Alphabet_UPPER_case", on figure 4.

The block "Shift_Register" represents the scheme from figure 2. To minimize the symbols among bigger schematic modules on figure 4, the approach with the counter and decoder is imported for input scheme of the last block called "LCD_Initialization". Their function is to set initialization

instructions, which states could be ordered to the "DL", "N", "F", "D", "C", "B" inputs. When the process is completed, active signal appears on output "Complete" and enables "Shift_Register" to write characters. In this case, the desired inscription is "MENU PARAMETERS", so "D0" is responsible for symbol "M", and his active signal is spread out for his relevant binary code. The output "Led3" is control signal for monitoring. [1, 2]

Block "DCM" is a digital clock manager which is a generator assembled by counters provides frequency 4Hz. Logic gate "AND3b1" stops the frequency when the whole line is displayed. The four "OR" gates extend the inputs for doubled symbols "A", "M", "R", and the tripled "E" one. Input "SPACE" is added to implant the empty space.

The sequence of data instructions for the example are systematically arranged, shown in Table 2. The description of parameters for instructions of initialization are shown in Table 1, b.

TABLE II
Example of writing "MENU PARAMETERS".

Step №	RS	R/W	DB7	DB6	DB5	DB4	DB3	DB2	DB1	DB0	Display	Description
1												Power supply
2	0	0	0	0	1	1	0	0	*	*		Function set
3	0	0	0	0	0	0	1	1	1	0		Display on/off
4	0	0	0	0	0	0	0	0	0	1		Display Clear
5	1	0	0	1	0	0	1	1	0	1	M_	Write data to CGRAM/DDRAM
6	1	0	0	1	1	0	0	1	0	1	ME_	Write data to CGRAM/DDRAM
7					. . .				. . .			
8	1	0	0	1	0	1	0	0	1	1	MENU PARAMETERS	Write data to CGRAM/DDRAM

IV. RESULTS

The approach allows writing easy a simple word, or a line, figure 5. It could be developed to more complicated system forming a menu. Every line, or lines, shown on the display, needs a register, which will activate the sequence of the letters. Well known is that they have control enable pin, so in a bigger designed system, the inscriptions on the LCD could be summoned by one bit.
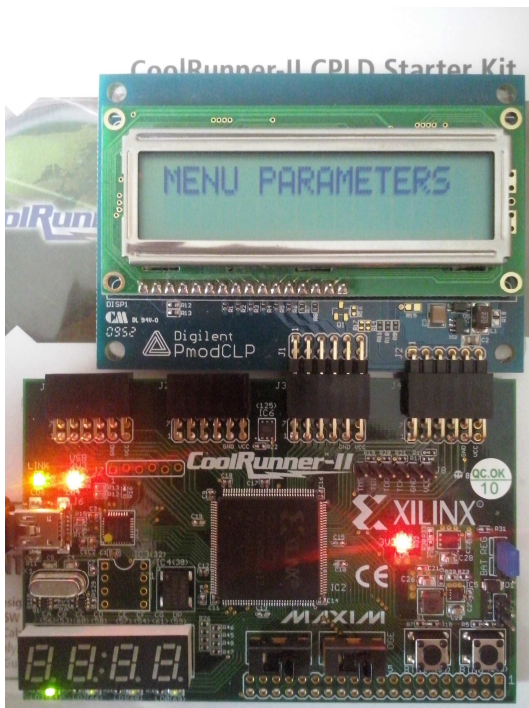
Fig. 5. Inscription is already shown

V. CONCLUSION

In this paper the mention approach for using dot matrix LCD 2x16 is based on a hardware solution. Owing to the purpose to show a simple word on the display, the project could be realized by discrete integrated circuits. Advantage is the low cost. Actually it is a difficult method to create a system menu, with adjacent submenus, with embedded different parameters and values, with options to adjust them. Disadvantages could be the large number of ICs, and the space they will require.

Another solution is the usage of a programmable logic device. The choice concerning to the program language, capacity of the product, functionality, and additional peripheral devices is variety and rich enough in nowadays.

One balanced solution according to the list of demanding is "CoolRunner-II" board. Produced by Digilent, based on Xilinx's platform XC2C256TQ144. It is a CPLD (complex programmable logic device) generation structure, i.e. a logic matrix consisting of "OR" and "AND" logic gates. There are more than 60 peripheral devices formed to extend board's functionalities, called "Pmods". A "PmodCLP" is a board with LCD controller "KS0066U" and dot matrix LCD 2x16, managed by parallel interface. Thus designed system - board with LCD, is a reliable configuration for developing a more complicated system. [3, 4]

For example, designed menu in this way could be useful for a control system for driving electrical motor. [5]

REFERENCES
[1]   Hitachi, "HD44780U", Dot Matrix Liquid Crystal Display Controller/Driver, 1999.
[2]   Samsung, "KS0066U", 16Com/40Seg Driver & Controller For Dot Matrix LCD.
[3]   Xilinx, "Xilinx CPLD Applications Handbook", 2006.
[4]   Todorova, V.D., Todorov, P.J., "Microelectronics. Integrated circuits", 2013.
[5]   Goranov, G.D., Hubenov, P.P., "Control system designed of 3 phase brushless motor based on programmable logic", "Mechanical and electrical engineering", 2014/02.
[6]   Kirchev, A., "Digital schematics", 2009.
[7]   Tokheim, R., "Digital electronics", 1999.

Using of mobile platforms for sensor nodes in Biomedical Wireless Sensor Networks

Georgi Stoyanov [1], Borislav Naidenov [2], Stela Kostadinova [3]

Abstract – **The rapid pace of technological advances in recent years has enabled a significant evolution and deployment of Biomedical Wireless Sensor Networks (BWSN) which have the important impact nowadays. Related technologies have variety of applications, and they are key enabling technologies of IoT (Internet of Things) in the field of e-health. IoT solutions, based on different mobile platforms, offering cost efficiency, flexibility and simplicity of development, can be used in BWSNs as sensor nodes or gateways. The current paper aims to summarize the challenges related with the collection, manipulation and exploitation of the data generated by these networks, using mobile platforms as sensor nodes. The biomedical data security, the type of communication and the sensor node's software, also will be considered in terms of increasing the efficiency of data transmission in BWSNs.**

Keywords – *Biomedical Wireless Sensor Networks (BWSN), mobile platforms, sensor node*.

I. INTRODUCTION

With the swift progress in wireless communication and semiconductor technology, the sensor networks are developing rapidly, covering areas of different applications, including in medicine. Rapidly growing number of various wearable health monitoring devices in recent years is used for a wide range of medical and healthcare solutions. These devices provide continuous and real time health monitoring to patients, users, doctors or medical centers and they are responsible for collecting, processing and transmitting data from a wide range of body-sensors such as: blood pressure, heart rate, respiration rate, electrocardiography (ECG) for the heart health, Electromyography (EMG) for the muscles control, electroencephalography (EEG) for the brain signals, oxygenation signals, and temperature. Some of the latest research in this area are focused on the wireless, mobile, and easy to wear solutions to make this monitoring more user friendly since the wires may limit the user activities and level of comfort and also influence significantly the measured results. Nevertheless, the technology advances in miniaturization and integration of microcontrollers, physical sensors, embedded radio interfaces, wireless networking and microfabrication have enabled a new generation of wireless sensor networks suitable for many biomedical applications [1].

Biomedical Wireless Sensor Networks (BWSN) are a subcategory of wireless sensor networks (WSNs) which operate in a pervasive manner for on body applications, especially for health and medical aims. This network is designed to work independently and to make the link between different medical sensors attached to the body or implanted in it. These sensors have wireless capability with a central sensor, giving it any information on the state of the body. Once receiving the information, they transmit it to a central controller, which processes it and forwards it to an external device, which displays the collected information. Biosensors can be replaced by a variety of devices, called mobile platforms, which can also be used as a central node or gateway of BWSN. Mobile platforms represent development boards, such as raspberry pi, arduino and intel galileo, which can be used for various applications in biomedicine, depending on the software and the necessary configurations [2]. The architecture of BWSN is illustrated in Fig.1.
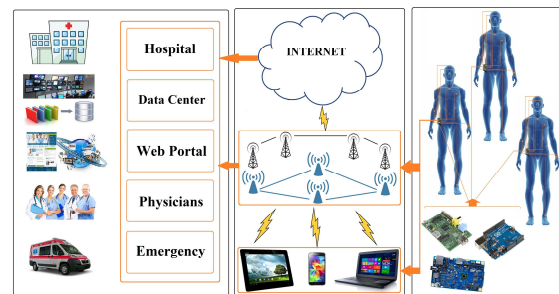
Fig. 1. Architecture of Biomedical Wireless Sensor Network

Actually the biomedical sensor nodes integrated into BWSN, can provide a novel information technology that will be able to support early detection of abnormal conditions and prevention of serious consequences. Today BWSN offer many advantages like the mobility of the users, flexibility and integration. Thus, today's body sensor networks are very useful to offer assistance to the users, patients, and also persons with disabilities. That's why a large number of novel biomedical sensors are proposed by different companies to cover the requirements of a wireless sensors node especially in terms of high accuracy, small scale factor, high integration, multisensory versatility and low-power consumption [3].

The aim of this paper is to consider the different types of biomedical sensors used in BWSN and categorized them according to various addictions. The characteristics of some mobile platforms are discussed and there is comparison between them and various wireless sensor nodes in order to examine their ussage as biomedical sensor nodes in BWSN.

[1]Georgi Stoyanov is with the Faculty of Electronics at Technical University of Varna, 1 Studentska Str, Varna 9010, Bulgaria, e-mail: georgi.stoyanovv@gmail.com.
[2]B. Naidenov is with the Faculty of Electronics at Technical University of Varna, 1 Studentska Str, Varna 9010, Bulgaria, e-mail: borna@abv.bg.
[3]S. Kostadinova is with the Faculty of Electronics at Technical University of Varna, 1 Studentska Str, Varna 9010, Bulgaria, e-mail: stela.kostadinova@gmail.com.

II. BIOMEDICAL SENSOR NODES

The key components of BWSN are sensors, because they connect the physical world with electronic systems. They are mainly used to collect the information about physiology and the surrounding environment. Some of the sensor nodes have a sensor as their main part, and they are responsible for processing information by format conversion, data storage, logical computing, and transmitting. One sensor node generally comprises a sensor module, processor module, wireless communication module, and power supply module [4]. The sensor module is responsible for collecting the status of measurands and converting data from physical quantities to electrical signals.

The building blocks of BWSN are sensor nodes and consist four basic elements shown in Fig. 2: the sensor unit, processing unit, communication and power units.
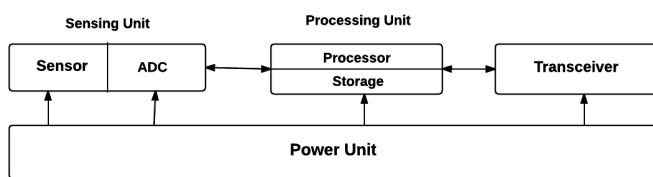
Fig. 2. Sensor Node architecture

Sensor nodes are the small, low power single board computers with a radio for wireless communication. Number and types of sensors depends on the applications. Sensor nodes collect and transfer data using four stages: collecting the data, processing the data, packaging the data and communicating the data [5].

Classification of Biomedical Sensors

In practical applications, the type of sensors and the number of sensors a BWSN system employs depend largely on the particular application scenario and system infrastructure [6]. Many different types of sensors can be used in BWSN, to complete the detection of physiology signals, human behavior, and the surrounding environment. Because of various application-specific requirements, the biosensors in BWSN can be of many types:

- *Depending on the type of measured signals:*

In [7] the authors divide sensors in BWSNs into two categories, according to the types of measured signals. The first category sensors collects signals continuously, placing more emphasis on real-time signal acquisition. The second category collect discrete time-varying physiology signals.

- *Depending on the types of data transmission media:*

One of the most commonly used biomedical sensors in BWSNs can be divided, depending on the types of data transmission media, into radio frequency identification devices (RFID), wireless sensors, which employ wireless communication technologies such as Bluetooth or Zigbee and Ultra Wideband (UWB). Removing wires completely will be an inevitable trend for BWSNs [8].

- *Depending on the deployment positions of sensor nodes:*

Biomedical sensors in BWSNs can be divided into three categories according to the deployment positions of sensor nodes [8]: sensors, which can be implanted into the body, such as a camera pill, called implantable; wearable sensors, such as accelerometers, temperature sensors and pressure sensors; sensors which can be used to recognize behaviors and collect information about the surroundings is placed surrounding peoplesuch as visual sensors.

- *Depending on the automatic adjustment ability:*

Biomedical sensors can be divided into two types, according to their automatic adjustment ability: self-adapting and non-self-adapting sensors. First type can automatically adjust processing method, order, and parameters, boundary conditions or constraints according to data characteristics, make themselves adapt to the statistical distribution and structural characteristics of the measured data, in order to get the best treatment effect. Second type are simple to design and need no consideration of self-adjusting function, and they are widely used in BWSNs at present.

- *Depending on the role in the network:*

Sensors in BWSNs can also be classified into three types based on their role in the network: Coordinator which is a gateway to the outside world or another BWSN; End Sensors which are only capable of performing their embedded application; Routers are intermediate sensors which have a parent sensor and a few child sensors through which they relay messages.

III. MOBILE PLATFORMS

The use of IoT standards in BWSNs allows the system designer to re-use application layer programs; presenting sensing data and actuation commands to software developers in a generic manner. The challenges however remains for BWSN designers operating at the gateway level and sensor network level.

A sensor node is a smart sensor that is capable of gathering sensory information, performing some processing, and communicating with other connected nodes on the network. Sensor platforms such as Arduino allow users to connect sensor and communications modules to a base platform. Gateways perform protocol translation between different networks. A gateway can operate at any network layer, and, unlike a router or a switch, a gateway can communicate using more than one protocol. PCs, servers, and M2M devices can function as gateways, although they are most commonly found in routers. In a sensor network, a gateway is responsible for interfacing the data from the sensor nodes to another network that uses a different protocol, and delivering commands back from that network to the nodes.

Mobile platforms, also called Single board computers (SBCs), that use a Linux based operating system provides the benefits listed below: Re-usable operating system drivers for networking, file storage and other peripherals; An abundance of software packages are readily available and actively maintained to accelerate functional software component development; Gateway devices can be monitored and

administered with existing IT infrastructure monitoring systems; IT network administrators already have the necessary skills to implement and maintain Linux-based devices; Gateway hardware can be exchanged for a more powerful solution in a domain specific application; improving the flexibility of the solution.

Using the system design and key benefits listed above the Raspberry Pi, BeagleBone Black and Arduino Uno (Fig. 6) was selected for comparing with sensor nodes used in BWSN.

Raspberry Pi

Arduino Uno

BeagleBone Black
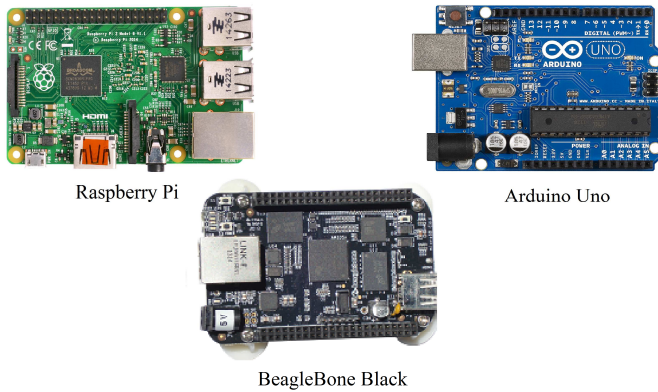
Fig. 6. Mobile Platforms

All these devices are popular mobile platforms for experimental use and run a Linux kernel based operating system. The logical separation of software and hardware components is a key requirement for scalability, cost and performance.

On the following table is shown a comparison between the mentioned mobile platforms and other similar boards.

TABLE I
THE COMPARISON BETWEEN MOBILE PLATFORMS

Features	Raspberry Pi - B	BeagleBone	Arduino	Panda Board	BeagleBone Black	Intel Galileo	MK802IIIS
SoC	Broadcom BCM2835	TI Sitara AM335x	Atmel ATmega328P	TI OMAP4460	TI Sitara AM335x	Intel Quark SoC X1000	Allwinner A10
Singlecore/ Multicore frequency	Single-core/ 700 MHz	Single-core/ 720 MHz	Single-core/ 16 MHz	Dual-core/ 1,2 GHz	Single-core/ 1 GHz	Single-core/ 400 MHz	Single-core/ 1 GHz
Networking	Ethernet	Ethernet	No	Ethernet/ Wi-Fi	Ethernet	Ethernet	Wi-Fi
RAM	512 MB	256 MB	2 KB	1 GB	512 MB	256 MB	1 GB
Price	0 - 49 $	50 - 79 $	0 - 49 $	400 $	50 - 79 $	50 - 79 $	50 - 79 $

Raspberry Pi is a small, powerful, cheap, hackable and education-oriented computer board. The platform contains a processor and graphics chip, program memory (RAM) and various interfaces and connectors for external devices [9]. Like any other computer, the Raspberry Pi also uses an operating system and the "stock" OS is a flavor of Linux called Raspbian. Linux, as a free and open source program, is a great match for Raspberry Pi [5].

BeagleBone Black is a single-board computer based on low-power Texas Instruments processors, using the ARM Cortex-A8 core. It is a small credit card-sized computer which can run an operating system such as Linux/Android 4.0. The main difference between it and Arduino is that it can run a small operating system, thereby practically converting it into a minicomputer that can run programs on these operating systems. BeagleBone is designed to function at a much higher l10el and it has far more processing capacity than Arduino.

Arduino is an open-source physical computing platform based on a simple microcontroller board, which can affect its surroundings by controlling lights, motors, and other actuators. The microcontroller on the hardware board can be programmed using the Arduino programming language and the Arduino Integrated Development Environment (IDE). Arduino supports two working modes, stand-alone or connected to a computer via USB cable [11].

IV. MOBILE PLATFORMS'S PERFORMANCES AND CONSTRAINTS

In next steps mobile platforms performances will be compared with following wireless sensor nodes [12]:

- MicaZ – is based on the Atmel ATmega128L which isa low-power microcontroller and runs MoteWorks from its internal flash memory.

- TelosB - bundles all the essentials for lab studies into a single platform including: USB programming capability, an IEEE 802.15.4 radio, a low-power MCU with extended memory and an optional sensor suite

- Iris – is used for enabling low power WSN.

- Cricket - is a location aware version of the popular MICA2 low-power Processor/Radio module.

- Lotus - is based on the NXP LPC1758, 32-bit ARM Cortex-M3 based microcontroller.

A. Size & Cost

The physical size and cost of each individual sensor node has a significant and direct impact on the ease and cost of deployment. Physical size impacts the ease of network deployment because smaller nodes can be placed in more locations and used in more scenarios. One of the main goals of every network is to collect data from as many locations as possible without exceeding fixed budget. A reduction in per-node cost will result in the ability to purchase more nodes, to deploy a collection network with higher density, and to collect more data. The comparison of size, weight and cost is given in Table II (the smaller values are better).

TABLE II
THE COMPARISON OF SIZE, WEIGHT AND COST

Name	Size (mm)	Weight (g)	Cost per node US$
Raspberry Pi	85.6*53.98*17	45	25-35
BeagleBone Black	86.3*53.3	39.68	45
Arduino (Uno)	75*53*15	30	30
MicaZ	58*32*7	18	99
TelosB	65*31*6	23	99
Iris	58*32*7	18	115
Cricket	58*32*7	18	225
Lotus	76*34*7	18	300

B. Power & Memory

The main goal of the biomedical sensor nodes is low power consumption in order to meet the multiyear application requirements. Also, algorithms and protocols must be

developed to reduce radio activity whenever possible. This can be achieved by using localized computation to reduce the streams of data being generated by sensors and through application specific protocols.

The CPU is the main component of the mobile platforms, responsible for carrying out the instructions of a computer program via mathematical and logical operations. Regarding the storage, the device should have sufficient memory in order to store the collected data.

Comparative analysis of platforms and sensor nodes's CPU, memory and power is given in Table III.

TABLE III
THE COMPARISON OF CPU, MEMORY* AND POWER

Name	Processor	RAM	External memory	Power
Raspberry Pi	ARM BCM2835	256-512 MB	2-64 GB	5V/ USB
BeagleBone Black	AM335x 1GHz ARM® CortexA8	512 MB	4 GB	5V/ USB
Arduino	ATMEGA8, ATMEGA168, ATMEGA328, ATMEGA1280	16-32 KB	32 KB	7-12V /USB
MicaZ	ATMEGA128	4 KB	128 KB	5 V
TelosB	TI MSP430	10 KB	48 KB	5 V
Iris	ATMEGA1281	8 KB	128 KB	5 V
Cricket	ATMEL128L	4 KB	512 KB	5 V
Lotus	ARM NXP LPC1758	64 KB	512 KB	5 V

(The higher value is better)

C. Operating System

Sensor nodes run embedded software that samples the physical environment, load data, aggregates and communicate with higher level (peers or gateways). Regardless of the hierarchical approach each sensor node as well as mobile platforms, still needs a program, and the most common approaches to programming each sensor node, is to either program it using some form of operating system or to choose a higher level of abstraction. The operating systems vary from traditional operating systems in terms of goals and technique and each system differs substantially in the approach to memory protection, dynamic reprogramming, thread model, real-time features, etc.

The comparison of used operating systems by the mobile platforms and sensor nodes is shown in Table IV.

TABLE IV
THE COMPARISON OF OPERATING SYSTEMS

Name	Operating system	Programming language
Raspberry Pi	Raspbian, Ubuntu, Android, ArchLinux, FreeBSD, Fedora, RISC OS	C, C++, Java, Phyton
BeagleBone Black	Linux Angstrom	Arduino
Arduino	No	Arduino
MicaZ	Tiny OS, Mote Runner	No
TelosB	Tiny OS, Sos, Mantisos	No
Iris	Tiny OS, Mote Runner	No
Cricket	Tiny OS	No
Lotus	Rtos, Tiny OS	No

V. CONCLUSION

BWSNs and the arranging sensors in, on and around the human body, realizes the detection of human action and physiological information, which has been widely used in the fields of health care, social welfare, sports, entertainment, etc. The ubiquitous network is coming with the method of taking human body as a part of the communication network. One of the main issues in BWSN is the design of sensor nodes, which should be paid to node size minimization and energy consumption reduction.

Mobile platforms's performances are compared with are compared with some popular sensor nodes on a general level by computing power, size and overall costs of the solutions. Based on performed analysis, it can be stated that Raspberry Pi has the best performances among considered mobile platfrms. The ultra-cheap-yet-serviceable computer board, with support for a large number of input and output peripherals, and network communication is the perfect platform for interfacing with many different devices and using in wide range of applications. In other words, the Raspberry Pi brings the advantages of a PC to the domain of sensor network, what makes it the perfect platform for interfacing with wide variety of external peripherals.

Even there are large differences between stated platforms in their ideal use cases, energy requirements, OS, etc., it can be noted that all of them can be very successively applied as biomedical sensor nodes.

REFERENCES

[1] F. Ali, "A Middleware to Connect Software Applications with Sensor Web", The International Journal of Technology, Knowledge and Society Volume 6, Number 5, pp. 27-35, 2010.
[2] Ullah S., Higgins H., Braem B. A Comprehensive Survey of Wireless Body Area Networks. Journal of Medical Systems, June 2012, Volume 36, Issue 3, pp 1065-1094.
[3] Robert Bogue, (2012) "Plessey launches range of unique electric field sensors", Sensor Review, Vol. 32 Iss: 3, pp.194 – 198.
[4] Ma, Q.; Hou, X.H. Study of wireless sensor node's structure. Sci. Technol. Inf. 2008, 24, 371–436.
[5] M. Richardson and S. Wallace, Getting started with Raspberry Pi, O'Reilly, USA, 2013.
[6] Liolios, C.; Doukas, C.; An Overview of Body Sensor Networks in Enabling Pervasive Healthcare. In Proceedings of the 3rd International Conference Greece, 23–25 June 2010.
[7] Gong, J.B.; Wang, R.; Cui, L. Research advances and challenges of body sensor network (BSN). J. Comput. Res. Dev. 2010, 5, 737–753.
[8] Chen, M.; Gonzalez, S.; Vasilakos, A.; Cao, H.; Leung, V.C.M. Body area networks: A survey. Mob. Netw. Appl. 2010, 16, 171–193.
[9] Frehill, P.; Chambers, D.; Rotariu, C. "Using Zigbee to Integrate Medical Devices", Engineering in Medicine and Biology Society, 29th Annual International Conference of the IEEE. Aug. 2007.
[10] M. Schmidt, Raspberry Pi – A Quick Start Guide, The Pragmatic Bookshelf, 2012.
[11] D. Uckelmann, M. Harrison, F. Michahelles, Architecting the Internet of Things, Springer 2011.
[12] M. Maurya, S. R. N. Shukla, "Current Wireless Sensor Nodes (Motes): Performance metrics and Constraints", Inter Journal.

A Universal Android Tourist Guide Using the GPS Technology

Đorđe Manoilov[1], Dušan B. Gajić[1], Radomir S. Stanković[1]

Abstract – **This paper describes an Android tourist guide for mobile devices, based on the global positioning system (GPS) technology. We present the architecture of the application and its typical use-case scenarios. Employment of the presented application on archaeological sites allows visitors to access additional multimedia content and easily navigate through on-site objects. Other existing systems, aimed at similar purposes, are mostly focused on audio content, while the presented application also features video and 3D material. The use of XML decouples the implementation from the content of the application, thus allowing an almost straightforward adoption of the presented application on different archaeological sites.**

Keywords – **Android, mobile programming, augmented reality, cultural heritage, tourist guide, global positioning system - GPS.**

I. INTRODUCTION

The application of information technologies (IT) creates new opportunities for better presentation of cultural and historical objects and museum exhibitions. Multimedia presentations and applications for mobile devices make historical sites and art exhibitions more accessible and more interesting to a wider audience. Through their use, each visitor can find information about an exhibition to the extent that suits their interests. Mobile devices, such as smartphones and tablets, recently became omnipresent in everyday human activities. Tourism industry followed this change.

Development of digital tourist guides based on various information technologies can greatly help in several aspects related to tourism. In addition to helping with the complex preparations before starting a trip, mobile applications can help in solving the problem of lack of information about cultural and historical attractions. Tourists have at their disposal a number of activities, such as visiting attractions or restaurants and shops, but they don't have enough information to make an informed decision. A similar problem may be encountered during visits to archaeological parks or large museums, due to the large number of offered exhibits and lack of information about the preferred visiting routes. Further, there is often a problem of timing certain activities and having correct information about working hours of institutions. Finding an optimal route to a desired location in an unfamiliar place is another task made easier with mobile devices. This task is especially difficult in large cities, where important attractions are often distant from each other.

[1]Đorđe Manoilov, Dušan B. Gajić, and Radomir S. Stanković, are with the CIITLab, Faculty of Electronic Engineering, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mails: manoilov88@gmail.com, dusan.b.gajic@gmail.com, radomir.stankovic@gmail.com.

The Android tourist guide for mobile devices presented in this paper is developed by the ARhiMedia group from the Faculty of Electronic Engineering in Niš, Serbia. The proposed application offers one approach to solving problems discussed in the previous paragraph. Solutions to some of these problems were discussed earlier in [1, 2, 3, 4], while this study focuses on solving the problem of choosing the optimal route. The idea for this application emerged during visits to archaeological parks Mediana, close to Niš, Serbia, and Iustinana Prima – Caričin Grad, near Lebane, Serbia.

The paper has the following structure. In Section II, we present the motivation behind the development of the application, as well as its architecture. Section III contains a brief description of the user interface and implementation of the application. Section IV presents the most important conclusions, as well as plans for future work.

II. THE APPLICATION

The mobile tourist guide presented in this paper, is developed to allow users easier navigation through archeological parks and objects of great cultural and historical importance which very often cover large areas. It also provides information about objects of cultural and historical significance, such as three-dimensional (3D) reconstructions of objects and audio and video content about them. The application also gives users an opportunity to dictate the pace of the tour and to select additional information in accordance with their interests and available time. The presented application is already used in the digital mini-museum of the Faculty of Electronic Engineering Niš, as a demo tourist guide through the Niš medieval fortress and the Roman site of Viminacium, Serbia.

A number of applications designed for similar purposes exist. An example is an Android application *Taj Mahal Official Tour*, intended for visiting the Taj Mahal in India [5]. This application contains an audio guide with a map that shows points on which visitors can stop and hear additional information. Another example is the *Smart Tour Guide* application, created by the Korean Tourism Organization, which contains interesting anecdotes collected from the archives of Korean history [6]. Both applications, as well as most of the other existing applications of similar purpose, are devoted primarily to audio content. The tourist guide presented in this paper puts more emphasize on video and 3D materials. Further, the presented application is developed so that the entire multimedia content is independent of the implementation. Therefore, the application can be applied to any location without additional programming, just by changing the configuration file and adding appropriate multimedia content.

A. The Application Architecture

The application is developed for mobile devices with Android operating systems. Position between points of interest and current position of the user is determined using global positioning system (GPS) and device's orientation sensors - accelerometer and compass. Functions for communication with GPS satellites and for determining orientation of the device are available within the Android application framework [7].

The application architecture is shown on Fig 1. The application consists of several forms: *Compass*, *Map*, *List*, *Video* and form for models *3D Preview*. Application logic and multimedia data that are used are completely separated. Each of the mentioned forms will be explained in detail later in the text. Data that are used in the application are packed in an expansion file, which is downloaded along with the application from the *Google Play* market and stored in the external memory of the device.
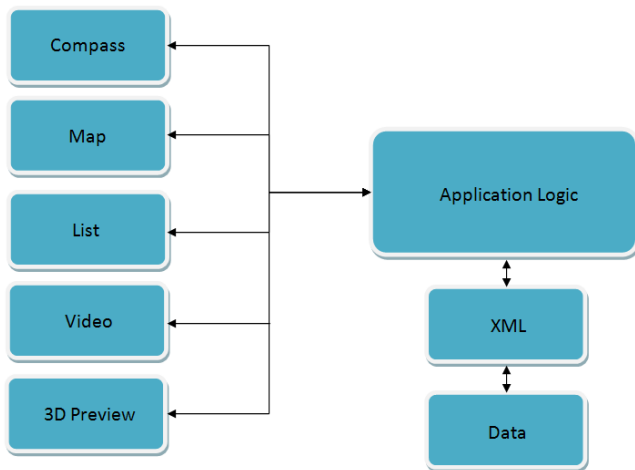
Fig. 1. The application architecture.

The maximum size of application that can be placed on the *Google Play* market is 50 MBs. Typically, there can be a large number of points of interest on archaeological parks. Therefore, there is a possibility that a lot of multimedia content might be necessary, so the application is implemented not to be size-limited. One possible solution are expansion files [8]. An expansion file contains all necessary information and is placed on *Google Play* along with the application. It is possible to add two expansion files, each with size up to 2 GBs. Conceptually each expansion file plays a different role. Main expansion file is the primary expansion file for additional resources that are necessary in application. Patch expansion file is optional, and it is designed for small changes in the primary expansion file. Both expansion file can be in any of the following formats: *ZIP*, *PDF*, *MP4*, etc. Independently of the type, *Google Play* renames them using the following scheme: *[main|patch].<version-of-expansion-file>.<package-name>.obb*.

After the download is complete, data are loaded into the application according to the XML configuration file. For our application, the XML configuration file is also a part of the

expansion file. The remaining content is classified into folders whit the tree structure shown in Fig 2.
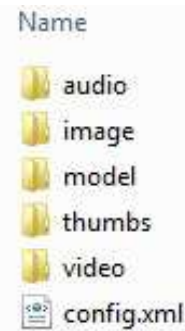
Fig. 2. Structure of the expansion file.

From the folder names, we can conclude that, for example, the *audio* folder keeps the necessary audio material, *model* folder stores the 3D models of reconstructed objects, etc. Materials to be loaded into the application for a particular object are determined by tags in the XML file. An example of a well-configured XML file is shown in Fig 3.

Fig. 3. A well-configured XML configuration file.

All important points in a site are placed within the <point> tag and they are all grouped within the root tag <points>. Every object contains the following tags: the <name> tag, which defines the name of object; <latitude> and <longitude> tags, which are used to define GPS coordinates of object; the

<radius> tag which represents radius around the object in meters, within which users can hear additional information about object. An audio file which will be played for an object is defined with the <file> tag. When a user is found within a defined radius of the object, he also has the possibility to see image of the object, video, or a 3D reconstruction of the object. This is all specified with the <img>, <video> and <model> tags, respectively. If there is a 3D reconstruction of a given object, it is necessary to export the model and its textures from the expansion file. After that, it is possible to successfully load it into the application. Files which need to be exported are defined using the <item> tags. All <item> tags are grouped within the <model_items> tag. For adding new points, its necessary to add the <point> tag, which contains all of the mentioned tags.

III. THE USER INTERFACE AND IMPLEMENTATION

The welcome screen of the application is shown in Fig 4a. In the top of the screen, a user can see the distance to the closest object, radius around, and the current accuracy of the GPS signals received from satellites. All distances are shown in meters, and calculations are based on the user's current GPS coordinates and the coordinates of the closest object in the archeological park.

In the case when GPS signal is not available or it is lost, user's last known location is used for calculations. This last known location is provided by a function called *getLastKnownLocation* [9]. Current GPS position of the user, which includes latitude and longitude coordinates, is stored in the device local memory using two double-precision floating point variables, one for each coordinate. Therefore, memory requirements for storing GPS data are measured in bytes, while the typical available device memory is in GBs.
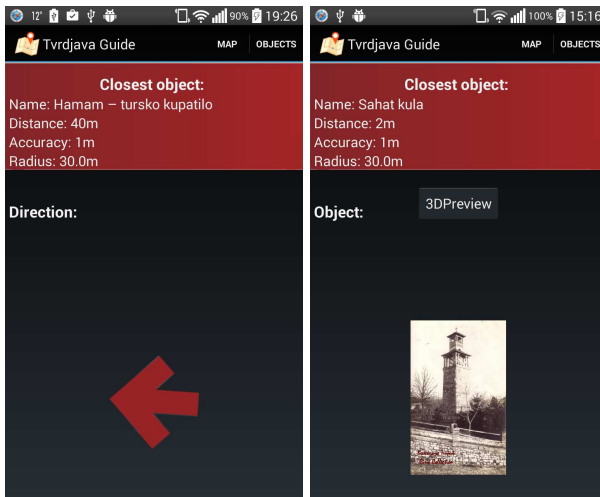
Fig. 4. a) Welcome screen of the application. b) The application screen shown when a user is within the object radius.

Besides these information, there is a compass in form of arrow in the bottom of the screen. Its purpose is to provide users easier positioning within the archaeological park. When a user walks through the park, this arrow is rotated so that its head shows the direction to the closest point.

The method of calculating the rotation angle is shown in Fig 5. In this figure, P1 is the current user location, P2 is the closest location of an object in the archeological park, α is an angle between P1 and P2 in relation to the North Pole, β is an angle of the phone, also in relation to the North Pole, and, finally, γ ($\gamma = \alpha - \beta$) is an angle for which it is necessary to rotate the arrow so that its head shows the direction of the closest point.
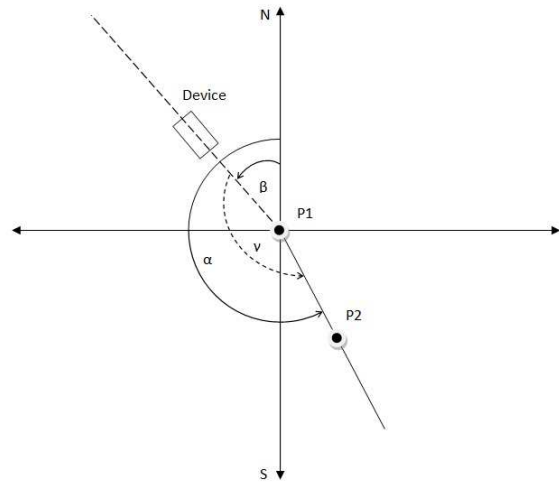
Fig. 5. Method for calculating the rotation angle.

When the user enters within the radius determined for an object, the application starts the corresponding audio file which is pre-defined in the XML file. This audio file can contain ambient sounds characteristic for an object, or additional information about it. Further, instead of an arrow, the application shows an image of the object. The main function of the image is to show to the user the reconstruction of object, if it is not preserved in its entirety, or to show an image of a mosaic, if it is, for protection, covered with sand. The previously described screen is shown in Fig 4b.

When a user is within the radius of an object, he can view a video clip about the object or a 3D model of its reconstruction. These two features are optional, and if they exist, buttons *3DPreview* and *Video*, respectively, are shown to the user.

By choosing the option *3DPreview*, the screen an example of which is shown in Fig. 7a is presented to the user. This application module is realized using *Metaio Mobile SDK*. *Metaio* is a tool which is primarily intended for developing of augmented reality applications for *Android* and *iOS* mobile devices, but it can be used for the implementation of this module [10].

With simple finger gestures, a user can rotate the model and thus view it from all sides. Zoom options are also implemented. In the top right corner, there is a model which represents *x*, *y*, and *z* axis, in order to provide users easier interaction with an object.

Selecting the *Video* option shows the user a form with a full-screen video about the object. A frame from an example video is shown in Fig 8.
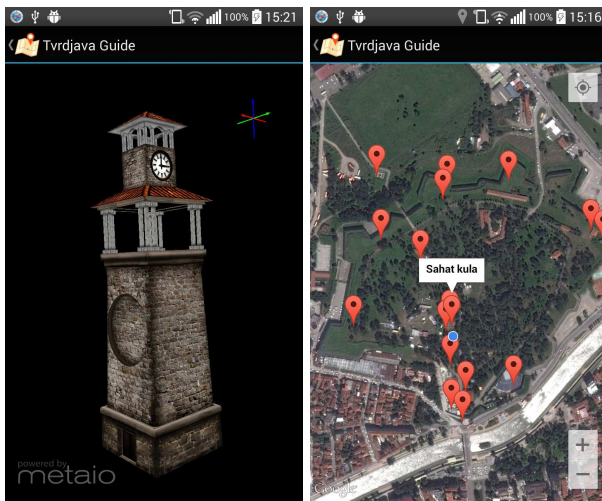
Fig. 6. a) 3D reconstruction of an object. b) Map of an archeological park.
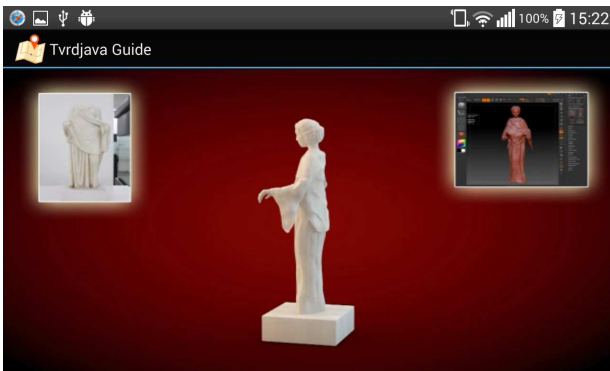
Fig. 7. Frame from a video content about an object.

On the top of the screen within the *Action* bar, there are two more options [11]. The first option is *Map*, which provides an overview of the entire archeological park and objects of interest. Objects are marked on the map as pins, and a pin which is the closest to the user is animated in order to be located easily. This form is realized using *Google Maps*. For our application, we used the *Google Maps API version 2* [12]. An example of using this functionality of the application is shown in Fig 7b. Another option within the *Action* bar shows a list of all objects of interest within the park.

All of the previously described application layouts are designed as responsive, which means that they are not dependent on neither screen resolution nor size.

IV. Conclusions

The application presented in this paper - *A universal tourist guide based on the GPS technology* - is implemented as an Android application for mobile devices. The application provides easier navigation through archeological sites using the GPS, accelerometer, and compass, which are available on almost all current mobile devices. Further, it provides textual information and audio and video contents about objects on the site. The presented application can be applied to any location without the need for additional programming, since its content is completely decoupled from its implementation.

Plans for future work include the implementation of an augmented reality module. The main functionality of this module will be to show the real-world environment augmented by 3D models of reconstructed objects in a common screen.

Acknowledgements

References

[1] R. Stanković, D. Tatić, M. Stošić, Đ. Manoilov, *ARhiMed*, technical solution - prototype, (in Serbian), Faculty of Electronic Engineering Niš, Serbia, 2014, available at: http://www.elfak.ni.ac.rs/ downloads/projekti/tehnicka-resenja/07-10-003-13/prijava.pdf, last access: Feb 8, 2015.

[2] D. Tatić, M. Stošić, Đ. Manoilov, R. Stanković, "Universal Mobile Cultural Heritage Guide Based on Android Technology", XII Conf. New Technologies and Standards: Digitization of National Heritage, Belgrade, October 2013.

[3] Đ. Manoilov, N. Gajić, M. Stošić, D. Tatić, "A Virtual Tour of the Mediana Archeological Park using Unity 3D Engine", XII Conf. New Technologies and Standards: Digitization of National Heritage, Belgrade, October 2013.

[4] Đorđe Manoilov, Dušan Gajić, "Razvoj virtuelne šetnje primenom programskog okruženja Unity i Microsoft Kinect-a", (in Serbian) in Proc. 20th YU INFO, Kopaonik, March, 2014, pp.271-274.

[5] Taj Mahal Official Tour, available at: https://itunes.apple.com/us/app/taj-mahal-official-tour/id531740560?mt=8, last access: March 29, 2015.

[6] Smart Tour Guide, available at: https://play.google.com/store/apps/details?id=kto.smarttour, last access: March 29, 2015.

[7] The Android Architecture, available at: http://www.tutorialspoint.com/android/android_architecture.htm last access: Feb 8, 2015.

[8] APK Expansion Files, available at: http://developer.android.com/google/play/expansion-files.html, last access: Feb 8, 2015.

[9] Android API – LocationManager, available at: http://developer.android.com/reference/android/location/LocationManager.html, last access: Jun 5, 2015.

[10] Metaio Mobile SDK, available at: http://www.metaio.com/sdk/, last access: Feb 8, 2015.

[11] Action Bar, available at: http://developer.android.com/guide/topics/ui/actionbar.html, last access: Feb 8, 2015.

[12] Google Maps API version 2.0, available at: https://developers.google.com/maps/documentation/android/, last access: Feb 8, 2015.

ORAL SESSION
COMPUTER SYSTEMS AND INTERNET TECHNOLOGIES II

Methodology for tools integration in the Online assisted Platform for Computer-aided design in communications

Galia Marinova [1] and Ognyan Chikov [2]

Abstract – **The paper describes the methodology for tools integration in the Online assisted platform for computer-aided design in communications Online-CADCOM, going through tools collection, verification, classification, estimation, passport definition and interconnections. A Knowledge-base containing tutorials for multitool task solution, projects and e-learning content is developed and supported in Online-CADCOM, as well as connections to standards, protocols and specifications of communication systems, economical(cost) estimates, prototyping and optimization are supported.**

Keywords – **Online CAD tools, Integrated online platform, communication system and circuit design.**

I. INTRODUCTION

Computer-aided design and Electronic Design Automation are entering a new era where online exposition and accessibility of tools are widely spread and integration approach, from specification to prototype realization, environmental influence and cost estimation are successfully applied. Observations on that topic are presented in the EDA360 vision of Cadence [2], where design creators are seen as integrators, providing application focused platforms, instead of chips. Engineering is considered in close collaboration, in a kind of an ecosystem, together with project and business management, green economy and education. Designs are application driven and with a high percentage of software impact. Software defined radio (SDR), Cognitive radio, Software defined networks (SDN) and Smart Grids illustrate the trend of Software defined hardware in communications system design. References [3, 4, 5] also point different aspects of tools integration for system design.

The Online assisted platform for computer-aided design in communications (Online-CADCOM) is as an attempt to provide designers and student with an instrument which allows them to take advantage of these new resources and their integration in order to perform application driven design and software defined hardware design. Online-CADCOM is a part of the platform OPTIMEK which concept is described in [1]. It manages links to online tools for computer-aided design of electronic and communication circuits and systems, which

[1]Galia Marinova is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: gim@tu-sofia.bg.
[2]Ognyan Chikov is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria.

are previously studied, estimated and provided with characterization passports. The paper describes the methodology for tools integration in Online-CADCOM, going through tools collection, verification, classification, estimation, passport definition and interconnections. A Knowledge-base containing tutorials for multitool task solution, projects and e-learning content, is developed and supported in Online-CADCOM, as well as connections to standards, protocols and specifications of communication systems and links to economical (cost) estimates, prototyping and optimization are supported.

First the methodology steps for tools integration in Online-CADCOM are described, then Web design and development approach for Online-CADCOM realization is presented and two design projects illustrate some advantages of Online-CADCOM.

I. METHODOLOGY STEPS FOR TOOLS INTEGRATION IN ONLINE-CADCOM

The methodology steps for tools integration in Online-CADCOM are:
- Tools collection and estimation,
- Tools classification,
- Tools characterization passports and
- Tools integration.

These steps are considered in details further.

A. Tools collection and estimation

The first step in Online-CADCOM is a large search and study of existing online tools for computer-aided design, estimation of their reliability followed by tests of their performance. Only tools that show reliability, that possess clear algorithms and theory fundament, that show good results in verification, are selected. Up to now more than 100 online tools for CAD in communications are studied. Some of the can be found on the Portal for online tools for Telecommunications: http://mircheva.free.bg/.

B. Tools classification in Online-CADCOM

The development of the platform Online-CADCOM manages the CAD tools based on Daniel-Gajsky Y model of design levels. Two main panels are defined on the home page of Online-CADCOM, each of them combining 2 design areas from the Y model and interconnections between them:

Panel 1: Structural and Behavioral areas from the Y model and Analysis and Synthesis features as interconnections between these two areas.

Panel 2: Physical/Geometrical/Topological and Behavioral areas, covering topology design and behavioral simulation based on electromagnetic and temperature simulators, using FEM, BEM, etc. methods. Structure extraction from topology is also classified in this panel.

In these Panels CAD tools are classified in categories. Categories of online CAD tools in Panel 1 are:

- RF and microwave design (*RF/MW*);
- Antenna design (*Ant*);
- Audio design (*Aud*);
- Analog design (*A*);
- Interface circuit design of ADC,DAC, etc. (*Int*)
- Digital design (*D*);
- Power supply design (*PS*).

Categories of online CAD tools in Panel 2 are:

- Printed circuit board design (*PCB*);
- Elements - resistors, capacitors, inductors, transformers, crystal oscillators, heat-sinks, etc.(*Elem*);
- Electric installation design (*EI*);
- Nanotechnology design (*Nano*);
- Outcome to prototype - Development boards, PCBs, FPGA/CPLDs, USRPs, Arduino controller, etc. (*P*).

Panels1 and 2 are shown on Figure 1. Online tools selected to be connected to Online-CADCOM are classified in some of the categories from Fig. 1.
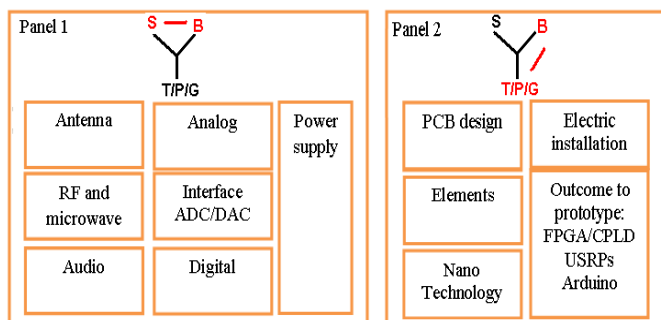
Fig.1. Panels on the home page of ONLINE-CADCOM

C. Tools characterization passports in Online-CADCOM

Each CAD tool to be considered in Online-CADCOM is characterized through 10 criteria defining a passport for each online tool as shown on Table I.

D. Tools connections in Online-CADCOM

Input/output connections of the tools considered permit to form sets of tools that solve complex design tasks in communications. Equivalence or application area overlapping, permit to enlarge multisolution synthesis in different application areas and to select optimal solution through verification tool estimates and/or cost estimates. Examples of equivalence or overlapping of application area of online tools are presented on Figures 2 and 3.

Figure 2 presents equivalence, functional area overlapping and connections of 4 online tools for filter design - 3 of the tools (Webench Filter Designer [7], Analog Filter Wizard [10] and FilterCAD [11]) perform active filter design and the forth one AADE [10] performs passive and quartz crystal filter design. Figure 3 presents Equivalence, Functional area overlapping and connections of Online tools for Switch mode power supply (SMPS) design Webench Power Designer [6] and PowerEsim [8].

III. WEB DESIGN AND DEVELOPMENT APPROACH FOR ONLINE-CADCOM REALIZATION

Besides the two Panels from Figure 1, Online-CADCOM develops and supports:

- Documents and links to Standards, Protocols and Specifications;
- Links to portals and platforms;
- Knowledge base containing E-learning content, Tutorials for complex task solutions in the multitool environment, Design Projects, Glossary;
- Economical estimation;
- Outcome to prototype;
- Links to Optimization tools.

These options are positioned in a separate left panel (Panel 3) on the Online-CADCOM home page.

Online-CADCOM is developed using HTML and PHP languages and MySQL for the Data base. The software architecture template MVC Framework is applied to separate Model, View and Controller parts. The Content management system permits the actualization of the platform independent of the web developer. The language CSS is used for the interface style.

IV. PROJECT DESIGN IN ONLINE-CADCOM

Two application driven designs in Online-CADCOM are presented to illustrate its performance. The first example consists in the design of a Bandstop filter with the online tool Webench Filter Designer which is verified through simulations in Cadence ORCAD Design suit 16.6. and the second example consists in the design of SMPS for a notebook using the online tools Webench Power Designer and PowerEsim.
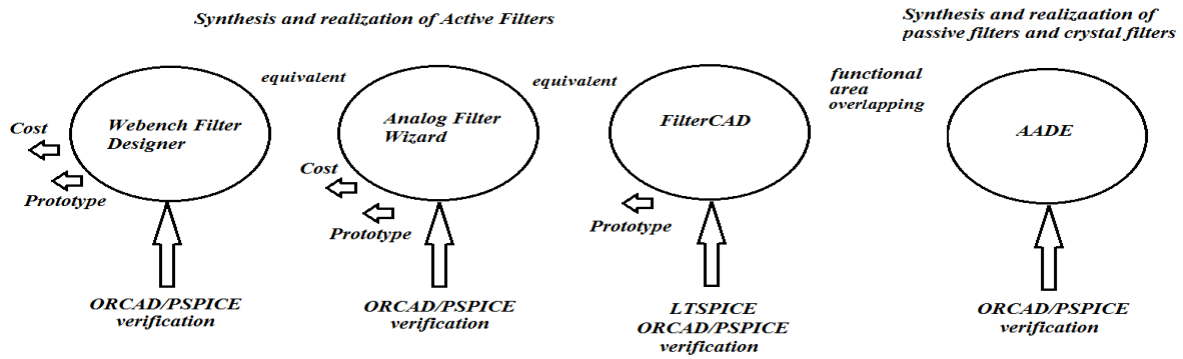
Fig.2. Equivalence, functional area overlapping and connections of online tools for filter design

TABLE I.
CHARACTERIZATION PASSPORT OF ONLINE CAD TOOLS

Type of Online CAD tool	Online calculator Online platform Module in Online Platform Free downloadable tool
Panel/ Category	Pane11 , 2 Categories from Fig.1 (Ant, Aud, RF/MW, A, Int, D, PS , PCB, Elem, EI, Nano, P)
Application area	Subcategories
Functions	Parameter calculation, Behavior, Synthesis, Analysis, Topology, Topology - Behavior Extraction
Levels of abstraction covered	Transistor, Logic, RTL, Architecture, System
Connections Input/Output	List of input and output connections to other online CAD tools
Verification tool	Simulator with high level of reliability which can verify the online tool results, ex.: ORCAD/PSpice, MATLAB, etc.
Equivalence or application area coverage	Sets of online CAD tools with application area equivalence or overlap
Qualitative Features	Theory fundament provided Friendly interface Graphical illustrations Traces and Waveforms building Animation 3D Author and Contact provided
Quantitative features	Number of parameters calculated Number of modules in a platform Number of circuits, ICs or topologies in a data base Number of models Number of elements Number of component providers
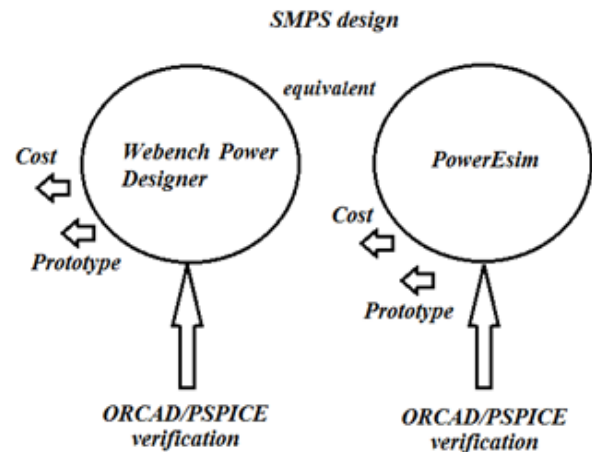
Fig.3. Equivalence, Functional area overlapping and connections of online tools for SMPS design

A. *Bandstop filter design with Webench Filter Designer with ORCAD/PSpice verification*

The specification of the filter to be designed is:
- Bandstop filter,
- Center frequency fc= 5kHz,
- Gain in the passband A=0dB,
- Minimal attenuation in the stopband ASb= -40 dB,
- Stopband bandwidth SB=100Hz,
- Passband bandwidth BWs=1000 Hz,
- Dual supply, Supply voltage ±12V.

Webench filter designer synthesizes a 6th order filter, with Linear phase 0.05° filter response, with three stages, with Bainter topology. The circuit is taken from the report file generated by Webench Filter Designer and it's presented on Fig. 4. Figure 5 presents Gain and Phase curves in frequency area, obtained in Webench Filter Designer. Then the design in verified through a project in ORCAD/Capture on Fig.6 and ORCAD/PSpice AC simulations of the Gain and the Phase, presented on Fig.7.

In Table II are shown the results for the filter parameters obtained from Webench Filter Designer simulations and from Cadence ORCAD 16.6 simulation and the difference between both estimations is calculated in percent.
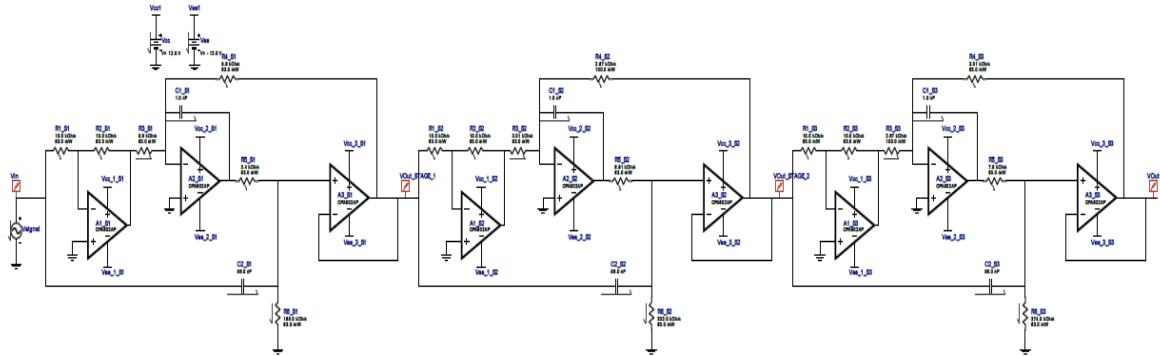
Fig.4. Bandstop filter synthesized in Webench power designer
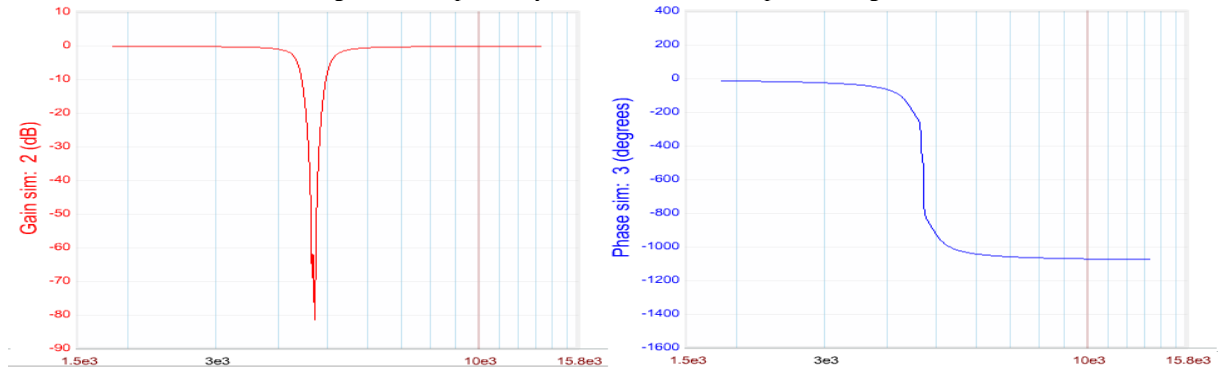
Fig.5. Gain and phase in frequency domain for the bandstop filter from Figure 4, obtained in Webench Power Designer tool
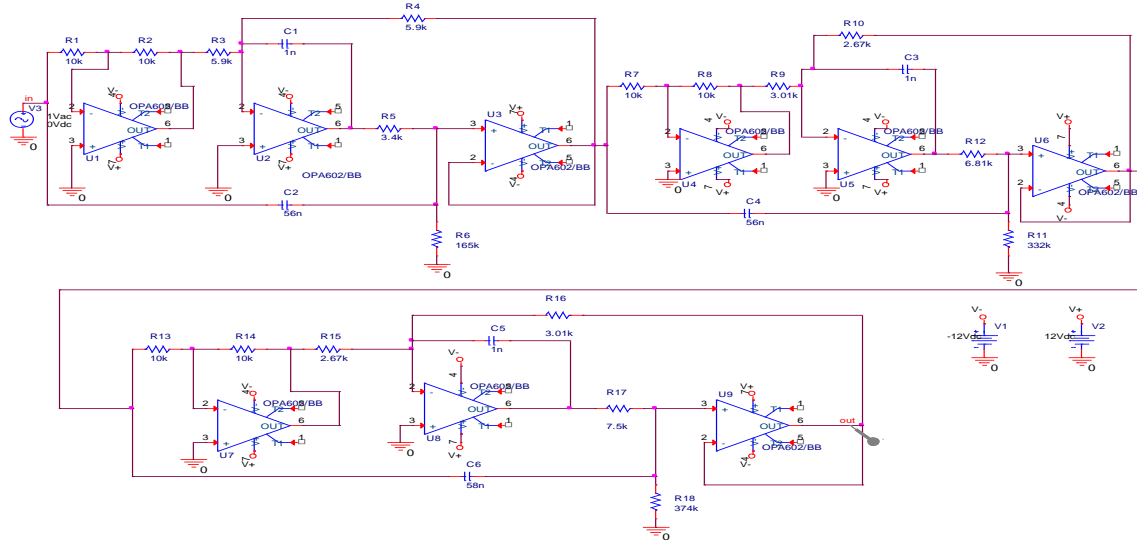
Fig.6. Electrical circuit of the Bandstop filter from Fig.4 in ORCAD Capture 16.6
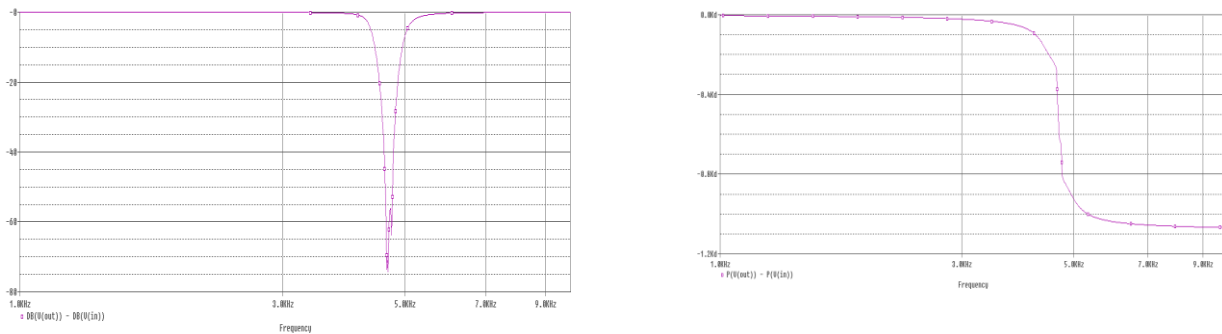
Fig.7. Gain and phase of the Filter from Fig.6, simulated in ORCAD/PSPICE 16.6

TABLE II.
COMPARISON OF PARAMETER ESTIMATIONS FOR THE FILTER FROM FIG.4 IN WEBENCH FILTER DESIGNER AND CADENCE ORCAD/PSPICE 16.6

Parameter of the filter	Estimation in Webench Filter Designer	Estimation in Cadence ORCAD/PSPICE 16.6	Difference in estimates in [%]
Center frequency fc	4.731 kHz	4.656 kHz	1.61%
Attenuation at fc - ASb	-77.4 dB	-74 dB	4.6%
Stopband SB	135 Hz	360 Hz	62.5%
Passband BWs	968 Hz	846 Hz	14.42%

Results obtained in Webench Filter Designer are closer to the specification than those obtained in Cadence ORCAD design suit 16.6, so it seems that Webench Filter Designer gives more optimistic results. The PSpice verification is one more guarantee for the results obtained and they might occur to be more realistic.

B. *Comparative design of SMPS for a notebook with Webench Power Designer and PowerEsim*

The specification of the SMPS for the notebook to be designed is:

The input voltage is an AC voltage with value   220-240 V and frequency 50 Hz, which makes for the   Vmin(Vin RMS)=156 V and Vmax(Vin RMS)=167 V .

The output voltage Vout is 15 V and the output current Iout is 3.75 A, the ambient temperature is 30º.

A minimal acceptable value of the SMPS efficiency coefficient is 65%.

The SMPS for the notebook is specified in PowerEsim and it proposes 5 possible solutions as shown on Fig.8. Table III presents the values of the efficiency coefficients for each of the solutions. Only 3 solutions have acceptable efficiency coefficient values, since they have values of the efficiency coefficient superior than 65% , these are the topologies - Buck PNP Buck DC-DC and Flyback AC/DC.
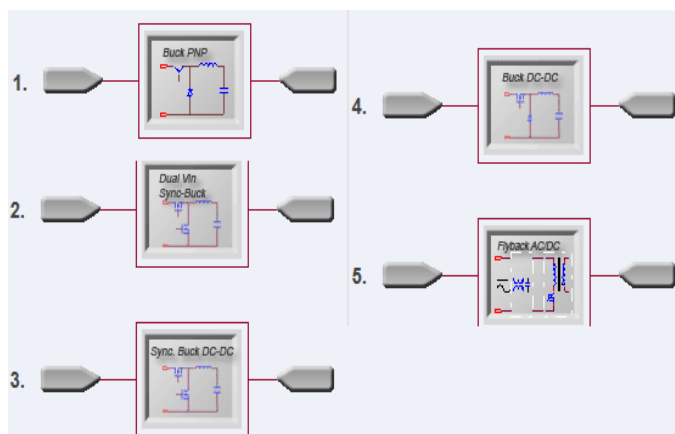
Fig. 8. SMPS designs for the notebook in PowerEsim

The topologies Dual Vin Sync Buck and Sync Buck DC-DC are rejected since their efficiency coefficients values are too low.

TABLE III.
TOPOLOGIES AND EFFICIENCY COEFFICIENTS OF THE SMPS DESIGNS IN POWERESIM.

SMPS topology in PowerEsim	Efficiency coefficient
Buck PNP	67%
Dual Vin Sync Buck	5%
Sync Buck DC-DC	33%
Buck DC-DC	87%
Flyback AC/DC	85.58%

The same specification is synthesized with the tool Webench Power Designer. An unique solution is proposed as shown on Fig.9. The efficincy coefficient obtained for the power supply designed is 86%. The circuit synthesized, which is taken from the Webench Power Designer report file is shown on Fig.10. Webench posts a message (*why other parts were not found*) that several hundreds of   ICs were tested for this specification and each of them doesn't meet the current or the voltage constraints in the specification (*Current exceeds limit* or *Voltage is outside of limit*) or they *are not configured for AC/DC Conversion*.

The comparison between these two online platforms for SMPS design shows that PowerEsim proposes more numerous solutions, most probably because this tool supports ICs from multiple providers and Webench supports mostly TI ICs. This makes PowerEsim more versatile and the possibility to consider both platforms which is provided by Online-CADCOM helps to increase the number of solutions and thus gives a wider choice for designers and users.

Part	Create	WEBENCH® Tools	Schematic	BOM Images	Design Considerations	BOM Footprint (mm2)	BOM Cost (1ku)	Eff (%)	BOM Count	Freq (kHz)	Vout p-p (mV)	Xover Freq (kHz)	Phase Margin (deg)	Topology	LDO	Temp (deg)	Iout Max (A)
LM5023	Open Design				AC-DC QR Current Mode PWM Controller	1778	NA	86%	44	67	67.72	NA	NA	Flyback	N	32°C	5.00

Fig. 9. Unique design solution for the SMPS for a notebook in Webench Power Designer
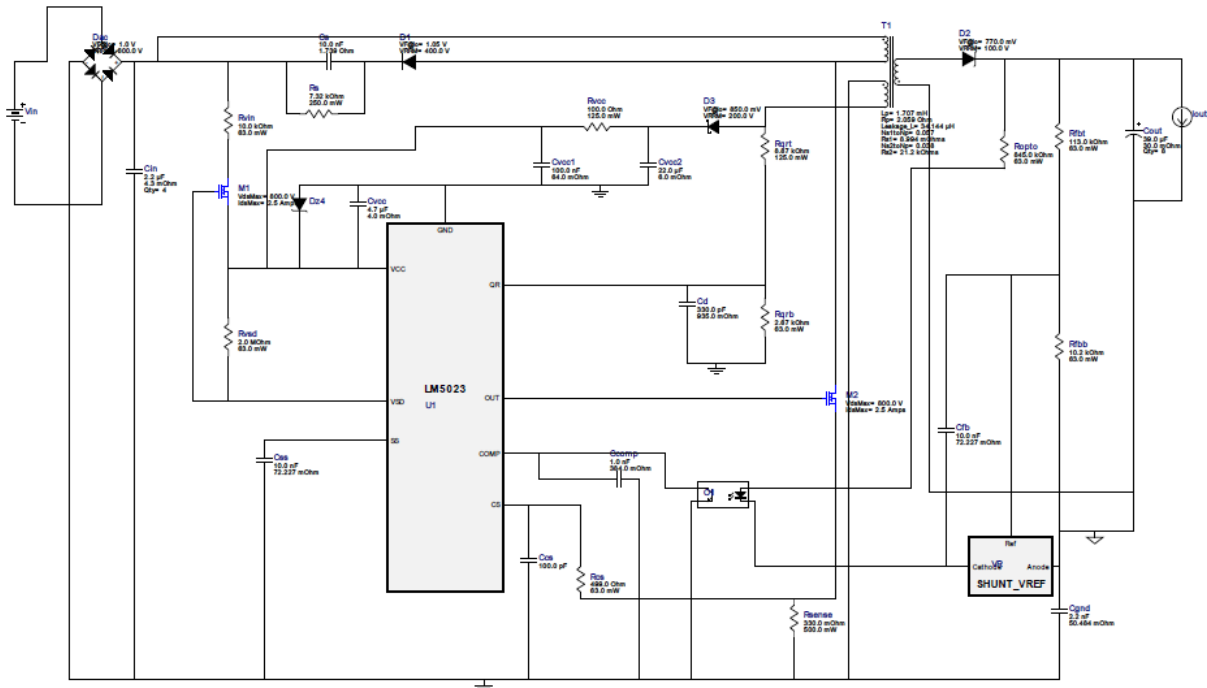
Fig. 10. Power supply circuit for the notebook generated in Webench Power Designer tool

V. CONCLUSION

The paper defines the main steps of the methodology for tools integration in the Online assisted Platform for Computer-aided design in communications and it illustrates some advantages for designers when using Online-CADCOM for developing high quality designs with reduces time cycle and outcome to prototype. Further results for Online-CADCOM structure and content, as well additional projects developed, will be presented in future works.

ACKNOWLEDGEMENT

REFERENCES

[1] G. Marinova, V. Guliashki, O. Chikov, "Concept of Online Assisted Platform for Technologies and Management in Communications – OPTIMEK", International Conference in Computer Science, Information System and Telecommunication, ICCSIST 2014, 7-9 November 2014, Durres, Albania, 2014

[2] Cadence, "Delivering on the EDA360 Vision", 2011, http://www.cadence.com/downloads/eda360/EDA360_Milestones.pdf

[3] Farrahi, A.H. , Hathaway, D.J., Wang, M., Sarrafzadeh, M., "Quality of EDA CAD tools: definitions, metrics and directions", ISQED 2000. Proceedings. IEEE 2000 First International Symposium on Quality Electronic Design, San Jose, USA, 20 -22 Mar 2000, pp. 395 – 405, 2000

[4] Satyandra K. Gupta, William C. Regli, Dana S. Nau, "Integrating DFM with CAD through Design Critiquing", Concurrent Engineering: Research and Applications, Vol. 2, N 2, 1994

[5] Zhen Mu, Combining "Multiple Simulations to 12.5 Gbps to Guide CAD Tool Use", DesignCon 2002

[6] Webench Power Designer, Texas Instruments http://www.ti.com/lsds/ti/analog/webench/power.page

[7] Webench Filter Designer, Texas Instruments http://www.ti.com/lsds/ti/analog/webench/webench-filters.page?DCMP=hpa_amp_opa1664_en&HQS=active-filter-designer-pr

[8] PowerEsim, http://www.poweresim.com/

[9] AADE Filter Design and Analysis http://aade.com/filter.htm

[10] Analog Filter Wizard, Analog Device http://www.analog.com/designtools/en/filterwizard/#/type

[11] FilterCAD, Linear Technology http://www.linear.com/designtools/software/

Assessment tool based on semantic content annotation

Dejan Todosijević[1] and Martin Jovanović[2]

Abstract – **This paper describes system for students' knowledge assessment which is based on the technologies of the Semantic Web. This framework is a follow-up of the development of DSi (Drag and Drop Semantic Interface) tool for e-learning with the important difference that document for semantic annotation of teaching materials serve as the basis for creating questions for knowledge tests. The aim of the project described in this paper is software that enables creation of individualized knowledge tests, depending on the predefined semantic relations between terms in the teaching material. The paper gives a brief overview of the previously implemented versions of the DSi framework. The architecture of the implemented system and a description of the user experience of using this system are presented.**

Keywords – **Knowledge Assessment, E-learning, Semantic Web, DSi.**

I. INTRODUCTION

In this paper we present an approach of using Web technologies (Web 2.0 and Semantic Web) to create framework for e-learning. This approach to learning is enhanced with additional true-false semantic relations between notions in order to use it as a knowledge assessment tool.

Use of the well-known DSi principle of semantic annotation of learning materials can be applied to assess students' knowledge. The aim of the application described in this paper is use of semantic layer as a basis for creating questions required for testing knowledge. Constantly adding new relations in semantic document is one of the advantages of this approach. That simultaneously increases the number of questions that can be used in knowledge tests.

E-learning can be defined as delivery of teaching materials to students using computer technology as a tool for transferring knowledge. In the basis of E-learning is individualized and dynamic approach to the creation of learning materials. This means that it is necessary to distribute teaching materials to student at the time when it is needed and in the quantity needed. On the other hand, the Semantic Web represents a vision of the Web as an environment that allows "machines" to automatically acquire, integrate and process data. And then based on these data programs can fully autonomously reach conclusions that will be presented to users on their request. [1] DSi concept of learning is positioned at the intersection of the fields of E-Learning and Semantic Web. The concept on which this framework is based allows completely individualized delivery of semantic

[1]Dejan Todosijević is with the University of Niš, Faculty of Electronic Engineering, Aleksandra Medvedeva 14, 18000 Niš, E-mail: todosijevic.dejan@gmail.com.

[2]Martin Jovanović is with the University of Niš, Faculty of Electronic Engineering, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mail: martin.jovanovic@elfak.ni.ac.rs.

relations depending on the terms in the teaching material that are crossed (crossing terms means making connection between terms using drag-and-drop action). Delivery of additional knowledge or semantic relations between terms in the teaching material is enabled on the user's request, using drag-and-drop action. This tool has been developed using Semantic web technologies. Semantic Web technology provides the ability of using semantic layer for the purpose of creating questions for knowledge tests. [2]

In this paper term semantic annotation is used for explaining the process of enabling users to get some extra information about the relations between notions in the teaching material. Semantic annotation can be defined as a specific metadata generation and usage schema, which is aiming towards enabling new methods for accessing information and extending the existing ones. [3] Also as an annotation assigns to an entity, which is in the text, a link to its semantic description. [4] In simpler terms it represents the process of enhancing entities, found in the text, with additional information about them or relations between them. Calling this annotations "semantic", "entity" or in some other way it is just matter of terminology. [3] DSi relies completely on this component, connecting data by meaning, and enabling users to get some extra information about relationships between two entities. Terms, which can be found in textual teaching material, are kept in semantic document as resources. These resources combined with additional information form structures similar to sentences called statement (subject-predicate-object). Statement structure is the basis of Semantic Web applications, so this kind of term annotation of textual teaching material comes very naturally and is to be expected.

The development of the DSi tool will be briefly described further in the paper. Furthermore, implemented system for students' knowledge assessment will be described in detail and also the possibilities of using this tool.

II. DEVELOPMENT OF DSi TOOL

DSi (Drag and Drop Semantic Interface) is a system developed to support e-learning. It was developed in the CIITLAB laboratory, at the Faculty of Electronic Engineering and aims to improve the existing systems for e-learning. [5] This system, using the Semantic Web technologies, provides the users with the possibility of making connections (relations) between terms in textual teaching material. Making connections is achieved by simple and intuitive interface to the user, dragging one term onto another (drag-and-drop action). [6] [7] Framework is still under development, there have been developed several versions of the system in which the emphasis was on learning acceleration using this concept. In the first version (1.0 and also version 1.5) system provides one-way communication with the user. It is possible to drag and drop terms on each other and as a result of this action relations stored within the semantic document are presented to

the user. In version 2.0 two-way communication with the user is achieved, in addition to learning users can participate in the creation of semantic documents, adding new relations between terms. Versions 1.5 and 2.0 are recognized by the Faculty of Electronic Engineering in Nis as technical solutions in the field of software. [8] [9] Figure 1 shows a demo version of the application DSi 1.5.
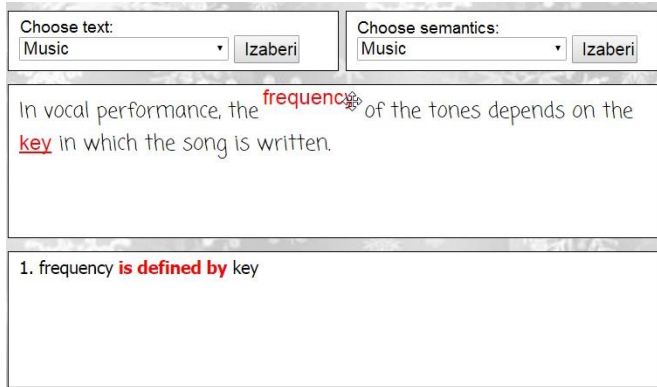
Fig. 1. Preview of demo application DSi 1.5

A system for testing DSi concept in real operating learning conditions has also been developed. This system monitors and records all user actions during the whole session of learning. It is essentially based on the version 1.5 of DSi framework with the addition of modules for tracking users' actions and for testing knowledge. This system has been designed as a Web application in which students are presented with learning materials and in addition with the possibility of dragging and dropping terms on each other. At the same time questions related to the lesson are presented to them. The basic idea was to create a learning environment in which there is a standardized knowledge tests in order to examine the impact of the proposed concept to learning acceleration. Figure 2 shows system for evaluating DSi tools.
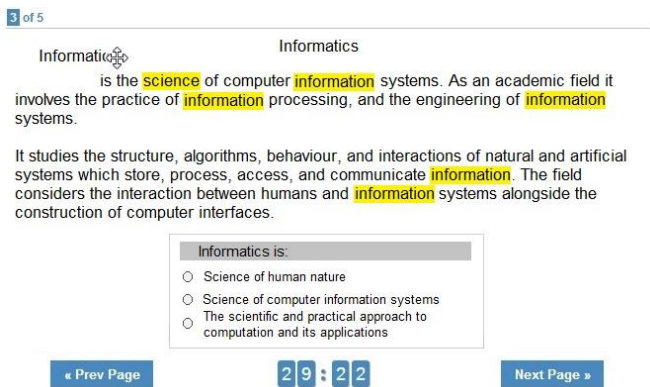
Fig. 2. System for evaluation of DSi tool

III. KNOWLEDGE ASSESSMENT

DSi-A (Drag and Drop Interface Semantic Assessment) is follow-up of the development of the DSi framework, with the important difference that in this version the concept of dragging and dropping terms on each other is used to evaluate knowledge and not just for learning. Approach to learning by dragging one term to another is combined with possibility of choosing one of the presented relations after crossing two terms. The user can select one of several presented relations. Action of selecting one relation means that user believes that selected relation is properly defined, which means that the other provided relations are incorrectly defined. This tool can be defined as a concept of evaluation of students' knowledge by combining the technologies of the Semantic Web (use of RDF for storing information, i.e. semantic relations in form of resources) and crossing terms in learning materials on a web page (drag-and-drop functionality). Use of these technologies form a unique test that is used for assessing students' knowledge based on pre-defined criteria. DSi-A is designed as a stand-alone Web application, but the concept is applicable to any system for e-learning, which allows preview of teaching materials as textual lessons.

This concept of using DSi tool was first introduced in [10]. As part of this paper the initial concept of knowledge assessment by using connections between terms stored in semantic document was presented. During the implementation of this framework there have been made some changes, specifically the manner in which the relations will be kept within the RDF document as well as the manner in which correct and incorrect relations will be classified for easing the process of checking the accuracy of user responses. As already has been said this system is still under development and is expected to make certain changes in further development.

Further in the paper we present the architecture of the implemented system and DSi assessment tool user experience.

The basic components of this application, as in previous versions of the DSi system are client and server sides. Client-side of the system is implemented in JavaScript, it is used to display teaching materials and to enable the functionalities needed for human-computer interactions. The server side is implemented in PHP, and is responsible for all application logic related to semantic documents and replying to requests received from the client.

It is planned that the system has two user groups: students (whose knowledge will be assessed) and teachers (who will create a semantic documents for testing). Given this division and previous versions of the application, architecture is divided into three subsystems: module for creating new semantic relations, module for the assessment of knowledge and a module for tracking user actions. In the subsystem for creating relations between terms the user is prompted to enter new relations and to define which of those relations are correct, the rest are declared faulty. Knowledge assessment subsystem relies on recognition of right or wrong-defined relations. Subsystem for monitoring user actions was first presented as part of a system for evaluating DSi tools for e-

learning in the paper [11]. This subsystem relies entirely on the architecture of mentioned system.
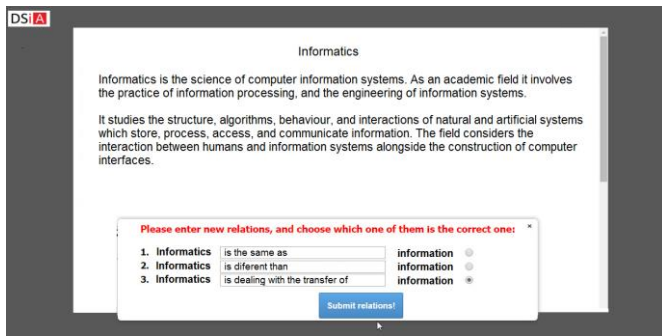
Fig. 3. Adding new relations to semantic document, the definition of correct/incorrect relations

The process of preparing teaching materials that will be used for knowledge assessment purposes, consists of preparing two documents: textual lessons and semantic document (RDF document with relations between terms). The lesson needs to be prepared in HTML format. It is kept on the server and is fully delivered to the client at startup testing. Semantic document is used for storing the connections between terms in the teaching material.

The teacher creates this document as a basis for testing students' knowledge. In this document, there are two types of relations between terms: correct and incorrect relations. With each of these relations information about whether it's correctly defined or not are also kept. After dragging and dropping the term on to another to user are presented relevant relations saved in this document. The entire document is never fully available on the client, but only some parts that are required on drag operation. In this way it is provided an additional level of security of semantic layer.

It is planned that the semantic documents preparation is done by teachers, people who are versed in the subject, i.e. people with sufficient knowledge of the matter, which will be presented to students in the form of lessons. The teaching material is presented in textual format, and all the words in the lesson are draggable and droppable. Defining new relations is possible by dragging and dropping one word on to another. After crossing two words (terms) a form for creating new relations is shown on the web page. It is possible to enter a certain number of relations and it is crucial to define which relation is correct. All other relations will be automatically declared as incorrect (Figure 3). System also gives the possibility of deleting the defined relations. If a teacher wants to delete relations already defined, crossing the same terms system will display those relations as well as the option to delete them.

Assessment of students' knowledge is planned as finding as many correct relations between key terms in prepared lessons, as well as the identification of those relations that are incorrectly defined. As mentioned above, for now it is possible to drag all words in the text, and the student needs to recognize which two terms can be related. Depending on the existence of terms crossed in the semantic document, response to student request will be formed. As a response to this action to them will be displayed a certain number of predefined relations. Figure 4 shows the layout of a web page after crossing two terms. The student can respond by selecting one of these relations as the only correct one from all relations presented. If student believes that he was offered only with faulty (incorrectly defined) relations he can answer that none of them is correct. Of course it is possible to cancel answering to this question if the student is not sure what a good answer is. Each of these actions will be stored in a database so that after the end of the session certain conclusions can be reached based on the data collected.
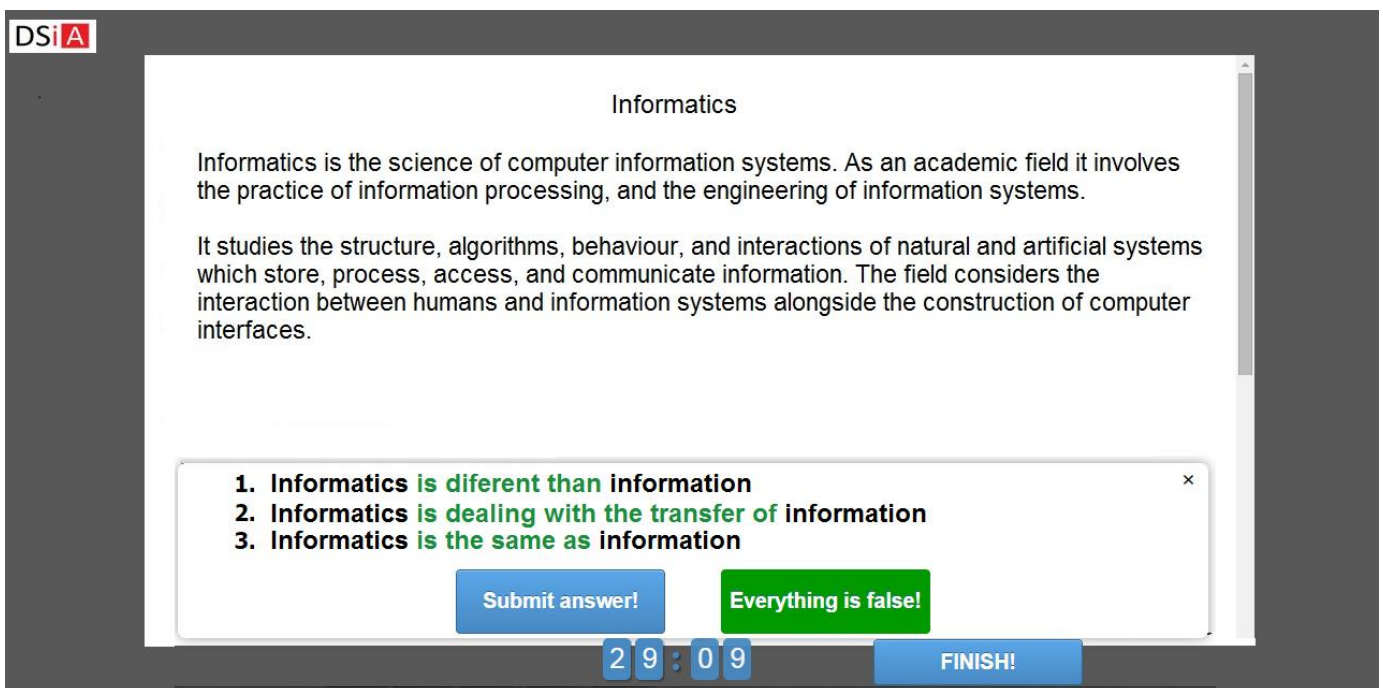
Fig. 4. Web page after crossing two terms, answering to the question presented

IV. Conclusion

DSi-A represents a tool which can be used for assessing knowledge of the individual and his progress in learning. User progress can also be compared with a certain predefined criterion. Collecting information about users while using the system and combining and analyzing those information can be used to assess their knowledge. Assessing students' knowledge is planned as finding as many correct relations between the key terms in the prepared lessons, as well as the identification of those relations that are incorrectly defined. This system also enables the preparation of semantic documents for use in the knowledge tests. The idea is that domain experts with sufficient knowledge of the subject participate in the preparation of these documents necessary for a knowledge test. Every test for each student is completely individualized and depends only on the selection of relations in the lesson. The questions that will be presented to students depends precisely on that selection and also on crossing those selected terms from the teaching material. The concept of such standardization of questions for knowledge tests provides the possibility of later expansion of the system in the form of the introduction of automated reasoning and creating new relations based on pre-defined rules.

Further development of the system will go in the direction of separation of terms (words) that may be the key terms, or the separation of those words that makes no sense to drag (and cross with other terms) in a separate semantic document. It is necessary to implement a simple interface that would allow domain experts to easily and quickly add these words in new semantic document. The idea is that when creating semantic document the user can identify which words does not make sense to drag and depending on this to create a local "base" of these terms. It is necessary to allow quick and easy search of the document. After that it is planned to develop a central repository of those words that makes no sense to drag. For knowledge tests for the specific lessons it would be necessary to take segments of that central repository containing words that appear within the test.

Acknowledgement

References

[1] V. Devedžić - Education and the Semantic Web, International Journal of Artificial Intelligence in Education 14 (2004), pages 39-65.

[2] http://www.w3.org/standards/semanticweb/

[3] A. Kiryakov, B. Popov, I. Terziev, D. Manov, D. Ognyanoff - Semantic annotation, indexing, and retrieval, Web Semantics: Science, Services and Agents on the World Wide Web (2004), pages 49-79

[4] H. N. Talantikite, Dj. Aissani, N. Boudjlida - Semantic annotations for web services discovery and composition, Computer Standards & Interfaces (2009), pages 1108-1117

[5] M. Jovanović, "Semantička nadgradnja nastavnog materijala u sistemima za elektronsko učenje", master's thesis, University of Nis, Faculty of Electronic Engineering, 2009.

[6] M. Jovanović - "Arhitektura konceptualno-orijentisanog alata za podršku učenju", Infoteh - Jahorina, Conference Proceedings, Vol. 6, Ref. E-IV-4, pp. 477-479, issued by Elektrotehnički fakultet Istočno Sarajevo, Jahorina, 2007, ISBN 99938-624-2-8.

[7] M. Jovanović, D. Todosijević, "DSI – Okvir za semantičko proširenje e-učenja", Yu Info - Kopaonik, Conference Proceedings, Kopaonik, Serbia, 2013, ISBN: 978-86-85525-11-7

[8] Version DSi 1.5 is recognized by the Faculty of Electronic Engineering in Nis as a technical solution, called "DSI (Drag and Drop Semantic Interface) - softverski okvir za semantičku dopunu tekstualnog nastavnog sadržaja u sistemima za e-učenje", authors: Martin Jovanović, Dejan Todosijević, Milena Stanković. Technical solution is recognized by the decision of the Academic Council of the Faculty No. 07/10-004/13-001, in Nis, 17.01.2013.

[9] Version DSi 2.0 is recognized by the Faculty of Electronic Engineering in Nis as a technical solution, called "Vizuelni uređivač ontologija (DSI 2.0) - softverski okvir za vizuelno uređivanje predikata kod grafova i ontologija u okviru Semantičkog Weba", authors: Martin Jovanović, Dejan Todosijević, Milena Stanković. Technical solution is recognized by the decision of the Academic Council of the Faculty No. 07/10-015/14-001, in Nis, 26.12.2014.

[10] M. Jovanović, V. Ognjenović, D. Todosijević, "Open Semantic Assessment: A Multiplied Choice Approach To E-Assessment", International Conference on Information Technology and Development of Education ITRO 2014 - Zrenjanin, Conference Proceedings, Zrenjanin, Serbia, 2014, ISBN: 978-86-7672-225-9

[11] D. Todosijević, M. Jovanović, "Sistem za evaluaciju semantičkog alata za e-učenje - DSI", Yu Info – Kopaonik, Conference Proceedings, Kopaonik, Serbia, 2014, ISBN: 978-86-85525-13-1

Evaluation of VoIP speech quality using neural network

Angel Garabitov[1] and Aleksandar Tsenov[2]

Abstract – **An important and unsolved problem is the automatic quantification of the speech quality. Particularly interesting is the ability of this task to be executed in real time. Many parameters affect the voice quality, but only subjective assessments are decisive. The article makes a proposal to create cost effective and efficient solution for prediction the subjective speech quality assessment. A neural network is used as a "customer". For training of the neural network, BYE Statistics is used. It gives the client's view for the voice quality. Result of some practical experiments is given in article.**

Keywords – **VoIP, quality of speech, neural network**

I. INTRODUCTION

The evaluation of the quality of a VoIP telephony is usually a difficult task that is performed by the statistical processing of the satisfaction of a group of users. In this paper is developed the idea for pseudo-subjective quality assessment of VoIP telephony. To evaluate the effects of various QoS-related parameters, we can use a neural network that can convert the multidimensional assessment criteria in simple relationships. For training the neural network are been given various QoS-related parameters such as PSQM +, end-to-end delay, packet loss in VoIP system. Results confirmed that this method can effectively evaluate the overall QoS level QoS of several related factors.

II. THEORETICAL BACKGROUND

A. Using artificial subjective quality rating to analyze perceived voice quality in VoIP

The determination of Subjective measurement methods is a difficult task. Subjective measurements of QoS are carried out by use of a group of people [1, 2, 3, 4]. A test phrase is recorded and then volunteers listen to it under various conditions. These tests are performed in special rooms, with background noise and other environmental factors, which are kept under control for test executions. Some examples are: conversation opinion test, hearing test review, an interview and survey test. There are also some disadvantages to this type of tests, such as the experience, the humor and the culture of the subject. These tests are very expensive and are not implemented in practice because of the large number of people needed and running truthful results obtained.

This type of test is used for unidirectional transmission, and it will have been carried out on transmission by examination of the recorded conversations or phrases. The goal of this test is to evaluate the performance of the individual connections and algorithms under different conditions. Some of the known hearing tests are as follows: ACR (Absolute Category Rating), DCR (Degradation Category Rating) and CCR (Comparison Category Rating).

B. Selecting an appropriate network topology

The education network in a specific order searches the training set. This order can be sequential or random. Some networks, that are trained without a teacher, just once to search the sample. Other networks that are trained with a teacher who browse some of the sample several times in the pass of learning. Even if they are successful at first glance, the behavior of the neural network is not always such that, what is required by the network. Assessment of quality of training of the neural network must be performed by examples that have not participated in its design. In addition, the number of test cases must be greater to achieve higher quality of the training of the network. It turns out that the testing of well-trained neural networks is a very difficult task. The assessments of the quality related parameters is possible with the BYE message.

C. Speech Quality

QoS measurement and speech quality prediction is an object of this research, which is a classical problem in VoIP networks. Upon examination of the problem, theoretically the most sense factors on speech quality are jitter, packets sent, lost packets, codec and quality.

All VoIP manufacturers [5] "Send In" the speech quality. This is shown in Fig. 1. As described, the input data of the neuron network are jitter (shortened to the BYE message with JI), packet sent (shortened to the BYE message with PS), packet loss (shortened to the BYE message PL) and codec type (shortened to the BYE message with EN).

The first variable is jitter. In packet-switched network, when a sequence of packets is send from point A to point B, each of the packets will require a slightly different time to reach the destination. The various maturities are not a problem if a web page is downloaded, but it is when we want to transfer a stream of data in real time.

Formally the jitter is defined as the statistical variance of the RTP data packet inter-arrival time. In the Real Time Protocol we measure jitter in timestamp units.

The first step to deal with jitter is to know its value. However, we do not need to calculate the exact value. In the RTP the receiving endpoint calculates an estimate with a simplified formula (a first-order estimator).

[1]Angel Garabitov is a former master student at the Faculty for German Engineering and Industrial Management Education at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: angel.garabitov@fdiba.tu-sofia.bg.

[2]Alexandar Tsenov is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: akz@tu-sofia.bg

The jitter estimate is sent to the other party using RTCP (Real Time Control Protocol).

The second and the third variables are Packet loss and Packet sent.
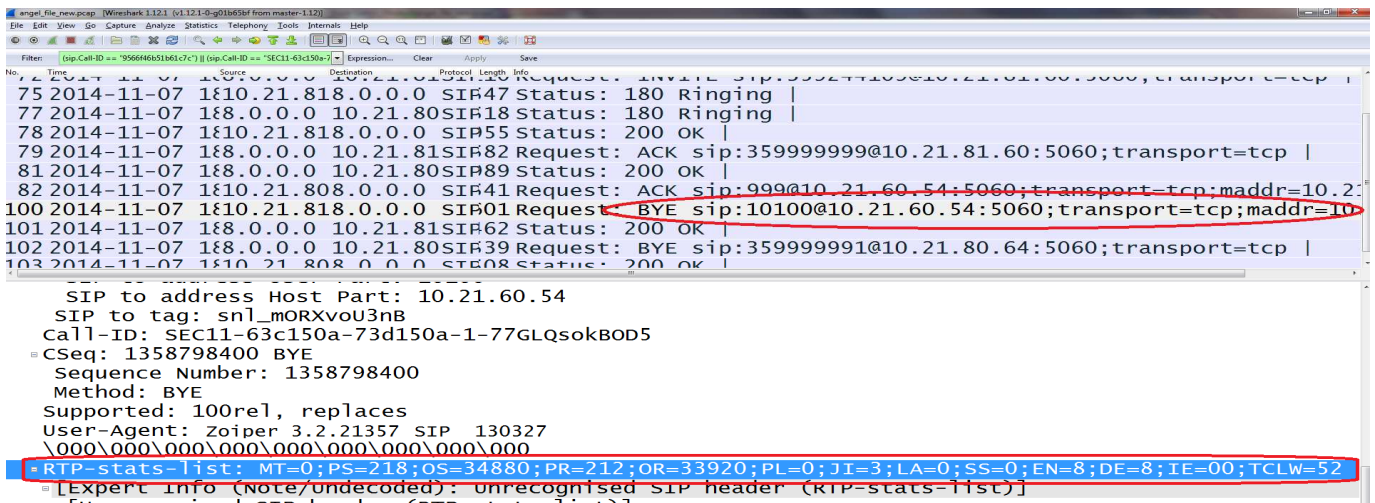
Fig. 1. BYE Message

Important and useful are the following measures:

The average proportion of packets lost during the measurement period. This definition is used in VQmon, RTCP XR and RFC3550 / RTCP.

The average proportion of packets discarded due to late arrival at the receive jitter buffer. This definition is used in VQmon and RTCP XR.
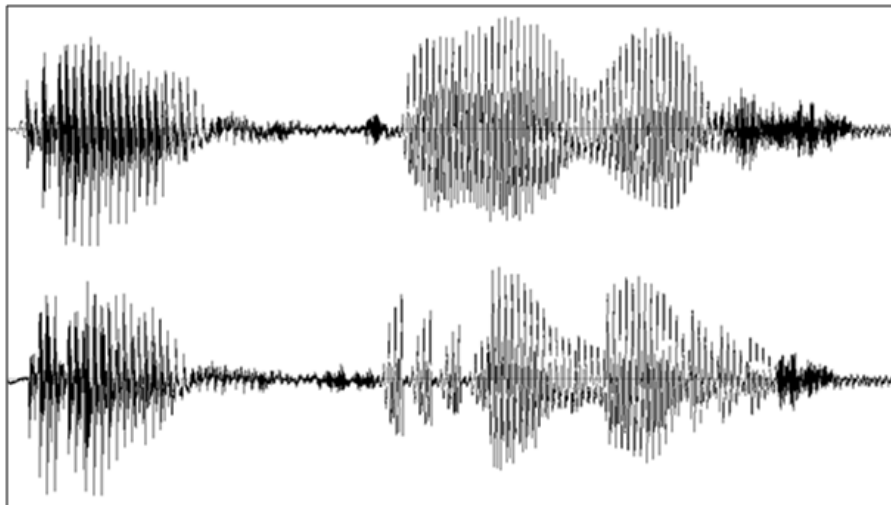
Fig. 2. Comparison between 5 and 20 ms Jitter

The average proportion of packets both lost and discarded during periods of low-density loss, and the average length of these periods. A "gap" is often defined as a time period during which density loss of less than 5 percent and packet loss isolated or widely dispersed. This definition is used in VQmon and RTCP XR.

VoIP receiving endpoints generally capture both packet loss and packet discard counter, so that you can determine for what type of problem you eliminate errors. Cisco phones include these statistics either in manual user interface on the phone set, or through their web interface. Avaya RTCP report these statistics, the suddenly requires a separate RTCP receiver to capture the statistics. As VoIP control software is updated, you may find these statistics available within the VoIP controller.
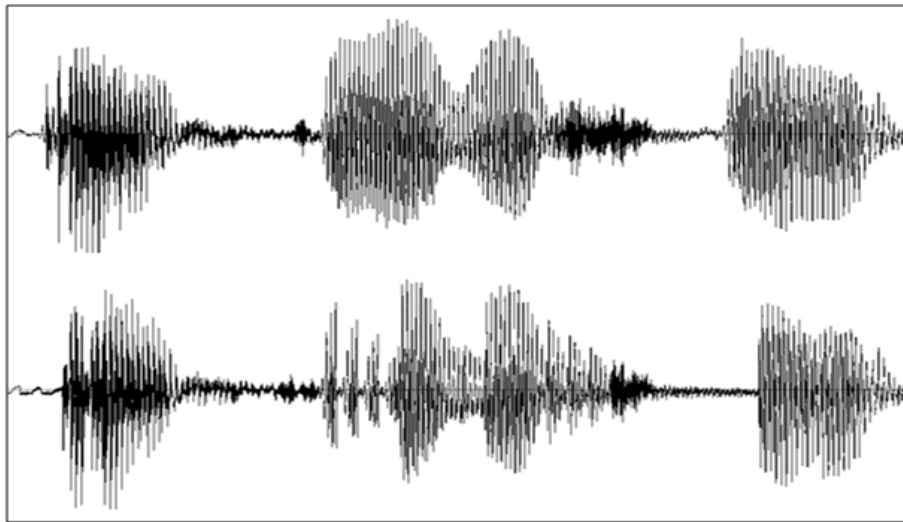
Fig. 3. Comparison between 2 % anf 40% Packet loss

The fourth variable is Codec. A codec converts the analog audio file into digital data packets and decodes them back at the receiver. Most codecs use compression of data at the same time in order to reduce the bandwidth required for transmission. Mostly used codecs are G.711noPLC, G.711PLC and G.729A.

III. PRACTICAL SOLUTION OF THE PROBLEM

A. Building the neural network

For construction of the neural networks, the application STATISTICA [6] is used. STATISTICA Automated Neural Networks (SANN) is one of the most modern and efficient neural networks application on the market. It offers numerous unique advantages to neural network experts and to new users.

Experts have a wide range of network types and training algorithms. New users are directed through the necessary procedures for creating neural networks via the Automated Network Search tool. There is the possibility to save the trained network in C code.

These are the real samples of voice quality in relation to any parameter of the VoIP speech quality (jitter, packet sent, lost packet and codec). A special equipment generates the simples, which allows the perception of the real voice quality at various jitter or lost packets.

After the construction and training of the neural network, all networks are tested. The best result are obtained by 6-10-1 MLP network with logistic activation function (Fig. 5).

To be sure, that this neural network solves the given task, it was tested by 150 results with known quality and input data (Fig. 4). We think that this number of tests should be enough because of the constant average of the deviation of the voice quality error. The error can be decreased with further train samples. Increasing the neuron number can also shrink the error.
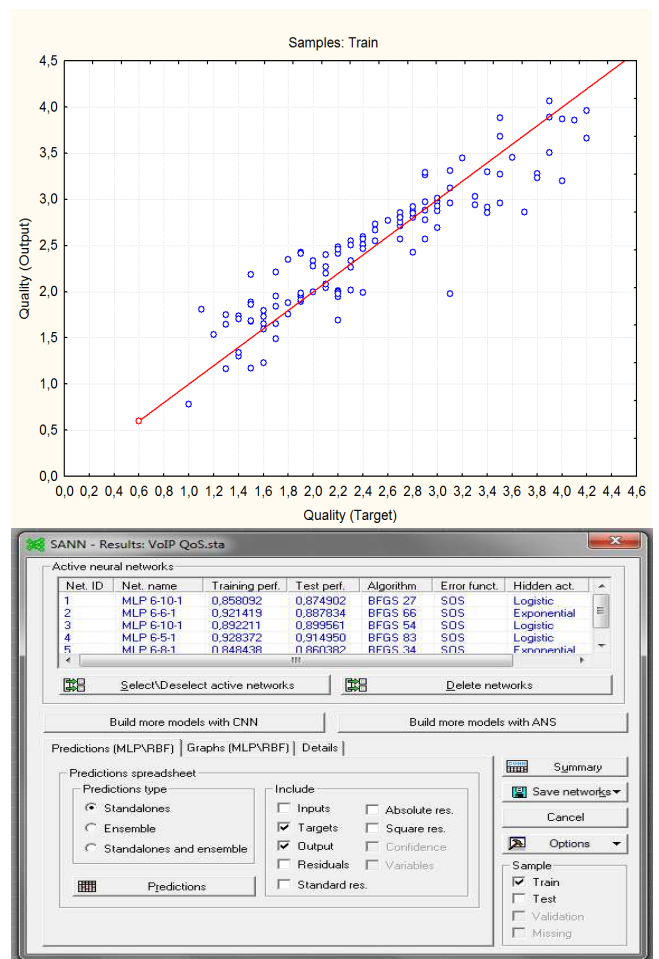
Fig 4. Quality (output) as a function of quality (target)

The trained network can be exported in C source code. This code can be inserted into the server later. If needed the network can be retrained with new data and the code recompiled.
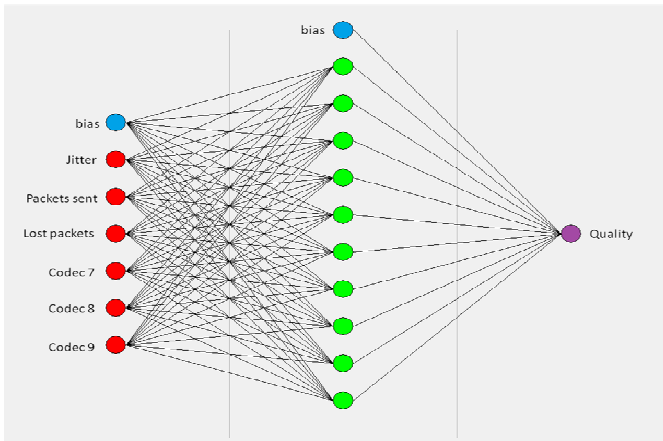
Fig 5. The used neural network

B. Results

All graphics for the representation of the network will be presented when using the best codes (G.711PLC) and constant number of sent packages. In Figure 6 is the 3D graphics speech quality depending on jitter and lost packets. On the x-axis is the jitter values are shown in milliseconds. The y-axis shows the variation of the lost packets. The third axis (z-axis) shows the speech-quality in MOS LQ. It is easy to notice how
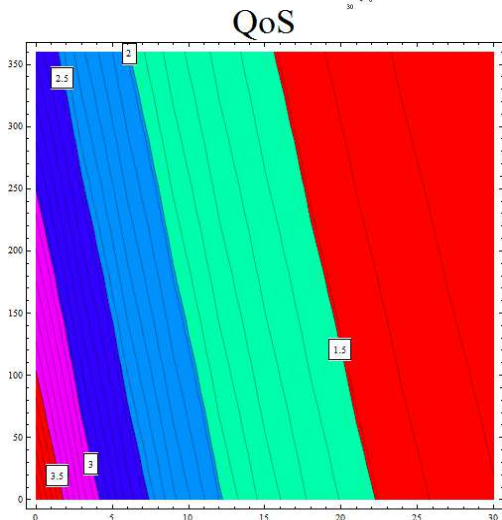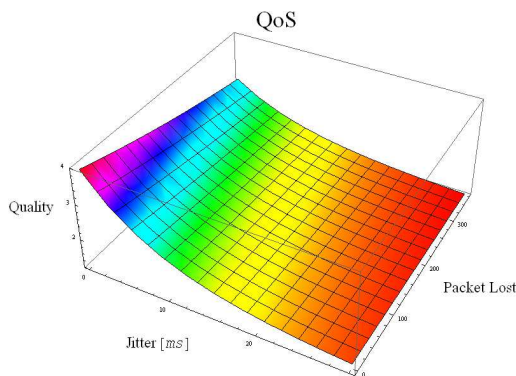
Fig 6. ANN of 500 sent Packets

with enlargement of jitter, the quality drops down. It is interesting that the lost packets does not affect the voice quality significantly.

The figure clearly shows that the parameters jitter and lost packets carry the vast information about call quality. Comparing the results obtained with the neural network, and those derived from the study of the subjective evaluation confirm this. Using BYE message receipt of digital expression of these two parameters is the least expensive way to obtain an objective assessment of quality. Slightest deviation from the subjecttive results are comparable with the confidence interval of the study.

IV. CONCLUSION

The proposed solution provides simple implementable QoS monitoring. As a result, this program is a cheap tool for QoS and VoIP network monitoring and for the VoIP network errors as well. This solution allows to predict the subjective user opinion of the voice quality only with the usage of statistics in the BYE-message and the corresponding OK message between the devices and the signaling server. The parameters considered by use of different codecs show, that the highest loss rate and the average loss burst size are provided by the packetization. In this study, a 0.874902 correlation coefficient between the prediction value and the result with the neural network as a subjective score is obtained,

V. FUTURE WORK

The use of a neural network, which assesses the voice quality subjectively, has a great advantage. The neural network module can easily improve the the quality of service by the use of the accumulated statistical data. In the updated module only the weights of the neurons are different from the original network. For update, only the weights in activating functions should be changed. After training the network, whose subjective evaluation is highly precious, corresponds to the customer evaluation.

REFERENCES

[1] A. Lakaniemi, J., "Subjective VoIP speech quality evaluation based on network measurements", Rosti, V. I. Raisanen IEEE ICC 2001;

[2] H. Koumaras, F. Liberal, L.Sun, "PQoS Assessment Methods for Multimedia Services", IGI Global publication, ISBN: 978-1-59904-820-8, July 2008.

[3] A. Lakaniemi, J., "Subjective VoIP speech quality evaluation based on network measurements", Rosti, V. I. Raisanen IEEE ICC 2001;

[4] ITU-T Recommendation P.800, "Methods for Subjective Determination of Transmission Quality", 1996;

[5] http://www.cisco.com/en/US/docs/ios-xml/ios/voice/cube_sip/configuration/15-2mt/voi-report-end-cal.html,CISCO, 09.2012

[6] http://www.statsoft.com/, 5.12.2014

Analysis of cryptographic protection of block cryptographic algorithms on the base of the theoretical digital stability

Ivan Ivanov [1], Rumen Arnaudov [2] and Stella Vetova [3]

Abstract – **the following report presents analysis of cryptographic protection of most used symmetric block cryptographic algorithms nowadays on the base of the theoretical digital stability.**

The theoretical digital stability is assessment reached on condition that the cryptographic algorithm is qualitative and firm to cryptanalysis accelerated attacks but can be attacked only by brute force attack.

Keywords – **cryptanalysis, cryptographic attack, cryptography, cryptographic protection, cryptographic algorithms**

I. INTRODUCTION

Cryptanalysis (from Greek κρυπτός — hidden and analysis) — a science for the methods used to reach the original meaning of data after it has been ciphered without the availability of secret information (key) necessary for this purpose. In most cases this concerns the key revealing. In nontechnical sense, cryptanalysis break on the cipher. The term is led by the American cryptographer W. F. Fridmann in 1920.

The results of the cryptanalysis of concrete cipher are called cryptographic attack on this cipher. The successful cryptographic attack fully discrediting the attackable cipher is called.

Cryptanalysis methods [1, 2, 3, 4, 5]
1. Classical cryptanalysis:
• Frequency analysis;
• Kasiski examination;
• Method index matches;
• Method of index mutual matches.
2. Cryptanalysis of symmetric algorithms:
• Differential cryptanalysis;
• Linear cryptanalysis;
• Integral cryptanalysis;
• Statistical cryptanalysis;

[1]Ivan Ivanov is with the College of Telecommunications and Posts of Sofia, 1 St. Mladenov Blvd, Sofia 1100, Bulgaria, E-mail: ivanivanov@hctp.acad.bg.
[2]Rumen Arnaudov is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: ra@tu-sofia.bg.
[3] Stella Vetova is with Bulgarian Academy of Sciences, Acad. G. Bonchev St. Sofia 1113, Bulgaria, E-mail: vetova.bas@gmail.com

3. Cryptanalysis of asymmetric algorithms:
• Decision tasks decomposition numbers of multipliers;
• Solvation of tasks with discrete logarithms;
4. Other methods:
• Attack "Birthday";
• Attack "Man in the middle";
• Brute force attack;

As far as the cryptographic protection realizes using block and stream cryptographic algorithms whose concrete realizations are simply defined by the interaction keys, the goal of the cryptanalysis reduces to revealing those interaction keys which are used in the cipher process[6].

Another assessment for the stability of the cryptographic protection exists. It is distinguished for a quantitative expression – theoretical digital stability. This assessment is important mainly for the theoretical estimation of the cryptographic algorithm but can be used by the user too to get information about the quality (grade) of the cryptographic algorithm (cryptographic equipment).

The theoretical digital stability is assessment reached on condition that the cryptographic algorithm is qualitative and firm to cryptanalysis accelerated attacks but can be attacked only by brute force attack [7].

In reality, the opponent does not have information only for the concrete key for interaction which is used in the cipher process. Therefore, the cryptanalysis is brought to search of the used key by the method of brute force that is through all the keys for interaction.

II. METHOD

Whilst in the cryptanalysis methods of the by brute force there is consecutively testing of the interaction keys from the set {Ki} with size N – key, for the assessment of the stability of the cryptographic algorithms the formula (1) for the mean time for interaction key revealing can be used:

$$\overline{T_{cal}} \approx \frac{Nn_{op}S_{min}}{6B}10^{-7}, [years], \qquad (1)$$

where nop is the number of computer operation necessary for 1bit information processing, Smin [bit] – minimum length of the glaring sample, and B [computer operation co/s] is the performance of the processing in the technical base for cryptanalysis.

The interaction key number N is defined by formula (2):

$$N = 2k, \; [\text{number.}], \qquad (2)$$

Where κ is the length of the basic key used by the cryptographic algorithm.

The experiment is accomplished over the most used symmetric block cryptographic algorithms nowadays such as: DES, K=64 bit, Triple DES, K=112Bit, IDEA, K=128 bit, AES, K=128 bit, AES, K=192 bit, AES, K=256 bit, Blowfish, K=32 bit, Blowfish, K=448 bit, RC5, K=128 bit, RC2, RC5, K=512 bit, RC2, RC5, K=1024 bit.

III. RESULTS

On the base of formula (2), is obtained (table 1, 2 and 3) the dependency of the theoretical digital stability $\overline{T_{ycm}}$, the interaction key number N in ncoSmin=10^2, 10^5 and 10^8 and different meanings of the performance of the B = 10^9 co/s; 10^{12} co/s; 10^{15} co/s; 10^{20} co/s; 10^{30} co/s.

Table I. Result in $n_{co}S_{min}=10^8$

N, num.	B, co/s	T_{cal}, years	N, num.	B, co/s	T_{cal}, years
$0,18*10^{20}$	10^9	$3*10^{10}$	$5,19*10^{33}$	10^9	$8,7*10^{24}$
	10^{12}	$3*10^{7}$		10^{12}	$8,7*10^{21}$
	10^{15}	$3*10^{4}$		10^{15}	$8,7*10^{18}$
	10^{20}	$3*10^{-1}$		10^{20}	$8,7*10^{13}$
	10^{30}	$3*10^{-11}$		10^{30}	$8,7*10^{3}$
$3,4*10^{38}$	10^9	$5,7*10^{29}$	$6,28*10^{57}$	10^9	$1.02*10^{49}$
	10^{12}	$5,7*10^{26}$		10^{12}	$1.02*10^{46}$
	10^{15}	$5,7*10^{23}$		10^{15}	$1.02*10^{43}$
	10^{20}	$5,7*10^{18}$		10^{20}	$1.02*10^{38}$
	10^{30}	$5,7*10^{8}$		10^{30}	$1.02*10^{28}$
$1,16*10^{77}$	10^9	$1,93*10^{68}$	$7,27*10^{134}$	10^9	$1,21*10^{126}$
	10^{12}	$1,93*10^{65}$		10^{12}	$1,21*10^{123}$
	10^{15}	$1,93*10^{62}$		10^{15}	$1,21*10^{120}$
	10^{20}	$1,93*10^{57}$		10^{20}	$1,21*10^{115}$
	10^{30}	$1,93*10^{47}$		10^{30}	$1,21*10^{105}$
$1,34*10^{154}$	10^9	$2.23*10^{145}$	$1,8*10^{308}$	10^9	$3*10^{299}$
	10^{12}	$2.23*10^{142}$		10^{12}	$3*10^{296}$
	10^{15}	$2.23*10^{139}$		10^{15}	$3*10^{293}$
	10^{20}	$2.23*10^{134}$		10^{20}	$3*10^{188}$
	10^{30}	$2.23*10^{124}$		10^{30}	$3*10^{178}$

Table II. Result in $n_{co}S_{min}=10^5$

N, num.	B, co/s	T_{cal}, years	N, num.	B, co/s	T_{cal}, years
$0,18*10^{20}$	10^9	$0,3*10^{8}$	$5,19*10^{33}$	10^9	$8,7*10^{21}$
	10^{12}	$0,3*10^{5}$		10^{12}	$8,7*10^{18}$
	10^{15}	$0,3*10^{2}$		10^{15}	$8,7*10^{15}$
	10^{20}	$0,3*10^{-3}$		10^{20}	$8,7*10^{10}$
	10^{30}	$0,3*10^{-13}$		10^{30}	$8,7*10$
$3,4*10^{38}$	10^9	$0,57*10^{27}$	$6,28*10^{57}$	10^9	$1.02*10^{46}$
	10^{12}	$0,57*10^{24}$		10^{12}	$1.02*10^{43}$
	10^{15}	$0,57*10^{21}$		10^{15}	$1.02*10^{40}$
	10^{20}	$0,57*10^{16}$		10^{20}	$1.02*10^{35}$
	10^{30}	$0,57*10^{6}$		10^{30}	$1.02*10^{25}$
$1,16*10^{77}$	10^9	$1,93*10^{65}$	$7,27*10^{134}$	10^9	$1,21*10^{123}$
	10^{12}	$1,93*10^{62}$		10^{12}	$1,21*10^{120}$
	10^{15}	$1,93*10^{59}$		10^{15}	$1,21*10^{117}$
	10^{20}	$1,93*10^{54}$		10^{20}	$1,21*10^{112}$
	10^{30}	$1,93*10^{44}$		10^{30}	$1,21*10^{102}$
$1,34*10^{154}$	10^9	$2.23*10^{142}$	$1,8*10^{308}$	10^9	$3*10^{296}$
	10^{12}	$2.23*10^{139}$		10^{12}	$3*10^{293}$
	10^{15}	$2.23*10^{136}$		10^{15}	$3*10^{290}$
	10^{20}	$2.23*10^{131}$		10^{20}	$3*10^{185}$
	10^{30}	$2.23*10^{121}$		10^{30}	$3*10^{175}$

Table III. Result in $n_{op}S_{min}=10^2$

N, num.	B, co/s	T_{cal}, years	N, num.	B, co/s	T_{cal}, years
$0,18*10^{20}$	10^9	$0,3*10^{5}$	$5,19*10^{33}$	10^9	$8,7*10^{18}$
	10^{12}	$0,3*10^{2}$		10^{12}	$8,7*10^{15}$
	10^{15}	$0,3*10^{-1}$		10^{15}	$8,7*10^{12}$
	10^{20}	$0,3*10^{-6}$		10^{20}	$8,7*10^{7}$
	10^{30}	$0,3*10^{-16}$		10^{30}	$8,7*10^{-3}$
$3,4*10^{38}$	10^9	$0,57*10^{24}$	$6,28*10^{57}$	10^9	$1.02*10^{43}$
	10^{12}	$0,57*10^{21}$		10^{12}	$1.02*10^{40}$
	10^{15}	$0,57*10^{18}$		10^{15}	$1.02*10^{37}$
	10^{20}	$0,57*10^{13}$		10^{20}	$1.02*10^{32}$
	10^{30}	$0,57*10^{3}$		10^{30}	$1.02*10^{22}$
$1,16*10^{77}$	10^9	$1,93*10^{62}$	$7,27*10^{134}$	10^9	$1,21*10^{120}$
	10^{12}	$1,93*10^{59}$		10^{12}	$1,21*10^{117}$
	10^{15}	$1,93*10^{56}$		10^{15}	$1,21*10^{114}$
	10^{20}	$1,93*10^{51}$		10^{20}	$1,21*10^{109}$
	10^{30}	$1,93*10^{41}$		10^{30}	$1,21*10^{99}$
$1,34*10^{154}$	10^9	$2.23*10^{139}$	$1,8*10^{308}$	10^9	$3*10^{296}$
	10^{12}	$2.23*10^{136}$		10^{12}	$3*10^{293}$
	10^{15}	$2.23*10^{133}$		10^{15}	$3*10^{290}$
	10^{20}	$2.23*10^{128}$		10^{20}	$3*10^{185}$
	10^{30}	$2.23*10^{118}$		10^{30}	$3*10^{175}$

On the base of formula (2), we reach the dependency between the basic cryptographic keys length K and the possible number of combinations for the same N.

In K=64 bit, $N = 2^{64} = 0,18*10^{20}$;
In K=112 bit, $N = 2^{112} = 5,19*10^{33}$;
In K=128 bit, $N = 2^{128} = 3,4*10^{38}$;
In K=192 bit, $N = 2^{192} = 6,28*10^{57}$;
In K=256 bit, $N = 2^{256} = 1,16*10^{77}$;
In K=448 bit, $N = 2^{448} = 7,27*10^{134}$;
In K=512 bit, $N = 2^{512} = 1,34*10^{154}$;
In K=1024 bit, $N = 2^{1024} = 1,8*10^{308}$;

Fig. 1, 2, and 3 illustrate the dependency of the theoretical digital stability $\overline{T_{ycm}}$, the interaction key number N in $n_{co}S_{min}=10^2$, 10^5 and 10^8 and different meanings of the performance of the $B = 10^9$ co/s; 10^{12} co/s; 10^{15} co/s; 10^{20} co/s; 10^{30} co/s.
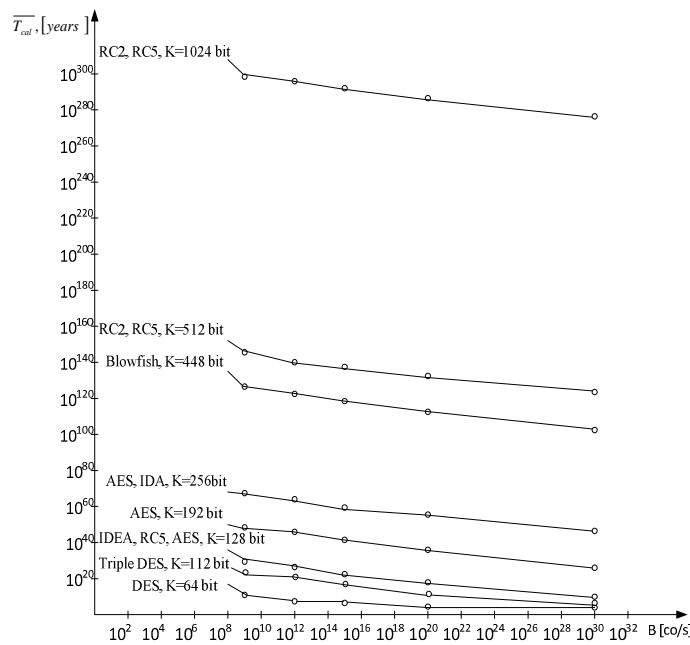
Fig. 1. Graphics of the dependency for the different algorithms and $n_{co}S_{min}=10^8$.

The graphics (fig.1) shows that the theoretical digital stability goes down increasing the performance of the process in the technical base for cryptoanalysis.

For DES algorithm, in K = 64 bit, and processing capacity of the modern computers (4÷5 GHz processors) the theoretical digital stability is low as for values $B = 10^{20}$ op/s and $B = 10^{30}$ op/s, it is respectively : $\overline{T_{cal}} = 0,3*10^{-11} = 95$ µs; $\overline{T_{cal}} = 0,3*10^{-1} = 946080$ s. = 11 days, it is insignificant.

For the cryptographic algorithms where K≥128 bits, guaranteed cryptographic protection is applied.
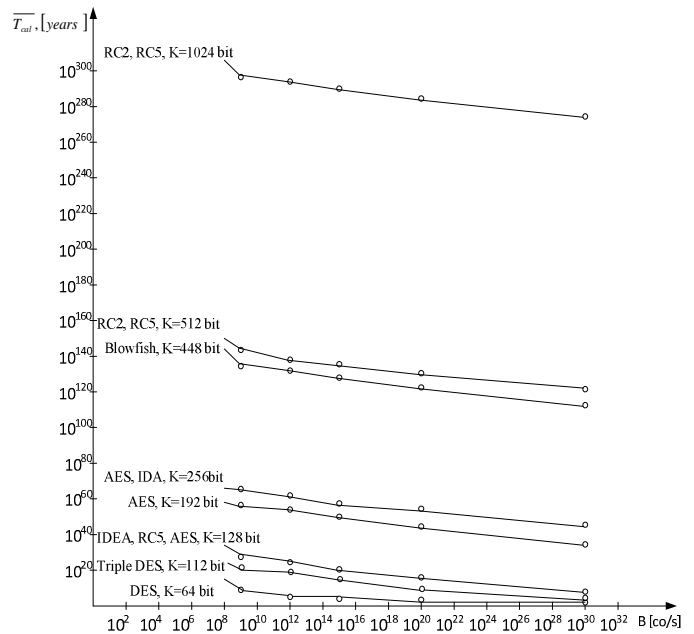
Fig. 2. Graphics of the dependency for the different algorithms and $n_{co}S_{min}=10^5$.

Fig. 2 depicts the dependency for the different algorithms and $n_{co}S_{min}=10^5$. It shows that for the DES algorithm in K = 64 bits, the theoretical digital stability is low as for values of $B = 10^{20}$ op/s and $B = 10^{30}$ op/s, it is respectively: $\overline{T_{cal}} = 0,3*10^{-13} = 0,95$ µs; $\overline{T_{cal}} = 0,3*10^{-3} = 9460,8$ s.= 2,63 hours, it is insignificant.

For the cryptographic algorithms where K≥128 bits, guaranteed cryptographic protection is applied.
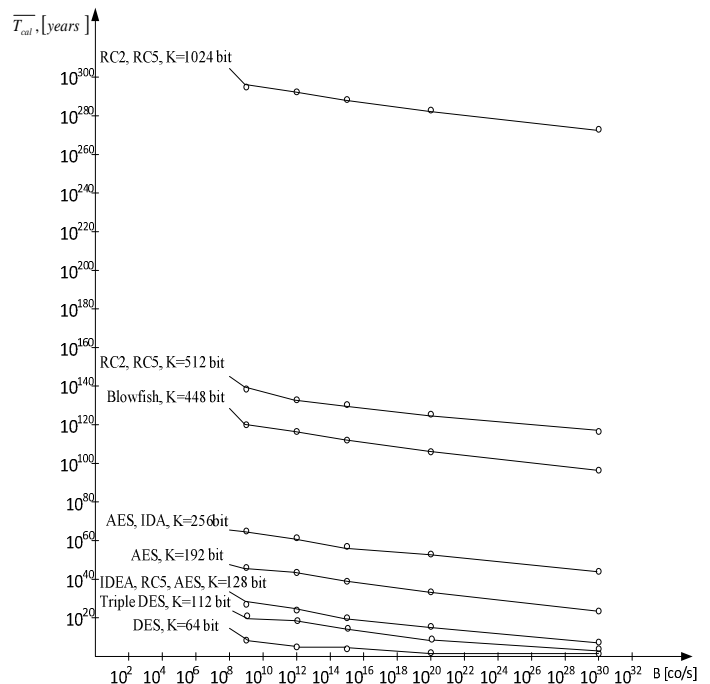
Fig. 3. Graphics of the dependency for the different algorithms and $n_{co}S_{min}=10^2$.

From Fig. 3 it is seen that for the DES algorithm in K = 64 bits, the theoretical digital stability is low as for values of B = 10^{15} op/s, B = 10^{20} op/s and B = 10^{30} op/s, it is respectively: $\overline{T_{cal}}$ = 0,3*10^{-16} = 0,95 ns; $\overline{T_{cal}}$ =0,3*10^{-6} = 9,46 s, $\overline{T_{cal}}$ = 0,3*10^{-1} = 946080 s. = 11 days, it is insignificant.

For the cryptographic algorithms where K$\geq$128 bits, guaranteed cryptographic protection is applied.

IV. CONCLUSION

The accomplished analysis is based on three different Smin [bit] – minimal lengths of the glaring sample: 10^2 bit, 10^5 bit и 10^8 bit.

On the base of the presented analysis and graphics, it is clearly seen that the theoretical digital stability of the tested cryptographic algorithms including the mean time of key revealing goes down increasing the performance of the processing in the technical base for cryptanalysis.

For the cryptographic algorithms using key length shorter than 128 bits (DES, Triple DES), the mean time for key revealing is very short whilst those algorithms using key length with or more than 128bits are noticeable for good and high guaranteed stability of cryptographic protection..

REFERENCES

[1] Bauer F. Decrypted secrets methods and maxims of cryptology, Springer, 2007.
[2] Joye M., Tunstall M. Fault analysis in cryptography, Springer, 2012.
[3] Katz J., Lindell Y. Introduction to Modern Cryptography, Second Edition (Chapman & Hall/CRC Cryptography and Network Security Series), CRC Press, 2014.
[4] Schneier B. Appliet Cryptography Protocols, Algorithms, and Source Code in C, Wiley, 2013.
[5] Ferguson N., Schneier B., Kohno T. Cryptography Engineering: Design Principles and Practical Applications, Wiley, 2010.
[6] Sokolov A. Shangin B. Information protection rzpredelenyh corporate networks and systems, DMK Press", M., 2002.
[7] Ivanov I. Laboratory experiments on security and protection of information and administration and protection of communication and computer networks, VU CTP, Sofia 2013.

ORAL SESSION
DIGITAL IMAGE PROCESSING

Comparison of Different Methods for the Estimation of Text Skew

Darko Brodić [1], Zoran N. Milivojević [2], Ivo R. Draganov [3] and Viša Tasić [4]

Abstract – **The paper analyzes different methods for the evaluation of the text skew. It incorporates the widely used vertical projection method as well as the moment based and log-polar cross-correlation method. The comparison among methods is based on the printed text samples. The test consists of the document image samples in the standard resolution of 300 dpi. The analyzed algorithms show different skew accuracy. However, the method with the smallest accuracy deviation proved the advantage over the other methods. Furthermore, this contributes to its robustness in applications.**

Keywords – **document image analysis, horizontal projection profiles, initial skew rate, moments, skew estimation.**

I. INTRODUCTION

The identification of the object skew in the image is one of the most important tasks in digital image processing. Accordingly, the text skew estimation is a key step in the document image analysis, too. Its existence could cause the optical character recognition system (OCR) failing. It is due to the system sensitivity to any skew appearance in the text. Hence, the text skew estimation represents the crucial step in OCR [1].

The printed text is characterized with the regularity in shape [1]. It means that the letters have the similar sizes and the distance between text lines are adequate. It enables that between line spacing is sufficient to split text lines. Furthermore, the orientation of the text lines is similar, which forms the uniform text skew. All above attributes represent relatively predicted characteristics. Hence, they simplify the procedure of the skew estimation for the printed text.

It should be noted that the text skew occurrence is unavoidable. It is an implication of the digitization process. Existing methods for the estimation of the text skew can be classified as [1]: (i) Projection profiles method, (ii) K-nearest neighbor clustering method, (iii) Hough transforms method, (iv) Fourier transformation method, (v) Cross-correlation method, and (vi) Other methods.

[1]Darko Brodić is with the Technical Faculty in Bor, University of Belgrade, Vojske Jugoslavije 12, 19210 Bor, Serbia, E-mail: dbrodic@tf.bor.ac.rs, mjevtic@tf.bor.ac.rs.
[2]Zoran N. Milivojević is with College of Applied Technical Sciences, Aleksandra Medvedeva 20, 18000 Niš, Serbia, E-mail: zoran.milivojevic@vtsnis.edu.rs.
[3]Ivo R. Draganov is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: idraganov@tu-sofia.bg
[3]Viša Tasić is with the Mining and Metalurgy Institute, Department of Industrial Informatics, Zeleni Bulevar 35, 19210 Bor, Serbia, E-mail: visa.tasic@irmbor.co.rs

In this paper, we decided to analyze widespread technique for text skew estimation based on a projection profiles calculation along with moment based method and newly proposed log-polar cross-correlation method. Furthermore, their estimation of the detected text skew is analyzed and compared.

Organization of this paper is as follows. Section 2 describes different methods for text skew detection. Section 3 defines the experiments. Section 4 presents and analyzes the results. Section 5 gives the conclusions.

II. PROPOSED METHODS

Below is presented each of tested algorithm for text skew detection. They are: (i) horizontal projection profiles, (ii) moment based algorithm, and (iii) log-polar cross-correlation algorithm.

Projection profile method

The horizontal projection profile method extracts features from the projection profiles of text lines. Horizontal profile gives the sum of the black pixels perpendicular to the x axis. It is represented by the vector P_h of size N defined by:

$$P_h[j] = \sum_{i=1}^{M} B(i,j) \qquad (1)$$

where **B** is the binary image featuring $i = 1, ..., M$ rows and $j = 1, ..., N$ columns. This vector is called a density histogram of horizontal black pixels.

Consequently, the valleys of the horizontal projection correspond to background areas of the binary image. Furthermore, the amplitude and the frequency of the projection are maximized when the skew of the text is zero. Based on this characteristic, the image is rotated through a full range of angles. Finding peak in the horizontal projection gives the estimated angle of the text skew [2]. This circumstance is shown in Fig. 1.

Moment based method

Moment defines the measure of the pixel distribution in the image. It identifies global image information concerning its contour. Moments of the binary image $B(i,j)$ are [3]:

$$m_{pq} = \sum_{i=1}^{N} \sum_{j=1}^{M} i^p j^q , \qquad (2)$$
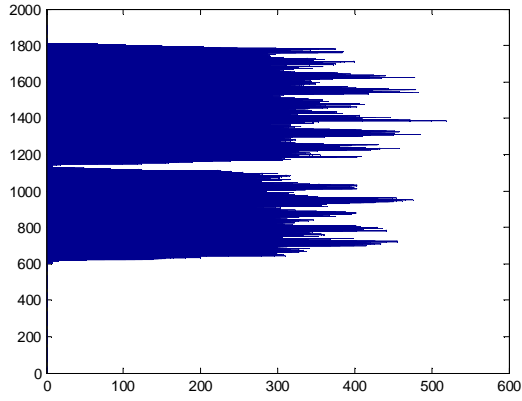
(a)
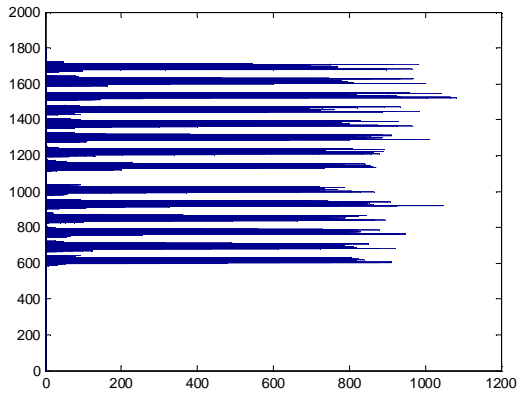
(b)

Fig. 1 Horizontal projection profiles method: (a) with skewed text, (b) with unskewed text

where p and $q = 0, 1, 2, 3, ..., n$, and n represent the order of the moment. From eq. (2) central moments μ_{pq} for binary image $B(i, j)$ can be calculated as:

$$m_{pq} = \sum_{i=1}^{N} \sum_{j=1}^{M} (i - \overline{x})^p (j - \overline{y})^q . \qquad (3)$$

Some of the important image features are obtained from the moments. One of them is the object orientation defined by the angle θ. It represents the angle between the object and the horizontal axis. It can be estimated by minimizing the function [4]:

$$S(q) = \sum_i \sum_j \left[(i - \overline{x})\cos(q) - (j - \overline{y})\sin(q) \right]^2 . \qquad (4)$$

The minimization gives the following values [3-4]:

$$q = \frac{1}{2}\arctan\left(\frac{2m_{11}}{m_{20} - m_{02}} \right). \qquad (5)$$

All these features characterize the separate object. Hence, this method has been utilized regularly for the printed text

skew evaluation. In that case, the skew evaluation of the each text object is similar or equal to all others.

Log-polar cross-correlation algorithm

The log-polar cross-correlation algorithm is based on the log-polar transformation. It emulates the principle of primate eye retina [5]. Each point in Cartesian space maps into equivalent point in the log-polar space. Firstly, the Cartesian space coordinates x and y are converted into polar coordinates radius r and angle θ. Their mapping is as follows [5]:

$$r = \sqrt{x^2 + y^2} , \qquad (6)$$

$$q = \arctan \frac{y}{x} . \qquad (7)$$

Furthermore, the log transformation is obtained as:

$$r = \ln r = \ln \sqrt{x^2 + y^2} . \qquad (8)$$

Hence, the (x, y) coordinates in Cartesian space are mapped into (ρ, θ) coordinates in log-polar space.

The log-polar transformation is nonlinear and non-uniform. Nonlinearity has been introduced by polar mapping, while non-uniform sampling is the result of logarithmic scaling [6].

Let consider the digital form of the transformation. For the binary image **B**, the center point of transformation is given as $B(m_c, n_c)$. The radius is assigned as R. It ensures that the maximum number of pixels is included within the reference circle of the conversion. Accordingly, the center of the circle is given as [7]:

$$m_c = \frac{M}{2} , \quad n_c = \frac{N}{2} . \qquad (9)$$

Furthermore, the image is converted into the polar coordinate system. This way, the binary image **B** has been transformed into the polar domain (r, θ) [6]:

$$r = \sqrt{(i - m_c)^2 + (j - n_c)^2} , \ 0 \le r \le R , \qquad (10)$$

$$q = \arctan\left(\frac{j - n_c}{i - m_c} \right), \ 0° \le q \le 360° , \qquad (11)$$

where $i = 1,…,M, j = 1, …, N$. The log-polar transformation is given as (ρ, θ), where ρ is given from eq. (8).

In the log-polar domain, the text image matrix **B** and referent object matrix **E** are marked as **BC** and **EC**. Cross-correlation shows a similarity measure between two images. It is [8]:

$$CC(q) = C_{coeff}\left(\mathbf{BC}, circshift(\mathbf{EC}, q) \right), \qquad (12)$$

where **ECS** is circshift(**EC**, θ) and C_{coeff} (**BC**, **ECS**) ia given as [8]:

$$C_{coeff} = \frac{\sum_r \sum_q (BC_{rq} - \overline{BC})(ECS_{rq} - \overline{ECS})}{\sqrt{\left(\sum_r \sum_q (BC_{rq} - \overline{BC})^2\right)\left(\sum_r \sum_q (ECS_{rq} - \overline{ECS})^2\right)}} . \quad (13)$$

If the images are more alike, then cross-correlation function $CC(\theta)$ will tend to approach 1.

The identification of the rotation in the image space is a complex task. In the log-polar space, the rotation is mapped into translation. The translation in the direction of one axis is an easy task to solve. Suppose that a referent object is rotated in the space domain. If it is cross-correlated with the text image for the different angles, then it will be read out as the translation in the log-polar space. The first objective is the selection of the correct referent object as a template. The right choice can be the ellipse. It is a suitable object because it can overlap text efficiently. However, the ellipse has to be normalized according to the text image dimension. Furthermore, the ellipse is split into left and right half part from the center point of the transformation. This way, the parts of the ellipse are matched with the text image by the cross-correlation. Hence, they establish the left and right skew estimation. Unlike the other methods, the log polar transformation identifies two skews: left and the right one.

III. EXPERIMENTS

The primary goal of the experiment is the evaluation of the algorithm used for the text skew estimation. Hence, it evaluates the algorithm's performance in the skew tracking domain. Experiments for the printed text consist mostly of the synthetic datasets. The most important one is the test of the single line printed text. This test consists of the single line printed text rotated for the angle θ from 0° to 60° in the steps of 5° around x-axis [9]. Typical text sample is shown in Fig. 2.
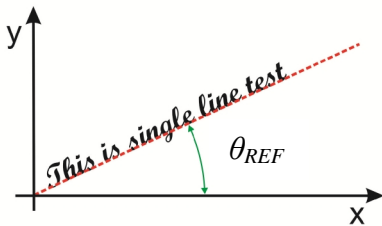
Fig. 2 Skew angle θ_A vs. referent angle θ_{REF}

For the evaluation of the method robustness, all text samples are given in the standard resolution of 300 dpi. The results are evaluated by the absolute, i.e. error. Absolute deviation is given as:

$$\Delta q_A = |q_{REF} - q_A|, \quad (14)$$

where θ_{REF} is the reference skew of the input text sample, while θ_A is the skew of the text sample obtained with tested algorithm.

IV. RESULTS AND DISCUSSION

The test results for the different algorithms are given in Table I.

TABLE I
ABSOLUTE SKEW DEVIATION FOR DIFFERENT METHODS

Angle range	Text skew estimation methods		
	Horizontal Projection method	Moment based method	Initial Skew rate method
θ_{REF} (°)	$\Delta\theta_{A,VPA}$	$\Delta\theta_{A,MA}$	$\Delta\theta_{A,LCCA}$
0	2.5	0.0000	0
5	0.1	0.0122	0
10	0	0.0089	0
15	0	0.0331	0
20	1.8	0.0435	0
25	0	0.0377	0
30	0	0.0417	0
35	0	0.0554	0
40	0	0.0248	0
45	0	0.0027	0
50	0	0.0175	1
55	0	0.0330	1
60	0	0.0545	1

The graph of the skew angle obtained by tested algorithms θ_A vs. reference angle θ_{REF} is shown in Fig. 3.
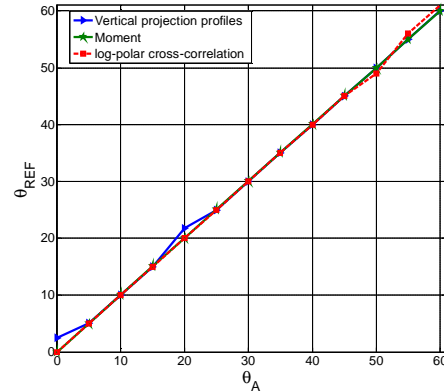
Fig. 3 Skew angle θ_A vs. referent angle θ_{REF}

Presented graph describes the response of the algorithm. If the algorithm has a correct response, then the graph will be linear. This means that the algorithm adequately identifies the text skew angle.

The vertical projection algorithm has the largest deviation (Fig. 3 - See blue line for reference). Log-polar cross-correlation algorithm has a linear response for the angles up to

45°. However, higher values of the angle θ_{REF} contributes to the nonlinearity (Fig. 3 - See red line for reference). In contrary, the moment based algorithm has uniform and linear response with a very small deviation in the whole range of angles. Hence, this method identifies the text skew angle with the smallest deviation. Fig. 4 shows the comparison of the absolute deviation results.
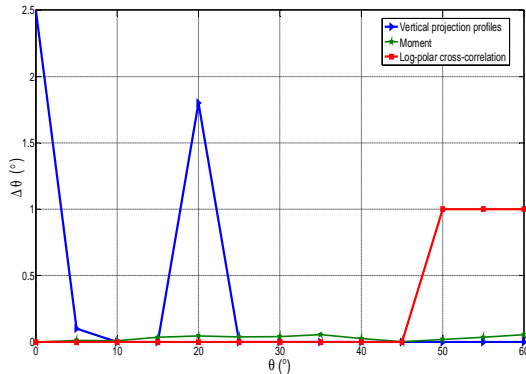
Fig. 4 Skew angle θ_A vs. reference angle θ_{REF}

The horizontal projection algorithm has the worst result. Its deviation is up to 2.5°. From the above, the log-polar cross-correlation algorithm has no deviation for the angles up to 45°. For the larger angle, it is up to 1°. Moment based algorithm has the uniform absolute deviation of up to 0.055°. This result is characterized as a pretty correct one. Summarized results for the tested algorithms are given in Table 2.

TABLE II
SUMMARY OF THE OBTAINED RESULTS FOR DIFFERENT METHODS

Angle range	Text skew estimation methods		
$0° \leq \theta \leq 60°$	Horizontal Projection method	Moment based method	Initial Skew rate method
$\Delta\beta$	< 2.5°	< 0.05°	< 1°

Taking into account the results from Tables 1-2, the moment based algorithm has the best result. It has a uniform absolute deviation and *RLHR*. Log-polar cross-correlation method is a promising one. However, it needs some adaptation. It is especially true for the angles higher than 45°. The horizontal projection algorithm has the worst results. Hence, it needs a serious adaptation.

However, tested algorithms have some additional obstacles in the application. Moment based algorithm has sensitivity to the noise. Hence, the additional cleaning of the image is the indispensable. Log-polar transformation is a very sensitive to the determination of the center point of transformation. If it is incorrect, then transformation will fail. Hence, it needs a special study on criteria for the center point determination. Vertical projection algorithm, in its initial form, has a problem with the correct text skew identification. In addition, its application is an easy one. Compared to the other methods, it is the slowest and error prone method. From above facts, the

moment based algorithm is the best choice for the correct text skew estimation of the binary objects.

V. CONCLUSIONS

This paper gives the critical analysis, and comparison between different text skew estimation techniques, which include horizontal projection, moment based and log-polar cross-correlation method. All methods are tested with the document image samples in the standard resolution of 300 dpi. Analyzed algorithms show a different skew accuracy. The horizontal projection algorithm shows an obvious error prone. The log-polar cross-correlation algorithm has a satisfactory correctness. However, it needs some adaptation as well as a clear criteria for determination of transformation center point. Though it has obstacles, the moment based method has a clear advantage over the other methods. The most important is that it estimates correctly the text skew. Furthermore, it is computationally non-intensive method.

ACKNOWLEDGEMENT

REFERENCES

[1] A. Amin, S. Wu, "Robust skew detection in mixed text/graphics documents," Proceedings of 8th International Conference on Document Analysis and Recognition, pp.247–251, 2005.

[2] R. Manmatha, N. Srimal, "Scale space technique for word segmentation in handwritten manuscripts," Proceedings of 2nd International Conference on Scale Space Theories in Computer Vision, pp.22–33, 1999.

[3] P. Shivakumara, G. H. Kumar, H. S. Varsha, S. Rekha, M. R. Rashmi Nayaka, "A new moments based skew estimation technique using pixels in the word for binary document images," Proceedings of 8th International Conference on Document Analysis and Recognition, pp.151–156, 2005.

[4] G. Kapogiannopoulos, N. Kalouptsidis, "A fast high precision algorithm for the estimation of skew angle using moments," Proceedings of Signal Processing, Pattern Recognition, and Applications, pp.370–143, 2002.

[5] V. J. Traver, A. Bernardino, "A review of log-polar imaging for visual perception in robotics," Robotics and Autonomous Systems, vol.58, no.4, pp.378–398, 2010.

[6] S. Zokai, G. Wolberg, "Image registration using log-polar mappings for recovery of large-scale similarity and projective transformations," IEEE Transaction on Image Processing, vol.14, no.10, pp.1422–1434, 2005.

[7] M. K. Bhowmik, et all., "Classification of log-polar-visual eigenfaces using multilayer perceptron," Proceedings of 2nd International Conference on Soft computing, pp.12–23,2008.

[8] D. Brodić, Z. Milivojević, "Log-polar transformation as a tool for text skew estimation," Electronics and Electrical Engineering, vol.19, no.2, pp.61–64, 2013.

[9] D. Brodić, D. R. Milivojević, Z. N. Milivojević, "Basic test framework for the evaluation of text line segmentation and text parameter extraction," Sensors, vol.10, no.5, pp.5263–5279, 2010.

Testing Oculus Rift Virtual Reality Headset Applicability to Medical Assistive Systems

Ivo R. Draganov[1], Nikolay N. Neshov[2], Darko Brodic[3]

Abstract – **In this paper are shown results from initial testing of the Oculus Rift virtual reality headset in relation to basic movements of the human head. Number of subjects were given series of instructions through the virtual environment with a moving cursor to follow from left-to-right, top-to-bottom and from horizontal-to-vertical-to-horizontal position of their heads. Mean and deviation values are obtained for rotation angles in space by the embedded sensors which show the possibilities of using the headset in some medical assistive systems.**

Keywords – **Oculus Rift, Virtual Reality, Medical Assistive Systems**

I. INTRODUCTION

Virtual reality opens up a full new frontier in assisting motor disabled patients during both their rehabilitation and implementing ordinary duties. Although not intended for professional purposes a number of companies, such as Canon, Sony, Motorola, Philips, Olympus, Epson and others, presented to the market low priced head mounted displays for virtual reality at relatively low price. In this study one of the most popular such headsets is used for probing its applicability in applications for medical assistive systesms.

In [1] Reinke et al. describe a newly developed prototype of virtually reality based system for investigation of peripheral vestibular disorders. Considering mainly the spatial orientation of the human head to the gravitational axis of the Earth which is crucial for proper human posture, gait and large amount of motor activities they combine the visual and vestibular information in perception along the vertical to estimate the type and degree of a vestibular disorder. The proposed system consists of Oculus Rift display and dedicated software where a certain object is presented to the patient and then he/she is asked to move his/her head to a position where the object appears vertical. Medical personnel can select the type of target object from a database and gets as a feedback the relative position of the head at end point creating a record in another database for further examination and possibly comparison over a long period of time with results from previous tests. The authors consider all target objects to retain tolerably good quality along all tests with the different movements of the head and the latency of the sensor to be adequately low preserving the desired accuracy.

Wood [2] proposed a model for more advanced system incorporating virtual reality in the form of gaming chair. His goal was to achieve motion based on events from control software and user input along 3 axis of movement. There is a resemblance with flight simulators in functionality about this system but at much more lower cost. A syntactic feedback was provided to the user by Oculus Rift as additional feature in comparison to already developed such systems while preserving safety at first place and assuring precision, longevity, usability, and compactness.

Another practical solution in a simpler form for distance control was developed by Burström et al. [3] called Desktopulus. It is a system putting together the Oculus Rift and Leap Motion with software tools for building up a virtual workspace for office needs. Embedding natural gestures by the user another degree of freedom was added to the user. On a virtual desktop environment are placed a set of different tools as in traditional desktops which can be manipulated by grab, move and circle movements. High CPU power consumption was reported for smooth operation of the system. Hence some optimization procedures need to be further applied in order to make the system affordable to the mass user with lower hardware requirements.

Another application of the Oculus Rift was presented by Paolantonio et al. in [4]. It is a 3D virtual representation of drones' flights in web browsers using globes visualization technology Cesium.js. Processing telemetry data, digital elevation models, georeferenced images and mosaics by the GDAL software and OGC web services (WMS, WFS). The user can virtually feel the flight. The proposed system use two types of data storage – a raster and a vector one, first of which located directly over a file system and the second over a PostGIS. The web map service exploits mosaic and pyramid representations. The first one takes a set of georeferenced images focusing on an area of interest forming a new continuous scene. The second one is based on collection of multiple mosaics each one of which is connected to different zoom level. The input telemetry data is processed in three steps: shapefile creation leading to a layer of points, database upload as a table to vector data storage, and finally exposing the database to the server. The creation of 3D globes and 2D maps in the web browser is done by the JavaScript library Cesium.js. It uses WebGL with hardware-accelerated graphics and is efficient for dynamic-data visualization. When presenting the scene to the user it's taking the advantage of the Cesium Sensor plugin. A pyramidal cone is generated pointing to the current telemetry point in time where the projection angles are of 50° horizontally and 70° vertically. They are selected equal to the design parameters of the Canon PowerShot S110 camera most frequently used in these applications. Viewing the scenes by the Oculus Rift it was possible to move by user selection up/down, forward/backward, right/left and also to duplicate movement's

[1,2]Ivo R. Draganov and Nikolay N. Neshov are with the Faculty of Telecommunications, 8 Kliment Ohridski Blvd., 1000 Sofia, Bulgaria, E-mail: idraganov@tu-sofia.bg, nneshov@tu-sofia.bg
[3]Darko Brodic is with the University of Belgrade, Technical Faculty in Bor, V.J. 12, 19210 Bor, Serbia, E-mail: dbrodic@tf.bor.ac.rs

velocity while looking straight to the horizon or to overlook the surface from above.

Some other problems often met with the head mounted displays are the focus cues, accommodation and retinal-image blur which were addressed by Konrad in his study in [5]. Both the object space methods where the depth of field effects are made straight into the rendering pipeline and the image space methods where all images are rendered at focus and then by a depth map undergo partial blur are considered by the study. Since the latency is the most important factor in these systems image space methods being more time efficient at the loss of some accuracy have higher priority for improvement. Konrad propose the linear filtering by Gaussian function for adding depth of field. The filter is spatially variant with depth-dependent point spread function and filtering time is proportional to the size of it. Separability of the Gaussian filter is considered as a way of lowering execution time. Another solution suggested by Konrad is the use of adaptive bilateral filter with weights proportional to depth values for the neighboring pixels. When a foreground is being captured in focus the blurred background will not occlude a portion of it blurring it as well. Since the bilateral filter preserves edges intensity leakage is reduced and results are better than the Gaussian filter. Execution time in this case is higher. The last approach proposed by Konrad is based on natural heat diffusion. As the time passes by uneven heat distribution over a conductive object tends to diffuse to more even one which has a similarity with image blur. Using the already developed mathematical apparatus consisting of differential equations for the heat diffusion it is possible to easily model the blur process. Computational time is highest for this approach but results are most promising.

In [6] it is given an extensive comparison about the capabilities between the Oculus Rift DK2 and Project Morpheus (PM) head mounted sets by Goradia et al. While the Oculus Rift is supported by PCs and mobile devices the PM is intended to work with PlayStation console only. The first set has 7" panel size against the 5" of the latter. Both sets have native resolution of 1920pixels along the horizontal and 1080 pixels along the vertical which gives resolution per eye of about 960x1080 pixels. The Oculus Rift DK2 is made by using OLED technology while the PM uses LCD. The maximum refresh rate of the PM is less than 2 ms typical for the Oculus Rift. Apart from the gyroscope and accelerometer the Oculus Rift possesses also a magnetometer. The inertial update rate for both units is 1000 Hz and the field of view of the PM is 90° using also the PlayStation camera - with 10° less than the Oculus Rift's equipped with a near infrared CMOS camera. The outputs are 1USB 2.0 and 1 HDMI where only the PM has 1 headphone output intended for 3D audio. Both sets are ready to be used in industrial type applications with the appropriate software as the one described in [7].

In the following Section II a prototype of a new medical assistive system for motor disabled patients is proposed with a protocol for testing the Oculus Rift inside it. Then in Section III some experimental results are presented from this testing followed by a conclusion in Section IV.

II. MEDICAL ASSISTIVE SYSTEM FOR MOTOR DISABLED PATIENTS AND EXPERIMENTAL SETUP

The medical system proposed in this paper is intended for assistance of patients with fully immobilized lower limbs and partly lively upper limbs. The main purpose of the system is to help the motor disabled patient with rehabilitation procedures where a series of training tasks should be accomplished by both hands. Each hand must be equipped with a sensor glove and the subject must wear the Oculus Rift. The exercises that need to be performed are stored in an indexed form in a database. Using a scene composition software module ghostly images of the hands of the patient are virtually generated inside a full training scene and then passed along to the virtual reality render for viewing through the Oculus Rift. Trying to repeat each movement of the "target hands" with his/her own the disabled person produces head's and hands' motions detected by the sensor gloves (using sensor gloves controller) and the Oculus Rift (through the accelerometer, gyroscope, and magnetometer) which are analyzed through another software module called movement analysis. Either at the same place or from a distance by remote connection a specialist on rehabilitation can get in real time or off-line the results of the exercise program. They are presented in the form of spatial differences by position at given time markers for both hands separately considering their volumetric properties.

The experimental setup includes a healthy individual (subject) that was presented a test pattern (marker) corresponding to the higher human eye resolution both horizontally and vertically rather than diagonally. He needed to track its movement first from Upper-Left to Down-Right position and then from Upper-Right to Down-Left position for a period of 11 sec. For each experiment the rotation angles (Fig. 1) around the three coordinate axes were measured for 10 repetitions in total.
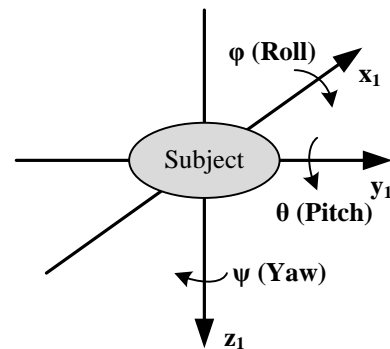
Fig. 1. Reference coordinate system for the subject

The resulting position and orientation from each movement of the head could be represented by a series of vectors with the following structure – $\{[t_0, x_0, y_0, z_0], [t_1, x_1, y_1, z_1], …, [t_k, x_k, y_k, z_k], …, [t_N, x_N, y_N, z_N]\}$ where t_k is the current moment of time for estimating the position and $\Delta t = t_k - t_{k-1} = 1$ sec for $N = 11$ sec in this case for the position. The orientation is given by $\{[t_0, q_{00}, q_{10}, q_{20}, q_{30}], [t_1, q_{01}, q_{11}, q_{21}, q_{31}], …, [t_k, q_{0k}, q_{1k}, q_{2k}, q_{3k}], …, [t_N, q_{0N}, q_{1N}, q_{2N}, q_{3N}]\}$ where $q_{0k}, q_{1k}, q_{2k}, q_{3k}$ are the components of the quaternion Q found from the roll (θ), pitch (φ), and yaw (ψ) as suggested in

[6]. It could be found in general form for arbitrary spatial rotation by considering at first the yaw rotation (Fig. 2) according to (1):

$$R_1^{s1} = \begin{bmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix}. \qquad (1)$$
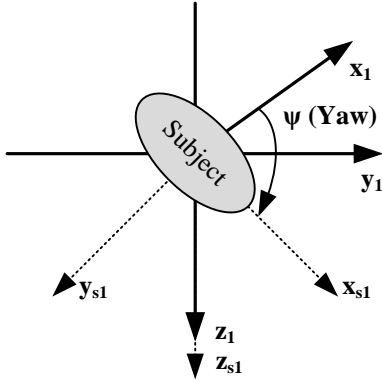
Fig. 2. Yaw rotation component of the subject

Taking both yaw and pitch rotation together (Fig. 3) a more detailed transform is obtained using (2):

$$R_{s1}^{s2}(\theta) = \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix}, \qquad (2)$$

to the form of:

$$R_1^{s2}(\theta,\psi) = R_{s1}^{s2}(\theta) R_1^{s1}(\psi). \qquad (3)$$
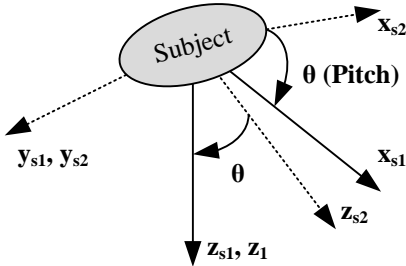
Fig. 3. Simultaneous yaw and pitch rotation of the subject

Finally combining the roll rotation based on (4) transformation matrix:

$$R_{s2}^2(\varphi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & \sin\varphi \\ 0 & -\sin\varphi & \cos\varphi \end{bmatrix}, \qquad (4)$$

a full description of the whole rotation (Fig. 4) could be written in general form according to:

$$R_1^2(\varphi,\theta,\psi) = R_{s2}^2(\varphi) R_{s1}^{s2}(\theta) R_1^{s1}(\psi). \qquad (5)$$
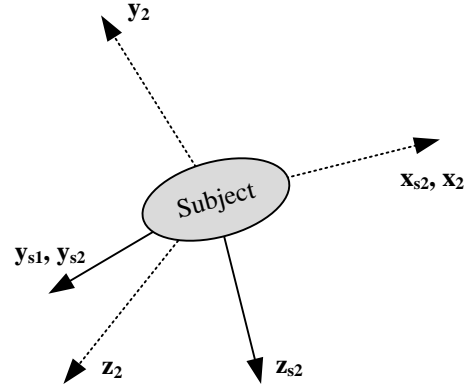
Fig. 4. Yaw, pitch and roll rotations combined

III. EXPERIMENTAL RESULTS

The average values of the pitch, yaw and roll for moving from Upper-Left to Down-Right position of the head are given in Fig. 5.
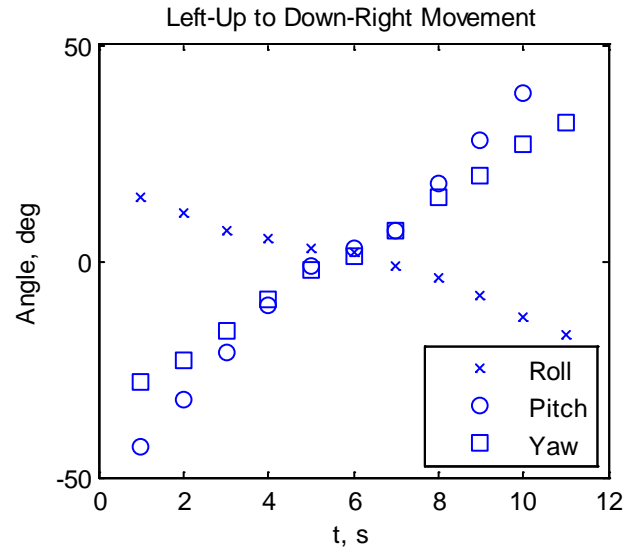
Fig. 5. Pitch, yaw and roll change when performing Upper-Left to Down-Right motion

The analysis of the angles change when performing the Upper-Left to Down-Right movement about their average magnitudes reveal most increase for the pitch (Fig. 5). It starts from -43° and ends up at 50° changing relatively smoothly. It is followed by yaw change from -28° up to 32° and with least decrease is the roll – from 15° down to -17°. During the medium time intervals the diversion from linearly changing values was around 2-3° at most cases.

Somewhat the opposite is the situation with all three angles when performing Upper-Left to Down-Right movement. Here the changes are also relatively smooth with pitch increasing from -47° to 52°, followed by decreasing of yaw from 27° to -33 and with smallest deviation is the roll – from -16° to 14°.
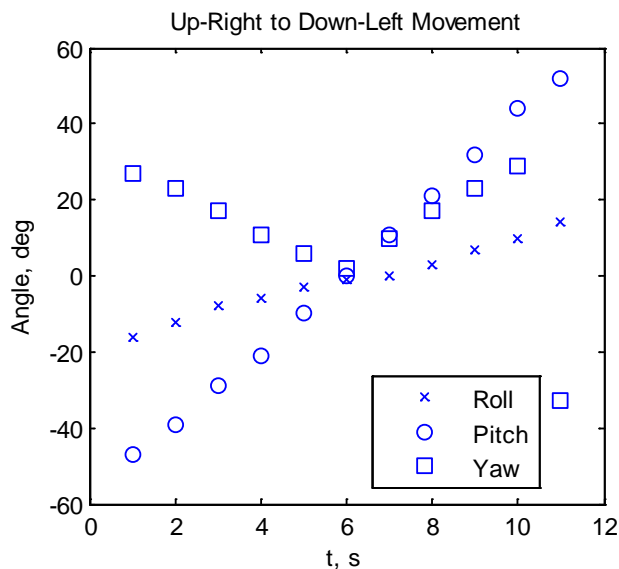
Fig. 6. Pitch, yaw and roll change when performing Upper-Right to Down-Left motion

All these results reveal that the movement analysis module should not be implemented in a straight-forward manner. All deviations for each angle should be carefully estimated for every patient at the time of initial calibration of the system. In order to have objective assessment of the results from therapy from the medical personnel all future deviations should be compared to initial ones and proper conclusions made on that basis.

IV. CONCLUSION

In this paper a testing of the Oculus Rift virtual reality headset was presented based on analysis of the Upper-Left to Down-Right and Upper-Right to Down-Left motion of a human head to evaluate the precision and time efficiency in relation to its applicability to a newly proposed medical assistive systems. The results are promising which provides a reliable displaying mean at low cost for virtual involvement of patients during their rehabilitation.

ACKNOWLEDGEMENT

REFERENCES

[1] W. Reinke, V. Marozas, I. Uloziene, M. Šileikaite, R. Liutkevičiene, and A. Lukoševičius, "A prototype of virtual reality based system for investigation of peripheral vestibular disorders", In Proc. of the Conference "Biomedical Engineering", vol. 18, no. 1, pp. 38-41, 2014.

[2] S. Wood, "Oculus Rift gaming chair", *Instructor*, vol. 12, no. 3, 2014, available at https://redmine.lcseecloud.net/attachments/download/790/Design%20Proposal.docx .

[3] J. Burström, K. Kullbrandt, Z. Juto, and O. Wahlström, "Desktopulus", available at http://pervasivecomputing.se/M7012E_2014/projects/Deskotoplus/Desktopulus-report.pdf, May 2015.

[4] M. Paolantonio, C. Fernández, M. Latorre, A. García, "3D virtual representation of drones' flights through Cesium.js and Oculus Rift", *In Proc. of the 9as Jornadas de Sig Libre,* Girona, Spain, March 26-27, 2015, available at http://www.sigte.udg.edu/jornadassiglibre/wp-content/uploads/2014/09/JSLart_Cesium_DEIMOS.pdf .

[5] R. Konrad, "Perceptually accurate depth of field rendering", Stanford, 2015, available at: http://web.stanford.edu/class/ee368/Project_Spring_1415/Proposals/Konrad.pdf .

[6] I. Goradia, J. Doshi, and L. Kurup, "A Review Paper on Oculus Rift & Project Morpheus", International Journal of Current Engineering and Technology, vol.4, no.5, pp. 3196-3200, Oct. 2014.

[7] V. Georgieva, P. Petrov, "An Approach for Position Detection of Industrial Objects", Proceedings of ICEST'13, Vol. 2 pp.581-584, 2013.

Facial Expression Recognition Based on Constrained Local Models and Support Vector Machines

Nikolay Neshov[1], Ivo Draganov[2], Agata Manolova[3]

Abstract – **This paper presents a face expression recognition algorithm using Constrained Local Model (CLM). CLM is facial alignment method that is based on Active Shape Models (ASM) and Active Appearance Models (AAM). It takes the advantages of both of them and gains high accuracy. To distinguish different expression states, we use CLM model parameters that describe shape deformation in a compact form. These parameters form feature vectors for training Kernel Support Vector Machine (KSVM) classifier. The experimental results over Cohn-Kanade Extended Facial Expression (CK+) database show improvement of the recognition rate in comparison to some existing methods, suggested by other authors.**

Keywords – **Constrained Local Model (CLM), Support Vector Machines (SVM), Expression Recognition (ER), Emotion Estimation, OpenIMAJ**

I. INTRODUCTION

Facial expression analysis reflects the emotional state of people and hence provides useful information about their personality and psychopathology. The implementation of ER system can play an important role wherever humans interact with machines. This would be of help in various vital purposes, such as security, entertainment, health care, robotics, society, etc. For example, in health care, it becomes useful to integrate a facial ER module in existing surveillance system, which constantly observes patients and analyses their emotional states. Once pain presence is detected (which causes deformations in facial expression) the system would automatically notifies the doctor.

In this paper we describe a facial expression recognition system that is based on Constrained Local Model (CLM) [1] approach. The last one is built by hybridization of two parametric modelling techniques: Active Shape Models [2] and Active Appearance Models [3]. In order to develop our system we utilized the CLM implementation available in OpenIMAJ [4] (Open Intelligent Multimedia Analysis for Java) which is released as an open source under the BSD license. This model can fit a statistical shape to a facial area (detected in image) which is used to locate feature points. CLM is a form of so-called Point Distribution Model (PDM) that consists of non-rigid shape and rigid global transformation parameters. Unlike PDM, which models the appearance of the whole face, CLM takes into account local patches around landmarks of interest. This leads to more generalizability because there is no need to model the complex appearance of the whole face [5].

[1,2,3] Nikolay Neshov, Ivo Draganov and Agata Manolova are with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: nneshov@tu-sofia.bg, idraganov@tu-sofia.bg, amanolova@tu-sofia.bg,

The set of model parameters controls displacement of feature key-points according to the deformation of facial expression. When we align the shape to detected face, we construct a feature vector of these parameters which is further used to determine desired expression.

II. ALGORITHM DESCRIPTION

The main steps of suggested ER algorithm can be described using the diagram in Fig. 1.
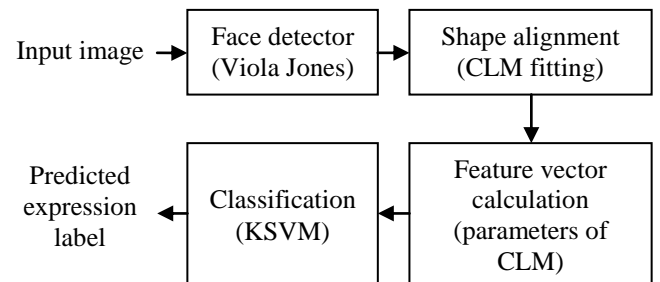
Fig.1. Overall block diagram of suggested ER alogorithm

A. Face detector

The input image is converted from color to grayscale and further examined for human faces, using Viola-Jones face detector [6]. The Viola-Jones approach provides competitive object detection rates in real-time.

B. Shape alignment

Once face location is found, the algorithm proceeds with CLM fitting strategy. In order to detect landmarks, CLM models non-rigid shape variations linearly and composes it with a global rigid transformation, placing the shape in the image frame [7]:

$$\mathbf{x}_i = s\mathbf{R}(\bar{\mathbf{x}}_i + \mathbf{\Phi}_i\mathbf{q}) + \mathbf{t}, \qquad (1)$$

where $\bar{\mathbf{x}}_i$ denotes the mean shape associated to $\mathbf{x}_i$ of the PDM's i-th landmark. $\mathbf{\Phi}_i$ are eigenvectors associated to $\mathbf{x}_i$. The set $\mathbf{p} = \{ s, \mathbf{R}, \mathbf{t}, \mathbf{q} \}$ represents PDM parameters, which contains: a global scaling s, a rotation $\mathbf{R}$, a translation $\mathbf{t}$ and a set of non-rigid parameters $\mathbf{q}$. The objective of CLM can then be interpreted as maximizing the likelihood of the model parameters such that all of its landmarks are aligned with their corresponding locations on the object in an image. Assuming conditional independence between

landmark detections, the optimization problem can be represented in the following way:

$$\arg\max_{\mathbf{p}} \prod_{i=1}^{N} p(\{l_i = aligned\} \mid \mathbf{x}_i), \qquad (2)$$

where l_i denotes whether $\mathbf{x}_i$ is aligned ($l_i = 1$) or not ($l_i = 0$) and N is the number of local features (landmarks). The optimization strategy used is based on Regularized Landmark Mean-Shift [7] which has been shown as a simple and efficient technic due to the reduction in computational calculations and avoiding the local optima. The implementation of CLM in OpenIMAJ allows detection of 62 facial landmarks.

C. Feature vector calculation

Recall Eq. 1 (by disregarding global scaling, rotation and translation parameters), the weight vector $\mathbf{q}$ for the eigenvectors $\mathbf{\Phi}_i$ can describe essential variations of the mean shape. The fitting process in OpenIMAJ computes 24-dimensional parameter - $\mathbf{q}$. This parameter is used as a feature vector in our work.

D. Classification

Finally the feature vector is classified using trained Support Vector Machine (SVM) classifier with Gaussian kernel (Kernel SVM - KSVM) [8].

III. EXPERIMENTS AND RESULTS

For the experimental evaluation of the proposed ER alogorithm we used Cohn-Kanade Extended Facial Expression Database (CK+) [9]. It contains 118 subjects, annotated by 7 expressions: anger, contempt, disgust, fear, happy, sad and surprise. Expression recognition module evaluation was performed using a leave-one-person-out (LOPO) methodology to separate training and testing parts. We also applied 5-fold cross-validation method to find out the optimal kernel SVM parameters. The experimental results of classification accuracy are shown in Table 1.

TABLE I
EXPRESSION RECOGNITION CLASSIFICATION - CONFUSION MATRIX

%	anger	con-tempt	disgust	fear	happy	sad	sur-prise
anger	**51,11**	0,00	28,89	4,44	2,22	8,89	4,44
contempt	11,11	**16,67**	0,00	33,33	5,56	27,78	5,56
disgust	11,86	0,00	**77,97**	0,00	8,47	0,00	1,69
fear	4,00	4,00	4,00	**60,00**	12,00	8,00	8,00
happy	0,00	1,45	5,80	1,45	**91,30**	0,00	0,00
sad	17,86	14,29	3,57	10,71	0,00	**39,29**	14,29
surprise	0,00	2,41	1,20	2,41	0,00	3,61	**90,36**
Avg.	60,96						

It is obvious that expressions with large displacement of landmark locations (e.g. happy and surprise) resulted in more than 90,3% correct classification. For the worst case (contempt), the accuracy is 16,67%. Looking at the results reached by algorithms suggested by another authors (e.g. Active Appearance Model (AAM) that utilizes shape and texture information [9]) it can be seen that in our work the average recognition accuracy is higher (from 50,3% to 60,96%).

Following below are some typical visual examples of incorrectly recognized emotions. In Fig. 2 is shown a face expressing disgust with all feature points detected correctly but still the feature vector falls into the class representing anger. This is the most typical case for this type of detector – where the preprocessing phase proceeds normally and the classification stage produce a faulty result.

Fig.2. Disgust emotion confused with anger while all feature points are detected correctly

Another example for misclassification is shown in Fig. 3.

Fig.3. Anger emotion confused with disgust where eyes' feature points are incorrectly detected

There the detection of eyes' contours is off-center in relation to their real position. In this case the anger is confused with disgust. There are other examples of this displacement for other emotions. The reason is the low contrast of the eyes'

area due to the nature of the lighting conditions. It's the second most frequently met case leading to errors.

In Fig. 4 is shown considerable displacement of the mouth's feature points covering only about a half of the targeted area. The reason is the misleading effect from the lighter area of the teeth right below the darker upper lip and followed by the darker area of the tongue which easily is mistaken with the lower lip. This case appears with almost the same frequency as the latter.

Fig.4. Surprise emotion confused with fear where lips' and partially face countor around the chin feature points are detected incorrectly

In Fig.5 is shown similar effect but here the feature points outline an area higher than expected.

Fig.5. Sadness emotion confused with fear where lips' feature points oclude an area higher than the mouth and right part of the face contour is slightly translated to left

The eyebrows' feature points form curves sometimes laying below (Fig.6) or above (Fig.7) the eyebrows themselves. It causes misclassification of emotions of different types but more rarely than the previous misdetections. In some cases (Fig. 8) there is no error introduced by this effect.

A more rare case is when the mouth's feature points fall below almost the entire lower lip or partially covers it (Fig.9).

Fig.6. Surprise emotion confused with sadness where eyebrows' feature points fall below the upper boundaries of the eyebrows themselves. Left part of the face contour is slightly displaced to left
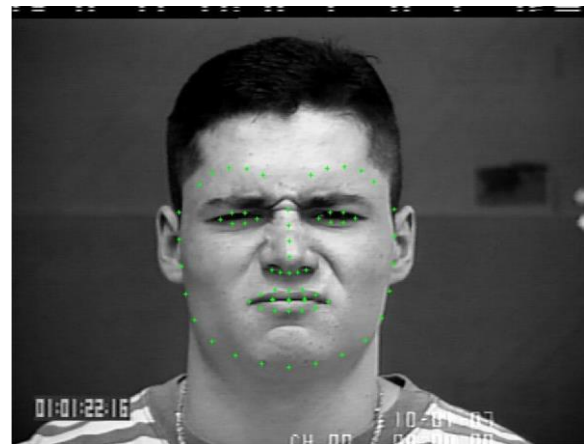
Fig.7. Disgust emotion confused with anger where eyebrows' feature points arised above the upper boundaries of the eyebrows themselves. The other features are correctly detected

Fig.8. Correctly recognized disgust emotion while eyebrows' feature points are well above the upper line of the eyebrows themselves. Eyes' contours are slightly displaced below as well

Fig.9. Sadness emotion confused with surprise where mouth's feature points are considerably shifted belowc overing only portion of the lower lip

The opposite case when mouth is detected high above its original position (Fig. 10) is also not that frequently taking place.

Fig.10. Sadness emotion confused with surprise where mouth's feature points are considerably shifted belowcovering only portion of the lower lip

Fig.11. Sadness emotion confused with contempt – eyes' contours are slightly folded where the lower part is lifted up by around1/3 of the eye's height

In this case it's possible to have some of the nose points also incorrectly located which additionally minister to misclassification of the emotions, e.g. of sadness with surprise.

In Fig. 11 is given probably the rarest case where eyes' feature points do not extend enough to the full height of the eyes themselves. Again some emotions may not be recognized correctly.

The close examination of all these cases (Fig.2-Fig.11) may very well help in the further development of the tested approach to get more accurate results.

IV. CONCLUSION

In this paper we presented an algorithm for expression classification that can distinguish seven expressions. Comparing the results over CK+ database with some other state-of-art methods we got improvement of accuracy rate more than 10%.

ACKNOWLEDGEMENT

REFERENCES

[1] J. Saragih, S. Lucey, J. Cohn, Face Alignment through Subspace Constrained Mean-Shifts, IEEE International Conference on Computer Vision (ICCV), 2009.

[2] B. T. Cootes, C. J. Taylor, D. H. Cooper, J. Graham. Active shape models their training and application. Comput. Vis. Image Underst., vol. 61 no. 1, pp. 38–59, Jan. 1995.

[3] T. Cootes, G. Edwards, C. Taylor. Active appearance models. IEEE Transactions On Pattern Analysis And Machine Intelligence, vol. 23, no. 6, pp. 681–685, June 2001.

[4] J. Hare, S. Samangooei, D. Dupplaw. OpenIMAJ and ImageTerrier: Java libraries and tools for scalable multimedia analysis and indexing of images. In Proceedings of ACM Multimedia 2011, MM '11, pp. 691–694. ACM, 2011.

[5] T. Baltrusaitis, P. Robinson, L. P. Morency, 3D constrained local model for rigid and nonrigid facial tracking. In: CVPR, 2013.

[6] P. Viola, M. J. Jones, Robust real-time face detection, Int. J. Comput. Vis., vol. 57, no. 2, pp. 137–154, 2004.

[7] J. M. Saragih, S. Lucey, J. F. Cohn. Face alignment through subspace constrained mean-shifts. IEEE 12th International Conference on Computer Vision, ICCV 2009, Kyoto, Japan, September 27 - October 4, 2009. IEEE. pp. 1034-1041, 2009.

[8] K. Crammer, Y. Singer, On the Algorithmic Implementation of Multiclass Kernel-based Vector Machines, J. of Machine Learning Research 2, pp. 265–292, 2001.

[9] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, I. Matthews, The Extended Cohn-Kanade Dataset (CK+): A complete expression dataset for action unit and emotion-specified expression. Proceedings of the Third International Workshop on CVPR for Human Communicative Behavior Analysis (CVPR4HB 2010), San Francisco, USA, pp. 94-101, 2010.

Extreme Learning Machines for Real-Time Image Classification

Stevica S. Cvetković[1], Miloš B. Stojanović[2], Saša V. Nikolić[3], Goran Z. Stančić[4]

Abstract – **In this study we investigate possibilities for application of Extreme Learning Machines (ELM) to the problem of image classification. We start by extraction of Local Binary Pattern (LBP) descriptor of the image. It is widely used global image descriptor characterized by compactness and robustness to illumination and resolution changes. Classification is done using recently introduced specific single layer neural networks called Extreme Learning Machines (ELM). Our tests on a standard benchmark dataset consisting of thousand images classified in ten categories, has shown high accuracy of results while executing almost instantly during tests (< 0.1ms).**

Keywords – **Image classification, Local binary patterns, Neural networks, Extreme learning machines.**

I. INTRODUCTION

Automatic content-based image classification is an important problem in computer vision research. The goal of an image classification system is to assign a category with the most similar visual content, to the given query image. Visual similarity between images is measured using robust and compact image descriptors (features).

There is a large set of visual descriptors available in the literature [1]. The choice of the descriptor essentially affects the overall performance of the classification system. Local Binary Pattern (LBP) is one of the most widely used descriptor due to robustness to resolution and lighting changes, low computational complexity, and compact representation [2, 3, 4]. The second crucial part of the system is machine learning technique to be applied for classification of descriptors. Support Vector Machine (SVM) is the most widely used machine learning technique for image classification purpose [5, 6]. In this study we investigate application of Extreme Learning Machines (ELM) [8, 9, 10, 11] for image classification, as an alternative to the commonly used SVM. ELM is a single hidden layer feed-forward neural network (SLFN), which overcomes an important drawback of

[1]Stevica S. Cvetković is with the University of Niš, Faculty of Electronic Engineering, Aleksandra Medvedeva 14, 18000 Niš, Serbia, e-mail: stevica.cvetkovic@elfak.ni.ac.rs

[2]Miloš B. Stojanović is with the College of Applied Technical Sciences Niš, Aleksandra Medvedeva 20, Niš 18000, Serbia, e-mail: milos.stojanovic@vtsnis.edu.rs

[3]Saša V. Nikolić is with the University of Niš, Faculty of Electronic Engineering, Aleksandra Medvedeva 14, 18000 Niš, Serbia, e-mail: sasa.nikolic@elfak.ni.ac.rs

[4]Goran Z. Stančić is with the University of Niš, Faculty of Electronic Engineering, Aleksandra Medvedeva 14, 18000 Niš, Serbia, e-mail: goran.stancic@elfak.ni.ac.rs

traditional artificial neural networks (ANNs) - their slow learning speed. It increases training speed by randomly assigning weights and biases in the hidden layer, instead of iteratively adjusting its parameters by gradient based methods. As well as minimizing training error, ELM finds smallest norm of output weights and hence have better generalization performance then gradient based training algorithms, such as backpropagation. Furthermore, it can "naturally" handle multi-class classification problem using the architecture of multiple output nodes equal to the number of classes.

In the rest of the paper we first describe the process of LBP descriptor extraction. Then we give an overview of ELM for multi-class image classification. Finally, experimental evaluation and conclusion are presented.

II. LOCAL BINARY PATTERNS (LBP)

Local Binary Pattern (LBP) is a popular descriptor that initially captures the local appearance around a pixel. LBP descriptor of the complete image is then formed as a histogram of quantized LBP values computed for every pixel of the image. It was introduced in [4] for the texture classification problem, and extended to general neighborhood sizes and rotation invariance in [2]. Since then, LBP has been extended and applied to variety of applications [3].

For a given image *I*, the local LBP descriptor centered on pixel *I(x, y)* is an array of 8 bits, with one bit encoding each of the pixels in the 3×3 neighborhood (Fig 1.). Each neighbor bit is set to 0 or 1, depending on whether the intensity of the corresponding pixel is greater than the intensity of the central pixel. To form the binary array, neighbors are scanned starting from the one to the right, at position *I(x+1, y)*, in anti-clockwise order.

a) Pixel intensities			b) Thresholded diff.			c) LBP
158	229	229	0	1	1	
126	**214**	198	0		0	**01100010**
211	250	212	0	1	0	

Fig. 1. Example of a LBP extraction process for central pixel of intensity 214.

If we use 3×3 neighborhood, there are 256 possible basic LBPs. Using an extension from [2], this can be further reduced into a smaller number of patterns (58), and implemented in a simple rotation-invariant descriptor. The extension is inspired by the fact that some binary patterns occur more frequently than others.

To describe the complete image, the quantized LBP patterns are grouped into histograms. The image could be divided into blocks, with a histogram computed for every block and concatenated to form the final descriptor. In our method we used only one image block, i.e. a global histogram is computed for the complete image.

To include image details at multiple scales, we extracted LBP histograms over the original image and several times resized image. Resizing is done to the half width and height of the original image using bicubic interpolation method. Final descriptor is formed by concatenation of the previously extracted descriptors at several scales. In our method we used 3 scales, forming a 3×58=174 dimensional image descriptor.

III. EXTREME LEARNING MACHINES (ELM)

Let us define N training examples as $(\mathbf{x}_j, \mathbf{y}_j)$ where $\mathbf{x}_j = [x_{j1}, x_{j2}, ..., x_{jn}]^T \in \mathbf{R}^n$ denotes j-th training instance of dimension n and $\mathbf{y}_j = [y_{j1}, y_{j2}, ... , y_{jm}]^T \in \mathbf{R}^m$ represents j-th training label of dimension m, where m is the number of classes. LBP image descriptor from previous section will further be denoted as $\mathbf{x}_j$, while $\mathbf{y}_j$ will denote m dimensional vector of binary class labels with value "1" denoting membership to the class. SLFN with activation function $g(x)$ and L hidden neurons could be defined as:

$$\sum_{i=1}^{L} \beta_i g(\mathbf{w}_i \cdot \mathbf{x}_j + b_i) = \mathbf{f}_j, j=1,\ldots,N \quad (1)$$

where $\mathbf{w}_i = [w_{i1}, w_{i2}, ... , w_{in}]^T$ denotes the vector of weights which connects the ith hidden neuron and all input neurons, $\beta_i = [\beta_{i1}, \beta_{i2}, ... , \beta_{im}]^T$ is the weight vector which connects i^{th} hidden neuron and all output neurons, and b_i is the bias of the ith hidden neuron. By ELM theory [8], $\mathbf{w}_i$ and b_i can be assigned in advance randomly and independently, without a priori knowledge of the input data. The ELM network structure is presented in Fig 2.

| n input neurons | L hidden neurons | m output neurons |
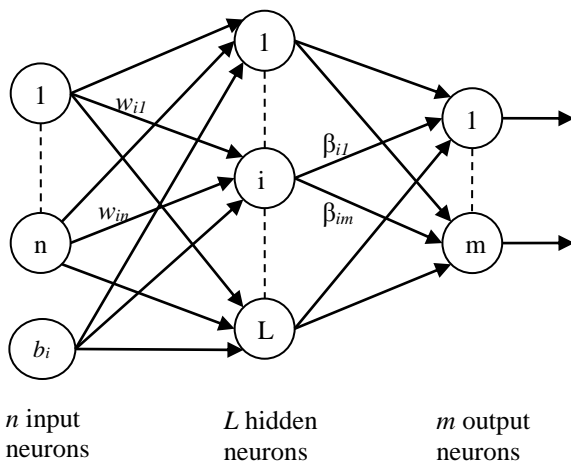
Fig. 2. Structure of the ELM network

SLFN in (1) should satisfy $\sum_{i=1}^{L} \|\mathbf{f}_i - \mathbf{y}_i\| = 0$, i.e., there exist β_i, $\mathbf{w}_i$ and b_i such that:

$$\sum_{i=1}^{L} \beta_i g(\mathbf{w}_i \cdot \mathbf{x}_j + b_i) = \mathbf{y}_j, j=1,\ldots,N \quad (2)$$

If we denote as $\mathbf{H}$ a hidden layer output matrix of the ELM; the ith column of $\mathbf{H}$ represents the ith hidden neuron's output vector regard to inputs $\mathbf{x}_1$, $\mathbf{x}_2$, ..., $\mathbf{x}_N$.

$$\mathbf{H} = \begin{bmatrix} g(\mathbf{w}_1 \cdot \mathbf{x}_1 + b_1) & \cdots & g(\mathbf{w}_L \cdot \mathbf{x}_1 + b_L) \\ \vdots & \cdots & \vdots \\ g(\mathbf{w}_1 \cdot \mathbf{x}_N + b_1) & \cdots & g(\mathbf{w}_L \cdot \mathbf{x}_N + b_L) \end{bmatrix}_{N \times L} \quad (3)$$

and

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{bmatrix}_{L \times m} \text{ and } \mathbf{Y} = \begin{bmatrix} \mathbf{y}_1^T \\ \vdots \\ \mathbf{y}_N^T \end{bmatrix}_{N \times m} \quad (4)$$

Then the equivalent matrix form of (2) can be represented as:

$$\mathbf{H}\beta = \mathbf{Y} \quad (5)$$

The output weights are then computed by finding the unique smallest norm least-squares solution of the linear system (5) as:

$$\beta = \mathbf{H}^{\dagger}\mathbf{Y} \quad (6)$$

where $\mathbf{H}^{\dagger}$ represents the Moore-Penrose generalized inverse of the $\mathbf{H}$.

For a given training set T={$(\mathbf{x}_j, \mathbf{y}_j)$}| $\mathbf{x}_j \in \mathbf{R}^n$, $\mathbf{y}_j \in \mathbf{R}^m$, $j = 1, ..., N$} with N instances of n-dimensional descriptors, sigmoid activation function $g(x)$, and hidden number of neurons L, ELM algorithm for classification problems can be summarized as follows:

Training:
 (a) Assign random input weights $\mathbf{w}_i$, and biases b_i, $i = 1,..., L$.
 (b) Compute the hidden layer output matrix $\mathbf{H}$ using (3).
 (c) Compute the output weights β using (6).

Testing:
 (a) Compute the hidden layer output matrix $\mathbf{H}_{test}$ for instances from the test set, using (3)
 (b) Compute the output matrix $\mathbf{Y}_{test}$ according to (5) using the β obtained in step 3 of the training.
 (c) For every row in $\mathbf{Y}_{test}$ (i.e. every test instance), compute a class label as the index of the maximal value in that row.

IV. EXPERIMENTAL EVALUATION

To test the proposed method for image classification, we used publicly available Corel1000 dataset [7]. It consists of

1000 images classified into following 10 categories: *Africa people, Beach, Buildings, Buses, Dinosaurs, Elephants, Flowers, Horses, Mountains and Food*. An example image for every category is presented in Fig. 3. The dataset is characterized by large variations of images inside a category, what makes this dataset close to the real world image classification scenario.

a)   b)

c)   d)
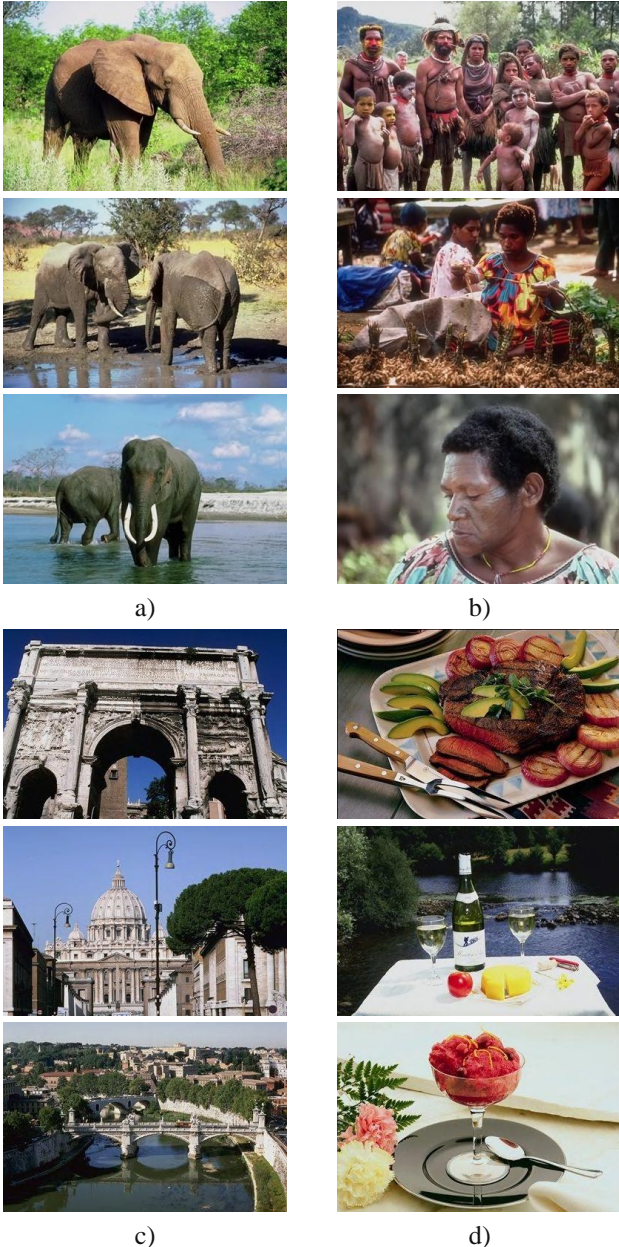
Fig. 3 Example images of several categories from *Corel1000* dataset:
a) Animals, b) Africa people, c) Buildings, d) Food.

For the tests, we implemented a method in MATLAB and used it to measure the classification accuracy and speed. Only grayscale image information is used by first converting an image into *YCbCr* color space and using only *Y* channel for LBP extraction. Final descriptor is formed by concatenation of the LBP histograms extracted at 3 scales (original + 2 downsampled). To achieve correctness of results, tests were repeated for 50 times over random partitions of every

category, where we randomly selected 50 images for training and other 50 for testing. Classification accuracy results are presented in Fig 4.
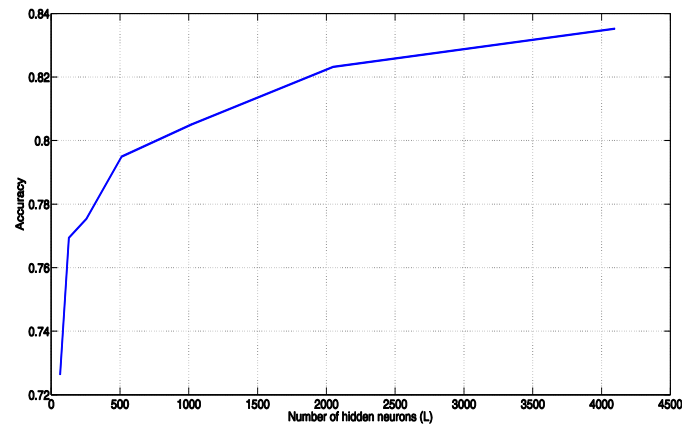
Fig. 4. Classification accuracy on Corel1000 dataset depending on the number of hidden neurons in ELM.

We further measured average training and testing time of the method on an *Intel Core i5 2.9GHz* computer. Results are presented in Fig. 5 and Fig. 6. Training time for all 500 images is only about 1 second (Fig. 5.), while test image classification is done instantly (< 0.1ms). These results demonstrate high performances in terms of training and test speed on this dataset.
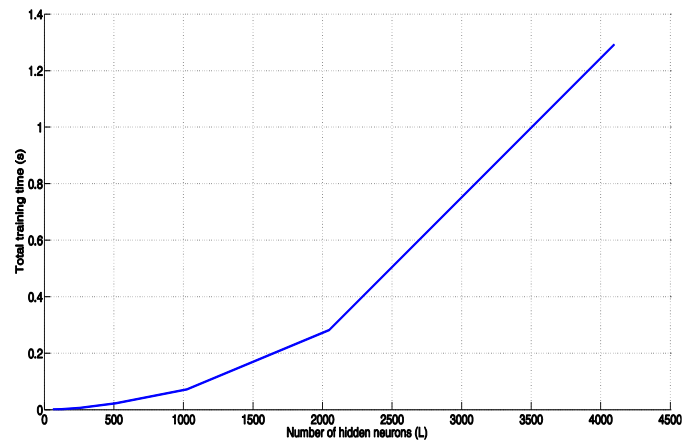
Fig. 5. Total ELM training time for all 500 training images.
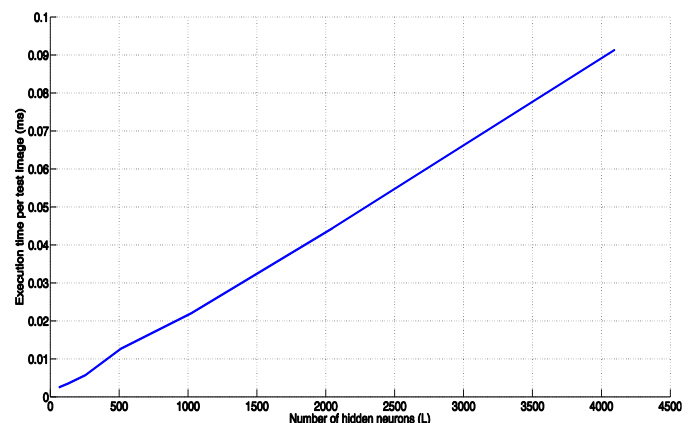
Fig. 6. Average ELM test time per image (in milliseconds).

It can be observed that increased number of neurons constantly improves classification results at the cost of increased training/testing time. Top achieved results are near to 83.52% for L = 4000 neurons in ELM hidden layer. For practical applications of the proposed method, one should experiment with values L>2000.

In order to compare results of the ELM with other common classification techniques, we measured accuracy of the Linear SVM and kernelized RBF SVM [12], on the same dataset. Linear SMV accuracy was 81.64%, while RBF SVM reached 83.49%. It can be noted that ELM outperforms Linear SVM in terms of accuracy, having the similar algorithm complexity. On the other side, ELM reaches results comparable to the kernelized SVM, while operating significantly faster during the training and testing.

V. CONCLUSION

In this paper we presented results of our research in the field of automatic image classification. Standard LBP image descriptor is used in combination with fast and powerful ELM classifier. Average accuracy of around 84% is acceptable result for rapid image classification. It can be concluded that combination of LBP descriptor with ELM classifier is reasonable choice for image classification applications. ELM classifier could be used as an alternative to the commonly used SVM. In the future, we plan to investigate performance of other types of image descriptors combined with ELM classifier, particularly integration of color and texture descriptors.

REFERENCES

[1] Xin Zhang, Yee-Hong Yang, Zhiguang Han, Hui Wang, and Chao Gao, "Object class detection: A survey," *ACM Computing Surveys,* 46, 1, Article 10, July 2013.

[2] T. Ojala, M. Pietikäinen and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with Local Binary Patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 24(7), pp. 971-987, 2002.

[3] M. Pietikäinen and G. Zhao, "Two decades of local binary patterns: A survey," In: E Bingham, S Kaski, J Laaksonen & J Lampinen (eds) Advances in Independent Component Analysis and Learning Machines, Elsevier, 2015.

[4] T. Ojala, M. Pietikäinen, and D. Harwood, "A Comparative Study of Texture Measures with Classification Based on Feature Distributions", *Pattern Recognition*, vol. 29, pp. 51-59, 1996.

[5] S. Lazebnik, C. Schmid, and J. Ponce, "Beyond Bags of Features: Spatial Pyramid Matching for Recognizing Natural Scene Categories," *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 2169–2178, 2006.

[6] C.-C. Chang and C.-J. Lin, "LIBSVM: A Library for Support Vector Machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, pp. 27:1–27:27, May 2011.

[7] James Z. Wang, Jia Li, Gio Wiederhold, "SIMPLIcity: Semantics-Sensitive Integrated Matching for Picture Libraries," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 9, pp. 947-963, 2001.

[8] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, no. 1–3, pp. 489–501, 2006.

[9] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme Learning Machine for Regression and Multiclass Classification," IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics, vol. 42, no. 2, pp. 513-529, 2012.

[10] L. L. C. Kasun, H. Zhou, G.-B. Huang, and C. M. Vong, "Representational Learning with Extreme Learning Machine for Big Data," *IEEE Intelligent Systems*, vol. 28, no. 6, pp. 31-34, December 2013.

[11] G. Huang, G.-B. Huang, S. Song, and K. You, "Trends in Extreme Learning Machines: A Review," *Neural Networks*, vol. 61, no. 1, pp. 32-48, 2015.

[12] C.-C. Chang and C.-J. Lin, "LIBSVM: A Library for Support Vector Machines," ACM Transactions on Intelligent Systems and Technology, vol. 2, no. 3, pp. 27:1–27:27, May 2011.

Analysis of Class Matrices for Complex Hadamard Transform

Rumen P. Mironov

Abstract – **An analysis of class matrices for complex Hadamard transformation (CHT) is presented. The concept of a family of CHT's is introduced and such a family is proposed to represent discrete signals and systems. Several properties of this family are outlined. The generation of the transformation matrices commences from the basic Walsh-Hadamard transformation matrix.**

Keywords – **digital signal processing, Complex Hadamard Transform, orthogonal transforms.**

I. INTRODUCTION

In this article, a class matrices for discrete orthogonal transformations with elements which are integer-valued complex numbers and may be considered as systems of complex Walsh functions are introduced. These transforms may be useful in applications where the need for complex-valued discrete orthogonal transforms arises, such as digital signal processing, spectral analysis, pattern recognition, digital coding, computational mathematic and etc. These systems of functions and transformations are called complex Hadamard transforms (CHT's) and are confined to four complex values (± 1 and $\pm j$). Dimensionality reduction in computation is a major signal processing application. Stated simply, these transform coefficients that are small may be excluded from processing operations, such as filtering, without much loss in processing accuracy. In the literature, there exists another transformation based on four-valued complex Walsh functions, called the "complex BIFORE transform" [1], [2]. For real-valued input data, the complex BIFORE transform reduces to a BIFORE or Hadamard transform whose bases are Walsh functions. The common definition of the complex BIFORE transform is based on a recursive formula defining one class of complex Hadamard matrices that involves diagonalization of higher order matrices and multiple Kronecker products. The unified complex Hadamard transforms (UCHT's) have recently been considered as a tool in spectral approach to logic design [3], [4]. Like its predecessors, the UCHT's show similar properties and characteristics [5]. The idea of using complex-valued rather than integer-valued transformation matrices for spectral processing of Boolean functions by using Perkowski

linearly independent logic is considered for the first time in article [6]. By increasing still further the number of possible different entries in the transformation matrices with complex numbers, one can expect the reduction of their spectral representation, especially if both the original functions and their spectra are presented in the form of some kind of decision diagrams

In particular, the Walsh–Hadamard transform is one of many UCHT matrices introduced here. Some of the UCHT matrices have a unique half-spectrum property (HSP). There are general fast algorithms from the representation of transform matrices in the form of layered Kronecker matrices. In addition, constant-geometry fast algorithms with in-place architecture are also available for the new transforms. The complex BIFORE transform instead has only fast transform without constant geometry algorithm. The existence of constant-geometry fast butterflies is suitable for efficient very large-scale integration (VLSI) implementation. The introduced UCHT's may be used for various applications, where the Walsh-Hadamard transform has already been used [7], [8], [9], [10], [11]. Generally, the UCHT's may be classified as the integer-valued and complex integer valued transforms. The integer-valued and complex integer-valued matrices have elements confined to two (± 1) and four complex numbers (± 1 and $\pm j$), respectively. Comparing the complex integer valued UCHT's between themselves, those that possess HSP will be advantageous as they require half of the spectral coefficients for analysis. However, it should be pointed out that if the functional data are real numbers, the existence of the HSP in complex integer valued UCHT's has no additional storage advantage compared to the integer-valued counterparts (e.g., Walsh-Hadamard transform) [4]. But, the complex integer-valued transforms are suitable for problems with complex-valued functions and for such functions, the CHT's with half spectrum property is the most compact representation. Some UCHT's are simply systems of complex Walsh functions while others become q-valued Chrestenson functions for q = 2 or 4 [12], [13]. It is then obvious that the UCHT's can be used in different applications of complex Walsh functions and Chrestenson functions in processing of multiple-valued functions, especially for the case of four-valued functions.

From the Complex Hadamard Transform (CHT), several complex decisions diagrams are derived and analysis of more general CHT properties for 1D and 2D signals are investigated [10], [11].

In this paper, the concept of a family of CHT's is introduced and such a family is proposed to represent discrete signals and systems. Several properties of this family are

[1]Rumen P. Mironov is with the Faculty of Telecommunications, Technical University of Sofia, Boul. Kl. Ohridsky 8, Sofia 1000, Bulgaria, E-mail: rmironov@tu-sofia.bg

outlined. The generation of the transformation matrices commences from the basic Walsh-Hadamard transformation matrix. All members of UCHT's may be produced by the defined direct matrix operator and recursively generated to higher dimension matrices by a single Kronecker product. It must be noted that although the basis functions in the definition that generates all UCHT matrices are discrete Walsh functions, each member of the newly defined UCHT fulfills requirements of complex Hadamard matrices; there are altogether 64 such different matrices that are introduced in this section, all of which are generated by one unifying formula.

II. MATHEMATICAL DESCRIPTION

In the definitions of existing discrete orthogonal transforms, the elements of transformation matrices normally consist of discrete values of +1 and -1; or generalizations that permit values of $e^{2\pi m j/q}$ for a prime q, $j = \sqrt{-1}$; which leads to a complete orthonormal system known as the Chrestenson system [12], [133].

The transformation matrices are defined by a set of basis discrete valued functions. To ensure that no information is lost in the resulting spectrum, orthogonality in the transformation matrix is essential. This requires zero correlation between pairs of different basis functions. In general, [H] is an orthogonal NxN matrix with real entries, when:

$$[H]^T = [H]^{-1} \ . \tag{1}$$

Then, the following relation is fulfilled:

$$[H].[H]^T = [H]^T.[H] = N.[I] \ , \tag{2}$$

where [I] is the identity matrix.

In the complex domain [H] is orthogonal if:

$$[H].[H]^* = [H]^*.[H] = N.[I] \ , \tag{3}$$

where $[H]^*$ is the represents the complex conjugate transpose of [H], $|det[H]| = N^{1/2N}$ and [H] is said to be a CHT. The resulting matrix [H] can be easily used as a binary, ternary, or quaternary transform as any two, three, or all four elements in the transformation matrix can be used for coding of two-, three-, or four-valued logic functions respectively. In addition, with an appropriate coding of the original function, the UCHT may be used as a multiple-valued transform.

All UCHT matrices of size 2x2 can be separated into two groups of 32 basic matrices dependent on the existence of the HSP [4]. These transformation matrices are listed in Table 1.

The transformation core matrix for any UCHT is defined as:

$$[H]_1^{\{r\}} = [W]_1 \circ \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} \circ [1 \quad r_3] \ , \tag{4}$$

where $[W]_1$ is the Walsh–Hadamard transform matrix of order 2 and matrix operator $\circ$ is defined as:

$$[A \circ B] = \{a(k,l) \quad b(l)\}$$

$$\tag{5}$$

and $\quad [A \circ C] = \{a(k,l) \quad c(k)\}.$

Here [A] is an r x c matrix, such as $[A] = \{a(k,l)\}$, where $0 \le k \le r$, $0 \le l \le c$. $a(k,l)$ is the current element of matrix [A] at row k and column l; $[B] = \{b(l)\}$ is 1 x c row matrix and $[C] = \{c(k)\}$ is an r x 1 column matrix.

As is shown in [4] the following properties of $\circ$ are derived:

$$(A \circ B) \otimes^n = (A \otimes^n) \circ (B \otimes^n)$$
$$(AB^T) \circ C = (A \circ C)B^T$$
$$(A \circ B_1) \circ B_2 = A \circ (B_1 \circ B_2) \qquad . \tag{6}$$
$$(A \circ B \circ C)^T = (A \circ B)^T \circ C^T$$
$$\overline{(A \circ B \circ C)} = \overline{(A \circ B)} \circ \overline{C} = \overline{A} \circ \overline{B} \circ \overline{C}$$

TABL.1

$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$	$\begin{bmatrix} j & j \\ 1 & -1 \end{bmatrix}$	$\begin{bmatrix} -1 & -1 \\ 1 & -1 \end{bmatrix}$	$\begin{bmatrix} -j & -j \\ 1 & -1 \end{bmatrix}$
$\begin{bmatrix} 1 & j \\ 1 & -j \end{bmatrix}$	$\begin{bmatrix} j & -1 \\ 1 & -j \end{bmatrix}$	$\begin{bmatrix} -1 & -j \\ 1 & -j \end{bmatrix}$	$\begin{bmatrix} -j & 1 \\ 1 & -j \end{bmatrix}$
$\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & -j \\ 1 & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}$	$\begin{bmatrix} -j & j \\ 1 & 1 \end{bmatrix}$
$\begin{bmatrix} 1 & -j \\ 1 & j \end{bmatrix}$	$\begin{bmatrix} j & 1 \\ 1 & j \end{bmatrix}$	$\begin{bmatrix} -1 & j \\ 1 & j \end{bmatrix}$	$\begin{bmatrix} -j & -1 \\ 1 & j \end{bmatrix}$
$\begin{bmatrix} 1 & 1 \\ j & -j \end{bmatrix}$	$\begin{bmatrix} j & j \\ j & -j \end{bmatrix}$	$\begin{bmatrix} -1 & -1 \\ j & -j \end{bmatrix}$	$\begin{bmatrix} -j & -j \\ j & -j \end{bmatrix}$
$\begin{bmatrix} 1 & j \\ j & 1 \end{bmatrix}$	$\begin{bmatrix} j & -1 \\ j & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & -j \\ j & 1 \end{bmatrix}$	$\begin{bmatrix} -j & 1 \\ j & 1 \end{bmatrix}$
$\begin{bmatrix} 1 & -1 \\ j & j \end{bmatrix}$	$\begin{bmatrix} j & -j \\ j & j \end{bmatrix}$	$\begin{bmatrix} -1 & 1 \\ j & j \end{bmatrix}$	$\begin{bmatrix} -j & j \\ j & j \end{bmatrix}$
$\begin{bmatrix} 1 & -j \\ j & -1 \end{bmatrix}$	$\begin{bmatrix} j & 1 \\ j & -1 \end{bmatrix}$	$\begin{bmatrix} -1 & j \\ j & -1 \end{bmatrix}$	$\begin{bmatrix} -j & -1 \\ j & -1 \end{bmatrix}$
$\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$	$\begin{bmatrix} j & j \\ -1 & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix}$	$\begin{bmatrix} -j & -j \\ -1 & 1 \end{bmatrix}$
$\begin{bmatrix} 1 & j \\ -1 & j \end{bmatrix}$	$\begin{bmatrix} j & -1 \\ -1 & j \end{bmatrix}$	$\begin{bmatrix} -1 & -j \\ -1 & j \end{bmatrix}$	$\begin{bmatrix} -j & 1 \\ -1 & j \end{bmatrix}$
$\begin{bmatrix} 1 & -1 \\ -1 & -1 \end{bmatrix}$	$\begin{bmatrix} j & -j \\ -1 & -1 \end{bmatrix}$	$\begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix}$	$\begin{bmatrix} -j & -j \\ -1 & -1 \end{bmatrix}$
$\begin{bmatrix} 1 & -j \\ -1 & -j \end{bmatrix}$	$\begin{bmatrix} j & 1 \\ -1 & -j \end{bmatrix}$	$\begin{bmatrix} -1 & j \\ -1 & -j \end{bmatrix}$	$\begin{bmatrix} -j & -1 \\ -1 & -j \end{bmatrix}$
$\begin{bmatrix} 1 & 1 \\ -j & j \end{bmatrix}$	$\begin{bmatrix} j & j \\ -j & j \end{bmatrix}$	$\begin{bmatrix} -1 & -1 \\ -j & j \end{bmatrix}$	$\begin{bmatrix} -j & -j \\ -j & j \end{bmatrix}$
$\begin{bmatrix} 1 & j \\ -j & -1 \end{bmatrix}$	$\begin{bmatrix} j & -1 \\ -j & -1 \end{bmatrix}$	$\begin{bmatrix} -1 & -j \\ -j & -1 \end{bmatrix}$	$\begin{bmatrix} -j & 1 \\ -j & -1 \end{bmatrix}$
$\begin{bmatrix} 1 & -1 \\ -j & -j \end{bmatrix}$	$\begin{bmatrix} j & -j \\ -j & -j \end{bmatrix}$	$\begin{bmatrix} -1 & 1 \\ -j & -j \end{bmatrix}$	$\begin{bmatrix} -j & j \\ -j & -j \end{bmatrix}$
$\begin{bmatrix} 1 & -j \\ -j & 1 \end{bmatrix}$	$\begin{bmatrix} j & 1 \\ -j & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & j \\ -j & 1 \end{bmatrix}$	$\begin{bmatrix} -j & -j \\ -j & 1 \end{bmatrix}$

In Eqs. (6) the matrices $[B_1]=\{b_1(l)\}$ and $[B_2]=\{b_2(l)\}$ are 1 x c row matrices and $\otimes^n$ denotes right-hand side Kronecker product applied n times.

The basis complex Hadamard matrices of order 2^n $(n>2)$ can be received as the Kronecker product of a number of identical "core" matrices of order 2^{n-1} in the following way:

$$\left[CH_{2^n}\right]=\begin{bmatrix}[CH_{2^{n-1}}] & [CH_{2^{n-1}}] \\ [CH_{2^{n-1}}] & -[CH_{2^{n-1}}]\end{bmatrix}. \qquad (7)$$

Using the basic forward one-dimensional complex Hadamard transform for n=2 from the input signal vector $\vec{X}=[x_1,x_2,x_3,........,x_N]$, the output spectral vector - $\vec{Y}=[y_1,y_2,y_3,........,y_N]$ is received by the equations [10]:

$$\left|\begin{array}{l}\vec{Y}=[CH_N]\,\vec{X} \\ \vec{X}=\dfrac{1}{N}[CH_N]\,\vec{Y}\end{array}\right. \quad \text{for:} \quad \left|\begin{array}{l}\vec{Y}=\{y(u)/u=\overline{0,N-1}\} \\ \vec{X}=\{x(v)/v=\overline{0,N-1}\}\end{array}\right. \qquad (8)$$

From the above equations the following mathematical properties can be established:

$$\left|\det[CH_N]\right|^2=N^N \qquad (9)$$

$$[CH_N][CH_N]^*=N[I] \qquad (10)$$

$$[CH_N]^{-1}=\frac{1}{N}[CH_N]^* \qquad (11)$$

$$[CH_N][CH_N]^t=[CH_N]^t.[CH_N]=N[I] \qquad (12)$$

The common results, obtained from the one dimensional Complex Hadamard Transform can be generalized for two-dimensional Complex Hadamard Transform. In this case the 2D signals (images) can be represented by the input matrix [X] with the size NxN. The result is a spatial spectrum matrix [Y] with the same size. The corresponding equations for the forward and the inverse 2D CHT are:

$$\left|\begin{array}{l}[Y]=[CH_N][X][CH_N] \\ [X]=\dfrac{1}{N^2}[CH_N][Y][CH_N]\end{array}\right. \qquad (13)$$

The symmetry of CHT coefficients allows 2D CHT to be accomplished in two steps. The first one is 1D CHT for every row the image and the second one is 1D CHT for the columns. This difference of transformation makes easier the calculations and the symmetry guarantees that the correlations between image elements in horizontal and vertical direction will influence in the same way the determination of transformed elements. The same considerations can be made for two steps calculation of the inverse 2D CHT.

III. Experimental Results

For the analyses of spectral distribution between the coefficients of 2D CHT, constructed with the different base matrices of order 2, a test image "LENNA", shown in Fig.1, with size 512x512 and 256 gray levels is used. This image is transformed by the 2D CHT with kernel 16x16. By this way the input image is divided on 1024 sub-images with size 16x16 and is calculated by MATLAB 6.5 program. In Fig.2a and Fig. 2b the averaged amplitude frequency spectrums of all sub-blocks for two-dimensional Complex Hadamard Transform and two dimensional real Hadamard Transform respectively, are shown. On Fig. 2c the averaged phase frequency spectrum calculated for all sub-blocks, for two-dimensional Complex Hadamard Transform, is shown.

Fig.1. Test image "LENNA" (512x512 pixels and 256 gray levels).
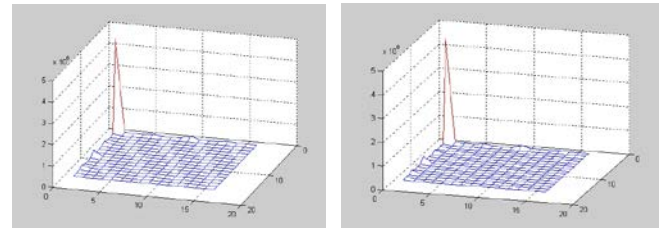
a) 2D CHT amplitude spectrum;   b) 2D HT amplitude spectrum;
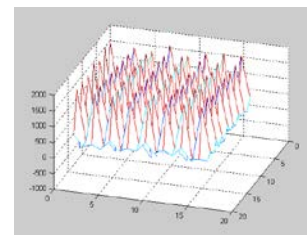
c) 2D CHT phase spectrum;

Fig. 2. Averaged amplitude and phase spectrums of image "LENNA"

IV. Conclusion

A class of Complex Hadamard matrices is presented.

The CHT is based on the mapping of 4-valued integers into the unit circle of the complex plane with elements strictly in the set {1;-1; j;-j}. Under the various permutations of the integers, there exist some conditions which will lead to the

transform being mapped to and being orthogonal in the complex domain. This has been identified as the family of UCHT's, as the mapping of the multiple-valued transforms into the complex domain will result in square basis matrices which satisfy the Hadamard's determinant equation in the complex domain. Intuitively, Walsh–Hadamard being an integer-valued transform is merely a special case of the UCHT's.

Another advantage of the presented transform is the existence of not only fast algorithms based on layered Kronecker products that can be represented by a series of strand matrices (which is similar to the complex BIFORE transform), but also a constant geometry fast algorithm that is well suited to VLSI hardware implementation. In such architecture, only one butterfly stage has to be implemented and the processed data can be fed back to the input to be processed by the same circuitry.

The general principles of complex matrices construction of high order for 1D and 2D transforms are given. The basic properties of CHT are discussed. The obtained amplitude spectrums for CHT and HT are practically identical and show that both can be used in similar applications.

The presented Complex Hadamard Transform can be used in digital signal processing for spectral analysis, pattern recognition, digital watermarking, coding and transmission of one-dimensional and two-dimensional signals.

Signal parameters in many DSP applications are estimated using the Fourier power spectrum. However, computing the Fourier transform is relatively complicated and there are applications for which it is important to achieve hardware savings, even at the expense of some loss in parameter estimation accuracy, as is the case in satellite radar altimetry. The Walsh–Hadamard transform is used for such an application. Also, it is well known from the literature that the fast Walsh–Hadamard transform can be efficiently used for the calculation of the DFT for implementing adaptive filters and for DFT spectrum filter realizations. The usual frequency-domain FIR filtering problem can be easily converted into a Walsh frequencydomain filtering problem, and similar structure results in a possible alternative for infinite-impulse response filter implementations. An efficient method for implementation of a class of isotropic quadratic filters using the Walsh–Hadamard transform was also proposed. Advantages of the 2-D Walsh–Hadamard transform, also known as S or sequential transform, in lossless image compression are well known. An integrated-circuit chip implementing 2-D Walsh–Hadamard transform has been implemented for commercial applications by Philips Corporation. As the Walsh–Hadamard transform is one of the UCHT's, it is thus believed that the important properties of the UCHT's presented in this article may also be of interest to researchers working in the above-mentioned areas where the standard Walsh–Hadamard matrices had been applied.

V. ACKNOWLEDGEMENT

VI. REFERENCES

[1] K. R. Rao, N. Ahmed. "Complex BIFORE transform," *Int. J. Syst. Sci.*, vol. 2, pp. 149–162, 1971.

[2] N. Ahmed, K. R. Rao. *Orthogonal Transforms for Digital Signal Processing*, Springer-Verlag Berlin, Heidelberg, 1975.

[3] B. Falkowski, S. Rahardja. "Complex Spectral Decision Diagrams", *Proc. of the 26th Int. Symposium on Multiple Valued Logic*, Vol. ISMVL'96, 1996.

[4] S. Rahardja, B. Falkowski. "Family of Unified Complex Hadamard Transforms", *IEEE Trans. on Circuits and Systems-II: Analog and Digital Signal Processing*, Vol. 46, No. 8, August 1999, pp.1094-100.

[5] S. Rahardja, B. Falkowski. "Complex Composite Spectra of Unified Complex Hadamard Transform for Logic Functions", *IEEE Trans. on Circuits and Systems-II: Analog and Digital Signal Processing*, Vol. 47, No. 11, November 2000.

[6] B. J. Falkowski and S. Rahardja, "Calculation and properties of fast linearly independent logic transformation", *IEEE Trans. Circuits Syst. II*, vol. 44, pp. 646–655, Aug. 1997.

[7] A. D. Poularikas. *The Transforms and Applications Handbook*, Second Ed., CRC Press, 2000.

[8] W. K. Pratt. *Digital Image Processing*, John Wiley&Sons, N.Y., 2001.

[9] B. Falkowski, S. Rahardja. "Complex Hadamard Transforms: Properties, Relations and Architecture", *IEICE Trans. Fundamentals*. Vol. E87-A, No.8, August 2004.

[10] R. Kountchev, R. Mironov. "Audio Watermarking in the Phase-Frequency Domain", *XL Intern. Scientific Conference on Information, Communication and Energy Systems and Technologies*, ICEST'2005, Nis, Serbia and Montenegro, 2005.

[11] R. Mironov, R. Kountchev. "Analysis of Complex Hadamard Transform Properties", *XLI International Scientific Conference on Information, Communication and Energy Systems and Technologies*, ICEST 2006, 26 June – 1st July, Sofia, Bulgaria, 2006, pp.173-176.

[12] R. S. Stankovic, M. R. Stojic, and M. S. Stankovic, "Recent developments in abstract harmonic analysis with applications in signal processing", in Science. Belgrade, Yugoslavia: Science Publisher, 1996.

[13] L. A. Zalmanzon, "Fourier, Walsh, and Haar transforms and their application in control, communication and other fields", in Science. Moscow, U.S.S.R.: Science Publisher, 1989.

[14] Aye Aung, Boon Poh Ng, S. Rahardja. "Conjugate Symmetric Sequency-Ordered Complex Hadamard Transform". *IEEE Transactions on Signal Processing*, July 2009, Vol. 57, No. 7, pp. 2582-2593.

ORAL SESSION
ENERGY SYSTEMS AND EFFICIENCY

Analysis of the Grounding System of the Mines with Surface Exploitation

Nikolche Acevski [1] , Mile Spirovski [2] , Metodija Atanasovski [3] and Blagoj Stevanoski [4]

Abstract – **In the event of a single-phase short circuit on the 110 kV side of the switchgear "Oslomej", the power is distributed to all earth electrodes. During the above process, potentials significantly higher than the potential of the soil may occur on the metal parts of the mine equipment. The purpose of this paper is to perform an analysis of the 6 kV network in the mining power complex of "Oslomej", thereby calculating the parameters of the grounding system (touch and step voltages) by using empirical formulas and appropriate software.**

Keywords – **Grounding system (GS), touch and step voltages.**

I. INTRODUCTION

Thermal power plant Oslomej is complex technological and technical unity consist by mine with surface exploitation of minerals and thermal power plant along with object located in the area. Thermal power plant through 110 kV switchgear is connected with the HV transmition system of R. M, with three power lines, towards substation 110 kV – Kicevo, substation 110 kV – Vrutok, substation 110 kV Samokov. On the 110 kV switchgear is connected generator G1 with block transformer AT1 through wich is conected to power grid. Also in the thermal power plant are located transformer BT1 and transformer BT2, whose role is to supply the electricity consumption it the power plant, asynchronous motors, including coal mills, smoke ventilator, cooling water pump, power pump, fresh air ventilator, compressors etc. In the thermal power plant Oslomej is performed complex GS, set as grounding mesh that connects all electrical equipment in the plant, affecting the entire area inside the fence, with ancillary facilities. Mutually connected are grounding of elements in TPP Oslomej and parts of the switchyard 110kV plant in the TPP, 110kV power plant and conveyors for supply of coal from the mine to the plant storage. At first it is performed modeling of the GS network with equivalent schemes of the elements, then using the software package Neplan 5.2., the short circuit current injected into GS is calculated. Characteristics of the groundings are analyzed using the program Zazem.xls, MrezZaz.xls, while for calculating of the potentials and currents of GS, step and touch voltages, appropriate program developed with software package Matlab R2011a is used. In Section 2, the GS is described.

[1]Nikolche Acevski, is with the Faculty of Technical Sciences, 7000 Bitola, R. Macedonia, e_mail: nikola@acevski.com.

[2]Mile Spirovski, is with the Faculty of Technical Sciences, 7000 Bitola, R. Macedonia, e_mail: mile.spirovski@uklo.edu.mk.

[3]Metodija Atanasovski, is with the Faculty of Technical Sciences, 7000 Bitola, R. Macedonia.

[4]Blagoj Stevanoski, is with the Faculty of Technical Sciences, 7000 Bitola, R. Macedonia.

II. GROUNDING SYSTEM

A. Modelling mesh grounding systems of substations

Substations 110 / 35 kV and 35/6 kV, Figs.1 and 2, are set grounding mesh that connects all electrical equipment inside the fence. The grounding resistance and potential differences of touch and step are calculated as in [1], [3], Tables I and II.
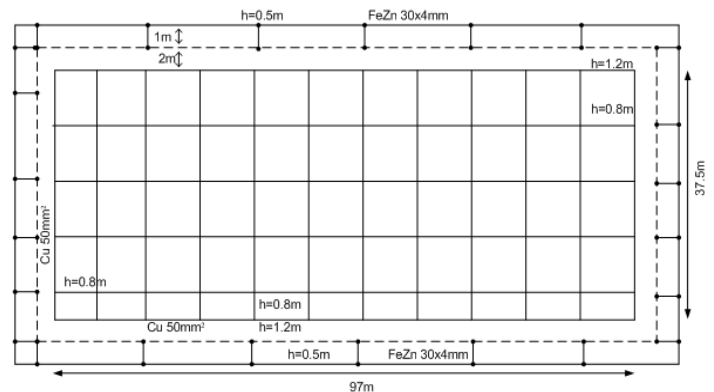
Fig.1. Mesh grounding of the substation 110/35 kV Oslomej

TABLE I

CALCULATION OF GROUNDING RESISTANCE, POTENTIAL DIFFERENCES OF TOUCH AND STEP IN PERCENT

According to:	Rz (Ω)	ΔEd %	ΔEc %
Zazem.xls	0,85	47,12	5,30
Nahman [1]	0,82	31,31	4,78
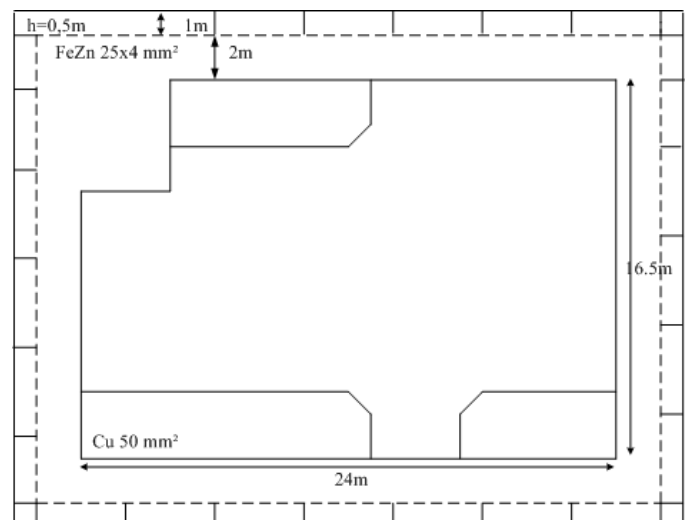
Fig.2. Mesh grounding of the substation 35/6 kV mine Oslomej

According to:	R_Z (Ω)	ΔEd %	ΔEc %
Zazem.xls	2,59	43,79	10,50
Nahman [1]	2,53	18,73	6,29

B. Modelling of TS 6/0,4 kV and auxiliary groundings

Each TS 6 / x kV will be modeled with transverse mounted active resistance to ground *R*, calculated by [3]. Auxiliary groundings of the various types of mining facilities and machines are modeled in an identical manner as the grounding TS 6 /x kV. It should be noted that cases are possible when more diverse plated grounding are connected to the same grounding spot. In that case, the replacement scheme GS will emerge as active resistors connected in parallel.

C. Modelling the surface grounding

If there are *rails and strips placed on the surface of the earth*, resistance to ground of the sliding bar or a length *l* and an equivalent diameter *d*, placed on the surface of the earth, with the average specific resistance of soil *ρ*, will be:

$$R = \frac{\rho}{\pi \cdot l} \ln \frac{2l}{d} \qquad (1)$$

If the bar is on one end tied to the GS and thus is free of its other end, then it will be presented with a transversely mounted active resistance *R*. The picture we are given is a replacement scheme of the track when the two ends are connected to different plated grounding, Fig.3.
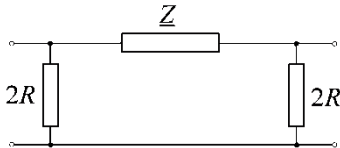
Fig. 3 - replacement scheme

The value of the parameter $\underline{Z}$ is calculated by the relation:

$$Z = \left[\frac{\rho_{Fe} \cdot l}{S} + 0,05 \right] + j \left[0,1445 \log\left(\frac{2D_e}{d} \right) + 0,0157 \mu_r \right] \quad (2)$$

ρ_{Fe} [Ω·mm²/km] - specific resistance of iron

l [km] - the length of rail / bar;
S [mm²] - cross section of rail / bar;
μ_r - relative magnetic permeability of the material;

$$D_e = 658\sqrt{\rho / f} \qquad (3)$$

d [mm] - equivalent diameter of the bar / rail, with approximately valid $d = 1,128 \cdot \sqrt{S}$.

Excavators and other mining machinery and buildings that have large land area, with its caterpillars achieves good electrical contact with the ground. Thus, in case of short circuit in a 110 kV network, where they will get some potential, through the said contact lead the current in the ground and in their current environment creates a field that can be dangerous for people who are in their immediate nearby. For the purposes of modeling the equivalent scheme GS, they can be treated as concentrated grounding that the appropriate place of the equivalent scheme would introduce an active resistance *R*. Analyses show that the land surfaces of the excavators, machine with caterpillars and other devices that have contact with the ground surface, can be satisfactorily modeled by a suitable mesh GS, mounted horizontally on the ground of a certain small depth h. Moreover reticulated GS should have the same geometry as the geometry of land area of the dredge / device, and modeling can only be successfully performed if land surface is replaced by a dense network of horizontal strips, placed on a small distance between (the example D = 50 cm), buried in shallow (eg *h* = 5 cm).

D. Modelling 6 kV cables and characteristics of their models

6 kV cables EPN78 used for distribution of electricity in 6 kV input power network in the area of the mine, despite the power conductors (three phase conductors), have more and signal conductors or a conductive sheath. Signal conductors (i.e conducting sheath) are isolated in terms of land and phase conductors, and the emergence of single-phase short links in 110 kV network through them the transfer potentials in the GS of the mine. Thus, each cable, watched with its return route across the ground, may be presented with a I-replacement scheme, ie an ordinal impedance $Z = z \cdot l$. Thus, i.e longitudinal impedance per unit length in this case i will be calculated using known formulas Carson-'s, [1]:

$$\underline{z} = r + jx = \left(\frac{1000}{k \cdot s} + 0,05 \right) + j \log \frac{D_e}{D_S} \qquad (4)$$

k [Skm/mm²] - conductivity specific material from which is made cable signals;
S [mm²] - total cross section of signal cables;
De [m] - the equivalent depth of return current path across the ground (soil);
ρ [Ω · μ] - average value of the specific resistance of the terrain through which the cable passes;
Ds [m] - own middling geometric distance of the cable.

$D_S = \sqrt[3]{r_p \cdot (D/2)^2}$ - the signal cable has three conductors, placed in the vertices of an equilateral triangle with side *D* / 2.

r_p is the radius which is indicated by the conductor signal cable, and *D* is the outside diameter of the conducting sheath.

	cable	D(mm)	Ds(mm)	$r(\Omega m)$	$x(\Omega m)$
1	EpN78 3x35+3x10	54	11,4	0,688	0,710
2	EpN78 3x50+3x16	60	13,2	0,679	0,700
3	EpN78 3x70+3x16	65	14,3	0,673	0,695
4	EpN78 3x95+3x25	67	15,5	0,669	0,690

E. Modelling of overhead lines with protective rope

Overhead lines also participate in transfer of currents and potentials in case of short circuit in the 110kV system, and because any such line should be represented in an appropriate manner so-called "- replacement scheme". Protective rope is of the type FeIII50mm^2, and a range of power lines $lj = 300$m.

Impedance of the protective rope for a span is calculated:

$$\overline{Z_{r1}} = (r_j + jx_j) \cdot l_j = (1{,}82 + j0{,}36)\Omega \qquad (5)$$

r_j, x_j -active and inductive resistance of the rope per km length

$$r_j = 0{,}05 + \rho_j \frac{10^3}{S_j} = 6{,}05\,\Omega/km \qquad (6)$$

$$x_j = 0{,}1445 \cdot \log \frac{2 \cdot De}{d_j} + 0{,}016 \cdot \mu_r = 1{,}22\,\Omega/km \qquad (7)$$

dj - diameter of the rope $Sj = 50$mm^2 $dj = 0{,}009$ m.

De - equivalent distance between the protective rope and phase conductors of power lines (m).

Rs - grounding resistance of each of the columns, which for 110 kV is the average power lines: $Rs=10\Omega$. According to [1] the equivalent impedance of the grounding is:

$$\overline{Zek.v} = \sqrt{\overline{Z_1} \cdot Rs} - \frac{\overline{Zr_1}}{2} = (2{,}97 + j1{,}61)\Omega \qquad (8)$$

Self impedance, mutual impedance between protective rope and phase conductors of power line and reduction factor are:

$$\underline{Z_{mi}} = 0{,}05 + j \cdot \log \frac{D_e}{D_m} = 0{,}05 + j2{,}24 \qquad (9)$$

$$\underline{Z_{si}} = r \cdot jx = 6{,}05 + j1{,}22 \qquad (10)$$

$$r = 1 - \frac{\underline{Z_{mi}}}{\underline{Z_{si}}} = \frac{0{,}05 + j2{,}24}{6{,}05 + j1{,}22} = 0{,}9231 - j0{,}3542 \qquad (11)$$

III. Single-phase current short circuit at 110 kV – model in neplan

During the analysis of the GS, the problem with transfer potentials and protection from hazardous touch and step voltages, we take the single phase short circuit 110kV side of the plant switchyard as the most suitable for this kind of analysis. In case of single phase short circuit in switchgear 110/35/6 Elektrana, current is distributed to all GS as grounding of 110 kV poles, grounding of 110/35/6 Elektrana and GS of the mine. Using the software package NEPLAN 5.2.2 it is made electricity grid of MV of TPP Oslomej and it is estimated the value when single phase circuit happened on the 110 kV. Elements modeling is according to the standard models for the analysis of short-circuit, Figs.4 and 5, Table IV. For calculation of voltages in GS current value injected in the GS in case of a short circuit side is $I_k^{''} = 11\,017$ A.
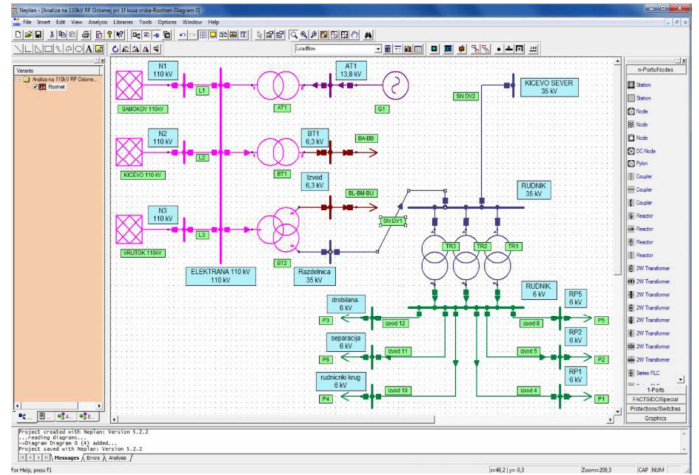
Fig. 4. Modeling the TPP Oslomej with NEPLAN

TABLE IV
CURRENT VALUE WHEN SINGLE-PHASE SHORT CIRCUIT HAPPENED ON 110 KV SIDE

Fig. 5. Modeling the 6 kV network in the mine Oslomej

IV. RESULTS

TABLE V

VOLTAGES AND CURRENTS IN THE GROUNDING SYSTEM

N	Grounding	U(V)	Ed(V)	Ec(V)	Iz(A)
1	RP110kV	1920,5	376,1	154,8	7435,9
2	TS35/6kV	382,8	167,6	40,2	247,8
3	RSO	321,4	82,7	28,9	30,9
4	US3	214,0	73,1	27,0	65,1
5	RP7	217,5	66,1	38,4	58,1
6	RSO	213,4	54,9	19,2	20,5
7	RP5	27,4	14,3	3,3	5,0
8	RSO	22,2	5,7	2,0	2,1
9	ARS3500	19,4	9,4	6,8	6,6
10	OT2	5,1	3,2	0,3	14,6
11	RSO	24,9	6,4	2,2	2,4
12	SRS401	21,2	8,2	3,1	9,3
13	BRS1200	18,9	9,0	2,3	12,4
14	ST2	22,9	13,8	1,4	46,7
15	RSO	97,7	25,1	8,8	9,4
16	RP4	27,0	14,1	3,3	4,9
17	UE4	20,9	11,5	1,8	24,4
18	RSO	15,8	4,1	1,4	1,5
19	SX400	14,7	5,5	4,9	3,7
20	PV31800	12,5	5,6	1,7	6,8
21	RSO	23,6	6,1	2,1	2,3
22	ES10/70	22,6	13,4	9,6	3,8
23	RP2	86,5	45,3	10,5	15,8
24	US7	81,9	45,1	7,0	95,2
25	RSO	78,7	20,3	7,1	7,6
26	ES6/45	76,7	47,0	14,3	12,9
27	RP1	83,7	43,8	10,1	15,3
28	RSO	70,2	18,1	6,3	6,7
29	OT1	47,0	29,5	2,5	134,1
30	ZP2500	53,7	25,9	18,8	18,2
31	ZT1	12,8	7,7	0,8	26,1
32	JE1	4,4	2,7	0,3	9,1
33	RSO	8,6	2,2	0,8	0,8
34	SX400/1	8,7	3,3	2,9	2,2
35	PVP1800	8,9	4,6	0,9	7,8
36	SRS400	8,8	3,7	1,3	3,8
37	BRS1400	8,9	4,3	1,0	5,9
38	US6	75,9	41,9	6,5	88,3
39	US5	67,6	37,3	5,8	78,6
40	TS110kV	274,0	68,5	6,9	601,8
41	TS110kV	274,0	68,5	6,9	601,8
42	TS110kV	274,0	68,5	6,9	601,8
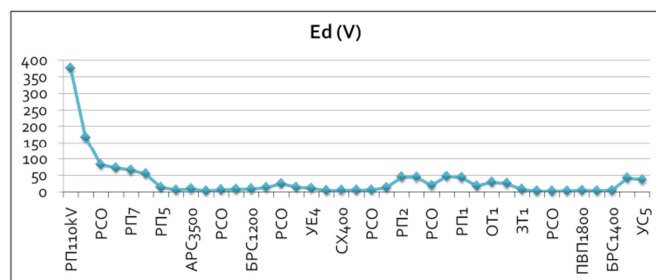
In the Table V, voltages of the groundings, potential differences of touch and step and currents in the groundings are given, calculated by software Matlab R2011a. Critical points in the network are given by Figs. 6 and 7. According to calculations touch and step voltages, are in accordance with existing regulations [8], by which inside the plant are limited to 300V, during off $t = 0.1$sek, except in node 1, RP 110 kV. In this node, we should put vertical electrodes, or/and salt into the soil around the grounding to get better performances.
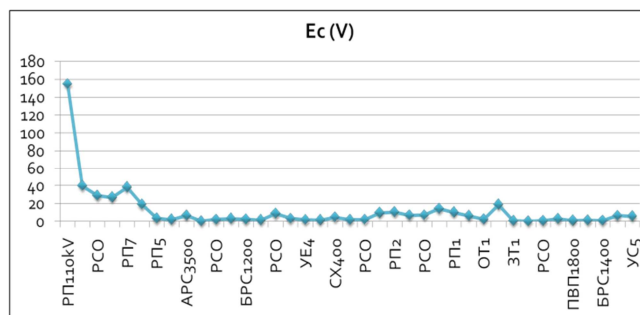
Fig.7. Step voltages at the critical points in the network

V. CONCLUSION

For analysis of the GS, the problem with transferred potentials and protection from hazardous touch and step voltages, we take the single phase short circuit on the 110kV side as the most suitable for this kind of analyses. At first it is performed modeling of the GS network with equivalent schemes of the elements, then using the software package Neplan 5.2. the short circuit current injected into GS is calculated. Characteristics of the groundings are analyzed using the program Zazem.xls, MrezZaz.xls., while for calculating of the potentials of GS, currents in the branches of the GS, step and touch voltages is used appropriate program developed with software package Matlab R2011a. The results obtained with calculations can be used in determining the risk of excessive voltages and critical points in the network and according to the results can be propose appropriate technical solutions for proper protection and proper building of the GS.

REFERENCES

[1] R. Ackovski, „Zazemjuvaci i zazemjuvacki sistemi vo elektroenergetskite mrezi", FEIT, Skopje, septemvri 2008.
[2] M. Zlatanovski, „Zazemjuvanje i Zastitni Merki", Skopje 2007.
[3] N. Acevski, "Prilog kon modelite za resavanje i analiza na zazemjuvaci i zazemjuvacki sistemi", Skopje, 2003.
[4] D.Rajicic, R.Taleski, „Metodi za Analiza na EES" Skopje 1996.
[5] Software package NEPLAN 5.2.2. http://www.neplan.com
[6] I. Pesev, R. Simitcev, K. Tulev, "Tehnicko izvestuvawe na nadvoresno zazemjuvanje na TE Oslomej", Skopje 1977.
[7] V. Dimcev, ''Studija za ispituvanje na otpor na zazemjuvanje i gromobranska zastita vo REK Oslomej'', Skopje 2009.
[8] Pravilnik za tehnicki normativi za elektoenergetski postrojki so nominalen napon nad 1000V, Sluzben list na SFRJ br.4/1974.

Fig.6. Touch voltages at the critical points in the network

Overcurrent Protection Analysis of Distribution Networks with Dispersed Generation

Metodija Atanasovski, Ljupco Trpezanovski, Blagoja Stevanovski and Nikolce Acevski[1]

Abstract – **The goal of this paper is to research and analyze the problem with distribution feeder protection selectivity when dispersed generation (DG) is present in the network. Typical example of distribution network (DN) is created and complete protection which is used for distribution feeders, buses and transformers in HV/MV station is modelled. The protection setting and testing of its selectivity is performed in conditions without DG. Then, DG is connected on the network. Set of problems with DN protection selectivity are identified. Several useful conclusions and solutions for overcoming the problems are presented at the end of the paper.**

Keywords – **Overcurrent protection, Distribution network, Dispersed generation.**

I. INTRODUCTION

DG by definition is not centrally planned and dispatched, so its commitment in generation of active and reactive power is out of control of power system operators. The installed capacity of DG has significant impact on which voltage level it will be connected. Typical voltage levels for DG connection are distribution ones from 400 V up to 110 kV [1]. DG has significant technical and economic impact on DN as it alters the power flows in the network from unidirectional to bidirectional, and by that DG affects network losses, voltage conditions and short circuit currents. Also DG changes the operating conditions and selectivity of installed protection devices into DN.

The goal of this paper is to research and analyze the problem with distribution feeder protection selectivity when dispersed generation (DG) is present in the network. Typical example of DN is created and complete protection which is used for distribution feeders, buses and transformers in HV/MV station is modelled. The protection setting and testing of its selectivity is performed in conditions without DG. Then, DG is connected on the network. Two types of generators are used in connected DG: synchronous and asynchronous generator. Set of problems with DN protection selectivity are identified. It is shown in the paper that DG connection causes serious problems with distribution feeder protection selectivity, especially in case when the fault occurred on neighbouring feeder with that where DG is connected. Island operation of DG is also analyzed as hazardous state which can be a case when fault is on the same feeder with DG independently upstream to the root or downstream to network

ends from DG point of connection. Several useful conclusions and solutions for overcoming the problems are presented at the end of the paper.

The paper is composed of five sections. Section II elaborates the general concept of distribution feeders protection and the principles of protection setting. Section III is theoretical background of DG impact on protection of distribution feeders. Section IV is study case analysis of DG impact on protection. Section V is conclusion.

II. DISTRIBUTION FEEDERS PROTECTION

DNs are usually operated as radial networks. Protection from short circuit is mostly realized with overcurrent protection (OCP) with three modules: instantaneous OCP I>>, time delayed OCP I> and OCP for ground fault I0>. Complete functionality and setting of this protection has to be achieved for typical operating conditions of the network. OCP has two setting dimensions: current and time. Calculations for protection settings are performed for steady state operating conditions of DN. Power flow results for maximum load are mostly used for current settings of OCP. Short circuit calculations are used for testing OCP sensitivity and selectivity. Minimum short circuit current is calculated in protective device zone of operation.

Maximum value of current I_{max} through the protected element is used as basis for OCP current setting. This value has to be less or equal to element rated current I_n.

Instantaneous OCP I>> is usually located on primary side of transformer HV/MV (used as replacement or fast reserve of differential protection), at the beginning of feeder and also at longer feeders for increase of selectivity can be installed in few location downstream the feeder. Time setting of instantaneous OCP is zero by definition. When it is installed at the beginning of the feeder, then its current setting is calculated as:

$$I_{set} = k_m \cdot I_{max} \cdot \left(1/CT\right) \qquad (1)$$

Where k_m – is setting coefficient (usually has value 6), CT – is current transformer ratio.

Current setting of time delayed OCP I> is calculated with following expression:

$$I_{set} = k \cdot I_{max} \cdot \left(1/CT\right) \qquad (2)$$

Where: k – is setting coefficient (usually has value 1.5).

More about protection settings of DN elements can be found in reference [2].

[1]Metodija Atanasovski, Ljupco Trpezanovski, Blagoja Stevanovski and Nikolce Acevski are with the Faculty of Technical Scieneces-Bitola University St. Kliment Ohridski - Bitola, address: Makedonska falanga No. 33, 7000 Bitola. E-mail: metodija.atanasovski@tfb.uklo.edu.mk.

III. DG IMPACT ON DN PROTECTION

DG connection increases short circuit currents and has influence on voltage condition in DN. Also DG changes operation of installed protective devices. If fault occurs as depicted on Fig. 1, besides from the network, short circuit current will be fed from DG also, that results with greater impedance measured by line distance protection (if it is installed) and shortening of protected zone. OCP will measure increased current in comparison without DG. This problem can be overcome with usage of modern protection equipment which is sensitive and has possibility of precise setting [3].

However, at usage of this protection equipment, problem can occur during fault on neighboring feeder, which leads to false operation. During protection setting in DN with DG, it has to be taken into consideration that protection should function correctly when by any cause one or all DG will not be connected on the network.

Except mentioned problem, another one is possible when network is out of service and DG continues to work (islanded work). Islanded operation is not allowed by distribution companies. Also during any protection setting individually, network grounding and type of DG has to be taken into account for obtaining normal operation in any situation.

Feeders with DG can be out of service although there is no fault on them. This is most often case when fault is on feeders supplied from the same substation as feeder with DG (Fig. 2). In case of fault at location K, current will flow from the network and from DG. Fault location will be cleared by protection R_1. However, current I_{DG} will flow through protection R_2 and if it has enough great value and protection R_2 is not directional, it will put out of service feeder with DG although there is no fault on it. This is a problem with protection selectivity which is especially expressed in situation when fault location and DG connection are near to substation HV/MV. This phenomenon is influenced by the type of generator used at DG technology. Synchronous generators significantly contribute to the problem with selectivity. Asynchronous generators impact on selectivity is less expressed. DG with inverters practically has no influence on protective devices in the network.

Protection selectivity can be achieved with its precise setting. If the value of problematic protection setting current cannot be changed, problem can be solved with appropriate time setting of protection operation on neighboring feeder. But, it should be done without exceeding thermal capability of network elements. If problem of selectivity at unidirectional OCP cannot be solved with mentioned solutions above, then directional protection can be installed.

IV. STUDY CASE ANALYSIS

Theoretically analyzed problems with OCP selectivity in previous section will be illustrated on study case of typical DN (Fig. 3), which is modeled with complete protection used for distribution feeders, buses and transformers in substation HV/MV1/MV2. DN protection is set and tested in conditions without DG, because it is originally passive. The network is taken from reference [2], where complete network data can be

found. DN model and complete analysis are done with software package NEPLAN 5.3.5 [4]. Analyzed DN is radial and consisted of substation 110/20/10 kV/kV/kV with two 3-windings transformers 110/20/10 kV, with rated power 31,5/31,5/10,5 MVA. Secondary side 20 kV of transformers is with grounded neutral through active resistance of 40 Ω. Used OCP in DN are shown on Fig. 3 (indicated with numbers from 1 to 9). They are used for protection of feeders, bus-bars and transformers. Table I summarizes current and time settings of each OCP. DN is consisted of three feeders. Feeder 1 supplies two nodes with three 20/0.4 kV transformers each 1 MVA, feeder 2 supplies three nodes with eight 20/0.4 kV transformers each 1 MVA and feeder 3 supplies two nodes with two 20/0.4 kV transformers each 1 MVA.
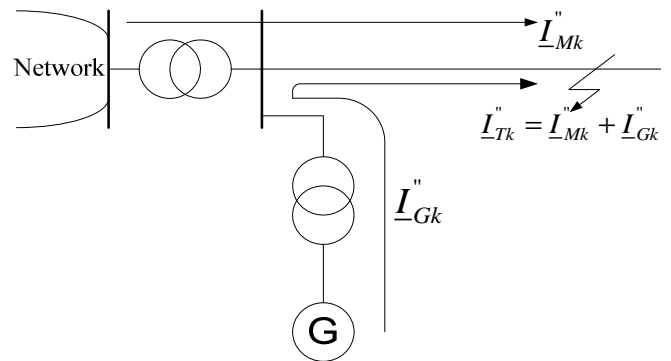
Fig. 1. Short circuit in DN with DG
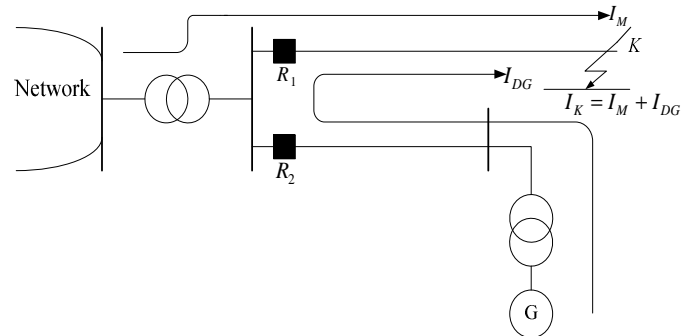
Fig. 2. Distribution feeder selectivity problem in DN with DG
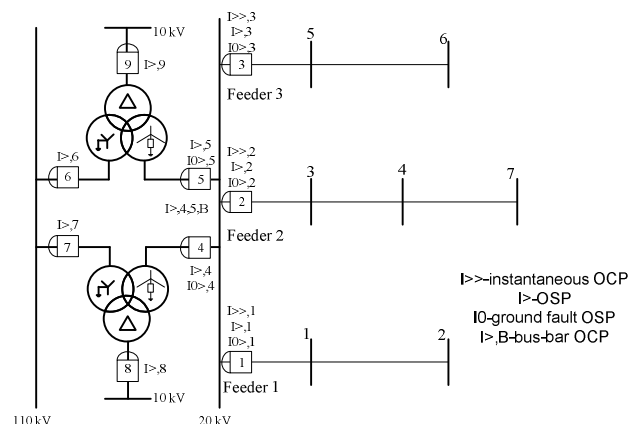
Fig. 3. Study case DN with OCP

110 kV network is modeled with Thevenin equivalent of power system with resistance 4.3 Ω and reactance of 8.07 Ω. For executed analysis it is assumed that equivalent parameters for sub-transient, transient and steady state period are equal. Same parameters are used for inverse and zero sequence. This is regular assumption for short circuit calculations in DN (except for calculations in the substation (HV/MV1/MV2).

TABLE I
SETTING OF USED OCP IN DN

OCP	Setting A/s			OCP	Setting A/s		
	I>>	I>	I0>		I>>	I>	I0>
1	540/0	150/0,5	60/0,2	6	-	260/2	-
2	1380/0	360/0,5	60/0,2	7	-	260/2	-
3	345/0	90/0,5	60/0,2	8	-	900/1,5	-
4	-	1400/1	200/0,5	9	-	900/1,5	-
5	-	1400/1	200/0,5	B	-	4000/0,1	-

Then, DG with synchronous generator with rated power 3 MVA and voltage 0.4 kV is connected to DN. Connection is realized with two transformers 0.4/20 kV with power 1.6 MVA. DG location is changed for investigation its impact on installed OCP in DN.

Four cases of problems with OCP selectivity are depicted on Fig. 4, with different DG and three phase fault location. Namely at Fig. 4a, location of three-phase fault is on half of line 3 and DG is connected on node 1. Through protection 1 will flow 570 A which lead to false operation of OCP I>> on feeder 1. In this case for selectivity only protection 2 has to operate.

Fig. 4b shows three phase fault on half of line 2 and DG is on node 5. Through protection 3 460 A will flow, which lead to false tripping of OCP I>> on feeder 3. In this case for selectivity only protection 1 has to operate.

Fig. 4c shows three phase fault on half of line 5 and DG is on node 2. Through protection 1 will flow 560 A, which lead to false operation of protection I>> on feeder 1. In this case for selectivity only protection 3 has to operate.

Fig. 4d shows case of three phase fault on half of line 4, DG is connected on node 6. Through protection 3 will flow 430 A, which lead to false tripping of OCP I>> on feeder 3. In this case for selectivity only protection 2 has to operate.

For analyzed study case it can be concluded that with DG connection with synchronous generator at any node on feeders 1 and 3, there is a problem with OCP selectivity for fault on neighboring feeder. But, when DG is connected at nodes on feeder 2, there is no problem with OCP selectivity on this feeder, when fault is on neighboring ones. This is because of the fact that this feeder supplies eight transformers TC 20/0,4 kV each 1 MVA and the setting of instantaneous OCP I>>,2 on this feeder is very high 1380 A. Practically when three phase fault occurs on neighboring feeder, this protection cannot be activated with short circuit current generated from DG.

Analysis is also performed in case when connected DG is with asynchronous generator with rated power 3 MVA and

voltage 0.4 kV. Connection is realized with two transformers 0,4/20 kV with power 1.6 MVA.

On Fig. 5 are shown two cases for OCP selectivity problem with different DG and three phase fault location. Namely, at Fig. 5a three phase location is on the half of line 2 and DG is connected on node 5. Through protection 3 will flow 370 A which lead to false operation of OCP I>> on feeder 3. In this case for selectivity only protection 1 has to operate.

At Fig. 5b, fault location is on half of line 4 and DG is connected on node 6. Through protection 3 will flow 355 A, which will lead to false tripping of OCP I>> on feeder 3. In this case for selectivity only protection 2 has to operate.

According to performed analysis with DG connection with asynchronous generator, it can be concluded that at any node on feeder 3, there is a problem with OCP selectivity on this feeder for fault on two neighboring feeders. Connection of DG at nodes on feeders 1 and 2 does not lead to problems with OCP selectivity for fault on neighboring feeder.
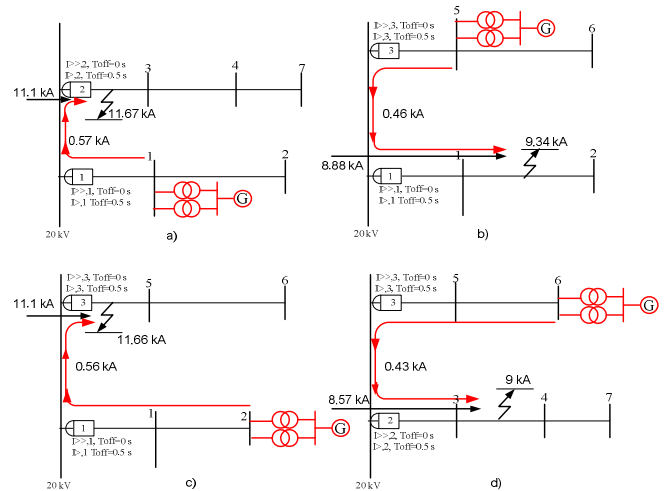
Fig. 4. OCP selectivity with different DG (synchronous generator) location and three phase fault location
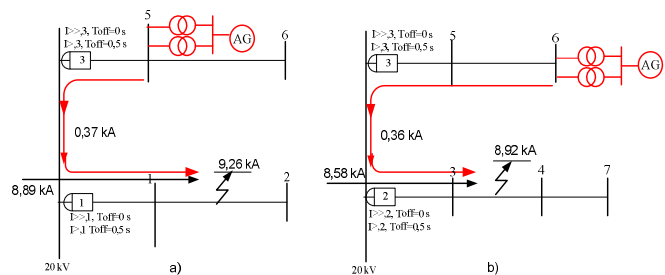
Fig. 5. OCP selectivity with different DG (asynchronous generator) location and three phase fault location

Another problem with DG connection on DN is islanded operation. This type of operation is usually not allowed by distribution companies. Islanded operation occurs in cases when fault is on the same feeder with DG. Using analyzed DN, several cases of possible islanded operation can be simulated with synchronous or asynchronous generator. Fig. 6 depicts two such cases for illustration of possibility and

hazardous nature of islanded operation of DG with synchronous generator.

On Fig. 6a is shown example of islanded operation when fault is happening downstream from DG connection to network ends. Feeder 1 is observed for three phase fault on the half of line 2. The current that flows from the network is very high and it will cause operation of instantaneous I >> OCP of the feeder. However, DG will continue to supply the short circuit despite protection tripping. On Fig. 6b is shown case of islanded work when short circuit occurs upstream from DG connection to network root. Feeder 2 is subject of interest where DG is located in node 3 and three phase fault is on the half of line 3. The current that flows from the network is very high and it will cause operation of instantaneous I >> OCP of the feeder. However, DG will continue to supply the short circuit despite protection tripping.

Fig. 7 illustrates two possible cases of islanded operation of DG with asynchronous generator. On Fig. 7a is shown example of islanded work when for short circuit is located downstream from DG connection to network ends. Feeder 1 is analyzed where DG is connected on node 1 and three phase fault is happening on the half of line 2. The current that flows from the network is very high and it will cause operation of instantaneous I >> OCP of the feeder. However, DG will continue to supply the short circuit despite protection tripping. Same conclusions can be obtained for the case shown on Fig. 7b.
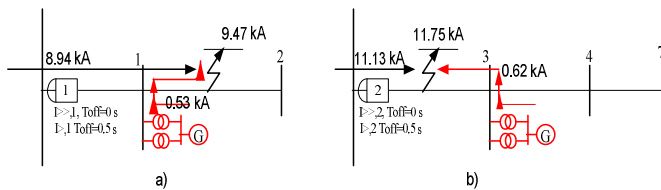
Fig. 6. Simulation of islanded operation of DG with synchronous generator
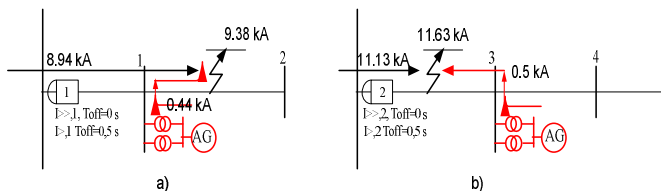
Fig. 7. Simulation of islanded operation of DG with asynchronous generator

According to performed simulations, it can be concluded that DG increases values of short circuit currents in average of 5-10 %.

V. CONCLUSION

DG connection in DN needs serious analysis of existing OCP. It is shown in the paper that DG causes serious problems with distribution feeder protection selectivity, especially when fault is located on neighboring feeder with the one with DG. Protection selectivity problem can be resolve with precise setting. If the problem with selectivity at unidirectional OCP persists, solution is installation of directional OCP, which will operate only for current directions from the feeding substation to network feeders.

Performed analysis on the used test DN, imposes conclusion that if feeder where DG is connected is more loaded in steady state (supplies more HV/MV transformers) in that case is mostly possible will be no problem with protection selectivity for faults on neighboring feeders. This is due to the fact that instantaneous OCP current setting in this case will have high value. Simultaneously, DG connection on heavily loaded feeders is beneficial for other technical conditions in the network, such as: reduction of losses, improvement of voltage profile especially in periods of maximum load, reduction of HV/MV transformer loading etc [5].

Islanded operation of DG as hazardous state can be a case when fault occurs on the same feeder with DG, independently upstream to network root or downstream to its ends from DG location. This state is not allowed by distribution companies. Because of this, for each DG is mandatory to install islanded work protection or so called "loss of mains protection".

Also it can be concluded that DG increase values of short circuit currents. This imposes need for mandatory check of breaking current of commutation equipment and elements dimensioning before DG connection in DN.

REFERENCES

[1] Atanasovski, M., Taleski R., "Power Summation Method for Loss Allocation in Distribution Networks with DG", IEEE Trans., Power Systems, vol. 26, no. 4, pp. 2491-2499, November 2011.
[2] Dragan Popovic, Dusko Bekut, Valentina Treskanica. *Specijalizovani DMS algoritmi,* DMS Group, Novi Sad 2004.
[3] Vladica Mijailovic. *Distributivni izvori energije,* Beograd 2011
[4] NEPLAN User's Guide V5, BCP, www.neplan.ch, (license Faculty of Technical Sciences-Bitola).
[5] Atanasovski, M., Taleski R., "Energy Summation Method for Loss Allocation in Distribution Networks with DG", IEEE Trans., Power Systems, vol. 27, no. 3, pp. 1433-1440, August 2012.

New Calculation Method for Power Loss Allocation in Radial Distribution Networks Without Dispersed Generation

Ljupco Trpezanovski[1], Metodija Atanasovski[2] and Tomce Mijovski[3]

Abstract – **In this paper is presented a new method for calculation the power loss allocation caused by electricity consumers in radial distribution network, without dispersed generation. The method is based on the application of network load-flow calculation results. With known voltage drop and current decomposition of each branch a simple procedure for calculation the complex power losses in the branches is applied. The power loss allocation analyze is conducted on an example of radial network with the proposed new method and other two known methods.**

Keywords – **Radial distribution network, power loss allocation, methods for loss allocation.**

I. INTRODUCTION

In the past two decades a lot of countries performed technical and economic restructuring in their power systems (PS) management. These changes in governing, exploitation and planning of the PS development from technical and economic aspects are known as deregulation in the PS. Nowadays, deregulation enables open market and trading on concurrent base in electricity production and delivery. With implementation of deregulation and market economy the processes for electricity production, transmission and distribution are divided in several, separate independent companies. Each of the participants on the electricity market, as are electricity: producers, consumers, transmission and distribution operators, in certain way participate in creating and covering the costs for the power and energy losses. Deregulation of electricity sector and application of market economy in the management of electricity, actualize the need for calculation the loss allocation and costs covering for energy losses during the energy transmission and distribution. It is a difficult problem to make right and fair allocation of the costs for losses covering, among the all participants in the electricity market. For solving of this problem, significant attention should be pay on defining algorithms and methods which fulfil several conditions. Some of the principals for defining methods for loss allocation [1] are recalled and stressed in the paper. A loss allocation technique should be:

– easy to understand and based on real data of the network;
– carefully designed to avoid discrimination between user;
– able to recover the total amount of the losses;
– consistent with the rules of competitive electricity markets;
– economically efficient, avoiding cross-subsidization between users;
– able to send out economic signals aimed at increasing the efficiency of the network;
– able to provide correct signals concerning the size and location of loads;
– applicable to different situation, e.g. following the time evolution of the generator and load patterns.

Therefore in the past few years several methods for calculation of loss allocation in the networks caused by electricity producers and consumers are reported in the relevant literature. These methods can be divided according two criterions. The first one is according to the type of the network and the methods can be for transmission or distribution networks. The second one is according of the approach on the problem for loss allocation and the methods can be with marginal, average and actually approach.

In this paper a new method for power loss allocation caused only by the loads is presented. This method is with actual approach and is concern for radial distribution networks (RDN) without dispersed generation (DG) sources (explained in section II). The technique of the proposed method is based on the results of the load-flow calculation given in section III. Because during the calculations are used voltage drops and decomposition of the currents in the network branches, the method proposed in sections IV and V, is named as *Voltage Drop Current Decomposition Power Loss Allocation* (ΔUCDPLA). The loss allocation is calculated for a real RDN without DG with the proposed new method and two other methods reported in [4] and [5]. In the section VI, the results of the conducted analysis are compared and shown in table. The conclusions are drawn in the section VII.

II. RADIAL DISTRIBUTION NETWORK PROPERTIES AND MODELLING

For easier understanding and defining the algorithm of the proposed new method for power loss allocation, all line (branch) capacitances are neglected. Thus, overhead and cable RDN lines are presented only with serial impedances. One line diagram of the part with *n* branches and (*n*+1) nodes of middle voltage (MV) level RDN is shown on Fig. 1. The node marked with index 0 is with known voltage magnitude, held on constant value U_0 with HV/MV transformer regulation.

[1]Ljupco Trpezanovski is with the Faculty of Technical Sciences at University St. Kliment Ohridski – Bitola, I. L. Ribar bb. 7000 Bitola, Republic of Macedonia, E-mail: ljupco.trpezanovski@uklo.edu.mk.

[2]Metodija Atanasovski is with the Faculty of Technical Sciences at University St. Kliment Ohridski – Bitola, I. L. Ribar bb. 7000 Bitola, Republic of Macedonia.

[3]Tomce Mijovski is the postgraduate student on the Faculty of Technical Sciences at University St. Kliment Ohridski – Bitola.

Fig. 1. One line diagram of MV radial distribution network without distributed generation.

Phase angle of this voltage is taken 0^0 and it will be a reference angle for the phase angles of the voltages in the rest n nodes of the network. Ordering of nodes and branches indexes is according to the rule of oriented ordering [2]. The nodes are indexed with numbers from 0 to n. The indexes of the branches are marked with circled numbers from 1 to n. Voltage, active and reactive power and current of the load connected in any node, have the same index as the index of the corresponding node. Because the RDN is without DG the directions of the branch currents are uniformly ordered from the node 0 to the loads in other nodes. The branch currents are signed with indexes of two numbers. The first number is equal with the node index where the current enters in the branch (branch beginning node). The second number is equal with the node index where current exits from branch (branch end node), which is the same number of the branch index.

Actually, the loads are connected on the low-voltage (LV) sides on the distribution transformers which MV sides are connected in one of the RDN nodes. If the losses in the distribution MV/LV transformers are neglected it is allowed to consider that the loads are directly connected on the MV nodes. This approximation and way of loads connection on the RDN are shown in the left bottom part on Fig. 1.

The following quantities are known for the network from Fig. 1:

− voltage in the MV node 0 in the substation: $\underline{U}_0 = U_0 \angle 0^0$;

− serial branch impedances: $\underline{Z}_g = R_g + jX_g$, g=1, 2,....., n;

− apparent power of the loads: $\underline{S}_i = P_i + jQ_i$, i=1, 2,....., n.

If there are no loads in some nodes that nodes can be treated as the nodes with connected loads with zero power.

For each node i (i=1, 2, ..., n) there is a set of branches $\{G_{0i}\}$ forming the path through which starting from the node 0, the node i can be reached. Through the branch g (g=1, 2, ..., n) flow the current of the load connected at the branch end node i (according to oriented ordering i=g) and currents of the loads connected in the nodes with index $i > g$ which can be reached on forward direction through the branch g. The set of these nodes for the branch g is signed as $\{J_{gi}\}$.

III. LOAD-FLOW CALCULATION FOR RDN

The proposed method for power loss allocation is based on the load-flow calculation results. For a working scenario with known values of the branch serial impedances and load powers, taking into account that the voltage U_0 is also known, a load-flow calculation can be conducted for RDN. As a very proper load-flow method for radial networks is taken *Current Summation Method* (CSM) [3]. This is an iterative method based on multi-calculations "forward" and "backward" until the accuracy of calculated values is achieved. The procedure "backward" starts from the node with index n and goes down to the node with index 0. During this procedure the currents of the loads and currents in all branches are calculated. The procedure "forward" starts from the node with index 0 and goes up to the node with index n. During this procedure the voltages from new iteration in all nodes are calculated. When the accuracy of the values of node voltages is achieved the calculation procedure stops. The results obtained by CSM load-flow analyse are:

− node voltages in RDM: $\underline{U}_i$ for i=1, 2, …, n;

− load currents in the nodes: $\underline{I}_i$ for i=1, 2,....., n and

− total branch currents for branches g=1, 2,....., n:

$$\underline{I}_{pg} = \sum_{i\in\{J_{gi}\}} \underline{I}_i , \qquad (1)$$

where p ($p<g$) is the beginning node of the branch g and $i \in \{J_{gi}\}$, for nodes $i \geq g$ which can be reached in forward direction through the branch g.

IV. POWER LOSS ALLOCATION IN BRANCHES

With known voltages in all nodes i=0, 1, 2, …, n of RDM, the voltage drop $\Delta \underline{U}_g$ on each branch g=1, 2,….., n can be easy calculated. Therefore, for the branches from Fig. 1, the voltage drops are: $\Delta \underline{U}_1 = \underline{U}_0 - \underline{U}_1$ for branch 1, $\Delta \underline{U}_2 = \underline{U}_1 - \underline{U}_2$ for branch 2 and generally for any branch k with the beginning node i:

$$\Delta \underline{U}_k = \underline{U}_i - \underline{U}_k, \quad k>i. \tag{2}$$

Using the branch g current $\underline{I}_{pg}$ and voltage drop obtained by Eq. (2), the complex power loss $\Delta \underline{S}_g$ for any branch g=1, 2,….., n is obtained by the Eq. (3):

$$\Delta \underline{S}_g = \sqrt{3}\Delta \underline{U}_g \cdot \underline{I}_{pg}^*. \tag{3}$$

If the branch current in Eq. (3) is expressed by Eq. (1), than the total power losses in the branch g are expressed as:

$$\Delta \underline{S}_g = \sqrt{3}\Delta \underline{U}_g \cdot \left(\sum_{i \in \{J_{gi}\}} \underline{I}_i \right)^*, \tag{4}$$

or:

$$\Delta \underline{S}_g = \sum_{i \in \{J_{gi}\}} \left(\sqrt{3}\Delta \underline{U}_g \cdot \underline{I}_i^* \right). \tag{5}$$

In the sum of Eq. (5) the first addend for i=g represents the power losses caused by the current $\underline{I}_g$ which is the current of the load connected in the node g. This addend is expressed as:

$$\Delta \underline{S}_g^{(g)} = \sqrt{3}\Delta \underline{U}_g \cdot \underline{I}_g^*, \quad g \in \{J_{gi}\} \tag{6}$$

where with the superscript (g) is marked that the losses are caused by the current from load connected in node g. The rest addends in the sum of Eq. (5) represent the losses caused by currents $\underline{I}_i$ of other loads $\underline{S}_i$ connected in nodes i and supplied with electricity through the branch g. In generally these addends are given with Eq. (7):

$$\Delta \underline{S}_g^{(i)} = \sqrt{3}\Delta \underline{U}_g \cdot \underline{I}_i^*, \quad i \in \{J_{gi}\}, \quad i>g. \tag{7}$$

The conjugate values of the load currents in Eqs. (6) and (7) can be expressed through the node voltages (where the loads are connected) and corresponding load powers. According to the suggested replacements the power loss allocation in branch g, caused by the load connected at the end of branch g

or any of the rest loads connected in nodes i supplied through this branch, the Eqs. (6) and (7) are given in form:

$$\Delta \underline{S}_g^{(i)} = \frac{\Delta \underline{U}_g}{\underline{U}_i} \cdot \underline{S}_i = \frac{\Delta \underline{U}_g}{\underline{U}_i} \cdot (P_i + jQ_i), \quad i \in \{J_{gi}\}, \quad i \geq g. \tag{8}$$

The total power losses in any branch g of the RDN, can be calculated as a sum of power losses in that branch allocated to the load connected to the end of the branch and other loads which are supplied through concerning branch, or as:

$$\Delta \underline{S}_g = \sum_{i \in \{J_{gi}\}} \Delta \underline{S}_g^{(i)}, \quad g=1, 2, …, n. \tag{9}$$

From above explanations it is obvious that power loss allocation calculations for any branch of the RDN without DG, can be easy conducted with known node voltages and load powers.

V. NEW POWER LOSS ALLOCATION METHOD

The total power losses in the RDN: $\Delta \underline{S}_N = \Delta P_N + j\Delta Q_N$ are obtained as a sum of the power losses in each branch calculated by Eq. (9). With Eqs. (10) are presented apparent, active and reactive power losses, separately for the entire network:

$$\Delta \underline{S}_N = \sum_{g=1}^{n} \Delta \underline{S}_g,$$

$$\Delta P_N = Re\left(\sum_{g=1}^{n} \Delta \underline{S}_g \right), \quad \Delta Q_N = Im\left(\sum_{g=1}^{n} \Delta \underline{S}_g \right). \tag{10}$$

Obtaining the total power loss allocation in the entire network, separately from each load connected in node i and signed as $\Delta \underline{S}_i^{\Sigma} = \Delta P_i^{\Sigma} + j\Delta Q_i^{\Sigma}$, can be performed on two ways. The first way is for summation of power loss allocation by the load i in all branches $g \in \{G_{0i}\}$, applying the equations:

$$\Delta \underline{S}_i^{\Sigma} = \sum_{g \in \{G_{0i}\}} \Delta \underline{S}_g^{(i)}, \quad \Delta P_i^{\Sigma} = Re\left(\sum_{g \in \{G_{0i}\}} \Delta \underline{S}_g^{(i)} \right),$$

$$\Delta Q_i^{\Sigma} = Im\left(\sum_{g \in \{G_{0i}\}} \Delta \underline{S}_g^{(i)} \right), \quad \text{for } i=1, 2,….., n. \tag{11}$$

According to the second way, the total power loss allocation in the entire network, separately from each load connected in node i can be calculated directly with Eqs. (12), applying voltage drops on all branches $g \in \{G_{0i}\}$, voltage and power(s) of the load connected in node i:

$$\Delta \underline{S}_i^{\Sigma} = \frac{\underline{S}_i}{\underline{U}_i} \cdot \sum_{g \in \{G_{0i}\}} \Delta \underline{U}_g = \frac{(P_i + jQ_i)}{\underline{U}_i} \cdot \sum_{g \in \{G_{0i}\}} \Delta \underline{U}_g. \tag{12}$$

If Eq. (12) divides on real (*Re*) and imaginary (*Im*) parts, the total active and reactive power loss allocation by the load in node *i* are:

$$\Delta P_i^\Sigma = Re\left(\frac{\underline{S}_i}{\underline{U}_i} \cdot \sum_{g \in \{G_{0i}\}} \Delta \underline{U}_g \right),$$

$$\Delta Q_i^\Sigma = Im\left(\frac{\underline{S}_i}{\underline{U}_i} \cdot \sum_{g \in \{G_{0i}\}} \Delta \underline{U}_g \right), \text{ for } i=1, 2,..., n. \quad (13)$$

Usually, for practical analysis the interest is focused on active power loss allocation of each load, expressed in percentage of the total active power loss in the RDN. These values are calculated with the Eq. (14):

$$\Delta p_i^\Sigma = \frac{\Delta P_i^\Sigma}{\Delta P_N} \cdot 100 [\%], \text{ for } i=1, 2,..., n. \quad (14)$$

According to the used quantities and procedures for new method technique definition, the method is named as *Voltage Drop Current Decomposition Power Loss Allocation* or shortly ΔUCDPLA.

VI. APPLICATION OF THE NEW ΔUCDPLA METHOD

The proposed new method is tested on real overhead radial distribution network [2], shown on Fig. 2. This test RDN consists of 5 lines (branches), 6 nodes (including the node with index 0) and 5 loads. All lines are from the same bare conductor Al/Fe 35/6 with $r = 0,8353$ Ω/km and $x = 0,37$ Ω/km. The line lengths are: $l_1 = 2,5$ km, $l_2 = 6,0$ km, $l_3 = 0,4$ km, $l_4 = 0,5$ km and $l_5 = 0,6$ km. Voltage in node 0 is $\underline{U}_0 = (10,25 + j0)$ kV.
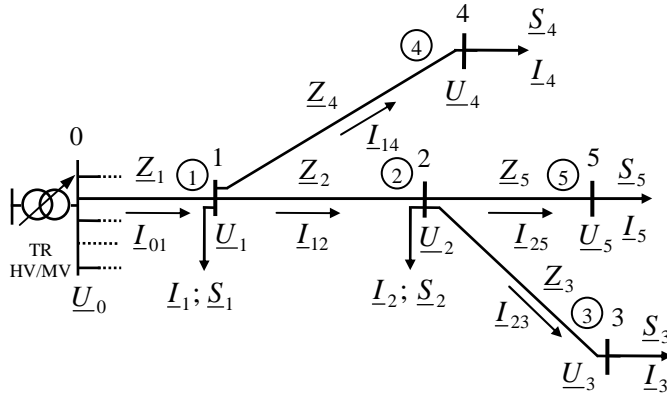
Fig. 2. Middle voltage test RDN without DG.

The maximum values in kVA are taken for the powers of the node loads. The powers are: $\underline{S}_1 = (400+j130)$, $\underline{S}_2 = (180+j60)$, $\underline{S}_3 = (140+j40)$, $\underline{S}_4 = (200+j70)$, $\underline{S}_5 = (100+j30)$. Three different methods are used for power loss allocation calculations on the test RDN from Fig. 2. The first one is *Power Summation Method for Loss Allocation* (PSMLA) [4], the second is *Branch Current Decomposition Loss Allocation* (BCDLA) [5] and the last is the proposed method ΔUCDPLA in section V.

Some of the obtained loss allocation results conducted by three methods are given in Table I. In the first column are assigned the node number with proper connected load. In other columns for each method and each node (load) active power loss allocation results in absolute value in kW and in % are given. The last row shows total power loses in RDN.

TABLE I
POWER LOSS ALLOCATION FOR THE LOADS OF TESTED NETWORK.

Method	PSMLA		BCDLA		ΔUCDPLA	
Node *i* - Load *i*	ΔP_i^Σ [kW]	Δp_i^Σ [%]	ΔP_i^Σ [kW]	Δp_i^Σ [%]	ΔP_i^Σ [kW]	Δp_i^Σ [%]
$1 - \underline{S}_1$	13,644	38,97	9,509	27,17	9,508	27,17
$2 - \underline{S}_2$	8,964	25,60	8,788	25,11	8,725	24,93
$3 - \underline{S}_3$	5,387	15,39	6,820	19,49	6,902	19,72
$4 - \underline{S}_4$	4,428	12,65	4,982	14,24	4,933	14,10
$5 - \underline{S}_5$	2,587	7,39	4,896	13,99	4,926	14,08
Σ loss	35,01	100	34,995	100	34,994	100

Comparing the results in Table I, it can conclude that the results obtained with the new method are correct.

VII. CONCLUSION

In this paper is presented a new effective, robust and fast method for power loss allocation in RDN without DG sources. The method is based on load-flow results application and procedure for branch current decomposition on current addends which belong to the loads. Method application is demonstrated on a test RDN. The obtained results for power loss allocation with a proposed new method are verified through their comparison with the results obtained with already published methods. Furthermore, this method can be a base for establishing a new energy loss allocation method.

REFERENCES

[1] P. M. Costa, M. A. Matos, "Loss Allocation in distribution networks with embedded generation", *IEEE Transactions on Power Systems*, vol. 19, no. 1, pp. 384-389, Feb. 2004.

[2] Д. Рајичиќ, *Вовед во дистрибутивни електроенергетски системи*, Скопје, Универзитет Св. Кирил и Методиј – Скопје, Електротехнички факултет, 1995.

[3] D. Shirmohammadi, H. W. Hong, A. Semlyen, G.X. Luo, "A Compensation-Based Power Flow Method for Weakly Meshed Distribution and Transmission Networks". *IEEE Transactions on Power Systems,* Vol. 3, No. 2, pp. 753-762, 1988.

[4] M. Atanasovski, R. Taleski, "Power Summation Method for Loss Allocation in Radial Distribution Networks With DG". *IEEE Transactions on Power Systems*, Vol. 26, No. 4, pp. 2491–2499, November 2011.

[5] E. Carpaneto, G. Chicco, J. S. Akilimali. "Branch Current Decomposition Method for Loss Allocation in Radial Distribution Systems With Distributed Generation". *IEEE Transactions on Power Systems,* Vol. 21, No. 3, pp. 1170-1179, August 2006.

Wind Turbine Fault Ride Through in Weak Power System

Blagoja Stevanoski[1] Nikola Acevski[2] and Metodija Atanasovski[3]

Abstract – **This paper deals with dynamic characteristics of the first DFIG wind farm which is connected to the transmission power system of Republic of Macedonia. Matlab/Simulink software is used to present the fault ride-through ability of the wind farm. The DFIG wind turbine and the power system are modelled with real characteristics, like in the initial project. Results from tests short circuit on 110 kV voltage level buses will be simulated in order to define how the DFIG will react under these circumstances and which actions should be undertaken in the power system to improve power system stability.**

Keywords – **DFIG, wind turbine, dynamic characteristics, fault ride-through, grid code, Macedonian power system**

I. INTRODUCTION

Wind power has widely proved to be one of the most competitive and efficient renewable energy sources with a most favorable technical and economic prospects. This is due to the existence of non-exploited wind resources and to the fact it is a clean energy with a reduced cost of installation and maintenance. Wind turbines, which typically are centralized in wind farms, are constantly planned and commissioned. The produced electrical power from wind is steadily increasing and large wind generation units have capacities comparable to conventional power plants. As a consequence, wind power has reached significant influence on the power production and penetration levels imposing new challenges to the Transmission System Operators (TSO) [11]. The process of high wind energy penetration requires the impact analysis of this new technology in power systems. Impact of wind energy on power systems is related to security, stability, power quality and operation of power system. As a consequence grid operators require them to participate in grid voltage support in steady state as well as during faults. Many grid codes contain dynamic requirements like the fault ride through (FRT) ability of generating units. Additionally, voltage support during grid fault due to the injection of short time reactive current is requested in certain grid codes. FRT capability and voltage support allow a secure operation of the power system.

Among the several wind energy technologies, a generation of variable speed wind turbines present many advantages compare to the fixed speed wind turbines. These wind energy conversion systems are connected to the grid through voltage source converters to make variable-speed operation possible. The doubly fed induction generator (DFIG) wind turbines are nowadays more widely used especially in large wind farms. The main reason for the popularity of the DFIG is their competitive cost and performance and ability to supply power at constant voltage and frequency while rotor speed varies.

Following the fact that most country intend to use renewable energy sources in their power systems, our case study came as result of the initial project of implementing wind power energy into the electrical power system of Republic of Macedonia [13]. In the initial project, authorities have planned to implement DFIG because of the fact that it will drastically decrease the costs for investors, it will enable control of the torque and as last, it will increase the efficiency of wind extraction. Therefore, the paper includes dynamic simulation results of DFIG-based wind farm connected to a weak power system (the installed capacity of the Macedonian power system is about 1350 MW) during grid disturbances. The simulations are made by using the simulation platform Matlab/Simulink utilizing its SimPowerSystem toolbox.

II. WIND TURBINE MODEL DESCRIPTION

Several generator types are in use for wind power applications today. The main distinction can be made between fixed speed and variable speed wind generator types. The most widely used variable speed wind generator concept is doubly fed induction generator. DFIG can combine fast control due to the high switching frequency of the power electronics and moderate costs of converter system [2].

The DFIG is a wound rotor induction generator with a voltage source converter connected to the slip rings of the rotor [10]. The number of pole pairs varies between two and three. Therefore slow speed wind turbine is coupled to the induction generator through a mechanical shaft system, which consists of a low-speed and high-speed shafts and gearbox between. The stator winding is coupled directly to the grid and the rotor winding is connected to the grid via AC/DC/AC converter. The mechanical power generated by the wind turbine drives the DFIG, which feeds electrical power into the grid through the stator and rotor windings. In order to produce electrical power at constant voltage and frequency to the utility grid over a wide operation range, the power flow between the rotor winding and the grid must be controlled. Therefore, there are two PWM converters: a rotor side converter (RSC) and a grid side converter (GSC) connected back-to-back by a dc link capacitor. The converters allows controlling the amplitude, frequency and phase angle of the rotor voltage. This enable variable speed operation of the DFIG, which can be used to adopt the generator speed according wind speed to increase the wind power utilization . The speed range of the generator is about ± 30% of the synchronous speed. Thus the speed of the generator is

[1]Blagoja Stevanoski is with the Faculty of Technical Sciences, Makedonska falanga 33, Bitola 7000, Macedonia, E-mail: blagoj.stevanovski@uklo.edu.mk.

[2]Nikola Acevski is with the Faculty of Technical Sciences, Makedonska falanga 33, Bitola 7000, Macedonia, E-mail: nikola.acevski@uklo.edu.mk.

[3]Metodija Atanasovski is with the Faculty of Technical Sciences, Makedonska falanga 33, Bitola 7000, Macedonia, E-mail: metodija.atanasovski@uklo.edu.mk

decoupled from grid frequency [5].

The voltage source converters are usually equipped with Insulated Gate Bipolar Transistors (IGBT) in a standard three-phase bridge configuration. The IGBT are controlled by pulse width modulation signals from a digital signal processor. The rotor circuit and so the converter circuit can be dimensioned for 20-30% nominal active power of wind turbine. The power flow through the rotor can be bidirectional. The power flows into the rotor when the wind turbine operates at sub synchronous speed, with low mechanical input power. The rotor power flow reverses at super-synchronous speed. Thus, with high mechanical input power, part of power is fed to the grid through the stator and part through the back-to-back converter. For protecting the components against overvoltage's and overcurrent's a DC-link chopper and a rotor crowbar are included. The crowbar circuit is used to short circuit the rotor side convertor in order to protect it from over current in the rotor during transient disturbances. The chopper protects the DC-link against excessive voltage following grid faults.

The operation of the DFIG wind turbine is regulated by a control system, which generally consists of two parts: the electrical control of the DFIG and the mechanical control of the wind turbine blade pitch angle. Control of the DFIG is achieved by controlling the RSC and GSC, as shown in Fig. 1.
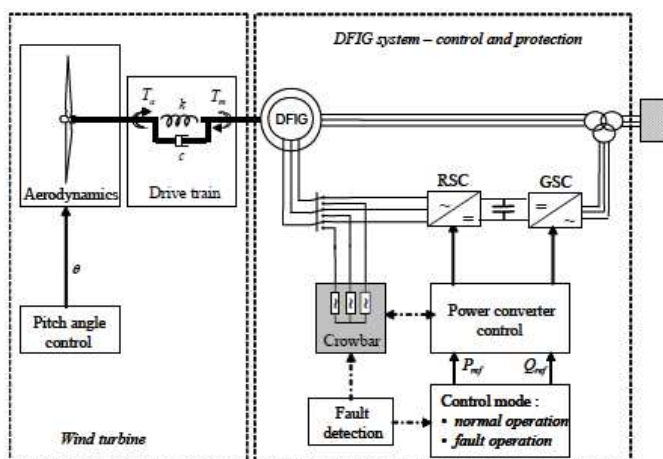

Fig. 1. DFIG wind turbine configuration and control

A. Wind turbine during grid fault

At the occurrence of a fault in the network, high currents are induced in a DFIG. Therefore, the RSC switches must be blocked to avoid their damage and at the same time DFIG wind turbine can be protected by using a different methods. A standard solution is short-circuiting the rotor circuit through an external resistance called crowbar resistance. Modern DFIG wind turbine can also be protected with a DC chopper. At the event of a short-circuit, the RSC switches are blocked and the rotor current is led into the DC-link capacitor through anti-parallel diodes of the RSC. In this way the rotor circuit is rapidly demagnetized and the RSC can be restarted when the

rotor current and DC-link voltage decrease below a certain value.

The high crowbar resistance causes the ac component of the symmetrical short-circuit current to decay more rapidly and after some periods, the short-circuit current of a DFIG is made up predominantly of a dc component. Someone authors have been proposed to calculate the symmetrical short-circuit current in the same manner as done for a SCIG (squirrel cage induction generator), but incrementing the value of the rotor resistance in the expression for the rotor transient time constant by the value of the crowbar resistance. When the RSC is reconnected it control the stator current and DFIG may be looked as a constant current source.

The RSC controls independently the active and reactive power injected by the DFIG into the grid in a stator flux dq-reference frame [10]. The q-axis current component is used to control the active power using a maximum power tracking. The d-axis current component is used to control the reactive power exchanged with the grid, which in normal operation is set to zero. In case of disturbance, if induced current in the rotor circuit is not high enough to activate the over-current protection, the RSC is set to inject reactive power into to grid in order to support the voltage restoration.

The objective of the GSC is to maintain the voltage at the DC link between both converters. In case of disturbance, the GSC is set to inject reactive power into the grid, whether the RSC is blocked or is kept in operation. As for the RSC, the control of the GSC is performed using d and q-axis current, but instead of rotating with the stator flux, the axis rotates with the grid voltage.

III. GRID FAULT RIDE-THROUGH REQUIREMENTS

The fault ride-through requirements has been imposed in order to avoid significant loss of wind turbine production in the event of grid faults. At the beginning, wind turbine were only required to disconnect from the grid when a grid fault was detected. However, with the increased capacity of the wind turbines in the power system, such a disconnection of wind turbines could generate problems in the control of frequency and voltage in the system, and as worst case a system collapse. Therefore, the large increase in the installed wind capacity in transmission systems necessitates that wind generation remains in operation in the case of network disturbances. For avoiding scenarios where generators are disconnected during grid faults, transmission system operators have developed voltage and time duration profiles that define requirements of a generator to "ride through" grid faults without disconnection [18]. Voltage dips following a fault in the system are expected to be above this curve. The duration of the voltage dip is dependent on the speed of the protection system. Since TSO in different countries may have different protection rules, the grid fault ride-though (GFRT) requirements are specific to each national grid code. The voltage-time curves from Macedonian national grid code is shown in Figure 2. A wind turbine is required to remain connected to the grid if the voltage at the point of common coupling (PCC) during fault always remains above the voltage-
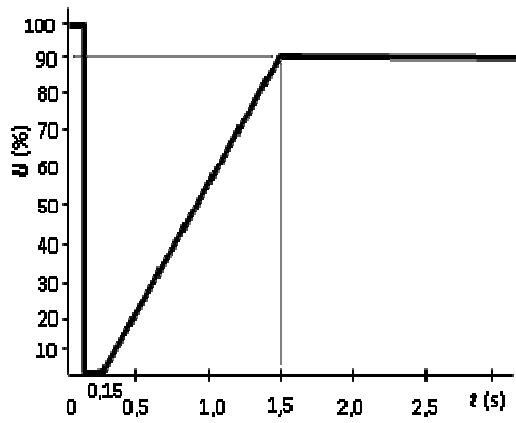
Fig. 2 Voltage requirement at PCC during a fault

time curve given in the grid code. The 150 ms delay shown in the figure accounts for the normal operating time of protection relays. Besides remaining connected, a wind turbine may also be required to provide reactive current support to the grid during voltage dip, to support and faster restore the grid voltage. Wind turbine must support voltage with additional reactive power during voltage collapse. Voltage regulator must, within 40 ms after fault recognition, to deliver reactive power at low voltage side of the block transformer of at least 2% of the rated current for each percent of the voltage dip. If necessary, it should be possible to deliver reactive power to at least 100% of the nominal current [11].

IV. SIMULATION

The simulation performed in this work concern the first wind farm in Macedonia, located near southeastern town of Bogdanci [15]. The rated power capacity of the wind farm is 36,8 MW which is installed at first phase. There are 16 wind turbines (Siemens SWT-2,3-93), each one of them with installed power of 2,3 MW equipped with doubly fed induction generators. The wind farm is connected through step up transformers 20/110 kV followed by a classical 110 kV AC overhead line to a transmission network. For the studies the point of common coupling (PCC) is defined at a busbar of the connected grid. In this study case a simplified Simulink model of the transmission power system is used with appropriate short circuit power. A block diagram of the model used in this simulation is presented on Fig. 3.
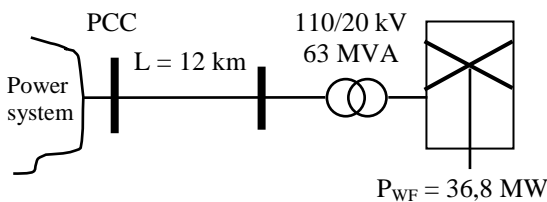
Fig. 3 Wind farm connection on power system

Two scenarios have been performed in order to investigate how wind farm grid ride-through in case of near fault and fault below transmission network.

A. GFRT for fault at the point of common coupling

The first test was made under symmetrical and unsymmetrical short circuit with a duration of 150 ms on the 110 kV bus near the wind farm, at the point of common coupling (PCC). It is assumed that the wind turbine generators remain connected to the network during the whole fault duration. The turbines works in two modes. The first mode is VAR Regulation while the other mode is the Voltage regulation. As first test the wind farm works in Voltage regulation mode. The three phase short circuit is applied at first second of simulation and lasts until 1.15 sec. Simulation with different voltage controller gain were made and the effect of the voltage support of the wind farm during the grid disturbances at the PCC is analyzed. The transient behavior of the voltage and reactive power are shown in figure 4 and 5.
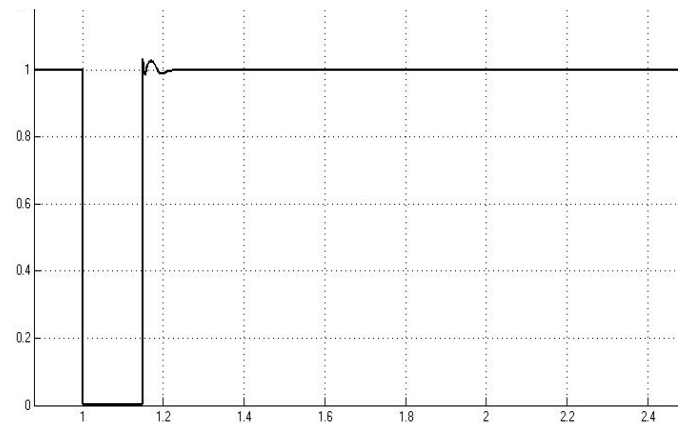
Fig. 4 Voltage at PCC during three-phase fault
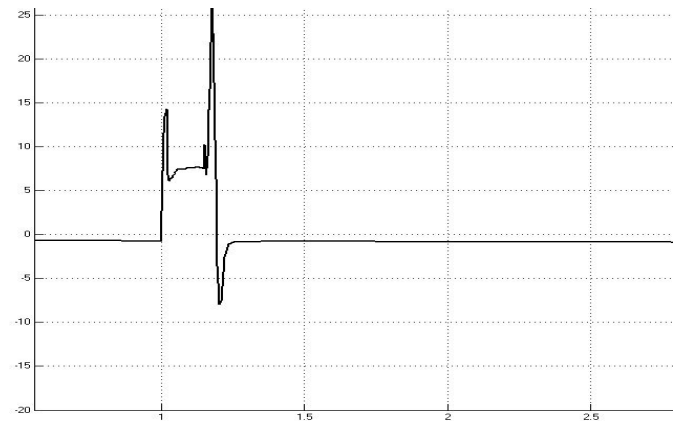
Fig. 5 Reactive power (MVAr) injected by wind farm

B. GFRT for remote fault

The remote fault below on transmission system is simulated by applied voltage dip of 70% at the busbar connecting wind farm to the grid for 300 ms. Results of simulations concern voltage at PCC and reactive power support are shown in figure 6 and 7.
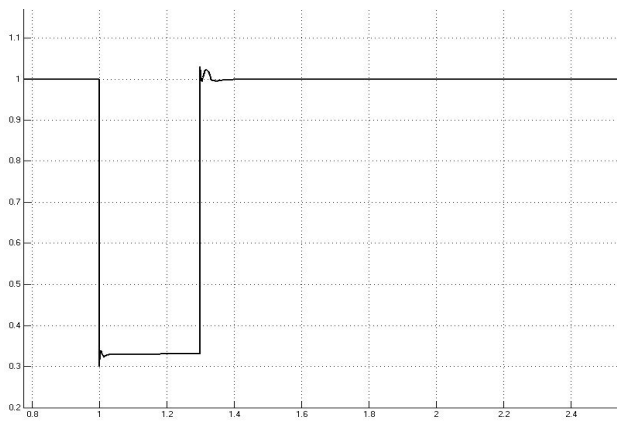
Fig. 6 Voltage at PCC during remote fault
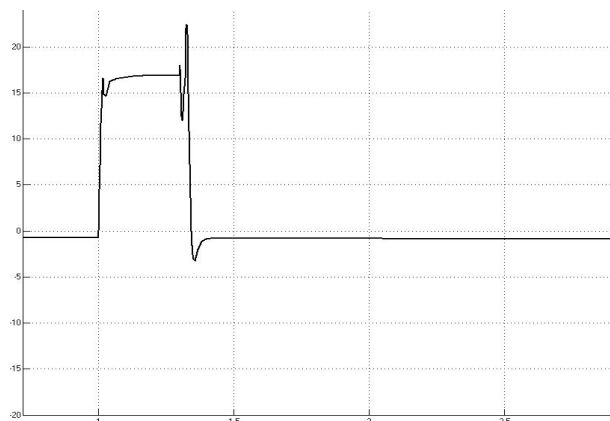
Fig. 7 Reactive power (MVAr) injected by wind farm during remote fault

V. CONCLUSION

High penetration of wind turbines imposes a big challenge to the safe operation of power system. This work has shown possibilities and limitations of DFIG concerning FRT capability during grid fault. Simulation have shown that the wind turbines are able to provide a considerable contribution to grid voltage support during short circuit periods by using their fast voltage control without additional compensation devices. The reactive power injected by the wind farm would be enough to accomplish grid code requirements and in accordance with technical requirements prescribe by the Macedonian transmission system operator (MEPSO). The Voltage regulation mode is more tolerant unlike the other mode, where the turbine trips and disconnects from the network.

REFERENCES

[1] V. Akhmatov, "Analysis of dynamic behavior of electric power systems with large amount of wind power", Ph.D. Thesis, TU of Denmark, 2003.

[2] F. Sulla, "Fault Behavior of Wind Turbines", PhD dissertation, Lund University, Lund, Sweden, 2012.

[3] S. M. Bolik, "Modelling and Analysis of Variable Speed Wind Turbines with Induction Generator during Grid Fault", Aalborg University, Denmark, 2004.

[4] I. Erlich, H.Wrede, C. Feltes,"Dynamic Behavior of DFIG-Based Wind Turbines during Grid Faults", IEEE, pp. 1195-1200, 2007

[5] T. Neumann, C. Feltes, I. Erlich,"Response of DFIG-based Wind Farm Operating on Weak Grid to Voltage Sags", IEEE, 2011.

[6] A. Perdana, "Dynamic models of Wind Turbines", PhD dissertation, Chalmers University of Technology, Götegorg, Sweden, 2008.

[7] T. Sun, "Power quality of Grid-Connected Wind Turbines with DFIG and Their Interaction with the Grid", PhD dissertation, Aalborg University, Denmark, 2004.

[8] J. T. G. Pierik, Y. Zhou and P. Bauer, "Wind farm as Power Plant", TU Delft, 2008.

[9] B. Babypriya and R. Anita, "Modelling, Simulation and Analysis of Doubly Fed Induction Generator for Wind Turbines", Journal of Electrical Engineering, Vol. 60, No2, pp. 79-85, 2009.

[10] C. Rahmann, H.J. Haubrich, L. Vargas and M.B.C.Salles, "Investigation of DFIG with Fault Ride-Through Capability in Weak Power Systems", IPST2009, Kyoto, Japan, 2009.

[11] KEMA, "Study for integration of wind power plants in the Macedonian transmission system", Macedonian ministry of economy, 2011.

[12] F. Iov, A.D. Hansen, P. Sorensen and F. Blaabjerg, " Wind Turbine Blockset in Matlab/Simulink", RISO&Aalborg University, Denmark, 2004.

[13] B.Chitti Babu and K.B. Mohanty, "Doubly-Fed Induction Generator for Variable Speed Wind Energy Conversion Systems-Modeling& Simulation", International Journal of Computer and Electrical Engineering, Vol. 2, No.1, pp. 141-147, 2010.

[14] A.D. Hansen and G. Michalke, "Voltage grid support of DFIG wind turbines during grid faults", EWEA, pp. 93-97, 2007.

[15] ***, "Bogdanci Wind Park Project", Main Design, Siemens Energy Sector, Wind Power Division, 2012.

[16] K. Singh, C. Das, "Analysis of DFIG Wind Turbine System during Different Types of Grid Fault", International Journal of Innovative Technology and Exploring Engineering 2014.

[17] M.M. Rojas, A. Sumper, O.G. Bellmunt, R.V. Robles, E.V. Naranjo "Comparison of interconnection and operation requirements for wind farms", *Wind Expo LAWEA*, 2008.

[18] E.ON Netz "Grid Code: high and extra high voltage", *Technical report, E.ON Netz*, 2006

Analysis of Hydrology as Input Parameter for Financial Income Calculation from Hydro Power Plant

Goce Bozinovski [1] and Atanas Iliev[2]

Abstract – **The basic idea of this paper is to provide statistical analysis of the input parameters of hydrology through a series of hydrological inflow data for a long period of years with monthly distribution. The hydrological data is for the last 45 years (from 1970 until 2014) which covers extremely wet and extremely dry hydrological periods. This is important in order to get relevant statistics of expected values with appropriate probabilities of occurrence. The proposed methodology is applied on the real example: HPP Spilje which is part of the hydro power system of Crn Drim river in the Republic of Macedonia.**

Keywords – **Statistical analysis, HPP, probabilities, water inflow, expected values, probabilities of occurrence.**

I. INTRODUCTION

Hydroelectric power plants (HPP) are high investment projects that require serious and comprehensive analysis in terms of energy benefits and financial effects of the project. Hydrological data and especially water inflows of the basin are the most important parameters in energy analysis, as well as technical parameters of HPP units. This affects on the expected production of electricity and on the benefits obtaoined from electricity generated in the hydro power plant.

The results from analysis can be applied in determining the expected hydropower production, or appropriate expected revenue of the electricity generated from the HPP.

II. HYDROLOGICAL ANALYSIS OF HPP SPILJE

The hydro power system of Crn Drim consists of two hydro power plants HPP Globocica and HPP Spilje and three water reservoirs Ohrid lake, Globocica lake and Debar lake. The system is in operation in the last 50 years. The paper is focuses on hydrology data for HPP Spilje, which is on Debar lake with water accumulation of 210 million m^3. The installed turbine flow for each of the 3 units is 36 m^3/s, or the total turbine flow is 3x36 m^3/s = 108 m^3/s. The average gross head is 90 m, and the total installed power of the generators are 3x28MW = 84 MW. Average yearly production of electricity from HPP Spilje is 300 GWh.

Based on the hydrological data for the last 45 years in the Debar Lake which is the reservoir for HPP Spilje, the average annual value of the water inflow is 47,48 m^3/s.

[1]Goce Bozinovski is with the AD ELEM-Skopje, Branch HES "Crn Drim"-Struga, Plostad na revolucijata bb, 6330 Struga, p.fax 83, R. Macedonia, e-mail: goce_boz@yahoo.com
[2]Atanas Iliev is with the Faculty of Electrical Engineering and Information Technologies – Skopje, Karpos II, 1000 Skopje, R. Macedonia,e-mail: ailiev@feit.ukim.edu.mk
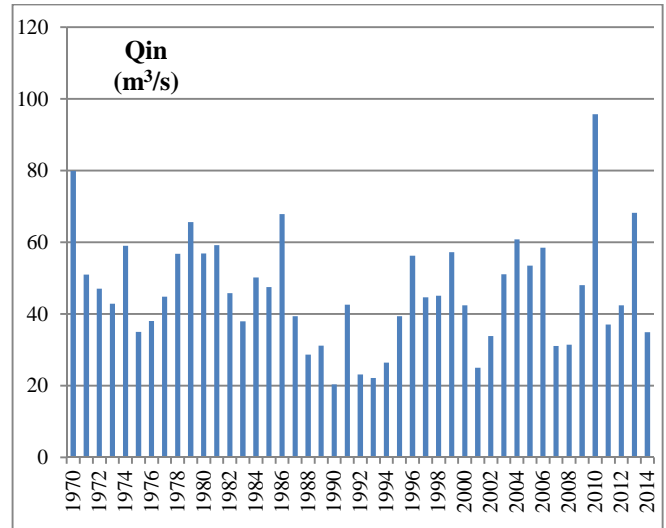
Fig. 1. Average yearly water inflow for the period 1970-2014 in Debar lake

Factor of water inflow is the ratio of average yearly inflow of water and the average inflow for the whole range of years (T).Factor of water inflow for each year (t) is calculated by the relation:

$$K_{in,year}(t) = \frac{Q_{year}(t)}{Q_{year}^{averager}} = \frac{Q_{year}(t)}{\frac{1}{T}\sum_{t=1}^{T} Q_{year}(t)} \qquad (1)$$

Factor of generating energy (production factor), is the ratio between the annual production for a year and the average production of a whole range of years (T). Factor of generating energy for each year can be obtained from the relation:

$$K_{gen,year}(t) = \frac{W_{year}(t)}{W_{year}^{average}} = \frac{W_{year}(t)}{\frac{1}{T}\sum_{t=1}^{T} W_{year}(t)} \qquad (2)$$

The coefficient of correlation between water inflow factor and production of electricity factor for the period 1970-2014 is 0,974. Therefore, it indicates a strong positive linear correlation between water inflow and electricity production. The average total water inflow in a year is approximately 6,5 times larger than accumulation space of Debar lake. HPP Spilje is the seasonal type of the hydro power plant and presents the indicator for possibility of additional unit(s) for installation of the power house.

III. PROBABILITY OF WATER INFLOW AND GENERATION OF ELECTRICITY

According to the monthly water inflow data for the period of 45 years (1970-2014), as well as from data of HPP operating regime (turbine discharge and electricity generation), we can obtain probabilistic parameters.

The main parameters for analysis are:
- Q_{in} (m^3/s) - water inflow (monthly or in a year)
- W_{in} (GWh) - inflow energy corresponding to the water inflow Q_{in}
- Q_{tur}(m^3/s) - turbine water discharge (monthly or a year)
- W_{gen} (GWh) - generated energy corresponding to the turbine discharge Q_{tur}

According the data for the parameters of 45 years for each month (45x12=540), the probability of having the expected values in the intervals between minimum and maximum values are shown in Tab.1.

$$P (X_{min}<X<X_{max}), \text{ where } X=(Q_{in}, W_{in}, Q_{tur}, W_{gen}) \quad (3)$$

TABLE I
PROBABILITY P(X_{MIN}<X<X_{MAX}) FROM THE DATA FOR THE PERIOD OF 45 YEARS (1970-2014)

Probability of hydrology inflow Water inflow and inflow energy			
0<Qin<20	0.00	0<Win<10	0.13
20<Qin<30	0.34	10<Win<20	0.39
30<Qin<40	0.19	20<Win<30	0.26
40<Qin<60	0.25	30<Win<40	0.09
100<Qin<60	0.16	40<Win<50	0.07
100<Qin<200	0.06	50<Win<120	0.06

Probability of operation regimes Turbine discharge and generated energy			
0< Qtur <20	0.11	0< Wgen <10	0.12
20< Qtur <30	0.17	10< Wgen <20	0.35
30< Qtur <40	0.21	20< Wgen <30	0.29
40< Qtur <60	0.28	30< Wgen<40	0.13
60< Qtur <80	0.16	40< Wgen <50	0.08
80< Qtur<120	0.07	50< Wgen <100	0.02

The top side of the table shows the probability of inflow parameters as the water inflow and inflow expected energy corresponding to the water inflow. The inflow parameters are the main drive for the operating regimes for each hydro power plant. HPP Spilje can storage some of the water for 3-5 months depending on the hydrology and demand for generation of electricity in the grid. Therefore, on the bottom side of the table the operating parameters show the probability of turbine discharge and generated energy corresponding to turbine discharge. On Fig.2 and Fig.3 are presented the graphical overview of the numerical results from Tab.1.

The results of probability show that the expected inflow water over the turbine installed discharge is near 6% (Fig.2a), which correspond to spilling energy of the HPP Spilje.
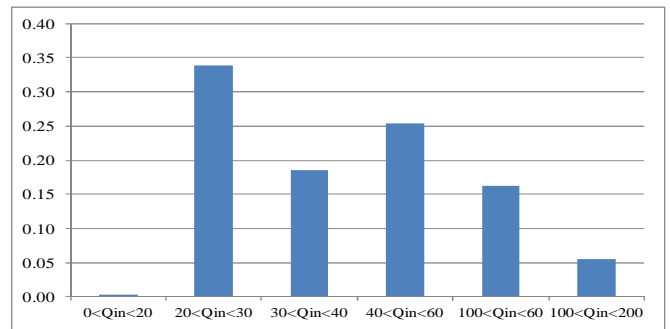
a) Water inflow probability
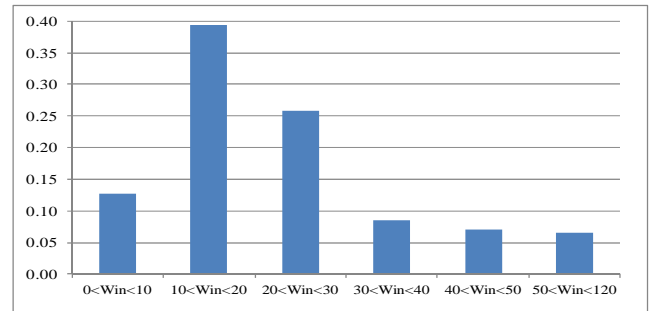
b) Energy inflow probability

Fig. 2. Probability of inflow parameters
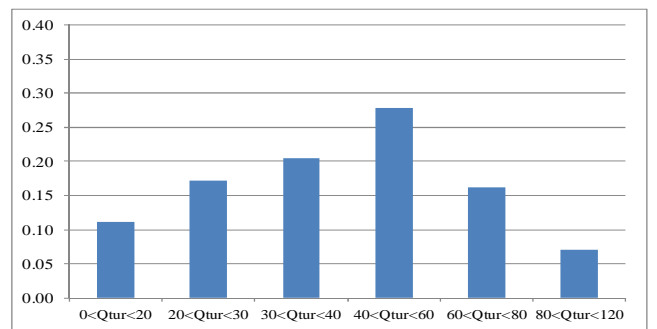
a) Turbine discharge probability
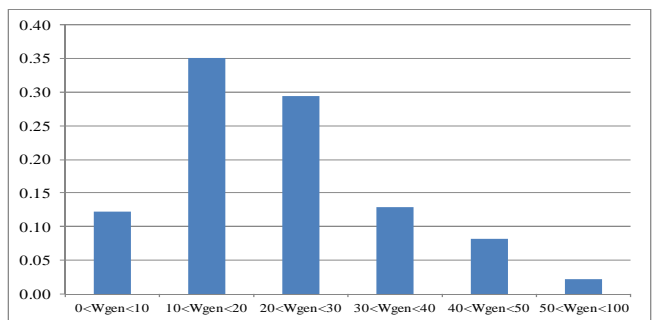
b) Generated energy probability

Fig. 3. Probability of operating regimes of the HPP Spilje

The turbine discharge distribution (Fig.3a) shows relatively good matches with normal distribution around the average inflow. The highest probability of generation electricity is around 10-30 GWh in a month (Fig. 3b) which is nearly half of the maximum monthly production (W_{max} =56 GWh in a month) for the gross head of 90 m.

IV. MONTHLY DISTRIBUTION OF THE PROBABILITY FOR INFLOW PARAMETERS

The same analysis can be done for monthly probability distribution for the inflow parameters (water and energy inflow), as well as for the values of operating regimes for the 45 years of data. The values of probability show the expected energy inflow, and other relevant parameters needed for operating of the HPP. The energy which is generated from the HPP can be divided into three parts, as which are base energy, peak energy and spilling energy.

$$W(t) = W_{base}(t) + W_{peak}(t) + W_{spill}(t) \qquad (3)$$

The expected base load and peak load energy can be approached by the time interval of low tariff (for base load) and high tariff (for peak load) periods of the day. For the Power System from ELEM Macedonia, the low tariff is Tlow=7 hours in a day and the rest Thigh=17 hours is the peak load period. Therefore, the maximum generated monthly energy of HPP Spilje is around Wmax=56GWh.

If the Win<[Thigh/(Tlow+Thigh)]*Wmax, then the HPP can operate in peak regime. For HPP Spilje, if the Win>(17/24)x56 = 40 GWh, then the plant is operating in base load regime (low tariff). Therefore P(40<Win<50) is the probability to have base load energy. Over 50 GWh of inflow energy, or P(50<Win<120) is the probability to have spilling energy.

Tab.2 gives the monthly probability of inflow energy which gives the expected energy inflow for each month, which is very important for the planning point of view.

TABLE II
PROBABILITY OF INFLOW ENERGY Win CORRESPONDING TO WATER INFLOW Qin

Probability Win (GWh)	Jan	Feb	Mar	Apr	May	Jun
0<Win<10	0.02	0.13	0.02	0.00	0.00	0.00
10<Win<20	0.42	0.27	0.31	0.04	0.09	0.38
20<Win<30	0.27	0.47	0.36	0.33	0.22	0.33
30<Win<40	0.16	0.04	0.13	0.18	0.13	0.16
40<Win<50	0.09	0.07	0.11	0.22	0.18	0.11
50<Win<120	0.04	0.02	0.07	0.22	0.38	0.02

Probability Win (GWh)	Jul	Aug	Sep	Oct	Nov	Dec
0<Win<10	0.18	0.20	0.33	0.29	0.29	0.07
10<Win<20	0.51	0.71	0.62	0.60	0.40	0.38
20<Win<30	0.31	0.09	0.04	0.07	0.29	0.31
30<Win<40	0.00	0.00	0.00	0.04	0.02	0.16
40<Win<50	0.00	0.00	0.00	0.00	0.00	0.07
50<Win<120	0.00	0.00	0.00	0.00	0.00	0.02

The last two rows have the values of probability to operate in base load (40<Win<50) or to have spilling energy (50< Win<120). These probability values for base load and spilling energy in each month are presented graphically on Fig.4.
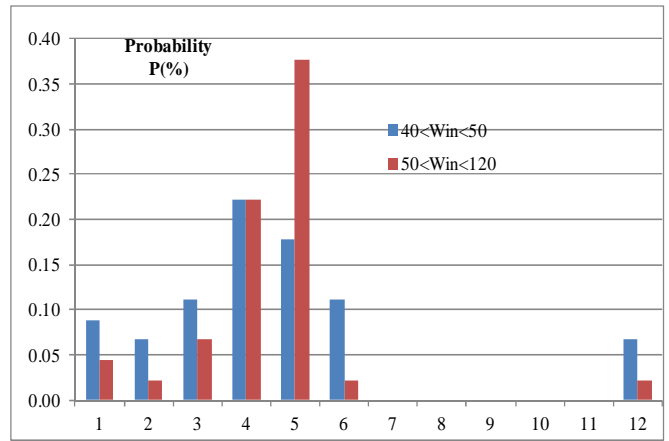
Fig. 4. Probability of operating in base load and spilling energy for each month in the year
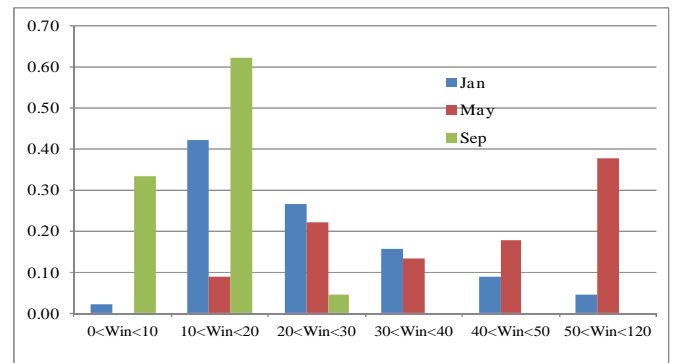
a) Expected inflow energy probability
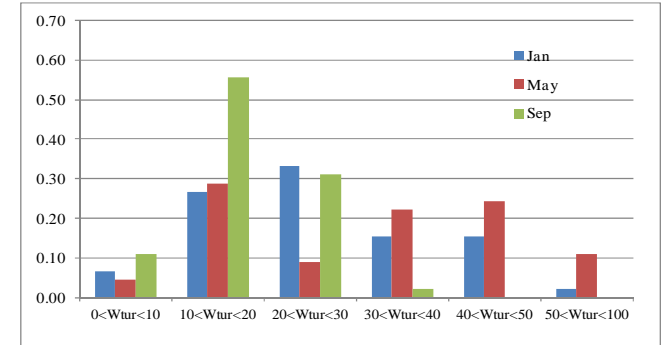
b) Generated energy probability

Fig. 5. Probability of inflow energy and generated energy for HPP Spilje for 3 typical months

The results show high probability to have spilling water and energy in April and May, and also high probability of operating in base load (from January until June).

Fig.5 presents the values of probability for inflow energy and generated energy for 3 selected months (January, May and September). The selected months are typically for winter period (January), spring period with highest inflow (May) and September with lowest inflow.

The results show that the represented months have different probability of operating regimes comparing with the inflow energy. January is typically winter month with middle inflow, and high generated energy which is needs of the request on

electricity market. May is typically the period with the highest water inflow, but the requests of generated energy are smaller because of low demand. The high generated energy is mainly for the reason to avoid spilling water. This is the indicator for additional unit(s) in order to have more hours peak load in a day and to avoid spilling. September is typically month with low inflow water and low demand of energy.

V. INFLUENCES OF ENERGY PARAMETERS TO THE FINANCIAL ANALYSIS OF HYDRO POWER PROJECT

The financial analysis is based on the economic and technical parameters of the hydropower project, with taking into account indicators such as energy and benefit from the HPP operation. The parameters needed as input variables for the analysis can be divided into 3 groups:
• Technical parameters (installed flow, number of units, head losses, the reservoir volume, the reservoir elevations)
• Energy parameters (installed power, electricity generation, tariff policy, transmission system limitation etc.).
• Economic parameters (investment, repayment period of the loan, construction period, discount rate, operating and maintenance costs, the cost of produced electricity, etc.).
The model can be described as a function of all three sets of input parameters as:

$$FINanalyse\_HPP = f(TEHNp, ENERGp, ECONp) \quad (4)$$

The technical parameters reflect the design of the project with all associated facilities: construction, mechanical and electrical design solutions. Energy parameters are closely related to the technical solution and the inflow water and operation modes of the HPP. Economic parameters as investment costs depend on the technical solution, time of construction, the discount rate and the rate of loans, prices of generated energy and power engaged, and the operation costs. After appropriate processing of the necessary input parameters, follows the financial calculations of economic indicators such as: NPV, B/C, IRR and PBP, which give a clear image for project feasibility.

Required input parameters of the model (technical data, energetic parameters and economic parameters) are:
• Installed capacity of the hydro power plant.
• Investments in power plant unit (includes all investments as material needs, equipment, installation, construction, etc.).
• Expected electricity production or CF-Capacity Factor.
• Economic parameters: time of repayment of the loan in years, a grace period of years, the rate of return of the loan, inflation rate, discount rate for all cash flows etc,.
• Operating life of the plant, number of years of operation for economic profit.
• Price of electricity in high and low tariff
• O&M costs of the plant are divided into fixed costs which do not depend on generated electricity (salary costs of employees and utility costs) and variable costs which are mostly for equipment maintenance and repair and depend on electricity production.
• Additional benefits such as subsidies for environmentally friendly facility (CDM project).

The annual or monthly income depends on electricity production in high and low tariff:

$$Income = \sum_{t=1}^{T} \left(Chigh(t) \cdot W_{HIGH}(t) + Clow(t) \cdot W_{LOW}(t) \right) \quad (5)$$

The data for water inflow and the operating data of energy generation, can give expected probability for certain value of generation electricity corresponding of technical and operating condition of the HPP.

VI. CONCLUSION

The hydro power plants are complex facilities which require large investment activities, which should certainly make a precise analysis of economic valuation. Taking into account the future uncertainties, the investors face risks of different nature, such as: financial risks, stock risks, political risks, technological risks and others.

Each project individually may have satisfactory financial performance and be appropriate positive and attractive for investment, or on the other side to have unacceptable economic parameters for investment.

The data of inflow parameters can give energy expectations of the HPP, which is necessary for operational modes of the hydro power plant as the followings:
• HPP operating for 24-hour demand in high and low tariff.
• Impact of water inflow or the distribution of electricity production in the long-term period of 30-40 years (various hydrological conditions of extremely dry, average and wet monthly inflows).
• Energy losses because of spilling energy in extremely wet periods.
• Expected lower revenue in extremely dry years.

Therefore, the contribution of the input data for water inflow, as well historical operation data of the hydro power plant is very important in order to calculated expected income from electricity generation. On the other hand, these analysis are very important for improve the technical performances of the HPP in increasing the installed turbine flow in existing units, or to install additional unit(s) in the power plant.

REFERENCES

[1] Olsson, M., Söder. L.: "Hydropower Planning including Trade-off between Energy and Reserve Markets", Published in Proceedings of IEEE Bologna PowerTech 2003 June 23-26th Bologna Italy.
[2] Amir Pasha Zanjani Nasab; Financial Analysis of Small-Hydro Power Project in Malaysia from the Investor Perspective, 2012 IPCBEE Vol.33 (2012) Singapore.
[3] A.Iliev, V. Fustik: Hydro Project Benefit Calculation in a Fuzzy Sense, Proccedings of Second Balkan Power Conference, Beograd, 19-22 June 2002, pp.247-252.
[4] A. Iliev, V. Fustik: A new Fuzzy Based Method for Calculation of Electricity Production from Hydroelectric Projects, 4-th Balkan Power Conference, 25-28 May 2004, Sarajevo, Bosna i Herzegovina, p.p. 137-143.
[5] Technical documentation of hydro energy system of Crn Drim.

ORAL SESSION
ENGINEERING EDUCATION

Engaging Students in Learning through Collaborative Cloud Technologies

Tania Vasileva [1], Vassiliy Tchoumatchenko [2], Mariana Manoeva [3]

Abstract – **Formal education is challenged by emergent trends highlighting students' needs to develop competencies and abilities to use technologies for collaborative knowledge creation and innovation. The paper considers a case of transforming bachelor degree course towards promoting students' knowledge work competencies by using the trialogical learning approach and computer cloud technologies to increase student motivation.**

Keywords – **Knowledge work practices, Collaborative learning, Cloud technologies.**

I. INTRODUCTION

Today's students will have to tackle jobs that are profoundly different from existing ones. They will be employed in positions representing modern knowledge work in technology-rich environment. Corporations are using Web 2.0 technologies and cloud tools to allow geographically dispersed large groups of workers to collaborate. In order to manage changes in the society and in the work life, new types of competencies are needed, such as collaborative learning, networking, working in multidisciplinary teams, cultural awareness, self-leadership and flexibility.

Formal education is expected to support students in acquiring necessary competencies and in developing abilities to use technologies for collaborative knowledge creation and innovation. Prevalent pedagogical methods and practices do not usually support these new challenges because the focus is on content learning rather than on fostering higher-order knowledge work competencies. [1]. Pedagogical methods are still largely based on well-defined problems with known solutions. In results students are reported to leave higher education with underdeveloped abilities to collaborate, manage their work processes, use computers, or solve open-ended problems [2].

New strategies are needed in educational systems for introducing pedagogical models that address learning of necessary competences and abilities students to use modern computer cloud technologies for collaborative knowledge creation. Web 2.0 technologies and sites are becoming an integral part of youth culture: today's youth create, share content with other creators, and communicate via blogs, wikis and digital video. Educators can tap students' enthusiasm and creativity to shape and carry out their education agenda while preparing students for the workforce of the future. They need to find practical way to bridge conventional educational practice with new technologies in order to harness students' engagement as they address instructional objectives through creation of content.

To answer these challenges the KNORK (Promoting Knowledge Work Practices in Education) project [3] aims at developing pedagogical models and technology to support collaborative practices in technology-rich environment.

The paper considers efforts done in the Department of Electronics at the Technical University - Sofia to reconstruct pedagogical practices in compulsory course of Semiconductor devices to promote collaborative team work within computer cloud environment. The results from pilots conducted with several students group from two Faculties are also discussed.

II. COURSE REDESIGN

A. Educational Problem

The Semiconductor device course is a basic compulsory course delivered to the huge amount of students from several faculties in 3-th semester of bachelor degree study. Before restructuring the pedagogical practices used in our teaching, we have carefully reviewed our course, its positive outcomes and drawbacks. Currently, to the students in the laboratory are given many unrelated tasks they perform in groups of 3-4 persons. Each student should individually prepare a separate report on the outcome of the practical work. Teacher guides individual student when needed.

This way of conducting training allows some students just to attend in classes without being actively involved in the tasks during the semester. Teachers cannot assess the progress of students as they evaluate the final product of their work. Since the multiple tasks are the same for all students most of them just copy the reports from their colleagues without understanding the material. Because assessment is based on individual final product, the teacher has thoroughly to conduct face-to-face examination of each student in order to evaluate him correctly.

The problem was how to restructure the Semiconductor Devices course in order to: obtain better students' knowledge and competencies, encourage better systematic training during the semester, stimulate circuit design and simulation for project verification, and transfer the initiative towards student-teacher direction. The educational challenge was haw:

- To increase the commitment and motivation of students

[1]Tania Vasileva is with the Faculty of Electronic Engineering and Technologies at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: tkv@tu-sofia.bg.

[2]Vassiliy Tchoumatchenko is with the Faculty of Electronic Engineering and Technologies at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: vpt@tu-sofia.bg.

[3]Mariana Manoeva is with the Faculty of Electronic Engineering and Technologies at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: manoeva@abv.bg

- To increase students' practical training, and
- To meet the requirements of business for: team work on common task, shared responsibility for the quality of the overall product; distribution of tasks in line with the specified deadline.

B. Trialogical Learning Approach

In order to meet these challenges and to resolve mentioned educational problems a new trialogical educational approach [4] was introduced with using cloud computing technologies, up-to-date communication tools for student-teacher connection, continuous monitoring and assistance students' activities.

Thrialogical approach emphasizes the organization of learner activities around shared objects that are created for some meaningful purpose or reason. This approach builds on the assumption that learning is not just individual knowledge acquisition (*monological*) or social interaction (*dialogical*), but activity is organized around transforming, or creating shared knowledge objects. A set of trialogical design principles was developed to guide the implementation of trialogical approach into pedagogical practices [5].

We decided to reconstruct the whole course giving the students opportunity to work collaboratively in group with clear role of each participant in a common work. In course redesign we have used examples and previous experiences from courses based on trialogical approach on learning [6], but in this case the problem was that the students are too many and have not enough engineering background to develop collaboratively three months long project.

Our efforts were aimed at changing the practical training to obtain better skills and competencies. Trialogical design principle was used to address:

- Team work on shared object (report)
- Continuous and prolonged work (within 2 weeks) before the laboratory work
- Strengthening the tasks of circuit design using devices' data sheets and simulation of the circuits, calculations of circuit's currents and parameters
- Continuous monitoring and teacher assistance in this process, providing help on request
- Reporting on the individual contribution of each team member to the overall project
- Respect to meet the deadline (after the prescribed date the project is locked for editing)

III. PILOTS ORGANIZATION

Laboratory exercises cover part of preparation and design and practical work. Team work is encouraged in collaborative development of common shared reports by using cloud computer technologies. The aim is to achieve systematic training during the semester and to stimulate design and simulation for design verification.

Instead of giving students many separate or loosely connected tasks we provide them with a large task, continuous working process, shared research plan and final presentation in groups. All group activities are organized around shared objects – collaboratively development of common project, and preparation of shared report. Project development in such practice permits for self-selected time and place allocation of the participants and teachers. Guidance is provided through systematic instructions and group work rules.

This approach permits for educational methods of direct student-educator contact that are not face-to-face, but are mediated through new communications technologies. Online communication allows students and academics to remain separated by space and time, but to sustain an ongoing dialogue.

Before the course starts are done: teams' formation; gmail accounts of all students; development of documentation templates with tasks to be done for all pre-lab project and final report; guidelines for students for practical sessions; LTspice tutorial. Students have access to learning materials in the Moodle LMS system and the Web site of the course.

Students have to prepare 8 Lab projects (4 two-week cycles with 2 lab project each) and practical measurement tasks in the semester, to pass two intermediate tests and the final exam test. The main phases in each cycle are:

- Pre-lab phase (Design & Analysis tasks – circuit design, parameter calculation, simulation)
- Face to face session (Discussion on common problems, faced by most students, answering difficult subject questions)
- Pre-lab tasks continue (Design & Analysis)
- Laboratory work (Practical measurement)
- Project finalization (final preparation of shared report including measured data)

The goal of pre-lab phase is students to be prepared in advance on devices, which they will be experimented with during the laboratory session. Teams have to develop shared reports, which include questions on device's mode of operation, basic volt-ampere characteristics, to design simple circuits with basic device applications and to calculate device parameters. A number of problems are assigned to be investigated using simulation tools. During this activities student use schematic capture & graphic drawing, simulation of characteristics to explore influence of different parameters (temperature, currents or voltages), calculation of device parameters, circuit design - currents, resistance, parameters, design verification by simulation.

Face-to-face session is intended for discussion of common problems, faced by many students.

During the lab measurements students perform practical work in the laboratory measuring devices in different mode of operation, different temperatures and signal frequencies. All measured data are filled in shared document by using students' smartphone or computers.

In project finalization phase students have to prepare final shared report including measured data, drawing graph on these data and parameters calculation. Reports also include conclusions, answering problem questions, concerning device application, which are not covered in the practical sessions.

Students' knowledge is evaluated continuously during the semester and by final exam test. The shared report grade is based on next criteria: material/organization, presentation,

depth of material, handling of questions, solving problems and conclusion on simulation and measurement results. Commenting activities and communications between students - teachers (emails, participation in discussions, pear comments) are also appreciated.

IV. COLLABORATIVE WORKSPACE ENVIRONMENT

The used environment consists of public cloud based services, combined in a way that supports collaborative electronic design reports development (see Figure 1). All participants had to register individual Google accounts. The teacher was responsible for creating a Google Docs document for each project report and sharing it with the team.

During the course we use cloud & communication tools and specific tool for simulation. Google Tools are used for collaborative development of a common shared object in the cloud – Google Drive, Google Docs, Sheets; Google calendar – to set deadlines and to monitor progress – assignments, intermediate stages reporting, deadline for submission of project.

Students work in teams of 2-3 persons and are required collaboratively to design, analyze and measure circuits with basic semiconductor devices. During the semester every team needs to prepare within fixed deadline 8 (two week long) reports, concerning features and characteristics of different semiconductor devices.

Completed document on the long-term group work is created in Google Drive as a shared document with the possibility of collaboration between the team members and comments from the teacher. In the shared space it is possible to upload files Word documents, graphics, pictures and other materials as well as measured data during practical exercised in the labs (there is Wi-Fi in the classroom and students are allowed to use their laptops or smartphones to access shared report).

As a Specific tools for analysis phase is used LTspice® – Free Circuit Simulation, Schematic Capture and Waveform Viewer Tool http://www.linear.com/designtools/software/.

Students were encouraged to ask for help or advice, via email, at any time and not to wait for the scheduled classes. For student–teacher communications are used Google applications: Gmail, Calendar, Drive и Google+. Usually they were getting a response during the same day. Announcements were made on a Google+ hangout and via email. Each class had a Google calendar with all relevant milestones and class schedules. For inter team communications students can choose their preferred tools (chat, conferences, e-mail, forum).

V. RESULTS

The pilots were conducted with 6 students group during autumn semester of bachelor degree courses in Electronics and Computer Systems and Technologies.

In order to describe their opinion and experience of the course students are asked to answer through SurveyMonkey https://www.surveymonkey.com/ to the following open questions: How would you characterize your overall experience in the course? What has been positive or impressive in the course? What has been challenging or disturbing in the course?

For the students this was their first course in which they work in team and they consider this very positive, challenging and useful to understand the benefits of working in collaboration. Most students appreciated the fact that they had to work in a group and share the work between the members by using technology.

The students appreciated the visibility of their contributions to the common work. They said that they have understood how important is the expertise and commitment of others when developing common products. They also noticed that during team work they started knowing their colleagues better contacts and even make new friends.
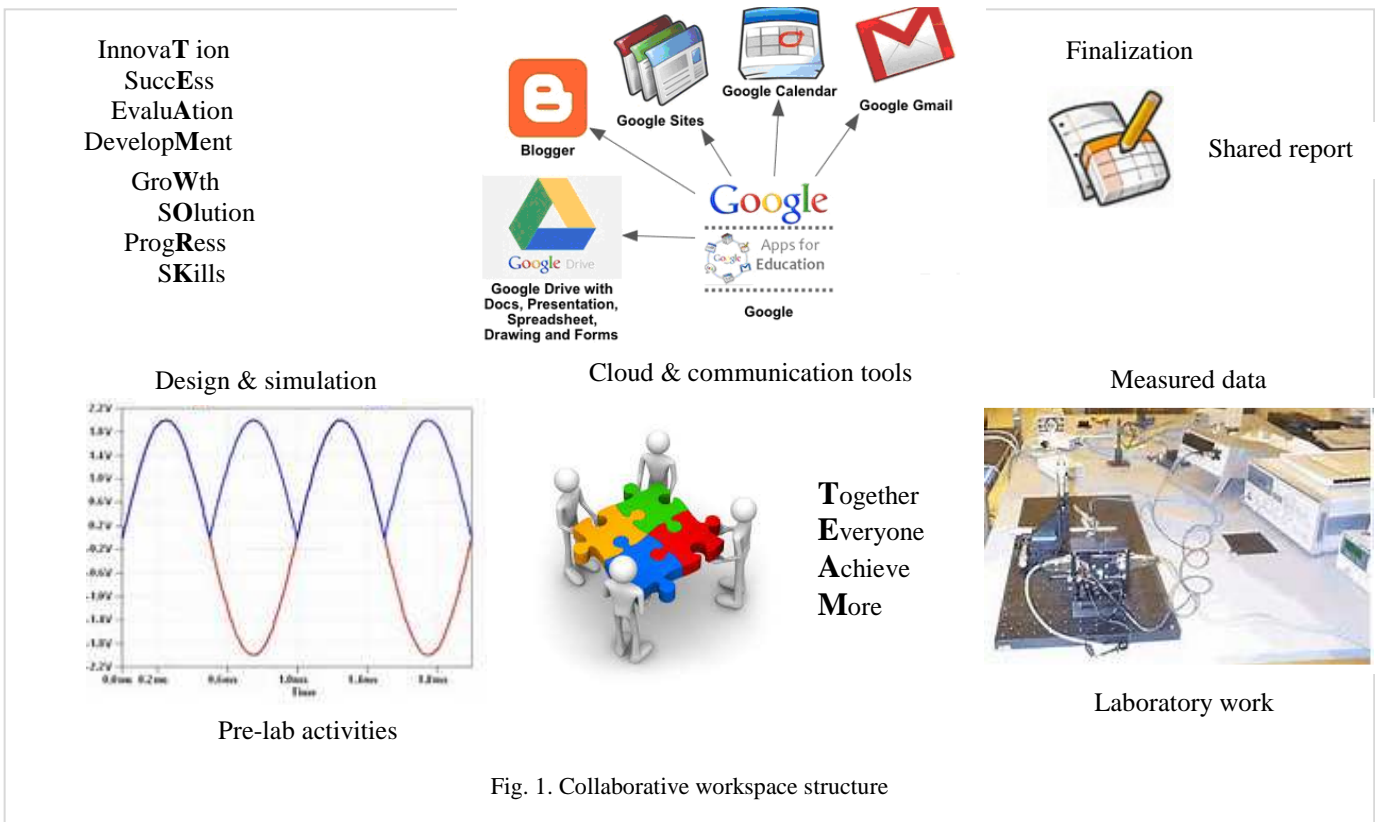

Fig. 1. Collaborative workspace structure

Innovative way of working in teams using up-to-date digital technologies was appreciated. The positive aspects identified from students are mainly related to the possibility to know and learn new tools, to study in an innovative and engaging way, to have immediacy support from teachers by receiving timely feedback and help. The immediacy of the help provided via email, compared to the scheduled face to face meeting, was cited as a major plus in the post-course surveys. Being able to receive a timely advice on their design problems was highly regarded. They consider positively the opportunity to work at any time at any place, which helps them to manage their free time in more effective way. Some students complain that part of the team does not work well and do not contribute to the quality of common work. Most of the students are satisfied with the new way of course delivering and declares that their expectations were exceeded.

Teachers adopted new pedagogical practices compared to previous courses: longitudinal work which also supported students' more in-depth focusing, students' collaboration for a shared outcome. According to teachers, the students learned knowledge work practices, such as information processing, analysis, presentation and sharing, longitudinal work, using digital tools and group work in general. The teachers felt that it is important that the new practices were successfully used for improving obligatory courses and for a large group of students. All teachers will continue to apply the practices and this course and into their other courses.

VI. CONCLUSION

The paper considers transforming bachelor degree course towards promoting students' knowledge work competencies by using the trialogical learning approach and computer cloud technologies to increase student motivation.

Introducing new technologies and paradigms in established engineering courses is always challenging. In addition to the core subject matter, students had to learn new tools and development workflows. In a whole, it has been a rewarding experience for both students and teachers. As positive results could be mentioned:

- Greater interest and involvement of the majority of students.
- Besides their knowledge on the subject students acquire skills to work in a team and to use advanced tools for collaboration and communication in the network.
- Better preparation for laboratory work – students are acquainted with the problems, with devices' mode of operation and the characteristics, which they will explore
- Avoiding the problem of copying reports from one to other and their delivery to the end of the semester (the project is locked after the deadline).
- Control of the process and the contribution of each participant – notes and comments of the teacher in total shared reports during its development in the Google docs' document.

As problems students mention difficulties in distribution work between team members, insufficient opportunity to learn from their mistakes and those of their colleagues, lack of habit

to comply with fixed deadlines for projects' submission (after the final date the project is locked for editing).

As a problems teachers reported:

- Difficulties in precise evaluation of personal contribution of each team member to common work
- Problems how to force lazy students to work well
- Extremely heavy-duty of assistants not only in classes but also in the preparation of assignments for individual & shared work and continuous consultations, monitoring and evaluation of many students' reports.

Based on upper mentioned outcomes we will try to improve the solution in next course release by dividing role between students in the team and to rotate these roles during the semester, by rewarding and punishing student by bonus points, contributing to their final score, by forcing the students to comment on each other's work throughout the course. We will try to reduce teaching load by giving students less number but bigger reports (which will be more easy to be monitored) and by minimizing face to face seminars through development of guides how students to use cloud tools.

ACKNOWLEDGEMENT

REFERENCES

[1] H. Muukkonen, M. Lakkala, J. Kaistinen, G. Nyman, „Knowledge Creating Inquiry in a Distributed Project management course", Research and Practice in Technology-Enhanced Learning, no. 5, pp. 73–96, 2010

[2] J. Duderstadt, "Engineering for a Changing World. A Roadmap to the Future of Engineering Practice, Research, and Education", The Millennium Project, 2008, Retrieved from http://milproj.dc.umich.edu/publications/EngFlex_report/download/EngFlex%20Report.pdf

[3] "Promoting knowledge practices in education" - KNORK Project, http://knork.metropolia.fi/website/, 2015

[4] S. Paavola, L. Lipponen, K. Hakkarainen, "Models of Innovative Knowledge Communities and Three Metaphors of Learning", Review of Educational Research, vol. 74 no. 4, pp. 557-576, 2004

[5] S. Paavola, M. Lakkala, H. Muukkonen, K. Kosonen, K Karlgren," The Roles and Uses of Design Principles for Developing the Trialogical Approach on Learning", Research in Learning Technology, vol 19, no. 3, pp. 233-246, 2011

[6] T. Vasileva, V. Tchoumatchenko, "Promoting Knowledge Practices in Electronic Education",Annual Journal of Electronics, vol. 8, pp. 158-161, 2014

Infrared Investigation on the Thermal Field in the Case of Influence of Low Frequency Magnetic Signals on the Human Body

Desislava Chocheva [1] and Kalin Dimitrov [2]

Abstract – **In this paper the subject of infrared thermal imaging cameras and their potential for medical applications are investigated. Two different medical physiotherapeutic appliances are reviewed in terms of heat emittance during a standard therapy session.**

Keywords – **infrared thermography, medical applications, therapeutic ultrasound, low frequency magnetic signals**

I. INTRODUCTION

Through the development of technics and technology, every day we are facing new perspectives in health care – both diagnostically and therapeutically. What is observed in the last few years is an increasingly more significant implementation of infrared thermal-imaging-based medical apparatus, generally in the diagnostic sphere. However the potential of this technology in contemporary medicine is still to be explored in further detail. The current paper is an attempt to broaden the view and deepen the understanding of some of the existing therapies through the analysis of data provided by a thermal camera.

II. CURRENT STATE OF THE PROBLEM

Currently, therapeutic ultrasound is widely used in physiotherapy, especially to facilitate the healing process in cases of ligament sprains, muscle strains, tendonitis, joint inflammation, plantar fasciitis, metatarsalgia, facet irritation, impingement syndrome, bursitis, rheumatoid arthritis, osteoarthritis, and scar tissue adhesion [1-6]. One of the primary reasons for its therapeutic abilities is the fact that stimulating the tissue beneath the skin's surface increases the temperature of the tissue, thus providing better blood circulation in the affected area.

The mechanism of low frequency magnetic signals therapy is based on a flow of electrical charges causing a net flow of ionic current for basic cellular restoration activities. Many studies show that healing effects are observed: cytoprotection of cells and the stimulation of growth factor synthesis [7]. Contemporary devices can generate signals with a different shape, frequency and length. The frequencies used are

[1]Desislava Chocheva is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail:des@motorsport.bg.

[2]Kalin Dimitrov is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: kld@tu-sofia.bg.

between 1 and 100 Hz, magnetic flux density being up to 100 mT.

III. EXPERIMENTS

The thermal camera used for data collection is FLIR E40, with thermal sensitivity of $< 0.07°C$, accuracy of $±2°C$ or $±2\%$ of reading and temperature range of $-4°F$ to $1,202°F$ ($-20°C$ to $650°C$) [1-4]. For maximum accuracy, the camera is fixed on a stand and movement of the object is avoided.

The general setting is shown in Fig. 1:

Fig. 1. Therapeutic procedure – thumb joints application

IV. RESULTS

To facilitate the data analysis, several measurement points and lines are used for the thermal images of both treatment methods. For the Ultrasound the points are SP1, SP2, SP3, SP4, SP5 and SP6, and the lines – L1, L2 and L3 (Fig.2):

Fig. 2. Thermal Image of therapy –
points and lines used for analysis

It is important to note that for each image the emissivity of human skin is considered through the coefficient $\varepsilon=0.98$ (Table 1):

TABLE 1. DATA COLLECTED
FROM ONE THERMAL IMAGE OF ULTRASOUND THERAPY

Measurements		°C
Sp1		30.9
Sp2		30.6
Sp3		31.9
Sp4		29.8
Sp5		31.3
Li1	Max	32.1
	Min	30.7
	Average	31.4
Li2	Max	31.8
	Min	30.3
	Average	31.0
Li3	Max	30.9
	Min	29.6
	Average	30.1
Parameters		
Emissivity		0.98
Refl. temp.		20 °C

During the Ultrasound therapy 34 thermal images were taken in even time-intervals. The data is collected by examining the thermal changes in each analyzed point/line in the sequences of thermal images. Figs. 3-7 show the detailed upward tendencies for all 5 analyzed points:
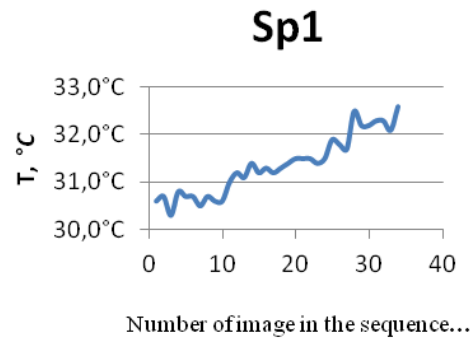
Fig. 3. Thermal changes in point Sp1
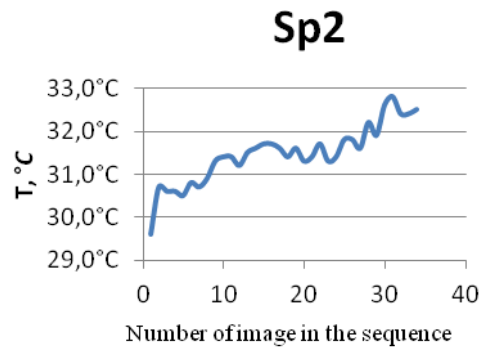
Fig. 4. Thermal changes in point Sp2
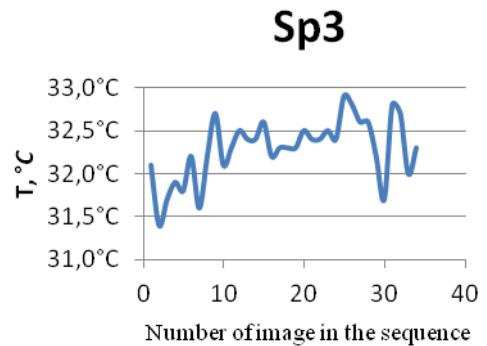
Fig. 5. Thermal changes in point Sp3
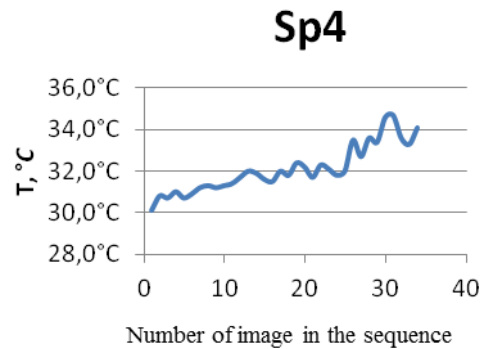
Fig. 6. Thermal changes in point Sp4
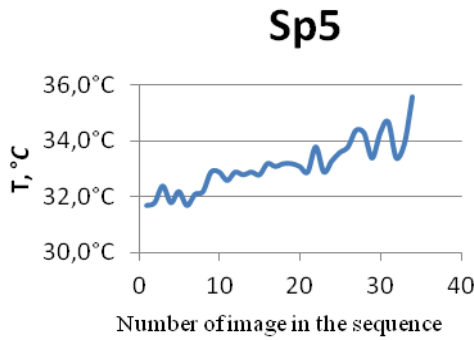
Fig. 7. Thermal changes in point Sp5

Fig. 8 shows the combined increase in temperature in all 5 examined points. What can be observed is the difference in the speed with which the temperature rises in the different points – for SP1 and SP2, which are chosen to be away from the direct field of application, the temperature increase is slowest, even though still significant. SP3 is on the head of the ultrasound appliance and the temperature there fluctuates a lot, with a total increase of below 1 degree. For SP4 and SP5, however, there is a more dramatic change – about 4 degrees rise throughout the ultra sound session.
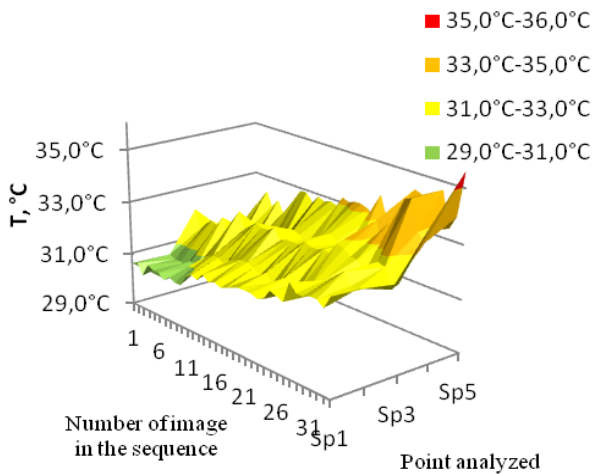
Fig. 8. Therapy – thermal changes in analyzed points (Sp1 – Sp5).

The next figures (Fig. 9, 10 and 11) show how the average temperature around the spot of application is changing through the therapy session. The upward tendencies are very obvious when the whole line is considered.

When all points in the lines are considered (finding the minimum, maximum and average), a more thorough image can be created explaining the entire thermal field in the close proximity of the spot of application and Fig. 12 shows the increase of temperature in further detail. This means that even though the ultrasound therapy is directly applied in a specific spot, it leads to a wider temperature rise, that goes as far as several centimeters from the point of application.
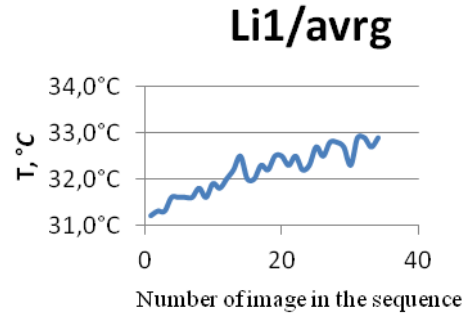
Fig. 9. Thermal changes in line Li1/avg
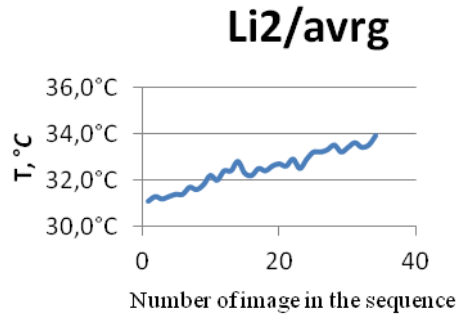
Fig. 10. Thermal changes in line Li2/avg
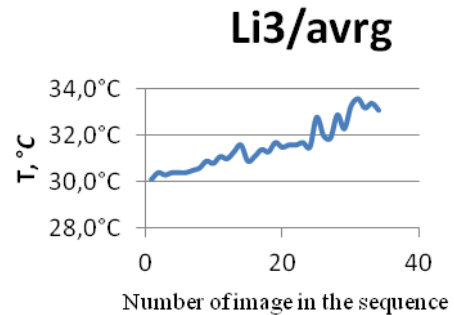
Fig. 11. Thermal changes in line Li3/avg
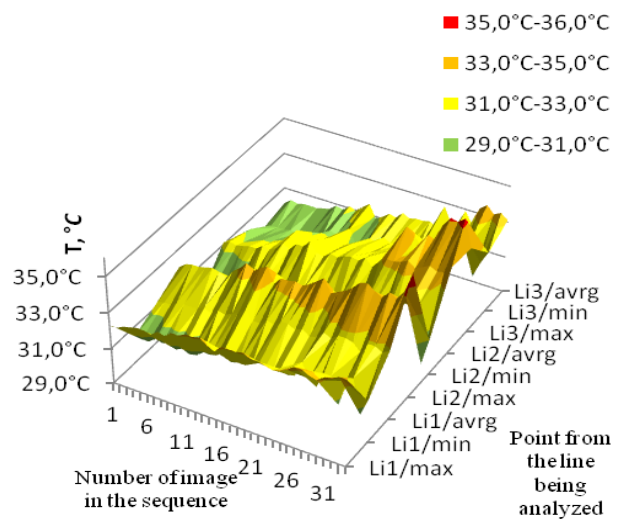
Fig. 12. Ultrasound Therapy – thermal changes in analyzed lines (the points in the lines that show min/max/avrg values of temperature)

In the case of magneto therapy, the infrared pictures were taken at intervals of 30 sec. for 15 minutes. The therapy duration was determined by the doctor in charge.
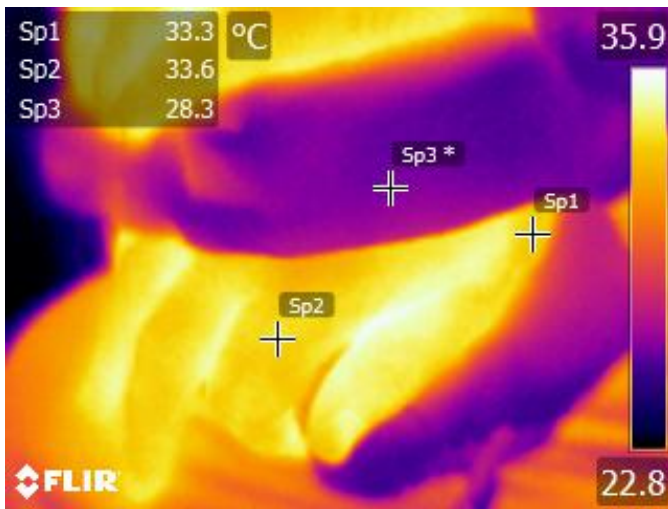
Fig. 13. An image from the magneto therapy series

Three points of interest were selected: Sp1 – in the area closest to the wrist; Sp2 – in the area of the index finger. Sp3 – on the textile wrapping of the magnet. We selected ellipses, not points, to eliminate probable human movements and noises. For the areas where the human skin is observed, a radiation coefficient 0.98 is selected and for the fabric-wrapped magnet a radiation coefficient 0.77 is selected.
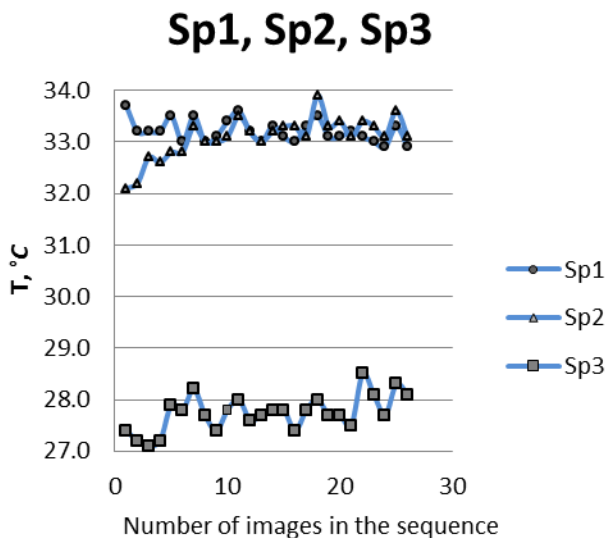
Fig. 14. Thermal changes in points Sp1, Sp2, Sp3

V. CONCLUSION

The analyzed data gives one more perspective of the countless applications of thermal imaging cameras for medical purposes. Such analysis can be very useful in providing a new understanding of widely-used technology in terms of safety and effectiveness. Since IR imaging is non-invasive and does not bear any risks, it can be implemented wherever it is useful

and possible. It carries a great potential also for diagnostic medicine – early tumor/cancer detection, diagnosing neuropathies, all kinds of inflammation, scanning large groups of people for high body temperature, veterinary purposes and many others. In the near future IR thermography might become a valid single diagnostic method for many diseases that involve intricate changes in the temperature field.

REFERENCES

[1] Childs P., (et. all),   Practical Temperature Measurement. Butterworth Heinemann, Reed Elsevier plc group, 2001.
[2] Darling, Charles R.; "Pyrometry. A Practical Treatise on the Measurement of High Temperatures." Published by E.&F.N. Spon Ltd. London. 1991.
[3] http://www.scigiene.com/productimages/429_INFRAREDTHE RMOMETERExplanationdoc.pdf - Scigiene Corporation, 1295 Morningside Avenueue, Unit 16, Scarborough, ON MIB 4Z4, Canada
[4] http://www.flir.com/cs/emea/en/view/?id=41372   -   FLIR Systems, Inc.
[5] Giridhar D., Robinson R., Liu Y, Sliwa J, Zderic V, Myers M. Quantitative estimation of ultrasound beam intensities using infrared thermography-Experimental validation, J Acoust Soc Am.,131(6):4283-91, 2012
[6] Van Haren F., Kadic L., Driessen J., Skin temperature measured by infrared thermography after ultrasound-guided blockade of the sciatic nerve, Acta Anaesthesiologica Scandinavica, vol. 57, Issue 9, pp. 1111–1117, 2013.
[7] D. Dimitrov, "Visualization of a Low Frequency Magnetic Field, generated by Girdle Coil in Magnetotherapy", Electronics and Electrical Engineering, No.6(78), 2007, pp.57-60

Computer visualization of space configuration of a low-frequency magnetic field using experimental data in physiotherapy and engineering education

Vladimir Nikolov [1], Kalin Dimitrov [2]

Abstract – **In the paper we demonstrate the step-by-step implementation of the application that allows us to calculate and visualize the magnetic field generated by out-of-line coils in arbitrary disposition. The application performs calculations by model equations in a given area. We have used various types of interpolations to reduce calculations. The program system is a useful tool for both the engineers and physicians who can choose an appropriate regime to increase the effectiveness of magnetotherapy treatment.**

Keywords – **low-frequency, medical applications, magnetic field, physiotherapy.**

I. INTRODUCTION

Physiotherapy, and particularly magnetotherapy, is one of widely used treatment methods. Therapeutic action of magnetic field is used when treating a broad spectrum of pathologies. The use of magnetotherapy improves the dynamics of the blood-vascular system and immunity. It should be noted that in numerous researches, no negative side-effects when using magnetotherapy were detected. The magnetotherapy method using low frequency field is in a good accordance with traditional medicine means.

The impact of the low-frequency magnetic field with low intensity may be considered as low and ultralow doses of radiation. The mechanism of ultralow doses effect is rather complex and at the moment there are no complete theoretical results. One of the known models explaining this phenomenon was proposed by L.N. Gall [4].

We have to solve the so-called inverse problem: given results of system functioning (or measurements) to analyze the system behavior, where "to analyze" means to obtain some characteristics and interpret them. To analyze the results of ultralow effects one can use bio tests (for example blood analysis, thermal imager data). It is the analysis of digital images that opens up possibilities to apply mathematical and computer methods to obtain comparative characteristics and classification signs of the images [2].

The axis of these coils can be parallel or not parallel. Usually a system for magneto-therapy contains one microprocessor unit which, according to the special software,

[1]Vladimir Nikolov is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: v_nikolov@tu-sofia.bg

[2]Kalin Dimitrov is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: kld@tu-sofia.bg

can switch different coils: define the number of "active" coils and their space dispositions so that they would be different in every moment during the process of therapy.

II. THEORY

1. Presentation for courses of bioengineering on the influence of a low-frequency magnetic field

In most cases, in practice the required space configuration of the magnetic field is created by means of one or more air coils, which are appropriately arranged in the space. For a coil there is a linear relationship between the magnetic flux density of the excited magnetic field and current in the coils. It is assumed that the environment, in which the space-time configuration of the magnetic field is considered, is linear. The resulting field is the superposition of the fields of the individual coils, which forms a more complex time-spaced magnetic field. We suppose that the environment around the coils is homogeneous and the relative magnetic permeability is constant, i.e. $\mu_0(x, y, z, t) = const$.

The results of mathematical, computer and experimental studies of the spatial configuration of the magnetic field excited by a cylindrical coil, one of the most common structures of the low-frequency excitation magnetic field, are described in [3]. Visualization was performed in the plane orthogonal to the base plane of the coil. Numerical experiments were also performed for two coaxial coils. In what follows we use Fig.1 [11] illustrating a base model in the
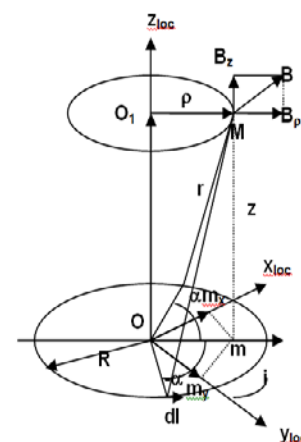
Fig.1 Local cylindrical and Cartesian coordinate systems of a coil

study of the spatial configuration of the magnetic field excited by a cylindrical coil.

2. Mathematical model for the base configuration

Every coil has its own coordinate system — a local cylindrical one. We denote this system by LC (with coordinates $\rho, z = z_{LC}$) and introduce the following notations:

i — current value in a single current loop;

R— radius of the current loop;

O — the origin of the cylindrical coordinate system of the coil;

M — an arbitrary point in which the magnetic induction is calculated;

m — the projection of the point M in the plane of current loop;

O_1 — the center of the circle lying in a plane parallel to the plane of the current loop and passing through the point M;

ρ — radius of the circle lying in the plane parallel to the plane of the current loop and passing through M;

dl — an elemental segment of the current loop;

2α — the central angle corresponding two symmetrical elemental segments;

r — the distance from any of endpoints of $2 dl$ segment to the point M.

It should be noted that on Fig.1 a current loop coincides with the base plane of the coil.

The vector of magnetic induction B in the point M has two components which can be calculated as the following [11]:

$$B_\rho = \frac{\mu_0 i}{2\pi} \frac{z}{\rho\sqrt{(R+\rho)^2+z^2}}\left(\frac{R^2+\rho^2+z^2}{(R-\rho)^2+z^2}L-K\right),$$

$$B_{zLC} = \frac{\mu_0 i}{2\pi} \frac{1}{\sqrt{(R+\rho)^2+z^2}}\left(\frac{R^2-\rho^2-z^2}{(R-\rho)^2+z^2}L+K\right),$$

$$(1)$$

where K and L are complete elliptic integrals of 1 and 2 sort as functions of k and

$$k^2 = \frac{4\rho R}{(R+\rho)^2 + z^2}.$$

An example of visualization of magnetic induction for one coil is given below.

3 Model for out-of-line coils

The real problem is to calculate and visualize in a 3D magnetic field (in a given space area) generated by several coils, which are in an arbitrary disposition [1, 5, 6].
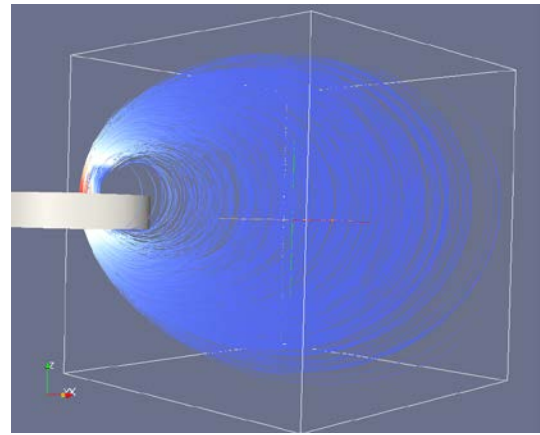
Fig.2. 3D visualization of the magnetic field for one coil. The given space area is bounded by parallelepiped.
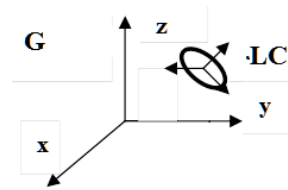
Fig.3 Global coordinate system (G) and local cylindrical system (LC) of a coil.

To do it, for a coil we consider a global coordinate system (G) with the origin O_G, a local rectangular coordinate system (LR) with the origin O_L and local cylindrical coordinate system of the coil. The calculation of the total magnetic field for several coils is performed in the area bounded by a rectangular parallelepiped. The area is divided into cells with the same size, which are also rectangular parallelepipeds. The calculation is performed in the nodes of the constructed mesh. The origin of G is supposed to be in the left bottom angle of the area.

Coils may be located both inside the area and outside it. Magnetic field values are calculated only inside the area. The direction of the coil axis is defined in accordance with the direction of current.

For every coil we use an auxiliary local rectangular coordinate system (LR) to reduce calculations, because the transition from one Cartesian coordinate system to another is coordinate shift and multiplication on a transition matrix. In this way for every coil the transition and inverse transition matrix can be calculated only once.

So, for given coordinates of a coil in G the sequence of calculation is the following:

1. $G \to LR$
2. $LR \to LC$, calculations by (1)
3. $LC \to LR$
4. $LR \to G$.

The magnetic induction should be calculated in many points around the coils. The number of these points depends on the resolution of computer images and the space configuration of the field of magnetic induction. It is clear that the increasing of resolution of computer images would need increasing of number of points. To optimize the running time, a variant of the algorithm using parallel calculations has been implemented: for every coil, the calculation of magnetic induction has been performed in a separate flow. The results of experiments are described in [1] and show that such an optimization leads to significant decreasing of the running time.

An example of visualization of results of calculations for 3 given out-of-line coils is shown below (fig.4).
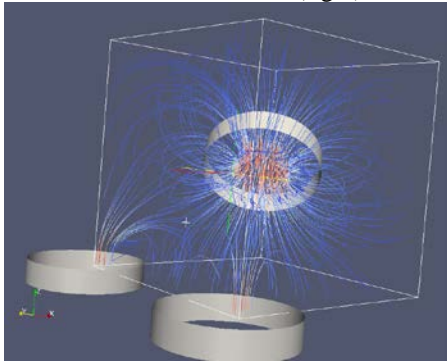
Fig.4. Magnetic field generated by 3 out-of-line coils. Parallelepiped shows the area where magnetic induction is calculated.

4 Interpolation as a method for reducing running time

The described algorithm of calculation of magnetic induction may be optimized by combining calculation in some base points of the area with interpolation in the rest of the points.

In [7, 8] we used values obtained by calculations as the base ones. We consider the following methods: 1) 3-dimensional linear interpolation, 2) one-dimensional quadratic Lagrange interpolation and 3) coordinate-wise cubic spline. Numerical experiments show that any of these methods reduces the running time considerably and does not lead to a loss of accuracy. For example, the results of experiments given in [7] show that the 3D linear interpolation is more effective than the 1D Lagrange interpolation and it decreases the running time nearly 9 times for the number of nodes 2.5×10^6.

5 Interpolation by using experimental data

For a special magnetotherapy device (magnet bed, fig.5) the solution the problem of magnetic field interpolation by using experimental data was presented in [9]. In such a configuration at any moment only one pair of coils is active. In the process of the system functioning every pair of coils is active by turns, and a movement of magnetic field occurs.
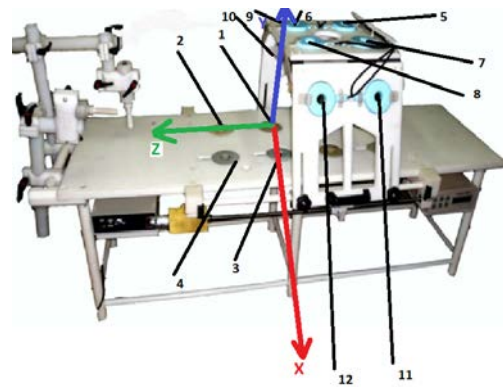
Fig. 5. Magnet bed, the coils and the coordinate system XYZ

For every given pair of coils the magnetic field was measured in the centers of coils and the middle of the interval between the centers. The measurement of magnetic induction is on the basis of a special algorithm implemented in a microprocessor. The values of magnetic induction are scalars (i.e. the module of the induction was measured). The interpolation of magnetic induction value in a given segment connecting centers of the coils (pair) and visualization of results have been performed. The order of the choice of pairs of coils is defined by a data table which also contains the coordinates of every measurement point.

We use one-dimensional quadratic Lagrange interpolation. As the electromagnetic induction value continuously depends on the point position, for any given pair of coils we at first calculate the induction value in all the points of the segment by a step of 1 mm (base points). Then we show the obtained distribution of the induction values in a special area in the interface in accordance with the following scheme (B denotes the induction value): $B \in [0, a]$ –blue, $B \in [a, b]$ –lightblue, $B \in [b, c]$ –yellow , $B \in [c, d]$ –orange, $B > d$ – red.

The user can choose the values of parameters a, b, c, d defining the color distribution. In this implementation the segment [0, d] is partitioned into 5 equal parts, where d is the maximal value of induction in the base points. Thus the diagram of color distribution is in agreement with the position of the point the user chooses as the interpolation one (fig.6).
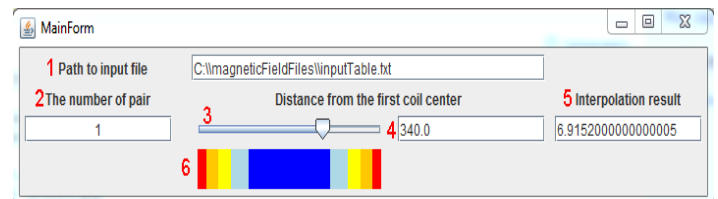
Fig.6. User interface for interpolation by experimental data.

The user gives the path to the input file (in .txt format) that contains experimental data. In field 2 one can choose a pair of coils (row number in the data table). Elements 3 and 4 are for input of the interpolation point. In 3 we select a point by using a slider, and when using 4 we should point a distance from the left end of the segment to the interpolation point. When changing data in 3 or 4 the result of interpolation is shown in 5 in real time. Note that the color distribution diagram is

drawn in element 6 immediately after the choice of coils; in doing so, the position of the slider corresponds the diagram.

III. RESULTS

For the visualization of calculation results, the Para view [10] package was used. One can compare the magnetic field intensity for various values of the current. As the magnetic induction linearly depends on the current (see (1)), we see that the intensity of the magnetic field in an area of interest increases gradually. The following pictures show the visualization of the field when we change the current from 0.5A to 2.5A (figs.7,8,9). It should be noted that Para view has a wide spectrum for transforming of obtained images, in particular it allows showing coordinates of a given point and the induction value.

IV. CONCLUSION

The paper is devoted to the presentation of methods of investigation of the influence of a low-frequency magnetic field on the human body. Usually a system for magnetotherapy contains one microprocessor unit which can switch different coils according to the special software during the process of therapy.

The visualization of the obtained magnetic field is a base for the engineering explanation of the process of therapy and for engineers designing magneto-therapy systems.

ACKNOWLEDGEMENT

Fig.7   I=0.5A

Fig.8   I=1A

Fig.9   I=2A

REFERENCES

[1] N. Ampilova, D. Dimitrov, B.Kudrin. Mathematical modeling of low frequency magnetic field in systems for magnetotherapy .Proc. 8 Int. Conf. CEMA13, 17-19 Oct. 2013, Sofia, Bulgaria. p.48-51.

[2] N. Ampilova, I. Soloviev, E. Gurevich. Computer Methods of Investigation of Ultralow Doses Effects in Biomedical Applications. Proceedings of the International Conference on Biomedical Engineering and Systems, Prague, Czech Republic, August 14-15, 2014. p.67-1–67-7. Available at http://avestia.com/ICBES2014_Proceedings/papers/67.pdf. ISBN : 978-1-927877-08-1.

[3] D. Dimitrov. Medical Systems for Influence of Electromagnetic Field on the Human Body (in Bulgarian), Sophia, Technical University, 2008.

[4] L.N. Gall. Physical principles of functioning of a live organism matter. SPb, 2014 (in Russian)

[5] B.Kudrin, A.Dimitrov. An algorithm for visualization of low-frequency magnetic signals in systems for magnetotherapy. Proc. 8 Int. Conf. CEMA13, 17-19 Oct. 2013, Sofia, Bulgaria. p.31-35.

[6] B.Kudrin, A.Dimitrov. Computer visualization of low-frequency magnetic signals in systems for magnetotherapy with variable parameters. Proc. 8 Int. Conf. CEMA13, 17-19 Oct. 2013, Sofia, Bulgaria. p.36-39.

[7] B.Kudrin, I.Soloviev On interpolation methods of low frequency magnetic field in systems for magnetotherapy. Proc. 9 Int. Conf. CEMA14, 16-18 Oct. 2014, Sofia, Bulgaria. p.154-157.

[8] B.Kudrin,V.Nikolov. Algorithms of interpolation and visualization of low frequency magnetic field using experimental data. Proc. 9 Int. Conf. CEMA14, 16-18 Oct. 2014, Sofia, Bulgaria. p.43-45.

[9] B.Kudrin,V.Nikolov. On the mathematical model of interpolation of low frequency magnetic field using experimental data. Proc. 9 Int. Conf. CEMA14, 16-18 Oct. 2014, Sofia, Bulgaria. p.80-84.

[10] ParaView Tutorial - http://paraview.org/

[11] K.Brandiski, G,Georgiev, V.Mladenov, R.Stancheva Theoretical Electrical Sofia, Bulgaria

ORAL SESSION
INFORMATICS AND COMPUTER SCIENCE

Parallel Computation of Fast Spectral Transforms of Logic Functions using the MPI Framework

Miloš Radmanović[1], Radomir S. Stanković[2], Dušan B. Gajić[3]

Abstract – **In many practical applications in signal processing, digital system design, logic design, pattern recognition, and related areas, it is often essentially important to be able to efficiently compute spectral transforms for logic functions. The corresponding Fast Fourier transform (FFT)-like algorithms for computation of various spectral transforms can be efficiently adapted to parallel computation platforms. In this paper, we investigate parallel implementation of two classes of FFT-like algorithms, the 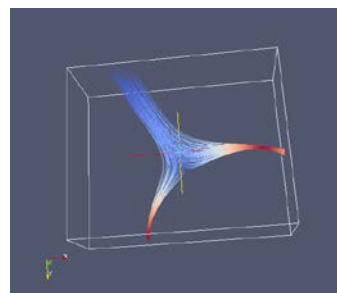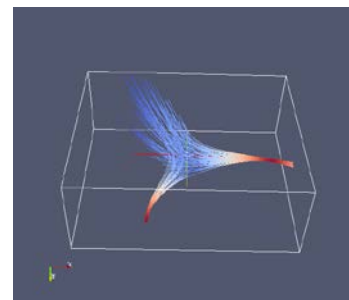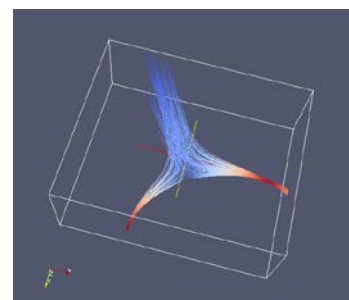Cooley-Tukey algorithms, and the constant geometry algorithms, for computing the Reed-Muller, the Walsh, and the arithmetic transforms on multicore Central Processing Units (CPUs) using the Message Passing Interface (MPI) framework. The paper also discusses certain specific parallel implementation styles of programming FFT-like algorithms on multi-core CPU platforms. Performance of the MPI implementations is compared with the classical C++ implementations for the single-core CPUs. It is shown that parallel MPI programming provides significant speedups for one of the considered implementation styles of the Cooley-Tukey algorithms. Other MPI implementations have a negative impact on the performance of both Cooley-Tukey and constant geometry algorithms.**

Keywords – **Logic functions, spectral transforms, FFT-like algorithms, multicore CPU, MPI.**

I. INTRODUCTION

Spectral transforms have many applications in signal encoding and processing techniques, synthesis, verification, and testing of logic circuits [1], [2], and many other areas [9], [10]. Due to rapidly increasing complexity of logic circuits and systems, in recent years, there has been a renewed study in spectral transforms for logic functions. For practical applications, it is often necessary to be able to efficiently compute these transforms. There is a variety of algorithms for efficient calculation of these transforms: FFT-like algorithms using truth vector representations of functions [3], fast tabular techniques, calculation algorithms through reduced representations of logic functions, and binary decision diagrams [2], [9]. The FFT-like computation has been one of

the most popular numerical methods applied in almost every field of science. The FFT-like computation can be executed much faster by using parallel processing [4], especially nowadays, as many supercomputing facilities are available to scientists and engineers across the world. Furthermore, the multicore desktop computers offer an inexpensive capability of parallel processing. Parallel computing on multicore CPUs enables parallel processing on commodity hardware. Only very recently the possibility of using multicore CPUs to solve complex problems in logic design has been explored by many researchers, for example in [5], [6].

Moreover, inspired by efficient execution of parallel problems in logic design and possibility of using multi-core CPUs platform, in this paper we investigate two classes of parallel FFT-like algorithms for computing spectral transforms of logic functions using the MPI framework. Particularly, in the case of the Cooley-Tukey algorithms, we investigate three implementation styles for the efficient parallel computation of the Reed-Muller, the Walsh, and the arithmetic transforms of logic functions using a multi-core CPU computation platform. Fast Reed-Muller, Walsh, and arithmetic transforms have the same time complexity of $O(N\log_2 N)$, where $N = 2^n$ is the size of the truth vector, and n is the number of variables in the function. These spectral transforms have different transform matrices that are Kronecker product representable.

The paper also investigates mappings of two distinct FFT-like algorithms, the Cooley-Tukey class and the constant geometry class algorithms, to the multi-core CPU computing model. There are many approaches for implementation of parallel Cooley-Tukey class FFT-like algorithms and they can be categorized into three styles for parallelizing butterflies of FFT: block data mapping, cycling data mapping, and all butterflies parallel mapping [7].

Performances of the MPI implementations of two classes and three programming styles of FFT-like algorithms for the efficient parallel computation of Reed-Muller, Walsh, and the arithmetic transform of logic functions are compared with the classical C++ implementations on the single-core CPU. The idea behind the selection of these transforms is to compare the performance of their implementations since they have FFT butterflies of different computational complexity.

The paper is organized as follows. Section 2 shortly introduces the fast spectral transforms for logic functions and illustrates by examples FFT-like algorithms for efficient parallel computation of the Reed-Muller, the Walsh, and the arithmetic transforms. In section 3, parallelization of FFT-like algorithms for these transforms is discussed. Section 4

[1]Miloš Radmanović is with the Faculty of Electronic Engineering, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mail: milos.radmanovic@elfak.ni.ac.rs

[2]Radomir Stanković is with the Faculty of Electronic Engineering, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mail: radomir.stankovic@gmail.com

[3]Dušan Gajić is with the Faculty of Electronic Engineering, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mail: dusan.gajic@ elfak.ni.ac.rs

presents experimental testing of mappings of these FFT-like algorithms with various data mapping styles to the multi-core CPU computing model using MPI framework. Section 5 offers some concluding remarks and directions for future work.

II. FAST SPECTRAL TRANSFORMS OF LOGIC FUNCTIONS

Spectral transforms of logic functions are an efficient tool in solving many tasks in logic design [2]. Spectral transforms defined by the Kronecker product representable transforms matrices have found many practical applications. The most common reason for this is existence of efficient calculation algorithms for these transforms.

In this paper, we discuss three different kinds of spectral transform of logic functions: the Reed-Muller, Walsh, and the arithmetic transforms. These transforms have different transform matrices that are Kronecker product representable.

The Reed-Muller transform [2] represents an important operator for obtaining AND-EXOR expressions of logic functions. The Reed-Muller transform matrix of order n, denoted by $R(n)$, is defined recursively as:

$$R(n) = \overset{n}{\underset{i=1}{\otimes}} R(1), \ R(1) = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}. \quad (1)$$

The arithmetic transform [2], which is also known as the integer Reed-Muller transform, was initially introduced to represent multiple-output functions by a single polynomial for the equivalent integer functions. The arithmetic transform matrix of order n, denoted by $A(n)$, has a recursive structure analogous to that of the Reed-Muller transform and is defined as:

$$A(n) = \overset{n}{\underset{i=1}{\otimes}} A(1), \ A(1) = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}. \quad (2)$$

The Walsh transform [6] is based on a set of orthogonal functions defined by J. L. Walsh which are an extension of a set of functions defined by H. Rademacher. Analogously to previous transforms, the Walsh transform matrix of order n in Hadamard ordering, denoted by $W(n)$, is defined as:

$$W(n) = \overset{n}{\underset{i=1}{\otimes}} W(1), \ W(1) = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \quad (3)$$

The spectrum of a logic function f given by truth vector $F = \left[f(0), f(1) \ldots f(2^n - 1) \right]^T$ is computed as:

$$S_f = T(n)F, \quad (4)$$

where $T(n)$ is any of the three matrices $R(n)$, $W(n)$, and $A(n)$, with computations performed in $GF(2)$ for the Reed-Muller transform, and in the set of rational numbers for the Walsh and the arithmetic transforms.

The FFT developed in signal processing for computing the Discrete Fourier transform (DFT) can be used to compute the coefficients in spectra of logic functions by varying just the basic kernels of the algorithms and ranges were the computations are performed [1], [2], [9]. The Reed-Muller, the arithmetic, and the Walsh transform matrices, expressed in (1), (2) and (3) respectively, can be factorized in different ways yielding different fast transform algorithms, the so-called FFT-like algorithms [8]. In this paper, we consider the Cooley-Tukey and the constant geometry algorithms for the Fast Reed-Muller transform (FRT), the Fast Arithmetic Transform (FAT) and the Fast Walsh Transform (FWT) algorithms. The selection was made in order to compare the performance of different parallel implementations on the multi-core CPU computation platform using MPI framework.

First, we consider the Cooley-Tukey (CT) algorithms, based on the Good-Thomas factorization which originates from the Kronecker product structure of the transform matrix [2]. Figure 1 shows the elementary butterflies operations (flow-graphs) for the Reed-Muller, the arithmetic and the Walsh basic transform matrices, respectively. In this figure, the solid and the dotted lines carry positive (+1) and negative (-1) weight, respectively. Note that, the operations for Walsh and arithmetic elementary butterfly are over integers, while the Reed-Muller butterfly uses $GF(2)$ operations.
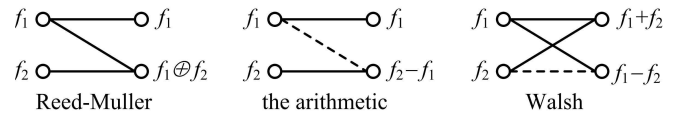
Fig. 1. The elementary butterfly operations for basic Reed-Muller, the arithmetic and Walsh transform matrices.

Figure 2 shows the flow graphs of the FFT-like Cooley-Tukey algorithm for the computation of the Reed-Muller spectrum of a three-variable function f given by the truth vector $F = \left[f(0), f(1) \ldots f(7) \right]^T$.
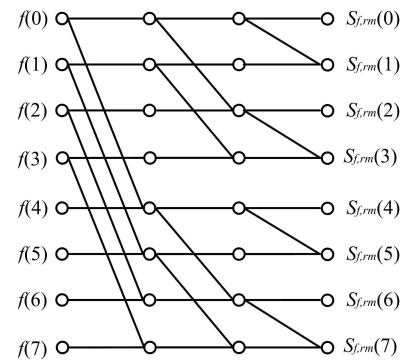
Fig. 2. The flow graphs of the FFT-like Cooley-Tukey algorithm for computing the Reed-Muller spectrum of a tree-variable function.

The constant geometry (CG) FFT-like algorithms are based on a factorization of the transform matrix into identical factor matrices [2]. Therefore, the indices of the butterfly operations are fixed for each step producing lower arithmetic complexity of algorithm. Because the results of butterfly

operations cannot be written in the same memory locations where the function is stored, implementations of this class of algorithms results in increased memory requirements compared to the Cooley-Tukey algorithms.

Figure 3 show the flow graphs of the FFT-like constant geometry algorithm for computing the Walsh spectrum of a three-variable function f.
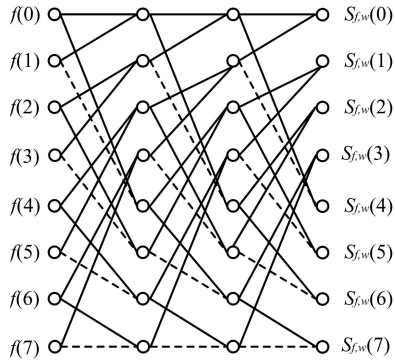
Fig. 3. The flow graphs of the FFT-like constant geometry algorithm for computing the Walsh spectrum of a three-variable function.

III. PARALLELIZATION OF THE FFT-LIKE ALGORITHMS OF SPECTRAL TRANSFORMS

The FFT-like algorithms have a large degree of parallelism in each step of the flow graph and according to this, their implementation on parallel computers has been well studied. A fundamental step in parallelizing the FFT-like algorithms on multicore CPUs is the mapping of array addresses to cores.

There are many approaches for implementation of parallel Cooley-Tukey class FFT-like algorithms on multi-processing elements and they can be categorized into three styles for parallelizing butterflies of FFT: block data mapping, cycling data mapping, and all butterflies in parallel mapping [7]. Figure 4 shows the flow graphs of the FFT-like algorithm of Cooley-Tukey class with various data mapping styles for the computation of Reed-Muller spectrum of a three-variable function f executed on two processing elements.

Block data mapping (BDM) style for parallelizing butterflies of Cooley-Tukey class of FFT-like algorithms (Figure 4a) uses an approach where the scope for parallelism increases with steps of the flow graph. In the first step, all the data is intertwined and all butterflies are performed on one processing element. In the second step, butterflies visibly break into two separate processing elements, and in the last stage there are two points of parallelization.

Cycling data mapping (CDM) style for parallelizing butterflies of Cooley-Tukey class of FFT-like algorithms (Figure 4b) uses an approach where the scope for parallelism decreases with steps of the flow graph. In the first step, all the butterflies are distributed individually to processing elements. In the second step, butterflies visibly break into two separate points of parallelization and, in the last step, there are four points of parallelization.

All butterflies in parallel mapping (BPM) style for parallelizing butterflies of Cooley-Tukey algorithms (Figure 4c) uses an approach where all butterfly computations can be performed in parallel. While this mapping style provides maximum scope of parallelism, there is a significant additional computation of memory addresses for each step producing higher overall arithmetic complexity of the algorithm.

Usually there is only one implementation style for parallel constant geometry class FFT-like algorithms, since the control flow is the same in each step.

Fig. 4. The flow graphs of the parallel FFT-like algorithm of CT class for computation of the FRT of the tree-variable function executed on two processing elements: (a) BDM, (b) CDM, (c) BPM.

IV. EXPERIMENTAL RESULTS

For multi-core CPU architectures, the model of parallel processing is based on a large number of processor cores with the ability to directly address into a shared RAM memory. The MPI framework has become a widely used standard, though not necessarily the best framework for parallel programming. For comparison purposes, we developed referent C++ and MPI implementations of FFT-like algorithms using two classes of algorithms and three implementation styles of the fast Reed-Muller, the fast Walsh, and the fast arithmetic transforms. The computations are performed on an Intel desktop PC working at 3.66 GHz with 12 GBs of RAM. The quad-core CPU that is used is an Intel i7 with hyper-threading, yielding 8 logical cores (threads).

We compared the performance of a multi-core CPU accelerated MPI implementations to a single-core CPU C++ implementations for a sample set of random logic functions. Since the computations are performed over vectors, processing times are independent of function values and, therefore, are performed using randomly-generated function truth vectors.

Table 1 shows computation performance of FFT-like algorithms using various classes and implementation styles discussed in the previous section. All times in the table are given in seconds. The data in the table are horizontally sorted in the increasing order of the number of functions variables.

TABLE I

COMPUTATION TIMES OF THE REED-MULLER, THE WALSH, AND THE ARITHMETIC TRANSFORM USING VARIOUS CLASSES OF FFT-LIKE ALGORITHMS AND IMPLEMENTATION STYLES ON SINGLE-CORE CPU AND MULTI-CORE CPU PLATFORM

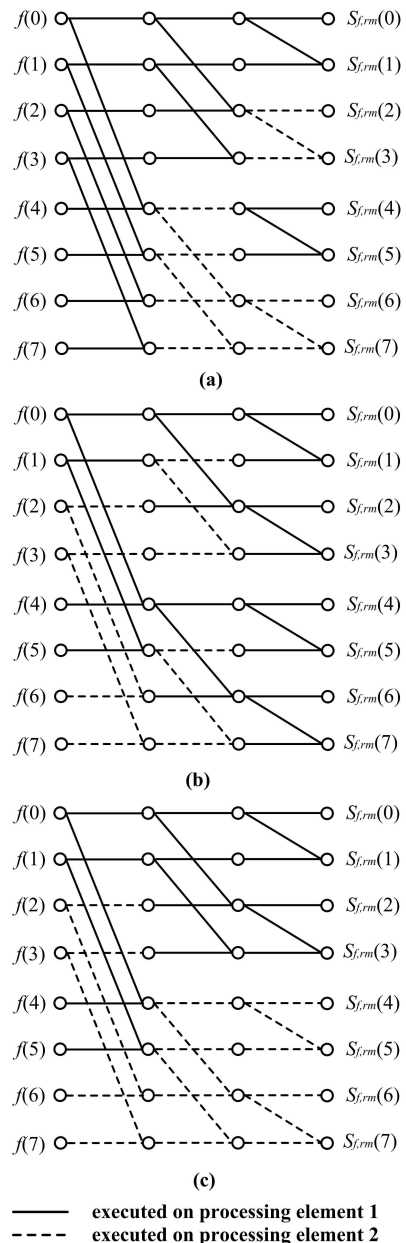| FFT-like algorithm				Computation time [s]				
Transform	Algorithm	Style	CPU cores	Number of variables n				
				25	26	27	28	29
FRT	CT		1	0.6	1.3	2.0	3.9	10
FRT	CG		1	0.9	1.8	4.0	6.5	13
FRT	CT	BPM	8	0.6	1.2	1.9	3.8	-
FRT	CT	BDM	8	0.1	0.3	0.4	0.7	-
FRT	CT	CDM	8	0.9	2.0	4.3	8.6	-
FRT	CG		8	0.6	1.4	3.0	-	-
FWT	CT		1	0.8	1.4	3.2	6.5	9.3
FWT	CG		1	1.2	2.2	4.5	6.9	19
FWT	CT	BPM	8	0.8	1.3	2.0	4.0	-
FWT	CT	BDM	8	0.2	0.3	0.6	1.2	-
FWT	CT	CDM	8	1.2	2.4	4.8	9.7	-
FWT	CG		8	0.7	1.4	2.9	-	-
FAT	CT		1	0.5	1.0	1.7	3.8	7.2
FAT	CG		1	1.0	1.9	3.8	7.9	16
FAT	CT	BPM	8	0.6	1.1	1.9	3.8	-
FAT	CT	BDM	8	0.1	0.2	0.4	0.7	-
FAT	CT	CDM	8	1.1	2.1	4.1	8.5	-
FAT	CG		8	0.6	1.3	2.9	-	-

From the data in Table 1, it can be seen that, on this multi-core CPU platform, for all considered transforms, the BMD style for the Cooley-Tukey algorithms significantly reduces computation times when compared to single-core CPU implementation. In the case of CDM style for the Cooley-Tukey algorithms, computation times are increased for about 50 to 60%. Computation times of butterflies in BPM style are very close to the computation times in single-core CPU implementations. Table entries with dashes indicate that the implementation failed to complete for that particular benchmark because of running out of memory.

The constant geometry FFT-like algorithms are not suitable for an MPI parallel processing configuration, because they increase inter-shared memory communication delay.

V. CONCLUSION

Computing power can be substantially increased through the exploitation of the inherent parallelism available in FFT-like calculations. However, an experimental performance analysis of the parallel FRT, FWT, and FAT algorithms has not been sufficiently investigated in a multi-core CPU environment. In this paper, we experimentally evaluated various implementations of fast algorithms for three different spectral transforms for logic functions. We analyzed the speedup obtained, by taking into account both the class of algorithms and the implementation styles.

It has been shown that the parallel MPI programming provides significant speedups only for block data mapping style of the Cooley-Tukey algorithms. Other implementation styles, for both Cooley-Tukey algorithms and constant geometry algorithms, have a negative impact on the performance of parallel program execution.

REFERENCES

[1] M. A. Thornton, R. Drechsler, and D. M. Miller, *Spectral Techniques in VLSI CAD*, Springer, 2001.

[2] M. G. Karpovsky, R. S. Stanković, and J. T. Astola, *Spectral Logic and Its Applications for the Design of Digital Devices*, Wiley, 2008.

[3] S. L. Hurst, D. M. Miller, and J. C. Muzio, *Spectral Techniques in Digital Logic*, Bristol, Academic Press, 1985.

[4] M. Frigo, S. G. Johnson, "The Design and Implementation of FFTW3", Proc. of the IEEE, vol. 93, no. 2, pp. 216-231, 2005.

[5] C. Brunelli, R. Airoldi, and J. Nurmi, "Implementation and Benchmarking of FFT Algorithms on Multicore Platforms", Proc. Int. Symposium on System on Chip, pp. 59-62, 2010.

[6] Y. Zhou, J. Zhang, and D. Fan, "Software and Hardware Cooperate for 1-D FFT Algorithm Optimization on Multicore Processors", Proc. Int. Conf. on Computer and Inf. Technology, vol. 1, pp. 86-91, 2009.

[7] E. C. Chu, A. George, *Inside the FFT Black Box: Serial and Parallel Fast Fourier Transform Algorithms*, CRC Press, 2000.

[8] J. Astola, R. S. Stanković, *Fundamentals of Switching Theory and Logic Design: A Hands on Approach*, USA, Springer, 2006.

[9] K. R. Rao, D. N. Kim, and J. J. Hwang, *Fast Fourier Transform – Algorithms and Applications*, Springer, 2010.

[10] A. Deb, S. Ghosh, *Power Electronic Systems – Walsh Analysis with MATLAB*, CRC Press, 2014.

A Performance Comparison of Computing LU Decomposition of Matrices on the CPU and the GPU

Dušan B. Gajić[1], Radomir S. Stanković[1], Miloš Radmanović[1]

Abstract – **This paper presents a comparison of time efficiency of different implementations of LU decomposition of matrices, performed on either central processing units (CPUs) or graphics processing units (GPUs). For processing on the CPU, we use the Eigen C++ linear algebra template library, Intel Math Kernel Library (MKL), and MATLAB (MATrix LABoratory). For performing LU decomposition on the GPU, we employ MATLAB's Parallel Computing Toolbox and Nvidia CUDA (Compute Unified Device Architecture) augmented with CULA (CUDA Linear Algebra) programming library. Processing times are compared using randomly-generated single-precision floating point matrices of size up to 16384×16384. The experiments show that the CUDA/CULA GPU implementation offers best performance for matrices of size up to 8192×8192. This implementation is on average 2.03 times faster than for the second-best performing implementation (MATLAB's Parallel Computing Toolbox on the GPU) and 11.82 times faster than the worst-performing implementation (Eigen on the CPU). For problem instances which cannot be stored in the global GPU memory (matrices larger than 8192×8192 on the used GPU), the LU decomposition is performed only on the multicore CPU, where Intel MKL proved to be 1.38 times faster than MATLAB and 2.72 times faster than Eigen.**

Keywords – **Performance comparison, LU decomposition, parallel computing, general purpose computations on graphics processing units, GPGPU.**

I. INTRODUCTION

The LU decomposition or LU factorization decomposes a square matrix into a product of a lower triangular unit matrix (unit matrix has all entries on the main diagonal equal to 1) and an upper triangular matrix [7, 15]. This method, introduced by Turing in 1948 [18], is widely used for problems such as solving systems of linear equations, inverting matrices, and computing matrix determinants, which are critical parts of many problems in science and engineering [7, 15, 16]. Therefore, efficient computation of LU decomposition is of considerable importance in scientific computing to the degree that this task is one of the standard benchmarks used to measure the performance of top supercomputers in the world [2].

Currently, two main computing platforms are available for performing numerical algorithms, such as LU decomposition that is discussed in this paper. These are the central processing unit (CPU) and the graphics processing unit (GPU). Current CPUs are still based on the single instruction, single data (SISD) von Neumann architecture, although they are now multicore [8]. GPUs have a single instruction, multiple data (SIMD) manycore architecture, which became programmable for non-graphics general-purpose algorithms only in the last ten years [1, 10]. Distinctions in computing architectures lead to different performance when implementing the same algorithms. Further, specific characteristics of architectures motivated development of various programming frameworks tailored to take advantage of some of these characteristics. Therefore, it is compelling to perform a comparison of different implementations of LU decomposition computed on CPUs and GPUs.

This paper presents a performance comparison of computing the LU decomposition using three different programming frameworks for the CPU and two for the GPU. The aim of the paper is to identify the computing platform and programming framework which produces the fastest LU decomposition of single-precision real matrices.

The paper is organized as follows. First, we briefly review the method of LU decomposition in Section 2. Section 3 provides basic information about the Eigen, Intel MKL, and MATLAB programming environments for the CPU, and MATLAB's Parallel Computing Toolbox and CUDA/CULA programming frameworks for the GPU. Section 4 offers some implementations details. The experimental settings and results are presented in Section 5. Section 6 summarizes the results of the research reported in this paper.

II. LU DECOMPOSITION

LU decomposition is a method for factorizing a square matrix **A** into a product of two matrices - a lower triangular unit matrix **L** and an upper triangular matrix **U**. The method of LU decomposition can be extended to non-square matrices by adding the requirement that **U** must be a row echelon matrix [15]. In the general case, partial row reordering (pivoting) of **A** before decomposition is performed when necessary to ensure existence and numerical stability of LU factorization [7, 16].

Using the LU decomposition, a square matrix **A** of size $N \times N$ is factorized as

[1]Dušan B. Gajić, Radomir S. Stanković, and Miloš Radmanović are with the University of Niš, Faculty of Electronic Engineering, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mails: dusan.b.gajic@gmail.com, radomir.stankovic@gmail.com, milos.radmanovic@elfak.ni.ac.rs.

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{bmatrix} =$$

$$= \mathbf{LU} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ l_{21} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{N1} & l_{N2} & \cdots & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1N} \\ 0 & u_{22} & \cdots & u_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{NN} \end{bmatrix}. \quad (1)$$

Example 1. A given (3×3) matrix **A**

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 3 \\ 4 & -1 & 3 \\ -2 & 5 & 10 \end{bmatrix}$$

can be factorized by using the LU decomposition as

$$\mathbf{A} = \mathbf{LU} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -3 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ 0 & 0 & 4 \end{bmatrix}.$$

The LU decomposition is chosen for the research reported in this paper because of its generality. If we take the problem of solving a system of linear equations as an example, the LU decomposition can be directly applied for solving any square system of linear algebraic equations of the form

$$\mathbf{Ax} = \mathbf{b}, \quad (2)$$

where **A** is a given $N \times N$ matrix with the coefficients of the system, **b** is a given vector with the N constant terms, and **x** is a vector with the unknown solutions to be computed. After applying the LU decomposition, the system becomes

$$\mathbf{LUx} = \mathbf{b}. \quad (3)$$

The lower triangular system $\mathbf{Ly} = \mathbf{b}$ is then solved by forward substitution. Subsequently, the upper triangular system $\mathbf{Ux} = \mathbf{y}$ is solved by back-substitution to obtain the solution **x** of the original system [15].

III. PROGRAMMING ENVIRONMENTS

In this research, the LU decomposition is computed using Eigen, Intel MKL, and MATLAB on the multicore CPU. On the GPU, we employ MATLAB Parallel Computing Toolbox and CUDA extended with the CULA programming library.

A. Eigen

Eigen is an open source C++ library for linear algebra which includes functions for vector and matrix operations, numerical solvers, and other related algorithms [4]. It offers clean and expressive interface and supports explicit vectorization for programs using instruction set extensions such as Intel's Streaming SIMD Extensions (SSE) [4]. However, it lacks an in-built support for multithreaded processing on multicore CPUs. Only some of the Eigen's functions can exploit parallelism using Open Multi-Processing (OpenMP) [5, 14].

B. Intel Math Kernel Libraries

Intel MKL is a library of mathematical functions for Intel and compatible CPUs [9]. It includes Basic Linear Algebra Subprograms (BLAS) and Linear Algebra PACKage (LAPACK) routines, fast Fourier transform (FFT) algorithms, and vectorized math functions. Functions implemented in Intel MKL are optimized for Intel multicore processors and allow automatic multithreaded execution on available CPU cores [9]. Intel MKL has compilers for C, C++, and Fortran, and it is available for Windows, Linux, and Mac OS X.

C. MATLAB

MATLAB is an interactive numerical computing environment and a programming language developed by MathWorks [11]. It allows numerical computations, data analysis and visualization, as well as algorithm implementation using a high-level programming language [11]. It also contains a large set of toolboxes for performing computations in mathematics, signal and image processing, computational finance, parallel computing, etc. Therefore, it is widely used by both engineers and scientists.

D. MATLAB Parallel Computing Toolbox

MATLAB's Parallel Computing Toolbox allows the use of multicore CPUs, GPUs, and clusters, for parallel processing of computationally intensive algorithms [17]. This toolbox includes special high-level data types, parallel loops, and numerical algorithms, which permit concurrent execution of programs, through translation and execution of code using Message Passing Interface on the CPU [Pacheco] and CUDA on the GPU [17].

For more details on parallel computing on CPUs, as well as MPI and OpenMP programming frameworks, used by Eigen, Intel MKL, and MATLAB for parallel processing, we refer to [8, 14].

E. CUDA/CULA

CUDA is a parallel programming architecture, framework, and language, developed by Nvidia, for the purposes of implementing and processing general-purpose algorithms on graphics processing units (GPGPU or GPU computing). It supports a massively-parallel programming model constructed around the high-throughput, high-latency GPU architecture.

CULA is a GPU-accelerated library for linear algebra, built on top of the Nvidia CUDA programming framework [3]. It is developed by EM Photonics as two separate tools for dense and sparse linear algebra – CULAdense and CULAsparse [3].

For more details on GPU computing, which attracted a fast-growing interest of researchers in recent years, as well as the CUDA programming framework, we refer to [1, 10, 12, 13].

IV. IMPLEMENTATION DETAILS

For the implementation of LU decomposition, we used the following approach.

We generated square matrices with random single-precision floating numbers using the *rand* function available in MATLAB and MATLAB's Parallel Computing Toolbox. For the same purposes in the Eigen, Intel MKL, and CUDA/CULA implementations, we used the *rand* function from *cstdlib*, with pseudo-random generator number seed set using *srand*(*time*(*NULL*)) function.

For computing the LU decomposition on the CPU, we used the *partialPivLu* function in Eigen, and the [L, U, P] = *lu(A)* command in MATLAB. In Intel MKL, we called the *LAPACKE_dgesv* routine, with matrix **A** set to be in the row major format. For performing the LU decomposition using MATLAB's Parallel Computing Toolbox on the GPU, we stored the matrix **A** in the *gpuArray* data structure [17]. In CUDA/CULA, we first transferred **A** to the GPU using pinned memory, in order to effectively use the PCIe bus between the CPU and the GPU [6, 10], and then called the *culaDeviceSgetrf* function from the CULA library.

V. EXPERIMENTS

A. Experimental Settings

The experiments were carried out using the computing platform and software presented in Table I. The LU decomposition was performed on matrices of size $N \times N$, for $N = 1024, 2048, 4096, 8192, 16384$. The elements of the matrices were single-precision floating point numbers, each represented by 4 bytes in the memory. The presented computational times represent average values for 10 program executions for each size of the input matrix.

TABLE I
THE EXPERIMENTAL PLATFORM AND SOFTWARE VERSIONS

CPU core frequency number of cores	*Intel Core i7-920* *2.66 GHz* *4*
Memory	*12 GB DDR3-2000 MHz*
Operating System	*Windows 7 Ultimate 64-bit*
GPU core frequency memory number of cores compute version	*Nvidia GTX 650 Ti* *900 MHz* *1 GB GDDR5 4.2 GHz* *384* *2.1*
Library/Software version	
Eigen MATLAB Intel MKL CUDA CULA	*3.2.4* *2015a* *11.2 (in Intel Parallel Studio 2015XE)* *6.5* *CULAdense R18*

B. Experimental Results

Results of the experiments preformed as described in the previous subsection are shown in Table 2. Dashed lines in the table indicate that the computation could not be performed for the corresponding matrix size. Figure 1 shows the same information graphically.

TABLE II
PROCESSING TIMES FOR DIFFERENT IMPLEMENTATIONS OF
LU DECOMPOSITION ON THE CPU AND THE GPU

	Processing time [ms]				
	CPU			GPU	
N	Eigen	MATLAB	Intel MKL	MATLAB GPU	CUDA/ CULA
1024	73	36	24	21	17
2048	485	254	204	98	55
4096	3409	1779	922	739	240
8192	24561	12843	5965	-	1230
16384	189188	96813	46374	-	-

When computing on the multicore CPU, the Intel MKL is the fastest implementation. This is due to the automatic multithreaded processing of MKL's numerical routines, if a multicore processor is present in the system, and its optimization for execution on Intel processors. The Intel MKL implementation was on average 1.38× and 2.72× faster than MATLAB and Eigen, respectively. The Eigen implementation is the slowest among the considered implementations because it lacks multithreading capability [5]. Therefore, the LU decomposition using Eigen was performed only on a single core of a quad-core CPU available in our test platform. Note that, for the largest considered matrices ($N = 16384$), we were able to perform the LU decomposition only on the CPU, due to the GPU memory limitations [6, 10].
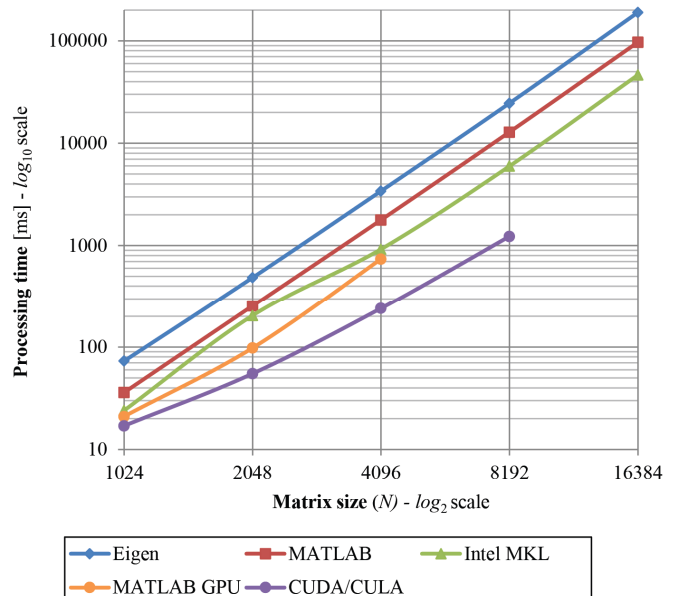

Fig. 1. Processing times for different implementations of
LU decomposition on the CPU and the GPU.

The CUDA/CULA GPU implementation is the fastest among all considered CPU and GPU implementations, with the exception of the largest considered matrices, which could not be computed on the GPU, due to the lack of memory. This is a major limitation for computing on the GPU, since it offers shortest processing times, but only for matrices which can be stored in its global memory. This restraint is even stronger when using the MATLAB Parallel Computing Toolbox GPU implementation, which could not handle matrices of size $N = 8192$, as a consequence of additional memory requirements for storing high-level MATLAB matrix representation (*gpuArray*).

The CUDA/CULA implementation is, for the considered matrix sizes, on average $2.03\times$ faster than the second-best implementation – the MATLAB Parallel Computing Toolbox on the GPU. The difference in speed between these two implementations increases with the size of matrix – from $1.23\times$ for $N = 1024$ to $3.07\times$ for $N = 8192$.

We can see that a significant difference in speed exists between the two GPU implementations, even though they both, on the low-level, use CUDA. This can be attributed to the additional time needed to translate high-level MATLAB code to CUDA, as well as to inefficiencies in implementation due to the automatic generation of CUDA code. On the other hand, the CUDA/CULA implementation uses CUDA functions and data structures directly and allows full control of the corresponding code. We can conclude here that the price to pay for faster program execution is extended CUDA/CULA program development time in comparison to using MATLAB's Parallel Computing Toolbox.

Further, when we compare the CUDA/CULA implementation with the considered CPU implementations, we can observe that it is, on average, $11.82\times$, $6.15\times$, and $4.34\times$ faster than the Eigen, MATLAB, and Intel MKL, respectively.

VI. CONCLUSIONS

In this paper, we presented a comparison of time efficiency of computing the LU decomposition using three different implementations for CPUs and two for GPUs. The experiments performed on matrices with randomly-generated single-precision floating-point numbers showed that the CUDA/CULA GPU implementation offers shortest computation times, but only for matrices which can be stored in the GPU global memory ($N \leq 8192$ on our test platform). The LU decomposition on larger matrices could be performed only on the available multicore CPU, where Intel's MKL is the fastest implementation due to multithreaded execution optimized for Intel processors.

We can conclude that, when processing time is a critical parameter, the LU decomposition should be performed on the GPU using highly-optimized linear algebra libraries like CULA, whenever matrices can be stored in the GPU's memory. In other cases, Intel's MKL is the best out of the considered solutions for performing the LU decomposition on multicore CPUs.

REFERENCES

[1] A. R. Brodtkorb, M. L. Sætra, T. R. Hagen, "Graphics processing unit (GPU) programming strategies and trends in GPU computing", J. of Parallel and Distributed Computing, vol. 73, no. 1, pp. 4-13, 2013.

[2] J. Dongarra, M. Faverge, H. Ltaief, P. Luszczek, "Achieving numerical accuracy and high performance using recursive tile LU factorization", Concurrency and Computation: Practice and Experience, vol. 26, no. 7, pp. 1408–1431, 2014.

[3] EM Photonics, CULA Programming Guide, available from: *http://www.culatools.com/cula_dense_programmers_guide/* [accessed 16 March 2015], version 18, April, 2014.

[4] Eigen C++ Template Library for Linear Algebra, available from: *http://eigen.tuxfamily.org/index.php?title=Main_Page* [accessed March 16, 2015], version 3.2.4, January 2015.

[5] Eigen and Multithreading, available at: http://eigen.tuxfamily.org/dox/TopicMultiThreading.html [accessed March 20, 2015].

[6] D. B. Gajić, R. S. Stanković, "Computing the Vilenkin-Chrestenson transform on a GPU", in J. of Multiple-Valued Logic and Soft Computing, Old City Publishing, Philadelphia, USA, vol. 25, no. 1-4, pp. 317-340, 2015.

[7] G. Golub, C. Van Loan, *Matrix Computations*, 3rd edition, The Johns Hopkins University Press, 1996.

[8] J. L. Hennessy, D. A. Patterson, *Computer Architecture: A Quantitative Approach*, 5th edition, Morgan Kaufmann, 2011.

[9] Intel Corporation, Intel Math Kernel Library Reference Manual, available from: *https://software.intel.com/sites/default/files/ managed/9d/ c8/mklman.pdf* [accessed March 20, 2015].

[10] D. Kirk, W. M. Hwu, *Programming Massively Parallel Processors: A Hands-on Approach*, Morgan Kaufmann, 2010.

[11] MathWorks, MATLAB, available from: http://www.mathworks.com/products/matlab [accessed March, 16, 2015], version 8.5 (R2015A), March, 2015.

[12] Nvidia, *Nvidia CUDA Programming Guide*, available from: *http://docs.nvidia.com/cuda/pdf/CUDA_C_Programming_Guid e.pdf* [accessed March, 16, 2015], version 6.5, August, 2014.

[13] J. Owens, M. Houston, D. Luebke, S. Green, J. Stone, J. Phillips, "GPU computing", Proc. of the IEEE, vol. 96, no. 5, pp. 279–299, 2008.

[14] P. Pacheco, *An Introduction to Parallel Programming*, Elsevier, 2011.

[15] D. Poole, *Linear Algebra - A Modern Introduction*, 2nd edition, Brooks/Cole, Thomson, 2006.

[16] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, 3rd edition, Cambridge University Press, 2007.

[17] J. W. Suh, Y. Kim, *Accelerating MATLAB with GPU Computing – A Primer with Examples*, Morgan Kaufmann – Elsevier, 2014.

[18] A. M. Turing, "Rounding-off errors in matrix processes", The Quarterly J. Mechanics and Applied Mathematics, vol. 1, part 3, pp. 287–308, September, 1948.

A Promethee-based Approach to Multi-Criteria Flexible Job Shop Scheduling Problem

Vassil Guliashki [1] and Leoneed Kirilov [2]

Abstract – In this paper an approach combining a genetic evolutionary algorithm with Promethee-based evaluation of generated solutions is introduced. Pair-wise comparisons of criteria values are used for calculation of fitness measure for all solutions in the current population. The proposed new Promethee-based approach gives a possibility for better quality evaluation of generated solutions and achieving better parent-selection mechanism in the genetic algorithm. It also reduces very much the efforts of Decision Maker (DM) in the choice of final compromise solution.

Keywords – **Multi-Criteria optimization, Job Shop Scheduling, Multi-Criteria Decision Analysis, Promethee method.**

I. INTRODUCTION

The job shop scheduling problem is well-known from operations research and computer science and is of high practical value with applications in many real-life situations [2, 13]. While first approaches in this area consider optimality of schedules for a single objective function, multi-objective formulations of the problem have become gradually of increasing importance [5]. A theory of multi-criteria scheduling is presented in [14]. A survey of methods for job shop scheduling using multi-criteria decision making is presented in [12]. In this paper the multi-criteria flexible job shop scheduling problem as an extension of the popular multi-criteria job shop scheduling problem is considered. During the last decades many researchers have devoted considerable efforts to developing evolutionary multi-criteria algorithms.

The problem of scheduling arises when planning and controlling the decision-making process of manufacturing and service industries. It can be schematized as follows: There is a number of N jobs to be executed. Each job consists of a given sequence of operations which needs to be performed using a number of M machines. All operations for each job must be performed in the order given by the sequence. Each operation demands the use of a particular machine for a given time. Each machine can process only one operation at a time. The goal is to find a schedule optimizing the above problem according to the given objective function (cost function, make-span, tardiness, maximal workload etc.). Scheduling consists of assigning each operation of each job a start time and a completion time on a time scale of the machine with the preference relations.

The most used in practice is the job shop scheduling problem. It is a difficult computational problem. Optimal solutions for job shop scheduling can be found in polynomial time if the number of jobs is 2, or if the number of machines is 2 and all jobs have 1 or 2 operations, or if the number of machines is 2 and all operations have duration 1. In all cases the problem obtained by incrementing the number of machines, jobs, operations or durations by 1, is NP-hard [6, 9]. Below are presented the basic formulations of the classical job shop problem (JSP) and of the flexible job shop problem (FJSP):

Job shop problem (JSP)

The JSP is formulated as follows: There is given a set of n jobs: $J_1, \ldots, J_n$, which have to be performed on m machines $M_1, \ldots, M_m$.

For each job there is given the operative consequence of the jobs composing this job. Namely:

$J_i = (O_{i,1}, \ldots, O_{i,j(i)})$, $j(i)$ is the number of operations for the corresponding job, $i=1,\ldots,n$.

It is well-known which operation on which machine should be executed. Therefore another formulation of this model is:

$$J_i = (M_{i,1}, \ldots, M_{i,j(i)}),$$

The processing times for each possible operation on each machine are known: $p_{i,k}$, $i=1,\ldots,n$; $k = 1, \ldots, j(i)$.

The optimal schedule according preliminary given criterion (criteria) has to be found. For example one criterion could be the minimization of make-span (time window) – C_{max}.

This is the most often used and chronologically the earliest developed model – see for example [6, 8].

Flexible job shop problem (FJSP)

This model represents an extension of the above job shop problem. Here each operation can be executed not only on one machine, but on a given subset of machines. This subset is naturally different for each operation. In other words, it is not a priori known which operation on which machine should be performed.

This model is closer to real life production situations and could be applied, when some or all machines are multi-functional (multitasking) – i.e. they could perform more than one operation (not at the same time) with corresponding different processing times. Among the first researchers suggesting this model are Bruker and Schlie – [3].

As noted in [4] the FJSP is a problem of high complexity and practical value, and it has been widely investigated for the last two decades. Researches on its multiobjective version started about ten years ago, but most studies focused on searching for the single optimal solution with respect to a certain aggregated objective. Research works aiming at obtaining the set of Pareto optimal solutions appeared during the recent three years.

There are two variants of FJSP – [11]:

[1]Vassil Guliashki is with the Institute of Information and Communication Technologies – BAS, "Acad. G. Bonchev" Str. Bl. 2, 1113 Sofia, Bulgaria, E-mail: vggul@yahoo.com

[2]Leoneed Kirilov is with the Institute of Information and Communication Technologies – BAS, "Acad. G. Bonchev" Str. Bl. 2, 1113 Sofia, Bulgaria, E-mail: l_kirilov_8@abv.bg

First, when each operation of each job can be executed on any, no matter which, machine. This case is relevant to the total / global flexibility (total flexible job shop problem – T-FJSP).

Second, even not each (but at least one) operation can be performed on any machine. This case refers to the partial flexibility (partial flexible job shop problem – P-FJSP).

Below is considered the multi-criteria FJSSP job problem, where n jobs J (J_i, $i \in \{1,2,\ldots,n\}$) should be processed on m existing machines M(M_k, $k \in \{1,2,\ldots,m\}$). The job J_i consists of n_i operations. For each of these operations (O_{ij}) a predetermined set of capable machines is considered (M_{ij}). One of the capable machines should be selected to perform the operation. The processing time and the start time of operation j (O_{ij}) of job J_i on machine k are denoted by p_{ijk} and t_{ijk} respectively. The assignment decision variables are denoted by X_{ijk}, and the completion time on machine k is denoted by C_k. The problem includes three criteria, which have to be minimized: the makespan (C_{max}), the critical machine workload (CWL) and the total work load of machines (TWL).

$$C_{max} = \max \{C_K \mid k = 1,\ldots,n\} \qquad (1)$$

$$CWL = \max \{ \sum_{i=1}^{n} \sum_{j=1}^{n_i} P_{ijk} X_{ijk} \} \ k: 1,2,3,\ldots,m \qquad (2)$$

$$TWL = \sum_{i=1}^{n} \sum_{j=1}^{n_i} \sum_{k=1}^{m} P_{ijk} X_{ijk} \qquad (3)$$

For this multi-criteria FJSSP we make the following assumptions:

- There is a predetermined and fixed order for the operations of each job.
- There isn't assumed priority restriction among the operations of different jobs, as well as among the jobs.
- At the beginning (at time 0), jobs are released and machines are available.
- Move times between operations and setup times of machines are ignorable.
- Only one job can be processed on each machine at each specific moment and during the process, operations can't be broken off.

The paper is organized further as follows: In section 2 we give some information about the Promethee I method, which is used to estimate and reorder different alternatives. In section 3 we present the Promethee-based approach for multi-criteria FJSP optimization. An illustrative example is presented in section 4. In section 5 we consider some challenges connected with the use of this approach and draw some conclusions.

II. THE PROMETHEE-BASED ESTIMATION OF ALTERNATIVES

In this paper the PROMETHEE I method [10, 15] is used to estimate and reorder the different alternatives. It consists in pairwise comparisons. Let T be a finite set of possible alternatives. There are three possible relations between the alternatives **a** and **b**, where $a \in T$ and $b \in T$. These relations are:

preference, indifference and incomparability. They will be denoted by *Pr*, *Ind* and *Inc* respectively. Let be considered the particular criterion f(.), which has to be maximized. The effective choice between the alternatives is made interactively by the DM or by an analyst, according to their feeling of the intensities of preference between alternatives. In this connection the following parameters have to be fixed:

q - a threshold defining an indifference area,

v - a threshold, defining a strict preference area,

s - a parameter, which value lies between q and v.

Let d = f(a) – f(b). The preference function Pref(a,b) will be considered. It gives the intensity of preference of a over b in function of the deviation d.

A generalized criterion H(d) is associated to each criterion. Six different types of H(d) are defined in the PROMETHEE I method. In this paper the generalized criterion H(d) of type 5 is used, known as V-shape with indifference area. It is formulated as follows:

$$H(d) = \begin{cases} 0 & |d| \le q \\ (|d|-q)/(v-q) & q < |d| \le v \\ 1 & |d| > v. \end{cases} \qquad (4)$$

This criterion has been often used. For this type of H(d) the intensity of preference increases linearly between q and v.

Let the preference index π(a,b) of **a** over **b** over all the criteria be defined in the form:

$$\pi(a, b) = \sum_{l=1}^{k} w_l H_l(a,b), \qquad (5)$$

where w_l, $l=1,\ldots,k$; are weights associated to each criterion and $\sum_{l=1}^{k} w_l = 1$.

Here it is assumed that the weights w_l, $l=1,\ldots,k$; are equal. In this case π(a,b) is simply the arithmetic mean of all the intensities of preference $Pref_j(a,b)$, j = 1,…,k;. For each pair (a,b) the values π(a,b) are calculated. Then for each alternative $a \in T$ the positive outranking flow $\Phi^+(a)$ is calculated as follows:

$$\Phi^+(a) = \sum_{x \in T} \pi(a,x). \qquad (6)$$

In our problem the preference index π(b, a) = –π(a, b). The corresponding negative outranking flow in the problem (1)-(2) is symmetric to $\Phi^+(a)$ and has the opposite direction:

$$\Phi^-(a) = -\Phi^+(a). \qquad (7)$$

Hence the positive outranking flow is enough to expresses how each alternative is outranking all the others. The higher $\Phi^+(a)$, the better is the alternative. $\Phi^+(a)$ represents the power of a, it gives its outranking character.

To arrange all explored alternatives in an order according their preference over all the criteria (all the objectives) in the problem (1)-(2) the positive outranking flow $\Phi^+(a)$ is calculated for each explored alternative $a \in T$. Then the alternatives are reordered in a non-increasing order of their Φ^+ values.

III. THE PROMETHEE-BASED APPROACH TO MULTI-CRITERIA FJSSP

For problems with complex, non-smooth and multimodal objective function, where the information of objective

function derivatives is not accessible, have been developed global heuristic algorithms, as well as evolutionary (population based) algorithms, such like Genetic algorithms (see [7]), Scatter search, Ant systems and Particle swarm optimization algorithms. Their main common features are as follows:

- They memorize solutions (or characteristics of solutions) in a population of individuals. Each individual is associated with a feasible solution of the problem at hand.
- They include a generating solutions search procedure, which uses the information (implicitly) stored in the population.

Here is proposed an approach combining a genetic-type evolutionary algorithm with PROMETHEE I – estimation and ranking the generated solutions in the population (considering them as different alternatives), according the calculated $\boldsymbol{\Phi}^+$–values. The corresponding algorithm is called PBGALG.

To start the calculations the shifting bottleneck procedure (see [1]) is used to minimize the makespan C_{max} and the obtained best solution is used as initial solution in PBGALG.

To generate new feasible solutions (individuals) for the current population we use a *selection* operator and a *modified crossover* operator. For *selection* there aren't used two parent solutions, but only the best current solution is taken to be used for offspring generation. The best solution is chosen according the calculated $\boldsymbol{\Phi}^+$–values by means of Promethee - estimation.

The *modified crossover* is based on the representing the corresponding schedule (feasible solution) as a string of consecutive operations (O-string), executed on the given machines. In case more than one operation have one and the same start time t_{ijk} then the operation with lower job index and lower operation index precedes the operations with greater indices in the operation string. Then a number of l operations in this string are chosen randomly to generate l new feasible solutions. Here l depends on the population length. A check is performed if the corresponding operation can be executed on another machine, different from that one used in the current schedule. If there are other possible machines this operation is fixed to be performed on the machine, requiring the minimal processing time for the given operation. After that the operations after this operation in the operation string are reordered by means of the following heuristic to generate a new feasible schedule-solution.

Heuristic procedure:

Let we denote one operation by O_{ikj}, where first index shows to which job corresponds the operation, the second index gives the place of the operation in the operative consequence for the corresponding job and third index denotes the O-string position. Let v be the O-string position of operation, which should be fixed currently, and let s_{total} is the total number of all operations, which should be executed.

Basic cycle

For $j = v+1$, s_{total}

Among the operations still unassigned to machines find the minimal index: $k = \min\{k_1,\dots,k_i\}$

Machine-Selection:

Let the possible machines for operation $O_{i,k,j}$ are $m(1, i(k))$, ..., $m(r(i(k)), i(k))$ with the corresponding

summary processing times up to the moment $P(1, i(k))$, ..., $P(r(i(k)), i(k))$.

Select the machine $m(u, i(k))$ such that
$$P(u, i(k))+p_{i,k,j} = \min \{\ P(1, i(k))+p_{i,k,j},\ ..., P(r(i(k)), i(k))+p_{i,k,j}\ \}.$$

Select the machine for execution, which has the minimal processing time from the set of possible machines. Assign (fix) the operation $O_{i,k,j}$ to be executed on this machine.

end Machine-Selection

Actualization: $P(u, i(k)) = P(u, i(k)) + p_{i,k,j}$.

end j

In this way a new feasible solution (new schedule) is generated, which strictly differs from the parent solution.

The pseudo-code of PBGALG can be written in the following form:

Begin

Initialize the population P

Evaluate the individuals in P (by means of Promethee I – procedure)

Sort P according to the fitness value.

While *no stopping criterion is met, do*

Repeat: *Perform parent selection.*

Apply the modified crossover operator to generate offspring individuals.

Select individual s to survive.

If *s is better than one individual in P*

Replace the old individual in P by x_j. The check is performed starting by the worst individual.

EndIf

EndWhile

End

IV. AN ILLUSTRATIVE EXAMPLE

We consider the following test example.

3 machines - M_1, M_2, M_3

5 jobs - $J_1(O_1, O_2, O_3)$, $J_2(O_2, O_3, O_4)$, $J_3(O_1, O_2, O_3, O_4)$, $J_4(O_2, O_4)$, $J_5(O_1, O_2, O_4)$,

Note that some operations are identical.

The processing times are given in the (O, M) matrix:

	M_1	M_2	M_3
O_1	10	12	7
O_2	8	6	X
O_3	X	5	11
O_4	3	4	X

The symbol "X" denotes that the operation cannot be processed at the corresponding machine and vice versa.

According to the proposed algorithm we start the solution process generating an initial solution by the shifting bottleneck heuristic [1]. The obtained schedule is the following:

M1 (O31), (O32), (O12), (O24), (O34), (O54)
time: 10   18   26   29   32   35

M2 (O22), (O42), (O44), (O51), (O52),
time: 6   12   16   28   34

M3 (O11), (O23), (O33), (O13)
  time: 7   18   29   40

The number of operations is s_{total} = 15. The obtained makespan-value is C_{max} = 40.

The corresponding operation string is:

O22,O11,O31,O42,O23,O32,O44,O51,O12,O33,O24,O52,O13,O34,O54;

To create an initial population with l = 4, there are chosen randomly 3 operations: O23, O44 and O24 (correspondingly the 5-th, 7-th and 11-th in the string).

For each of them, there is available one possible other machine for execution (M2, M1 and M2).

Applying the above heuristic procedure we obtain the following new solutions:

1) M1 (O31), (O32), (O52), (O24), (O44), (O54)
      time: 10   18   26   29   32   35
    M2 (O22), (O42), (O12), (O23), (O33), (O34)
      time: 6   12   18   23   28   32
    M3 (O11), (O51),   (O13)
      time: 7   14   18-29

2) M1 (O31), (O32), (O44), (O24), (O34), (O54)
      time: 10   18   21   24   28-31 34-37
    M2 (O22), (O42), (O12), (O13), (O33), (O52)
      time: 6   12   18   23   28   34
    M3 (O11), (O23), (O51)
      time: 7   18   25

3) M1 (O31), (O32), (O12), (O52), (O54)
      time: 10   18   26   28-36   39
    M2 (O22), (O42), (O44), (O51), (O24) (O13), (O34)
      time: 6   12   16   28   32   37   41
    M3 (O11), (O23), (O33)
      time: 7   18   29

To evaluate the obtained solutions by the Promethee I - procedure, we calculate the following table of alternatives:

Alternative	*TWL*	*CWL*	*C_{max}*
1 (shift. bottl.)	109	40	40
2	92	35	35
3	92	34	37
4	107	41	41

This table is used by PBGALG to calculate the corresponding Φ^+–values through the Promethee – estimation. The following result is obtained:

Φ^+(1)= 0,078; Φ^+(2)=1,833; Φ^+(3)=1,74; Φ^+(4)=0,04

Hence the obtained solutions are ranked in the following order: Alternative 2, alternative 3, alternative 1 and alternative4. This result is presented to Decision Maker, who decides if the search process should be terminated or should continue. In case the Decision Maker is satisfied with at least one of the obtained solutions he/she can stop the calculations.

V. Conclusion

The presented Promethee-based approach is useful for solving multi-criteria FJSSP. It can facilitate very much the Decision Maker in the choice of best final solution.

One direction for further research is the possibility for determined choice of operations in the *modified crossover*.

One such rule, for example, could be: among the operations having minimal processing time on a machine different from the currently used that one with maximal difference between the current and the minimal processing time to be chosen first.

References

[1] Adams J., E. Balas and D. Zawack, (1988), "The Shifting Bottleneck Procedure for Job Shop Scheduling", *Management Science*, Vol. 34, No. 3, pp. 391-401.

[2] Blazewicz J., K. H. Ecker, E. Pesch, G. Schmidt, and J. Weglarz. *Scheduling Computer and Manufacturing Processes.* Springer Verlag, Berlin, Heidelberg, New York, 2. edition, 2001.

[3] Bruker, P., R. Schlie (1990) Job shop with multi-purpose machine, Computing, vol. 45, 1990, pp. 369-375.

[4] Chiang, Tsung-Che, Lin, Hsiao-Jou (2013) A simple and effective evolutionary algorithm for multiobjective flexible job shop scheduling, *Int. J. Production Economics*, 141, 87-98.

[5] Daniels, R., Incorporating preference information into multi-objective scheduling. *European Journal of Operational Research*, 77:272-286, 1994.

[6] Garey, M., D. Johnson, and R. Sethi. (1976) The Complexity of Flowshop and Jobshop Scheduling. *Mathematics of Operations Research*, 1(2):117–129, 1976.

[7] Goldberg D. E. Genetic Algorithms in Search, Optimization and Machine Learning, Addison Wesley, Reading, Mass, 1989.

[8] Lawler, E., J. Lenstra, A. Rinnooy Kan, and D. Shmoys. (1993) Sequencing and scheduling: algorithms and complexity. In A.H.G. Rinnooy Kan S.C Graves and P.H. Zipkin, editors, *Logistics of Production and Inventory*, volume 4, Chapter 9, pages 445–522. Elsevier, 1993.

[9] Lenstra, J. K., A. R. Kan, P. Brucker, (1977), "Complexity of Machine Scheduling Problems", *Annals of Discrete Mathematics*, vol. 1, 1977, pp. 343-362.

[10] Moca M. and G. Fedak "Using Promethee Methods for Multi-Criteria Pull-based Scheduling on DCIs", In *Proc. of the 8th IEEE Int. Conf. on eScience*, Chicago, USA, Oct. 2012, pp. 1-8.

[11] Motaghedi-Iarijani, A., Sabri-Iaghaie K., Heydari, M. (2010) Solving job shop scheduling with multi objective approach, *Int. J. of Industrial Engineering and Production Research*, vol. 21, No 4, pp. 197-209.

[12] Parveen S. and H. Ullah, (2010), "Reviev on Job-Shop and Flow-Shop Scheduling Using Multi Criteria Decision Making", *Journal of Mechanical Engineering*, Vol. ME 41, No. 2, December 2010, Transaction of the Mech. Eng. Div., The Institution of Engineers, Bangladesh, pp. 130-146.

[13] Pinedo, M. *Planning and Scheduling in Manufacturing and Services.* Springer Verlag, Berlin, Heidelberg, New York, 2005.

[14] T'kindt, V., J.-C. Billaut. *Multicriteria Scheduling: Theory, Models and Algorithms.* Springer Verlag, Berlin, Heidelberg, New York, 2002, 2006.

[15] Vincke, P., *Multiple Criteria Decision Aid*, John Wiley& Sons, New York (1992)

ORAL SESSION
MEASUREMENT SCIENCE AND TECHNOLOGY

Software Based Procedure for Estimation of Measurement Uncertainty Applied to Power Quality Measurement

Milan Simić[1], Dragan Denić[2], Goran Miljković [3] and Dragan Živanović[4]

Abstract – **Software supported procedure for calculation and analysis of standard measurement uncertainty components, used for electrical power quality measurements, is presented in this paper. Experimental measurement system is functionally based on the virtual instrumentation concept. Reference measurement signals are provided by means of LabVIEW based generator of standard three-phase voltage waveforms, including simulation of typical power quality disturbances, which is already described in previously published paper [1]. Device used for measurement of three-phase voltage basic parameters is standard power quality analyzer - Fluke 435 Series II. Developed software application controls measurement, recording, graphical presentation and statistical processing of obtained measurement results, regarding to basic parameters of reference three-phase voltage waveforms. Statistical processing of obtained measurement results includes possibility for graphical presentation of recorded time diagrams and histograms of measured voltage parameters. Estimations of standard measurement uncertainty components are performed according to basic demands of relevant international documents for estimation and expression of uncertainty in measurements.**

Keywords – **Measurement uncertainty estimation, Electrical power quality measurement, Virtual instrumentation software.**

I. INTRODUCTION

The primary goal of electrical power suppliers is to fulfill basic demands of their final customers under all possible conditions and environments. Consistently, in order to avoid the high cost of equipment failures, all customers have to make sure that they obtain an electrical power supply of satisfactory quality. Also, customer electrical equipment and devices must be capable of functioning as required, even in cases when small level network disturbances are present. This can be provided only if limit values of specific power quality parameters are exactly specified. These limit values can be defined by specific demands of individual customers in the power supply quality contract with suppliers, or generally by national and international electrical power quality (PQ) standards, such as European quality standard EN 50160 [2,3].

[1]Milan Simić is with the University of Niš, Faculty of Electronic Engineering, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mail: milan.simic@elfak.ni.ac.rs.

[2]Dragan Denić is with the University of Niš, Faculty of Electronic Engineering, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mail: dragan.denic@elfak.ni.ac.rs.

[3]Goran Miljković is with the University of Niš, Faculty of Electronic Engineering, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mail: goran.miljkovic@elfak.ni.ac.rs.

[4]Dragan Živanović is with the University of Niš, Faculty of Electronic Engineering, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mail: dragan.zivanovic@elfak.ni.ac.rs.

Previous documents define the main characteristics in low and medium voltage networks under normal operating conditions. For example, according to EN 50160, a customer is allowed to complain, if the voltage quality at the supply delivery point does not satisfy standard demands and requirements. As a consequence, standard PQ parameters should be constantly monitored, processed and analyzed. These procedures must be performed by means of advanced measurement instruments and equipment. Instruments for measurement and analysis of basic PQ parameters, such as Fluke 435 or Chauvin Arnoux C.A. 8334, based on complex digital processing functions, are capable to allow continuous power supply quality monitoring and control. It have been designed to measure relevant quality parameters and to perform software based statistical activities on single phase or three-phase power distribution networks, in order to verify compliance with relevant quality standards. Such instruments can be used as single units at selected points in the network, or alternatively several separated units can be combined in complete distributed measurement system, for monitoring and analysis of power distribution networks [4].

In order to satisfy specified parameters and measurement accuracy level, these measurement devices must be followed by the satisfactory metrological traceability. Verification and testing of such instruments must be performed in appropriate metrological laboratories. Reference devices, such as voltage and current calibrators, are available in various functional and constructive solutions. Calibrators generate reference signals with high accuracy levels, which correspond to the secondary standards, laboratory and industrial standards in metrological traceability assurance chain. There are calibration instruments for some specific types of PQ analyzers, such as solutions of multifunctional calibrators Fluke 5520A and Fluke 6100B [5].

Experimental software supported system for measurement and analysis of basic PQ parameters, presented in this paper, can be applied for estimation of measurement uncertainty components defined by the ISO Guide to the Expression of Uncertainty in Measurement. Procedure is functionally based on virtual instrumentation software. Standard PQ analyzer Fluke 435 Series II is applied for measurement of reference voltage parameters, provided by LabVIEW based generator of standard three-phase voltage waveforms defined according to requirements of European PQ standard EN 50160. Control software application enables statistical analysis and graphical presentation of obtained experimental measurement results.

II. CONFIGURATION OF EXPERIMENTAL MEASUREMENT SYSTEM

Hardware block configuration of experimental acquisition system, for measurement and software supported processing of standard PQ parameters, is presented in Fig. 1. Three-phase

Fig. 1. Block configuration of experimental system for measurement and software processing of standard PQ parameters

reference voltage waveforms, with nominal standard 230V RMS voltage values and 50Hz frequency, are provided using LabVIEW based generator of PQ waveforms and network disturbances. This signal generator is presented and described in the previously published paper [1]. It can generate various three-phase voltage test waveforms, including some special functions for simulation of typical PQ network disturbances characteristic for real power distribution networks. Voltage test waveforms and PQ disturbances can be defined according to the criteria of European quality standard EN 50160 [3].

Basic control front panel for generation and presentation of reference three-phase voltage waveforms is shown in Fig. 2. Each type of standard disturbances, for example voltage swell, voltage sag and high-order harmonic components, can be individually defined and generated using separate functional segments. Individual signal disturbances can be combined and unified in the form of final complex test waveforms. Shown front panel of signal generator includes primary control switch for selection of specific disturbance categories, such as: slow voltage variations, voltage swells, voltage sags, interruptions, voltage transients, high-order harmonics, voltage swells with harmonics and voltage sags with high-order harmonics [1].

Presented solution of signal generator provides definition and adjustment of basic signal parameters. These functions are enabled by number of control buttons and knobs, implemented on front panel of signal generator. In this specific example are presented six control knobs for definition of signal frequency, amplitude levels, disturbance start and stop times, maximum levels of voltage swell and voltage sag. Reference three-phase waveforms are generated with standard nominal frequency of 50Hz, signal phase differences of $2\pi/3$ rad and nominal RMS voltage values of 230V. Separate segment of control knobs on generator front panel is used for selection and adjustments of amplitude levels related to individual high-order harmonic components. Content of specific high-order harmonic can be precisely defined by number of control knobs for regulation of amplitude levels, from third to eleventh high-order harmonics.

Developed signal generator can be used for generation of reference waveforms applicable in testing of instruments for measurement and analysis of standard PQ parameters and network disturbances. As an example, in this specific case are tested three-phase PQ analyzer Fluke 435 Series II, previously shown in Fig. 1. Reference three-phase voltage waveforms must be sent directly to voltage inputs of tested instrument
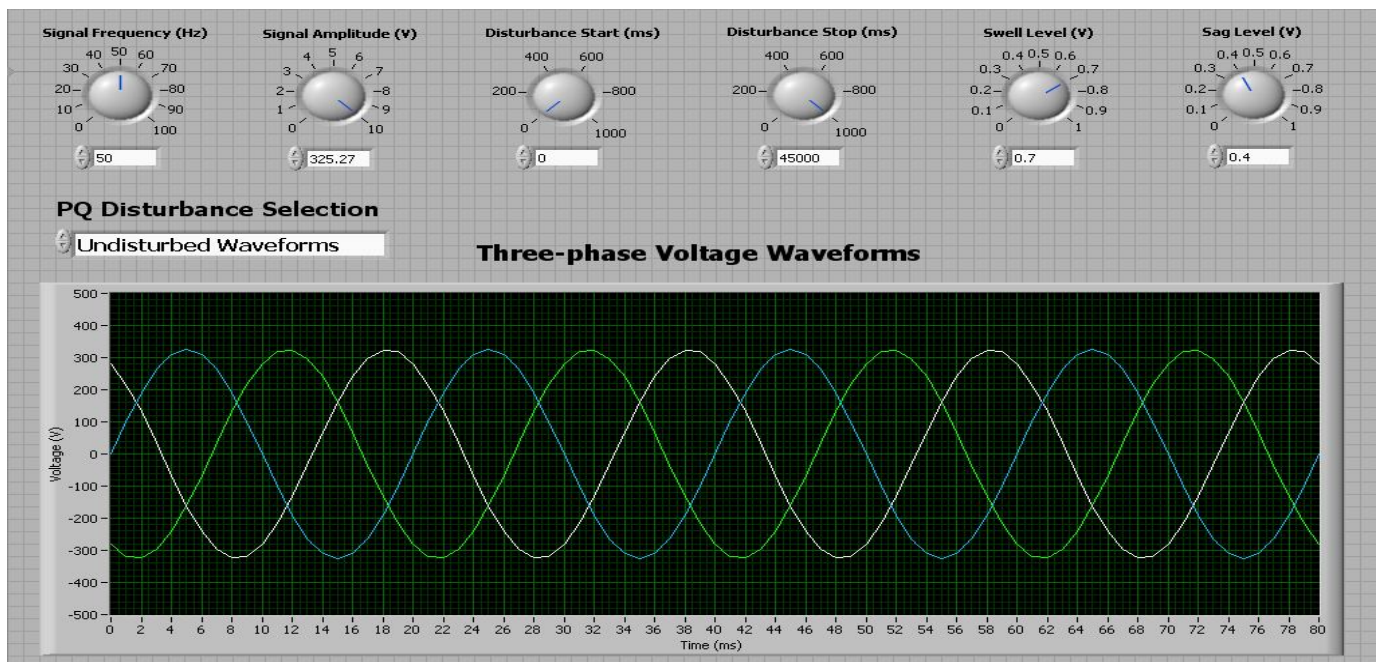

Fig. 2. LabVIEW control front panel of signal generator for presentation of standard three-phase voltage waveforms

Fluke 435. Direct two-way communication between tested instrument outputs and control computer is provided using standard USB interface. Measurement results obtained from instrument outputs can be simply transferred directly to the computer, recorded into system database and processed for graphical presentation. Some of characteristic experimental measurement results are shown in following paper segment.

III. PRESENTATION AND SOFTWARE ANALYSIS OF MEASUREMENT RESULTS

Experimental measurement system previously presented in Fig. 1 is software controlled using the virtual instrumentation concept. Application software, implemented in the control computer system, performs acquisition, recording, graphical presentation and statistical analysis of experimental results, previously obtained by measurement instrument Fluke 435.

Front panel of control software application, developed in LabVIEW environment, for graphical presentation of recorded time diagrams regarding to measured RMS voltage values per three phases of reference waveforms, is presented in Fig. 3. Beside recorded time diagrams of measurement data, control software application shows individual RMS voltage values, measured for three phases of reference voltage waveforms.

Software supported procedure for statistical processing and analysis of measured RMS voltage values is shown in Fig. 4. Specific statistical analysis includes two functional segments. First segment involves recording and presentation of statistical histograms for measured voltage values per phases. Second part of statistical analysis is calculation and presentation of standard measurement uncertainty components, defined by relevant international standards. Therefore, virtual instrument in Fig. 4 presents corresponding statistical histograms and estimated values of type A measurement uncertainty, type B measurement uncertainty and standard combined uncertainty per phases. Also, analysis includes detection and indication of minimum, maximum and average measured voltage values.

Presented average measured voltage values per phases are calculated as arithmetical means of individual measurement results per three phases, as is shown in a following equation:

$$V_{average} = \frac{1}{n}\sum_{i=1}^{n} V_i \qquad (1)$$

Statistical procedure applied for estimation of individual measurement uncertainty components is performed according to the recommendations of document Guide to the Expression of Uncertainty in Measurement [6], defined by International Organization for Standardization – ISO. Estimation of type A measurement uncertainty is performed according to statistical methods applied on obtained measurement results. Type A uncertainties are calculated by following square root equation for statistical standard deviation of measured voltage values:

$$u_A(V) = \sqrt{\frac{1}{n(n-1)}\sum_{i=1}^{n}\left(V_i - V_{average}\right)^2} \qquad (2)$$

Type B standard measurement uncertainties are estimated on the basis of maximum absolute errors, calculated from obtained measurement results per phases. In this specific case, considering maximum measured voltage values and nominal reference values of 230V, for type B standard measurement uncertainty values per phases are obtained following results:

$$u_{B1}(V) = \frac{V_{max\,1} - V_{nom}}{\sqrt{3}} = \frac{230.075 - 230}{\sqrt{3}} = 0.043V \qquad (3)$$

$$u_{B2}(V) = \frac{V_{max\,2} - V_{nom}}{\sqrt{3}} = \frac{230.060 - 230}{\sqrt{3}} = 0.034V \qquad (4)$$

$$u_{B3}(V) = \frac{V_{max\,3} - V_{nom}}{\sqrt{3}} = \frac{230.067 - 230}{\sqrt{3}} = 0.039V \qquad (5)$$
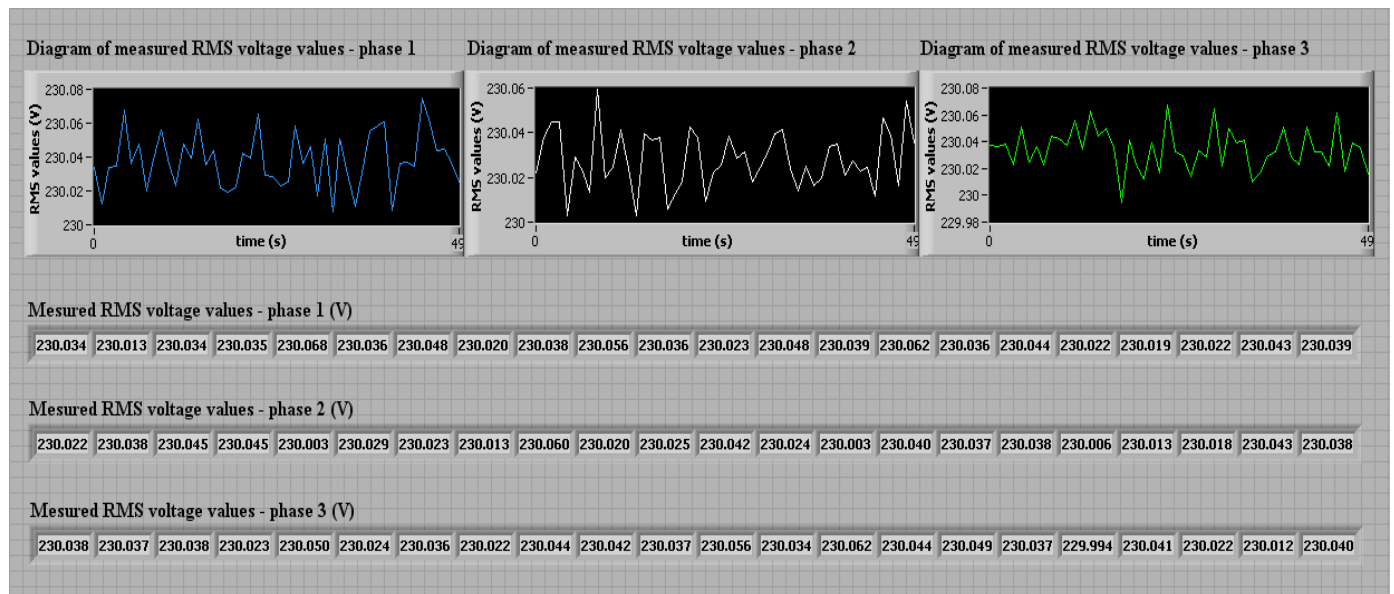

Fig. 3. Front panel of software application for presentation of time diagrams and measured RMS voltage values per phases
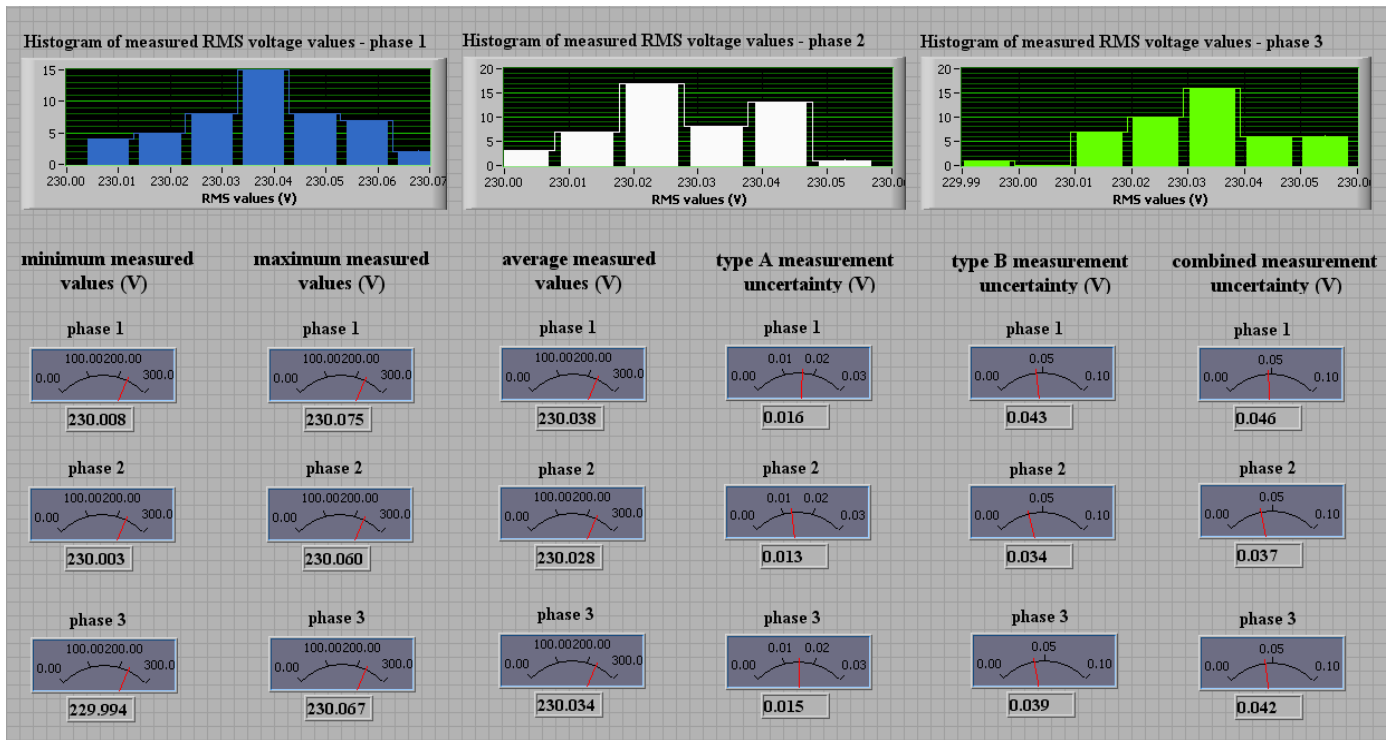
Fig. 4. Statistical processing and analysis of obtained measurement results performed in LabVIEW software environment

On the basis of previously estimated individual values of type A and type B measurement uncertainties, calculation of standard combined measurement uncertainties per phases is performed using the following square root equations [6]:

$$u_{C1}(V) = \sqrt{u_{A1}^2 + u_{B1}^2} = 0.046\,V \qquad (6)$$

$$u_{C2}(V) = \sqrt{u_{A2}^2 + u_{B2}^2} = 0.037\,V \qquad (7)$$

$$u_{C3}(V) = \sqrt{u_{A3}^2 + u_{B3}^2} = 0.042\,V \qquad (8)$$

Similar procedure for software analysis of measurement results and estimation of standard uncertainty values can be applied for others important PQ parameters, such as: signal frequency, signal phases, high-order harmonic components, power factor and others parameters. Some of characteristic measurement results will be presented in following papers.

IV. CONCLUSION

Possibility of using the virtual instrumentation software in procedure for statistical processing of PQ measurement results and estimation of standard measurement uncertainty values is described in this paper. Measurement of basic parameters for three-phase voltage waveforms are performed by standard PQ analyzer Fluke 435 Series II. Reference three-phase voltage waveforms, with standard parameters, are generated using the generator of PQ waveforms based on virtual instrumentation concept. Experimental measurement system is controlled by developed software application, which performs presentation of measurement results, presentation of recorded statistical histograms, indication of minimum and maximum measured voltage values, calculation of average measured values and standard measurement uncertainties. Calculations of type A measurement uncertainty, type B measurement uncertainty and combined uncertainty values are performed in accordance with recommendations of relevant document Guide to the Expression of Uncertainty in Measurement, prescribed by International Organization for Standardization – ISO.

ACKNOWLEDGEMENT

REFERENCES

[1] M. Simić, D. Denić, D. Živanović, D. Taskovski, V. Dimcev, "Development of a Data Acquisition System for the Testing and Verification of Electrical Power Quality Meters", JPE – Journal of Power Electronics, vol. 12, no. 5, pp. 813-820, 2012.

[2] E. F. Fuchs, M.A.S. Masoum, *Power Quality in Power Systems and Electrical Machines*, Academic Press, USA, February 2008.

[3] *Power Quality Application Guide, Voltage Disturbances, Stand. EN 50160*, Copper Development Association, 2004.

[4] L. F. Auler, R. D'Amore, "Power Quality Monitoring Controlled through Low-Cost Modules", IEEE Trans. on Instrumentation and Measurement, vol. 58, no. 3, pp. 557–562, 2009.

[5] *5520A - PQ Power Quality Option for the 5520A Calibrator – Technical Data and Specifications,* Fluke Corporation, Test Equipment Depot, Melrose, USA, 2007.

[6] *ISO Guide to the Expression of Uncertainty in Measurement*, International Standard Organization, Geneva, Switzerland,1993.

Audio system calibration using LabVIEW

Liljana Docheva [1] Ivo Dochev [2] and Marin Chushkov[3]

Abstract – **In this paper audio system calibration with LabVIEW is given. The aim is to acquire, analyze and present audio data. It's flexibility gives additional possibilities for custom audio measurements applications. Using LabVIEW software and hardware module enables the advantage of the personal computer processing power can be taken to perform complex signal analysis.**

Keywords – **LabVIEW, audio system calibration, ribbon microphone.**

I. ENABLES INTRODUCTION

Sound is very important for the satisfaction of the consumer using video system. Good surround sound audio contribute for increasing of video system quality. However this is one of the most difficult tasks. One require high-quality signal acquisition, acoustic anechoic chamber, extensive analysis, and graphical presentations.

The Number of devices which contain a microphone expands, therefore the test methods for audio devices are becoming more prevalent. The complexity of the devices continues to increase. This is the reason for the need of more complicated microphone calibration systems.

The flexibility of LabVIEW Virtual instrumentation gives additional possibilities for custom audio measurements applications. One is used to acquire, analyze and present audio data. This paper explains how audio system calibration in LabVIEW environment is made.

II. AUDIO MEASUREMENT SYSTEMS ARCHITECTURE

Usually audio measurement systems begin with microphone for transducer that generates electrical signals according to the audio signals. To measure these electrical signals and input them to a computer for processing, data acquisition system is necessary. Dynamic signals such as audio signals require high-resolution and high-dynamic-range digitizing devices. NI LabVIEW software provides a dynamic signal acquisition hardware making it suitable for building audio calibration system. Thus, using software and hardware module, the advantage of the processing LabVIEW power of personal computer can be taken to perform complex signal
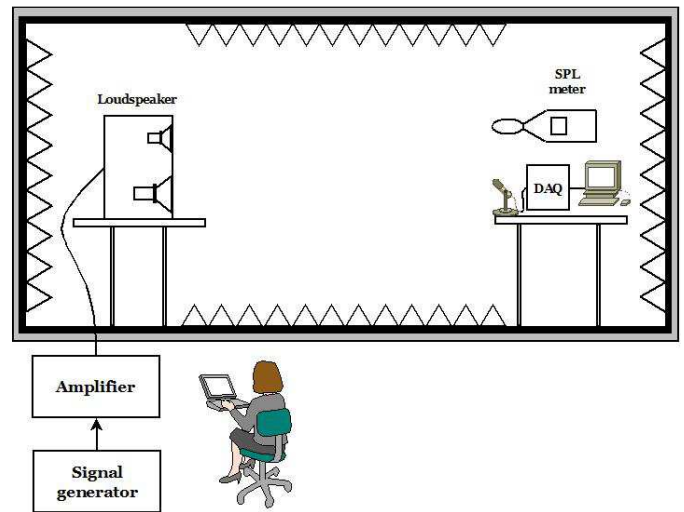
[1] Liljana Docheva is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: docheva@tu-sofia.bg.

[2] Ivo Dochev is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, idochev@tu-sofia.bg.

[3] Marin Chushkov is with the Bulgarian Institute of Metrology, Petar Mutafchiev 2 str., Sofia 1000, Bulgaria.

Fig. 1. The equipment used for audio system calibration using LabVIEW

Fig. 2. Test setup

analysis.

The equipment used for audio system calibration with LabVIEW is shown in figure 1. One includes signal generator, amplifier, loudspeaker, SPL meter, microphone, DAQ and computer. To obtain accurate results for the measurements acoustic anechoic chamber is needed. The test setup is shown in figure 2 and figure 3.
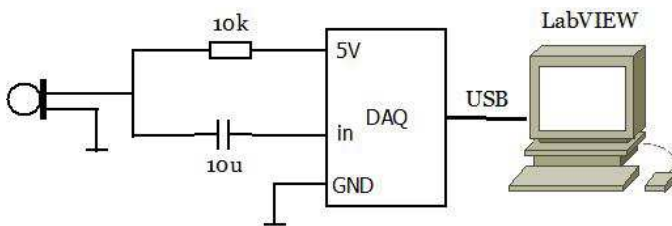
Fig. 3. Test setup

Fig. 4. Circuit design

Figure 4 presents a detailed circuit design for microphone calibration using LabVIEW data acquisition system and personal computer.

There have been made a series of measurements. The results are summarized. Based on these results correction factors for the microphone characteristic improving are proposed.

III. AUDIO SYSTEM CALIBRATION

In order to be tested and calibrated a ribbon microphone the above described system is used. The ribbon microphone amplitude frequency characteristics is measured with the equipment, shown in figure 1. The results of the measurements that were made before the amplitude frequency characteristic equalization are given in Table 1. Based on these results relative correction factors for the microphone characteristic improving are proposed. The relative correction factors are denoted with CF1 to CF5 (figure 10).

After applying relative correction factors the measurements were made again. The ribbon microphone amplitude frequency characteristic is measured with the equipment

shown in figure 1, but in LabVIEW virtual instrument correction factors are added. These results are given in Table 2.

TABLE I
MEASUREMENT RESULTS BEFORE CORRECTION

Frequency	Reference input level R_l	Power in Band P_B	Relative Correction factor C_F
20	68	-64	0,909090909
50	94	-38	0,909090909
100	94	-39	0,902255639
200	94	-38	0,909090909
300	94	-39	0,902255639
400	94	-40	0,895522388
500	92	-40	0,909090909
600	95	-37	0,909090909
800	92	-40	0,909090909
1000	95	-40	0,888888889
2000	93	-38	0,916030534
4000	90	-41	0,916030534
8000	92	-43	0,888888889
16000	87	-33	1
19500	88,4	-37,5	0,953137411
20000	87	-37	0,794701987

TABLE II
MEASUREMENT RESULTS AFTER CORRECTION

Frequency	Reference input level R_l	Power in Band P_B	Measured level
20	94	-40	94
50	94	-39	95
100	94	-40	94
200	94	-40	94
300	94	-41	94
400	94	-40	94
500	94	-40	94
600	94	-40	94
800	94	-40	94
1000	94	-41	94
2000	94	-40	94
4000	94	-40	94
8000	94	-40	94
16000	94	-42	95
19500	88,4	-37,5	95
20000	87	-37	94

Correction factors and relative correction factors are calculated by equations (1) and (2) respectively.

$$K = R_l - P_B \qquad (1)$$

$$C_F = \frac{K}{K_{16000\,Hz}} \qquad (2)$$

Where K is the correction factor , R_l is the reference input level, P_B is the power in band, K_R is the relative correction factor and $K_{16000Hz}$ is the correction factor for 16000 Hz frequency.

The LabVIEW block diagram is shown in figure 5. The LabVIEW version is 12.0f3. The data acquisition system input signals to the set of filters. For each frequency band a correction with the corresponding correction factor is made with aim to equalize the amplitude frequency characteristic. The next step is to performe octave analysis and spectral analysis to the obtained signal. Then virtual instrument *Power in Band* calculates the aggregate power in a specific frequency range which is determined by the user in VI settings. The LabVIEW front panel diagram is shown in figure 6. It includes the waveform graph, the signal spectral analysis and
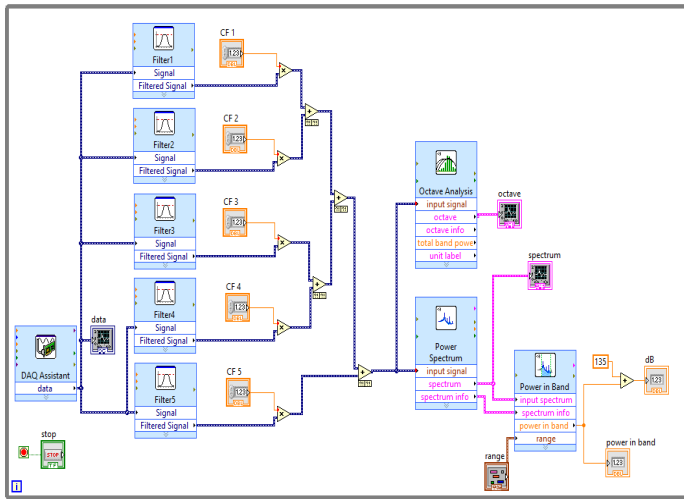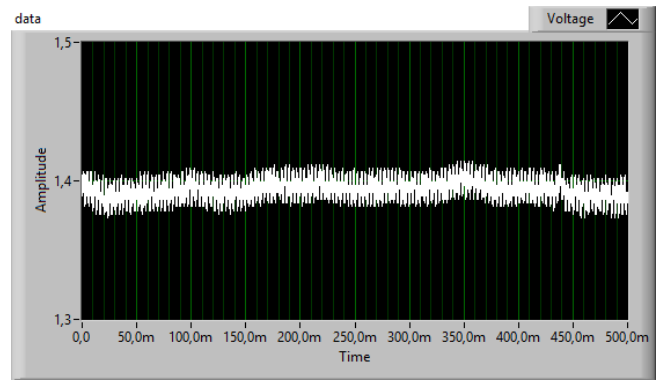

Fig. 5. LabVIEW block diagram
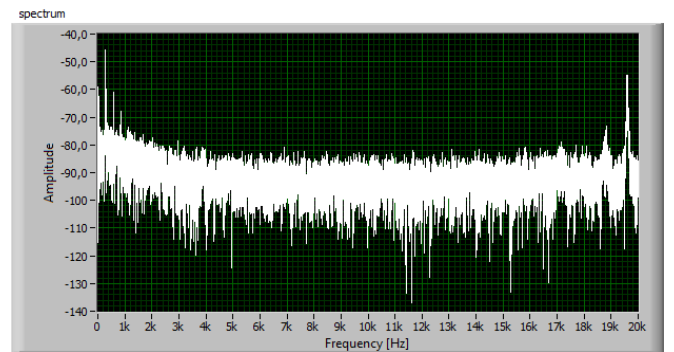

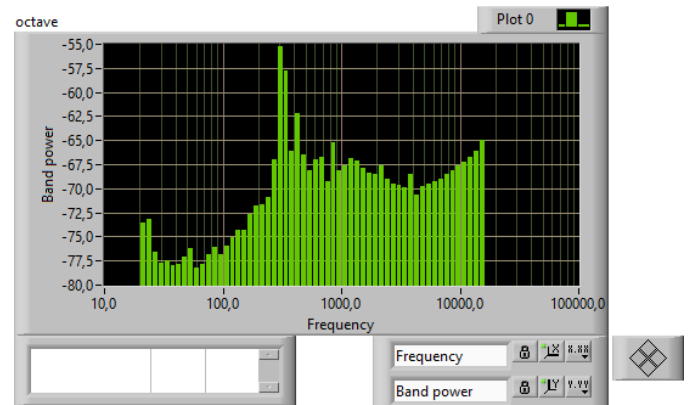Fig. 6. LabVIEW front panel

Fig. 7. Waveform graph

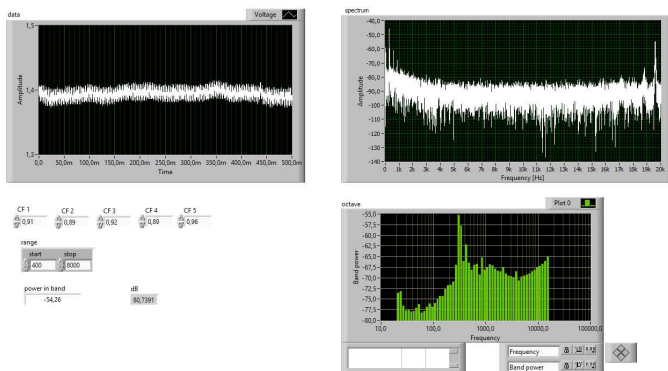Fig. 8. Signal spectral analysis

Fig. 9. Signal octave analysis

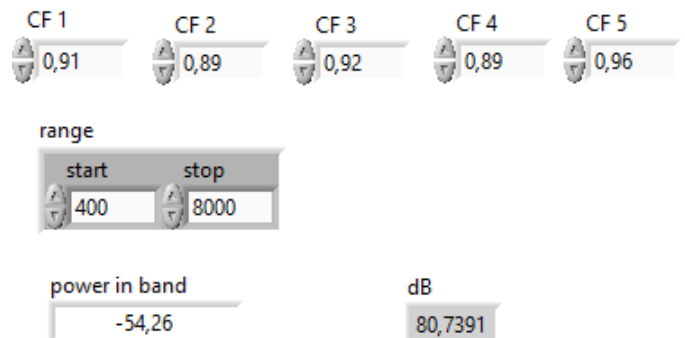octave analysis. For more detail presentation they are shown

Fig. 10. The relative correction factors

separately in figures 7, 8 and 9.

The virtual instrument *Octave Analysis* exhibits characteristic analogous to the response of the human ear. This is the reason for widely use of this technique for analyzing audio and acoustical signals.

The error before and after microphone amplitude frequency characteristic correction is given in table 3. This error is depicted in Figure 11. It can be seen that after microphone amplitude frequency characteristic correction the error is reduced.

TABLE III
ERROR CALCULATION BEFORE AND AFTER MICROPHONE AMPLITUDE FREQUENCY CHARACTERISTIC CORRECTION

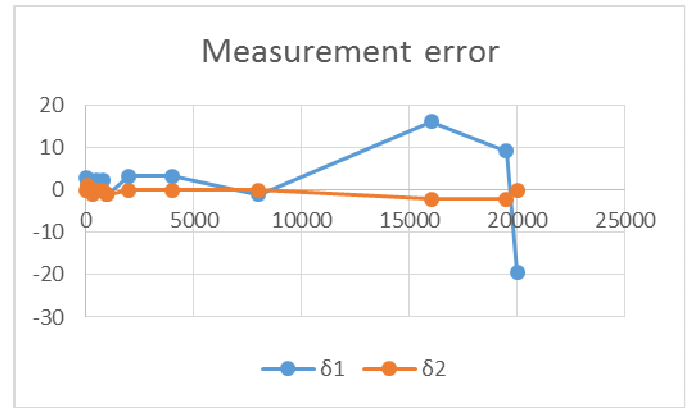| Frequency	δ1	δ2
20	2,941176	0
50	2,12766	1,06383
100	1,06383	0
200	2,12766	0
300	1,06383	-1,06383
400	0	0
500	2,173913	0
600	2,105263	0
800	2,173913	0
1000	-1,05263	-1,06383
2000	3,225806	0
4000	3,333333	0
8000	-1,08696	0
16000	16,09195	-2,12766
19500	9,162896	-2,12766
20000	-19,5402	0

Fig. 11 Illustration of Error before (blue) and after (red) microphone amplitude frequency characteristics correction

IV. CONCLUSION

In this paper a microphone calibration in LabVIEW environment is made. The ribbon microphone amplitude frequency characteristic is measured. There have been made a series of measurements. The results are summarized. Based on these results, correction factors for the microphone characteristic improvement are proposed. It is realized multipoint calibration microphone amplitude frequency characteristic by using LabVIEW. On this way error decreasing is achieved. After applying correction factors the measurements were made again. The included in the paper tables and figures illustrate the microphone amplitude frequency characteristic improvement.

REFERENCES

[1] Audio Measurements Made Easy with LabVIEW: http://www.ni.com/white-paper/3030/en/
[2] Making a Sound Pressure Measurement With NI LabVIEW: http://www.ni.com/tutorial/13601/en/

Audio signal processing system design using LabVIEW

Liljana Docheva [1]

Abstract – **In this article the audio signal fundamental tone measuring with LabVIEW is given. This is necessary for audio signal analysing, in particular for speech signals processing. The LabVIEW environment gives a possibility for creating custom audio measurements applications which interact with real-world signals. Using LabVIEW software and hardware module and personal computer can be performed complex signal analysis.**

Keywords – **LabVIEW, audio system, fundamental tone measurement.**

I. INTRODUCTION

The audio signal fundamental tone measuring is important process necessary for audio signal analysing. In some cases results may be improved through the use of preprocessing of the spectrum prior to fundamental tone measuring. The most commonly used methods are: Pre-emphasis and Masking. The first method consists in equalizing the spectrum in aim to flatten it. In the second method all peaks are rejected below an inaudibility threshold which is the maximum of the threshold of hearing. The reason is that, small peaks close to much larger peaks are often masked by the auditory system. Since it is simple to extract peaks in descending magnitude order, each removed peak can be replaced by its masking pattern, which elevates the assumed inaudibility threshold [1]. Measurement of the fundamental frequency is applied in a wide range of fields [4].

In this paper the audio signal fundamental tone measuring with LabVIEW is given. Two approaches are described: the first is real time fundamental tone measuring and the second is for vowels, recorded in „wave" format. For this aim individual vowels from man or woman spoken are used. It isn't need of preprocessing. The specifics of LabVIEW application for fundamental tone measuring are explained in the next section.

There are several ways to measure a basic tone. Commonly autocorrelation function is used. It is defined for a continuous signal by Eq. 1:

$$R(\tau) = \frac{1}{(t_{max} - t_{min})} \int_{t_{min}}^{t_{max}} s(t)s(t+\tau)dt \tag{1}$$

Where, s(t) is the sound waveform, τ is the delay time and [t_{min},t_{max}] is the integration interval [3].

LabVIEW software provides finding fundamental tone by the highest amplitude on the signal measuring [2]. The virtual instrument that can be used is Tone Measurement Express.

[1] Liljana Docheva is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: docheva@tu-sofia.bg.

Other virtual instruments, which are necessary and LabVIEW block diagram are described below.

II. LABVIEW APPLICATION FOR FUNDAMENTAL TONE MEASURING

Using LabVIEW software and hardware module enables solving the problem, described above, in two ways. One of them is to record the vowels in „wave" format, then they can be processed with aim to measure fundamental tone. LabVIEW block diagram is shown in figure 1. The LabVIEW version is 12.0f3.

Individual vowels from man or woman spoken are recorded as wave files. The required file is selected which ***svx_Get Wav List.vi*** and ***File Dialog*** virtual instruments. After reading the waveform with ***svx_Snd Read Waveform.vi*** virtual instrument the number of data elements in the waveform is calculated by ***Number of Waveform Samples*** virtual instrument. The selected file can be heard through ***Play Waveform*** virtual instrument. ***Tone Measurements Express*** virtual instrument performs audio quality measurements, which allow measuring of audio signal fundamental tone. The ***Tone Measurement Express*** virtual instrument finds the tone with the highest amplitude in the signal and calculates its amplitude and frequency. It can also to export a spectrum and additional tone analysis. For better performance, this virtual instrument can
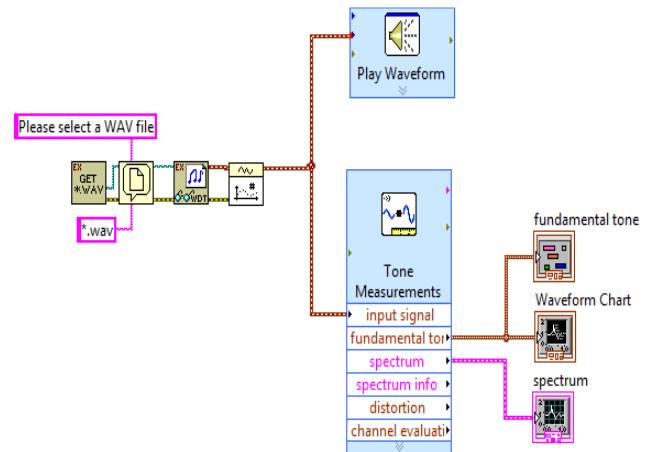

Fig. 1. LabVIEW block diagram. Fundamental tone measuring for recorded in „wave" format vowels

also narrow the search to a specified frequency band [2].

The fundamental tone is found for three „wave" format files. Figure 2 shows the LabVIEW front panel for man voice of sound „A". There can be seen its spectrum, the fundamental tone frequency and its amplitude.

Figure 3 shows the LabVIEW front panel diagram for woman voice of sound „A" - its spectrum, the fundamental tone frequency and its amplitude.

Figure 4 shows the LabVIEW front panel diagram for man voice of sound „I". The fundamental tone isn't found in real time. For real time fundamental tone measurement a different approach is needed.
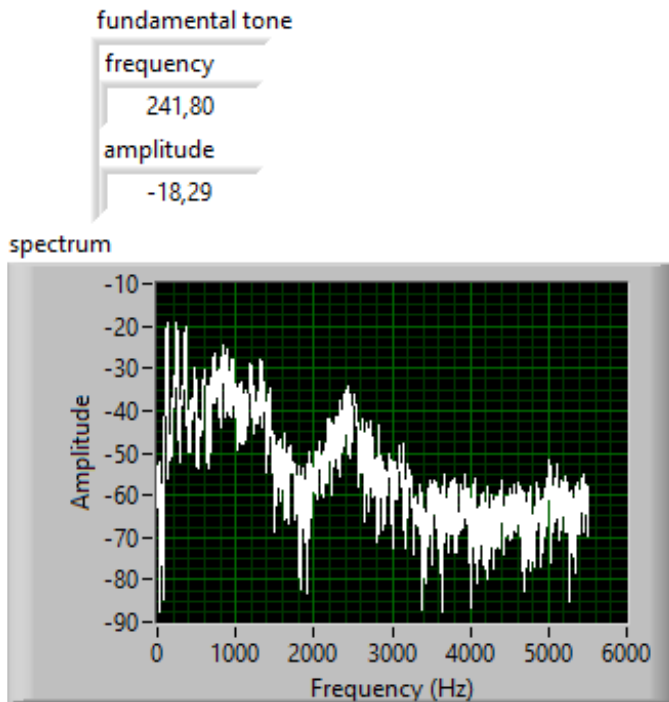
fundamental tone

frequency

241,80

amplitude

-18,29

spectrum

Fig. 2. LabVIEW front panel diagram for sound „A" – male voice

fundamental tone

frequency

346,84

amplitude

-18,94

spectrum
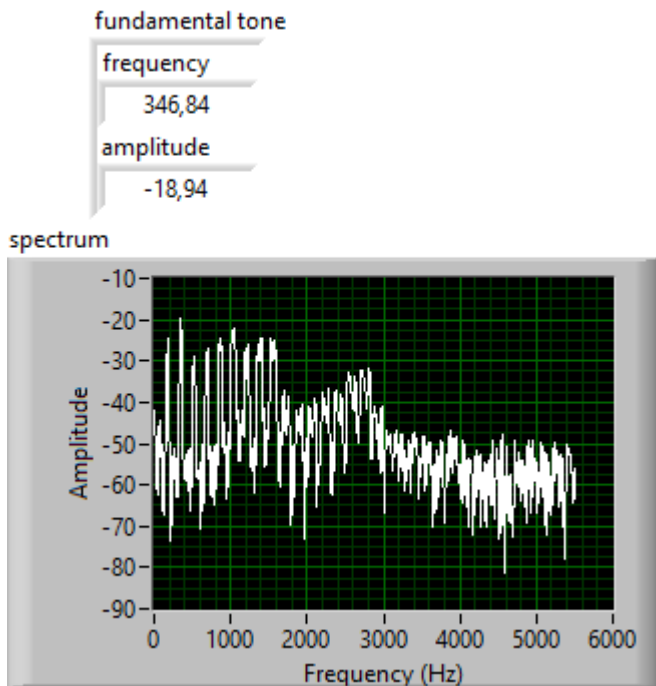
Fig. 3. LabVIEW front panel diagram for sound „A" – female voice

fundamental tone

frequency

301,33

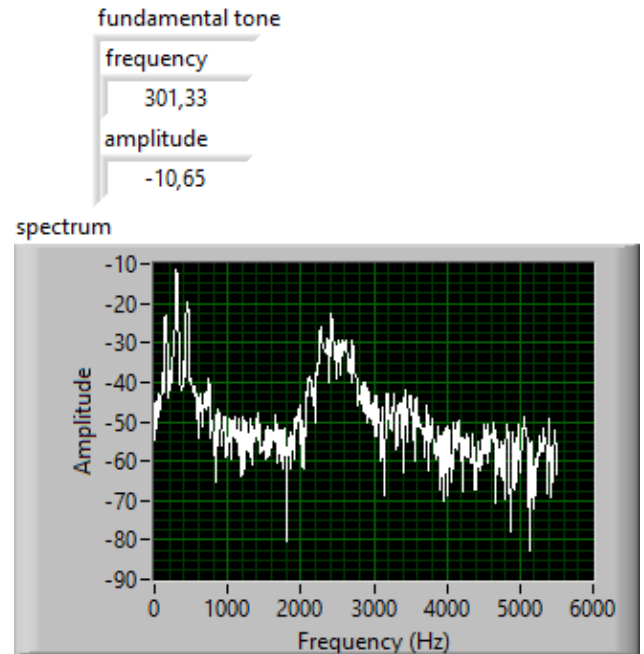amplitude

-10,65

spectrum

Fig. 4. LabVIEW front panel diagram for sound „I" – male voice

The block diagram for fundamental tone measuring in real time with LabVIEW is shown in figure 5. One includes microphone, DAQ and computer. The input signal and spectrum can be presented in graphics. Current frequency and amplitude values can be indicated too.

LabVIEW block diagram for fundamental tone measuring in real time is shown in figure 6. In this case data acquisition system is necessary. The data acquisition system and *DAQ Assistant* virtual instrument inputs signals from the microphone to the *Tone Measurements Express* virtual instrument. This allows audio signal fundamental tone measuring in real time.

The LabVIEW front panel diagram for real time

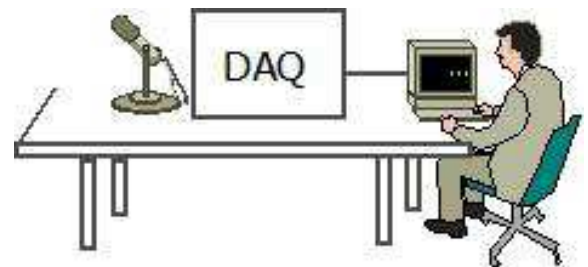Fig. 5. Block diagram for fundamental tone measuring in real time with LabVIEW.

fundamental tone measuring is shown in figure 7 and figure 8. The vowel "A" is produced by a woman (figure 7), a child (figure 8) and a man (figure 9). The fundamental tone frequency, its amplitude, data for channel evaluation and distortion are given. Furthermore the waveform graph and signal spectral analysis are given.
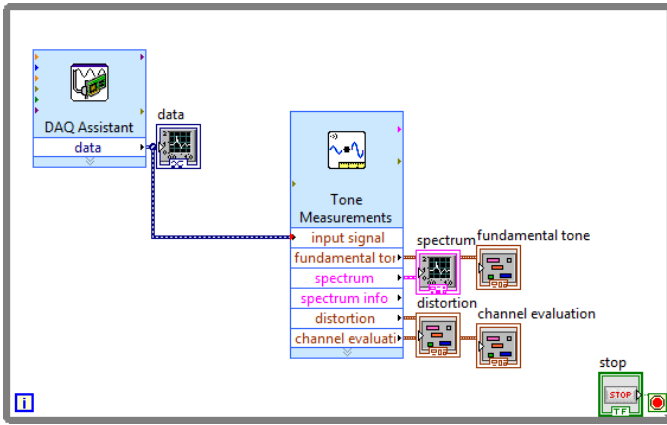
Fig. 6. LabVIEW block diagram.
Fundamental tone measuring for vowels in real time

Fig. 8. LabVIEW front panel diagram for fundamental tone
measuring in real time: sound „A" – children's voice

Figure 10, figure 11 and figure 12 show the LabVIEW front panel diagram for real time fundamental tone measuring for the same woman, man and child. The vowel "I" is produced. The fundamental tone frequency and its amplitude can be seen.

Fig. 9. LabVIEW front panel diagram for fundamental tone
measuring in real time: sound „A" – male voice

Fig. 7. LabVIEW front panel diagram for fundamental tone
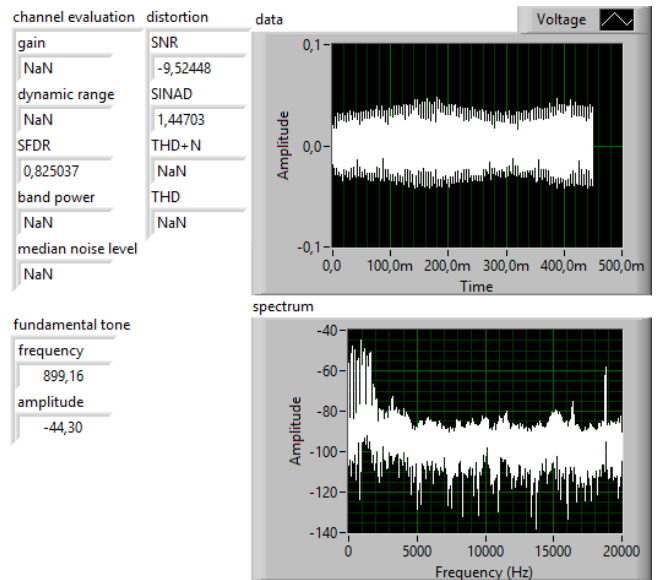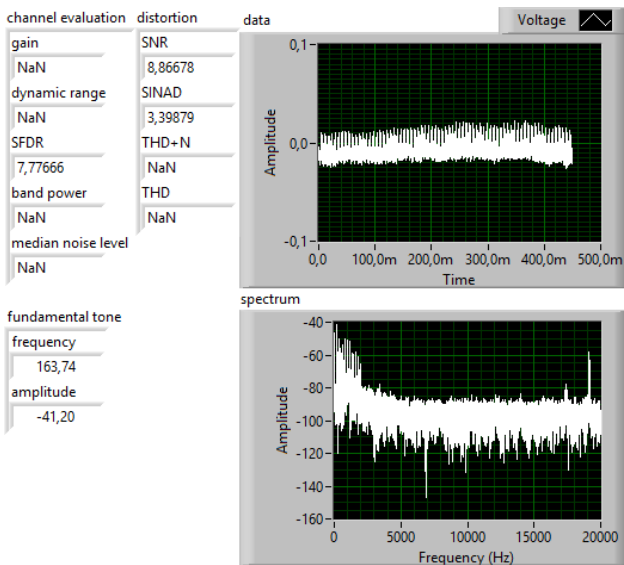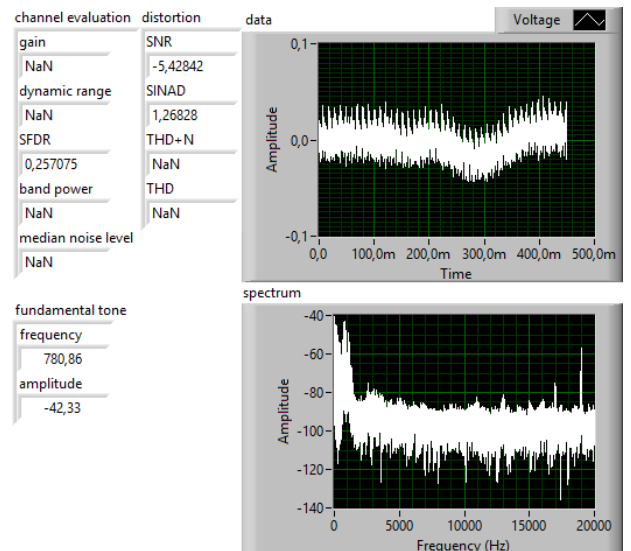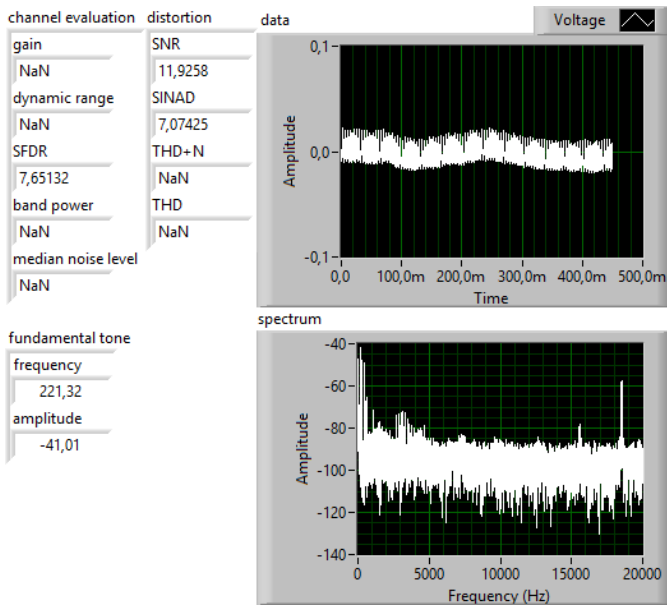measuring in real time: sound „A" – female voice

Fig. 10. LabVIEW front panel diagram for fundamental tone measuring in real time: sound „I" – female voice
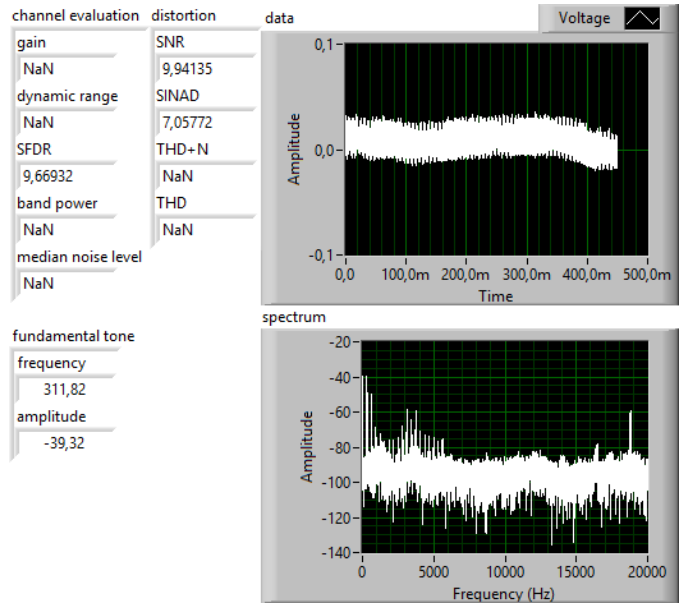

Fig. 1. LabVIEW front panel diagram for fundamental tone measuring in real time: sound „I" – children's voice

Two approaches are described: tone measuring for vowels recorded in „wave" format (figures 2, 3, 4) and real time fundamental tone measuring (from figure 7 to figure 12).

In both methods vowels spoken by different people have been applied to the system. Fundamental tone frequency and its amplitude are measured. The spectrum can be seen too. For the second approach - real time fundamental tone measuring - data for channel evaluation, data for distortion, waveform graph have been shown. The obtained important information will be used for further sound processing in audio systems and in education [5] which will be described in subsequent papers.

Fig. 11. LabVIEW front panel diagram for fundamental tone measuring in real time: sound „I" – male voice

III. CONCLUSION

In this paper an audio signal fundamental tone measuring with LabVIEW is given. The peculiarity of LabVIEW application for fundamental tone measuring are explained.

REFERENCES

[1] Fundamental Frequency Estimation from Sinusoidal Peaks - http://www.dsprelated.com/freebooks/sasp/Fundamental_Freque ncy_Estimation_Sinusoidal.html
[2] Audio Measurements Made Easy with LabVIEW: http://www.ni.com/white-paper/3030/en/
[3] Autocorrelation (for sound signals) - http://www.phy.mtu.edu/~suits/autocorrelation.html.
[4] Miletiev R., Iontchev E., Mihaylov E. , Yordanov R., Measurement of the Fundamental Frequency of the Tram Cart by Quad Sensor Inertial System, 24th National Scientific Symposium with International Participation "Metrology and Metrology Assurance", September 7-11, 2014, Proceedings of the Symposium, pp.228-233, Sozopol, Bulgaria.
[5] Yordanov R., New Trends and Methods in the Education of Microelectronics Students, XI-th International Conference „Challenges in Higher Education and Research in XXI Century", 4 – 7 June 2013, Sozopol, Bulgaria, Proceedings of the 11th Conference, vol.11, pp.33-35, Heron Press, Sofia, 2013.

Comparative Performance Studies of Laboratory Open IEEE 802.11b,g PTMP Links

José A. R. Pacheco de Carvalho[1],

Cláudia F. F. P. Ribeiro Pacheco[2], Hugo Veiga[3], António D. Reis[4]

Abstract – **The increasing importance of wireless communications, involving electronic devices, has been widely recognized. Performance is a crucial issue, leading to more reliable and efficient communications. Laboratory measurements were performed about several performance aspects of Wi-Fi (IEEE 802.11 b, g) Open point-to-multipoint links. Our study contributes to the performance evaluation of this technology, using new available equipments (V-M200 access points from HP and WPC600N adapters from Linksys). New detailed results are presented and discussed, namely at OSI levels 4 and 7, from TCP, UDP and FTP experiments: TCP throughput, jitter, percentage datagram loss and FTP transfer rate. Comparisons are made to corresponding results obtained for Open point-to-point links. Conclusions are drawn about the comparative performance of the links.**

Keywords – **Wi-Fi, WLAN, Open Point-to-Multipoint and Point-to-Point Links, IEEE 802.11bg, Wireless Network Laboratory Performance.**

I. INTRODUCTION

Contactless communication techniques have been developed using mainly electromagnetic waves in several frequency ranges, propagating in the air. Wi-Fi and FSO, whose importance and utilization have been recognized and growing, are representative examples of wireless communications technologies.

Wi-Fi is a microwave based technology providing for versatility, mobility and favorable prices. The importance and utilization of Wi-Fi has been increasing for complementing traditional wired networks. It has been used mainly in infrastructure mode. In this case an access point, AP, permits

[1]José Pacheco de Carvalho is with the Remote Detection Unit and the Physics Department at the University of Beira Interior, R. Marquês d'Ávila e Bolama, 6201-001 Covilhã, Portugal, E-mail: pacheco@ubi.pt.

[2]Cláudia Pacheco is with the Remote Detection Unit at the University of Beira Interior, R. Marquês d'Ávila e Bolama, 6201-001 Covilhã, Portugal, E-mail: a17597@ubi.pt.

[3]Hugo Veiga is with the Remote Detection Unit and the Informatics Centre at the University of Beira Interior, R. Marquês d'Ávila e Bolama, 6201-001 Covilhã, Portugal, E-mail: hveiga@ubi.pt.

[4]António Reis is with the Remote Detection Unit and the Physics Department at the University of Beira Interior, and with the Department of Electronics and Telecommunications/Institute of Telecommunications, at the University of Aveiro, 3810 Aveiro, Portugal, E-mail: adreis@ubi.pt.

communications of Wi-Fi devices (such as a personal computer, a wireless sensor, a PDA, a smartphone, a video game console, a digital audio player) with a wired based LAN through a switch/router. In this way a WLAN, based on the AP, is formed. Wi-Fi has reached the personal home, where a WPAN allows personal devices to communicate. Point-to-point (PTP) and point-to-multipoint (PTMP) setups are used both indoors and outdoors, requiring specific directional and omnidirectional antennas. Point-to-point and point-to-multipoint links use microwaves in the 2.4 and 5 GHz frequency bands and IEEE 802.11a, b, g, n standards [1]. The 2.4 GHz band has been intensively used and is having increasing interferences. Therefore considerable attention has been focused on the 5 GHz band where, however, absorption increases and ranges are shorter.

Nominal transfer rates up to 11 (802.11b), 54 Mbps (802.11 a, g) and 600 Mbps (802.11n) are specified. The medium access control is CSMA/CA. There are studies on wireless communications, wave propagation [2,3], practical implementations of WLANs [4], performance analysis of the effective transfer rate for 802.11b point-to-point links [5], 802.11b performance in crowded indoor environments [6].

Performance has been a very important issue, resulting in more reliable and efficient communications. In comparison to traditional applications, new telematic applications are specially sensitive to performances. Requirements have been pointed out [7], e.g.: for video on demand/moving images, 1-10 ms jitter and 1-10 Mbps throughputs; for Hi Fi stereo audio, jitter less than 1 ms and 0.1-1 Mbps throughputs.

Wi-Fi security is very important as microwave radio signals travel through the air and can be easily captured. WEP was initially intended to provide confidentiality comparable to that of a traditional wired network. In spite of presenting weaknesses, WEP is still widely used in Wi-Fi networks for security reasons, mainly in point-to-point links. More advanced and reliable security methods have been developed to provide authentication such as, by increasing order of security, WPA and WPA2.

Several performance measurements have been made for 2.4 and 5 GHz Wi-Fi open [8,9], and WEP links [10,11], as well as very high speed FSO [12]. In the present work new Wi-Fi (IEEE 802.11 b, g) results arise, using Open links, namely through OSI levels 4 and 7. Performance is evaluated through laboratory measurements of Open PTMP links, using available equipments. Comparisons are made to corresponding results obtained for Open PTP links.

The rest of the paper is structured as follows: Section II presents the experimental details i.e. the measurement setup

and procedure. Results and discussion are presented in Section III. Conclusions are drawn in Section IV.

II. EXPERIMENTAL DETAILS

The measurements used a HP V-M200 access point [13], with three external dual-band 3x3 MIMO antennas, IEEE 802.11 a/b/g/n, software version 5.4.1.0-01-9867 and a 100-Base-TX/10-Base-T Allied Telesis AT-8000S/16 level 2 switch [14]. Two PCs were used having a PCMCIA IEEE.802.11 a/b/g/n Linksys WPC600N wireless adapter with three internal antennas [15], to enable PTMP (three-node) links to the access point. In every type of experiment, interference free communication channels were used (ch 8 for 802.11 b, g). This was checked through a portable computer, equipped with a Wi-Fi 802.11 a/b/g/n adapter, running NetStumbler software [16]. No encryption was activated in the AP and the wireless adapters of the PCs. The experiments were made under far-field conditions. No power levels above 30 mW (15 dBm) were required, as the wireless equipments were close.

A laboratory setup has been planned and implemented for the measurements, as shown in Fig. 1. At OSI level 4, measurements were made for TCP connections and UDP communications using Iperf software [17]. For a TCP connection, TCP throughput was obtained. For a UDP communication with a given bandwidth parameter, UDP throughput, jitter and percentage loss of datagrams were determined. Parameterizations of TCP packets, UDP datagrams and window size were as in [10]. One PC, with IP 192.168.0.2 was the Iperf server and the other, with IP 192.168.0.6, was the Iperf client. Jitter, representing the smooth mean of differences between consecutive transit times, was continuously computed by the server, as specified by the real time protocol RTP, in RFC 1889 [18]. The scheme of Fig. 1 was also used for FTP measurements, where FTP server and client applications were installed in the PCs with IPs 192.168.0.2 and 192.168.0.6, respectively. Another PC, with IP 192.168.0.20, was used to control the settings in the AP.

The server and client PCs were HP nx9030 and nx9010 portable computers, respectively, running Windows XP Professional. They were configured to optimize the resources allocated to the present work. Batch command files have been re-written to enable the TCP, UDP and FTP tests. The results were obtained in batch mode and written as data files to the client PC disk. Each PC had a second network adapter, to permit remote control from the official IP University network, via switch.

III. RESULTS AND DISCUSSION

The PC wireless network adapter were manually configured, for IEEE 802.11 bg, with typical nominal transfer rates (1, 2, 5, 11 Mbps for 802.11b; 6, 9, 12, 18, 24, 36, 48, 54 Mbps for 802.11g). For every fixed transfer rate, data were obtained for comparison of the laboratory performance

of the PTMP and PTP links at OSI layers 1 (physical layer), 4 (transport layer) and 7 (application layer) using the setup of Fig. 1. For each standard and every nominal fixed transfer rate, an average TCP throughput was determined from several experiments. This value was used as the bandwidth parameter for every corresponding UDP test, giving average jitter and average percentage datagram loss.

At OSI level 1, noise levels (N, in dBm) and signal to noise ratios (SNR, in dB) were monitored and typical values are shown in Fig. 2. The main average TCP and UDP results are summarized in Table I, both for PTMP and PTP links. The statistical analysis, including calculations of confidence intervals, was carried out as in [19].

In Figs. 3-4 polynomial fits were made to the 802.11 b, g TCP throughput data for PTMP and PTP links, respectively, where R^2 is the coefficient of determination. It was found that, on average, the best TCP throughputs are for 802.11 g and PTP links (13.9+-0.4 Mbps). In Figs. 5-6, the data points representing 802.11 b, g jitter data for for PTMP and PTP links, respectively, were joined by smoothed lines. It was found that the best average jitter values are for 802.11 g and PTP links (3.1+-0.9 ms). In Fig. 7, percentage datagram loss data for 802.11 b, g are shown for PTMP links. It was found that average percentage datagram loss is not significantly sensitive to link type.

At OSI level 7 we measured FTP transfer rates versus nominal transfer rates configured in the access point and the PC wireless network adapter for IEEE 802.11 g as in [10]. The results show the same trends found for TCP throughput.

Generally it was found that, specially for TCP throughput, the data show performance degradations for PTMP links, where the AP experiences higher processing requirements so as to maintain links between PCs.
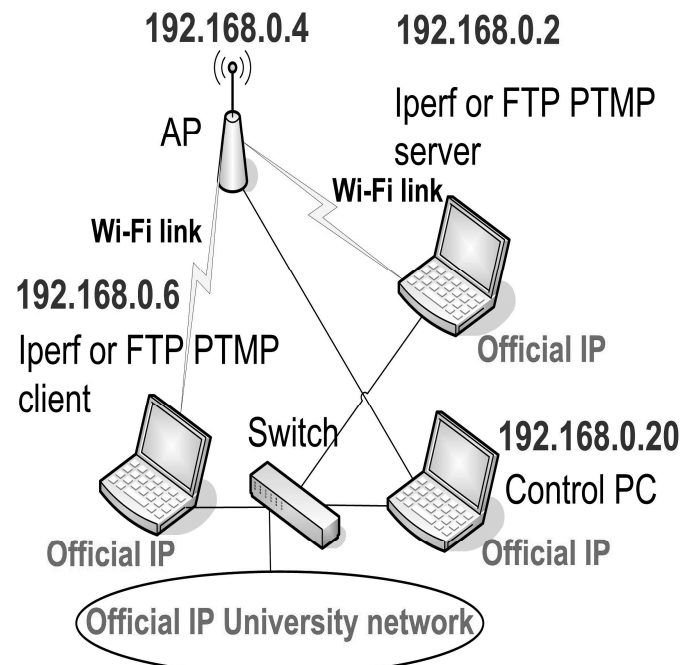

Fig. 1- Wi-Fi PTMP laboratory setup scheme.
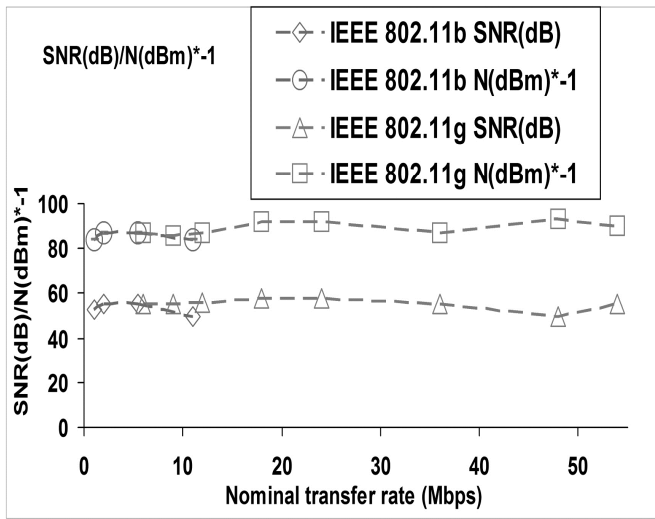
Fig. 2- Typical SNR (dB) and N (dBm); PTMP
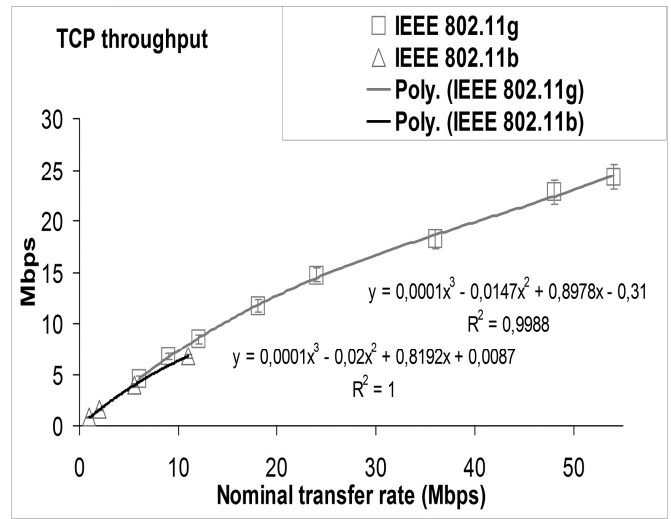
Fig. 4- TCP throughput versus technology and nominal transfer rate; PTP.

TABLE I
AVERAGE WI-FI (IEEE 802.11B,G) RESULTS;
OPEN PTMP AND PTP LINKS.

Link type	PTMP		PTP	
Parameter/ IEEE standard	802.11b	802.11g	802.11b	802.11g
TCP throughput (Mbps)	1.5 +-0.0	6.5 +-0.2	3.3 +-0.1	13.9 +-0.4
UDP-jitter (ms)	4.2 +-0.7	3.8 +-0.7	5.6 +-0.2	3.1 +-0.9
UDP-% datagram loss	1.3 +-1.0	1.3 +-0.1	1.2 +-0.2	1.5 +-0.1
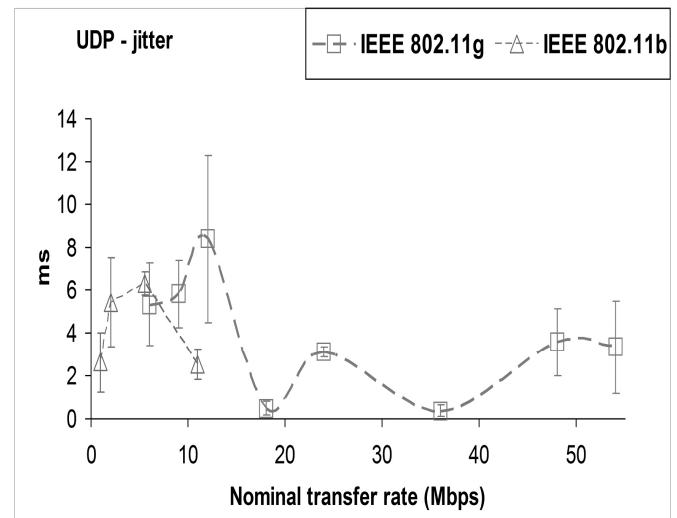
Fig. 5- UDP - jitter results versus technology and nominal transfer rate; PTMP.
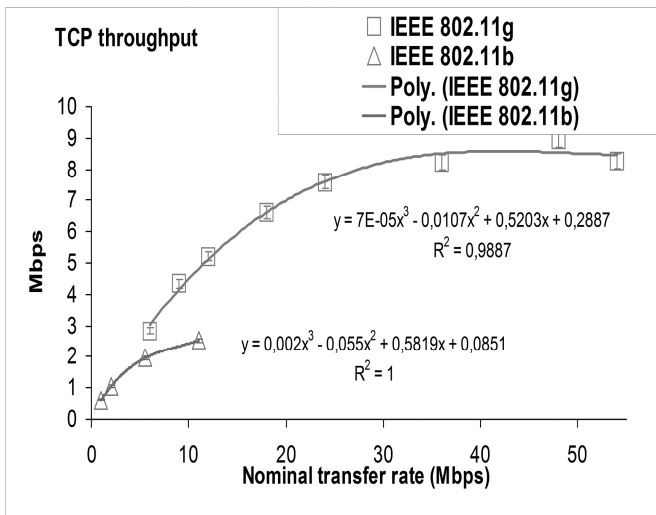
Fig. 3- TCP throughput versus technology and nominal transfer rate; PTMP.
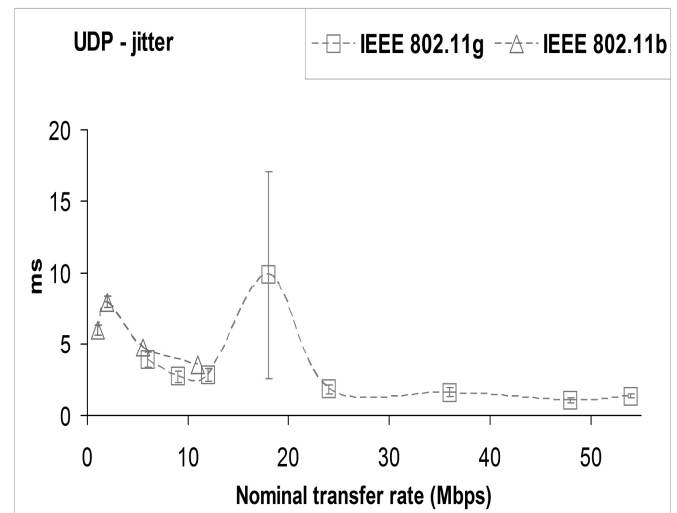
Fig. 6- UDP - jitter results versus technology and nominal transfer rate; PTP.
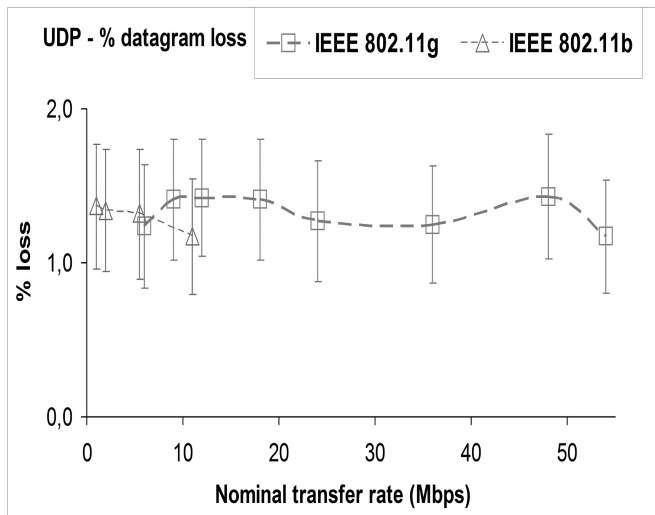
Fig. 7- UDP – percentage datagram loss versus technology and nominal transfer rate; PTMP.

IV. CONCLUSION

A versatile laboratory setup arrangement has been planned and implemented, that permitted systematic performance measurements of available wireless equipments (V-M200 access points from HP and WPC600N adapters from Linksys) for Wi-Fi (IEEE 802.11 b, g) in Open PTMP links.

Through OSI layer 4, TCP throughput, jitter and percentage datagram loss were measured and compared for each standard to corresponding results obtained for Open PTP links. It was found that, on average, the best TCP throughputs are for 802.11 g and PTP links. It was also found that the best average jitter values are for 802.11 g and PTP links. Concerning average percentage datagram loss data, it was found as not significantly sensitive to link type.

Generally it was found that, specially for TCP throughput, the data show performance degradations for PTMP links, where the AP experiences higher processing requirements so as to maintain links between PCs.

At OSI layer 7, FTP performance results have shown the same trends found for TCP throughput.

Further performance studies are planned using several equipments, topologies, security settings and noise conditions, not only in laboratory but also in outdoor environments involving, mainly, medium range links.

ACKNOWLEDGEMENT

REFERENCES

[1] Web site http://standards.ieee.org Web site; IEEE 802.11a, 802.11b, 802.11g, 802.11n, 802.11i standards; accessed 10 Jan 2014.

[2] J. W. Mark, W. Zhuang, Wireless Communications and Networking, Prentice-Hall, Inc., Upper Saddle River, NJ, 2003.

[3] T. S. Rappaport, Wireless Communications Principles and Practice, 2nd ed., Prentice-Hall, Inc., Upper Saddle River, NJ, 2002.

[4] W. R. Bruce III, R. Gilster, Wireless LANs End to End, Hungry Minds, Inc., NY, 2002.

[5] M. Schwartz, Mobile Wireless Communications, Cambridge University Press, 2005.

[6] N. Sarkar, K. Sowerby, "High Performance Measurements in the Crowded Office Environment: a Case Study", In Proc. ICCT'06-International Conference on Communication Technology, Guilin, China, 27-30 November 2006, pp. 1-4.

[7] E. Monteiro, F. Boavida, Engineering of Informatics Networks, 4th ed., FCA-Editor of Informatics Ld., Lisbon, 2002.

[8] J. A. R. Pacheco de Carvalho, P. A. J. Gomes, H. Veiga, A. D. Reis, "Development of a University Networking Project", in Encyclopedia of Networked and Virtual Organizations, Goran D. Putnik, Maria Manuela Cunha, Eds. Hershey, PA (Pennsylvania): IGI Global, pp. 409-422, 2008.

[9] J. A. R. Pacheco de Carvalho, H. Veiga, P. A. J. Gomes, C. F. Ribeiro Pacheco, N. Marques, A. D. Reis, "Wi-Fi Point-to-Point Links- Performance Aspects of IEEE 802.11 a,b,g Laboratory Links", in Electronic Engineering and Computing Technology, Series: Lecture Notes in Electrical Engineering, Sio-Iong Ao, Len Gelman, Eds. Netherlands: Springer, Vol. 60, pp. 507-514, 2010.

[10] J. A. R. Pacheco de Carvalho, H. Veiga, N. Marques, C. F. Ribeiro Pacheco, A. D. Reis, Wi-Fi WEP Point-to-Point Links-Performance Studies of IEEE 802.11 a,b,g Laboratory Links, in Electronic Engineering and Computing Technology, Series: Lecture Notes in Electrical Engineering, Sio-Iong Ao, Len Gelman, Eds. Netherlands: Springer, Vol. 90, pp. 105-114, 2011.

[11] J. A. R. Pacheco de Carvalho, C. F. Ribeiro Pacheco, A. D. Reis, H. Veiga, "Laboratory Performance Measurements of IEEE 802.11 b, g WEP PTP Links", Proc. WCE 2014 - World Congress on Engineering 2014, Imperial College London, London, England, 2-4 July 2014, pp. 724-727.

[12] J. A. R. Pacheco de Carvalho, Nuno Marques, H. Veiga, Cláudia F. F. P. Ribeiro Pacheco, A. D. Reis, "Contributions to Experimental Performance Studies of Wi-Fi and FSO Links", Proc. CISTI 2010-5th Iberian Conference on Information Systems and Technologies, Santiago de Compostela, Spain, 16-19 June 2010, pp. 221-226.

[13] Web site http://www.hp.com; HP V-M200 802.11n access point management and configuration guide; 2010; accessed 15 Jan 2015.

[14] Web site http://www.alliedtelesis.com; AT-8000S/16 level 2 switch technical data; 2009; accessed 10 Dec 2012.

[15] Web site http://www.linksys.com; WPC600N notebook adapter user guide; 2008; accessed 10 Jan 2012.

[16] Web site http://www.netstumbler.com; NetStumbler software; 2005; accessed 21 Mar 2011.

[17] Web site http://dast.nlanr.net; Iperf software; 2003; accessed 10 Jan 2008.

[18] Network Working Group. "RFC 1889-RTP: A Transport Protocol for Real Time Applications", http://www.rfc-archive.org; 1996; accessed 10 Feb 2008.

[19] P. R. Bevington, Data Reduction and Error Analysis for the Physical Sciences, Mc Graw-Hill Book Company, 1969.

Application of Virtual Instrumentation for Measuring the Angular Position and Velocity

Goran Miljković[1], Dragan Denić[2], Milan Simić[3] and Aleksandar Jocić[4]

Abstract – **Accurate measurement of angular position and velocity is required in many modern movable systems in industry and everyday life. Realized virtual instrument for measuring the angular position and velocity is tested and presented in this paper. It uses signals at the output of the electronic system for optical code reading from a glass code disc of 10-bit pseudorandom encoder, which has two code (pseudorandom) and one synchronization track. Presented virtual instrument is modular and adaptable, and can be modified and upgraded some of its elements. The use of computer and modular hardware, as well as the USB interface defines the limit frequency of the instrument.**

Keywords – **position measurement, pseudorandom position encoder, velocity measurement, virtual instrument**

I. INTRODUCTION

Virtual instrument is a combination of hardware and software elements, the most commonly used with a PC, which is accomplished classical instrument function [1, 2]. Functionally, virtual instruments, as computer-based instruments, can perform the tasks of measuring, monitoring, acquisition, simulation, testing, and so on.

Virtual instruments, usually designed in LabVIEW software environment, have properties of good flexibility, easy modification, hierarchy and modularity, easy connectivity with other instruments, the Internet, various data presentations, etc. Such an instrument could be later adapted, repaired, and extended its functionality. The three main parts of the virtual instrument software are the front panel, the block diagram, and the icon with the connector. Using graphical programming in realization of virtual instrument was removed some difficulties related to programming with textual programming languages, and shorten the time of realization of an instrument. Sources of measurement uncertainty with virtual instruments [3] that need to be addressed when determining the measurement uncertainty are: transducers, signal conditioning circuits, data acquisition,

application software for measurement, computer.

Absolute pseudorandom position encoders can be used to accurately measure the angular position and additionally the angular velocity in various fields of industry (machinery, motors, conveyors), elevators, telescopes, antennas, computer peripherals (printers, mice), and so on. The window of length n, which slides along the n-bit pseudorandom binary sequence (PRBS) allocates a unique code word at any time, which is used to determine the absolute position of the pseudorandom absolute position encoder [4, 5]. Unlike conventional absolute encoders, code words are now longitudinally arranged on the code track; with two consecutive code words differ in only one bit. Some of the advantages of absolute pseudorandom position encoders compared to conventional absolute encoders are they only have one code track on the code disk, regardless of the resolution encoder; then they are more reliable due to the possibility of using the method for detecting code reading errors; code reading can be implemented using one or two sensor heads as a serial code reading, etc.

The main functional parts of a pseudorandom absolute position encoder are code reading system [5, 6], the code scanning method [5, 7, 8], the code reading error detection methods [9] and the pseudorandom/natural code conversion methods [5, 8, 10]. The pseudorandom/natural code conversion can be done as parallel, which is the fastest way, serial or a combination of these two methods.

Angular velocity is an important parameter which usually needs to be regulated or monitored in different servo, mechatronic, robotic and precise production systems. One signal from synchronization track is used for angular velocity measurement. There are different methods for velocity measurement, and here is used the direct counting method (or M method) [11]. This method is based on counted pulses from the encoder in fixed-time intervals.

In the literature we can find realized virtual instruments for measuring the angular position and velocity based on the application of the incremental encoder signals, whereby the principle of determining the angular position considerably simpler than in the case of pseudorandom position encoder [12]. This paper presents a virtual instrument for measuring the angular position and velocity, based on signals from the code disk of pseudorandom absolute position encoder.

II. FUNCTIONAL PARTS OF VIRTUAL INSTRUMENT

The signals that are used for measuring the angular position and velocity in realized virtual instrument were obtained using an electronic system for optical reading from glass code disk of realized pseudorandom position encoder, Fig. 1. The code disk of pseudorandom position encoder has synchronization and two phase shifted 10-bit pseudorandom code tracks. The

[1]Goran Miljković is with the Faculty of Electronic Engineering at University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mail: goran.miljkovic@elfak.ni.ac.rs

[2]Dragan Denić is with the Faculty of Electronic Engineering at University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mail: dragan.denic@elfak.ni.ac.rs

[3]Milan Simić is with the Faculty of Electronic Engineering at University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mail: milan.simic@elfak.ni.ac.rs

[4]Aleksandar Jocić is with the Faculty of Electronic Engineering at University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mail: aleksandar.jocic@elfak.ni.ac.rs

signals obtained with an electronic system of optical code reading are fed to USB multifunction data acquisition card NI USB-6341, which is shown in Fig. 2. The two phase shifted signals from synchronization track are used to determine the direction of rotation as well as the determination of the moment reading pseudorandom code on two pseudorandom code tracks. One signal from synchronization track is also use for measuring the angular velocity. Signal from one pseudorandom code track is used to form a code word during movement in one direction, while the other signal, phase-shifted, from another pseudorandom code track is used for the second direction of movement of the shaft encoder.
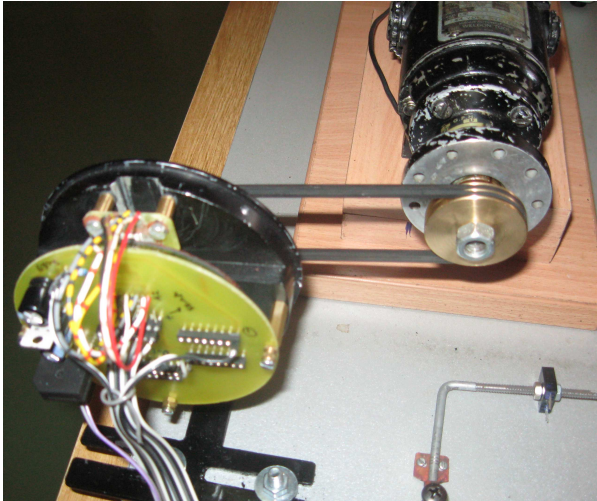

Fig. 1. The code disk with an electronic code reading system

Fig. 2. The external appearance of acquisition cards X series (made by National Instruments)

NI USB-6341 acquisition card has 16 analog inputs of 16-bit resolution with a sampling rate 500 kS/s, 2 analog outputs, 24 digital inputs / outputs, 4 counters, etc. The signals at the output of the electronic system for code reading are also observed and analyzed on a digital oscilloscope, Fig. 3. The first two digital TTL signals are from synchronization track, and the other two from two phase shifted pseudorandom code tracks.

M method or direct counting method is simple and widely used method for angular velocity measurement, and it is based on counting of pulses in fixed time interval. C_P denotes the number of counted pulses in fixed sampling period T_S. The measured velocity ω is estimated using

$$\omega = \frac{60 C_P}{P T_S}\,[rpm]\,, \qquad (1)$$

where P denotes the number of encoder pulses per rotation, and it can be changed in the front panel of the realized virtual

instrument, which depends on the applied encoder type. Measurement time is equal to the sampling period T_S, and it can be changed in the front panel. One counter is used for determination of sampling period and the other counter is used for counting the encoder pulses. This method has good accuracy at high speed.
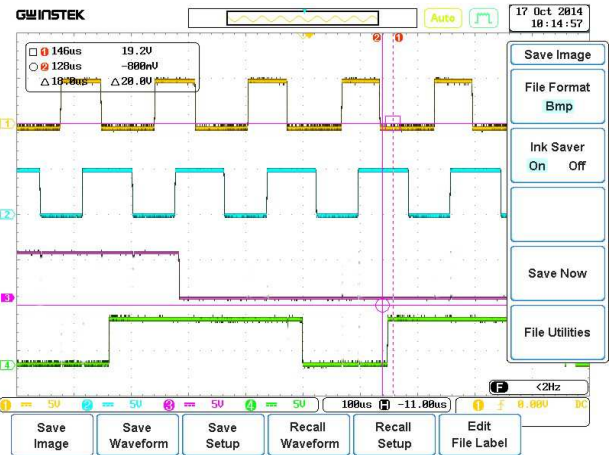
Fig. 3. Pseudorandom encoder signals on the screen of the digital oscilloscope

Using two phase shifted signals from synchronization track obtained information on the direction of rotation of the encoder using the same method as for incremental encoders. Based on the signals from pseudorandom code tracks the read 10-bit code word is formed in a bidirectional shift register, within virtual instrument block diagram. This code word is used for determination of position information, Fig. 4. Signal from one pseudorandom code track is used for one direction of rotation of the shaft encoder, and a second signal for the opposite direction of rotation of the shaft encoder. The disadvantage of using serial reading of pseudorandom binary code is that the after system initialization it takes some initial time until a first valid pseudorandom code word is formed. After that with each clock period gets new 10-bit code word that is compared to the previous code word differs only in one bit.

Fig. 4. Forming read code word in the registry

The read 10-bit pseudorandom code word formed in the bidirectional shift register in the block diagram of the virtual instrument, is converted from a pseudorandom to natural code using serial code converter base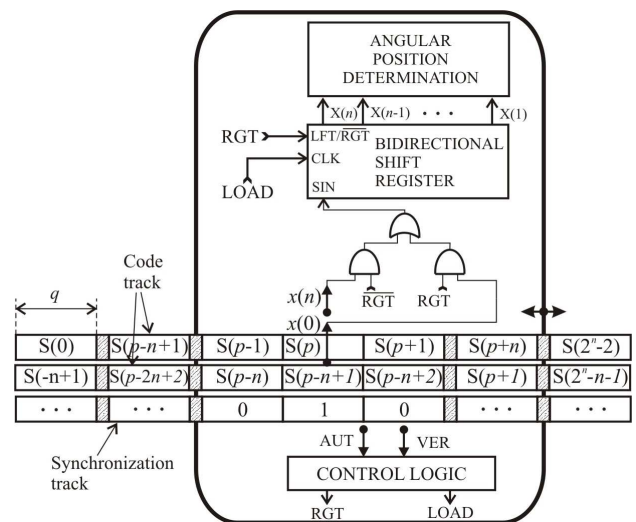d on the Fibonacci generator of 10-bit PRBS, Fig. 5. The code word, which was adopted as the initial code word is used as a reference with which to compare the read code word and the code word generated in the Fibonacci generator in each clock period. When the initial code word equalize with code word in Fibonacci generator of PRBS the 10-bit counter is stopped, and at its output provides information about the position in natural code. The feedback configuration of the Fibonacci generator is defined for 10-bit PRBS and can be found in the literature. This conversion process is here realized in the block diagram of virtual instrument and his speed depends on computer configuration.

III. FRONT PANEL OF THE VIRTUAL INSTRUMENT

Realized virtual instrument, Fig. 6, on its front panel displays in real time the signals from the synchronization track (channels A and B), as well as two signals from the pseudorandom code tracks (channels C and D). Information about angular position and angular velocity is also displayed. Angular position obtained in the opposite direction of rotation of the shaft encoder has a negative sign. The indication of the direction of the shaft rotation is also given. The input signals are recorded on the hard disk of the computer, for eventual their later processing or analysis, and on the front panel can be changed location and name of the recorded file. Also, parameters for input signal acquisition (buffer size, sampling rate) can be changed in the front panel.
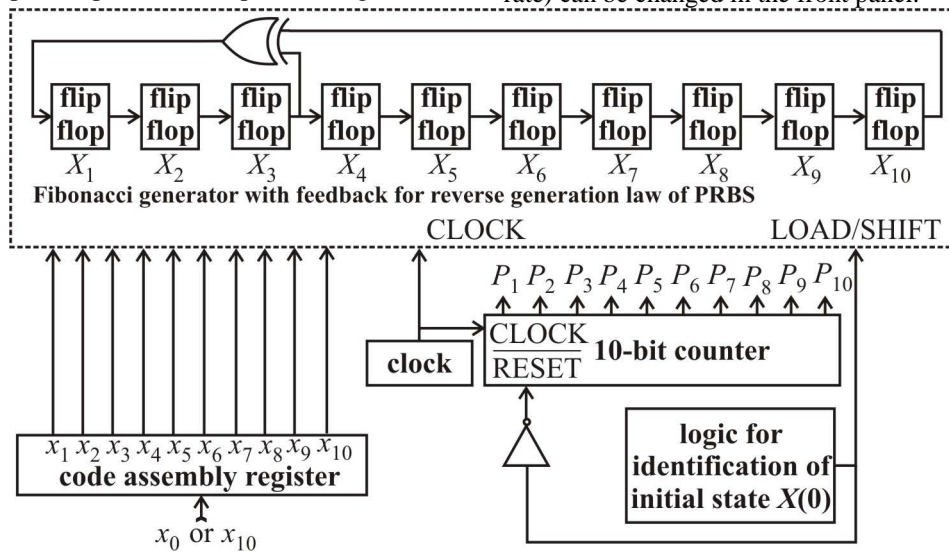

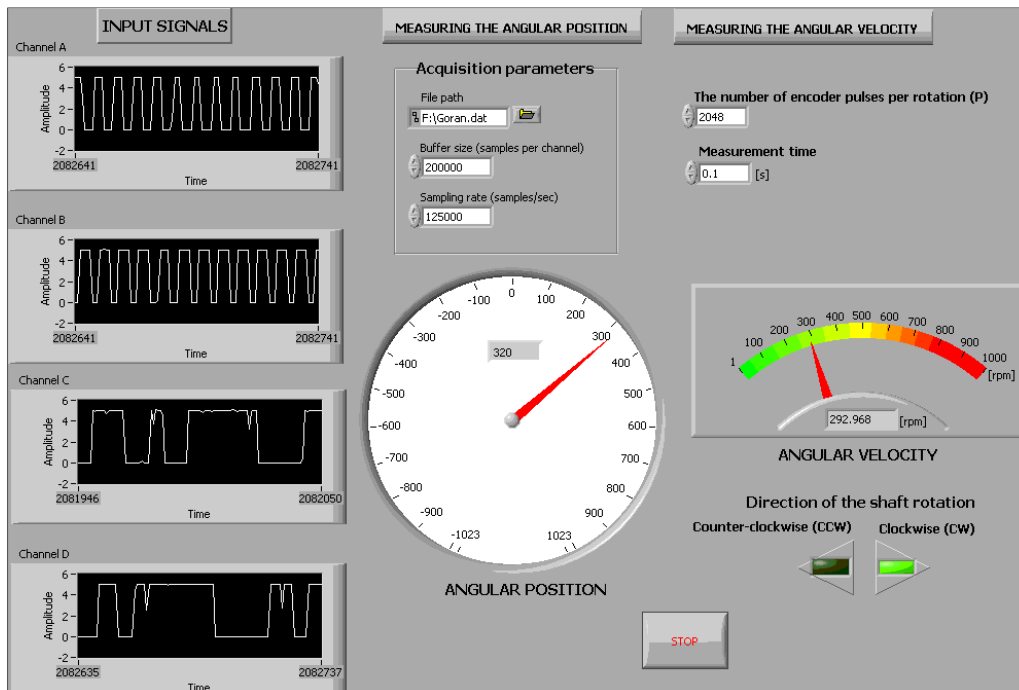Fig. 5. The serial pseudorandom/natural code converter

Fig. 6. Virtual instrument for measuring the angular position and velocity

Parameters, the number of encoder pulses per rotation, P, and measurement time, T_S, which are used for angular velocity determination can be also changed in the front panel of virtual instrument. This realized virtual instrument can be subsequently modified and changed, depending on the specific application, input signals, etc. This instrument is portable in case of using laptops. Its performance depends on the used acquisition card, as well as the used computer configuration. Electronic block of encoder is usually made using a microprocessor or programmable FPGA circuits for maximum performance. However, and this is realized measuring the angular position and velocity may very well find its practical application in various areas where you need precise positioning of different machines, robots, telescopes, cranes, etc. Graphical programming significantly reduces the time required for the implementation of such a system.

IV. CONCLUSION

Presented virtual instrument for measuring angular position and velocity represents a modern, modular, adaptive system which uses signals from code disk of pseudorandom position encoder. Realized virtual instrument uses a modular hardware, computer software, which allows you to expand the variety and upgrading performance. This virtual instrument allows the measurement of angular position and velocity based on the signals obtained from the electronic system of reading the code with encoder code disk and storing these signals, adjust the signal acquisition and observation signals in real time on the front panel of the instrument. The virtual instrument performance depends on the used computer configuration and modular hardware.

ACKNOWLEDGEMENT

REFERENCES

[1] V. Smiesko and K. Kovac, "Virtual instrumentation and distributed measurement systems", Journal of Electrical Engineering, vol. 55, no. 1-2, pp. 50 - 56, 2004.
[2] G.S. Georgiev, G.T. Georgiev, S.L. Stefanova, „Virtual instruments – functional model, organization and programming architecture", Information theories & applications, vol. 10, pp. 472-476, 2003.
[3] X. Jing, B. Xu, C. Wang, S. Yhu, G. Pu, S. Dong, "Evaluation of measurement uncertainties of virtual instruments", Int. J. Adv. Manuf. Technol., vol. 27, no. 11-12, pp. 1202 - 1210, 2006.
[4] F.J. MacWilliams, N.J.A Slone, "Pseudo-random sequences and arrays", Proceeding of IEEE, Vol. 64, No. 12, pp. 1715-1728, 1976.
[5] E. M. Petriu, J. S. Basran, "On the position measurement of automated guided vehicles using pseudorandom encoding", IEEE Trans. IM, vol. 38, no. 3, pp. 799-803, 1989.
[6] M. Arsić, D. Denić, "New pseudorandom code reading method applied to position encoders", Electron. Lett. vol. 29, pp. 893-894, 1993.
[7] D. Denić, G. Miljković, "Code reading synchronization method for pseudorandom position encoders", Sensor. Actuat. A-Phys. vol. 150, pp. 188-191, 2009.
[8] E.M. Petriu, J.S. Basran, F.C.A. Groen, "Automated guided vehicle position recovery", IEEE Trans. Instrum. Meas. vol. 39, pp. 254-258, 1990.
[9] D. Denić, M. Arsić, "Checking of pseudorandom code reading correctness", Electron. Lett. vol. 29, pp. 1843-1844, 1993.
[10] D. Denić, I. Stojković, "Pseudorandom/natural code converter with parallel feedback logic configuration", Electron. Lett. vol. 46, pp. 921-922, 2010.
[11] G. Miljković, D. Živanović, D. Denić, J. Lukić, M. Simić, "Virtual instrumentation used for angular velocity measurements and detection of shaft rotation direction", 11th International Scientific Conference UNITECH '11, Technical University of Gabrovo, Gabrovo, Bulgaria, 18-19 November 2011, vol. II, pp. 403-408, 2011.
[12] N. Patrascoiu A. Poanta A. Tomus, B. Sochirca, "Virtual Instrumentation used for Displacement and Angular Speed Measurement", International Journal of Circuits, Systems and Signal Processing, vol. 5, no. 2, ISSN: 1998-4464, pp. 168-175, 2011.

Dielectric Strenght Testing of Transformer Oil

Natasa Mojsoska [1] and Blagoja Stevanoski [1]

Abstract – **Main issue in this paper work is transformer oil. The function of liquid dielectric is to provide insulation and cooling. The various characteristics of transformer oil are describes such as physical, chemical and electrical. Dielectric strength, also called breakdown voltage BDV is one of electrical parameters that is very important, specially for transformer oil in use. Described is the process of purification of the oil sample, with the elements of the Purifying station for transformer oil. Measured is the dielectric strength of the transformer oil.**

Keywords – **Transformer, Transformer oil, Dielectric strength, Measurement.**

I. INTRODUCTION

The parameters of transformer oil are categorized as:

1. Electrical parameters: dielectric strength, specific resistance, dielectric dissipation factor.

2. Chemical parameters: water content, acidity, sediment content.

3. Physical parameters: interfacial tension, viscosity, flashpoint, firepoint.

The function of transformer oil is to provide insulation and cooling. Because of these oil properties associated with the composition, transformer oil should fulfill several requirements:

1. In the cold condition must be sufficiently rare to run across all active parts of the transformer.

2. It must not contain moisture because it reduces dielectric strength.

3. Does not contain any solid materials, because they reduce the insulating properties.

4. Do not contain acid and sulphur that eat away and destroy insulation.

5. Must have high thermal flammability.

6. Must be able to apply at low temperatures.

7. The new oil should have a dielectric strength of at least 80 kV / 2,5 cm, and the transformers in operation must have at least 30 kV / 2,5 cm.

II. CHARACTERISTICS OF THE OIL

A. Physical and chemical properties of oil

The purpose of testing is to determine the suitability of oil

[1]Natasa Mojsoska is with the Technical Faculty at University St Clement Ohridski of Bitola, Makedonska falanga 33, Bitola 7000, Macedonia, E-mail: natasa.mojsoska@tfb.uklo.edu.mk.
[1] Blagoja Stavanoski is with the Technical Faculty at University St Clement Ohridski of Bitola, Makedonska falanga 33, Bitola 7000, Macedonia, E-mail: blagoj.stevanovski@uklo.edu.mk.

application and assessment of its quality. Based on the results of the measurements it is possible to determine the corrective measures in terms of bringing the oil condition which is allowed for use.

Color and clarity - is not a critical feature, but it is useful in comparative terms (quick change of color can be a sign of accelerated aging and pollution).

Water content - water in the oil derived from the production itself, from the humidity of the atmosphere or it occurs as a product of aging paper (cellulose).

Acidity - it is a measure of free organic and inorganic acid present in the oil.

Sediment - it occurs as a result of the aging of the oil and paper insulation in the working life.

Neutralization number - it is a measure of the acidic compounds or oil pollution.

Interfacial tension - measured at the oil-water, and indicates soluble polar impurities and degradation products, the value is changing quite rapidly in the initial stage of aging, therefore the results should be interpreted in conjunction with other features.

Oxidation stability - is determined by measuring the aging of the product oil in the laboratory, thus measuring the content of sediment, volatile or soluble acids.

Viscosity - n few wards, viscosity of transformer oil can be said that is the resistance of flow, at normal condition. A good oil should have low viscosity so that it offers less resistance to the convectional flow of oil thereby not affecting the cooling of transformer. Low viscosity of transformer oil is essential, but it is equally important that, the viscosity of oil should increase as less as possible with decrease in temperature.

Flash point and fire point - Flash point of a volatile liquid is the lowest temperature at which it can vaporize to form an ignitable mixture in air. The flash point is an empirical measurement rather than a fundamental physical parameter. The Fire point is defined as the temperature at which the vapor continues to burn after being ignited. It is the lowest temperature at which, on further heating beyond the flash point, the sample support and combustion for five seconds.

B. Electrical parameters of oil

Specific resistance - This is another important property of transformer oil. This is measure of DC resistance between two opposite sides of one cm3 block of oil. Its unit is taken as ohm-cm at specific temperature. With increase in temperature the resistivity of oil decreases rapidly. Just after charging a transformer after long shut down, the temperature of the oil will be at ambient temperature and during full load the temperature will be very high and may go up to 90°C at over load condition. So resistivity of the insulating oil must be high at room temperature and also it should have good value at high temperature as well.

That is why specific resistance or resistivity of transformer oil should be measured at 27°C as well as 90°C.

Minimum standard specific resistance of transformer oil at 90°C is 35×1012 ohm–cm and at 27°C it is 1500×1012 ohm–cm.

Dielectric dissipation factor DDF is also known **as loss factor** of transformer oil. When a insulating materials is placed between live part and grounded part of an electrical equipment, leakage current will flow. As insulating material is dielectric in nature the current through the insulation ideally leads the voltage by 90°. Here voltage means the instantaneous voltage between live part and ground of the equipment. But in reality no insulating materials are perfect dielectric in nature. Hence current through the insulator will lead the voltage with an angle little bit shorter than 90°. Tangent of the angle by which it is short of 90° is called dielectric dissipation factor or simply tan delta of transformer oil.

Dielectric strength of transformer oil is also known as **breakdown voltage** of transformer oil or **BDV** of transformer oil. It refers to the ability of the oil to withstand dielectric stress.

Breakdown voltage is measured by observing at what voltage, sparking strands between two electrodes immerged in the oil, separated by specific gap. Low value of BDV indicates presence of moisture content and conducting substances in the oil

The dielectric strength of transformer oil at transformers over 10 MVA must control twice a year and at transformers under 10 MVA once a year.

Dielectric strength of an insulation material depends on pressure, temperature, humidity, electrode configuration and nature of applied voltage. Breakdown strength analysis of transformer oil gives effective results through which suitable dielectric material for the related high voltage applications can be explored. The Breakdown Voltage (BDV) test kit consists of two electrodes mounted on horizontal axis with 2.5 mm gap and enclosed in a glass chamber as shown in Figure 2. Electrodes used in test apparatus is of sphere-sphere electrode configuration since the sphere-sphere configurations provides uniform field distribution. Diameter of sphere electrodes is of 19.8 mm.

C. Structural analysis of the oil

Is made by the method of infrared spectrometry and provides information on the chemical structure of the oil. In case of accelerated aging of the oil can detected the degree of chemical degradation.

Example for physical - chemical analysis of the oil or characteristics, appearance or her unit are measured with specified application for a declaration of the property is given in the Table 1.

The criteria for the purpose in the table are based on experiences gained good insulating oils in the transformer factory, and comply with the requirements of IEC 422/89.

TABLE I
PHYSICAL AND CHEMICAL CHARACTERISTICS OF TRANSFORMER OIL

	measure unit	request
looks	clean	clean
collor (acc. DIN 51578)	-	-
neutralization value total acidity	mg KOH/g	<0.2(<0.3)
Interfacial tension at 20°C	mN/m	>25(>20)
Presence of oxidation inhibitor	-	present
Sediment value n-heptanol	-	absent
Dielectric strength (BDV electr. 2.5mm)	kV	>50(>40)
Factor of dielectric losses at 90°C	-	<0.1(<0.2)
Specific resistance at 90°C	GΩm	>10 (>1)
Water content	mg/kg	<20(<30)

III. TECHNOLOGY PURIFICATION OIL

A. Vacuum purifying transformer oil

The device works on the principle of eliminating the drying gas and oil in the vacuum chamber, unlike the older generation where they used spin.

With its features, and depending on the size of the device and its flow and size (power) of the transformer is possible in the short term to achieve a very high dielectric strength of oil. Device works in a completely closed system, and each contact of humans and animals oil environment is minimized.

The process of purifying transformer oil down into several phases that subsequently occur. The oil is first filtered by filter coarse impurities and heated to 55 ° C. The inherent drying chamber is under vacuum -1 bar resulting in a lower boiling point of water at approx. 40 ° C.

In dryer water and undesirable gases evaporate and drain through a vacuum pump, and oil further filters filters and returns to the transformer.

Quality oil clean transformer in the circular process through the device. This is done until the oil has reached a high dielectric strength.

The advantage of the technology of purification of oil is beneficial oil additives that are not separated from the oil. Extraction of oil additives significantly accelerates aging of the oil itself.

Examples of devices that will describe the technology of purification with this method is WH-020 of the famous Swiss company Micafil.

B. Preparation for treatment

Before you begin the process of purifying transformer oil necessary to ensure the workplace in accordance with the "golden rules to protect electricity", which includes the following:

1. It appears to be separated from the transformer voltage,

2. Prevent reconnection of the transformer (put insulation board)

3. indicator to check no voltage condition (indicator previously examined).

Fig. 1. Connecting the transformer with device by hoses

4. earthling and short merged transformer,
5. fence parts under voltage
After that the procedure can start.

The transformer is covered with a cover to keep the heat from the transformer oil circulating through the transformer machine and back, and also reduces the duration of the procedure. The cover is made only in terms of low temperatures.

Input and output hose device to circulate the oil in the process of purification is related to the power transformer, as shown in Fig. 1.

C. Power plug the device

Preparatory work includes planning of connecting power connector installed power of device connectivity and planning due to its remoteness Fig. 2.

Fig. 2. Connect the device to a power source and connectors with extensions

D. Preparation of the vehicle treatment

The transmission device for purifying transformer oil, which is located in the vehicle has a great advantage, because the test is done on-site.

E. Cleaning equipment for transformer oil

Purifying station for transformer oil type Micafil is given on Fig. 3.
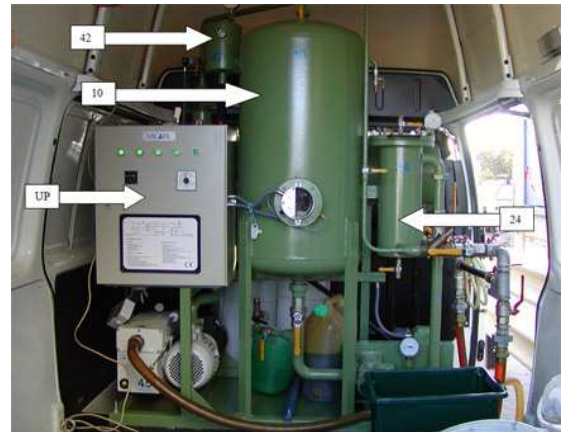The image is labeled with the following elements:

Fig. 3. View of purifying station inside

• 10 - drying chamber and degassing,
• 24 - 5 micron filters
• 42 - gas oil indicator suction vacuum pump oil,
• UP - Control Panel device.

Oil is brought through inlet hoses in the device, and then heated. It is expelled vacuum drying and removal of gases, which through a series of sieves is purified oil, water and gas and side drain through a vacuum pump. Warm oil with high quality returns to the transformer and wash it in a circular process 3-5 times. With help the pilot light lit inside the chamber and control the oil in the chamber, shown on the Fig. 4.a.

During treatment, it is possible to control the temperature of
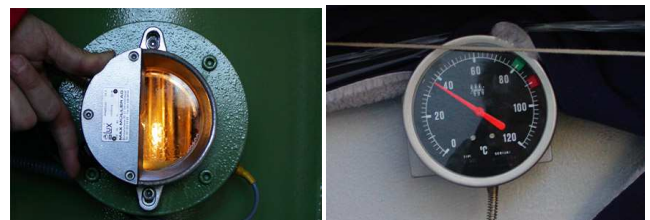
Fig. 4.a Control the flow of oil from the window of the chamber

Fig. 4.b Contact thermometer in power transformer

the oil which circulates through a power transformer with a contact thermometer (given on Fig. 4.b)

F. The taking of samples and measuring the dielectric strength of transformer oil

Before testing the transformer oil and measurement of dielectric strength with container were adjusted distance between the electrodes of 2.5 mm shown on Fig. 5.

Fig. 5. Calibration of distance between electrodes of 2.5 mm

G. Releasing oil from the device

The oil is discharged from the device to remove impurities from drainage pipes and to take a clean copy. A copy of the oil control is taken in dry weather, and if necessary the wet weather then take special measures for protection from spraying (e.g. protection from rain, deleted and dried floodgates, waterproof covers transport of the sample).

The frequency and method of sampling of transformer oil in order to verify its dielectric strength (penetration) must be carried out in accordance with existing standards and regulations.

Fig. 6. Realizing oil from device, Filling the container with the sample to measure the oil

IV. RESULTS OF MEASURING

Portable Tester for testing dielectric strength of transformer oil in photos is BAUR Dieltest DPO60.

Fig. 7 shows that the measured dielectric strength of the oil is **46,9 kV / 2.5 mm**, which is the result of **187,6 kV / 1cm.**

Fig. 7. Measuring of breakdown voltage BDV of the oil

REFERENCES

[1] "Investigations of Transformer Oil Characteristics", Chapter 2, http://shodhganga.inflibnet.ac.in:8080/jspui/bitstream/10603/16350/7/07_chapter%202.pdf

[2] T. Toudja, A. Nacer, H. Moulai, I. Khelfane and A. Debche, "Physico-chemical properties of transformer mineral oils submitted to moisture and electrical discharges ", ICREPQ '12, http://www.icrepq.com/icrepq'12/538-toudja.pdf

[3] A. Müller, M. Jovalekic and S. Tenbohlen, "Assessment of Oil Analysis Data For Medium Voltage Distribution Transformers", XVII International Symposium on High Voltage Engineering, Hannover, Germany, August 22-26, 2011

[4] http://www.baur.at/en/products/insulating-oil-testing/breakdown-detection/dpa-60-c

ORAL SESSION
RADIO COMMUNICATIONS, MICROWAVES, ANTENNAS

Continuous Revolution Method for Antenna Radiation Pattern Measurements

Miroslav Gechev [1], Kliment Angelov [2], Boyan Kehayov [3], Sava Denev [4], and Svetoslav Kremenski [5]

Abstract – **In this paper a new method for antenna radiation pattern measurements has been presented. An idea for continuously rotating the antenna and mathematical processing of the collected data is described. That gives the opportunity for minimizing the error and simplifying the measurement techniques. The results can be used in case of automatized low-cost antenna tests with improved accuracy level.**

Keywords – **Radiation pattern, Antenna measurements, Improved accuracy.**

I. INTRODUCTION

The radiocommunications, and more specifically the use of electromagnetic waves (EMWs) for the transmission of information, are an integral part of the everyday life of the modern person. The transmission and reception of EMWs is accomplished by using antennas, which have requirements for specific characteristics and parameters, which in turn are based on the specific use of the antenna in the various communication systems. Many analytical and simulation methods are used in the construction of the antennas, but still the need for experimental measurements exist. This is a fact, on one hand, due to the need of meeting and/or exceeding the specific technical requirements, and on the other hand, in order to create a better living environment for the population by reducing the unwanted electromagnetic radiation.

A key characteristic of each antenna is its radiation pattern (RP) [1] [2]. It shows how the antenna distributes spatially its transmission and/or reception of EMWs in/from the different directions. There are two types of RPs – directional and omnidirectional, and in some cases a more specific shape of the RP is desirable (e.g. in cell phones it is desirable to reduce the radiation in the direction of the head of the user [3]). The experimental determination of the RP is connected to the requirement for different spatial orientation of the antenna under test (AUT) for the specific equipment used for the measurement. Two degrees of freedom (rotation in two orthogonal axes) are required as well, but still in most of the cases a rotation in only one axis could be used. After measuring the RP on the first plane, the AUT is positioned orthogonally to the first position (plane) and a measurement of the RP on the second plane is performed. Interruption and manual set of the spatial position of the AUT leads to errors, avoiding which is especially critical when testing and measuring antennas that are highly directional and with narrow beams on its RP.

In this paper a method for automated measurement of antennas' RPs is proposed – it is based on rotation on one axis with multiple repetitions of spins and subsequent mathematical processing of the reported results in order to achieve higher measurement accuracy.

II. DESCRIPTION OF THE PROBLEM

A. Radiation Pattern Measurements

Problems associated with experimental tests of antennas are examined in numerous scientific studies, where determining the RP is one of the most commonly discussed issues [4][5][6][7]. Table I shows the basic criteria for classification of the measurements of RPs during the study of an antenna. Usually when conducting low-budget experiments with a relatively low degree of precision, stands with one degree of freedom are used – the rotation is on one axis only. Also, in such cases the spatial orientation of the antenna under test is determined manually and the measurements are conducted in outdoor conditions, rather than in an anechoic chamber that reduces any disruptive and/or reflected signals.

When it is necessary to carry out more high-tech measurements, which aim to increase the accuracy of the results, more complex mechanical positioning stands with two degrees of freedom are used (the rotation is on two orthogonal axes).

Also – it is very much appropriate to conduct the experimental measurements in an anechoic chamber. This would increase the accuracy of the test results, but it must be noted that the economic cost of such facility is very high and it depends on the working frequency bands the chamber is aimed to be applicable for.

The use of anechoic chamber is economically justified only when used regularly.

[1] Miroslav Gechev, PhD student is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: miroslav.gechev@gmail.com;

[2] Kliment Angelov, chief assistant professor, PhD is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria. E-mail: kna@tu-sofia.bg;

[3] Boyan Kehayov, assistant professor is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: bkehayov@tu-sofia.bg;

[4] Sava Denev is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria;

[5] Svetoslav Kremenski, student is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria.

TABLE I
RADIATION PATTERN MEASURING CLASSIFICATION

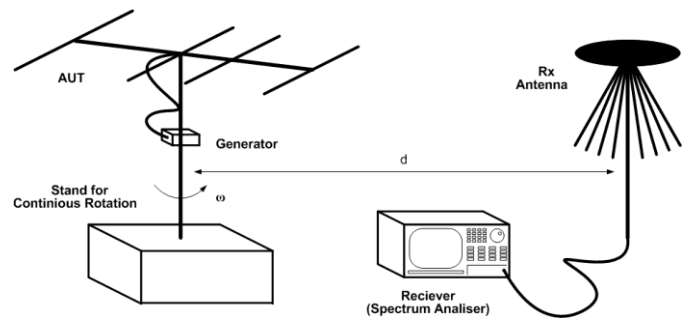| Number of Rotation Axes	One Axis
	Two Axes
Measuring Distance	Near Field Measurements
	Far Field Measurements
Position Angle Settings	Manual
	Mechanically Automatized
Antenna Mode	Transmitting
	Receiving
Minimization of the Interference	Free Space Measurements
	Anechoic Chamber

Fig. 1. Scheme of Test Stand

Another advantage of the aforementioned method is the possibility for a relative reading of the angle of the position of the AUT − θ. In the case of an absolute set of this positional angle, it is required to choose a neutral (starting) position at the beginning of the measurement in contrast to which to report the radiation in the other directions. In the proposed in this paper method, the angle θ can be reported after finding the average analytical dependence of the radiation pattern, keeping in mind the fact that its main lobe rises from the continuous rotation of the antenna with angular velocity ω. The minimal angular size in the reporting of the RP can be determined by applying the Nyquist principle at given angular velocity ω and sampling period of the received signal T_s:

$$\theta_{min} = 12.T_s.\omega.$$ (1)

In Equation (1) θ_{min} is obtained in degrees as T_s is in seconds and ω in rpm.

In terms of the distance of measurement d two cases can be considered [5] – if the measurements are carried out in the near field of the antenna or if carried in the antenna's far field. In the former case a more complex measurement equipment is needed but based on the obtained experimental results and with the help of analytical methods the far field radiation of the antenna can be determined [6][7].

C. Test Implementation

Based on the idea proposed in this paper a simplified measurement setup was built. The goal is to verify the reliability of the proposed method by using repetitive/multiturn measurement of the radiation pattern of a simple Yagi-Uda antenna and after the accumulation of enough data a mathematical analysis can be applied. The resulting averaged RP from the test then can be compared to the RP obtained by the simulation in order to determine the accuracy of the measurement.

Table II shows the implementation characteristics of the main elements of the test stand. The spectrum analyser GW Instek [8] model GSP-830 is used as a measuring receiver, which with the help of specialized proprietary software developed by the manufacturer allows for the data of the spectral image at specific intervals to be stored in the memory of a personal computer.

During the measurement itself the requirement to perform the measurement in the far field of the source is fulfilled [9][10]:

B. Theoretical Basis

Based on the overview of the methods for measuring the antenna parameters, it becomes clear that it is more relevant to seek accomplishment of a smaller error when using the more simpler and cost-efficient testing techniques. In the theory, in order to reduce the negative impact of random and sporadic errors, we resort to measurement repetitions and result averaging. In the case of experimentally determining the radiation pattern of the antenna under test, this repetition can be accomplished by repeated (multiturn) rotation of the antenna around the axis of rotation for the specific measurement. Fig. 1 shows an example scheme of the test stand based on the aforementioned concept.

If the angular velocity of rotation ω of the antenna is constant, then at a constant sampling period of the received signal a recurrent dependence would be obtained. From the data thus obtained and by making use of the mathematical analysis the analytical dependencies determining the RP based on the averaged data for each rotation period can be found.

$$d > \frac{2.D^2}{\lambda},$$ (2)

where D shows the maximal dimension of the antenna as λ shows the working wavelength.

TABLE II
TEST IMPLEMENTATION CHARACTERISTICS

AUT	Type	Yagi-Uda
	Number of Elements	Three
	Maximal Dimension	0,5 m
	Theoretical Directivity	7,52 dBi
Transmitter	Frequency	434 MHz
	RF Power	0 dBm
Stand for Continuous Rotation	Number of Axes	One
	Rotation Speed	~1,4 rpm
Measuring Distance		10 m
Receiving Equipment	Spectrum Analyser	GSP-830
	Receiving Antenna	Omnidirectional
	Sampling Speed	1 s^{-1}

III. RESULTS

Fig. 2 shows the received spectrum affected by the radiation of the AUT at specific point in time.
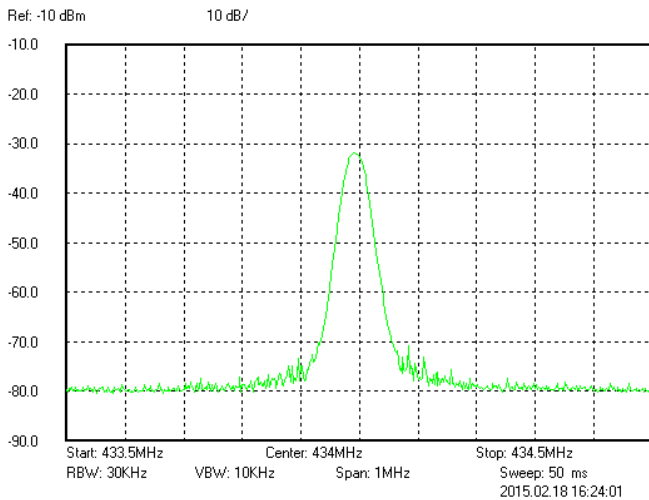

Fig. 2. Received Spectrum

Fig. 3 shows the measured change in time of the received by the spectrum analyser signal for the transmitter's transmission frequency. The shown duration of the signal covers several rotation turns of the AUT. In some of the turns more irregular changes were noticed that were caused by outside interference from various kinds. After averaging the RP by using the method of repetitive measurements, the impact of the outside interference shall be minimized.
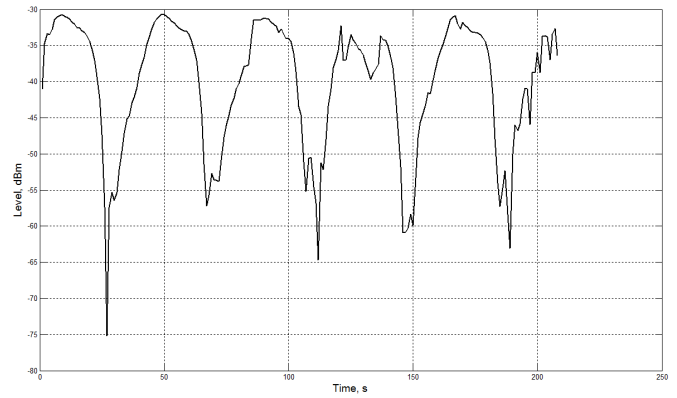

Fig. 3. Received Periodic Signal

Fig. 4 shows the averaged radiation pattern after processing the shown on Fig. 3 periodic signal using the MATLAB [11] environment.
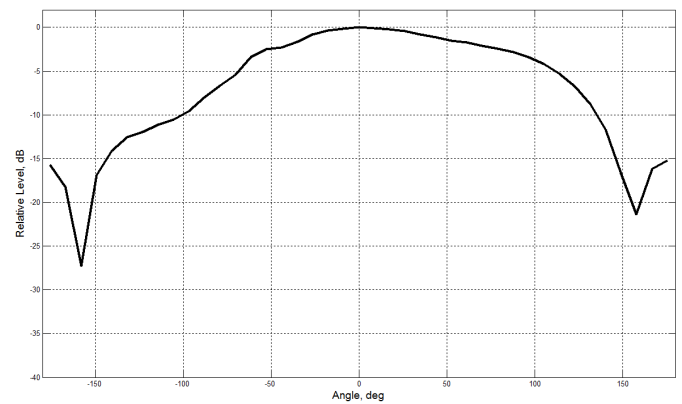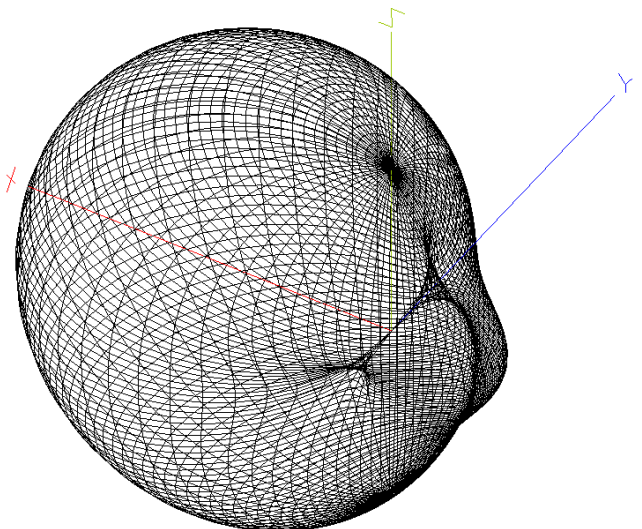

Fig. 4. Averaged Radiation Pattern

Fig. 5. Simulated 3D Radiation Pattern

Fig. 5 shows the resulting three-dimensional simulated radiation pattern of the antenna under test in free space.

The antenna's two-dimensional cross-section in the plane used in the performing of the experimental study is shown on Fig. 6.
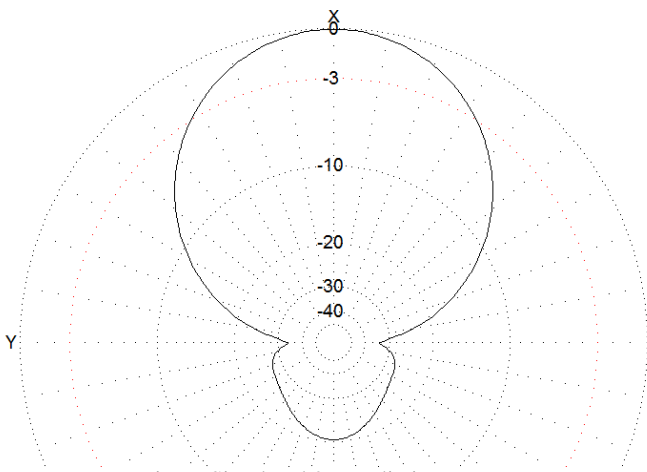
Fig. 6. Simulated 2D Radiation Pattern

Fig. 7 shows the experimentally obtained averaged radiation pattern. There can be seen the difference in the position of the directions with minimal radiation, as well as the angular width of the RP's elements. This is due to the poor isolation of the influence of the reflected signals during the implementation of the test.

In general there are significant similarities as well – the levels of backward radiation and those in the areas of minima are close. Also, the experimentally determined radiation pattern coefficient, which value is 7,57 dBi, is close to the theoretical one – 7,52 dBi. Also, there is a minimization of the interferences and it can be seen in the primary data of the measurement.
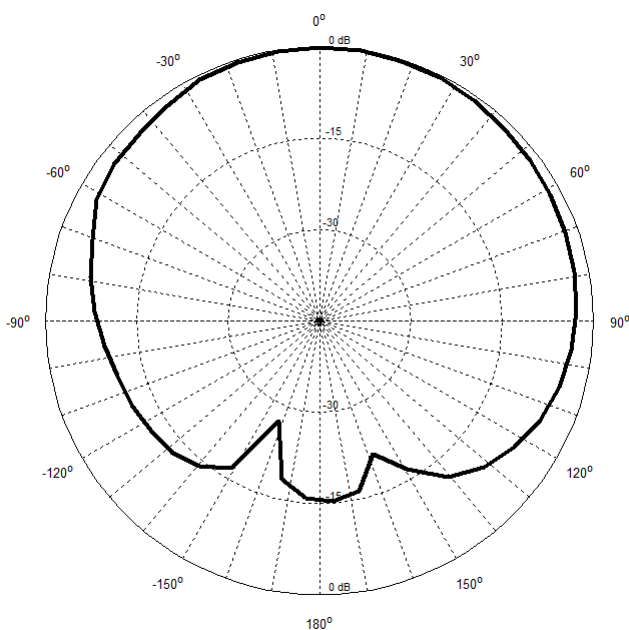

Fig. 7. Experimental Radiation Pattern in Polar Coordinates

IV. CONCLUSION

Based on the obtained results the following conclusions can be drawn:

- the method proposed in this paper for determining the radiation pattern and the subsequent analytical processing of the periodic data provides encouraging results for the applicability of the method in practical measurements;
- there is a minimization of interferences, seen in some of the antenna's revolutions, because of the process of mathematical processing of the collected measurement data;
- it is needed to look for additional optimizations, in order to enhance the method's accuracy.

ACKNOWLEDGEMENT

REFERENCES

[1] Johnson R. C., Antenna Engineering Handbook, 3-rd ed., McGraw-Hill, 1993.

[2] Collin R. E.,Antennas and Radiowave Propagation, McGraw-Hill, 1985.

[3] Schiavoni A., P. Bertotto, G. Richiardi, P. Bielli, SAR Generated by Commercial Cellular Phones – Phone Modeling, Head Modeling and Measurements, IEEE Transaction on Microwave Theory and Techniques, Vol. 48, No. 11 Part: 2, pp. 2064-2071, Nov. 2000.

[4] Bhavsar V., N. Blas, H. Nguyen, A. Balandin, Measurement of Antenna Radiation Patterns, EE117 Laboratory Manual, UC-Riverside, 2000.

[5] Icheln C., Methods Measuring RF Radiation Properties of Small Antenas, Helsinki University of Technology Radio Laboratory Publications, Teknillisen korkeakoulun Radiolaboratorion julkaisuja Espoo, October, 2001.

[6] Hansen J.E., Spherical Near-Field Measurements, IEE Electromagnetic Waves Series 26, 387p, London, UK, 1988.

[7] Yaghjian A.D., An Overview of Near-Field Antenna Measurements, IEEE Transactions on Antenna and Propagation, Vol. AP-34, No. 1, pp. 30-45, Jan. 1986.

[8] www.gwinstek.com

[9] Antenna Measurement Theory, Introduction to Antenna Measurement, Reprinted with the permission of ORBIT/FR Inc., www.orbitfr.com

[10] Foegelle M. D., Antenna Pattern Measurement: Concepts and Techniques, Reprinted from Compliance Engineering, Annual Reference Guide 2002 • Copyright © 2002 Canon Communications LLC

[11] MATLAB IMAGE PROCESSING TOOLBOX. User's Guide, "The Math - Works Inc.", 2000. www.mathworks.com

Wave Digital Approach in Characterization of T-junction Discontinuity in Microstrip Stub-line Structures

Biljana P. Sto–i

Abstract – **The purpose of this paper is to show that the one-dimensional wave digital approach can be used in analysis of stub-line structures. A stub-line structure, divided into uniform sections, can be efficiently modeled by wave digital network. Frequency response is obtained by direct analysis of formed block-based network. Also, a complete understanding of microstrip circuits requires characterization of T-junction discontinuities included in the circuits. Two different approaches to the modeling of the equivalent network of T-junction discontinuity are presented in this paper. Open T-resonator circuit, proving the response accuracy of the new technique, is given.**

Keywords – **Wave digital approach, wave digital networks, microstrip circuits, stub-line structures, T-resonator circuit.**

I. INTRODUCTION

The modeling of the planar structures by wave digital elements is based on well known theory of wave digital filters (WDFs). A detailed review of WDF theory is given in references [1-4]. A large variety of WDF-based techniques has been developed for a wide range of applications [2]. Recently, great effort has been placed on modeling of microwave planar structures are modeled by one-dimensional (1D) [5-8], [10-14], and by two-dimensional (2D) [5], [9] wave digital elements (WDEs). Standard WDEs, such as delay, adder, multiplier and adaptor, are used in design of wave digital networks (WDNs). The reader is refered to [7], [10-13] for a thorough introduction to WDNs.

Microstrip lines are widely used in microwave and RF circuits due to its superior compatibility to the circuit environment for easy integration and manufacturing. In practice, there does not essentially exist just one type of microstrip structures, but a whole variety of quite distinct subclasses (stepped-impedance filters, linearly tapered nonuniform transmission lines, stub-line structures, elliptic filters), each of which can again be divided into varies families, etc. This reflects the richness of WDNs, and the most appropriate one have to be chosen for structure at hand.

In the previously published paper [9], microstrip structure with stub is analyzed by use of 2D wave digital approach. In the paper [14], the authors have had mainly deal with 1D wave digital approach. In fact, in view of their geometry, stub-line structures are also natural candidates for 1D approach. A way for connecting segments in the structure with stubs is also developed and also represented in the Section II in order to address both readers with and without background in 1D wave digital approach.

Biljana P. Sto–i is with the University of Ni–, Faculty of Electronic Engineering, Department of Telecommunications, Aleksandra Medvedeva 14, Ni–, Serbia, (e-mail: biljana.stosic@elfak.ni.ac.rs)

Microstrip circuits inevitably incorporate transmission line discontinuities of one type or another. A complete understanding of microstrip circuits requires characterization of various discontinuities included in the circuit. Two different approaches to the modeling of the equivalent network of T-junction discontinuity are described in the Section III.

In the Section IV, simulation validation of the proposed modeling approach is provided by means of an open T-resonator circuit.

The main contribution of this paper is developed wave digital network model for modeling the equivalent circuit of T-junction discontinuity. The equivalent circuit elements can be determined by closed-form expressions or by carrying out EM simulations. Here, it is shown how WDN model can be derived starting with multiport capacitor in parallel branch [9].

II. MODEL OF THE STRUCTURE WITH STUB

In the paper [14], one-dimensional wave digital approach is shown to be suitable for the blocked-based analysis of microstrip stub-line structures.

The formed Simulink model of WDN given in [14] is depicted in Fig. 1. The blocks **TLine_1**, **Stub_2** and **TLine_3** represent uniform transmission line (UTL) segments. The blocks **ADP-S** and **ADP-L** represent two-port series adaptors, and block **ADP_T1S2T3** three-port parallel adaptor with port 2 being dependent.
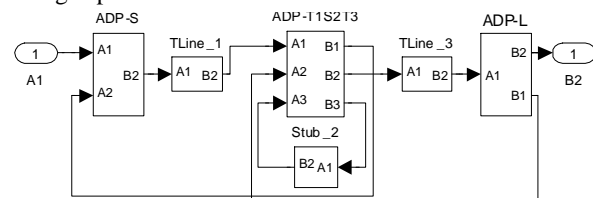
Fig. 1. Simulink model of the structure

III. MODELING OF THE T-JUNCTION DISCONTINUITY

Here, only discontinuities associated with the T-pattern resonator, such as open end and T-junction, are considered. Since discontinuity dimensions are usually much smaller that the wavelength in a microstrip, they can be modeled by lumped-element equivalent ciruits.

A T-junction discontinuity is very offten used in the microstrip circuits, such as microstrip stubs, power dividers, etc [15].

An open-end discontinuity occurs frequently in a number of circuits such as resonators, matching stubs, filters, and microstrip antennas. A closed form expressions for calculating the excess length of transmission line [16] are used here.

Fig. 2. (*a*) Symmetric T-junction, and
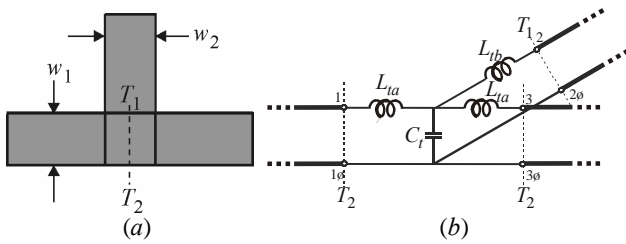(*b*) equivalent network of the discontinuity

For a symmetric T-junction structure depicted in Fig. 2*a*, the equivalent network is shown in Fig. 2*b*. It consists of the inductances L_{ta} and L_{tb} in the series branches, and the parallel capacitance C_S. The inductances and capacitance can be obtained by closed form expressions given in the papers [3-4]. Here, two different ways for modeling the equivalent circuit of the T-junction discontinuity are given. The modeling procedures are based on:

- decreasing line lengths, and
- a wave digital element.

A. Modeling by Decreasing Line Lengths

The simplest modeling approach of the equivalent discontinuity network involves changing line lengths. In order to absorbe discontinuity effects a new line lengts are counted. Corrections for T-junction effects are done as follows: the physical length of the series line which length is l_1 and width w_1 is decreased by $w_2/2$ value, i.e. it is decreased by one half of width of the stub line. The physical length of the stub line with length l_2 and width w_2 is decreased by $w_1/2$ value, i.e. it is decreased by one half of width of the series line [17].

The WDN of a T-resonator circuit with discontinuities is obtained in a few steps. First, new lengths of the structure segments are calculated. Then, the procedure described in the section II can obviously also be applied to the obtained structure. Finally, WDN of a nonuniform planar microstrip structure is the same as that one shown in Fig. 1. A number of sections in UTL segments differs from the previous one because the segments lengths are being changed.

This approach of modeling tee-junction discontinuity is very acceptable because there are not new blocks here. In the case of a microstrip stub-line structure the proposed method reduces drastically the computation time while giving acceptable accuracy.

B. Wave Digital Element - Direct Realization of the LC Elements of the Equivalent Discontinuity Network

In the other modeling approaches of the step discontinuity, new blocks have to be included in the WDN. A WDN depicted in Fig. 3 consists of three types of building blocks (*UE*, two-port adaptors and *WDE_Tjunction* element).
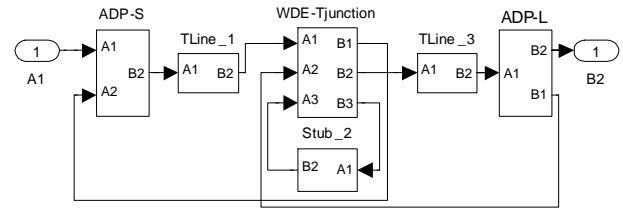

Fig. 3. Simulink model of the structure with discontinuties being modeled by block

As has been shown previously, T-junction equivalent network is approximated by inductors in the series branches and capacitor in the parallel branch. A starting point in realization of **WDE_Tjunction** element is the four-port network containing capacitor in parallel branch, Fig. 4.
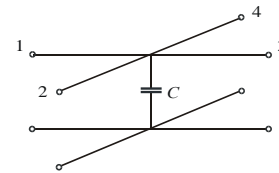
Fig. 4. A four-port C analog network
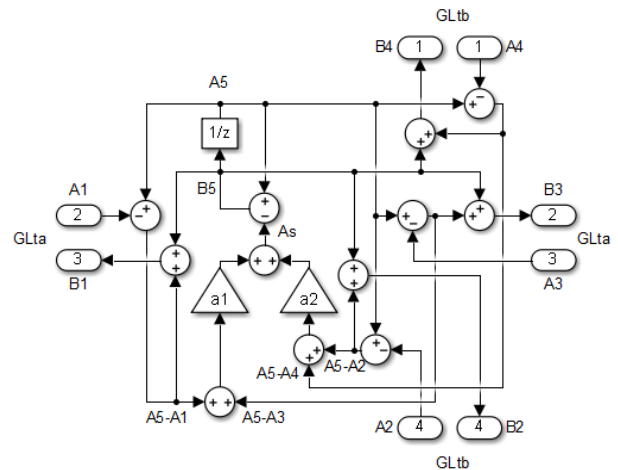
Fig. 5. Symmetrical four-port wave digital network of capacitor

The five-port parallel adaptor with dependent port five is used for modeling capacitor in parallel branch. In this way the four-port wave digital network of the capacitor in parallel branch (WC_Cen), shown in Fig. 5, is obtained [1-3]. The WC_Cen is symmetrical because of $L_1 = L_3$ and $L_2 = L_4$ for homogenous structures. This network corresponds to centrally placed segments. For this wave digital network, the equations for wave variables are

$$A_5 = z^{-1} \cdot B_5, \tag{3}$$

$$A_S = \alpha_1(2A_5 - A_1 - A_3) + \alpha_2(2A_5 - A_2 - A_4), \tag{4}$$

$$B_5 = A_5 - A_S, \tag{5}$$

$$B_k = B_5 + A_5 - A_k, \quad k = 1, 2, \cdots, 4 \tag{6}$$

where A_k and B_k are incident and reflected wave variables.

The multiplier coefficients can be expressed as follows

$$\alpha_1 = \frac{2 \cdot GL_{ta}}{2 \cdot GL_{ta} + 2 \cdot GL_{tb} + GC} \quad (7)$$

and

$$\alpha_2 = \frac{2 \cdot GL_{tb}}{2 \cdot GL_{ta} + 2 \cdot GL_{tb} + GC}, \quad (8)$$

where GL_{ta} and GL_{tb} are port conductances.

Fig. 6. *Simulink* model of **WDE_Tjunction** block
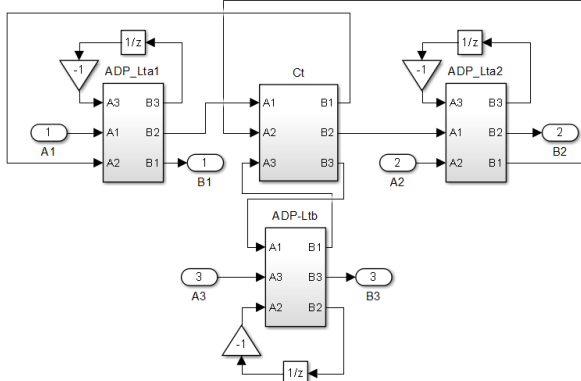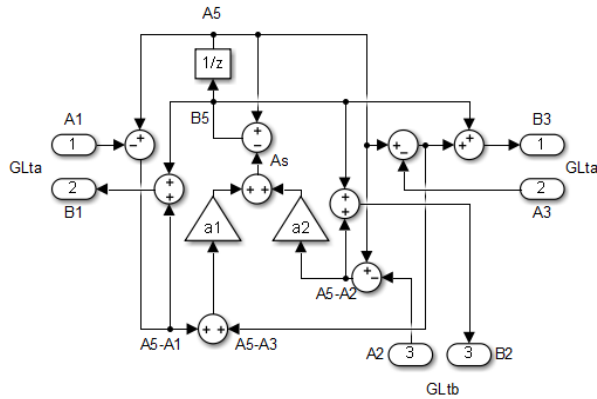

Fig. 7. *Simulink* WDN model of **WCt** block

The wave digital element of T-junction (**WDE_Tjunction** block) is shown in Fig. 6. It consists of one WCt element and three three-port series adaptors with wave digital elements of inductors (a single delay and a multiplier with coefficient -1) placed on their dependent ports. WDN of WCt element is given in Fig. 7. The equations for wave variables in z-domain are

$$B_k = \frac{(1 + z^{-1})[\alpha_1(A_1 + A_3) + \alpha_2 A_2]}{1 - z^{-1}[1 - 2\alpha_1 - \alpha_2]} - A_k, \quad k = 1,2,3, \quad (9)$$

and in t-domain are

$$b_k(n) = \alpha_1[a_1(n) + a_3(n) + a_1(n-1) + a_3(n-1)] +$$
$$+ \alpha_2[a_2(n) + a_2(n-1)] - a_k(n) + [1 - 2\alpha_1 - 2\alpha_2] \cdot \quad (10)$$
$$\cdot [a_k(n-1) + b_k(n-1)]$$

where $n = 0,1,2,\dots$, $a_k(-1) = b_k(-1) = 0$ and $k = 1,2,3$.

IV. T-RESONATOR CIRCUIT EXAMPLE

A microstrip structure with a single stub, depicted in Fig. *2a*, is used for verification of the proposed method. The structure is fabricated on CuFlon substrate with dielectric constant $\varepsilon_r = 2.17$ and width $h = 0.508\, mm$.

TABLE I
TRANSMISSION LINE PARAMETERS WITHOUT MODELED DISCONTINUITY

nv	d [mm]	w [mm]	Zc [Ohm]	Tv [ps]
1	30.0000	4.7100	21.9255	140.4001
2	30.0000	15.7600	7.5216	144.2831
3	30.0000	4.7100	21.9255	140.4001

TABLE II
TRANSMISSION LINE PARAMETERS WITH MODELED DISCONTINUITY

nv	d [mm]	w [mm]	Zc [Ohm]	Tv [ps]
1	22.1200	4.7100	21.9255	103.5217
2	27.9196	15.7600	7.5216	134.2776
3	22.1200	4.7100	21.9255	103.5217

T-resonator circuit is approximated by connection of 3 transmission lines with parameters given in Tables I and II. In this case, T-junction discontinuity is modeled by decreasing lengths of the line in the junction. The effect of the open stub is compensated by increasing length of the segment UTL2 [16].

A minimal number of sections for the given error, can be found as described in the papers [8] and [9] where multiple factor $q \geq 1$ is used. In the case where discontinuity is being modeled, for given error $n\_er = 0.01\%$, a total minimal number of sections in *WDN* is $n_t = \sum_{k=1}^{3} n_k = 122$. For segments UTL1 and UTL3, a number of sections n_k is 37, and for segment UTL2 is 48. A total delay for the digital model of the structure is $T_t = n_t \cdot T_{\min} / q = 341.3209\, ps$ where $q = 37$ is a multiple factor and $T_{\min} = \min\{T_1, T_2, T_3\} = 103.5217\, ps$ is a minimum delay. A total real delay of the structure is $T_\Sigma = \sum_{k=1}^{3} T_k = 341.3418\, ps$. A sampling frequency of the digital model of the planar structure for the chosen minimal number of sections is $F_s = n_t / T_t = 357.4130\, GHz$. In this case, a relative error of delay is $er = -0.0061\%$. Adaptor coefficients are $\alpha_S = -\alpha_L = 0.3903$ and $\alpha_1 = \alpha_2 = 0.4069$.

The S_{21} parameters of the structure at the frequencies from $300\, MHz$ to $6\, GHz$ are shown in Fig. 8. The transmission parameter was measured by a network analyser. A very good agreement between the simulated results was also obtained for many other structures of different shapes with T-junctions, which are not presented here. The measured result shows a small difference of the resonance frequency in comparison with the simulated result. This difference is caused by the microstrip T-junction and the open-end, which are not

described accurate enough by the models. This may be due to the closed form expressions which are used.

According to these findings, a simple 1D wave digital approach can produce results that are similar to the much more sophisticated methods. The simulation showed that the discontinuities of the microstrip T-resonator structure had a great effect on the resonance frequencies.
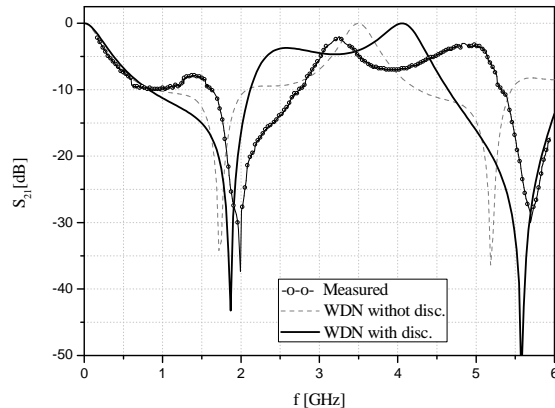

Fig. 8. The measured and simulated transmision parameters of the open T-resonator circuit

V. CONCLUSION

The primary purposes of this paper are twofold.

The first objective is to give an original and general method to characterize the behaviour of microstrip stub-line structures in both the frequency and time domains. A very simple block-diagram method of analysis of the WDN is used here. WDN is formed directly in the Simulink toolbox of the MATLAB environment. Frequency response is obtained by direct analysis of formed block-based wave digital network.

The second objective, but not the least, is the study of discontinuities associated with the T-pattern resonator. A simplified approach has been established to study the characterization of the microstrip discontinuities (open-end stubs and tee-junction). Two different ways of modeling T-junction discontinuity are given. Also, WDN model of the equivalent multiport circuit of T-junction is developed.

The broadband accuracy of the suggested procedure is validated by two examples realized in the microstrip line technique, such as an open T-resonator circuit. The results of the analysis obtained by the WDN have shown a very good agreement with those obtained by other programs mentioned above. The results of one T-resonator are compared with measurements performed by Vector Network Analyzer in the frequency range $300\,MHz\,-\,6\,GHz$.

The proposed approach can be used by microwave engineers because of the associated computational accuracy. It can be used for analyzing the transmission lines of various nonuniform shapes present in practice.

REFERENCES

[1] A. Fettweis, õDigital Circuits and Systemsö, *IEEE Transactions on Circuits and Systems*, Vol. CAS-31, No. 1, January, 1984, pp. 31-48.

[2] A. Fettweis, õWave Digital Filters: Theory and Practiceö, *Proc. IEEE*, Vol. 74, 1986, pp. 270-327.

[3] W. K. Chen, *The Circuits and Filters Handbook*, CRC Press, 1995 (Wave Digital Filters, pp. 2634-2661).

[4] M. V. Gmitrovi , *Microwave and Wave Digital Filters*, Faculty of Electronic Engineering, Ni–, 2007 (in Serbian).

[5] B. P. Sto–i , *Analysis of Planar Microwave Structures Modeled by Wave Digital Elements*, Doctoral thesis, Faculty of Electronic Engineering, University of Ni–, Ni–, September 2008 (in Serbian).

[6] M. V. Gmitrovi and B. P. Sto–i , õAnalysis of Planar Structures Modeled by Wave 1D Digital Elementsö, *14th Telecomm. forum - TELFOR 2006*, Serbia, Belgrade, November 21-23, 2006, pp. 418-421.

[7] B. P. Sto–i and M. V. Gmitrovi , õImplementation of Wave Digital Model in Analysis of Arbitrary Nonuniform Transmission Linesö, *Microwave and Optical Technology Letters*, Vol. 49, No. 9, September 2007, pp. 2150-2153.

[8] B. P. Sto–i and M. V. Gmitrovi , õEquivalent Thevenin Source Method as Tool for Response Computation of Wave Digital Structuresö, *8th Inter. Conference on Telecommunications in Modern Cable, Satellite and Broadcasting Services - TELSIKS 2007*, Serbia, Ni–, September 26-28, 2007, Vol. 1, pp. 203-206.

[9] B. P. Sto–i and M. V. Gmitrovi , öGenerating of Basic Wave Digital Elements for Modeling of Two-dimensional Planar structuresö, *XLII International Scientific Conference - ICEST 2007*, Macedonia, Ohrid, June 24-27, 2007, pp. 309-312.

[10] B. P. Sto–i and M. V. Gmitrovi , õModeling of Step Discontinuity in Microstrip Structures by using Wave Digital Approachö, *XLIII Intern. Scientific Conference on Information, Communication and Energy Systems and Technologies - ICEST 2008*, Serbia, Ni–, June 25-27, 2008, pp. 347-350.

[11] B. P. Sto–i and M. V. Gmitrovi , õWave Digital Approach ó A Theoretical Model of Step Discontinuityö, *16th Telecommunications forum - TELFOR 2008*, Serbia, Belgrade, November 25-27, 2008, pp. 539-542.

[12] B. P. Sto–i and M. V. Gmitrovi , õDirect Analysis of Wave Digital Network of Microstrip Structure with Step Discontinuitiesö, *The 7th WSEAS International on System Science and Simulation in Engineering – ICOSSSE'08*, Italy, Venice, November 21-23, 2008, pp. 25-29.

[13] B. P. Sto–i and M. V. Gmitrovi , õWave Digital Approach - A Different Procedures for Modeling of Microstrip Step Discontinuitiesö, *International Journal of Circuits, Systems and Signal Processing*, Vol. 2, No. 3, 2008, pp. 209-218.

[14] B.P. Sto–i and M.V. Gmitrovi , öBlock-based Analysis of Microstrip Structures with Stubs by use of 1D Wave Digital Approachö, *XLIV Intern. Scientific Conference, Information, Communication and Energy Systems and Technologies - ICEST 2009*, Bulgaria, Veliko Tarnovo, June 25-27, 2009, Vol. 1, pp. 23-26.

[15] MATLAB - The Language of Technical Computing, Version 7.6.0.324 (R2008a), February 10, 2008, The MathWorks Inc. 1984-2008.

[16] P. F. Combes, J. Graffeuil and J.-F. Sautereau, *Microwave Components, Devices and Active Circuits*, John Wiley & Sons, New York, 1987.

[17] B. P. Sto–i , õUsing *z*-variable Functions for the Analysis of Wave-based Model of Microstrip Stub-line Structureö, *Microwave Review*, Vol. 15, No. 2, December 2009, Serbia, pp. 6-11.

[18] Advanced Design Software 2002, Agilent Technologies 1983-2002, 395 Page Mill Road, Palo Alto, CA 94304, USA.

Circularly polarized 2×2 patch antenna array at 5 GHz

Sinisa Jovanovic, Sinisa Tasic and Predrag Manojlovic

Abstract – **This paper presents a simple design procedure for obtaining a uniplanar patch antenna array with circular polarization. The feeding network of the array is carefully designed for achieving very good matching at the main antenna port as well as for obtaining a precise signal equal amplitude dividing with the required phases for all the elements of the array. An experimental example for a 5 GHz central frequency is manufactured and measured for validating the proposed method.**

Keywords – **Patch antenna, circular polarization, printed antenna array, sequential feeding, uniplanar microstrip.**

I. INTRODUCTION

Circularly Polarized (CP) antennas are employed within various communications systems whose operating comprises position or orientation change of transmitting and receiving antennas such as aircraft and satellites communication, radar, remote control, telemetry, WLAN, WiMAX, and noninvasive microwave. Their usage eliminates the influence of transmitting/receiving antennas' mutual orientation on the receiving signal level and prevents the occurrence of excessive propagation loss due to cross polarization [1, 2].

Compared to antennas with linear polarization, CP antennas have a more complex set of requirements that have to be fulfilled during the design process. Among all conceivable practical realizations of CP antennas, printed CP antennas are particularly attractive due to their low mass, simplicity, reproducibility and simple manufacturing [3].

II. DESIGN THEORY

A. Basic Concept

Generally speaking, CP can be attained either by using CP radiating elements, or by using an array in which several linearly polarized elements form CP [4]. The design described in this paper combines both techniques to achieve better axial ratio (AR) purity over a broader frequency range [5-7]. In addition, and for the same reason, the two point fed patch is used rather than the perturbed single point fed patch. In that manner, a 2x2 array, composed of square shaped patches as shown in Fig.1, is proposed.

To achieve an ideal CP with AR=1 (0 dB) the signals at all eight ports (P1-P8) of the array should have equal amplitudes

and mutual phase shifts as indicated in Fig. 1. As long as this condition is fulfilled the AR value will be very close to 0 dB. However, with a realistic feeding network, the ideal CP could be achieved at a single frequency point only, while AR ≤ 3 dB, which is acceptable for practical use, could be achieved in a relatively narrow frequency range of a few percent.

B. Patch Element Initial Design

For a selected microstrip (MS) substrate (specified with ε_r and h) the optimal dimensions (W and L) of a rectangular patch having a resonance at frequency f_0 is given with the following well known equations:

$$W = \frac{c}{2f_0\sqrt{\frac{\varepsilon_R+1}{2}}} \tag{1}$$

$$L = \frac{c}{2f_0\sqrt{\varepsilon_{eff}}} - 0.824h\left(\frac{(\varepsilon_{eff}+0.3)(\frac{W}{h}+0.264)}{(\varepsilon_{eff}-0.258)(\frac{W}{h}+0.8)}\right) \tag{2}$$

$$\text{where: } \varepsilon_{eff} = \frac{\varepsilon_R+1}{2} + \frac{\varepsilon_R-1}{2\sqrt{1+12\frac{h}{W}}}$$

Equation (2) gives a very good initial size ($L \times L$) for a square-shaped two point fed patch with a relative error smaller than 2%, compared to results obtained by EM analysis, which can easily be corrected in second design iteration by scaling the array's overall dimensions. The array elements are separated by approximately $2\lambda_0/3$ along both axes.
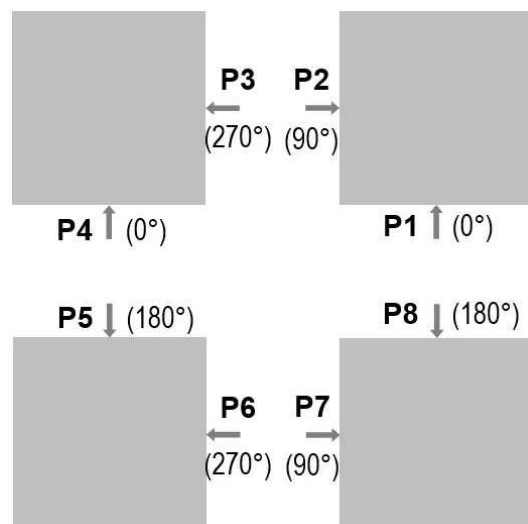

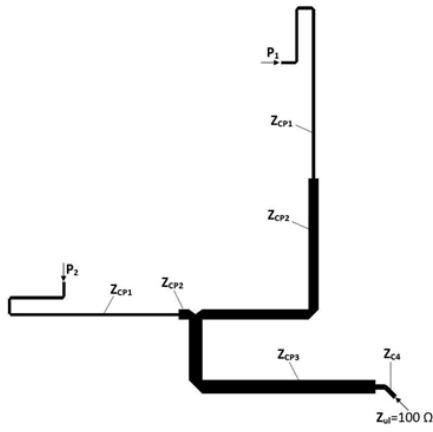Fig. 1. Principle diagram of a circularly polarized 2×2 patch array

Fig. 2. Patch Feeding Network

C. Patch Feeding Network Design

Every element of the array has an identical Patch Feeding Network (PFN) which is shown in Fig. 2. It provides two equal signals at the patch's ports (P1 and P2), with a phase difference of $\pi/2$. Another function of the PFN is to match the patch ports' input impedances (of about 220 Ω) to the impedances of the output ports of the Array Feeding Network (AFN), which is adopted to be $Z_{ul} = 100\,\Omega$. Fig. 2 shows one possible PFN configuration composed of microstrip lines (MSL) having characteristic impedances: $Z_{CP1} = 122\,\Omega$, $Z_{CP2} = 68\,\Omega$ and $Z_{CP3} = 58\,\Omega$, with the respective widths (in mm): $w_{CP1} = 0.12$, $w_{CP2} = 0.46$ and $w_{CP3} = 0.6$ for the selected MS substrate ($\varepsilon_r = 2.17$, $h = 0.254$ mm, $\tan\delta = 0.0009$ and $t = 0.017$ mm). The electrical length of MSL with Z_{CP1} and Z_{CP3} is $\pi/2$, which is also the difference of the longer and shorter MSL having Z_{CP2}.

D. Array Feeding Network Design

A layout of the AFN is shown in Fig.3. The feeding coaxial SMA connector is placed in the centre of the symmetry of the array, vertically to the MS ground layer. The central part of the AFN (shaded in grey) divides the input signal into four equal amplitude samples shifted successively by $\pi/2$. It consists of a short 50-ohm MSL followed by a cascade of $\lambda/4$ transformers
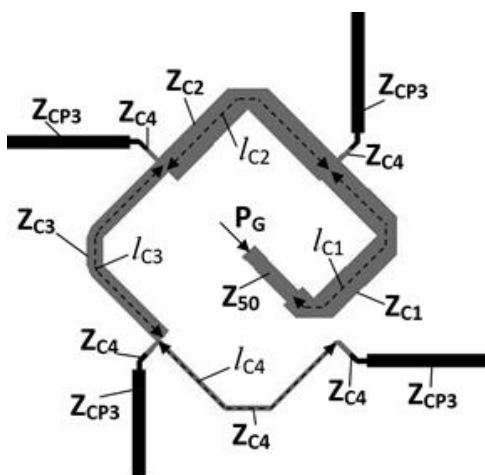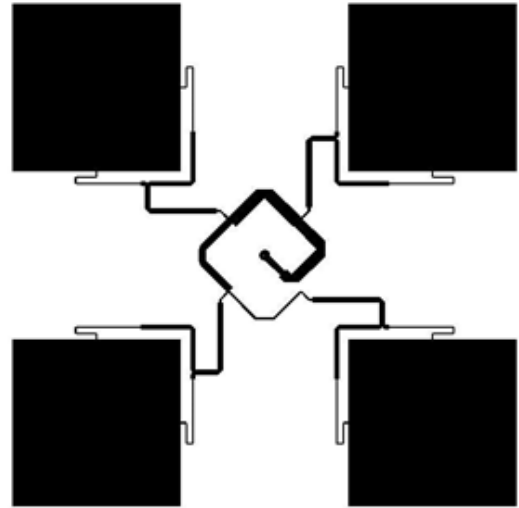

Fig. 3. Array feeding network

Fig. 4. Antenna Array Layout

with characteristic impedances: $Z_{C1} = 35.35\,\Omega$, $Z_{C1} = 33.33\,\Omega$, $Z_{C3} = 50\,\Omega$ and $Z_{C4} = 100\,\Omega$. The corresponding MSL widths (in mm) are $w_1 = 1.26$, $w_2 = 1.35$, $w_3 = 0.77$ and $w_4 = 0.21$. Each section of the cascade is followed by a short parallel MSL having $Z_{C4} = 100\,\Omega$, that connects AFN and one of the four PFNs. The MSL of AFN are meandered using an appropriate combination of 90° or 45° bends in order to accommodate MSL of different widths and slightly different lengths (l_{C1} to l_{C4}) into the available space between the patches as well as to precisely align the connections between AFN and four PFNs,. A complete layout of the proposed antenna array having right-handed (RH) CP is shown in Fig. 4, while its flipped version (either around x or y axes) corresponds to a left-handed (LH) CP antenna array. The layout is designed for a central frequency of 5 GHz which gives the patch dimension of 20×20 mm, patch spacing of 40 mm and overall array dimensions of 60×60 mm.

E. Simulation Results

The configuration from Fig. 4 is analysed with a full wave electromagnetic simulation (EMS) program. Fig. 5 shows the obtained results for the return loss (RL) at the input port of the array, which is better than 30 dB at a central frequency. RL better than 10 dB is obtained from 4.94 GHz to 5.07 GHz.
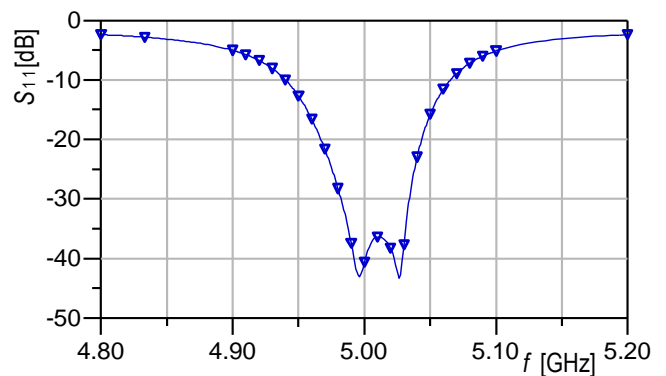

Fig. 5. Return loss at the input of antenna array (EMS results)
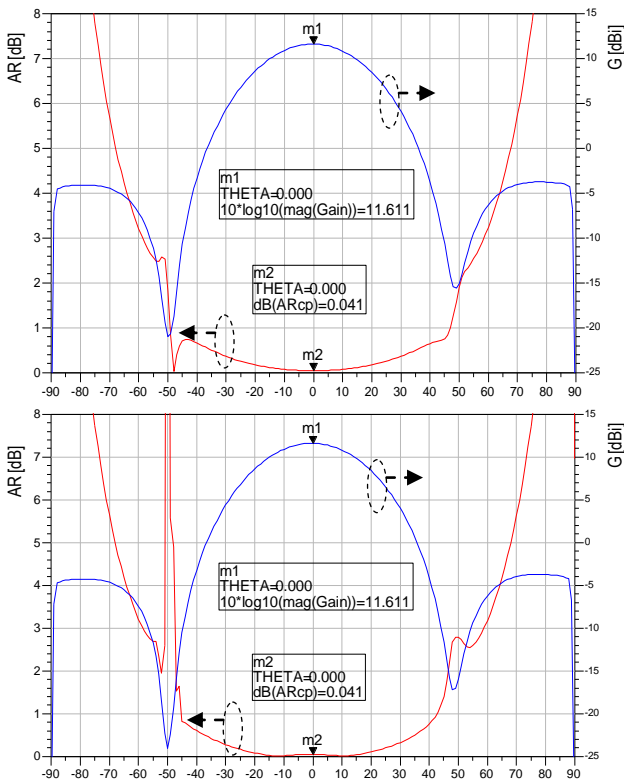
Fig. 6. Radiation pattern and AR in two orthogonal planes
at 5.01 GHz (EMS results)
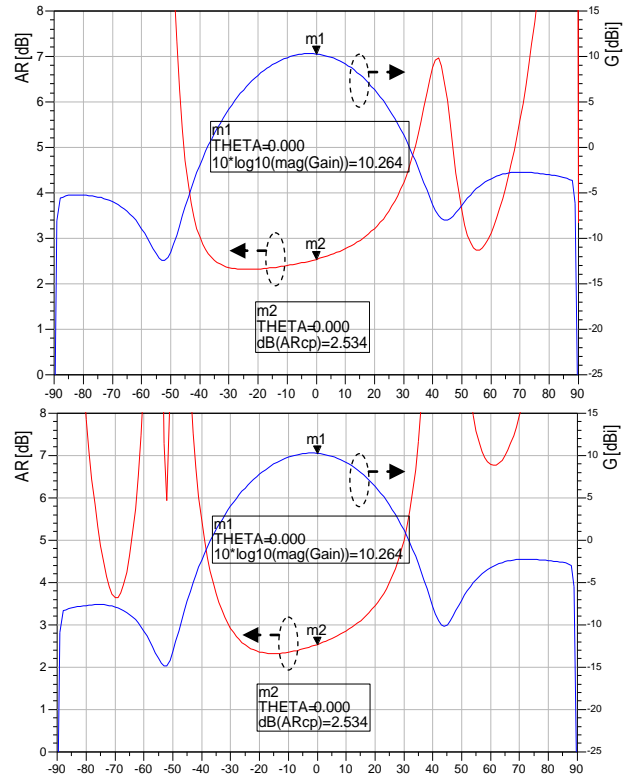
Fig. 8. Radiation pattern and AR in two orthogonal planes
at 5.05 GHz (EMS results)

Figs. 6 to 8 show the radiation pattern and AR of the antenna array from Fig. 4 in two mutually orthogonal planes that are also normal to the plane containing the array. Fig. 6 shows that the best results for both gain (G = 11.6 dBi) and AR ≈ 0 dB are obtained at a frequency of 5.01 GHz. At 4.98 GHz and 5.05 GHz, which frame the useful frequency range, the gain reduces to 10 dBi, while AR increases to 1.2 dB and 2.5 dB, respectively. The narrow range is caused by the selection of a relatively thin MS substrate that was required for MSLs with low Z_c within AFN. Radiation patterns at a central frequency have very good axial symmetry, which becomes slightly deformed toward framing working frequencies.
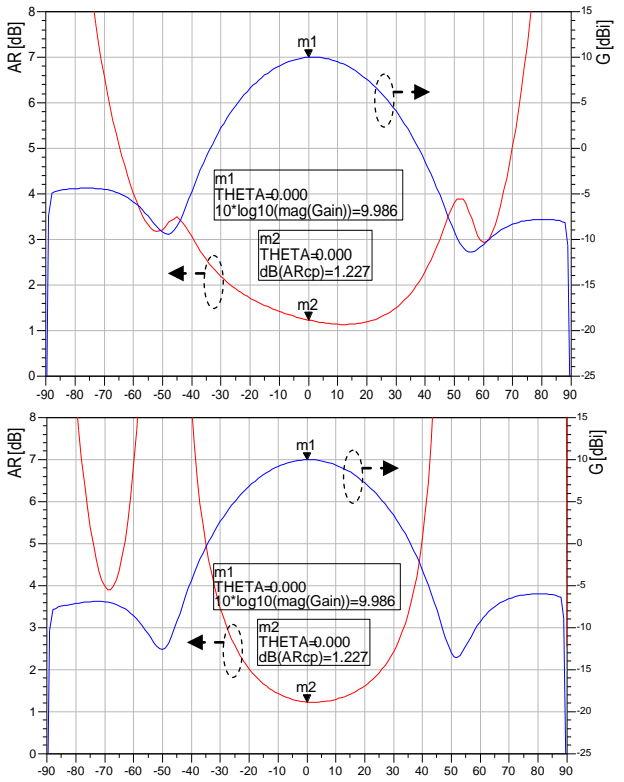
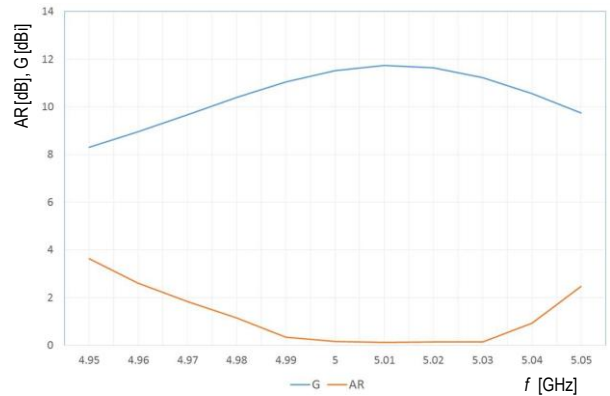Fig 9 shows the frequency characteristics of AR and the maximum gain and AR.

Fig. 7. Radiation pattern and AR in two orthogonal planes
at 4.98 GHz (EMS results)

Fig. 9. Gain and AR of antenna array (EMS results)

III. REALIZATION AND MEASURED RESULTS

As shown in Fig. 10, both the RH and LH versions of the antenna layout from Fig. 4 are fabricated using the standard lithographic procedure, for experimental validation of the proposed design. An identical pair of the RH version is realized for gain valuation by free space loss measurement, which showed a maximum gain of 10.5 dBi at 4.95 GHz,
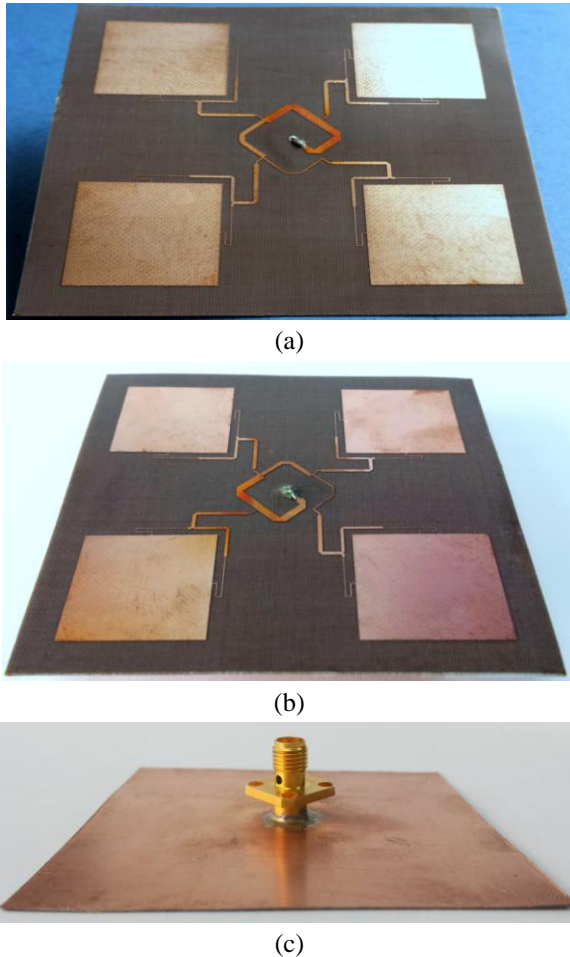

(a)

(b)

(c)

Fig. 10. Photo of realized 2×2 antenna array: (a) front side - RHCP; (b) front side - LHCP; (c) back side
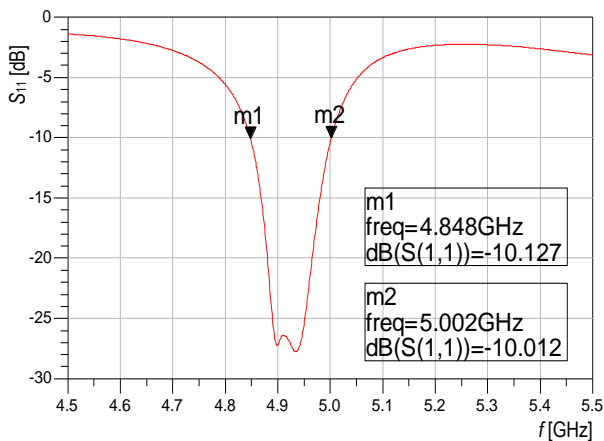

Fig. 11. Measured Return Loss of realized antenna array

while the LH antenna is used for cross-polarization isolation measurement, which was about 25 dB. AR is estimated from receiving signal level variation due to the rotation of one of the tested antennas, which showed the AR value of about 1dB in a 70 MHz wide range around the central frequency.

As shown in Fig. 11, the measured RL is better than 10 dB from 4.848 GHz to 5.002 GHz, whereas at the range centre, the RL is better than 25 dB. The measured frequency characteristics are shifted for about 1.5% toward lower frequencies relative to EM simulation results, which can be corrected in the next design iteration by suitable layout scaling.

IV. CONCLUSION

The presented solution enables the realization of circularly polarized antennas as a printed array realized in uniplanar microstrip technology. The parasitic radiation of the array feeding network is significantly mutually canceled due to its center-symmetric shape which prevents AR degradation. The proposed design and structure of the array feeding network, allows the achievement of excellent matching at the main input of the array with an RL better than 25 dB at the central working frequency. The selected values of λ/4 transformers' characteristic impedances allowed keeping the widths of the transmission lines within the available technological limits. The measured results are in very good agreement with the theoretical predictions with a slight frequency characteristics shift of 1.5% toward lower frequencies.

ACKNOWLEDGEMENT

REFERENCES

[1] G. Kumar, K.P. Ray: Broadband Microstrip Antennas, Artech House, Boston, 2003

[2] R. Garg, P. Bhartia, I. Bahl, A. Ittipiboon: Microstrip Antenna Design Handbook, Artech House, Boston, 2001.

[3] Nesic, A.; Jovanovic, S., "New printed slot antenna with circular polarization," Antennas and Propagation Society International Symposium, 1997. IEEE., 1997 Digest, vol.3, no., pp.1870,1873 vol.3, 13-18 July 1997

[4] Huang, J, "A technique for an array to generate circular polarization with linearly polarized elements," Antennas and Propagation, IEEE Transactions on , vol.34, no.9, pp.1113,1124, Sep 1986

[5] Tongbin Yu; Hongbin Li; Xinjian Zhong; Tao Yang; Weigang Zhu, "A wide bandwidth circularly polarized microstrip antenna array using sequentially rotated feeding technique," Antennas & Propagation (ISAP), 2013 Proceedings of the International Symposium on , vol.02, no., pp.743,745, 23-25 Oct. 2013.

[6] Kraft, U.R., "An experimental study on 2×2 sequential-rotation arrays with circularly polarized microstrip radiators," Antennas and Propagation, IEEE Transactions on, vol.45, no.10, pp.1459,1466, Oct 1997

[7] Aljibouri, B.; Sambell, A.J.; Sharif, B.S., "Application of genetic algorithm to design of sequentially rotated circularly polarized dual-feed microstrip patch antenna array," Electronics Letters , vol.44, no.12, pp.708,709, June 5 2008

An Experimental Setup for Switching Noise Measurement in Monolithic On-Chip Antennas

Hristomir Yordanov[1]

Abstract— **On-chip integrated antennas are of growing interest for the development of miniaturized sensors, as they allow for a reduction of the cost and dimensions of the autonomous systems. The amount of chip area dedicated to the antenna should be minimal. One way to achieve that is to share available chip metallisation, such as the ground supply plane, between the antenna and the on-chip digital circuitry. This can lead to interference between the circuit and the antenna. This paper presents an experimental setup for the measurement of this interference and for evaluation of the possible degradation of a wireless link between on-chip integrated antennas.**

Keywords— **Embedded antennas, Noise measurement, On-chip integrated systems.**

I. INTRODUCTION

The on-chip integrated antennas are an interesting alternative to wired chip-to-chip communications. They can provide high data rates, reduce system development time, and allow for system miniaturization, which is of great interest for the development of miniaturised sensors and sensor systems. Integrated antennas can be manufactured using standard or modified CMOS technology, where the modifications are in order to reduce the RF losses in the substrate [1], [2]. The CMOS technology allows for development of RF front end circuitry up to the millimetre wave band [3]. Therefore we can fabricate a single chip device, capable of acquisition, processing and wireless transmission of data.

The antenna integration should be chip area efficient in order to reduce the manufacturing costs. This can be achieved by sharing available CMOS metallisation between the digital circuitry and the antenna. For instance the CMOS ground plane could be shaped such that it also supports antenna mode currents [4]. This is a potential source for interference between the antenna and the digital circuitry. The interference goes both ways: the antenna currents can lead to an increased bit error rate of the digital circuit, and the transistor switching noises can induce RF power at the antenna terminals. In order to reduce the interference effects, the antenna operating frequency is designed in the millimetre wave range, well above the typical clock frequencies of the digital circuits. In this case the effect of the antenna currents on the digital transistors can be neglected due to their low-pass nature. This assumption has been experimentally proven by Bohorquez and O [5]. This paper is a step towards the experimental proof of the opposite effect. It provides a way to measure the RF power induced on the antenna terminals due to the switching currents of digital MOS transistors.

[1]Hristomir Yordanov is with the Electromagnetic Compatibility Lab at the Technical University of Sofia, Faculty of Telecommunications, Kliment Ohridski Blvd. 8, Sofia 1000, Bulgaria, E-mail: h_yordanov@tu-sofia.bg
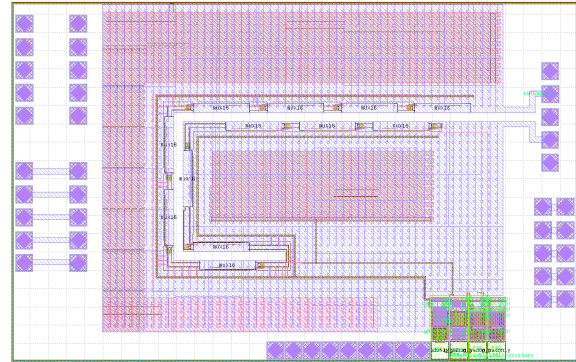
Fig. 1. Layout of the designed chip
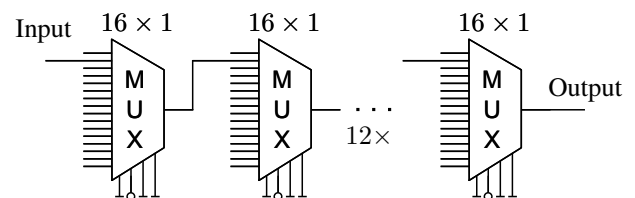
Fig. 2. A schematic of the implemented digital circuit. It consists of 12 multiplexors, hard-wired so that the input signal propagates to the output.

II. EXPERIMENTAL CHIP DESCRIPTION

The layout of the experimental integrated circuit is presented in Fig. 1. It consist of a L-shaped slot antenna, open-circuited at the port side, operating at 24 GHz. We have selected this frequency due to the limitations of the available $0.35\,\mu m$ CMOS technology, which does not allow transistor switching faster that 500 MHz. Therefore, in order to estimate the induced noise in a 60 GHz antenna due to a 1.3 GHz clock, we have designed a scaled prototype with the same ratio of clock to carrier frequency.

Full-wave electromagnetic models of the antenna show that the current density is highest along the antenna edges. This means that transistor switching currents generated in this area are expected to induce greatest RF power at the antenna terminals. Therefore the digital circuits which generate the interference are located only along the antenna edge. The circuitry consists of twelve 16-to-1 multiplexors, connected in series, as shown in Fig. 2. They are hard wired so that the input signal propagates through all of them and is routed to the output. The multiplexers consist of 270 transistors each. In the described configuration 32 transistors switch state and therefore draw current every time the input signal switches from high to low or from low to high, which is a representative

Fig. 3. Measurement setup for direct noise power measurement.

Fig. 4. Measurement setup for direct noise power measurement.

transistor usage efficiency for a typical digital circuit. The excitation is a series of alternating '1's and '0's with variable data rate. It is fed from the input pin to the upper-right multiplexer above the slot in Fig. 1, propagates through the rest 11 of them and exits from uppe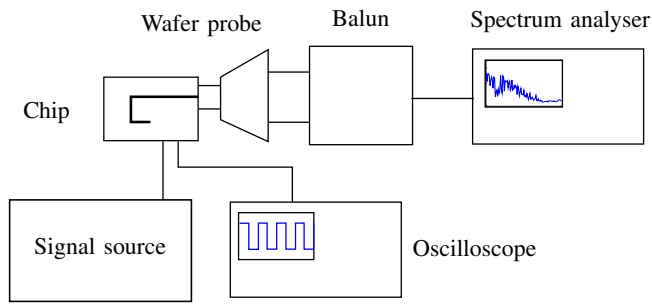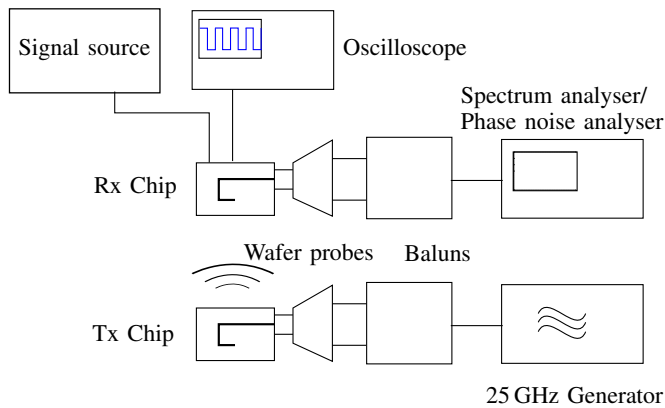r-right one below the slot, and is routed to the output pin below. The propagating signal is digital, with frequency up to 500 MHz.

The layout contains TLR calibration structures for measuring the antenna input impedance, which is required for noise power measurement calibration. The chip has 4 connected low-frequency pins, located on the bottom side, which are for Vcc, GND, Input and Output, and five RF pins on the right hand side, connected to the antenna.

III. MEASUREMENT SETUP

Two approaches for measuring the noise power are possible with the designed chip. First is the direct approach by connecting a spectrum analyser directly to the antenna port. This setup is presented in Fig. 3. The chip will be measured using a wafer probe, without bonding or housing. The signal source provides rectangular pulses with V_{ss} amplitude and variable frequency in the range $30 - 100$ MHz. The functionality of the digital part of the integrated circuit will be verified by monitoring the output using an oscilloscope. The on-chip antenna is connected to a spectrum analyser via a GSGSG wafer probe and a balun, coupling the symmetric antenna port to the unbalanced spectrum analyzer port. The balun has been realized using an in-house developed rat-race coupler.

The second approach is to use two on-chip antennas and transmit a CW signal over the link. We were looking for

changes of the input level and of the phase noise at the receiver side. The setup is presented in Fig. 4. We can also use this setup to evaluate the insertion loss of the channel and thus estimate the antenna gain and efficiency.

IV. EXPECTED RESULTS

The switching current due to the digital transistors is periodic, because the input signal, fed to the multiplexers is also periodic. Therefore it can be expected, that the induced RF power on the antenna terminals will have its peak power at frequencies close to the frequency of the digital signal. On the other hand, the length of the antenna is comparable to the wavelength of the wireless signal. This suggests that the currents due to each transistor along the antenna edge should be seen from the antenna terminal with a different phase. This suggests that we can expect a $1/f$ type of noise, which typically arises in similar scenarios. The experiment should confirm the assumption for $1/f$ noise and give an estimate on its rate of decay in frequency.

Additional information that can be obtained from the experiment is the evaluation of the antenna efficiency. It is expected to be low, because the standard CMOS substrate has very high losses in the RF frequency ranges, and no measures could be taken to reduce that losses, as standard CMOS technique is used for the chip manufacture. Furthermore, the insertion loss of a link between two antennas can be measured and an estimation of the capacity of such a link can be made.

V. SUMMARY

This paper presents an experimental set-up for measurement of the noise, induced in an on-chip integrated antenna due to the switching currents of the digital transistors, available on the same chip. The problem arises due to the high space coupling of the antenna and the digital circuitry, arising from the fact that the antenna uses the CMOS ground supply plane metallisation. The experimental chip is described, along with the proposed measurement setup, and the expected results have been described.

REFERENCES

[1] H. Yordanov and E. Angelopoulos, "On-chip integrated antennas on ultra-thin and on high-impedance Si substrate," in *Proceedings of the 33*rd *European Microwave Conference 2013*, Nuremberg, Germany, 2013, pp. 52–55.

[2] ——, "High efficiency integrated antennas on ultra-thin Si substrate," in *Proceedings of the IEEE Antennas and Propagation Society International Symposium, 2013*, Orlando, Florida, 2013.

[3] T. Yao, M. Q. Gordon, K. K. W. Tang, K. H. K. Yau, M.-T. Yang, P. Schvan, and S. P. Voinigescu, "Algorithmic design of CMOS LNAs and PAs for 60 GHz radio," *IEEE J. Solid-State Circuits*, vol. 42, no. 5, pp. 1044–1057, May 2007.

[4] H. Yordanov and P. Russer, "Area-efficient integrated antennas for inter-chip communication," in *Proceedings of the 30*th *European Microwave Conference 2010*, Paris, France, 2010.

[5] J. L. Bohorquez and K. K. O, "A study of the effect of microwave electromagnetic radiation on dynamic random access memory operation," in *Proc. 2004 IEEE Int. Symp. EMC, San Hose, CA*, vol. 3, Aug. 2004, pp. 815–819.

Inverse Electro-Mechanical ANN Model of RF MEMS Capacitive Switches - Applicability Evaluation

Zlatica Marinković[1], Ana Aleksić[1], Tomislav Ćirić[1],
Olivera Pronić-Rančić[1], Vera Marković[1], Larissa Vietzorreck[2]

Abstract – **The aim of this paper is to analyze applicability of the inverse electro-mechanical models of capacitive RF MEMS switches based on artificial neural networks (ANN). The analyzed model, proposed earlier, is developed to predict length of the fingered part of the switch for fixed length of the solid part of the bridge, resonant frequency and actuation voltage, directly without any optimization. Here, it is analyzed in which parts of the input space the model is applicable, and how to choose a physically meaningful combination of input values. Moreover, application of this model to determine bridge dimensions for the requested total length of the bridge is investigated as well. Finally, recommendations related to the application of the model are given.**

Keywords –**Artificial neural networks, RF MEMS, capacitive switch, inverse modeling, actuation voltage, resonant frequency.**

I. INTRODUCTION

In the recent years RF MEMS switches have become a competitive devices to their mechanical and electronic counterparts, as they have numerous advantages: they are very small, extremely linear, can be integrated and allow easy re-configurability or tunability of a system [1]. Due to the increasing application of RF MEMS switches in modern communication systems, there is also an increasing need for their accurate and efficient models. Standard simulations in commercial electromagnetic and mechanical simulators take a lot of time, making a design process time consuming [2]-[3]. To speed up the design process, alternative models based on artificial neural networks (ANNs) [4] can be used. Most of the developed neural models relate the switch electrical and mechanical characteristics and switch dimensions and operating conditions such as frequency and/or actuation voltage [5]-[10]. Since, the ANNs give response almost immediately, the simulation and optimization time is significantly reduced when these models are used. Moreover, inverse neural models aimed to determine the switch bridge dimensions for given requirements in electrical or mechanical domain were proposed [9]-[10]. Recently, authors of this paper proposed an inverse electro-mechanical model for determination of the switch dimensions for given electrical

[1] Z. Marinković, A. Aleksić, T. Ćirić, O. Pronić-Rančić, and V. Marković are with the University of Niš, Faculty of Electronic Engineering, Aleksandra Medvedeva 14, 18000 Niš, Serbia

E-mail: zlatica.marinkovic@elfak.ni.ac.rs, ana.aleksic@elfak.rs cirict@live.com, olivera.pronic@elfak.ni.ac.rs, vera.markovic@elfak.ni.ac.rs

[2] L. Vietzorreck is with the TU München, - Lehrstuhl für Hochfrequenztechnik, Arcisstr. 21, 80333 München, Germany, E-mail: vietzorreck@tum.de.

resonant frequency and actuation voltage [11]. It was shown that this model exhibits a very good accuracy of dimension determination. This paper deals with evaluation of the applicability of the mentioned model, with the main aim to determine the parts of the input space where physically meaningful output values are obtained.

The paper is organized as follows: after Introduction, in Section II the capacitive RF MEMS switch considered in this work is described. The analyzed inverse electro-mechanical ANN model is shortly described in Section III. The results of the analysis and discussion are presented in Section IV. Section V contains concluding remarks.

II. MODELED DEVICE

The considered device is an RF MEMS capacitive coplanar shunt switch, depicted in Fig. 1, fabricated at FBK in Trento in an 8 layer Silicon micromachining process [12].
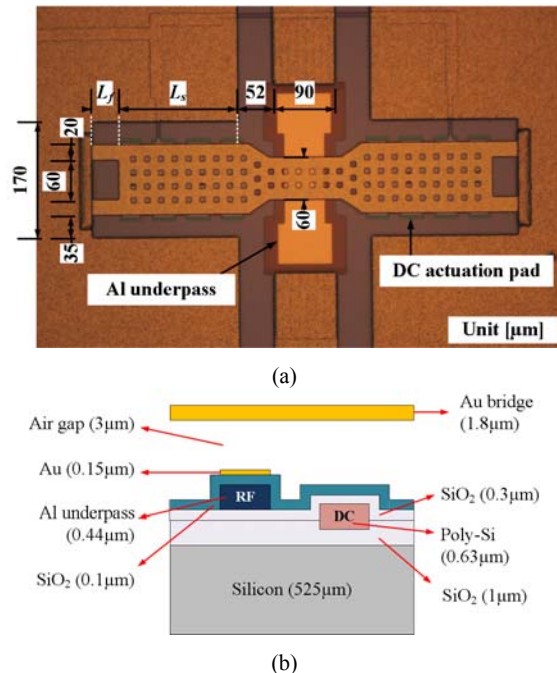

(a)

(b)

Fig. 1. a) Top-view of the realized switch; b) schematic of the cross-section with 8 layers in FBK technology [12]

The actuation voltage is determined as the instant voltage applied to the DC pads when the bridge comes down and touches a coplanar waveguide centerline, which is a pull-in voltage (V_{PI}). This is strongly related to the switch features and mechanical/material properties, such as a DC pad size and location, a bridge spring constant and residual stress, bridge shapes or supports, etc. The finger parts (corresponding to L_f in Fig. 1) are to control V_{PI}. If finger parts are long compared to the other parts, the bridge becomes flexible and the switch is easily actuated by a low V_{PI}. But this increases the risk of a self-actuation or an RF hold-down when the switch delivers a high RF power. And opposite, with the short finger parts, the switch needs a high V_{PI} to be actuated. Moreover, the switch resonant frequency, determining the switch operating frequency, depends on the bridge dimensions as well. Therefore, the lengths of the bridge solid and fingered parts (L_s and L_f) should be carefully determined considering a delivering RF power and a feasible DC voltage supply [1].

III. INVERSE ELECTRO-MECHANICAL ANN MODEL

In [11] an inverse electromechal model based on ANNs was poposed. It consists of an ANN traned to determine the length of the bridge fingered part (L_f) for given length of the bridge solid part (L_s) and desired resonant frequency (f_{res}) and actuation voltage (V_{PI}). ANNs are trained by using values of the resonant frequency and actuation voltage calculated in standard electromagnetic and mechanical simulators, or by using corresponding ANN models [10]. Once the ANN is trained properly, the length of the fingered part can be determined instantaneously, saving a lot of time comparing to the standard approach based on the optimizations in electromagnetic and mechanical simulators.
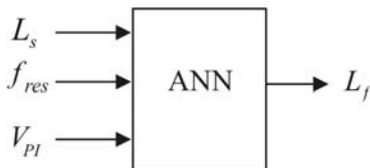
Fig. 2. RF MEMS ANN inverse electro-mechanical model.

IV. RESULTS AND DISCUSSION

The inverse electromechanical model developed for the RF MEMS switch described in Section II is an ANN having two hidden layers, containing 10 and 20 neurons. The details on model development and validation can be found in [11].
The model was developed for the following ranges of the switch geometrical parameters: L_s from 100 µm to 500 µm, and L_f from 0 µm to 100 µm. The corresponding ranges of resonant frequency and actuation voltage are approximately
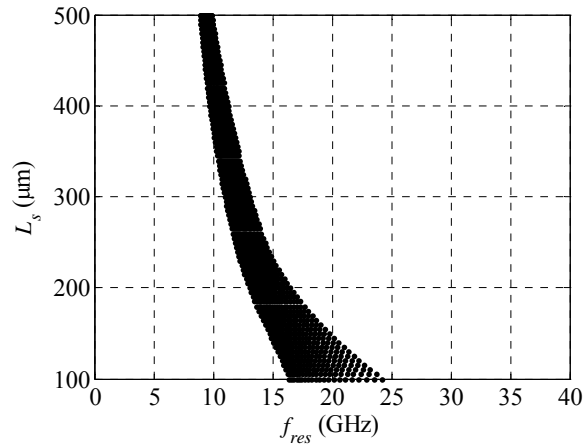
(a)
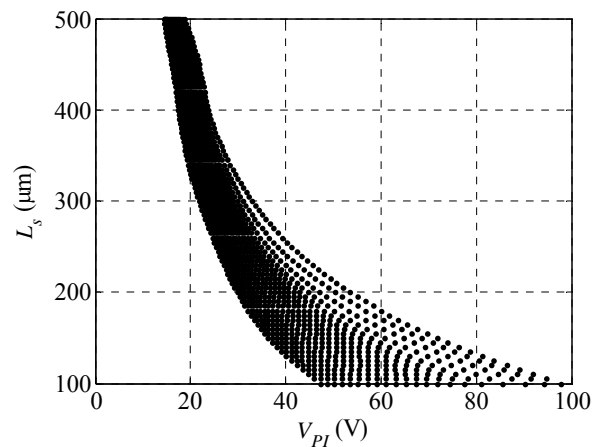
(b)

Fig. 3. Bridge solid part versus (a) resonant frequency; (b) actuation voltage.
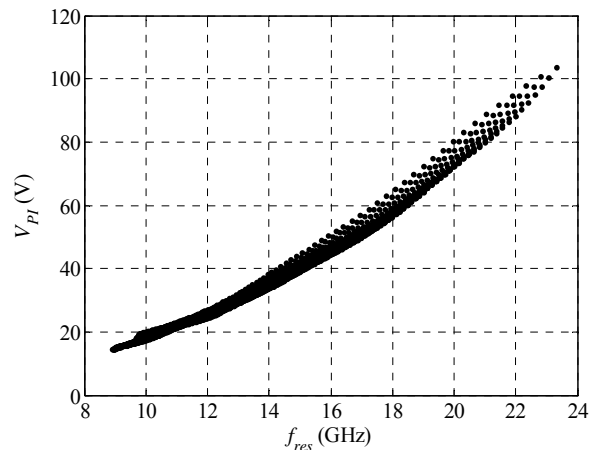
Fig. 4. Possible combinations of the actuation voltage and the resonant frequency for the considered range of bridge dimensions.

(8-24) GHz and (10-105) V, respectively. However, although the ANN model would give a response for every combination of values falling within the mentioned ranges, not every combination is physically meaningful, i.e. the model is not applicable in such cases. Therefore, here it will be analyzed in details which combinations of the input values are physically meaningful.

With that aim, the resonant frequency and actuation voltage have been calculated for different combinations of L_s and L_f in the above mentioned ranges by using the earlier developed neural models developed to determine resonant frequency [9] and actuation voltage [10] for the given bridge dimensions. Further, the bridge solid part is plotted versus the obtained f_{res} and V_{PI}, as shown in Fig. 3.

Moreover, in Fig. 4 there is a plot relating possible combinations of f_{res} and V_{PI} for the considered ranges of the bridge dimensions. In other words, these three plots represent parts of the input space referring to the physically meaningful input combinations. Therefore, the first step in determining the dimensions is to check if the input combinations are valid. This will be illustrated on the following example. Let the desired resonant frequency be 12 GHz. From Fig. 4, it can be seen that the corresponding actuation voltage values are from 24 to 28 V. The possible values of L_s for the resonant frequency of 12 GHz are from 250 µm to 350 µm, according to Fig. 3a. This range has to be compared to the range of possible values for a chosen V_{PI} values falling in the mentioned range from 24 to 28 V. Let the V_{PI} be 25 V. According to Fig. 3b, the corresponding L_s values are from 270 µm to 400 µm. Therefore, for the resonant frequency 12 GHz and actuation voltage 25 V, the possbile L_s values are in the range from 270 µm to 350 µm.

However, often the bridge part lengths should be determined for the fixed total length of the bridge ($L_t = L_s + L_f$). This model can be used to determine the dimensions also in such a case. As in the previous case, a desired resonant frequency can be achieved only for certain values of L_t, within the necessary range of the actuation voltage. By using the same data used for making plots shown in Fig. 3, the corresponding L_t is plotted versus f_{res} and V_{PI} and presented in Fig. 5. These plots are useful to estimate the possible values of L_t. As an example, with the total length of 300 µm it is possible to achieve only frequences between, approximately, 12 and 13 GHz, with the corresponding voltages in a range around 30 V. Also, the possible combinations of f_{res} and V_{PI}, plotted in Fig. 4, should be kept in mind. Once the total length is verified to be valid for the desired resonant frequency and the corresponding necessary actuation voltage, L_s and L_f can be easily found graphically, determined by the intercept point of the functions $L_f = f_{inv\_ANN}(L_s, f_{res}, V_{PI})$ obtained by the inverse electromechanical model and $L_f = L_t - L_s$, both for the L_s from 0 µm to L_t. This will be illustrated on the following example. Let the desired resonant frequency be 14 GHz, the actuation voltage 35V and the desired total lenght 260 µm. For L_s from 0 µm to 260 µm,

L_f is calculated and plotted in Fig. 6a as a black line. $L_f = L_t - L_s$ is a linear function plotted as a blue line. There are two intercept points, one around L_s =110 µm and the other at L_s =200 µm. Considering the plot shown in Fig. 3a, values around 100 µm are not physically possible for the resonant frequency of 14 GHz, therefore this intercept point is not a possible solution. The plot from Fig. 6a is zoomed around the second intercept point and shown in Fig. 6b, i.e., Fig. 6b shows the physically possible values which are represented in Fig. 6a as thicker parts of plotted lines. It can be clearly seen that the corresponding dimensions are L_s =200 µm and L_f =60 µm.

(a)

(b)
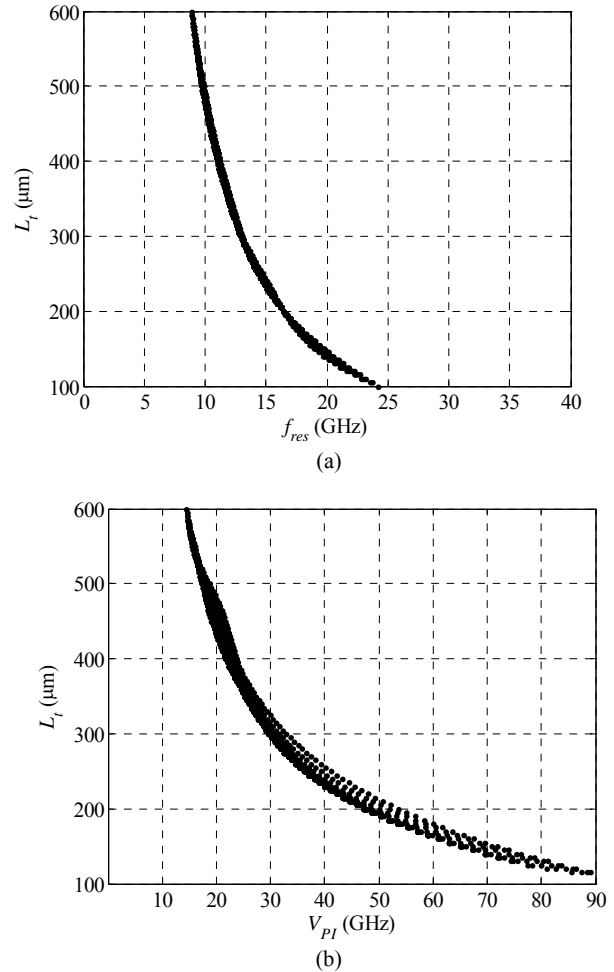
Fig. 5. Bridge total length versus (a) resonant frequency and (b) actuation voltage in the considered range of the bridge dimensions.

V. CONCLUSION

In this paper, an extensive analysis of the applicability of the earlier proposed electromechanical inverse model of RF MEMS switches aimed at calculation the length of the fingered part of the bridge has been done. Although the model is developed for a certain range of input space, not all

combinations of the input variables result in physically possible and meaningful values. Therefore, by using the previously developed neural models for determining the resonant frequency and actuation voltage, the plots defining the possible input combinations have been made. On an appropriate example, it was illustrated how to chose a valid input combination. Moreover, as sometimes it is requested to determine the lengths of the bridge part for the total bridge length given, the similar analysis has been conducted. In addition, a graphical approach based on using the considered neural model for obtaining the lengths efficiently for the given total length has been developed. This study can serve as a useful guide in a design of RF MEMS switches based on the switch inverse electro-mechanical neural models.
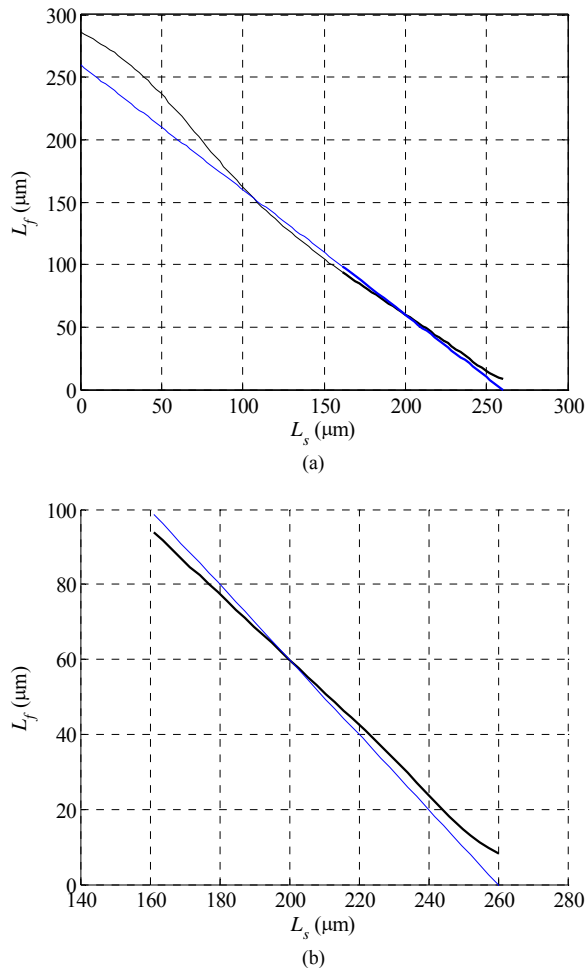

(a)

(b)

Fig. 6. L_s and L_f determination for the total length L_t =260 μm.

ACKNOWLEDGEMENT

REFERENCES

[1] G. M. Rebeiz, *RF MEMS Theory, Design, and Technology*. New York: Wiley, 2003.

[2] L. Vietzorreck, "EM Modeling of RF MEMS," 7th International Conference on Thermal, Mechanical and Multiphysics Simulation and Experiments in Micro-Electronics and Micro-Systems, EuroSime 2006.,Como, Italy, pp. 1- 4, 2006.

[3] R. Marcelli, A. Lucibello, G. De Angelis, E. Proietti, "Mechanical modelling of capacitive RF MEMS shunt switches," Symposium on Test, Integration & Packaging of MEMS/MOEMS 2009, MEMS/MOEMS '09, Rome, Italy, pp. 19 – 22, 2009.

[4] Q. J. Zhang and K. C. Gupta, *Neural Networks for RF and Microwave Design*. Boston, MA: Artech House, 2000.

[5] Y. Lee, D. S. Filipovic, "Combined full-wave/ANN based modelling of MEMS switches for RF and microwave applications," Proc. of IEEE Antennas and Propagation Society International Symposium, Washington, CD, USA, vol. 1A, pp. 85-88, 2005.

[6] Y. Mafinejad, A. Z. Kouzani, K. Mafinezhad , "Determining RF MEMS switch parameter by neural networks," *Proc. of IEEE Region 10 Conference TENCON 2009*, Singapore, pp. 1-5, 2009..

[7] Y. Lee, Y Park, F. Niu, D. Filipovic, "Artificial neural network based macromodeling approach for two-port RF MEMS resonating structures," *IEEE Proceedings of Networking, Sensing and Control*, March 2005, pp. 261 – 266.

[8] Y. Gong, F. Zhao, H. Xin, J. Lin, Q. Bai, "Simulation and Optimal Design for RF MEMS Cantilevered Beam Switch," *Proc. of International Conference on Future Computer and Communication* (FCC '09), Wuhan, China, pp. 84-87, 2009.

[9] T. Kim, Z. Marinković, V. Marković, M. Milijić, O. Pronić-Rančić, L. Vietzorreck, "Efficient Modelling of an RF MEMS Capacitive Shunt Switch with Artificial Neural Networks," *Proc. of URSI-B 2013 International Symposium on Electromagnetic Theory*, Hiroshima, Japan, pp. 550-553, 2013.

[10] Z. Marinković, T. Ćirić, T. Kim, L. Vietzorreck, O. Pronić-Rančić, M. Milijić, V. Marković, "ANN Based Inverse Modeling of RF MEMS Capacitive Switches", *11th Conference on Telecommunications in Modern Satellite, Cable and Broadcasting Services TELSIKS 2013*, Niš, Serbia, pp. 366-369, 2013.

[11] T. Ćirić, Z. Marinković, T. Kim, L. Vietzorreck, O. Pronić-Rančić, M. Milijić, V. Marković, „ANN based inverse electro-mechanical modeling of RF MEMS capacitive switches," *XLIX Scientific Conference on Information, Communication and Energy Systems and Technologies - ICEST 2014*, Niš, Serbia, vol. 2, pp. 127-130, 2014.

[12] S. DiNardo, P. Farinelli, F. Giacomozzi, G. Mannocchi, R. Marcelli , B. Margesin, P. Mezzanotte, V. Mulloni, P. Russer, R. Sorrentino, F. Vitulli, L. Vietzorreck, "Broadband RF-MEMS based SPDT", *Proc. European Microwave Conference 2006, Manchester*, UK, pp. 1727 – 1730, 2006.

ORAL SESSION
SIGNAL PROCESSING

Interactive design of digital filter using LabVIEW

Valentina Markova [1]

Abstract – **LabVIEW is powerful and flexible programming language for data acquisition, analysis and presentation of results. This paper introduces LabVIEW as intuitive tools for design, analysis and implementation of digital filters. It allows students and researchers to design and analyse DSP systems easily and for a shorter time, compared to Matlab. The aim of this paper is to present a set of VIs for design of digital filters with different interactive design options.**

Keywords – **Digital filter, LabVIEW, Virtual instruments.**

I. INTRODUCTION

LabVIEW is a graphical programming language, which has become prevalent among researchers and engineers in the industry during the last years. LabVIEW programs are called virtual instruments (VIs) and have two parts - a front panel and a block diagram. The front panel is the user interface. The block diagram contains the program code [1]. In case of filter design the VI reads the desired parameters of the filters entered by the user on the front panel and determines its characteristics or filter coefficients [2].

The NI Digital Filter Design Toolkit (DFDT) extends LabVIEW with tools for design, analysis and implementation of variety of IIR and FIR filters. It contains very useful teaching tools, which allow immediate feedback after specifying the desired filter requirements. The toolkit also includes a set of Filter Analysis VIs for evaluation of all filter characteristics – frequency and phase response, impulse and step response, group delay and zero/pole placement [3]. LabVIEW follows a dataflow execution model, which makes the learning process easier. Working within LabVIEW environment has many advantages, such as the ability to perform filter testing with simulated signals or real signals, acquired by real data acquisition device or sound card.

The DFDT includes a comprehensive set of design algorithms ranging from classical (Butterworth, Chebyshev, window, etc.) to modern optimizing options, including Remez exchange and least Pth norm [4].The Special Filter Design VIs helps for designing special digital filters like IIR notch filters, IIR comb filters, maximally flat filters, narrowband filters and delay compensators. The toolkit also includes the Multirate Filter Design VIs appropriate for design, analysis and implementation of single-stage and multistage multirate filters.

The aim of this paper is to present a set of VIs for design of digital filters with different interactive design options. The

[1]Valentina Markova is with the Faculty of Electronics, Department of Telecommunications at Technical University of Varna, 1 Studentska str., Varna 9010, Bulgaria.
E-mail: via@tu-varna.bg.

effect of the repeated interactive design changes is that the designer has the possibility to observe how different design parameters affect the filter performance.

II. DESIGN OF FIR FILTERS WITH LABVIEW

Finite impulse response (FIR) filters, also known as non-recursive filters operate only on the current and the past input values of the signal. There are some advantages in using FIR filter, since it can be designed with exact linear phase and the filter structure is always stable with quantized filter coefficients [5]. The price for this linearity of the phase response of the filter is the higher computational requirements.

FIR filters perform a convolution of the filter coefficients with a sequence of input values and produce an equally numbered sequence of output values, as shown in Eq.1:

$$y_i = \sum_{k=0}^{n-1} h_k x_{i-k} \qquad (1)$$

where x is the input sequence, y – the output sequence, and h is the FIR coefficients.

The DFDT facilitates the design of the digital filter by specifying its parameters, as well as by offering numerous display options. The design of a FIR filter is based on a direct approximation of the specified magnitude response. The toolkit includes a variety of VIs and functions for the design of FIR filters. The linear phase FIR filters can be designed using either Remez exchange method or windowed Fourier series. An example of FIR windowed design is shown on Fig.1.
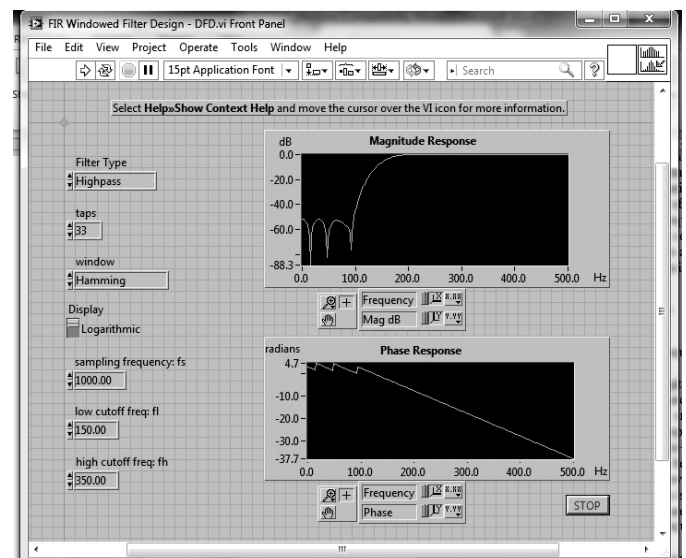

Fig.1. FIR Windowed Filter Design VI

Using FIR Windowed Filter Design VI, the design of the filter can be specified by entering the filter specifications (sampling frequency – fs, low cutoff frequency – fl, high cutoff frequency – fh and window function).

The design results are displayed immediately in the shape of magnitude and phase response. The filter information also contains the order of the filter. The DFDT supports variety of smoothing windows (Hanning, Hamming, Triangular, Kaiser-Bessel and so on).

The other VI - FIR Windowed coefficients VI, generates the set of filter coefficients.

The second design method - object of the current manuscript is the equiripple linear phase FIR design, using Remez exchange algorithm. Using DFD Remez Design VI (Fig. 2), it is possible to design a variety of filter types by adjusting the filter specifications by changing the passband and stopband parameters or the number of the taps. Equiripple design equally weights the passband and stopband ripple and produces filters with a linear phase characteristic.
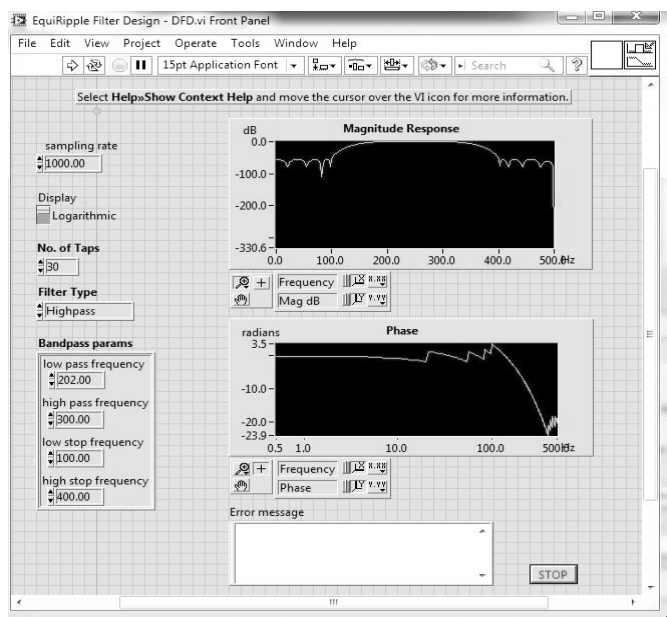
Fig.2. EquiRipple FIR Filter Design VI

The block diagram of the described VI is presented in Fig.3.
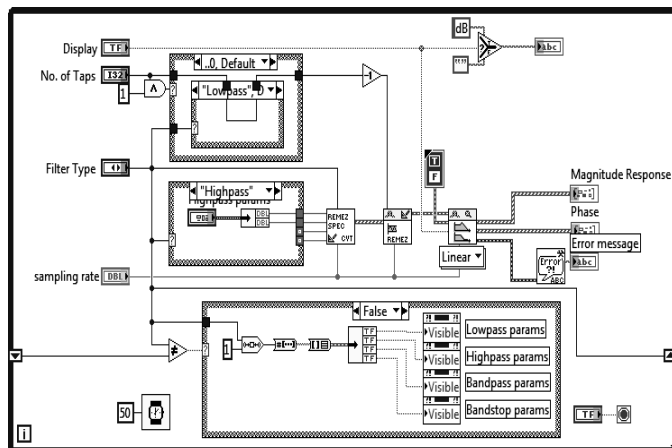
Fig.3. The screenshot of the block diagram of EquiRipple FIR Filter Design VI

III. DESIGN OF IIR FILTERS WITH LABVIEW

Infinite impulse response (IIR) filters, also known as recursive filters, operate on the current and the past input values as well as the current and the past output values of the signal. The IIR filters can achieve the same level of attenuation as the FIR filters, but with less number of coefficients. Therefore, an IIR filter can provide a significantly faster and more efficient filtering than an FIR filter.

DFDT offers a big variety of design options. In the first group VIs the design of the filter can be performed by entering the filter specifications. The filter characteristics or its coefficients are directly available in forms defined by used VI. There is also an option for designing of a filter by entering directly the transfer function by the user in the front panel. The final design option includes defining of the filter characteristics, by placing the poles and zeros of the transfer function in the pole zero diagram.

Express VIs are special tools in DFDT used for quick creation of digital filters by interactively specifying the filter parameters.

A. Design of IIR filter by specifications

The most common design methods are based on classical analogue filter approximations –Bessel, Butterworth, Chebishev, Inv. Chebishev and elliptic.

Using IIR Digital Filter VI (Fig.4), the designer can design a variety of filter types (Lowpass, Highpass, Bandpass, Stopband) by changing the design method, or by adjusting the filter specifications, by entering different values of the input parameters in the Front Panel of the VI. The available design methods are Elliptic, Butterworth, Bessel [6], The filter order is estimated automatically. The execution of the VI returns the magnitude and phase responses of the filter.
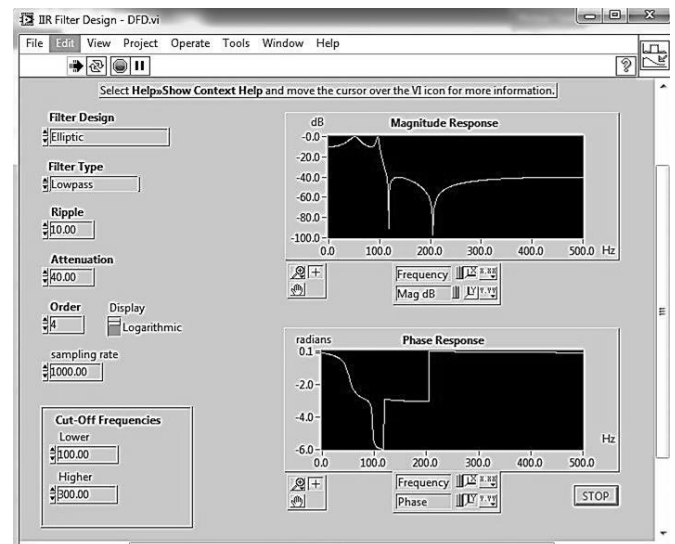
Fig.4. The screenshot of the Front panel of IIR Filter Design VI

B. Design of IIR filter from transfer function

DFDT also has VI for creating filters by entering directly the filter coefficients, as presented in Fig.5.

Fig.5. The screenshot of the Front panel of IIR Filter Design VI

Fig.6. The screenshot of the Front panel of IIR Filter Design VI

Running this VI returns the magnitude response and the used structure for realization of the filter. The design procedure in Fig.6 is following:

1. All the values of the coefficients in the numerator and the denominator of the filter are entered by user in two arrays in the Case Structure [nihepl].
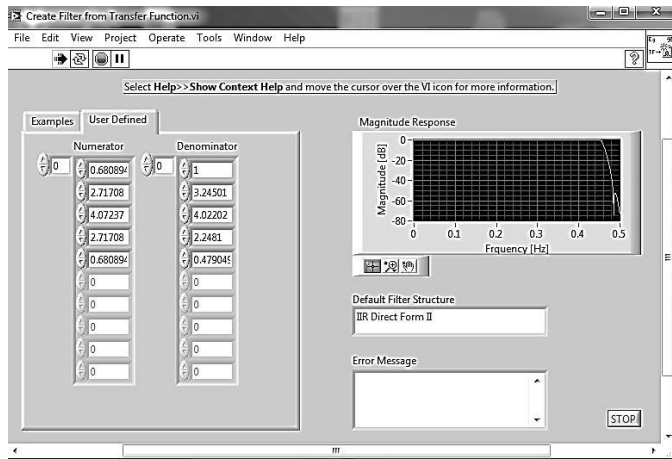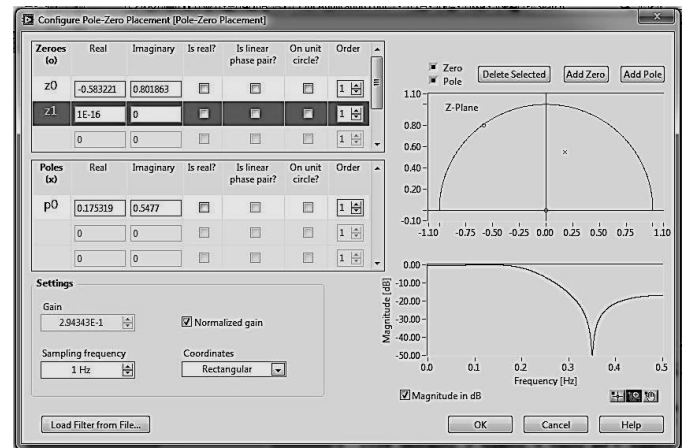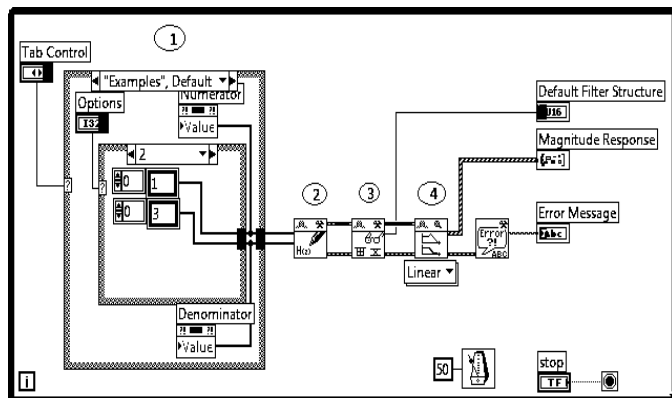
2. The next function creates a filter from the corresponding transfer function.

3. The next step is to retrieve the filter structure.

4. The final step is plotting the frequency responses of the filter by using waveform graph.

C. Design of IIR filter from Pole Zero Placement

Express VIs are very useful tools in DFDT, that help for the quick creation of digital filters. For instance, with Pole-Zero Placement VI you can interactively place zero and pole on the z-plane by entering the exact values or clicking on the pole-zero graph to move or place them with the mouse. Every change of the mentioned above parameters, affects immediately the shape of the magnitude response shown on fig. 7.

Fig.7. The screenshot of the Front panel of IIR Filter Design VI

IV. RESULTS

Design and implementation of digital filters with LabVIEW DFDT consist of several interactive steps. Some of the advantages of LabVIEW are, that the design parameters can be changed at the time of execution of the program and the execution results are immediately displayed. The result of the constantly repeating (with different input parameters) graphical design procedures is that the designer can see how these parameters affect the filter performance.

In this section, the efficiency, flexibility and interactivity of the LabVIEW environment is demonstrated by an example.

Example: Design a lowpass digital filter with following specifications: passband frequency – f_{pass} = 3.4kHz, stopband frequency– f_{stop} = 3.8kHz, passband ripple – 0.1, stopband attenuation – 60dB.

For design, analysis and implementation of the filter, the VI Design Filter step by step is used. The interface supports the following design methods: for FIR – Kaiser window, the Dolph-Chebyshev window and equiripple, for IIR – elliptic, Chebyshev, Inverse Chebyshev, and Butterworth.

The design starts with entering the filter specifications in "design a new filter" window, shown in fig.8. Initially, with the design method set to equiripple FIR, the order of the filter fulfilled the required parameters is 46. The respective characteristics of the filters such as the impulse response, step response, phase response, magnitude response, group delay, pole-zero diagram are immediately plotted in the window "Design result". Analysis of these characteristics is possible by using the dropping menu (marked as 1 on fig.8).

However, it is possible to select other design methods and to get immediately the new filter order, as well as the actual filter shape. For the given specification, as it is expected the elliptic design requires the minimum filter order – 5. Some of the result characteristics of the filter are shown in fig.9.

With the previous design method, the FIR filter requires higher filter order, but in behalf of it has linear phase response. The choice between both filter classes (IIR or FIR linear phase) depends on all given criteria. With the present example it is ease to observe the performance of both filter classes. This flexibility of the design process, as well as the

option for constant change of the parameters and analysis of the respective new characteristics, makes this software perfect for education. The main advantages and disadvantages of every specific design method and respectively the structure used for its realisation is visualised in a simple and understandable manner.

Fig.8. The screenshot of the Front panel of Filter Design VI

a) Group delay of Lowpass IIR filter

b) Impulse responseof Lowpass IIR filter

Fig.9. The screenshot of the Front panel of IIR Filter Design VI

The coefficients of the obtained filter are directly available in forms such as direct I and II, cascade, lattice form and so on. The filter coefficients that are result of direct form I realization of the Elliptic IIR design, are presented on fig.9, where b[i] are forward coefficients and a[j] − reverse coefficients of the from Eq.2.

$$H(z) = \frac{\sum_{j=0}^{m} b_m z^{-m}}{1 + \sum_{i=1}^{n} a_n z^{-n}} \qquad (2)$$

The final step is the implementation of the best filter using a general-purpose PC, a DSP or in an FPGA.

Fig.8. The screenshot of the Front panel with the Filter's coefficients

V. CONCLUSION

LabVIEW is a flexible, intuitive graphical event-based programming language. LabVIEW programs are called virtual instruments. In this paper a set of VIs for design of digital FIR and IIR filters with different interactive design options are given. Some of the advantages of LabVIEW are, that the design parameters can be changed at the time of execution of the program and the execution results are immediately displayed. The effect of the repeating interactive design changes is that the designer has the possibility to observe, how different design parameters affect the filter performance. The interactive tools are excellent teaching tools, because they provide immediate feedback after specifying the desired filter parameters.

REFERENCES

[1] J. Jerome, *Virtual Instrumentation using LabVIEW*, PHI Learning Pvt., 2010

[2] Y. Singh, S. Tripathi and M. Pandey, "Analysis of Digital IIR filter with LabVIEW", International Journal of Computer Applications, Vol. 10, No. 06, pp.23-30, 2010.

[3] National Instuments, LabVIEW Digital Filter Design Toolkit User Manual, 2012.

[4] National Instuments, Interactively Design Digital Filters with LabVIEW, 2014.

[5] S. Mitra, *Digital signal processing*, A Computer based approach, 4th edition, Sept. 2010.

[6] P. Podder, Md. M. Hasan, Md. R. Islam, M. Sayeed, "Design and implementation of Butterworth, Chebyshev –I and Elliptic filter for speech signal analysis", Int. Journal of Computer Applications, Vol.98, No.7, July2014.
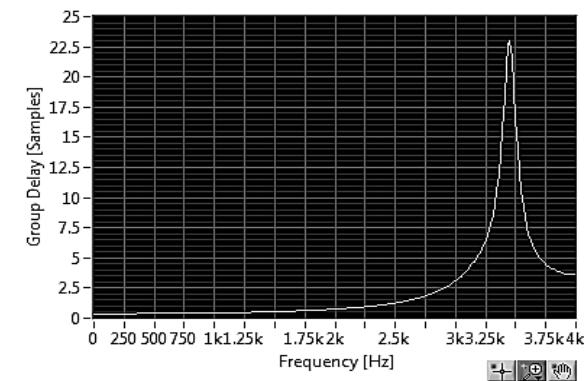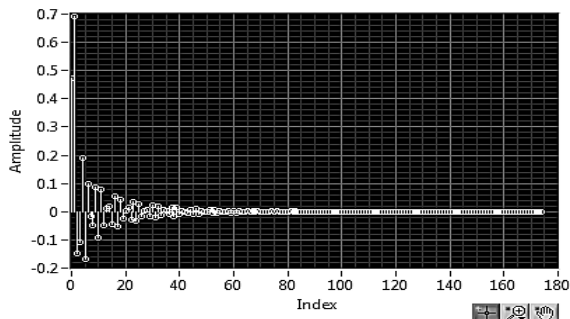
Background Noise Effects Reduction in Swept Sine Measurements of a Room Impulse Response

Dejan Ćirić [1], Marko Janković [2] and Aleksandar Pantić [3]

Abstract – **Main disturbances in acoustic measurements come from noise and distortion limiting the accuracy and reliability. This is also valid for the room impulse response measurements carried out even by the sophisticated techniques such as swept sine. In order to reduce the negative effects of background noise, a method based on known spectrum of the excitation swept sine is developed and analysed in this paper. Applying this method, it is possible to significantly reduce noise in the non-causal part of the impulse response, but also the noise effects in the causal part can be reduced. In this way, the impulse response dynamic range can be increased.**

Keywords – **Noise, Swept sine measurements, Room impulse response, Dynamic range.**

I. INTRODUCTION

Acoustic measurements are an important tool for acoustic investigations, analysis of acoustical problems or for creation of experimental references in theoretical and numerical approaches [1]. The measurements of a tested system Impulse Response (IR) has a very specific position in acoustics, and in some cases, such as in room acoustics, play a crucial role. The IR describes the linear transmission properties of a system able to transport or transform energy in a certain frequency range [2]. The accuracy of IR measurements is limited by the instrumentation, but also by the disturbances among which distortion and noise stand out [3].

The influence of noise is basically related to its interference with useful signal [4] reducing the dynamic range of the measured response. This negative effect depends to a certain extent on the technique applied for the measurement. There are some alternatives that can be applied during the measurements in order to reduce the noise effects. Another option is to apply an adequate post-processing method.

This paper focuses on a specific post-processing method developed for reducing the noise effects in the room IR measurements by the swept sine technique [2,5]. It is based on the fact that the excitation signal is deterministic and that the spectral content of its time segments of short enough duration is narrowband. The developed method can be applied to both transient and background noise, although only the latter one is

considered here. A special attention is paid to the consequences of the noise effects reduction and influence of some parameters of the developed method to its efficiency.

II. NOISE EFFECTS IN ROOM IMPULSE RESPONSE MEASUREMENTS

Noise is present in almost all environments, and as such, it is an inherent part of any acoustic measurement. Two most important types of noise relevant for the acoustic measurements are background (ambient) and transient noise. Background noise can be defined as overall noise present in an ambient where the measurements are done. It is actually noise (airborne, structure-borne and instrument noise) generated by all sources not related to a particular sound that is of interest.

The main consequence of noise during a room IR measurement is a decrease in IR dynamic range, that is, Signal-to-Noise Ratio (SNR). This negative noise effect could be generally reduced by increasing the excitation signal level. Unfortunately, this will also increase the distortion that could again reduce the dynamic range in some measurement techniques [2]. Another possibility for increasing the dynamic range (SNR) is the averaging of multiple responses measured under the same conditions. It is well known that the level of uncorrelated noise is reduced by 3 dB by every doubling the number of averaged responses [2].

When not a room IR is considered, but the decay curve obtained by the backward integration of this IR [6], then various methods have been proposed for the noise compensation. They include the method where the period of integration is limited, the method based on subtracting an estimated noise energy level from a response or the method where room IRs from two separate measurements are multiplied before the integration [7].

The consequences of noise depend on the technique applied for a room IR measurement. In the techniques such as Maximum Length Sequence (MLS) or swept sine [2], these consequences are not that serious as in some other (e.g. single impulse method). The swept sine technique is considered to be a technique enabling high SNR [2,5].

However, the distribution of noise artifact in this technique is somehow unique. If we look at the spectrogram of an IR measured by the swept sine technique presented in Fig. 1(a), we can see a specific distribution in the time-frequency domain. When this IR is filtered in octave or third-octave bands, there is a range of approximately stationary noise artifact shifted along time axis depending on the frequency. At lower frequencies, it is shifted to the right (towards the end of the IR) (see Fig. 1(b)), and at higher frequencies it is shifted in the opposite direction (see Fig. 1(c)).

[1]Dejan Ćirić is with the Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mail: dejan.ciric@elfak.ni.ac.rs.

[2]Marko Janković is with the Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mail: marestudio2004@gmail.com.

[3]Aleksandar Pantić is with the Knauf Insulation d.o.o Belgrade, Industrijsko naselje Belo polje bb, 17530 Surdulica, Serbia, , E-mail: aleksandar.pantic@knaufinsulation.com.
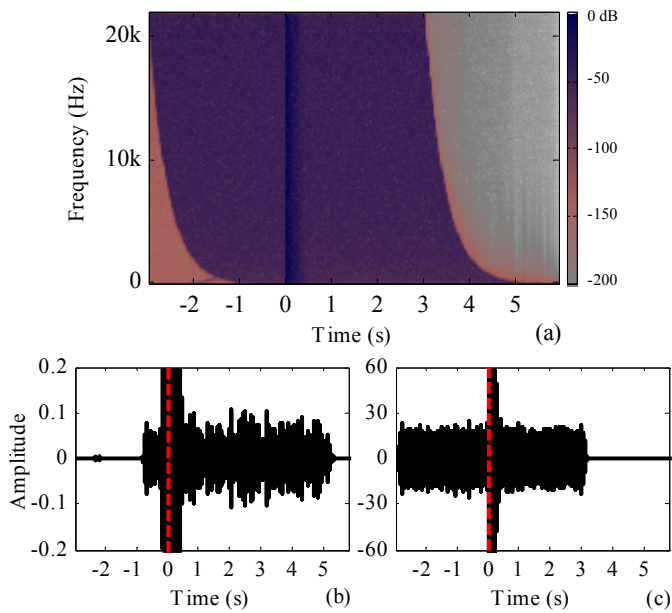
Fig. 1. Room IR with present background noise: (a) spectrogram, and zoom view of the part of IR with noise filtered in third-octave bands at (b) 100 Hz and (c) 16 kHz.

III. METHOD FOR NOISE EFFECTS REDUCTION

The idea based on which a method for noise effects reduction is developed here is rather simple one and can be described as follows. The excitation in the swept sine technique is a deterministic sine signal whose frequency is swept in time either linearly or exponentially [2,5]. This signal is completely known. It can be divided into shorter (time) segments. Any of these segments contains the spectral components only in a limited frequency range. There is no excitation energy outside that range. Its position on the frequency axis depends on the position and duration of the time segment. A segment closer to the swept sine beginning has a frequency range at lower frequencies. Also, a segment with longer duration has generally wider frequency range.

The excitation swept sine can be divided into segments in different ways. One of them is to create the segments of equal duration as done in Fig. 2(a). The corresponding frequency ranges of these segments will not have equal width (bandwidth) - the bandwidth at lower frequencies will be smaller than at higher frequencies (see Fig. 2(c)). Another possibility is to divide the swept sine into segments of equal bandwidth (see Fig. 2(d)). The segments obtained in such a way will have different durations (in time), as presented in Fig. 2(b). The swept sine can be divided into segments by combining two previously mentioned possibilities: some segments will have equal duration while the others will have equal bandwidth. One more alternative is to create the segments of different durations and different bandwidths.

If a room response to the excitation swept sine is divided into segments in the same way as done for the swept sine, the frequency bandwidths of the coincident segments in the swept sine and room response to this signal can be considered to be the same. Now, for evry of these segments, it would be

necessary to pass (filter) the spectral components from the bandwidth of that segment, and not pass (filter out or significantly attenuate) all the components outside the bandwidth. This could be done by bandpass filtering. This feature was already applied for the transient noise removal in the swept sine measurements [8]. However, in case of background noise reduction or removal, it would require design of a number of bandpass filters with rather strict specifications that could cause certain problems. This is why a somewhat different approach (although similar to filtering) is chosen here.
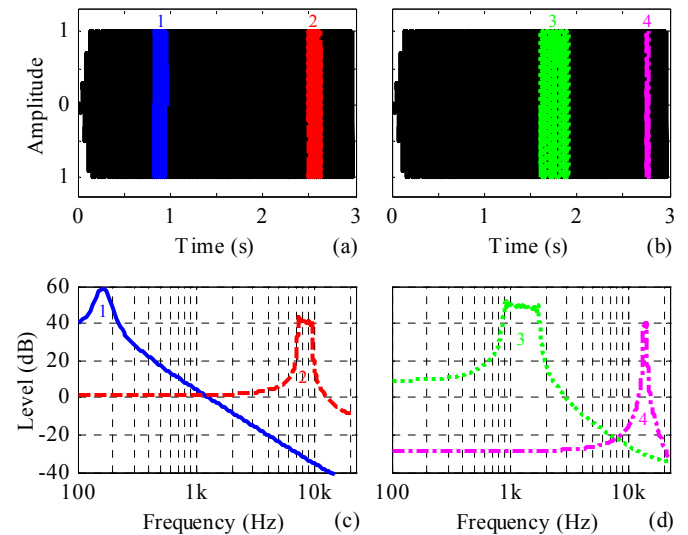
Fig. 2. The exponential swept sine divided into segments of (a) equal duration, and (b) equal frequency bandwidth, together with the spectra of these segments: (c) spectra of the segments from (a), and (d) spectra of the segments from (b).

Reduction or removal of the noise components outside the considered bandwidth (bandwidth of the particular swept sine segment) is realized in the way presented in Fig. 3. First, the excitation swept sine and response to this signal (but only up to L representing the length of the swept sine) are divided into segments according to one of the alternative possibilities described above. The corresponding segments in these signals are coincident in time.

These segments are then transformed to the frequency domain. The spectrum of each of the swept sine segments is first normalized to its maximum value, and then modified. This is done so that its value becomes 1 in the target bandwidth (bandwidth of the swept sine segment or somewhat wider), while the values outside this bandwidth remain unchanged. The normalized and modified spectra of the swept sine segments are multiplied by the spectra of the corresponding segments of the response to the swept sine. In this way, the corrected spectrum for each of the response segments is obtained. It should be noticed that the phase of the frequency counterparts of the response segments remain unchanged. The corrected segments from the frequency domain are then transformed back to the time domain. Summing of all corrected segments in the time domain leads to the corrected response with reduced noise.
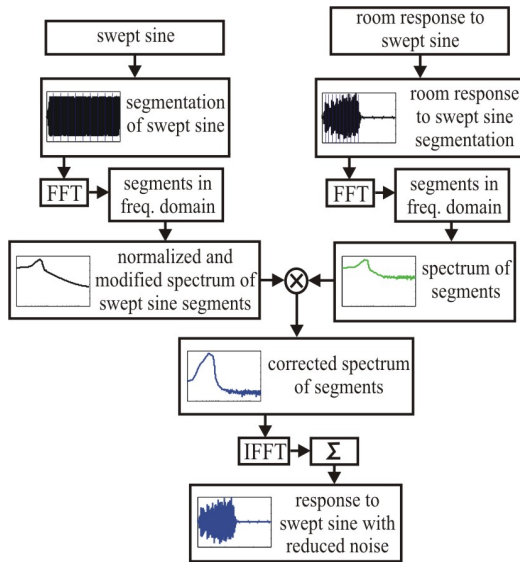
Fig. 3. Flow diagram of the developed method for background noise reduction.

IV. RESULTS

The method for background noise effects reduction is applied here to a simulated room IR generated by the model described in [9]. The measurement procedure is simulated by using adequate equations. An exponential swept sine of length (*L*) equal to that one of an MLS signal of degree 17 covering the range from 20 Hz to 22.05 kHz is chosen to be an excitation signal. The sampling frequency is 44.1 kHz. Since the room IR is set to be noiseless, background noise of controlled level is added to the room response to the excitation swept sine.

The developed method is also tested on a number of measured room IRs, but due to the lack of space these results will be presented elsewhere.

The spectrogram of the extracted room IR when original (not-reduced) noise is present is already shown in Fig. 1(a). When there is no noise in the room response up to *L*, which is a theoretical target in a noise removal procedure, the spectrogram has a shape as in Fig. 4(a). This theoretical target can be hardly reached.

Although all previously mentioned segmentations of the swept sine and response to this signal are implemented, only the results for the third segmentation are presented here. In this segmentation, a part of the segmented signals (3/4 of the signal) is divided into segments of equal (time) duration, while the rest part is divided into segments of equal (frequency) bandwidth. The numbers of segments with equal duration and equal bandwidth are the same. The effects of some parameters of the noise reduction procedure are analysed including the width of the passband and number of segments.

The bandwidth of each segment of the swept sine (target bandwidth) can be determined as a difference between the highest and lowest frequency of that particular segment (lower and upper cutoff frequencies - f_{c_low} and f_{c_up}). In case of the exponential swept sine, these frequencies can be calculated by

$$f_{c\_low} = A_1 \cdot f_1 \cdot e^{\frac{t_{s\_start}}{T_{ss}} \cdot \ln \frac{f_2}{f_1}} \qquad f_{c\_up} = A_2 \cdot f_1 \cdot e^{\frac{t_{s\_end}}{T_{ss}} \cdot \ln \frac{f_2}{f_1}} \quad (1)$$

where f_1 and f_2 represent the starting and ending frequency of the swept sine, T_{ss} is the swept sine duration, A_1 and A_2 are the constants shifting the cutoff frequency in a desired direction, while t_{s_start} and t_{s_end} represent the segment start and end, respectively, in time in reference to the start of the swept sine.
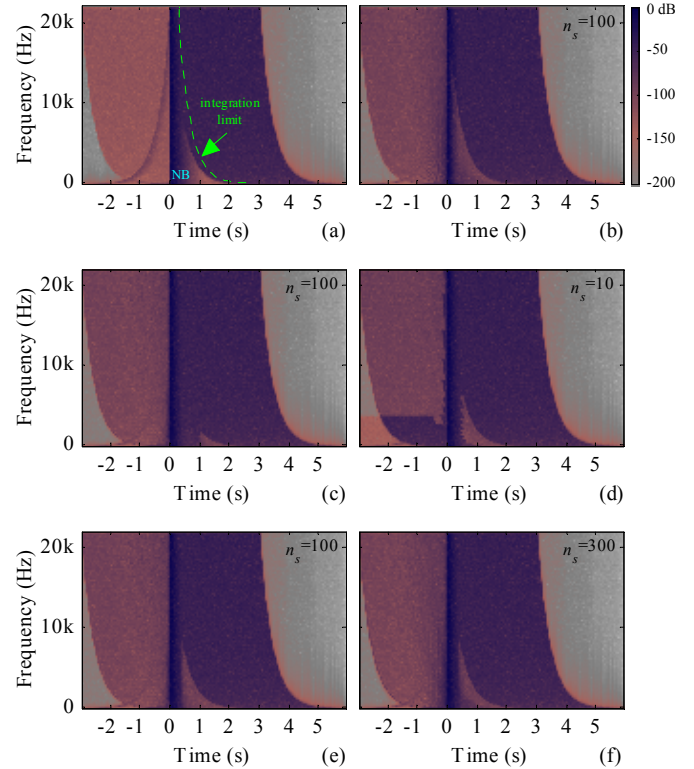
Fig. 4. Spectrograms of the room IRs extracted (a) without noise up to *L*, with noise reduced by the developed method using the bandwidth limits from Eq. (1) for (b) A_1=1 and A_2=1, and wider bandwidths ((c) A_1=0.1 and A_2=1.05, (d)-(f) A_1=0.42 and A_2=1.05) for the number of segments (n_s) given in the upper corner.

When the target bandwidths (applied for the described modification of swept sine spectrum) are equal to the swept sine segment bandwidths (when A_1=1 and A_2=1 in Eq. (1)), the developed method reduces noise significantly in the non-causal part of the IR (up to *L,* that is, 0 s). This is also the case for a part of the causal room IR (from the direct sound (*L* or 0 s) to the end of response), as presented by the spectrogram in Fig. 4(b). Unfortunately, at the same time, a part of the IR is also removed with noise. This can be prevented by widening the target bandwidth, that is, by setting the constant A_1 to be smaller than 1, and A_2 to be greater than 1. By increasing the constant A_2 beyond 1, the left edge of vertical noise band around the IR (denoted by NB in the spectrogram in Fig. 4(a)) is shifted to the left. In the same way, by reducing the constant A_1 below 1, the right edge of noise band is shifted to the right (see Fig. 4(c)), and this widening of the target bandwidth is of higher importance. By choosing adequate values for constants A_1 and A_2, there will be no loss of the room IR, and the

reduction of noise could still be significant. This will not be the case if the target bandwidth is too narrow or too wide. The constant A_2 can take a value only slightly larger than 1, e.g. 1.05. Regarding the constant A_1, the best results (no loss of the IR and the greatest noise reduction) are obtained for an optimal value of this constant, which is in the analysed example approximately 0.42 (see Fig. 4(e)).

Regarding the number of segments (n_s), for rather small n_s, the vertical noise band in the spectrogram can have stepwise shape and noise present at lower frequencies can be larger, as presented in Fig. 4(d). On the other hand, if n_s is too large, for example, above 300 or 500 in the analysed case, there is a sort of additional noise that increases overall noise in the extracted IR (see Fig. 4(f)). This implies that there is also an optimal value of this parameter, and in the analysed case it is about 100 segments.

The effects of noise reduction can be also observed by analyzing the decay curves. They are obtained by the backward integration of the room IR. The upper limit of integration is set to be on the left edge of noise part as shown by the dashed line in Fig. 4(a). It is calculated as

$$t_{\text{int\_lim}} = T_{ss} \cdot \left[1 - B \cdot \left(\ln \frac{f_c}{f_1} \right) \middle/ \ln \frac{f_2}{f_1} \right] \qquad (2)$$

where B is the constant shifting the integration limit (here $B=0.9$ shifting slightly the integration limit towards the IR start), and f_c is the central frequency of the considered (e.g. third-octave) band. The decay curves obtained by integrating the IRs from Fig. 1(a) (with noise), 4(a) (the target curves), 4(b) and 4(e) (with reduced noise) are presented in Fig. 5.
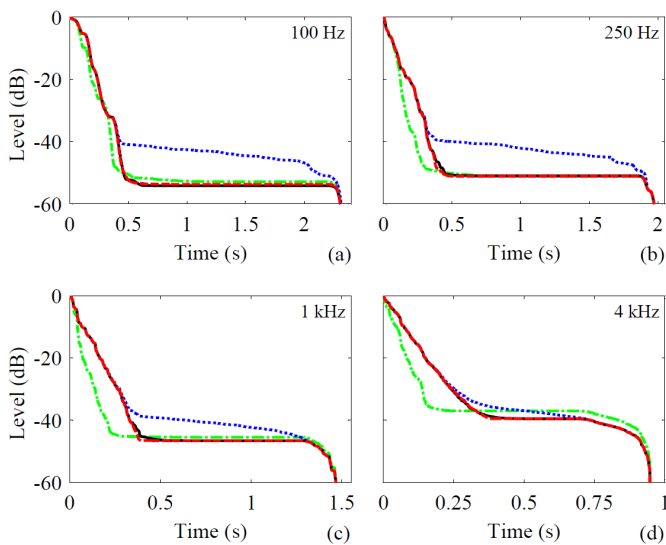

Fig. 5. Decay curves in the third octave bands obtained by the backward integration of the IRs from: (···) Fig. 1(a), (—) Fig. 2(a), (-·-) Fig. 4(b), (---) Fig. 4(e).

Due to the mentioned loss of the part of IR from Fig. 4(b), the decay curves for this IR are useless. However, the optimal application of the proposed noise reduction procedure illustrated in Fig. 4(e) leads to a significant increase in dynamic range, which in most curves at lower frequencies is even larger than 10 dB. These decay curves are almost identical to the target curves (obtained without noise in the response up to L). By increasing the frequency of the considered third-octave band, the reduction of noise effects becomes smaller (compare, e.g. Fig. 5(a) or (b) with 5(d)). So, the improvement in dynamic range at 4 kHz becomes of the order of several dB and further decreases at higher frequencies.

V. CONCLUSION

The proposed method for background noise effects reduction in the swept sine measurements of a room IR is rather efficient in the non-causal part or the IR where noise is significantly reduced. This can be of importance when the distortion products are analyzed. The proposed method also reduces noise in a part of the causal IR increasing the decay curve dynamic range. The increase is greater at lower frequencies and monotonically decreases towards higher frequencies. However, this is a desirable feature since the dynamic range is typically problematic at lower frequencies.

REFERENCES

[1] G. Müller, M. Moser, *Handbook of Engineering Acoustics, Berlin*, Springer, 2013.

[2] S. Müller, "Measuring Transfer-Functions and Impulse Responses", in *Part I: Acoustic Signals and Systems* edited by Finn Jacobsen in *Handbook of Signal Processing in Acoustics, Volume 1* edited by D. Havelock, S. Kuwano, and M. Vorländer, New York, Springer Science, 2008.

[3] D. G. Ćirić, M. Marković, M. Mijić, D. Šumarac-Pavlović, "On the Effects of Nonlinearities in Room Impulse Response Measurements with Exponential Sweeps", Appl. Acoust., vol. 74, no. 3, pp. 375-382, 2013.

[4] S. V. Vaseghi, *Advanced Digital Signal Processing and Noise Reduction,* Chichester, England, John Willey & Sons Ltd, 2006.

[5] A. Farina, "Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique", 108th Convention of the Audio Engineering Society, Abstract in J. Audio Eng. Soc, vol 48, p. 350, 2000, Paris, France, 2000.

[6] M. R. Schroeder, "Integrated-Impulse Method for Measuring Sound Decay without Using Impulses", J. Acoust. Soc. Am., vol. 66, no. 2, pp. 497-500, 1979.

[7] M. Karjalainen, P. Antsalo, A. Mäkivirta, T. Peltonen, V. Välimäki, "Estimation of Modal Decay Parameters from Noisy Response Measurements", J. Audio Eng. Soc., vol. 50, no. 11, pp. 867-878, 2002.

[8] A. Farina, "Advancements in Impulse Response Measurements by Sine Sweeps", 122nd Convention of the Audio Engineering Society, Vienna, Austria, 2007.

[9] D. Ćirić, M. Milošević, "Transient Noise Influence in MLS Measurement of Room Impulse Response", Acta Acust united Ac, vol. 91, no. 1, pp. 110-120, 2005.

Modeling of 3D Impedance Tube with a Complex Termination Impedance using Finite Element Method

Tsvetan Nedkov[1]

Abstract – **Impedance tubes are used to measure the acoustic impedance of a sound absorbing material and are essentially one-dimensional wave guides, with a source at one end and the acoustic load – typically the test sample of the absorptive material – placed at the other end. There are two common methods by which the impedance of the material is measured in an impedance tube. The first involves a moveable microphone that traverses the length of the tube. This method is the older and arguably simpler of the two methods but is slow. The second method is known as the "two-microphone" or "transfer function" method and will be the focus of the method employed in this section. In this study will be modeled a 3D impedance tube with a complex termination using finite element method software and the results will be compared with experimental data.**

Keywords – **Impedance tube, 3D modelling, Finite element method, Complex impedance termination.**

I. INTRODUCTION

Characterizing the absorptive properties of acoustic materials is critical for understanding their behavior when deployed in engineering applications. Mainly for building and industrial acoustic or high frequency noise control absorptive materials are often used. This engineering solution is known as passive noise control and is applied in room acoustics, enclosures, boxes, silencers and barriers. Different types of absorptive materials are available and mainly used are produced from stone or glass wool, technical polyurethane foams and wood based fiber plates.

The most important acoustic parameter for porous materials is absorption coefficient. According to it, these materials are classified as absorptive or reflective. The absorption coefficient can be defined as the relationship between the acoustic energy that is absorbed by a material and the total incident impinging upon it. This coefficient is limited between 0 for non absorbent materials and 1 for totally absorbent. Measurement methods have been developed and standardized in order to enable the study of the acoustic properties of different materials in order to use these data in acoustic projects [1]. There are two types of measurement of absorptive materials: in reverberation room described in EN ISO 354:2003 [2] and measurement in impedance tube described in EN ISO 10534:1 [3] and 10534:2 [4].

Using impedance tube method accurate acoustic absorbing

[1]Tsvetan Nedkov is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: tsvetan_nedkov@abv.bg.

measurements can be provided for normal incident sounds waves only [5]. So with the methods described by [2] or [3] it is possible to determine the normal incidence absorption coefficient and the specific acoustic surface impedance. At low frequencies the impedance tube method will not give accurate results because the sample must be attached air tightly and at the same time the sample must be able to vibrate freely [6].

II. THEORY

With the application of the two-microphone method the impedance, complex reflection coefficient, and the normal-incidence sound absorption coefficient can be determined. Regarding the one-dimensional wave guide shown in Fig. 1, it is excited by plane wave source on the left side ($x = -L$) while it is terminated with a complex impedance, Z, at the opposite end ($x = 0$). Two pressure sensors (microphones) are located at $x = -X_1$ and $x = -X_2$ and are used to determine the magnitude of the forward and backward traveling waves, from which the impedance, flow resistivity, and sound absorption coefficient may be determined [7]. It is assumed that there are no losses along the length of the tube and that only plane waves propagate. The origin of the system is at the right end of the tube and the wave incident on the specimen will be traveling in a positive x direction.
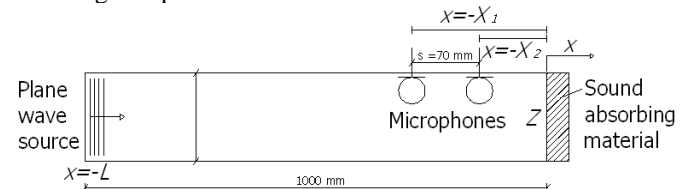
Fig. 1. Impedance tube with source at one end ($x = - L$) and impedance, Z, at the opposite ($x = 0$). The two microphones are used to estimate the termination impedance using the two-microphone technique.

The sound absorption coefficient of a material, α, is defined as the ratio of sound power absorbed by a surface to the incident sound power. For a plane wave at normal incidence, the sound absorption coefficient is given by [8]:

$$\alpha = 1 - |r|^2 \qquad (1)$$

r is the complex sound reflection coefficient and is defined as follow equation:

$$r = \frac{p_r}{p_i} \qquad (2)$$

p_r is reflected pressure, p_i is the incident pressure.

The incident and reflected waves can be written as [7]:

$$p_i = p_0 e^{j(\omega t - kx)} \quad \text{And} \quad p_r = r p_0 e^{j(\omega t + kx)} \qquad (3)$$

.

where p_0 is the amplitude of the incident wave, rp_0 is the amplitude of the reflected wave, $\omega = 2\pi f$ is the angular frequency, and k is the wave number. The total sound pressure, p_t, at a location x in the tube is the sum of the forward and backward traveling waves:

$$p_t = p_0 e^{j(\omega t - kx)} + r p_0 e^{j(\omega t + kx)} \qquad (4)$$

and the total particle velocity, u_t, is given by:

$$u_t = \frac{p_0}{\rho_0 c_0}(e^{j(\omega t - kx)} - r e^{j(\omega t + kx)}) \qquad (5)$$

.

The complex transfer function between the two microphones located at $x = -X_1$ and $x = -X_2$ is:

$$H_{12} = \frac{p_1}{p_2} = \frac{p_0 e^{j(\omega t + kX_1)} + r p_0 e^{j(\omega t - kX_1)}}{p_0 e^{j(\omega t + kX_2)} + r p_0 e^{j(\omega t - kX_2)}} , \qquad (6)$$

which can be solved for r giving the reflection coefficient as a function of microphone positions and the transfer function between the two as [9] :

$$r = e^{2jkX_2} \frac{H_{12} - e^{-jk\Delta X}}{e^{jk\Delta X} - H_{12}} , \qquad (7)$$

where $\Delta X = X_2 - X_1$. Hence, knowing the locations of the microphones, the complex sound reflection coefficient, and thus sound absorption coefficient, can be determined directly from the pressure transfer function between two microphones. At the same time, the reflection coefficient can be represented as:

$$r = \frac{p_r}{p_i} = \frac{Z - Z_0}{Z + Z_0} , \qquad (8)$$

where $Z_0 = \rho_0 c_0$ is the characteristic impedance of the fluid and Z is the specific acoustic impedance of the absorbing surface.

The specific acoustic impedance at any point, x, is defined as the ratio of the total pressure and total particle velocity and is given by [10]:

$$Z(x) = \frac{p_t(x)}{u_t(x)} = \rho_0 c_0 \frac{1 + r e^{2jkx}}{1 - r e^{2jkx}} \qquad (9)$$

The specific acoustic impedance ratio (or normalized acoustic impedance) is defined as:

$$\frac{Z}{Z_0} = \frac{p_t}{p_0 c_0 u_t} = \frac{1 + r e^{2jkx}}{1 - r e^{2jkx}} \qquad (10)$$

The impedance ratio as $x = 0$ is therefore

$$\frac{Z}{Z_0} = \frac{1 + r}{1 - r} \qquad (11)$$

And the complex reflection in terms of the impedance ratio is:

$$r = \frac{Z/Z_0 - 1}{Z/Z_0 + 1} \qquad (12)$$

In terms of a real (resistive) part and an imaginary (reactive) part, the specific acoustic impedance ratio (also known as the normalized specific acoustic impedance) is:

$$\frac{Z}{Z_0} = R + jX , \qquad (13)$$

where R is resistance and X is reactance. The equivalent admittance is:

$$\frac{Z}{Z_0} = \frac{1}{R + jX} = \left[\frac{R}{R^2 + X^2}\right] + j\left[\frac{-X}{R^2 + X^2}\right] \qquad (14)$$

It can be shown that the sound absorption coefficient in terms of the real and imaginary parts of the impedance ratio is given by:

$$\alpha = \frac{4R}{(R^2 + X^2) + 2R + 1} = 1 - |r|^2 \qquad (15)$$

III. MODELING

A 3D configuration of the test set-up as described in 10534:2 [4] was modeled using FEM software Ansys. It was used a linear 3D brick acoustic fluid element FLUID 30 defined by 8 nodes [11]. The acoustic body of the tube is presented with two parameters: speed of sound in air and volume density. At the end of the tube is inserted Harmonic Mass Source (Loudspeaker) generating plane wave and on the opposite side Acoustic attenuation surface (the tested specimen from PU elastic foam). At two different positions (Fig. 2) are placed microphones. It was provided three different models of the mesh size showed in Table I. As the observed frequency range is between 244 Hz to 2 kHz, so the mesh element size of Model 1 is 0.01 m to ensure that there

are at least 16 elements per wavelength at the highest frequency of interest. A mesh sensitivity analysis was carried out and the maximum length must satisfy the equation:

$$E_l = \frac{c}{8f}, \tag{16}$$

where E_l is the maximum element length, c is the speed of sound in air and f is the highest frequency of excitation.

For Model 2 was applied equation (16) for every one-third-octave band in the software simulation. For Model 3 was applied equal mesh size of 0.014 m for frequency region 250 – 400 Hz in reason of save computation time and for comparison with Model 2.

TABLE I
ELEMENT SIZE DISTRIBUTION WITH RESPECT TO FREQUENCY AT ONE-THIRD-OCTAVE BANDS

Frequency	Model 1	Model 2	Model 3
250	0,01	0,17	0,014
315	0,01	0,14	0,014
400	0,01	0,11	0,014
500	0,01	0,086	0,086
630	0,01	0,068	0,068
800	0,01	0,054	0,054
1000	0,01	0,043	0,043
1250	0,01	0,034	0,034
1600	0,01	0,027	0,027
2000	0,01	0,021	0,021

The upper limiting frequency f_u is calculated from the follow condition [9]:

$$d < 0.58\lambda_u \tag{17}$$

The lower limiting frequency f_l is defined using the follow conditions:

$$s_0 > 0.05\lambda_l. \tag{18}$$

And the follow condition must be satisfied:

$$f_u.s_0 < 0.45c_0. \tag{19}$$

Taking into account described conditions for distance *s =70 mm* the following limiting frequencies can by calculated [11]:

$$f_u < 2193 \text{ Hz and } f_l > 244 \text{ Hz};$$

The parameters used in the analysis are presented in Table II.
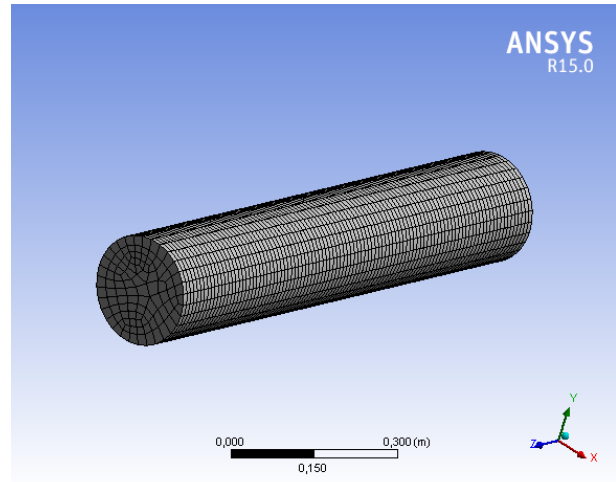
Fig. 2. Element size distribution of Model 1

TABLE II
PARAMETERS USED IN THE SIMULATIONS

Description	Unit	Value
Impeadance tube length	m	1
Impeadance tube width	m	0,1
Speed of sound c_0	m/s	343,24
Density ρ	kg/m^3	1,2041
Acoustic flow	kg/s^2	1
Porosity φ	-	0,99
Tortuosity α_ϖ	-	1,02
Viscous charact. length Λ	m	0,000078
Thermal charact. length Λ	m	0,000192
Thickness of PU elastic foam	m	0,05
Flow resistivity σ	Nm^{-4}s	12569
Density of frame ρ_1	kg m^{-3}	8,85

IV. COMPARISON BETWEEN CALCULATED AND MEASURED RESULTS

To verify the correctness of the model are used measurements provided in the book of Allard and Atalla [7]. Three different mesh densities were applied in order to study the mesh sensitivity in calculating the absorption coefficient with respect to frequency. Comparing α coefficient values shown in Fig. 3, it can be seen that the mesh size has effect on the obtained results. The highest differences for all the frequencies were exhibited for a mesh size of 0.01 m (Model 1) followed by Model 2 even though comply with equation (16) for a frequency range of 244 to 400 Hz, the values obtained at these frequencies were poor compared to Model 3 mesh size tested.

When the absorption coefficient from a mesh size of 0.01 m for a frequency of 244 to 400 Hz were compared to higher mesh sizes of 0.017 m, 0.14 m, 0.11 m and 0.014 m, an average difference of α = 0.08 was observed with a lowest difference of α = 0.03 at 400 Hz and highest of α = 0.1 at 250 Hz. Therefore for a frequency range of 244 to 400 Hz a mesh size from Model 3 can yield satisfactory results compared to a mesh size of Models 1 and 2.

It was found that the best results for a frequency range of 250 – 2000 Hz were observed at mesh sizes of Model 3.
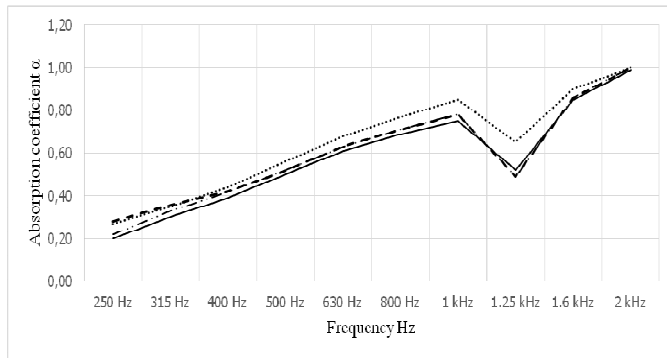
Fig. 3. Absorption coefficient: a) —— measured results; b)········ mesh size Model 1; c)- - - mesh size Model 2; d) —·· mesh size Model 3

V. CONCLUSION

From the provided simulations it's obvious that the usage of FEM models is applicable when measurement equipment like impedance tube is not available. Difficulty may occur in the possibility of collect correct parameters like viscous and thermal characteristic length of the tested absorbing specimen. For future work it can be observed frequency range extended from 63 to 244 Hz and between 2000 to 4000 Hz and the optimal mesh sizing for related frequencies.

ACKNOWLEDGEMENT

REFERENCES

[1] Sn. Pleshkova. Basics of Video and Audio. Publishing TU Sofia, 2012
[2] ISO EN 354:2003. Acoustics: Measurement of sound absorption in a reverberation room.
[3] ISO EN 10534-1 - 1996. Acoustics: determination of sound absorption coefficient and impedance in impedance tubes. Part 1: Method using standing wave ratio.
[4] ISO EN 10534-2 - 1998. Acoustics: determination of sound absorption coefficient and impedance in impedance tubes. Part 2: Transfer function method.
[5] L. Beranek, I. Ver, "Noise and Vibration Control Engineering: Principles and Applications", 1[-st] edition, New York, 1992.
[6] K. Vissamraju. "Measurement of absorption coefficient of road surfaces using impedance tube method". Master thesis. Faculty of Auburn University, Alabama, 2005.
[7] C. Howard, B. Cazzolato, "Acoustic Analyses Using Matlab and Ansys", CRS Press, 2014.
[8] J. Allard, N. Atalla, "Propagation of Sound in Porous Media: Modelling Sound Absorbing Materials", John Wiley and Sons, second edition, 2009.
[9] G.Muller, M. Moser, "Handbook of Engineering Acoustic". Spingers, 2013.
[10] D. Bies, C. Hansen, "Engineering Noise Control: Theory and Practice" Spon Press, London, UK, fourth edition, 2012.
[11] ANSYS, Element Reference Manual, ANSYS, SAS IP, 2011.

Simulation of Microphone Array for Sound Localization using Human Binaural Hearing Model

Viktor Hristov[1], Snejana Pleshkova[2] and Alexander Bekiarski[3]

Abstract – **There are proposed many microphone array structures and algorithms, which are tested and can be characterized with advantages and disadvantages in some concrete applications. Here in this article is proposed to apply and test a combination of a simple two microphone array structure and algorithm for modelling and simulating binaural human perception. The most important thing in this proposal is the simulation of human ears with two microphones and to choose their type to have the beam forming diagram of sensitivity similar each other. Also in simulation algorithm is proposed to apply the elements of human binaural hearing to as much as is possible to known the complex and not fully understand the processes in the human brain on auditory perception and in particular human perception of the direction from which sound waves arrive. The results from simulations are presented to show the characteristics of sound source localization (SSL) achieved with proposition to apply human binaural hearing model for the purposes of sound direction of arrival and localization.**

Keywords – **Human binaural hearing model, Sound source localization (SSL), Microphone arrays.**

I. INTRODUCTION

Investigation of the human hearing system and attempts for its simulation are one the most popular themes nowadays. If the sound signals are received from a person it is easy to determinate the proper direction only with the help of the human ears and human brain. But if the human is replaced with mobile robot - combination of a proper microphone array structure and the corresponding sound localization algorithm is needed. Both working together are important to achieve the appropriate accuracy of direction of sound arrival and the corresponding SSL. [1]

Microphone array structure and chosen algorithm working together are important to achieve the appropriate accuracy of direction of sound wave arrival and the corresponding SSL. A number of factors affect the spatial aspects of how a sound is perceived. The "Duplex theory" was the first comprehensive analysis of the physics of binaural perception [2, 3, 4, 5 and 7]. If the imaginary or artificial head as model of the human head in a mobile robot human hearing model of mobile robot microphone array had completely spherical and uniform surface, the interaural time difference (ITD) produced by the

sound source that arrives from an azimuth of θ radians can be approximately described with the following equation using diffraction theory [6]:

$$\tau = \left(\frac{a}{c}\right) 2\sin\theta, \qquad (1)$$

where
τ is the interaural time difference (ITD);
a - the radius of the imaginary of human head;
c = 343 m/s - the speed of sound waves in the air.

The goal of this article is to propose, apply and test a combination of a simple two microphone array structure and algorithm for modelling and simulating binaural human perception.

II. DEVELOPMENT OF SSL GEOMETRICAL MODEL

The geometrical model of microphone array based on the human binaural hearing is proposed in Fig.1. It is based on the human hearing aid and binaural SSL. It consists of two microphones M_1 and M_2, with the identical parameters. They are situated on the short distance d between them, with the value similar to the distance of the ears in the human head. The positions of the two microphones are shown in Fig. 1 as coordinates x_{m1}, y_{m1} and x_{m2}, y_{m2}, respectively (assuming x as horizontal, y as vertical coordinate and the origin of coordinate system in the left upper corner in Fig.1). Sound source, producing corresponding sound source signal S, (usually a speaking person or speaker producing sound waves from an audio system) is positioned on the other end of the room. It can be described with the coordinates (x_{src} and y_{src}).

Usually the number of sound waves arriving to the human ears or to each of the microphones in the microphone array is also infinite in real situations, but it is very difficult or impossible to describe the summary sound wave arriving to each of microphones in the microphone array. Therefore, in this article based on human hearing model, a simplified geometric model (Fig. 1) is proposed for sound waves propagation to each of two microphones in a microphone array, assuming the existence only of one direct and two reflected sound waves arriving to each the left and to right microphones, respectively. In Fig. 1 are shown for simplicity only the trajectories of direct and two reflected sound waves arriving from the sound source to the left standing microphone M_1 in microphone array. The corresponding trajectories lengths of the direct, two reflected sound waves arriving to each the left microphone M_1 and reflection angles are indicated in Fig. 2 in the following way:

- l_d direct sound wave trajectory length;

- $l^L = l_1^L + l_2^L$ first reflected sound wave trajectory

[1]Viktor Hristov is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: hristov.viktor@gmail.com

[2]Snejana Pleshkova is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: snegpl@tu-sofia.bg

[3]Alexander Bekiarski is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: aabbv@tu-sofia.bg

- θ_1 reflection angle for first reflected sound wave;
- $l_2^R = l_{21}^R + l_{22}^R + l_{23}^R$ second reflected sound wave trajectory;
- θ_2 angle of the for second reflected sound wave.

The direct and two reflected sound waves with the lengths l_d, l^L and l_2^R produce the corresponding sound signals S_d, S_1 and S_2 at the position of the left microphone M_1, arriving from the sound source S. The explained above simplified geometric model (Fig. 1) can be considered also for the right standing microphone M_2, assuming, that it is necessary only to change the index L with the index R for the corresponding indication of trajectories lengths of the direct, two reflected sound waves arriving to each the right microphone M_2 and reflection angles.
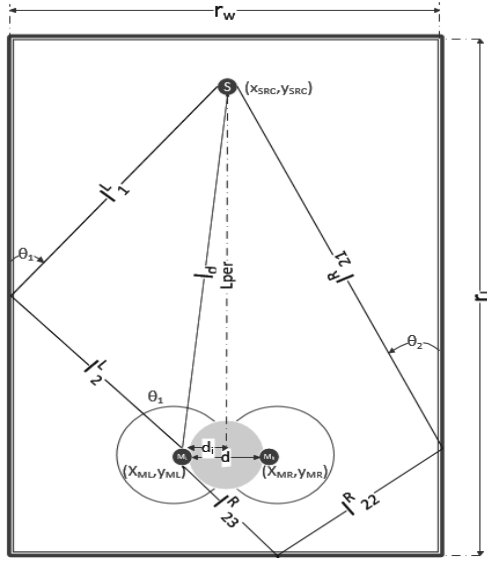

Fig.1. The proposed geometrical model of microphone array

Needed geometric parameters, shown in Fig. 1, are room dimensions (length – r_l and width r_w), distance between microphones – d, sound source coordinates (x_{src} and y_{src}) and positions of the two microphones (x_{m1}, y_{m1}, x_{m2} and y_{m2}). Some of these geometric parameters, for example x_{m2} and y_{m2}, can be calculated in the following way from x_{m1} and y_{m1}, if they are defined:

$$x_{m2} = x_{m1} + d, \qquad (2)$$
$$y_{m2} = y_{m1} = y_m, \qquad (3)$$

where

x_{m2} and x_{m1} are x-coordinates of two microphones M_1 and M_2; y_{m2} and y_{m1} - y-coordinates of two microphones M_1 and M_2; d - distance between them.

For the calculations of the values of the lengths l_d, l^L and l_2^R of direct and two reflected sound waves, is necessary to known the following additional geometric parameters:

$$L_{per} = |y_{m1} - y_{src}| \qquad (4)$$
$$d_i = |x_{m1} - x_{src}| \qquad (5)$$

where

L_{per} is the vertical component of the distance between sound source S and microphones M_1 and M_2 if it is assumed the placement of the two microphones M_1 and M_2 in a horizontal line, i.e. the equation (3) is satisfied;

d_i - the horizontal component of the distance between sound source S and microphone M_1.

Using equations (4) and (5) is possible to calculate the distance L_{di} (for i=1,2) between sound source S and each of microphones M_1 and M_2 as follow:

$$L_{di} = \frac{1}{2}\sqrt{d_i^2 + 4L_{per}^2} \quad \text{for i=1,2} \qquad (6),$$

The angle of arrival θ_0 of the direct sound wave, for example to the microphone M_1, can be calculated using equations (4) and (6):

$$\theta_0 = \arcsin(\frac{L_{per}}{L_d}) \qquad (7)$$

In the similar way, using the ordinary geometric relations is possible to calculate also each of the angles of arrival θ_n to the microphones M_1 or M_2 of reflected sound waves after n=1,2,3…. number of reflections from the walls of the room (Fig. 2). Therefore, the resultant signals S_{M1} and S_{M2}, received from microphones M_1 and M_2, respectively are the sum of corresponding signals produced by the arrived direct and reflected sound waves to each of them, after n=1,2,3…. number of reflections:

$$S_{M1} = k_0^1 S_d^1 + k_1^1 S_1^1 + k_2^1 S_2^1 + ... + k_n^1 S_n^1$$
$$S_{M2} = k_0^2 S_d^2 + k_1^2 S_1^2 + k_2^2 S_2^2 + ... + k_n^2 S_n^2, \qquad (8)$$

where

$k_0^1, k_1^1, k_2^1, ..., k_n^1$ and $k_0^2, k_1^2, k_2^2, ..., k_n^2$ are the attenuation indexes of the corresponding parts (direct and reflected) sound signals produced by direct and reflected sound waves arrived to each of the microphones M_1 and M_2.

The attenuation indexes are important for description of the interaural intensity difference (IID) in algorithms of sound localization. Indirectly in the resultant signals S_{M1} and S_{M2} in equation (8) exist as the lengths of the reflected sound wave trajectories and corresponding time difference of arrival, as estimation of interaural time difference (ITD).

The values of the defined in equation (8) attenuation indexes $k_0^1, k_1^1, k_2^1, ..., k_n^1$ and $k_0^2, k_1^2, k_2^2, ..., k_n^2$ can be determined (in decibels) from microphone beamforming diagram of chosen type of microphones M_1 and M_2, for the corresponding values of angles of arrival θ_n to the microphones M_1 or M_2 of direct and reflected sound waves after n=0,1,2,3…. number of reflections from the walls of the room.

For simplification, in the determination of the values of the attenuation indexes $k_0^1, k_1^1, k_2^1, ..., k_n^1$ and $k_0^2, k_1^2, k_2^2, ..., k_n^2$, is assumed to ignore the losses in reflection and absorption in the walls of the room.

The resultant signals S_{M1} and S_{M2} of microphones M_1 and M_2, derived by equation (8) can be used as input audio information in development and testing the algorithms of SSL. Simulating the appropriate situations and scenarios of positions for different cases for sound waves reflections from the room walls and in room constructions and dimensions is possible.

The proposed simplified planar geometrical model of sound waves propagation from sound source to the corresponding microphones in microphone array.

III. SIMULATION OF SSL GEOMETRICAL MODEL

The simulations proposed to test the proposed simplified planar geometrical model of sound waves propagation from sound source to the corresponding microphones M_1 and M_2 in microphone array presented in Fig.2.

It is based on the human binaural hearing model, arranged following the algorithm and then realized as the corresponding Matlab program.
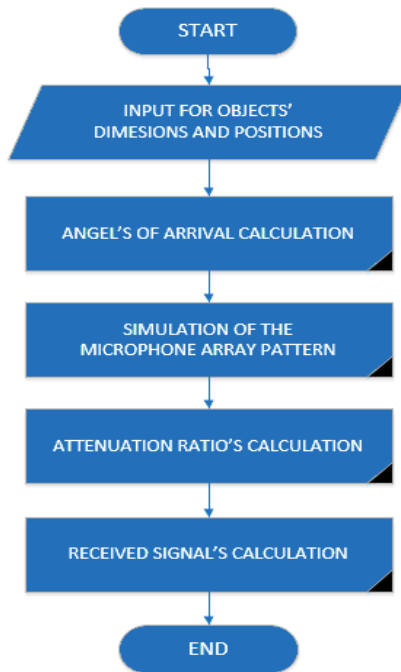

Fig.2. Simulation model is arranged as an appropriate algorithm.

In the first part of the algorithm and the corresponding Matlab program are defined some necessary input parameters for room dimensions - length – r_l and width r_w, also the distance between microphones – d, coordinates of sound source (x_{src} and y_{src}) and positions of the two microphones (x_{m1}, y_{m1}, x_{m2} and y_{m2}).

With these defined parameters the different angles of arrival are calculated depending on the situations if there is no reflection, there is one reflection or there are two reflections.

The generation of microphone beamforming diagram of microphones M_1 and M_2 is done using the existing in Matlab class of microphones as elements of basic class Microphone array. In this case is chosen to use the class "Custom

Microphone Element" as the type of microphones M_1 and M_2, which allow to define their beamforming diagram of the type "cardioid", similar to the human ear hearing sensitivity and with appropriate orientation like in the human ears on human head. These conditions are involved in the following module of Matlab program and the definition of beamforming diagram of the left (microphone M_1) and the right (microphone M_2) orientation is prepared with chosen "+" or "-" in the lines 4 and 5 in the module of Matlab program:

```
h = phased.CustomMicrophoneElement;
h.PolarPatternFrequencies = [500 1000];
h.PolarPattern = mag2db([...
0.5±0.5*cosd(h.PolarPatternAngles);...
0.6±0.4*cosd(h.PolarPatternAngles)]);
resp = step(h,[500 1500 2000],[0 0;40 50]');
plotResponse(h,500,'RespCut','Az','Format','Polar');
```

From the generated beamforming diagrams of the left microphone M_1 and the right microphone M_2, shown in Fig. 3 is possible to obtain the information about the values of attenuation indexes $k_0^1, k_1^1, k_2^1, ..., k_n^1$ and $k_0^2, k_1^2, k_2^2, ..., k_n^2$ in the correspondence of calculated from equation (7) angles of arrival θ_n to the microphones M_1 or M_2 of direct and reflected sound waves after n=0,1,2,3.... number of reflections from the walls of the room. The values of attenuation indexes $k_0^1, k_1^1, k_2^1, ..., k_n^1$ and $k_0^2, k_1^2, k_2^2, ..., k_n^2$ are necessary in the equation (8) for definition of the resultant signals S_{M1} and S_{M2} received from microphones M_1 and M_2 as the sum of corresponding signals produced from the arrived to each of two microphones direct and reflected sound waves.

IV. EXPERIMENTAL RESULTS

Experimental results after the simulations carried out with the proposed and developed simplified planar geometrical model of sound waves propagation from sound source to the corresponding microphones M_1 and M_2 in microphone array based on the human binaural hearing model are presented in the following way.

The initial parameters, needed for the calculations with equation (7) of the angles of arrival θ_n to the microphones M_1 or M_2 of direct and reflected sound waves after n=0,1,2,3.... are room dimensions (r_l=4000 mm and width r_w=6000 mm), sound source coordinates (x_{src}=2000mm and y_{src}=500 mm) and positions of the two microphones (x_{m1}=1900 mm, y_{m1}=5000 mm, x_{m2}=2100 mm and y_{m2}=5000 mm).

For example, in concrete simulation the values of the calculated angle of arrival of direct sound wave, i.e. without reflection, is $\theta_0 = 89$ degree (marked with red arrow in Fig.3), which correspond to the index of attenuation $k_0^1 = -6.02dB$ (Fig.3).

In the similar way, the angle of arrival of sound wave after one reflection is $\theta_1 = 49$ degree (corresponding to 131 degree in Fig. 3 and marked with green arrow) with attenuation

index $k_1^1 = -1.70 dB$. Also the angle of arrival of sound wave after two reflections is $\theta_2 = 22$ degree, which correspond to the attenuation index $k_2^1 = -28.01 dB$.
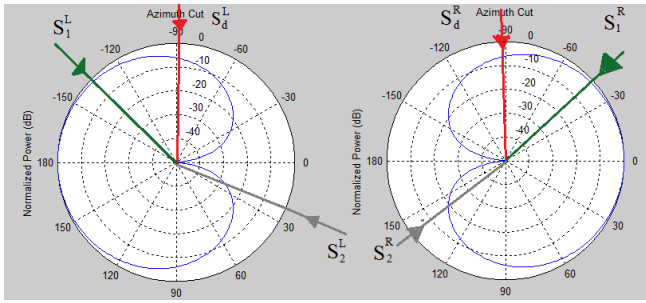

Fig.3. Experimental results

The used above values of the angles of arrival $\theta_0, \theta_1, \theta_2$ and the attenuation indexes k_0^1, k_1^1, k_2^1 are calculated executing the developed for these simulations Matlab program, based on the proposed algorithm, shown in Fig.2. The results in Table I are for these calculated from Matlab program angles of arrival $\theta_0, \theta_1, \theta_2$ and the attenuation indexes k_0^1, k_1^1, k_2^1.

The values of the attenuation indexes calculated by the Matlab program k_0^1, k_1^1, k_2^1 are substituted in the equation (8) for determine the resultant signal S_{M1} received from microphone M_1 as the sum of corresponding signals produced by the arrived to the microphone M_1 direct sound wave and reflected sound waves after n=1,2,3…. number of reflections:

$$S_{M1} = -6.02 S_d^1 - 1.57 S_1^1 - 28.01 S_2^1 \qquad (9)$$

TABLE I
THE CALCULATED IN MATLAB PROGRAM ANGLES OF ARRIVAL AND ATTENUATION INDEXES

Name ▲	Value	Min	Max
k0	-6.0206	-6.0206	-6.0206
k1	-1.5708	-1.5708	-1.5708
k2	-28.0138	-28.01…	-28.01…
theta0	89	89	89
theta1	49	49	49
theta2	22	22	22
xm1	1.9000	1.9000	1.9000
xm2	2.1000	2.1000	2.1000
xsrc	2	2	2
ym1	5	5	5
ym2	5	5	5
ysrc	0.5000	0.5000	0.5000

The similar equation to the equation (9) can be achieved for the right microphone M_2 using the similar results from Matlab program of the attenuation indexes k_0^2, k_1^2, k_2^2 substituted in the equation (8) for determine the resultant signal S_{M2} received from microphone M_2 as the sum of corresponding signals produced by the arrived to the microphone M_2 direct sound wave and reflected sound waves after n=1,2,3…. number of reflections.

V. CONCLUSION

The aim of the proposed in this article geometrical model of sound waves propagation from sound source with combination of a simple two microphone array based on the human binaural hearing model is fully achieved.

The equations derived from the proposed geometrical model lead to achieve the important descriptions of direct and reflected sound waves propagation after the defined simple cases of two numbers of reflections from the walls of the room, especially the resultant signals received from two microphones in the in the microphone array. These equations hold the information of interaural intensity difference (IID) as defined attenuation indexes and information of interaural time difference (ITD) as the lengths of the reflected sound wave trajectories. Used microphones are with similar receiving characteristics like the human ears and their type is chosen to have the beam forming diagram of sensitivity similar to human ears beam forming diagram of sensitivity. The results, derived as equations and simulations, will be used for further deep analyses and will be extended and complicated for the different positions of sound source, increasing the number of sound waves reflections.

REFERENCES

[1] J. Benesty, and J. Chen, "Study and Design of Different Microphone Arrays", 2013, pp. 15-31.
[2] Muller, R.S., Howe, R.T., Senturia, S.D., Smith, R.L., and White, "Microsensors", R.M, IEEE Press, New York, 1991, pp. 350-351.
[3] J. W. Strutt, "On our perception of sound direction", Volume 13, pp. 214–232, 1907.
[4] C. Trahiotis, L. R. Bernstein, R. M. Stern, and T. N. Buell, "Interaural correlation as the basis of a working model of binaural processing", Sound Source Localization, Springer Handbook of Auditory Research, Chapter 7, Springer-Verlag, Heidelberg, 2005, pp. 238–271.
[5] DeLiang Wang, Guy J. Brown, "Binaural sound localization", Chapter 5, Book "Computational Auditory Scene Analysis: Principles, Algorithms, and Applications", John Wiley & Sons, September 29, 2006, 147-187.
[6] M. Bodden, "Modelling human sound-source localization and the cocktail party effect." Acta Acustica, Volume1, 1993, pp. 43-55.
[7] M. Wu, D. L. Wang, and G. J. Brown, "A multipitch tracking algorithm for noisy speech", IEEE Transactions on Speech and Audio Processing, volume 11(3), 2003, pp. 229–241.

Short Time Fourier Transform for Power Disturbances Analysis

Mitko Kostov[1], Blagoj Gegov, Metodija Atanasovski and Mile Petkovski

Abstract – **The Short-Time Fourier transform is described as a detection method of power supply disturbances. Compared to the standard Fourier transform, it provides time data in frequency domain. The most common power disturbances are listed and three simulation models are analyzed to describe the advantages and drawbacks of the Short-Time Fourier transform.**

Keywords – **Power disturbances, Short-Time Fourier Transform, LabVIEW, Simulink.**

I. INTRODUCTION

Power systems are susceptible to numerous disturbances. Appropriate analysis is needed in order to describe when the disturbances appear and what their intensity is. One possibility is to detect them in the frequency domain. Effective detection and removal of such disturbances is essential to the proper functioning of power systems.

Fourier transform is the most commonly used mathematical tool that transforms the signal from time-domain to frequency-domain [1]. The problem arises when time data is needed in the frequency-domain. In time-domain frequency information cannot be obtained and vice-versa, in frequency-domain time data is not available.

One solution to this issue is the Short-Time Fourier transform [2]. The method consists of sliding a window function over the signal and taking its Fourier transform for every half window length. This way the time information is available in frequency-domain with a resolution that depends on the width of the window function.

Other methods include using the discrete wavelet transform (DWT) for detection of power disturbances [3-6]. This offers improved time and frequency resolution with multiple levels of decomposition of the signal.

This paper considers classification of power disturbances and an analysis and practical implementation of the Short-Time Fourier transform for their detection. The paper is organized as follows. After the introduction, the Short-Time Fourier Transform is summarized in Section 2. Section 3 classifies power disturbances. Matlab Simulink models that simulate two types of power disturbances are presented in Section 4. Section 5 concludes the paper.

[1]Mitko Kostov, Blagoj Gegov, Metodija Atanasovski and Mile Petkovski are with the Faculty of Technical Sciences at St. Kliment Ohridski University, Bitola, Ivo Lola Ribar bb, Bitola 7000, Macedonia, E-mail: mitko.kostov@uklo.edu.mk.

II. FOURIER TRANSFORM

The Fourier transform (FT) is the most commonly used tool for analyzing signals in frequency-domain. It is described with the following integral:

$$F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-j\omega t}dt, \ \omega \in (-\infty, \infty) \qquad (1)$$

The result of the Fourier transform of a signal represents its frequency spectrum as illustrated in Fig. 1.

It is easily noticeable that the base frequency of the signal is detected accurately and it is equal to 50Hz. But, if the frequency of the signal is not of the interest, but rather the time when the interruption occurs in the signal, Fourier transform does not provide any time-data in frequency domain; hence it is not possible to know when the interruption occurred.

III. SHORT-TIME FOURIER TRANSFORM

The Short-Time Fourier Transform (STFT) was developed as a solution to the lack of time data in frequency-domain. A signal can be considered stationary in a given time interval. Then the signal can be divided in an n-number of stationary parts. Fourier transform can be performed to each of these stationary parts. Depending on their width; certain phenomena that appear in them can be localized with a certain time or frequency resolution.

The procedure for performing STFT is as follows: the window function is set in t=0 and it is multiplied with the signal. Every value outside the window will have a value of 0. The FT of the product of the signal and the window function is calculated. The result is the FT of the part of the signal that is covered by the window. After that the window is moved by t' seconds and the procedure is repeated until the end of the signal is reached.

Mathematically, STFT is described with the following expression:

$$STFT_x^{(\omega)}(t,f) = \int [f(t) * \omega_p^*(t-t')e^{-j2\pi ft}dt \qquad (2)$$

where $STFT_x^{(\omega)}$ denotes the Short-Time Fourier transform of the signal, $f(t)$ is the analyzed signal, ω_p is the window function (usually a symmetric function that has a non-zero value during a short time interval) and t' denotes the translation along the time axis.
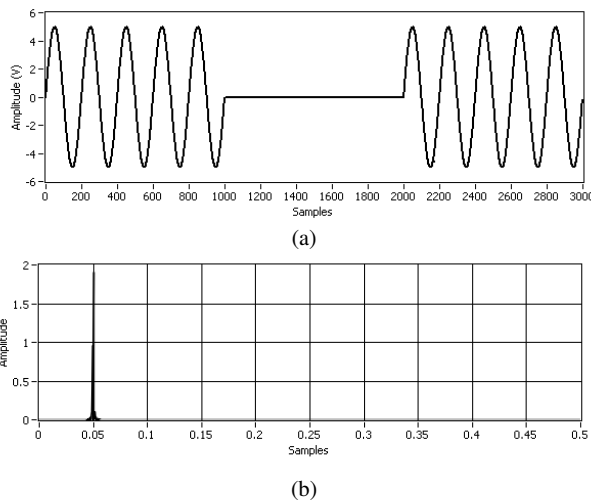
(a)

(b)

Fig. 1. Signal with interruption and its FT.

Fig. 2. Short-Time Fourier transform of the signal with an interruption.

Fig. 3. Gaussian window function.

The resolution of the analysis depends on the width of the window function. A wide window provides good frequency and bad time resolution. A narrow window provides good time and bad frequency resolution.

Fig. 2 illustrates the Short-Time Fourier Transform of the signal in Fig. 1.

Now, it is clearly visible that the interruption occurs around the 1000th sample and stops approximately at the 2000th sample of the signal. The issue now is that the base frequency of the signal is not clearly visible because of the narrow width of the window. Here a compromise is made in order to detect the moment of occurrence of the disturbance more accurately at the expense of the frequency resolution.

IV. CLASSIFICATION OF POWER DISTURBANCES

Power quality disturbances can be divided into seven categories based on wave shape: Transients, Interruptions, Sag/Undervoltage, Swell/Overvoltage, Waveform distortion, Voltage fluctuations, and Frequency variations [7].

Potentially the most damaging type of power disturbance, transients fall into two subcategories: Impulsive and Oscillatory. Impulsive transients are sudden high peak events that raise the voltage and/or current levels in either a positive or a negative direction. An oscillatory transient is a sudden change in the steady-state condition of a signal's voltage, current, or both, at both the positive and negative signal limits, oscillating at the natural system frequency.

An interruption as power disturbance is defined as the complete loss of supply voltage or load current. Depending on its duration, an interruption is categorized as instantaneous, momentary, temporary, or sustained.

Sag/Undervoltage (Swell/Overvoltage) is reduction (increase) of AC voltage.

Waveform distortion is defined as a steady-state deviation from an ideal sine wave of line frequency principally characterized by the spectral content of the deviation. There are generally five types of waveform distortion - DC offset, harmonic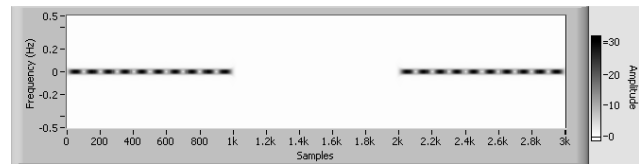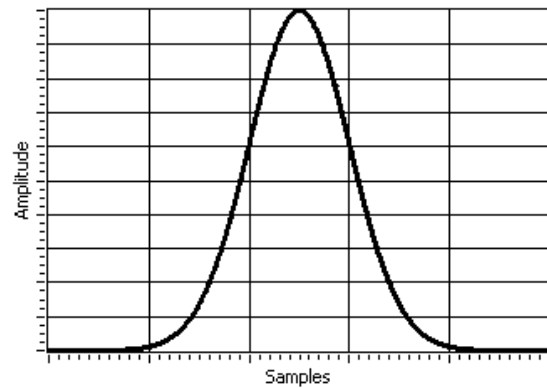s, interharmonics, notching and noise. *DC offset* is the presence of a DC current or voltage in an AC power system. A *harmonic* is defined as a sinusoidal component of a periodic wave or quantity having a frequency that is an integral multiple of the fundamental frequency. *Interharmonics* are defined as voltages or currents having frequency components that are not integer multiples of the frequency at which the supply system is designed to operate. *Notching* is a periodic voltage disturbance caused by normal operation of power electronics devices when current is commutated from one phase to another. *Noise* is unwanted distortion of the electrical power signals with high frequency waveform superimposed on the fundamental.

A *voltage fluctuation* is a systematic variation of the voltage waveform or a series of random voltage changes, of small dimensions. *Frequency variation* is extremely rare in stable utility power systems.

V. SIMULATIONS AND ANALYSIS IN FREQUENCY DOMAIN

Power disturbances can be localized using the Short-Time Fourier Transform. The resolution that they can be localized with depends on the width and the type of the window function.

Three Matlab/Simulink models are presented. The Short-Time Fourier Transform of the outputs of the simulations is performed using NI LabVIEW, using a Gauss window function (Fig. 3). The window function is described with the following expression

$$\omega(n) = e^{-\frac{1}{2}\left(\frac{n-(N-1)/2}{\sigma(N-1)/2}\right)^2} ; \ \sigma \leq 0.5 \qquad (3)$$

where N represents the width of the window in samples and σ is the time step. In the analysis the width of the window is constant and is equal to 100 samples. The time step is set to -1
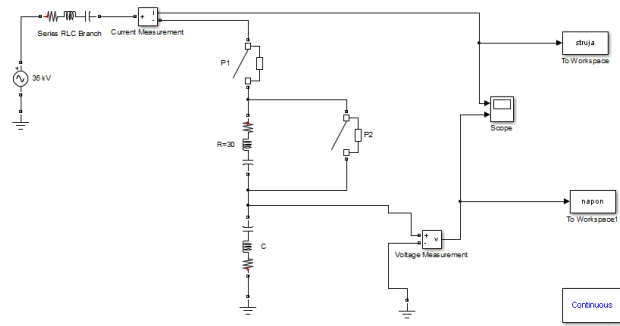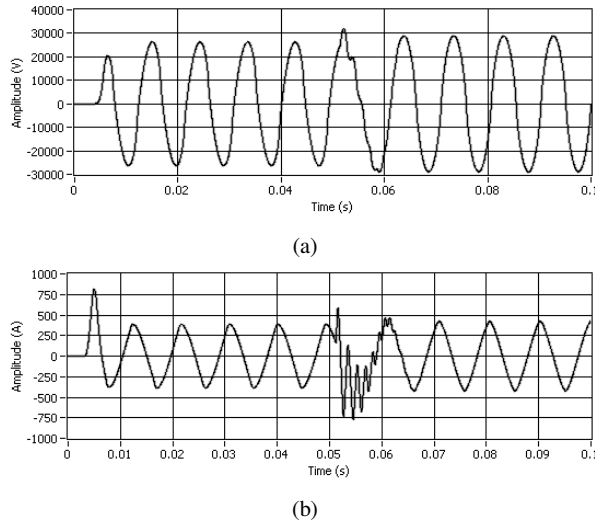
Fig. 4. Simulink model 1.

(a)

(b)

Fig. 5. Voltage and current for the Simulink model 1.

which means that LabVIEW decides which setting is best for the analyzed signal.

A. Simulation model 1

Fig. 4 illustrates a Matlab Simulink model that simulates tranzient-oscilatory power disturbance. It contains three serial connected RLC branches with parameters: 1) $R = 0.808\Omega$, $L = 2.6mH$, 2) $R = 30\Omega$, 3) $C = 46.8mF$. The simulation duration is 0.2s. The voltage in the third branch and the current in the circuit are measured. The switch P_1 turns on at the moment $t=0.02s$. The voltage and current have proper sinusoidal shape until the moment $t=0.12s$ when the switch P_2 turns on. Switching on of P_2 spoils the proper shape of the voltage and current for certain period, and after that the disturbances are weakened and the signals are stabilized, as it is illustrated in Fig. 5.

The moments when the switches P_1 and P_2 are turned on are not visibly detected by the Short-Time Fourier Transform (Fig. 4). The analysis for the current signal (Fig. 4b) provides better insight about the disturbance. The moments when the switches are closed is better visible. If we reduce the window width, the time resolution will improve, but the basic signal frequency will be indistinguishable.
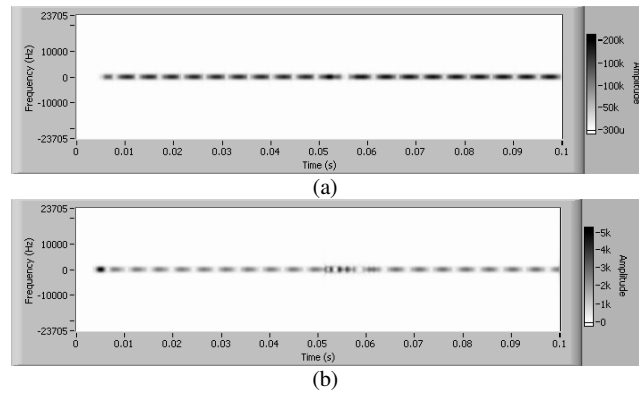

(a)

(b)

Fig. 6. STFT of (a) voltage and (b) current for the Simulink model 1.

B. Simulation model 2

The second Matlab Simulink model, shown in Fig. 7, simulates interruption in power supply (interruption power disturbance). It contains two serial connected RLC branches with parameters: 1) $R = 0.25\Omega$, $L = 15.8mH$ и 2) $C = 11mF$. The simulation duration is 1s. The voltage of the second branch is measured. The switch P turns on at the moment $t = 35ms$, hence the second RLC branch is short-circuited and this could corresponds to an interruption in the power supply. The switch turns off at the moment $t = 253ms$. Because of the reactive elements in the system, oscillations appear and then weackens after certain time. The voltage is shown in Fig. 8.

Using the Short-Time Fourier Transform the disturbance can be detected with a satisfactory time resolution (Fig. 9). The moment when the switch is turned on is described by coefficients with a high frequency and low amplitude. After the switch is turned off, the oscillatory transient disturbance that appears is illustrated by the decaying high frequency coefficients.

C. Simulation model 3

Fig. 10 illustrates a part of an electrical power transmition system. The simulation duration is 0.1s. The voltages of two lines V_a and V_b are measured. The switches P_1 and P_2 turn on at the moment $t = 4ms$. At the beginning the voltages have big oscillations which later weaken, and eventually the voltages obtain proper sinusoidal shape. This condition last until the switches turn off at the moment $t = 80ms$. The waveforms of the voltages U_a and U_b are given in Fig. 11.

The Short-Time Fourier transform of the voltages U_a and U_b in Fig. 12 illustrates that the moment when both switches are turned on. Again, the resolution is not perfect and the moment is illustrated with coefficients with a low amplitude and high frieqency (better visible at the analysis of the signal for U_b). The oscillatory transient disturbance then decays over time until it disappears completely.
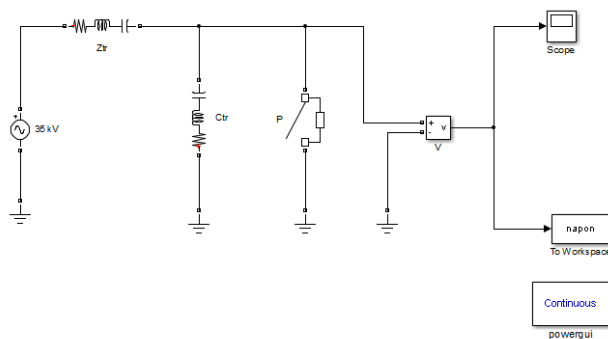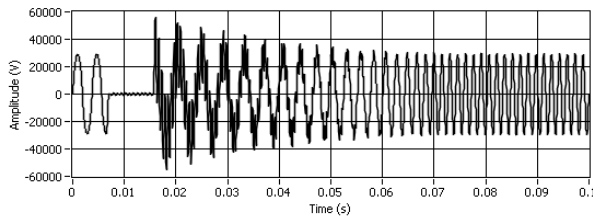
180

Fig. 7. Simulink model 2.
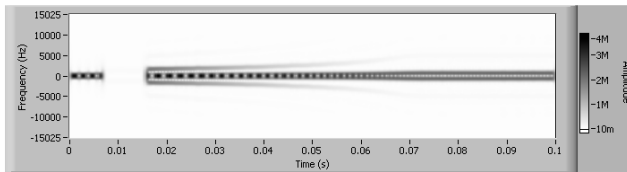

Fig. 8. Voltage for the Simulink model 2.

Fig. 9. STFT of the voltage for the Simulink model 2.

VI. CONCLUSION

The detection of power disturbances is essential towards their efficient removal. This ensures proper functioning of the power system and helps avoid workplace accidents.

STFT can be used to detect power disturbances with a somewhat satisfactory time resolution. We can never obtain a perfect solution, due to the fixed width of the window function. The ideal result would present a compromise between the time and frequency resolution.

An improvement to this method would be the discrete wavelet transform [1]. It possesses an improved resolution, due to the decomposition of the signal in different levels with different wavelet widths. Compared to the discrete wavelet transform, STFT provides better results when analyzing harmonic disturbances.

REFERENCES

[1] G. Strang and T. Nguyen, *Wavelets and Filter Banks*. Wellesley-Cambrige Press, 1996.

[2] P.P. Vaidyanathan, *Multirate Systems And Filter Banks*, Prentice Hall Signal Processing Series, 1992.

[3] Sudipta Nath, Arindam Dey and Abhijit Chakrabarti, "Detection of Power Quality Disturbances using Wavelet Transform", World Academy of Science, Engineering and Technology 49 2009.
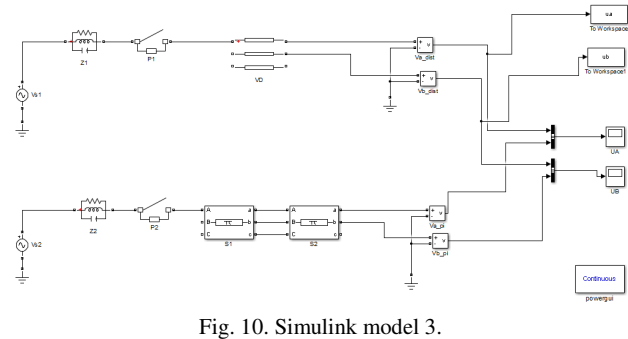
Fig. 10. Simulink model 3.
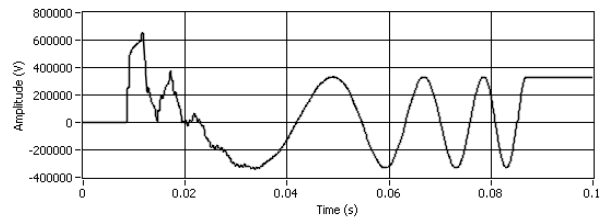
(a)
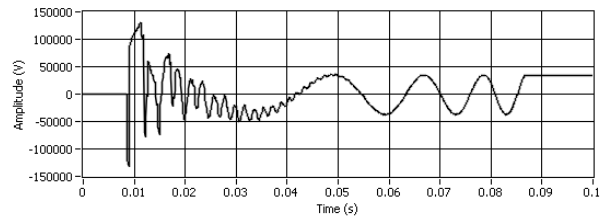
(b)

Fig. 11. Voltages U_a and U_b for the Simulink model 3.
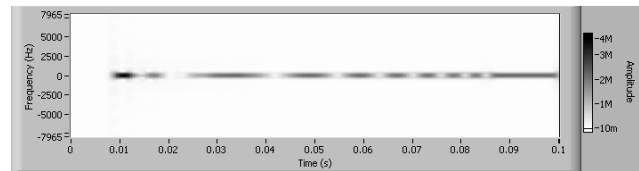
(a)
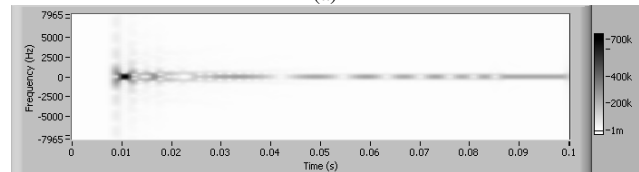
(b)

Fig. 12. STFT of the voltages for the Simulink model 3.

[4] T.Lachman, A.P.Memon, T.R.Mohamad, Z.A.Memon, "Detection of Power Quality Disturbances Using Wavelet Transform Technique", Int. Journal for the Advancement of Science&Arts, vol. 1, no. 1, 2010.

[5] A.M. Gaouda, S.H. Kanoun, M.M.A. Salama, A.Y. Chikhani, "Wavelet-Based Signal Processing Techniques for Disturbance Classification and Measurement".

[6] M.Kostov, B.Gegov, M.Atanasovski, M.Petkovski, C.Mitrovski. "Power Disturbances Simulation and Analysis in Wavelet Domain". ICEST 2014 Proceedings of papers, vol.1, pp.157-160. Nis, Serbia, June 25-27, 2014.

[7] Joseph Seymour, Schneider Electric "The Seven Types of Power Problems", Revision 1.

ORAL SESSION
TELECOMMUNICATION SYSTEMS AND TECHNOLOGY

Network management architecture and models in Internet of Things

Miroslav Slavchev[1] and Aleksandar Tsenov[2]

Abstract – **The main domain of interest in this paper is the implementation of Internet of Things (IoT) solutions in order to meet the ever-increasing demands on modern network management. As a result, the work presents a conceptual framework in order to demonstrate how to incorporate IoT technologies and to elaborate on the requirements and possible advantages. This article also provides examples with concrete IoT solutions and their possible application in the framework as a final proof-of-concept.**

Keywords – **Internet of Things, network management, architectural framework, out-of-band model**

I. INTRODUCTION

Internet of Things (IoT) has presented some emerging technologies [1], which have the potential to provide valuable alternatives and/or additions to common elements of current network management solutions. However, we do not discuss this topic. Thus, the aim of this work is to develop a framework for the seamless integration of IoT in the problem domain.

In general terms, the approach to network management is based on a set of elements, which were used to solve tasks since the first stages of the internet evolution. Examples of these are SNMP, the manager-agent paradigm and in-band management. [2]

Despite modifications (such as SNMP v2 and v3 [3]) in order to meet new demands, these elements impose limitations that may become problematic in the context of modern requirements. A simple example is the shift from reactive to pro-active networking [4], preconditions for which are having accurate enough information about the momentary regional state of the network and low latency by decision-making. The paradigm will result in a delay for data propagation between the agents and a remote manager and it will create severe competition between generated management data (due to finer polling) and primary traffic as well.

However, recent developments have also lead to new technologies and principles, with IoT being the last major step in the evolution of internet. Its applications in various fields, such as the management of households, traffic systems and city resource grids, are already in the focus of research. [5] But the possible benefits of this new design, oriented towards frequent collection, aggregation and distribution of small data

[1]Miroslav Slavchev is a Master-Student at the Faculty of German Engineering Education and Industrial Management at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: miroslav.slavchev@fdiba.tu-sofia.bg.

[2]Alexandar Tsenov is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: akz@tu-sofia.bg

volumes with reduced overhead, is not been discussed in terms of applicability for the tasks of network management.

II. MAIN CONCEPT

A. Key requirements

Seamless integration – in order to be commercially viable, the framework must not be mutually exclusive with existing solutions. It should behave as an optional extension, which is applicable to already existing systems. Integration should require minimal expense and modification (preferably only by configuration).

Modularity – For further reduce of transition costs, a modular approach should be applicable. It must be possible to delegate only certain tasks from an existing solution, without having to incorporate other (non-related) elements of the framework. E.g., one should have the option to delegate accounting data collection to an out-of-band module of the framework without having to migrate the fault-management or other tasks. This can also be beneficial in the sense to incrementally incorporation of the framework in an existing solution, further simplifying the migration procedure into manageable steps.

Out-of-band support – in-band management results in competition for resources between business and management tasks. This coupling introduces complexity in the development of a network, as addressing both problem domains simultaneously and compromising the design should between types of requirements. Using the primary network as a media to deliver management messages introduces unreliability. When managed resources cannot accept new messages (e.g. due to a deliberate attack against the system), their embedded agents cannot communicate with the manager. As a result, none of the orders required for restoring the functional state of the network can be delivered.

Interference – In the scenario, that physically separate primary and management networks should coexist together, the possible loss of data (and reduced utilization) due to interference between the two media should be taken into account.

Layered architecture – the design of the presented framework must consider the dynamic nature of emerging technologies and concepts. This translates directly into a requirement for a layered architecture, which abstracts the concrete specifics of a given implementation, lowering coupling between different elements and the dependencies of the global solution.

Decentralized logic – in order to reduce latency and overhead, responsibilities should be distributed over a chain of command. In comparison to already existing manager-agent

solutions, the more complex tasks will remain based in the manager, but problems which require only local situation awareness should be delegated closer to the agents in order to optimize the process. Enabling communication between agents without dependency on the manager is a prerequisite for logic decentralization.

Scopes – local processing does not exclude the possibility to propagate collected data further up the chain for more complex analysis. E.g, when a fault is detected, agents can collect and share data in order to undertake a simple strategy against the problem. These measures may include a set of steps in direct response to the threat. But they may also asynchronously send the same data to the manager in order to analyze the cause of the problem and possibly apply further actions. Introduction of scopes will allow for finer control and aggregation by such propagation.

B. Concept

Fig. 1 depicts the possibility of presenting auxiliary agents as an intermediating layer between agents and managers. First is the common solution – in-band management of network elements (NE) via embedded Agents (AA) and a remote manager (M).
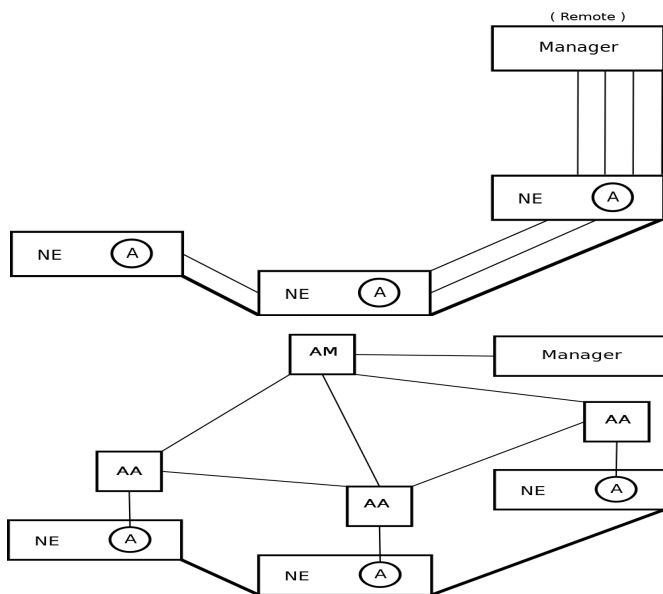

Fig. 1. Model extension via auxiliary elements

The conceptual framework model extends this system via auxiliary agents (AA), directly connected with the original agents. Simpler management and inter-AA communication tasks are incorporated in an embedded logic module (Fig. 2). An auxiliary manager (AM) can be used in order to accommodate more complex policies and coordinate AAs.

Since this is an expansion of an already existing solution, it is known in advance what functionalities the manager exposes to agents (and vice-versa) in order to cover a set of tasks. Therefore, these functionalities can be abstracted into an interface (quasi-manager) and exposed by the auxiliary agent,

allowing to extend the chain of command from (NE-A)-M to (NE-A)-AA-M. From an agent's perspective, messages will be sent to and received from a AA which behaves in the same manner as the manager. On the grounds of this, no changes in the inner logic of agents will be necessary. The same applies to remote managers as well. However, this does not eliminate or prohibit in any way the coexistence between a direct A-M and an A-AA-M chain. Fig. 2 illustrates another option for modification of the chain, where AAs can be used as alternatives to Agents. In this case, a subset of the agent's functionalities are embedded in the AA so that it can independently monitor and interact with the managed resources.

This presents the model from a view based on the management aspect. However, concerns should also be distributed in a manner, which is natural to IoT in order to utilize some important advantages. Agents can embody the sensor and actuator roles, as they already incorporate equivalent functionalities in the traditional manager-agent paradigm. AAs will provide the most important asset – device-to-device ("agent-to-agent") connectivity, and allow for out-of-band communication between local elements.
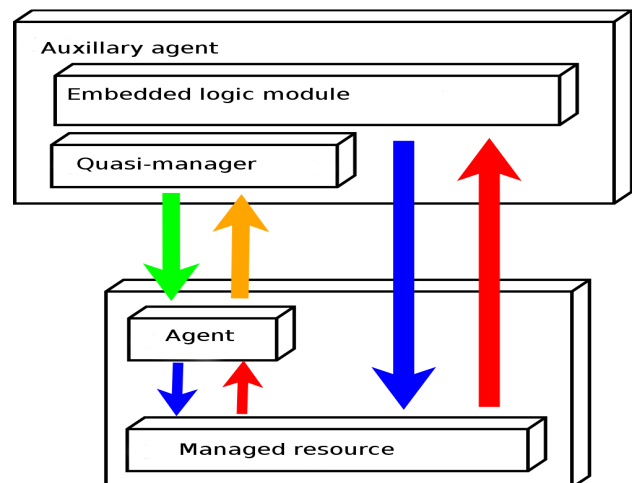
Fig. 2. Network Element– Agent – Auxiliary Agent

Auxiliary managers, if necessary, can be used also as coordinators for larger secondary networks.

The next point of interest is communication within the boundaries of the auxiliary network and between it and the central manager.

Some IoT protocols are more suitable than others for specific tasks. Nevertheless, all of them are designed to support rapid exchange of small data amounts, which is appropriate for most management problems (with the exception of large configuration file distribution). Therefore they are applicable for the purposes of "agent-to-agent" communication. In case some of the data should be further propagated to the remote manager over a standard TCP/IP network, protocol mapping may be incorporated in the AA logical module. If AMs are available, they are the most appropriate candidate for a gateway, adding collection,

filtering and aggregation capabilities before data is converted from one protocol to another. This will also reduce the impact of management traffic on network performance in other sectors which have only in-band support.

III. PRACTICAL APPROACH

A. Management network architecture

The management network itself needs some form of administration. Since a thorough discussion of this problem is not possible within the boundaries of this article, only a generalization will be presented. Based on the aforementioned networking view of the conceptual model, AAs are to be included in one or more groups, subordinated to local auxiliary managers. The responsibilities of the AM also include providing access to AAs to external elements (e.g the remote manager or other AMs).

Devices should also be uniquely identifiable. Regarding their task, AAs need an identity only in the scope of the management group they participate in. Thus, when AAs from two different groups need to exchange information, their full identities are required. For the far more likely scenario of direct communication between AAs within the same group, the overhead of the fully-qualified global identity is neither practical, nor obligatory. Instead, a shorter form of identification can be used. Upon evaluation of the optimal size scalability should be taken into consideration: smaller sizes will directly reduce the overhead by communication between devices from the same group, but it will also reduce the maximum number of elements one subgroup can hold. This will lead to a higher number of groups in order to cover a larger management network. Which also leads to increased probability of communication between AAs from separate groups, requiring the full representation form of both elements. In conclusion, optimal size of the shorter form for smaller domains is 1 Bytes, but mid- and large-scale networks will require 2 Bytes.

A. PHY layer and MAC - 802.15.4

The choice of a physical layer must be carefully considered. The aforementioned motivation for an out-of-band model may be severely compromised by radio frequency interference. RFI may cause significant packet loss or temporally deferred transmissions in both networks, which reintroduces the problem of competition for resources.

802.15.4, by specification, is supported in three following modes: 868 – 868.6 MHz/single channel/20 kbps, 902-928 MHz/10 channels/40 kbps and 2400-2483.5 MHz/16 channels/250 kbps. Of these, only for the last one exists a possibility toS overlap with the standard ISM bands. As prior research suggests, the coexistence between 802.15.4 and 802.11 is an achievable task, but the following measures are highly recommendable in order to minimize risk of RFI: Frequency Division Multiple Access (FDMA) providing 16 channels, FHSS (Frequency Hopping Spread Spectrum) balancing the load between channels and CSMA/SA (Carrier Sense Multiple Access with Collision Avoidance), also known as "listen before talking".

These strategies are advantageous in providing that simultaneous transmissions in the two networks will not overlap. However, in the scenario that a node from one networks detects that all possible channels are currently occupied, it must defer transmission to a later time period (hence the name "listen before talking"). This may cause unfairness. In order to reduce it and in accordance to the nature of management traffic, 802.15.4 packets can be limited to a size of 127 Bytes. The transmission of a single packet will require 50, 25 or 4 ms, depending on the chosen 802.15.4 mode.

The maximal overhead of the 802.15.4 MAC protocol is 39 Bytes (and 88 Bytes for payload). Of these, three fields are optional and thus candidates for optimization: destination address, source address and auxiliary security header.

The auxiliary header can be either 14 Bytes or fully omitted. However, out-of-band management provides full control over the underlying resources and therefore security is obligatory and cannot be sacrificed in order to achieve smaller overhead. But, as already described, destination and source addresses can be limited only to 2 Bytes. These 16 Bytes mean 41.03% overhead reduction and 18.19% increased payload.

B. Internet layer – IPv6

IPv6 introduces 40 Bytes of maximal overhead. However, a similar optimization is applicable here.

Some fields are obsolete. As the solution will be based only on IPv6, the 4 Bits for version field can be used for other tasks. Flow label (20 Bits) is experimental and not fully supported. It does not provide functionalities that are beneficial for the proposed framework. Finally, source and destination fields, 128 Bits each, can be fully omitted. The motivation behind this decision is that IPv6 is used only for the purposes of communication within the management network boundaries and the PHY layer contains enough information for addressing.

It is also possible to reduce the size of other fields. Halving the traffic class field (4 Bits) will provide 4 values and sufficient control over priorities. With less than 104 Bytes of payload, which is representable by 7 Bits. The "payload length" field is 16 Bits, but it can be reduced only to 11, as with future versions a larger 802.15.4 packet size may become available. Considering the relatively low distance between AAs and other AAs or AMs, 3 Bits will be enough for the hop limit field (currently 8 Bits).

In terms of overhead, these decisions lead to a reduction from 40 to 3 Bytes, which is 92.50%. The payload is increased from 64 to 101 Bytes (57.81%).

C. Transport layer – UDP

The source and destination port fields, 8 Bits each, can be reduced to 2 Bits. The management network relies on IoT devices which are constrained and more than 4 ports will not be required. The length field can be reduced from 16 to 12

Bits (considerations are similar to IPv6). With the checksum remaining unchanged (16 Bits), overhead is reduced from 16 to 4 Bytes (75.00%) and payload is increased from 85 to 97 Bytes (18.82%).

D. Application layer – MQTT-SN and CoAP

The MQTT (Message Queuing Telemetry Transport) protocol provides functionalities appropriate for remote telemetry querying. MQTT-SN (MQTT for Sensor Networks) is adjusted to the constraints of the IoT stack, but overhead here can be optimized further. Length and Type fields can be reduces to 11 and 5 Bits respectively. Different MQTT lifecycle tasks, such as automatic network discovery and topic (un)subscription require up to 16 Bytes. Most important is the topic publication task – in this case, overhead is 7 Bytes, leaving up to 90 Bytes for payload.

CoAP (Constrained Application Protocol) provides access capabilities to auxiliary elements based of web services. GET can be mapped to SNMP GET_REQUEST. PUT and SNMP SET_REQUEST are similar. DELETE can be used to remove objects from the MIB. The overhead here is up to 24 Bytes (8 Bytes for source and destination address, since this protocol will be used to provide remote access and 8 Bytes are required to support the core functionalities). This reduces payload to 73 Bytes.

E. Inner data representation

SNMP can be mapped to an inner protocol as a final optimization step as shown on Fig. 3. Three Bits are enough to represent different message types (such as TRAP or GET_REQUEST). Five Bits can be reserved for future usage.

The MIB Object ID is internally represented by 8 Bytes in a "Parameter" field. 1 Byte is enough to represent 256 discrete values, which are enough to cover information such as the current device temperature or port load. Thus the "value" field can contain between 1 and 4 Bytes. For certain events (such as device shutdown) the <parameter, value> tuple can be omitted and a simple code can be used instead. 2 Bytes are enough to represent 65,000 codes.

In order to support messages with different content (variable data length), three extra fields can be applied to every (code, parameter, value) triplet. Thus, 2 Bits are enough to demarcate the code length in Bytes (00-10 for 0-2 Bytes). The parameter field can be presented via 2 Bits, applying the following schema: 00 – no parameter, 01 – 2 Bytes, 10 – 4 Bytes, 11 – 8 Bytes parameter field. As different data types

may require finer control over the value field length, 4 Bits can be used to represent its length (again in Bytes). Since the size-descriptors and real data fields are logically connected, they can be viewed as a "data group" entity, which is between 2 and 17 Bytes long.

To sum up, 2 Bytes are enough to represent the SNMP type and future options. This leaves a payload of 71 (CoAP) or 88 (MQTT-SN) Bytes. Within these limitations, a data group can be fitted easily. This leaves enough space to embed several data groups within a single packet in order to reduce the overhead from other layers and emulate BULK_* operations.

IV. RESULTS

The given practical approach covers all layers, required to implement and fulfill the requirements set in chapter II. It also provides a motivated set of decisions in order to minimize overhead from standard protocols. The possibility of coexistence (and collaboration) between standard (802.11) and IoT networks is proved via a constructional approach. In the practical approach no step depends on the particular management task that is to be executed and therefore this solution is widely applicable.

V. CONCLUSION

IoT is capable of improving and replacing current network solutions and also prepared to meet future requirements. The technology stack is still under a very dynamic development and this article presents a solution based only on current state of the art. But the framework has the potential to utilize future developments in the field of IoT.

REFERENCES

[1] Haller S., S. Karnouskos, C. Schroth "The Internet of Things in an enterprise context", Vienna: Springer (Berlin-Heidelberg), pp. 14-28, 2008
[2] Pavlou G., "OSI Systems Management, Internet SNMP and ODP/OMG CORBA as Technologies for Telecommunications Network Management", Telecommunications Network Management: Technologies and Implementations, pp. 63-109, IEEE Press, 1998
[3] Stallings W., "SNMPv3: A security enhancement for SNMP", IEEE Press, 1998
[4] Bush S. F., A. B. Kulkarni, "Active Networks and Active Network Management", Kluwer Academic, New York, 2001
[5] Tan L., N. Wang, "Future Internet: The Internet of Things" in 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE), August 2010
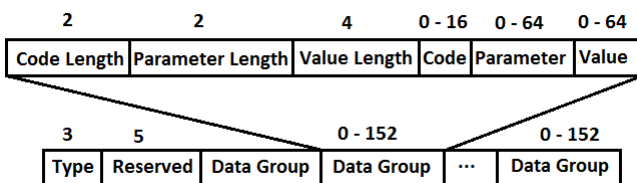
Fig. 3. Inner protocol (fields length in bits)

Analysis on Autonomic Characteristics of Quality of Service Management in EPS

Ivaylo Atanasov[1]

Abstract – **Network evolution is driven by deregulated markets, converged services, open competition, virtualisation and new business models. Efficient operation of such networks calls for "autonomic" or self-management functionality embedded into the network in order to maintain operational costs under control. The paper presents analysis on the autonomic characteristics of Quality of Service management functionality in the Evolved Packet System (EPS). Four basic levels of abstraction are considered in the context of standardized functions and future extensions.**

Keywords – **Evolved Packet Core, Policy and Charging Control, Self-management, Autonomic behaviour.**

I. INTRODUCTION

Network evolution features convergence between information technologies and communications, new business models for service provisioning, explosion of multimedia services, and it is driven by new IP based broadband technologies. Such networks require efficient operations and business support. Due to diversity of technologies and business models, and the operators' requirement for lower operating costs, some level of "autonomics" should be embedded into network equipment and operations support systems. This requires embedding self-optimization, self-healing, and self configuration features [1], [2]. Autonomics and self-management technologies are considered as an intelligent solution for network operators to control their own environment in an efficient and effective way [3], [4].

Evolved Packet System (EPS) covers the radio access, the core network and the terminals that comprise the overall mobile system. In EPS, simple and effective Quality of Service (QoS) mechanisms for multi-vendor mobile broadband deployments are defined. Such QoS mechanisms allow service and subscriber differentiation and dynamic, policy-based control on authorization and usage of bearer resources intended for IP-based multimedia traffic [5], [6].

The paper presents an analysis on standardized QoS management mechanisms in EPS and using different levels of abstraction identifies the autonomic characteristics of policy-based QoS and charging control.

The paper is organized as follows. Next section presents in brief the Generic Architecture of Autonomic Networks (GANA) as defined by ETSI in (ETSI GS AFI 002). The third section presents different levels of autonomic behavior of QoS management mechanisms in EPS. The fourth section

[1]Ivaylo Atanasov is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, e-mail: iia@tu-sofia.bg.

considers self management characteristics of future optimizations of QoS control in EPS. The conclusion summarizes the contributions.

II. THE GENERIC AUTONOMIC NETWORK ARCHITECTURE REFERENCE MODEL

Autonomic behavior is characterized by self management capabilities such as self-configuration, self-healing, self optimization, self-organization and self-protecting. It consists of a set of actions triggered by a decision making element (DE) on a Managed Entity (ME) following policies to achieve a goal. The DE drives a control loop to regulate the behavior of the ME exposing features like learning, reasoning, planning and cognition.

In GANA, DE, ME and control loops are designed in four level of abstraction: protocol level, function level, node level and network level (ETSI GS AFI 002). The lowest protocol level is associated with the self management functionality implemented intrinsically within a single network protocol such as TCP or OSPF. The function level abstracts autonomic mechanisms and protocols associated with a particular network function, such as autonomic routing and autonomic fault management. In the node level, the autonomic properties are exposed by a node or system, which means that the main node's DE has access to the views exposed by the lower level DEs operating on abstracted networking functions. The highest level of self-management functionality is the level of the network's overall functionality and behavior. Fig. 1 shows the generic model of an abstract autonomic networked system.
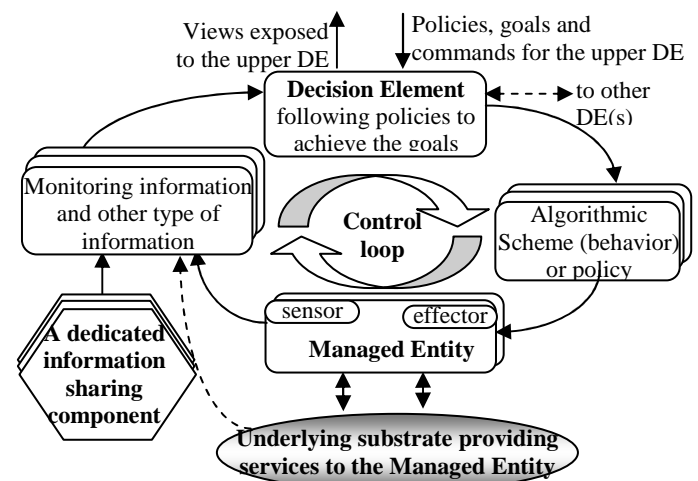

Fig.1 Generic Model of Abstract Autonomic Networked System (ETSI GS AFI 002)

The adopted hierarchy allows some decisions to be taken autonomously at different levels of control and complexity

and offers flexibility towards controlling/resolving decision conflicts and priorities.

The QoS management mechanism in EPS allows dynamic service-aware control on bearer resources used for multimedia packet traffic which is integrated with charging control (3GPP TS 23.203). The next section studies the standardised QoS management functionality and provides abstraction of autonomic characteristics onto the four levels of GANA referenced model.

III. Autonomous Behaviour of QoS Management Mechanism

The EPS bearer represents the level of granularity for QoS control in EPS and provides a logical transmission path with well-defined QoS properties between the mobile device and the network. The QoS management in EPS allows authorization of and usage monitoring of EPS bearers.

A. QoS management in GANA protocol level

The signaling protocol used for QoS management is Diameter. In EPS, the protocol is used for Authentication Authorization and Accounting functions. It does not provide intrinsic autonomic behavior in the context of QoS management, as no QoS control decision logic is embedded in it. But the protocol itself provides autonomous behavior in the context of Diameter request routing and peer discovery. In EPS, a logical entity called Diameter Routing Agent (DRA) is used during the selection/discovery of the function responsible for policy based QoS control. DRA ensures that all Diameter sessions established over the respective reference points for a certain IP session reach the same function responsible for policy based QoS control when multiple functions have been deployed in a Diameter realm.

The reason for excluding the intrinsic behavior related to QoS management in Diameter protocol is that the protocol has become too hard to manage due to intrinsic decision logic in other contexts that may interact in an undesired way with the decision logic of other protocols. This means that QoS management decision logic is implemented at a level higher, i.e. outside the individual protocol.

B. QoS management in GANA function level

QoS management in EPS is based on policies (3GPP TS 23.203). Policy control refers to gating control, QoS control, application detection control, and usage monitoring control. The gating control refers to blocking or allowing of packets, belonging to a service data flow to pass through between desired endpoints. QoS control is the capability of authorization and enforcement of the maximum QoS that is authorised for an IP flow(s). With application detection control feature, it is possible to request the detection of specified application traffic and receive notifications on the start and stop of application traffic. Usage monitoring control allows the operator to enforce dynamic policy decisions based on the total network usage in real time.

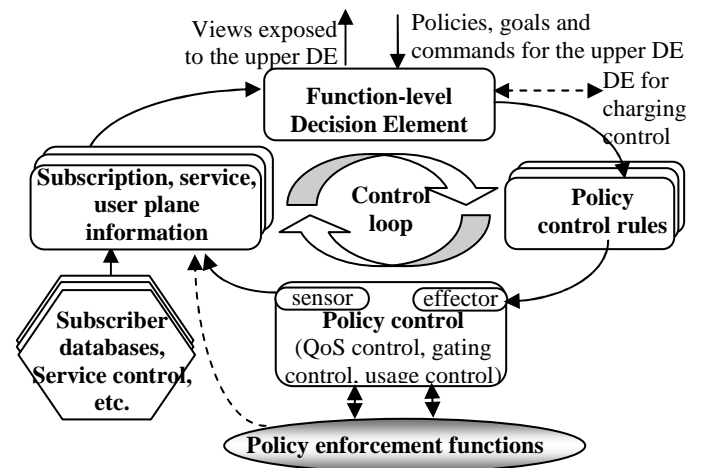Fig.2 illustrates the mapping of policy control functionality onto GANA functional level.

Fig.2 A control loop specific to an abstracted policy control function

The DE implements the decision logic for policy control. The policy control DE uses subscription information from subscriber databases for an IP session at establishment or a gateway control session at establishment. The subscription information may include user profile configuration indicating whether application detection and control should be enabled. The subscription information is retained as it is relevant for policy control decisions until the IP session termination and the gateway control session termination. The policy control DE uses also service information (e.g. acceptable bandwidth, usage threshold). If access network data are used for policy control decisions, congestion level and access technology information may also be included. The policy control DE makes decisions encoded into policy control rules. A dynamic policy control rule may include information related to service data flow detection, information related to policy control (i.e. gating control and QoS control) and usage monitoring control. The policy control rules may be dynamically changed during the session, based on reporting of usage information.

Another function that exposes autonomicity is the charging control function. Charging control includes means for both offline and online charging. The charging information can affect in real time the services being used. The interaction between policy control and charging control ensures coherent charging between IP connectivity access network and service control and is called Policy and Charging Control (PCC). The credit management applies for online charging and it enables policy control based on subscriber spending limits. The information about usage of resources may be used for offline charging. With PCC, the rules generated by the DE include charging information also.

C. QoS management in GANA system level

In the GANA node level of self-managing (autonomic) properties, the lower level DEs operating on the level of abstracted networking functions (such as Policy control DE and Charging control DE) become some of the managed

automated tasks of the main DE of the system (node). This means the node's main DE has access to the "views" exposed by the lower level DEs and uses its knowledge of the higher level (system and network level objectives/goals) to influence (enforce) the lower level DEs to take certain desired decisions, which may in turn further influence or enforce desired behaviors on their associated managed resources.

Fig.3 illustrates the mapping of PCC functionality onto GANA system level. In the context of collaborative autonomic functions performed by each functional entity, the whole PCC functionality may be considered at a system level functionality.
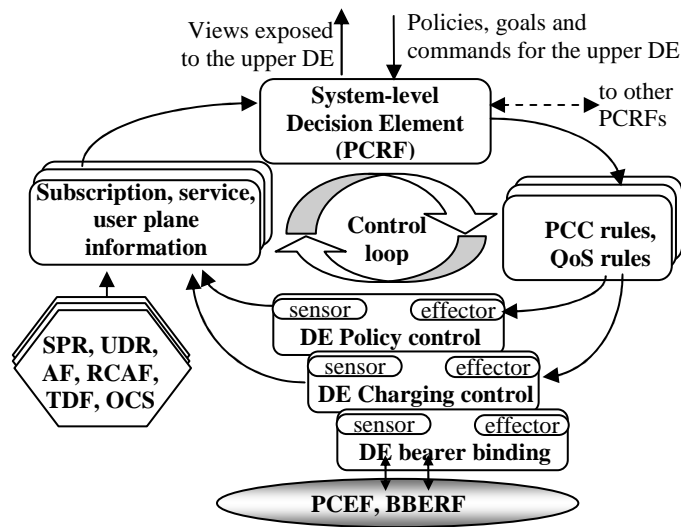

Fig.3 Mapping of PCC functionality onto GANA node-level

The Policy and Charging Control Function (PCRF) appears to the main DE. It receives session information from the access network and subscription information from subscriber databases. The PCRF takes the available information and based on the policies configured by the network operator, creates service-session-level policy decisions. The decisions are provided to the PCEF and the BBERF.

The Policy and Charging Enforcement Function (PCEF) enforces policy decisions (gating, maximum bit rate policing) received form the PCRF and also notifies the PCRF about user- and access network-specific events. The PCEF performs measurements of user plane and reports usage of resources to the PCRF. The Bearer Binding and Event Reporting Function (BBERF) supports a part of the PCEF functions such as bearer binding and event reporting to the PCRF. The Application Function (AF) interacts with applications and services requiring dynamic PCC. Using the application level signaling, the AF extracts session information and provides it to the PCRF. The AF may receive notifications form the PCRF about events occurred in the user plane in case of active subscription. The RAN Congestion Awareness Function (RCAF) reports congestion information in the Radio Access Network (RAN) user plane. This allows the PCRF to take the RAN user plane congestion status into account for policy decisions. The Subscription Profile Repository (SPR) and the User Data Repository (UDR) are databases that store PCC-related subscription data. The Traffic Detection Function

(TDF) is a functional entity that provides ADC functionality using packet inspection. It reports detected applications to the PCRF. The OCS is a credit management system for prepaid charging. The PCRF interacts with the OCS to check out credit and report credit status.

In the context of autonomic management, all the SPR, UDR, AF, RCAF, TDF and OCS provide dedicated monitoring and subscription related information to the PCRF required to make PCC decisions. The PCRF provides the PCC rule to the PCEF.

D. QoS management in GANA network level

Fig.4 illustrates the mapping of QoS management functionality onto GANA network level.
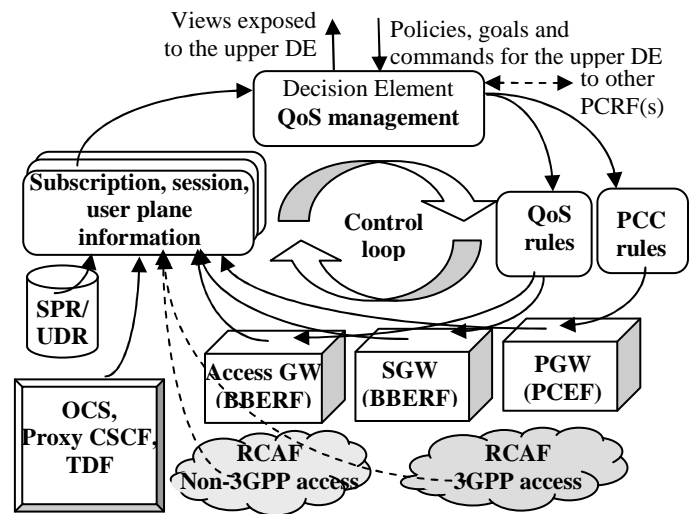
Fig.4 Mapping of QoS management functionality onto GANA network-level

In EPS, the decision logic for QoS management is distributed between different nodes. The Serving gateway (SGW) and Packet Data Network Gateway (PGW) handle the IP packets flowing to and from the mobile devices. The SGW implements the decision logic for mobility between 3GPP access and LTE (Long Term Evolution). The PGW is the point of interconnection to external IP networks and in addition to IP address allocation it implements the decision logic for charging, packet filtering, and policy-based control of user specific IP flows.

EPS supports different mobility protocols depending on which access technology is used. GPRS Tunneling Protocol (GTP) or Mobile IP-based protocols may be used between SGW and PGW. When GTP is used between SGW and PGW, the bearers are terminated in the PGW, which can use bearer procedures to control EPS bearers. This model is known as "on-path" model because the QoS/bearer signaling using GTP follows the same path as the user plane. In this model, the PCRF controls the QoS by providing the QoS information to PCEF in the PGW. When a Mobile IP-based protocol is used towards PGW, the decision logic for QoS and bearer reservation procedures is implemented closer to the Radio Access Network and PGW is not involved in policy control.

For 3GPP family of accesses, these procedures take place in the SGW. For other types of access, the Access GW in the access network is involved in policy control. This model is known as "off-path" model. In the "off-path" model the PGW is involved in charging control.

For IP Multimedia Subsystem (IMS), the AF corresponds to the Proxy Call Session Control Function (CSCF). For non-IMS service, the AF may be a video streaming server or Open Service Access/ Parlay X gateway.

For roaming scenarios, the Home PCRF located in the subscriber home network is not allowed to directly control a policy enforcement entity in the visited network that serves the roaming subscriber. Control of allowed services and the authorization of EPS bearers are always handled by the Home PCRF. For roaming scenarios, the Visited PCRF may accept or reject, but not change, policy decisions coming from the home network. This allows the visited operator to control the usage of EPS bearers in its access network.

IV. FUTURE ENHANCEMENTS OF QoS MANAGEMENT IN EPS

Current research work studies a possible enhanced functionality of PCC related to interworking with the Access Network Discovery and Selection Function (ANDSF) in EPS [7], [8], [9]. The ANDSF contains data management and control functionality and allows the network operator to define policies that prioritize between different access technologies if several non-3GPP access networks are available (3GPP TS 23.402). It can provide the mobile device with three types of information: access network discovery and selection information, intersystem mobility policies, and intersystem routing policies. The ANDSF may interact with policy and charging control to assist the resource reservation in re-selection and handover procedures.

In the GANA network level, the autonomic behavior of QoS management may be extended as shown in Fig.5, where Home Subscriber Server is a subscriber database.
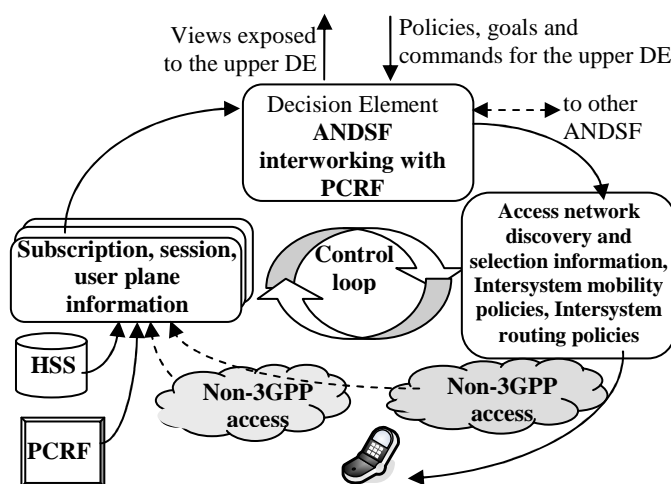

Fig.5 Abstract Autonomic Networked System with interaction between access network discovery and selection function and policy and charging control

V. CONCLUSION

The paper presents a study on autonomous behaviour of QoS management mechanisms in the Evolved Packet System. The standardized policy and charging control functionality is analysed and autonomic characteristics are abstracted in the four levels of the Generic Autonomic Network Architecture.

In the protocol level, the Diameter protocol used for PCC signalling exposes autonomicity during selection/discovering of the policy and charging control function. In the function level, different QoS management functions like policy control and charging control have autonomic characteristics in making decisions on authorization of EPS bearers. Combining different QoS management functions allows abstraction in a system level. The overall QoS management mechanism which considers the distributed network functionality exposes self-manageability in the network level. Further optimization of autonomic characteristics of QoS management in the network level may be achieved by interworking with access network discovery and detection functions.

REFERENCES

[1] W. Wendong, Q. Qinglei, G. Xiangyang, H Yannan, Q. Xirong. "Autonomic QoS management mechanism in Software Defined Network", *Communications*, China, vol.11, issue 7, pp.13-23, 2014.

[2] M. Zouari, C. Diop, E. Exposito. "Multilevel and Coordinated Self-management in Autonomic Systems based on Service Bus", *Journal of Universal Computer Science*, vol. 20, no. 3, pp.431-460, 2014.

[3] N. Huber. "Autonomic Performance-Aware Resource Management in Dynamic IT Service Infrastructures", dissertation, 2014, Available at:http://www.uni-wuerzburg.de/fileadmin/10030200/ user_upload/descartes/research/phdthesis_submission_nikolaus_ huber.pdf

[4] R. M. da Silva Bezerra, J. S. B. Martinis, Scalability issues in network Self-Management: a Pertitioning Approach towards Scalable Autonomic Management Computation, *International Journal of Innovative Computing, Information and Control*, vol.10, no.6, pp. 2143-2156, 2014.

[5] O. Ahmed, Z. Choukair. "QOS-Apcvs: An Enhanced EPS-IMS PCC Architecture Proposal to Improve Mobile Service Dependability", Electrical & Computer Engineering: An International Journal (ECIJ). vol. 2, no. 4, pp.29-36, 2013.

[6] G. Gómez, J. Lorca, R. García, Q. Pérez. "Towards a QoE-Driven Resource Control in LTE and LTE-A Networks", *Journal of Computer Networks and Communications,* vol.2013, 2013, pp.1-15.

[7] S. Frei, W. Fuhrmann, A. Rinkel, B. Ghita. "Improvements to Inter-system Handover in the EPC Environment", NTMS'2011, Conference proceedings, pp 1-5, Paris, France, 2011.

[8] Lonsethagen, H. et al. "Multilink network architecture", Report, Celtic project CP5-013, "MARCH – Multilink architecture for multiplay services", http://projects.celtic-initiative.org/march/ march/UserFiles/file/CP5-013-MARCH-D5_2-final.pdf, 2012.

[9] I. Atanasov, E. Pencheva, R. Dobrev. "Study on Access Network Discovery and Selection Based on Intelligent Resource Management in EPS", International Scientific Congress, technical University of Varna, Conference Proceedings, Varna, Bulgaria, pp.166-171, 2012.

Application of an Econometric Method for Forecasting Cellular Mobile Traffic

Suzana Miladic[1], Goran Markovic[2] and Valentina Radojicic[3]

Abstract – **This paper presents an econometric traffic forecasting model that is applied to the cellular mobile network of one telecom operator in the Republic of Srpska (BiH). As a suitable metric of traffic volume we chose the total monthly traffic minutes. The model deals with several explanatory variables affecting the changes to the traffic volume. These variables are measured by using available historical data and by applying regression analysis for estimation the model parameters. Through the numerical results the short-term forecasted results of cellular mobile traffic are given.**

Keywords – **Forecasting, econometric model, cellular mobile traffic.**

I. INTRODUCTION

A correct forecast of the future traffic needs, based on reliable and relevant input historical data, is the basis for the planning and upgrading of telecommunication networks. Generally, forecasting in the field of telecommunications involves estimation the number of subscribers as well as the traffic volume.

Forecasting as a process involves several steps: first it is necessary to define the problem or purpose of forecasting, then to collect the data, choose the suitable method of forecasting, estimation of necessary parameters and finally make the analysis and perform necessary calculations. Today, changes in market dynamics, especially rapid growth of the Internet and the number of mobile telephony users encourages telecommunications operators to respond the technological challenges. Consequently, the network planning methodology has to consider the market and demand forecasting in the first stage.

If the forecast is not reliable as a result may appear insufficient or excessive dimensioning of certain network capacity, resulting in operator losses and reducing the quality of services for the end users [1]. One of the most important parameters for the process of market and demand forecasting is the forecasted network traffic.

There are many forecasting methods. Quantitative or statistical and qualitative or non-statistical methods represent one possible classification. Quantitative methods process the

data using statistical principles and have a numerical output while qualitative methods are based on other principles such as extrapolation, reasoning, etc., so the output may not be numerical. This paper deals with one quantitative traffic forecasting method and its application on real traffic data.

The methods of traffic forecasting are defined by the ITU-T recommendations E.506 and E.507 [2], [3]. Some important methods include the following: curve fitting models (linear, parabolic, exponential, logistic, Gompertz trend), smoothing models (moving average, Holt, Holt-Winter), autoregression models, ARIMA (*Autoregressive Integrated Moving Average)*, regression, Kalman and econometric models. Models for total traffic forecasting based on increasing the number of subscribers and models for point-to-point traffic forecasting are given in [4]. Forecasting network traffic by means of neural networks and linear models is presented in [5].

The paper is organized as follows. In the second section we formulate the problem, describe the forecasting model and discuss its parameters. The model is applied to real data in the third section, and the obtained traffic forecast for the cellular mobile operator is presented. Concluding remarks and further work are given in the forth section.

II. THE FORECASTING MODEL

Inspired by [2], [3] and [6] we selected an econometric model to perform forecasts of total mobile telephony traffic for one telecom operator in the Republic of Srpska (BiH). Econometric model defines a set of explaining variables from which the unit of interest (traffic) is calculated according to given mathematical model. The model parameters have been determined through regression analysis. As a suitable metric of the traffic volume we chose monthly traffic minutes (denoted by *T*) which refers to the number of subscribers times the minutes of use per subscriber and month (denoted by MoU). The relevant historical data and the number of observations required for this model are obtained from sources [7] and [8].

A. Modelling the variables

The model considers following three independent or explanatory variables: the spending power *S* in the market (the market is specified over country area or network traffic area), the actual service cost *C* and the market penetration *P*. These variables affect the changes to the traffic volume *T,* which represents the dependent variable [6].

It is shown that the demand for mobile telecommunications services appears to be relatively elastic. It means that subscribers are highly sensitive to service price, which has

[1]Suzana Miladic (Ph.D student) is with the Faculty of Transport and Traffic Engineering at University of East Sarajevo, Bosnia and Herzegovina, e-mail: miladics@hotmail.com

[2]Goran Markovic is with the Faculty of Transport and Traffic Engineering at University of Belgrade, Serbia, e-mail: g.markovic@sf.bg.ac.rs

[3]Valentina Radojicic is with the Faculty of Transport and Traffic Engineering at University of Belgrade, Serbia, e-mail: valentin@sf.bg.ac.rs

direct impact on traffic volume. Obviously, the traffic volume depends on the spending power of population. Usually, as a measure of a spending power, the Gross National Product (GDP) per capita is used.

The costs of mobile service, C_t includes various parameters such as: subscription fees, usage tariffs, the cost of a terminal, etc. In this paper, we made a simplified assumption that the operator's revenues of telephony traffic service are based on paid traffic per minute. The structure of the traffic includes pre-paid, post-paid and VPN mobile users. The revenues are measured as average revenues per user and month (ARPU) and have been expressed in national monetary units (BAM).

The mobile service penetration rate, α, is the percentage of the relevant country's population that has purchased a mobile service during the time under study. Service life cycles may go through several stages: Introduction, Growth and Maturity stage. After the rapid sales during the Growth stage, the market starts to be saturated as there are fewer new subscribers. When service sales reach peak values and market becomes saturated, the penetration rate becomes very low. It could be modeled by the penetration function $P(\alpha)$ as follows [6]:

$$P(\alpha) = \left[1 - \cos(\alpha^2 \pi)\right]/2 \qquad (1)$$

B. Mathematical background

Considering all variables indicated above, the main form of the traffic forecasting model could be described as follows [6]:

$$T_t = e^{b_0} \cdot S_{t-1}^{b_1} \cdot C_t^{b_2} \cdot P(\alpha_t)^{b_3} \qquad (2)$$

Transforming Eq. 2 into linear form, the model thus becomes:

$$\ln T_t = b_0 + b_1 \ln S_{t-1} + b_2 \ln C_t + b_3 \ln P(\alpha_t) \qquad (3)$$

Finally, if we introduce shift variables, it is obtained:

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 \qquad (4)$$

The variables are indexed by time, t, and the model parameters b_0, b_1, b_2, b_3 are estimated using the ordinary least square (OLS) technique. The parameters b_1, b_2, b_3 can be interpreted as elasticity parameters indicating the impact of each explanatory variable or how much they attribute to the variation of dependent variable. This could be shown by index of determination R^2 and adjusted coefficient of determination R^2 (adj) which provides better estimate of the actual value of R^2, if we talk about small sample. These coefficients range from [-1,1]. The closer the marginal values, better the model is. The multiple linear regressions have been applied here, since the model is nonlinear in variables (because of the logarithms) and linear in parameters.

The model provides forecasts of the dependent variable value based on the changes in the values of explanatory variables.

III. NUMERICAL RESULTS

In this section, we have applied the model described above to the mobile network of one telecom operator in the Republic of Srpska. The historical data considering the GDP values and population size are given in Table I. The relevant statistical data for the chosen mobile operator obtained from their annual reports, are given in Table II. MoU and ARPU are given in real terms while the total monthly mobile telephony traffic (minutes) and subscribers are given in millions.

TABLE I
RELEVANT DATA FOR THE REPUBLIC OF SRPSKA [8]

Year	Republic of Srpska	
	GDP	Population
2009	5 739	1 435 179
2010	5 805	1 433 038
2011	6 073	1 429 668
2012	6 006	1 429 290
2013	6 146	1 425 549

TABLE II
RELEVANT DATA FOR THE CHOSEN OPERATOR [7]

Year	OPERATOR				
	Traffic	MoU	ARPU	Sub.	P(α)
2009	50.74	41.2	15.4	1.231	0.78
2010	59.94	44.6	16.2	1.344	0.82
2011	61.15	44.4	15.3	1.377	0.80
2012	72.01	50.9	13.7	1.415	0.82
2013	77.98	54.34	13.4	1.435	1

We used the direct method to determine the model parameters. Since we made assumption that traffic variable T depends on several significant factors, it is necessary to forecast these factors (explanatory variables), firstly. Considering the historical data and by applying the OLS technique, the regression equations could be written, which are the basis for computing these values for the next few years (Table III). Therefore, we have: $x_1 = 5.650 + 0.101t$; $x_2 = 16.75 - 0.65t$; $x_3 = 0.712 + 0.044t$.

TABLE III
FORECASTED VALUES OF EXPLANATORY VARIABLES

Year	t_i	GDP (x10³)	ARPU	P(α)
2014	6	6.256	12.85	0.98
2015	7	6.357	12.20	1.02
2016	8	6.458	11.55	1.06
2017	9	6.559	10.90	1.11

Forecasting of monthly traffic minutes for mobile telephony, considering the forecasted values of significant factors, is done by using the Eq. 4, whereby the shift variables are: $lnT=y$; $lnS=x_1$; $lnC=x_2$ and $lnP(\alpha)=x_3$. The parameters b_0, b_1, b_2, b_3 are estimated using the system of normal equations:

$$nb_0 + b_1 \sum_{i=1}^{n} x_{1i} + b_2 \sum_{i=1}^{n} x_{2i} + b_3 \sum_{i=1}^{n} x_{3i} = \sum_{i=1}^{n} y_i \quad (5)$$

$$b_0 \sum_{i=1}^{n} x_{1i} + b_1 \sum_{i=1}^{n} x_{1i}^2 + b_2 \sum_{i=1}^{n} x_{1i}x_{2i} + b_3 \sum_{i=1}^{n} x_{1i}x_{3i} = \sum_{i=1}^{n} x_{1i}y_i \quad (6)$$

$$b_0 \sum_{i=1}^{n} x_{2i} + b_1 \sum_{i=1}^{n} x_{1i}x_{2i} + b_2 \sum_{i=1}^{n} x_{2i}^2 + b_3 \sum_{i=1}^{n} x_{2i}x_{3i} = \sum_{i=1}^{n} x_{2i}y_i \quad (7)$$

$$b_0 \sum_{i=1}^{n} x_{3i} + b_1 \sum_{i=1}^{n} x_{1i}x_{3i} + b_2 \sum_{i=1}^{n} x_{2i}x_{3i} + b_3 \sum_{i=1}^{n} x_{3i}^2 = \sum_{i=1}^{n} x_{3i}y_i \quad (8)$$

After deduction of sums and incorporating them into the equations, we have got:

$$5b_0 + 8,92b_1 + 13,47b_2 - 0,87b_3 = 20,77 \quad (9)$$

$$8,92b_0 + 15,91b_1 + 24,02b_2 - 1,65b_3 = 37,08 \quad (10)$$

$$13,47b_0 + 24,02b_1 + 36,3b_2 - 2,36b_3 = 55,91 \quad (11)$$

$$-0,87b_0 - 1,65b_1 - 2,36b_2 + 0,19b_3 = -3,56 \quad (12)$$

Solving this system of equations, we have determined the necessarily model parameters. Using the SPSS Statistics 17.0 tool, we have calculated the coefficient of determination R^2 and adjusted coefficient of determination R^2 (adj). The obtained results are given in Table IV.

TABLE IV
FITTED MODEL PARAMETERS

Republic of Srpska Operator	b_0 Const.	b_1 S_{t-1}	b_2 C_t	b_3 $P(\alpha_t)$	R^2	R^2(adj)
	13.568	0.077	-3.539	0.107	0.967	0.942

In terms of the model parameters, the first one, b_1 shows that, if the GDP is increased for 1 % the traffic will increase for 0.077 million of minutes. The second one, b_2 shows that, if the ARPU is changed for 1 % the traffic will change for 3.539 %. This value is inversely proportional to the traffic value and therefore the Pearson correlation is negative. Finally, the third one, b_3 shows that, if the penetration effect is higher for 1 %, the traffic value will be higher, too. Constant b_0 shows the value of T if there were no independent variables.

It could be noted that R^2 (adj) is more than satisfactory for the presented model and explain the 94,2 % of the variation in total demand (traffic) concerning with independent variables (elasticities denoted by b_1, b_2, b_3).

In the next step, we used the obtained parameters to forecast the monthly traffic minutes, using the historical data (2009-2013) to fit the model. If we apply this model for forecasting the monthly traffic minutes of mobile telephony, for the years 2014, 2015 and 2016, we will get the following values (in millions of minutes) 98.54; 126.17 and 153.95, respectively. Fig. 1 summarizes the results.
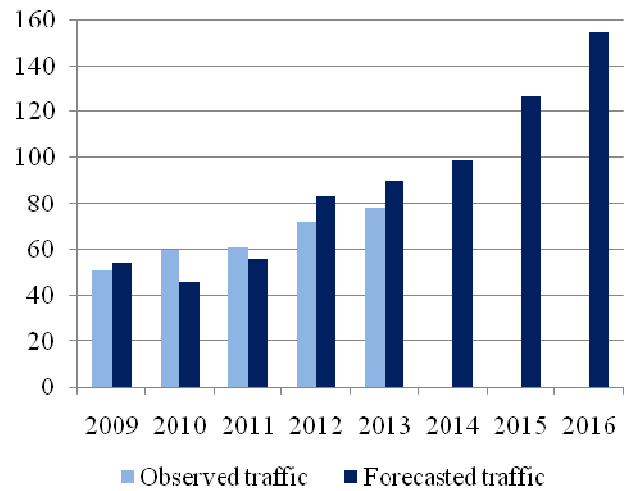

Fig. 1. Forecasted and observed monthly traffic volume (in millions of minutes)

It could be seen, that the forecasted traffic values are very similar to the obseved traffic values, which means that the proposed model can be used for forecasting the cellular mobile traffic. For a number of observations, using SPSS, we have determined the regression line with a deviations, shown in Fig. 2.
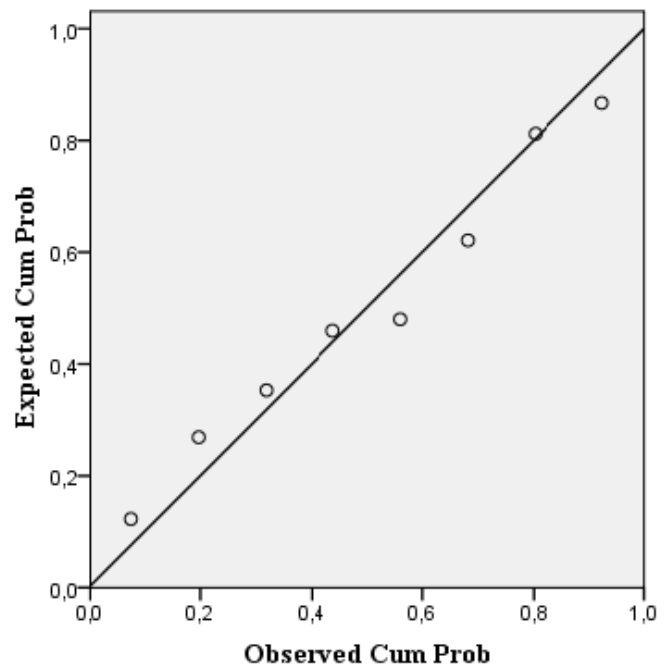

Fig. 2. The plot area of a regression

The observed data from 2009-2013 show that the revenues of the operator regarding to mobile telephony, since 2011, are slightly reduced due to the presence of smart phones on the mobile market. The forecasted values for 2014-2016 show that the volume of mobile cellular traffic is constantly increasing since the number of subscribers is increased every day.

IV. CONCLUSION

Traffic forecasting is one of the main tasks that precedes to telecommunication network dimensioning or upgrading and management of network resources. This paper presents one econometric model for traffic forecasting. It demonstrates that it is possible to identify a few essential variables which could be used to obtain the reliable forecast of the cellular mobile traffic. The spending power, service costs and market penetration are certainly the main factors affecting the changes to the traffic volume.

The model gives accurate forecast and could be used for some other traffic types, for example: Internet traffic, fixed telephony etc. Besides this, future work could include the data from more operators, since BiH has three dominant operators which provide mobile telephony service. At this moment, only the data for one operator were publicly available.

Additionally, it could be considered more operators from different countries, in order to do some comparative analysis of the influencing factors on the predicted traffic value.

It is important to mention, that the quality of obtained results depends on the reliable historical data, the right choice of forecasting model as well as the precise estimation of relevant model parameters.

ACKNOWLEDGEMENT

REFERENCES

[1] V. Radojicic, A. Kostic-Ljubisavljevic, "Telecommunication network design under traffic forecast uncertainty", PosTel, Conference Proceedings, pp. 281-291, Belgrade, Serbia, 2005.

[2] ITU-T Recommendation E.506 (rev.1): "Forecasting international traffic", ITU-T Geneva, 1992.

[3] ITU-T Recommendation E.507: "Models for forecasting international traffic", ITU-T Geneva, 1993.

[4] H. Leijon, "Basic forecasting theories: A brief introduction", ITU, 1998.

[5] I. Klevecka, "Forecasting network traffic: a comparison of neural networks and linear models", Proceedings of the 9th International Conference (RelStat'09), pp. 170-177, Riga, Latvia, 2009.

[6] A. Arvidsson et al., "Forecasting cellular traffic: an econometric approach", International Teletraffic Congress, Beijing, Blekinge Institute of Technology, 2006 available on: http://www.i-teletraffic.org/fileadmin/ITCBibDatabase/2005/arvidsson05.pdf

[7] Mtel, *Annual Report*, Banja Luka, 2009-2013, available at: http://www.mtel.ba/article/819

[8] Institute of statistics, *Statistical bulletin: National Accounts,* no. 7, Banja Luka, 2014, available at: http://www2.rzs.rs.ba/static/uploads/bilteni/nacionalni_racuni/Bilten_Nacionalni_Racuni_7.pdf

Content and Service Provider Interconnection Charging based on Revenue-Sharing Concept

Branka Mikavica[1], Aleksandra Kostić-Ljubisavljević[2], Vesna Radonjić Đogatović[3]

Abstract – **In this paper the application of Revenue-Sharing concept for provider's interconnection is analysed. Two types of contracts are shown, static and dynamic. The goal of our research is to consider possible market share enlargement according to price reduction. Relevant parameters in this research are provider's reputation factor and customer's willingness-to-pay.**

Keywords – ***Revenue-Sharing, contract, price sensitivity, reputation factor***

I. INTRODUCTION

The primary aim of Internet was to enable communication between remote hosts. Broadband access development transformed allows customers to connect with all available content (e.g. customers search information over Google, watch videos on YouTube etc.) [1]. Content provisioning involves many technologies, such as Content Delivery Networks (CDNs) and peer-to-peer networks; and different undertaking, such as Service Providers (SPs) and Content Providers (CPs). CP handles the distribution of online content including blogs, videos, music or files. In order to obtain connectivity and possibility for customers to purchase specific content from Content Providers, interconnection between Service Providers and Content Providers is necessary. Selection of proper charging strategy in such interconnection agreement is of great importance for providers. Improvement of market position and gathering enhanced revenues are often conflicted goals of providers' business strategy requiring compromises. Hence, providers negotiate in order to achieve satisfactory solution for involved parties. Although Cost-Based concept is widely recommended as the most appropriate charging concept, its complexity introduces the necessity for new, more feasible solutions. Revenue-Sharing concept is one of the alternative charging concepts, characterized with operational simplicity and possibility of rebalancing the returns of the providers when retail prices are distorted. In this paper, we address interconnection between Content Providers and Service Providers through Revenue-Sharing concept with the aim to improve providers' market position in long term by increasing customers' willingness-to-pay and their incentives to obtain specific content.

The paper is organized as follows. After introductory remarks, a brief literature review is given in Section II.

Section III presents problem statement where two approaches of Revenue-Sharing concept are introduced. Revenues of content and service providers are observed considering content popularity, customers' willingness-to-pay and providers' reputation factor. Numerical example is presented in Section IV. Concluding remarks are presented in Section V.

II. LITERATURE REVIEW

Determination of optimal Revenue-Sharing contracts between peering providers is presented in [2], for both symmetric and asymmetric retail prices. Relations among Internet Service Providers under network neutrality debate according to Revenue-Sharing peering and transit agreements are observed in [3]. Possibility of reorganization of revenue flows through an invoicing process that may benefit the mobile network operator more than the other involved parties is considered in [4]. Revenue-Sharing interconnection charging based on Service Level Agreement is observed in [5]. Game theory, especially Stackelberg setting, is employed addressing hierarchical decision structure. Amazon Internet portal has been investigating in that purposes in [6]. Non-cooperative and cooperative revenue sharing policies have been addressed in [7]. Non-cooperative policy might lead to unfair distribution of revenues among the providers. However, a charging strategy based on the weighted proportional fairness criterion stimulates cooperation among providers and can achieve higher profits for all involved providers. The effectiveness of Revenue-Sharing over other arrangements such as wholesale price contracts is observed in [8]. Revenue-Sharing as a charging strategy is widely investigated for supply chain coordination. Two-stage telecommunication supply chain under technology dependent stochastic demand has been analysed in [9], where Revenue-Sharing concept has been suggested.

III. PROBLEM STATEMENT

Let us consider full interconnection between several CPs and several SPs as shown on Fig. 1. Denote a set of CPs as $M = \{CP_1, CP_2, ..., CP_m\}$, a set of SPs as $N = \{SP_1, SP_2, ..., SP_n\}$, and a set of contents offered by all CPs as $Q = \{q_1, q_2, ..., q_s\}$. We assume each CP is connected with all SPs. They offer a subset of contents Q_i such that

$$\bigcup_{i=1}^{m} Q_i \subseteq Q, \ Q_i \cap Q_j = \varnothing, \forall (i,j) \in (1,...,m), i \neq j \quad (1)$$

For each content $q_k \in Q$ we define popularity factor,

[1]Branka Mikavica is with the Faculty of Transport and Traffic Engineering at University of Belgrade, Vojvode Stepe 305, 11000 Belgrade, Serbia, E-mail: b.mikavica@sf.bg.ac.rs

[2]Aleksanda Kostić-Ljubisavljević is with the Faculty of Transport and Traffic Engineering at University of Belgrade, Vojvode Stepe 305, 11000 Belgrade, Serbia, E-mail: a.kostic@sf.bg.ac.rs

[3]Vesna Radonjić Đogatović is with the Faculty of Transport and Traffic Engineering at University of Belgrade, Vojvode Stepe 305, 11000 Belgrade, Serbia, E-mail: v.radonjic@sf.bg.ac.rs

$$\Psi_{q_k} = \frac{\lambda_{q_k}}{\sum_{k=1}^{s} \lambda_{q_k}}, \Psi_{q_k} \in (0,1) \qquad (2)$$

where λ_{q_k} presents estimated number of requests for content q_k. Total number of estimated requests for all contents is expressed as $\sum_{k=1}^{s} \lambda_{q_k}$.
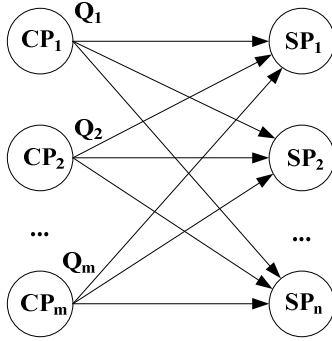

Fig. 1. Content and Service Provider Interconnection

In order to obtain appropriate Revenue-Sharing contracts between Content and Service Providers, such that provide market stability and ensure customers protection by reducing incentives of SPs to increase their retail prices, we analyse two types of contracts, static and dynamic. Static Revenue-Sharing concept defines fixed, predetermined, share of revenue that SPs obtain from provisioning CP's content on retail market. The aim of dynamic Revenue-Sharing contract is to stimulate providers to improve their market share, to strengthen market position and thus increase their revenue, rather than by price enhancement. Assume that total customer population, denoted as X, is fixed during observed time interval. Relevant parameters in these contracts are customers' willingness-to-pay specific content $q_k \in Q$ by the SP$_j$' retail price, $p_j^{q_k}$, and SPs' reputation factor. Customers' willingness-to-pay refers to the share of customers ready to pay given service/content by the retail price $p_j^{q_k}$. Regarding to the price sensitivity, two types of customers can be distinguished, more and less price sensitive. Thus, customers' willingness-to-pay mathematically can be expressed as follows:

$$W_j^{q_k} = \begin{cases} \dfrac{\alpha - \beta\left(p_j^{q_k}\right)}{N_{q_k}}, \text{ for more sensitive customers} \\ \dfrac{\rho\left(p_j^{q_k}\right)^{-\varphi}}{N_{q_k}}, \text{ for less sensitive customers} \end{cases} \qquad (3)$$

Parameters in equation (3), α, β, ρ and φ depend on customers' social and economic status and content substitutability and popularity, as well. Willingness-to-pay is inversely proportional to the total number of SPs that offer content q_k at the market, denoted as N_{q_k}.

Content price reduction leads to enhancement of customers' willingness-to-pay for both types of customers. As a result, number of SPs' customers will increase. Thus, one of the most important providers' business goals, improved market share, will be satisfied. Another important parameter in proposed models is SPs' reputation factor. SP$_j$'s reputation factor is denoted by r_j, $r_j \in (0,1)$. This value is established on the basis of long term business existence of SP$_j$ on the market, and $\sum_{j=1}^{n} r_j = 1$. This means that value for SP's reputation factor is normalized and sum of the values for all SPs equals 1. We assume that reputation factor of all SPs on the observed market is known. Higher reputation factor is reason why a number of customers are willing to pay higher price for a specific content. We assume that Quality of Service (QoS) requirements are satisfied by all SPs. Revenue of SP$_j$ from provisioning content to the customers and CP$_i$'s revenue in accordance with the static revenue sharing contract can be, respectively, written as follows:

$$R_j^S = \sum_{i=1}^{m} \sum_{q_k \in Q_i} \left(1 - \Phi_{i,j}^{S,q_k}\right)\left(p_j^{q_k} \Psi_{q_k} r_j W_j^{q_k} X\right) \qquad (4)$$

$$R_i^S = \sum_{j=1}^{n} \sum_{q_k \in Q_i} \Phi_{i,j}^{S,q_k} p_j^{q_k} \Psi_{q_k} r_j W_j^{q_k} X \qquad (5)$$

$\Phi_{i,j}^{S,q_k}, \Phi_{i,j}^{S,q_k} \in (0,1)$ presents fixed portion of generated revenue that SP$_j$, by the contract, pays to CP$_i$ for provisioning content q_k under static Revenue-Sharing contract.

Dynamic Revenue-Sharing contract defines flexible portion of revenue that SP pays to the CP, depending on SP's retail price. SP$_j$'s revenue share paid to CP$_i$ under dynamic revenue sharing contract can be expressed as follows:

$$\Phi_{i,j}^{D,q_k} = \left(1 + \Delta p_j^{q_k}\right) \cdot \Phi_{i,j}^{S,q_k} \qquad (6)$$

In equation (4), $\Delta p_j^{q_k}$ presents variation of SP$_j$'s retail price $p_j^{q_k}$. Thus, revenue of SP$_j$'s from provisioning content to the customers and CP$_i$'s revenue under dynamic Revenue-Sharing contract can be, respectively, written as follows:

$$R_j^D = \sum_{i=1}^{m} \sum_{q_k \in Q_i} \left(1 - \Phi_{i,j}^{D,q_k}\right)\left(p_j^{q_k} \Psi_{q_k} r_j W_j^{q_k} X\right) \qquad (7)$$

$$R_i^D = \sum_{j=1}^{n} \sum_{q_k \in Q_i} \Phi_{i,j}^{D,q_k} p_j^{q_k} \Psi_{q_k} r_j W_j^{q_k} X \qquad (8)$$

IV. NUMERICAL EXAMPLE

Let us consider the situation in when Revenue-Sharing concept is applied as a relevant interconnection agreement between CP and SPs. There are two CPs and two SPs at the market. Each CP offers single content differentiated from the one offered by other CP. Providers negotiate in order to

determine revenue share on the Revenue-Sharing basis. We assume that one SP is a new entrant at the observed market. It has having lower reputation factor compared to the other SP. Regarding to the customers' price sensitivity, two scenarios are assumed. The first scenario refers to the situation when only one SP increases retail price for more popular content and services related to that content, while other SP remains its price for the same content at the same level. The second scenario considers price reduction by the SP_1, offering higher price for less popular content. This situation is common when promotions and discounts are being applied. Values for parameters in the relation for customers' willingness-to-pay are specified according to assumed market situation. We assumed that values of relevant factors for calculation of revenues are the following: $X = 200000$, $\Phi_{i,j}^{S,q_k} = \Phi^S = 0.5$, $p_1^{q_1} = 100$, $p_1^{q_2} = 110$, $p_2^{q_1} = 90$, $p_2^{q_2} = 100$, $r_1 = 0.65$, $r_2 = 0.35$, $\Psi_{q_1} = 0.3$, $\Psi_{q_2} = 0.7$. All obtained revenues are expressed in monetary units [MU].

Fig. 2. presents revenue of SP_2 obtained applying static and dynamic Revenue-Sharing concepts for more price sensitive customers. Revenue of SP_2 obtained applying static and dynamic Revenue-Sharing concepts for less price sensitive customers is shown on Fig. 3.
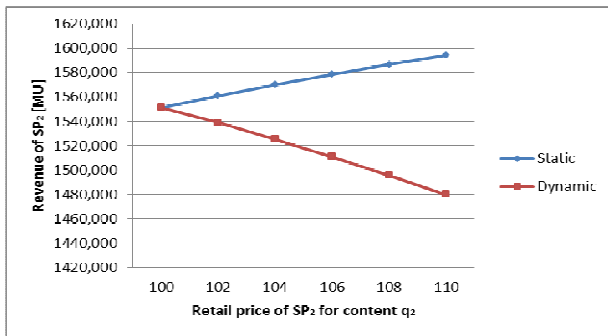

Fig. 2. Revenue of SP_2 under static and dynamic Revenue-Sharing concept for more price sensitive customers

Static Revenue-Sharing concept increases SP_2 revenue when retail price for more popular content increases, while dynamic reduces for both more and less price sensitive customers, as shown on Fig. 3. Considering dynamic revenue sharing contract, SP_2 has no incentive to increase its retail price. It can be noted that revenue approximately remains at the same level for less price sensitive customers for both more
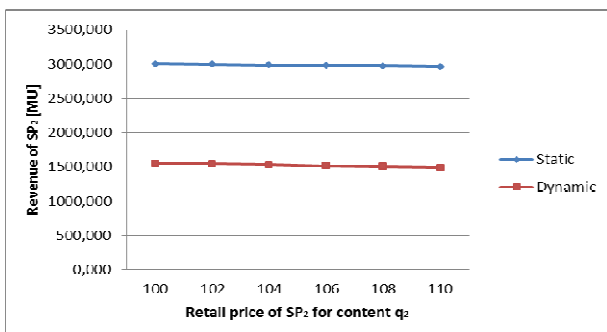

Fig. 3. Revenue of SP_2 under static and dynamic Revenue-Sharing concept for less price sensitive customers
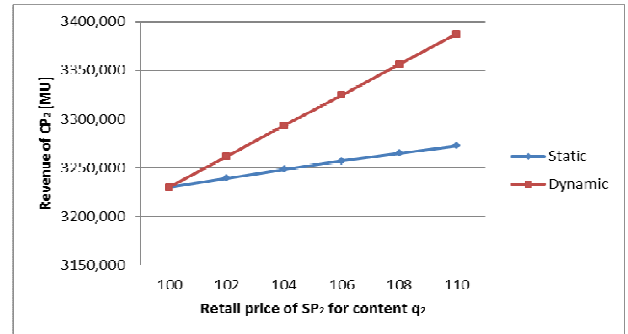
and less price sensitive customers.

Fig. 4. Revenue of CP2 under static and dynamic Revenue-Sharing concept for more price sensitive customers

Fig. 4. presents revenue of CP_2 offering more popular content q_2, when SP_2 increases its retail price for that content, for more price sensitive customers. Obtained results for less price sensitive customers are shown in Fig. 5.
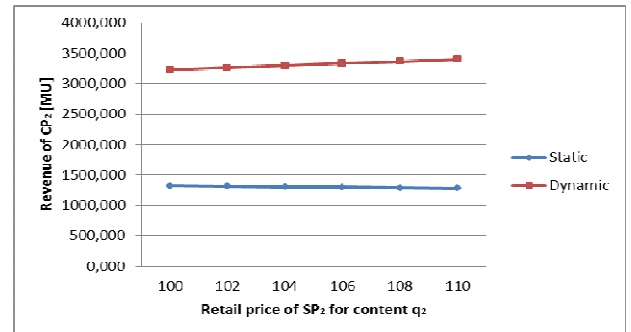

Fig. 5. Revenue of CP2 under static and dynamic Revenue-Sharing concept for less price sensitive customers

For more price sensitive customers, dynamic revenue sharing contract enhance CP_2's revenue. However, for less price sensitive customers, revenue is nearly at the same level, but high above revenue under static Revenue-Sharing concept.

Situation when retail price for less popular content, q_1, decreases in order to attracts more customers according to static and dynamic Revenue-Sharing concept for more price sensitive customers is shown on Fig. 6. SP_1 decreases its retail price for content q_1 by the level of SP_2's retail price. Since content q_1 is less popular, SP_1 is looking for appropriate concept that will increase customers' interest for that content, but to increase its revenue, as well. Dynamic Revenue-
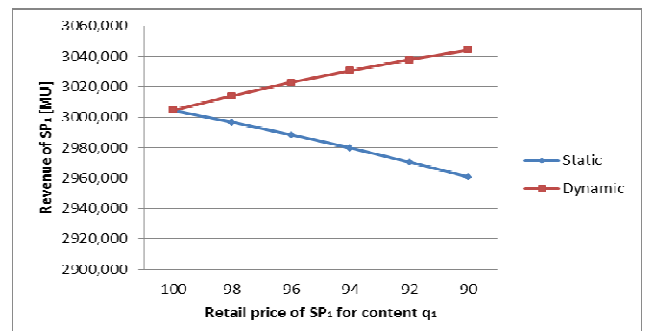

Fig. 6. Revenue of SP1 under static and dynamic Revenue-Sharing concept for more price sensitive customers

Sharing concept increases SP_1's incentives to decrease retail price for less popular content q_1, thus enhancing its revenue.

Fig. 7. depicts situation when SP1 decreases retail price for less popular content, q_1, for less price sensitive customers. Obtained results are nearly the same as for more price sensitive customers.
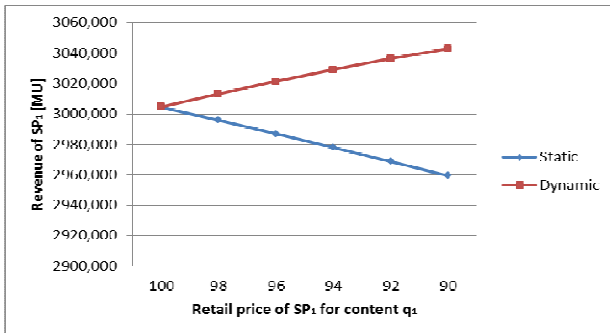

Fig. 7. Revenue of SP1 under static and dynamic Revenue-Sharing concept for less price sensitive customers

Revenue of CP_1, offering less popular content, when SP_1 decreases its retail price for content q_1 is shown on Fig. 8. Both static and dynamic Revenue-Sharing concepts decrease revenue of CP_1. Regarding static concept, revenue slowly decays in comparison with dynamic.
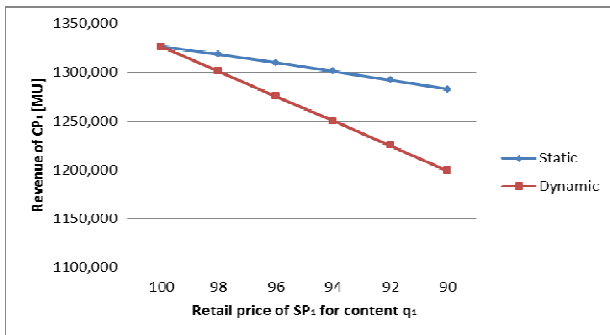

Fig. 8. Revenue of SP1 under static and dynamic Revenue-Sharing concept for more price sensitive customers

Fig. 9. shows revenue of CP_1 when SP1 decreases its retail price for less popular content, in the case of less price sensitive customers. Retail prices enhancement ensures greater revenue for CP. However, it leads to reduction of customers' willingness-to-pay and reduction of market share in long term.
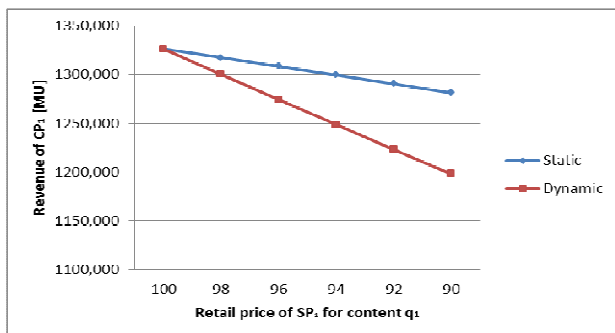

Fig. 9. Revenue of SP_1 under static and dynamic Revenue-Sharing concept for less price sensitive customers

V. CONCLUSION

This paper analyses possibility of application of Revenue-Sharing concept for charging interconnection between Content Providers and Service Providers on the given market. Two types of this concept are observed, static and dynamic. Static Revenue-Sharing concept establishes fixed portion of revenue that Service Provider pays to Content Provider. Dynamic Revenue-Sharing concept depends on retail price and involves fixed portion of revenue that SP pays to CP, but involves variable part which reflects retail price variation. Depending on service and hence, on content popularity, providers' reputation factor and customers' willingness-to-pay, revenues of specific provider are being analysed. The aim of dynamic revenue sharing contract is to enlarge customer base by price reduction, and thus improve providers' market position. It was shown that proposed dynamic Revenue-Sharing concept enables great incentives to reduce retail prices and increase revenue on the improved market share rather than price enhancement.

ACKNOWLEDGEMENT

REFERENCES

[1] G. Carofiglio, G. Morabito, L. Muscariello, I. Solis, M. Varvello, "From Content Delivery Today to Information Centric Networking", Computer Networks, vol. 57, no. 16, pp. 3116-3127.

[2] J. Y. Kim, Y. Lim, „The Revenue-Sharing Rule for Interconnection Charges", The Japanese Economic Review, vol. 55, no. 3, pp. 298-310, 2004.

[3] P. Coucheney, P. Maille, B. Tuffin, "Network Neutrality Debate and ISP Inter-relations: Traffic Exchange, Revenue Sharing, and Disconnection Threat", Economics Research and Electronic Networking, vol. 15, no. 3, pp. 155-182, 2014.

[4] A. K. Chakravarty, A. S. Werner, "Telecom Service Provider Portal: Revenue Sharing and Outsourcing", European Journal of Operational Research, vol. 215, no. 1, pp. 289-300, 2011.

[5] M. D. Stojanovic, A. M. Kostic-Ljubisavljevic, V. M. Radonjic-Djogatovic, "SLA-controlled interconnection charging in next generation networks", Computer Networks, vol. 57, no. 11, pp. 2374–2394, 2013.

[6] Y. Wang, L. Jiang, Z.J. Shen, "Consignment Contract with Revenue Sharing", Management Science, vol. 50, no. 1, pp. 34-47, 2004.

[7] L. He, J. Walrand, "Pricing and Revenue Sharing Strategies for Internet Service Providers", IEEE Journal on Selected Areas in Communications, vol. 24, no. 5, pp. 942–951, 2006.

[8] G.P. Cachon, M.A. Lariviere, "Supply Chain Coordination with Revenue-Sharing Contracts: Strengths and Limitations", Management Science, vol. 51, no. 1, pp. 30–44, 2005.

[9] E. Canakoglu, T. Bilgic, "Analysis of a Two-Stage Telecommunication Supply Chain with Technology Dependent Demand", European Journal of Operational Research, vol. 177, no. 2, pp. 995-1012, 2007.

Performance Analysis of a Power Saving Mechanism in WLANs

Kiril Kassev

Abstract – **Traditional WLANs consume significant amount of power to contend to the shared wireless medium as well as to periodically broadcast special beacons for identifying themselves to mobile nodes within the service area. Improving energy efficiency can be achieved by adopting power saving mechanism if no data is transferred for a particular amount of time. This paper aims at investigating such a mechanism, employed in WLANs. The performance analysis is carried out by analytical modelling under different service time distributions and inactivity timer values using a semi-Markov process. Numerical results demonstrate the trade-off relationship between energy consumption and cumulative wake-up delay performance.**

Keywords – **Energy consumption, Performance analysis, Power saving, Wireless access networks, WLAN.**

I. INTRODUCTION

Energy consumption in communication networks is a crucial issue, which is paid an increasing attention. The global communication infrastructure comprises various kinds of network domains implemented by different technologies. It is reported that more than 3 % of the worldwide energy consumption is driven by the ICT sector, and it is expected to increase in future [1]. Nowadays wireless communications are fundamental part of the modern networks. Such wireless access networks are widely deployed worldwide in order to meet customers' demands for broadband access to a rich variety of services. Over the years many efforts have been made towards network throughput enhancement, which is related to the energy consumption increase. Energy efficiency in telecommunication networks is a still open issue, which aims at reducing energy consumption while meeting QoS criteria [2], [3]. With the growing popularity of broadband wireless access technologies, research activities are focused on reducing energy consumption in battery-driven mobile devices [4]. Moreover, the worldwide mass deployment of 3G and 4G cellular networks requires adequate actions aimed at energy consumption reduction, since more than 50 % of total energy is consumed by the radio access network [5]. It largely depends on the mobile stations (MSs) density in the area covered by the base station.

Heterogeneous wireless access networks are considered to be a promising solution for supporting dense and hyper-dense traffic loads and providing cost-effective services. Different technologies coexist in the same geographical area and often

Kiril Kassev is with the Faculty of Telecommunications, Technical University of Sofia, 8 Kliment Ohridski Blvd., 1756 Sofia, Bulgaria, E-mail: kmk@tu-sofia.bg

cells of small sizes are installed either indoor or outdoor for better coverage and cellular traffic offload. The latter is an area of great interest and is usually realized by WLANs (Wi-Fi AP) [6].

WLANs are widely available in urban areas either deployed by operators as commercial hotspots or deployed by users for private or business usage. A common weakness is that they consume significant amount of energy as periodically broadcast special beacons for identifying themselves to MSs within the service area [7]. Thus, lack of data activity would lead to energy waste. This paper aims at investigating a power saving mechanism applied to WLANs with respect to traffic characteristics.

The rest of the paper is organized as follows: Section 2 describes the system and the analytical model for performance analysis of a power saving mechanism incorporated in a WLAN. Numerical studies in Section 3 reveal the influence of various parameters on system performance. Some conclusion remarks and suggestions for future work are presented in Section 4.

II. SYSTEM MODEL AND PERFORMANCE ANALYSIS

The system under investigation is represented by a Wi-Fi Access Point (AP) operating in infrastructure mode. Medium access control is done by the CSMA/CA protocol at the data link layer. AP operation can be in one out of three possible states (Fig. 1). During active state (denoted as *Active*) the AP transfers data with MSs after establishing packet data sessions. The total number of active sessions of all the MSs is a random variable. The AP is intended to serve various kinds of data traffic. Thus, sessions arrivals and completions are modeled as an M/G/1 queuing model. After completing the last active session, the AP turns to *Idle ON* state where an inactivity timer is launched. If a new session arrives before timer expiration, the process turns back to *Active* state for session service and inactivity timer is reset to its initial value. In order to avoid energy waste a power saving mechanism is incorporated. If there is not either MSs within the AP coverage or data activity for a predefined amount of time (inactivity timer expires), the AP switches to low power state (denoted as *Idle OFF*), and turning off the beacons. The interface will be able to serve a user request for a new connection (session) by employing an additional wake-up mechanism based on reverse paging and low-power radio module [8]. Such module uses different technology and operates on separate channel to avoid interference. In such case the AP directly moves to the *Active* state for session service.
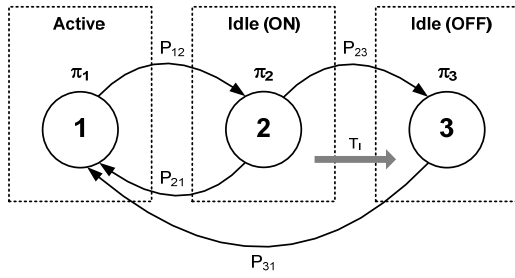
Fig. 1. State transition diagram of the AP operation

The performance analysis is based on the following assumptions with respect to traffic flow characteristics. Offered traffic A is generated by a Poisson arrival process. The active MSs within the AP coverage generate sessions with mean arrival rate λ_s. Time duration of the AP in *Active* state follows the busy period of M/G/1 queuing system. The holding time follows an arbitrary time distribution. During the busy period at least one session is served. The value of inactivity timer is denoted by T_I and is assumed to be constant (deterministic time intervals). Since not all the states of the stochastic process follow the exponentially distributed residence times, the AP behavior can be analyzed as a semi-Markov process. The focus is on steady-states probabilities calculation and proportion of time the process is in a given state.

Suppose that the process can be in any one of three possible states. Each time it enters the state i it remains there for a random amount of time. If an event occurs the process makes transition into state j with probability P_{ij}. Based on the state-transition diagram the transition probability matrix $P = [P_{ij}]$ of the semi-Markov chain is given as

$$P = \begin{bmatrix} 0 & P_{12} & 0 \\ P_{21} & 0 & P_{23} \\ P_{31} & 0 & 0 \end{bmatrix}. \tag{1}$$

Let π_i $(i \in \{1,2,3\})$ denotes the stationary (limiting) probability that the semi-Markov process is in state S_i $(i \in \{1,2,3\})$. In general, π_i is derived by using the balance equation [9]

$$\pi_i = \sum_{j=1}^{3} \pi_j \cdot P_{ji}, \quad i = 1, 2, 3 \tag{2}$$

$$\sum_{i=1}^{3} \pi_i = 1$$

By solving balance equations and using normalization constant, the stationary probabilities can be obtained as

$$\pi_1 = \frac{1}{P_{12} + P_{12} \cdot P_{23} + 1}$$

$$\pi_2 = P_{12} \cdot \pi_1 = \frac{P_{12}}{P_{12} + P_{12} \cdot P_{23} + 1} \quad . \tag{3}$$

$$\pi_3 = P_{12} \cdot P_{23} \cdot \pi_1 = \frac{P_{12} \cdot P_{23}}{P_{12} + P_{12} \cdot P_{23} + 1}$$

Since the process spends an expected time $\overline{t_i}$ in state i, the proportion of time the process is in state i, and hence the AP's steady state probabilities, should be a weighted average of the π_i, i.e.

$$P_i = \frac{\pi_i \cdot \overline{t_i}}{\sum_{j=1}^{3} \pi_j \cdot \overline{t_j}}, \quad i = 1, 2, 3 , \tag{4}$$

where π_i is given as the solution to (3) [9].

The probability that a new session begins before T_I expiration, and hence the state transition probability P_{21}, is

$$P_{21} = 1 - \exp(-\lambda_s \cdot T_I) = \Pr[t < T_I] . \tag{5}$$

Then, the probability of entering the power saving state is

$$P_{23} = 1 - P_{21} = \exp(-\lambda_s \cdot T_I) = \Pr[t \geq T_I] \tag{6}$$

In both *Active* and *Idle OFF* states exit transitions are driven by one event only – session completion and session arrival, respectively. Thus,

$$P_{12} = P_{31} = 1 . \tag{7}$$

During data activity the mean number of active sessions is obtained by the Polaczek-Khintchine formula for arbitrary service time distributions [10], [11]. For M/G/1 model the time duration $\overline{t_1}$ in *Active* state is derived by the Little's theorem

$$\overline{t_1} = \frac{[(c^2)^2 - 1]A^2 + 2A}{2\lambda_s(1 - A)}, \tag{8}$$

where c^2 is the coefficient of variation, which is normalized dimensionless measure for the irregularity of the holding time distribution. It is defined as the ratio between the standard deviation σ and the mean value τ of the session's holding time. The larger c^2, the more irregular (dispersive) is the service time distribution. Typical values of c^2 for M/M/1 and M/D/1 systems are 1 and 0 respectively.

The mean residence time $\overline{t_2}$ in *Idle ON* state is governed by the inactivity timer value T_I. If a new session occurs before T_I expires, the timer is cancelled and *Active* state is entered. Otherwise the AP enters the power saving mode at *Idle OFF* state. Therefore $T_I = \min(T_{21}, T_I)$ is obtained as

$$\overline{t_2} = E[\min(T_{21}, T_{23})] = \int_0^{\infty} \Pr[\min(T_{21}, T_I) > t]dt =$$

$$= \int_0^{T_I} \Pr[T_{21} > t]dt = \tag{9}$$

$$= \int_0^{T_I} \exp(-\lambda_s t)dt = \frac{1 - \exp(-\lambda_s T_I)}{\lambda_s},$$

where $E[X]$ is the expected value operator [12].

The mean residence time in the *Idle OFF* state follows the session arrival rate (Fig. 1)

$$\overline{t_3} = \frac{1}{\lambda_s} \,. \tag{10}$$

Based on the values of steady states probabilities, obtained by (4), the energy consumption of the AP can be calculated. If AP's power consumption in state i is denoted as φ_i, its mean value is derived by

$$\overline{E} = \sum_{i=1}^{3} \varphi_i \cdot P_i \,. \tag{11}$$

The energy consumption is generally expressed in Joules (J), by having knowledge on power consumption for a unit of time.

As stated above, in order to decrease energy consumption AP turns off its radio interface during power saving mode. When a new session arrives, AP must be woken up and move to active state. Unfortunately, this procedure introduces an additional delay since AP must allocate and assign a new IP address to the MS initiated the data session. Therefore, there is a trade-off between energy saving and the delay to obtain a new IP address by DHCP when AP wakes up. This could be quantitatively expressed by the cumulative average delay $\overline{D_c}$, which is defined as

$$\overline{D_c} = D_{OFF} \cdot \frac{P_3}{\overline{t_3}} \,, \tag{12}$$

where D_{OFF} is the delay to initiate a new session in case AP has been in *Idle OFF* state.

III. NUMERICAL RESULTS

This section presents numerical results of the performance analysis the power saving mechanism applied to a WLAN.

The basic unit of time is one hour. It is assumed the AP serving rate μ_s of 200 sessions per hour ($\mu_s = 200\,h^{-1}$). In general, the service (holding) time distribution is assumed to be an arbitrary. Different systems can be analyzed by setting the coefficient of variation c^2. The power consumption in all the states is as follows: $\varphi_1 = 3,604\,\text{W}$, $\varphi_2 = 3,604\,\text{W}$, and $\varphi_3 = 2,637\,\text{W}$. These are average values obtained by measuring on a prototype with a popular Wi-Fi AP [8]. The complete wake-up time comprises the following processes – an application request; a message sending to the AP, which let the interface to be activated; DHCP process starting and completing after assigning an IP address to the MS. Thus, the delay to initiate a new session D_{OFF} is 10 s., as measured in [8].

It is expected that the energy consumption will increase as the inactivity timer value increase, as well, since the probability of staying at *Idle ON* state increases. For a fixed value of inactivity timer, the energy consumption depends on the offered traffic, in terms of Erlangs (Fig. 2). It increases as the traffic load is rising due to frequent transitions into *Active*

state. On the other hand, the coefficient of variation c^2 may have a clear impact, especially for high traffic loads and small values of inactivity timer (right-hand side of Fig. 2). In order to maintain energy consumption constant an admission control mechanism should be adopted. It could be seen that dispersive service time distributions can lead to throughput degradation.

The design of power saving mechanisms for WLANs requires proper setting of the inactivity timer value, since a number of parameters influence on the system performance. Fig. 3 reveals how energy consumption changes with respect to the inactivity timer value and the dispersive nature of the service time distribution. For a fixed value of offered traffic flow, the more higher value of inactivity timer results in increase of energy consumed by the AP due to less number of transitions into power saving (*Idle OFF*) state. Again, these values are considerable for higher traffic loads, since more transitions into *Active* state occur. For sessions with constant (deterministic) service times ($c^2 = 0$) AP achieves the best energy efficiency performance.
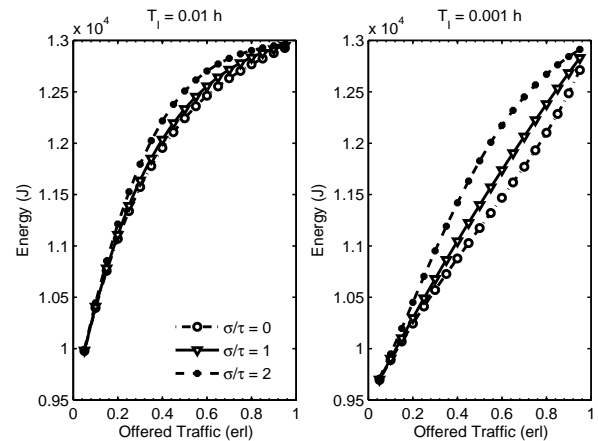

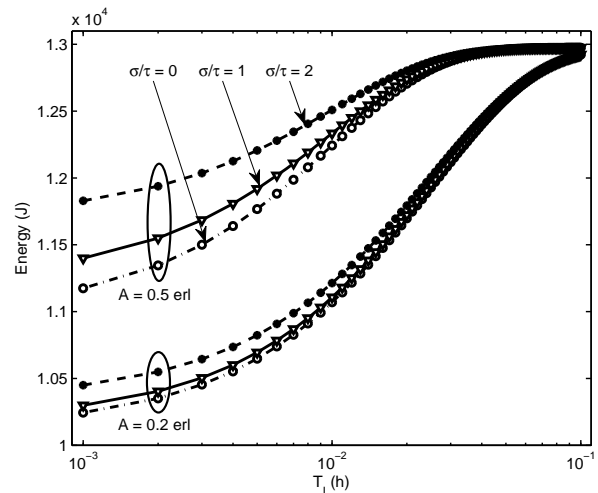Fig. 2. Energy consumption variation as a function of offered traffic and coefficient of variation

Fig. 3. Inactivity timer value influence on the energy consumption
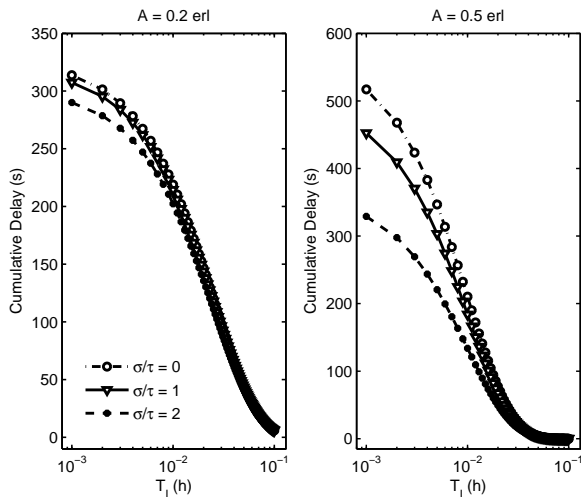
Fig. 4. Inactivity timer value influence on the cumulative delay

As shown, the smaller inactivity timer value, the greater energy efficiency can be achieved. Another interesting aspect of the analysis is the assessment of the average cumulative delay, quantitatively expressed by (12) and depicted on Fig. 4. From the results on both Fig. 3 and Fig. 4 is evident that there is a trade-off between the energy consumption and the cumulative delay. For small values of inactivity timer reasonable energy consumption is achieved (Fig. 3), but frequent transitions into power saving state incur long delays for session reestablishment. Heavy traffic loads and large variations of service time distributions are unfavorable. Thus, an appropriate setting of inactive timer value is necessary in order to balance the above mentioned trade-off.

IV. CONCLUSION

In this paper an analytical evaluation of a power saving mechanism in WLANs has been carried out. The system performance under different service time distributions and inactivity timer values is investigated. The tradeoff between energy consumption and cumulative wakeup delay has been demonstrated, which is important for optimal setting of the inactivity timer value under specific traffic load and service time distributions.

REFERENCES

[1] G. Y. Li et al., "Energy-Efficient Wireless Communications: Tutorial, Survey, and Open Issues", IEEE Wireless Communications, vol. 18, no. 6, pp. 28-35, 2011.

[2] G. Han et al., "Green Radio: Radio Techniques to Enable Energy-Efficient Wireless Networks", IEEE Communications Magazine, vol. 49, no. 6, pp. 46-54, 2011.

[3] Y. Cui et al., "A Survey of Energy Efficient Wireless Transmission and Modeling in Mobile Cloud Computing", Journal on Mobile Networks and Applications (Springer), vol. 18, no. 1, pp. 148-155, 2013.

[4] M. A. Hoque, M. Siekkinen, J. K. Nurminen, "Energy Efficient Multimedia Streaming to Mobile Devices — A Survey", IEEE Communications Surveys & Tutorials, vol. 16, no. 1, pp. 579-597, 2012.

[5] D. Fen et al., "A Survey of Energy-efficient Wireless Communications", IEEE Communications Surveys & Tutorials, vol. 15, no. 1, pp. 167-178, 2013.

[6] S. Dimatteo et al., "Cellular Traffic Offloading through WiFi Networks", IEEE 8th International Conference on Mobile Adhoc and Sensor Systems (MASS), Conference Proceedings, pp. 192-201, Valencia, Spain, 2011.

[7] A. S. Tanenbaum, D. J. Wetherall, *Computer Networks – Fifth Edition*, Pearson Education Limited, 2013.

[8] I. Haratcherev, C. Balageas, M. Fiorito, "Low Consumption Home Femto Base Stations", IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications (IEEE PIRC), Conference Proceedings, pp. 1-5, Tokyo, Japan, 2009.

[9] S. M. Ross, *Introduction to Probability Models – Ninth Edition*, Elsevier Inc., 2007.

[10] S. Mirtchev, *Teletraffic Engineering*, Technical University of Sofia, 2002. (in Bulgarian)

[11] V. B. Iversen, *Teletraffic Engineering and Network Planning*, Technical University of Denmar, 2010.

[12] S. Fowler, R. S. Bhamber, A. Mellouk, "Analysis of Adjustable and Fixed DRX Mechanism for Power Saving in LTE/LTE-Advanced", IEEE International Conference on Communications (ICC), Conference Proceedings, pp. 1964-1969, Ottawa, Canada, 2012.

POSTER SESSION
COMPUTER SYSTEMS AND INTERNET TECHNOLOGIES

Analysis of the Factors which Influence on QoS in LTE Networks

Veneta Aleksieva [1]

Abstract – **LTE is widely used 4G technology, which provides simultaneously voice, data and video (with different priority) on networks. The major concept of the LTE network uses the QoS bearers technique that provides high performance in packet delivery based on prioritization of the traffic. In this paper is proposed multifactor analysis of the applicability to prioritize users in order to fit the bandwidth requirement, while satisfying the application needs.**

Keywords –**LTE, QoS, optimization of QoS in LTE.**

I. INTRODUCTION

According to published in February 2015 prognosis of Cisco System for the period 2014-2019 [7], the Global mobile data traffic will grow three times faster than global fixed IP traffic from 2014 to 2019. The Mobile data traffic will grow 10-fold from 2014 to 2019, a compound annual growth rate of 57%, and it will reach an annual run rate of 291.8 Exabytes by 2019, up from 30.3 Exabytes in 2014. The prognosis of Ericsson [12], in turn, suggests that by 2018 it is expected a growth of smart phones up to 4.5 milliard and growth of mobile video traffic by 60% annually. These estimates require the search of optimal 4G solutions in terms of QoS offered by telecom providers.

Requirements of IMT-Advanced [6] for 1 Gbit/s speeds for fixed and 100 Mbit/s for mobile users present two challenges for providers of wireless services:

1. Optimization of the dynamic selection of the best interfaces of multi-interface devices according to user requirements and limitations in the models of devices such as power consumption, user fees, and application specific requirements for QoS (delay, latency and throughput);

2. Scalability of the work of billions of devices on the wireless network.

Under these requirements, providers of 4G services choose between two advanced wireless technologies - LTE [9] or WiMAX [10], as the success of one or the other technology is determined by a number of factors such as:

- the effectiveness of the business model that the telecommunications companies follows,
- the economic indicators related to the implementation,
- provided by the technology options,
- user expectations for QoS of the offered services,
- the support of the QoS by the service providers.

[1]Veneta Aleksieva is with the Faculty of Computer Science at Technical University of Varna, 1 Studentska Str, Varna 9000, Bulgaria, E-mail: valeksieva@tu-varna.bg.

Equipment manufacturers have an impact too. The fact is that the manufacturers switch to hybrid networks WiMAX / LTE to provide several 4G technologies with the prospect of merging them in future LTE-Advanced. Initially the telecom operators deploy WiMAX as an extension of the already offered DSL-services to provide customers with greater coverage with high speeds. LTE is viewed as an option for higher bandwidth than the existing networks. Companies such as Vodafone, Verizon, China Mobile, AT & T, Nokia and Ericsson stake in their equipment of LTE, while Huawei rely on both the two technologies, predicting that in the countries of Africa and Latin America will long dominate WiMAX technology and equipment, maintenance of WiMAX / LTE will dominate in those markets. The beginning of hybrid schemes supporting WiMAX / LTE is placed in 2010 from companies such as Huawei, Vodafone, KDDI, UQ Communication and others.

Solutions are needed to improve QoS in terms of delays in larger loads and any packet loss. For communications to be successful, it is also essential to focus on network traffic prioritizes for different types of communication streams.

II. ESSENCE OF MANAGEMENT OF TRAFFIC BY PRIORITY IN LTE NETWORKS

Even with the developing of LTE technology the QoS for uplink is discussed by many authors [3,4,5,8,11]. After the first implementations of LTE the focus in the allocation of resources is shifting towards to the profit maximization and user satisfaction [12].

In 3GPP, the QoS Class Indicator (QCI) consist of basic classes, which are defined as 'default', 'expedited forwarding', and 'assured forwarding'. It means: expedited forwarding is used for 'strict' priority (video and voice), and 'assured forwarding' is used for business differentiation (e.g., weighted-fair priority).

In LTE network traffic management achieving QoS is considerably more complicated than in WiMAX. LTE is for users who are willing to pay a higher price, but get more bandwidth for communications or advantage in providing access to the network at peak load hours if introduced the prioritization. But the prioritization applies not only for the consumers but also for the services themselves.

In LTE network QoS is between end-user devices and PDN Gateway applying the 'bearers'. 'Bearers' is a set of network configurations to provide a special handling of traffic to its set prioritization. Their hierarchy is presented in Table I. Default bearer is established when the user equipment (UE) connects to the LTE network, while Dedicated bearer is established whenever must be set QoS for a specific traffic type (service) as VoIP, video and etc.

GBR (Guaranteed Bit Rate) provides guaranteed bandwidth and monitors two parameters in directions uplink and downlink:

- GBR- minimum GBR for EPS bearer,
- MBR- maximum GBR for EPS bearer,
- Non-GBR bearer does not provide guaranteed bandwidth and also monitors two parameters in directions uplink and downlink: A-AMBR-general maximum speed permitted for the entire non-GBR throughput for specific APN (Access Point Name) and UE -AMBR- overall maximum speed permitted for the entire non-GBR throughput for all of APN particularly UE.

TABLE I
HYERARHICAL OF LTE QoS

LTE QoS		
Dedicated Bearer		Default Bearer
Non-GBR	GBR	Non GBR
QCI 5-9	QCI 1-4	QCI 5-9
APN-AMBR	GBR	APN-AMBR
UE-AMBR	MBR	UE-AMBR
TFT	TFT	APN
ARP	ARP	IP Address
L-EBI	L-EBI	ARP

ARP (Allocation and Retention Priority) is used to decide whether the distribution of resources is to be modified according to the new bearer or to maintain the current distribution of the resource.

TFT (Traffic Flow Template) is associated with Dedicated bearer, while Default bearer may or may not have TFT. TFT defines rules based on the source or destination address or protocol used, so that the UE and the network know which IP packets to send over the individual Dedicated bearer.

L-EBI (Linked EPS bearer ID). Each dedicated bearer is always connected to one of the default bearers and L-EBI notifies the Dedicated bearer to which default bearer is connected.

In LTE networks for differentiation of QoS same as in WiMAX are applicable classes which here are called QoS Class of Identifier (QCI). They define the basic characteristics of the IP packet level, as presented in Table II.

TABLE II
QCI CLASSES IN LTE

QCI	Bearer Type	Priority	Delay of the Packet	Packets Loss	Example of Traffic Type
1	GBR	2	100ms	10^{-2}	VoIP
2		4	150ms	10^{-3}	Video call
3		3	50ms	10^{-3}	Real time games
4		5	300ms	10^{-6}	Video stream
5	Non-GBR	1	100ms	10^{-6}	IMS Signaling
6, 8, 9		6, 8, 9	300ms	10^{-6}	TCP based services – chat, ftp...
7		7	100ms	10^{-3}	Voice, video, interactive games

Then in the cell is applied a preemption algorithm, which allows high priority requesting bearers to displace low priority connected bearers in order to reduce the cell load. This algorithm coupled with a priority-based admission control can achieve low dropping and blocking probabilities.

III. EXPERIMENTS

The Functions for management of QoS in access networks are responsible for the efficient allocation of resources in a wireless interface. They are generally defined as the control algorithms of radio resources (Radio Resource Management) and incorporate power management (Power Control), control of the transfer connection (Handover Control), access control (Admission Control), managing load (Load Control) and the management packet (PS), but directly related to QoS level cell are the last three. They are used to ensure a maximum throughput for individual services. The aim is to achieve keeping the network throughput as high as possible at a small price of only a bit more handovers.

LTE uses multiple access technology (OFDMA) and the total bandwidth is divided into Resource Blocks (RBs) in the frequency domain. The Data is transmitted in the Transport Blocks (TB) in one transmission time interval (TTI) for 1ms. Each RB consists from 12 subcarriers (each of them is 15kHz). The frame is 10ms and divides into ten equal subframes. Each subframe contains 2 slots*0.5ms. Each RB is related to one slot in time. One TB is related to 1 subframe and it is the minimum unit to schedule. The serve rule is to find first space that can fit the TB. If there are not enough RBs in the current TTI, the scheduler tries to find resources in the next TTI. This strategy minimizes the response latency, which is the best practice for delay sensitive traffic.

In wireless radio networks, the base station should allow access of as many users as possible to increase revenue. On the other hand, the quality of service should be guaranteed in order to provide satisfactory service. The maximum number of users a base station can support is bound by the system bandwidth. Under the restriction of QoS, if the maximum bandwidth is achieved, new connection requests should be rejected.

Let capacity of cell is C. Then the load L of cell at time slot t is

$$L = \sum_{i=1}^{n} L_i(t) \quad (1)$$

where $L_i(t) = \frac{b_i^v}{b_i(c)}$ (2)

and $L \leq C$ (3)

If bearer j want to use the same cell in the same time slot t, this will be possible if

$$L + L_j(t) \leq \quad (4)$$

If this condition is not executed, this means that the resources in the current TTI are not enough and scheduler must reorder resources in the whole TTI window and must allocate resources in reserved bandwidth. In this case a task is to find spaces big enough for the new request. If there are not enough resources in the current TTI, the scheduler tries to find resources in the next TTI. This strategy minimizes the response latency, and thus is useful for delay sensitive traffic.

But this procedure is not applicable for beacon transmissions (it is sent among devices each 100ms), because of emergency information it conveys, therefore the reserved resource blocks exist to accommodate the temporary overload.

This means, that important factors for QoS are: the Modulation and Coding Scheme (MCS) which be used, the MAC Transport Block (TB) size, the allocation bitmap which identifies which RBs will contain the data transmitted by the eNB to each user, number of users and prioritization of users.

IV. RESULTS

In this approach it is used the Rapid Miner 6.0 to create a model of influence of multi-factors on QoS parameters in LTE network such as real throughput and drop ratio. The drop ratio is defined as the number of the rejected beacons to the number of the accepted beacons. For non-prioritization scheme, the beacons are rejected due to the cell overload. New arrival beacons can only be accepted after some users move out the service region, resulting in load reduction. For prioritization scheme, the rejected beacons are the ones that are removed by the congestion control algorithm.

The test data is obtained according to the values of TB size reported in [1], considering an equal distribution of the physical resource blocked among the users using Resource Allocation Type 0 as defined in Section 7.1.6.1 of [1].

Let TTI duration is 1ms. To calculate the throughput allocated to each user formulae (2) is used. The acceptable relative tolerance (standard deviation) is

$$\sigma = 0. \tag{5}$$

This tolerance is needed to take an account for the transient behavior at the beginning of the simulation.

The main part of DataSet is presented in the Table III.

TABLE III
CALCULATED DATA FOR MODEL OF QoS IN LTE NETWORK

Number of users	Throughput [Mbps] with MCS=22	Throughput [Mbps] with MCS=16	Throughput [Mbps] with MCS=12	Drop ratio with prioritization by distance	Drop ratio without prioritization
1	10.000	9.400	9.350	0.000	0.000
2	8.978	8.378	8.328	0.020	0.100
3	7.956	7.356	7.306	0.020	0.100
4	6.933	6.333	6.283	0.013	0.100
5	5.911	5.311	5.261	0.013	0.040
6	4.889	4.289	4.239	0.010	0.040
7	3.867	3.267	3.217	0.010	0.040
8	2.844	2.244	2.194	0.004	0.024
9	1.822	1.222	1.172	0.004	0.024
10	0.800	0.200	0.150	0.003	0.023

The input DataSet is partitioned into 10 subsets of equal size. Of the 10 subsets, a single subset is retained as the testing DataSet, and the remaining 9 subsets are used as training data set. The cross-validation process is then repeated 10 times, with each of the 10 subsets used exactly once as the testing data. Then results can be averaged to produce a single estimation. The learning processes usually optimize the model to make it fit the training data as well as possible.

The Cross-Validation operator predicts the fit of a model to a testing data. This can be especially useful when separate testing data is not present.

The model is presented on the Fig.1 and Fig.2. The Meta Data is presented on the Fig.3.
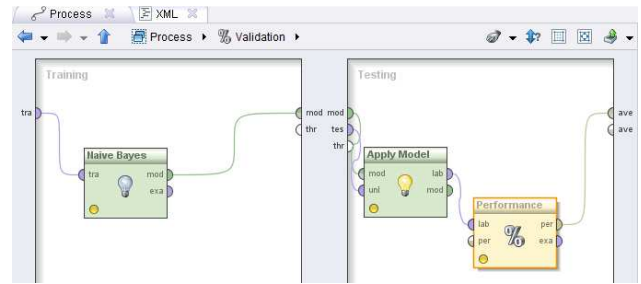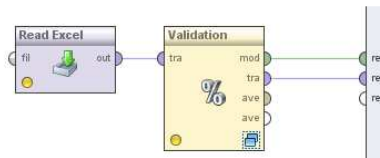

Fig. 1. X-Validation

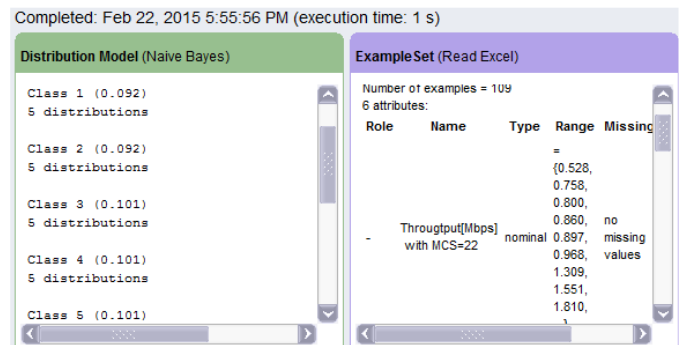Fig. 2. Validation of the Model of QoS in the LTE Network

Fig. 3. Meta Data of Model

The Fig.4, Fig. 5 and Fig.6 presented throughput in Mbps with different MCS – 22, 16 and 12 respectively. In each case when number of users grows up, the throughput decreases. When comparing values among these figures, it is possible to see that for the equal number of users when MCS is more, throughput is more too. When the MCS decreases, the scatter of measured data for throughput increases.
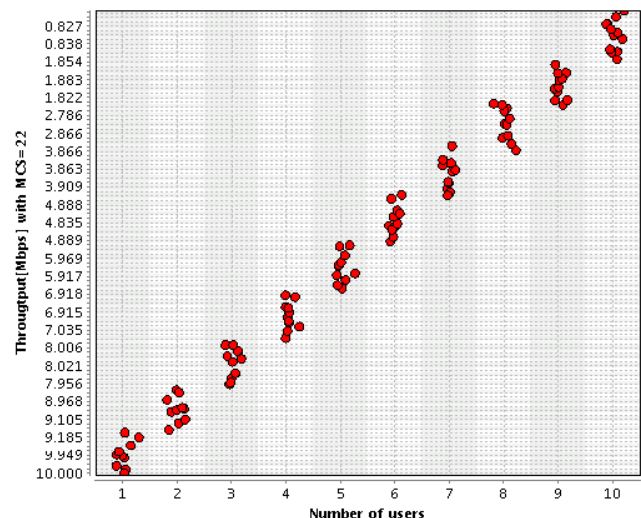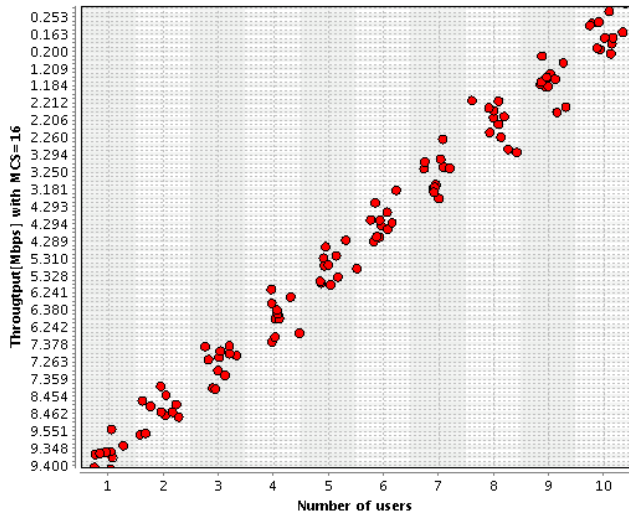

Fig. 4. Throughput with MCS=22
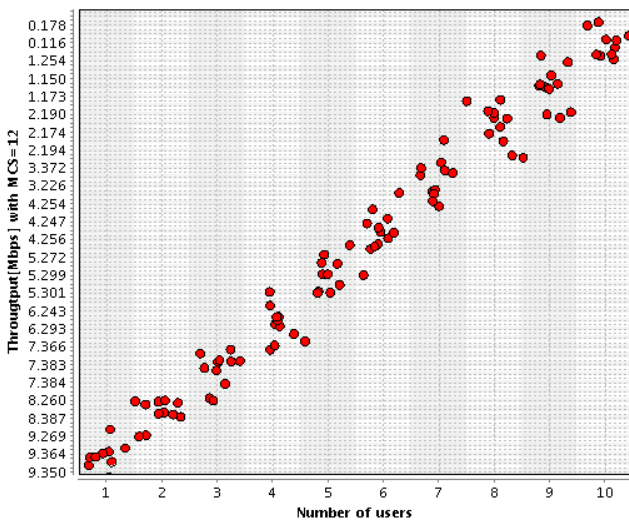
Fig. 5. Throughput with MCS=16

Fig. 6. Throughput with MCS=12

On Fig. 7 is presented the Drop ratio. It is possible to compare a/the number of drops when it is applied a/the prioritization, based on distance between the user and Base station and drops without prioritization.
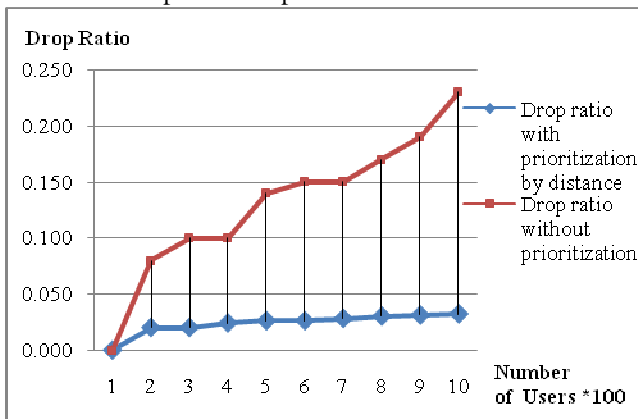

Fig. 7. Drop Ratio

When the number of users grows up, the numbers of rejected beacons grow up too. This is the reason to increase the drop ratio when the number of the users is increased.

The observed parameters degradation when connecting more users is related to the priority implemented scheduler, in which case the less priority queues may not be served in the case of network overload or congestion.

V. CONCLUSION

The main goal of researchers is creating a smart network which is flexible, robust and cost effective. Due to this reason QoS stays in focus in each network- wired, wireless or hybrid. In this paper is proposed an analysis of the applicability to prioritize users in order to fit the bandwidth requirement, while satisfying the application needs, based on analytical data for many factors as MCS which be used, the TB size, the number of users and the prioritization of users. There are presented two QoS parameters – throughput and drop ratio with/without prioritization. It was always assured a minimum transmission for all the service classes, although with different performances due to prioritization.

REFERENCES

[1] 3GPP, Tech. Specif. Group Radio Access Network; Physical layer procedures, 3GPP TS36.213, ver.12.4, 2015.

[2] Abu-Ali N., Abd-Elhamid M. Taha, Mohamed Salah, and Hossam Hassanein, " Uplink Scheduling in LTE and LTE-Advanced: Tutorial, Survey and Evaluation Framework," IEEE Communications Surveys & Tutorials, vol. 16, Issue:3, pp. 1239-1265, Aug. 2014.

[3] Al-Rawi M., R. Jantti, J. Torsner, and M. Sagfors, "On the Performance of Heuristic Opportunistic Scheduling in the Uplink of 3G LTE Net- works," in Personal, Indoor and Mobile Radio Communications, 2008. PIMRC 2008. IEEE 19th International Symposium, pp. 1 –6, 15-18 2008.

[4] Calabrese F., C. Rosa, M. Anas, P. Michaelsen, K. Pedersen, and P. Mogensen, "Adaptive Transmission Bandwidth Based Packet Scheduling for LTE Uplink," in Vehicular Technology Conference, 2008. VTC 2008- Fall. IEEE 68th, pp. 1 –5, 21-24 2008.

[5] Calabrese F., C. Rosa, K. Pedersen, and P. Mogensen, "Performance of Proportional Fair Frequency and Time Domain Scheduling in LTE Uplink," pp. 271 –275, May. 2009.

[6] Document IMT-ADV/1-E, Background on IMT-Advanced, ITU Radiocommunication Study Groups, ps. 3, 7 March 2008.

[7] Global - 2019 Forecast Highlights, 2019 Mobile Data Traffic, http://www.cisco.com/c/dam/assets/sol/sp/vni/forecast_highlights_mobile/index.html#~Region, February 2015.

[8] Liu F., X. She, L. Chen, and H. Otsuka, "Improved Recursive Maximum Expansion Scheduling Algorithms for Uplink Single Carrier FDMA System," pp. 1 –5, may. 2010.

[9] LTE: capacity and cell-edge performance improvements, February 2014, http://www.3gpp.org/.

[10] Ramadas K., R. Jain, WiMAX System Evaluation Methodology v2.1, WiMAX Forum, July 2008.

[11] Ruiz de Temino L., G. Berardinelli, S. Frattasi, and P. Mogensen, "Channel-Aware Scheduling Algorithms for SC-FDMA in LTE Uplink," in Personal, Indoor and Mobile Radio Communications, 2008. PIMRC 2008. IEEE 19th International Symposium, pp. 1 –6, 15-18 2008.

[12] Wi-Fi in heterogeneous networks, http://www.ericsson.com/res/docs/whitepapers/wp-wi-fi-in-heterogeneous-networks.pdf, June 2013.

Multi-threaded user and kernel-space library

Hristo Valchanov[1] and Simeon Andreev[2]

Abstract – **Developing of technology and the large range of possibilities offered by modern hardware allows the use of specialized high-performance approaches for the implementation of various software systems and algorithms. One of the most used and effective approach is the creation of multi-threaded software running in parallel on multiple processors. This paper presents the features of the implementation of multi-threaded library under Linux, which allows running both in user and kernel-space mode.**

Keywords – **Multi-threaded library, Threads, User-space, Kernel-space.**

I. INTRODUCTION

Development of technology and the large range of possibilities offered by modern hardware components allow the use of specialized high-performance approaches in the implementation of various software systems and algorithms. One of the most used and effective such approach is the creation of multi-threaded software running in parallel on multiple processors [7]. The actuality of this type of tasks and the need for new and improved modern applications argues mainly with the introduction of application oriented processors containing multiple independent cores (currently 6-8), which can operate independently of each other [2].

Every modern operating system supports multiprocessor operations. Important requirement for modern operating systems is the effective use of available hardware resources and provision them in the most appropriate way to user programs [1].

A current trend is using of multi-threaded libraries as an essential part of the most used and effective programming languages such as C ++, Java and others. The existence of different specific architectures makes this task difficult and sometimes even impossible for effective implementation. So there are many different standard libraries optimized for a particular architecture and operating system [5].

The paper presents the specifics of the implementation of multi-threaded library under Linux, which allows for operation both in user (user-space) and system (kernel-space) mode.

II. MOTIVATION

There are currently multiple implementations of multi-threaded libraries for different purposes. The GNU Portable Threads (Pth) [3] is a library created with the idea of a portable interface to a wide range of UNIX systems. The

[1]Hristo Valchanov is with the Department of Computer Science at Technical University of Varna, 1 Studentska Str., Varna 9010, Bulgaria, E-mail: hristo@tu-varna.bg.

[2]Simeon Andreev is with the Department of Computer Science at Technical University of Varna, 1 Studentska Str., Varna 9010, Bulgaria, E-mail: simeon.andreev90@gmail.com.

library Next Generation POSIX Threading (NGPT) [4] is developed by IBM for compatibility with POSIX standard for Symmetric Multi-Processing (SMP) machines. Linux Threads [8] is an implementation of the POSIX IEEE 1003.1c standard for Linux platforms. The library is based on the model one-to-one and operates in user mode. The new implementation of threads in Linux - Native Posix Thread Library (NPTL) [9] is also compatible with the POSIX standard, but it is based on a kernel-space model.

The implementations are optimized for machines with different architectures - uniprocessor, multiprocessor with shared memory, multiprocessor with distributed memory [6]. The modern processors are multi-core, which results in increased productivity. However, one should also take into account the fact that not every class of algorithms is subject to parallelism. There are algorithms for which has not yet found an effective multi-threaded implementation, or even inability to establish such one. On the other hand, there are many tasks for which the improvement in the speed of execution is great and justifies the additional difficulties that occur in multiprocessor operation. The software developer should have a choice of different functionality of multi-threaded libraries depending on the specifics of the various tasks to solve.

The presented library provides such flexibility. For this purpose a maximum identical user interface in both modes has been developed.

III. IMPLEMENTATION IN USER-SPACE MODE

A. Threads switching

The way of organizing threads switching by the dispatcher is essential for the effectiveness of multi-threaded library.

A possible approach for taking the processor from a user thread and providing it to the dispatcher is using an interrupt. In Linux the signals are convenient system for the realization of such user software interruptions. An interruption can be taken, for example, by a timer set to a specified interval. There are several problems making such a decision not very desirable. First, a very big advantage of the design of the entirely library in user-space, is the ability to fast switching between threads. The continuous use of signals and timers, however, requires switching to the kernel, which slows down repeatedly this usually fast operation. Second, storing of the user context that was interrupted by the timer is a difficult task. This is because when an interrupt function is executed, for some library functions it is not guaranteed to function properly.

Another possible approach is the dispatcher to work in a separate thread visible to the operating system, but this leads to other difficulties in implementation without providing the necessary effectiveness.

These are the main reasons why is using a non-preemptive dispatching in the proposed implementation of the user-space

library. This solution also provides an opportunity to clearly show the main advantages and disadvantages of threads creating entirely in user-space mode.

B. Threads states in user-space mode

Each thread is represented by a special structure (Thread Control Block - TCB), containing the necessary information for its management. This information is used by the dispatcher for scheduling and context switching. The dispatcher takes care of the management of the events, which is very important for proper operation of the library. The dispatcher is called whenever a thread gives the CPU to another thread or when it locks automatically to an event. Figure 1 shows the states that a thread can take during its existence.
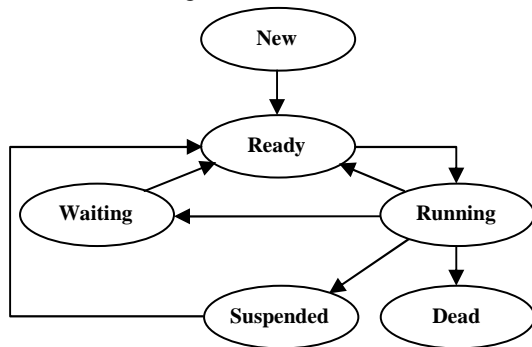
Fig.1. Threads states in user-space mode

Each of the states except *Running*, is represented with a priority queue of TCB of the threads which are in this state. A new created thread first is in the *New* state. From there, at the earliest opportunity the dispatcher can move it into the *Ready* state. On every dispatching, from the *Ready* queue is selected the most priority task which becomes *Running*. From the *Running* state a thread can go into any of the following states: *Dead* - if it finishes its execution, *Waiting* - if it is blocked to an event or into the *Ready* state otherwise. On every dispatching the threads in the *Waiting* queue are checked and moved to the *Ready* state if the appropriated events occur. In the *Suspended* state each thread can go if it is need not to be executed for a given period of time.

The non-blocking read/write operations, the thread sleeping for a specified time and the operations with synchronization primitives are entirely based on the event system.

IV. IMPLEMENTATION IN KERNEL-SPACE MODE

A. Threads states in kernel-space mode

The basic idea of this approach is to use the resources available to the operating system level for tasks dispatching. This allows for much better planning and using of resources of the entire system, but the disadvantage is that there is a slow switching between threads because of the need for system calls. The implementation of the threads in this mode is based on system call *clone()*. For the Linux kernel a thread

is stored in the same structure, which is used for a separate process. Although that the basic information is stored in the structures of the operating system, it is still necessary to maintain data for the thread in the user context too. Such information, for example, is for the starting function and its argument. Maintaining this duplicated information allows a simpler implementation of some library functions.

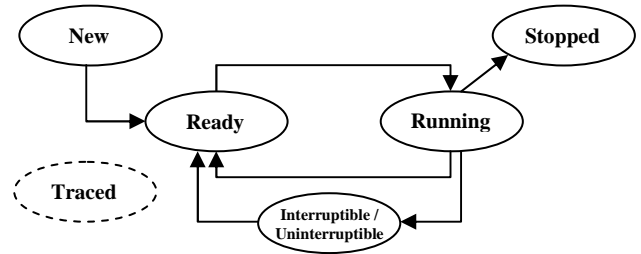Figure 2 shows a graph of the threads states in this mode.

Fig.2. Threads states in kernel-space mode

The thread states are following:
- *Running* - The thread is either in the queue for starting or its execution is in progress.
- *Interruptible* - The thread is temporarily suspended while is awaiting the fulfillment of some condition. It wakes up on the event occurrence or on receipt of a signal.
- *Uninterruptible* – The state is identical with the preceding, the only difference is that the signals have no effect on the thread.
- *Traced* - When a thread is traced in the system (debugging mode).
- *Stopped* - When the execution of the thread is terminated. This state is achieved upon receipt of certain signals.

B. Threads synchronization in kernel-space

In the current library implementation are developed two types of synchronization primitives - *spinlocks* and *mutexes*. The spinlocks provide synchronization based on active waiting. Although this approach is not particularly effective, implementation is extremely simplified.

The second type primitives are based on waiting to release the processor. They are more complicated to implement because they require the use of system calls to the kernel. This is necessary for the dispatcher to suspend the execution of the current thread and to choose another thread to continue its work. In the proposed library the mutexes are implemented in the most effective way through a mixed approach. Initially, an attempt to short active waiting is made. If after that the mutex is still busy, the kernel blocks the thread execution. For this purpose is used a special tool for basic access locking - Fast User-Space Mutex (*futex*) [10].

In general, for the proper operation of a futex it is need to allocate a system semaphore. In the current implementation are used special (private) version of futex, which are local to the process, thus saves checking the semaphore internally in the kernel. This implies that the implementation is faster than that provided by the NPTL library where mutexes have a much more complex structure, slowing their use.

V. APPLICATION PROGRAMMING INTERFACE

The library provides an application programming interface that is maximum identical in both modes. The interface includes several groups of data types and library functions with different purpose such as:

- Types of data to work with system structures;
- Functions for initialization and completion of the work with the library;
- Functions for creating and destroying threads;
- Synchronization functions;
- Function for blocking and waiting;
- Functions for access to information for a separate thread.

VI. EXPERIMENTAL STUDY AND RESULTS

The testing of multi-threaded library was made with regard to the two basic types of system load:

- CPU bounded processes. These are processes that perform many and long time calculations.
- I/O bounded processes. These are processes that perform intensive system calls to the operating system. For example, saving in the files, waiting for an input from the user, using network communications.

The fastest execution for CPU bounded processes is expected to be achieved when the dispatcher uses a maximum possible period between switching of two threads. This allows saving of information and maximum use of processor caches. These processes almost always work without interruption, using all their time determined by the dispatcher. An optimal policy for them is to be allowed to work a long time, whereas it is not necessary their frequent starting.

On the other hand, the I/O bounded processes do not require long periods to use the CPU. This is because they rarely utilize it fully and often block alone. An optimal policy for them is the dispatcher to run them as much as possible more often for work.

Comparisons are made with the most common multithreading library *pthreads* and in particular with its most recent implementation in Linux - NPTL.

The experimental platform is based on Intel i7 2600K, running at 4.0GHz clock speed. The processor uses "HT" technology and can execute 8 threads simultaneously. Each core has its own cache memory: L1 - 64KB, L2 - 256KB and shared L3 cache size 8MB. The memory is DDR3 SDRAM - 8GB, 1600MHz, dual channel mode. The operating system is 64-bit Arch Linux distribution with kernel version 3.8 and the compiler - gcc 4.8.1 and glibc 2.17.

The first test involves matrix multiplication, as an example of a task comprising multiple calculations. There are created 5 separate threads, each of them multiplies 20 times a square matrices of size 100x100 elements. Figure 3 shows the obtained results.

As we expect, the *pthreads* and kernel-space implementations cope best with the test. The results are close, but the advantage is for *pthreads* library. From this test is shown the great disadvantage of user-space libraries - they can

work on only one processor. The remaining two libraries take advantage of the available 8 cores. The test excludes blocking of threads.
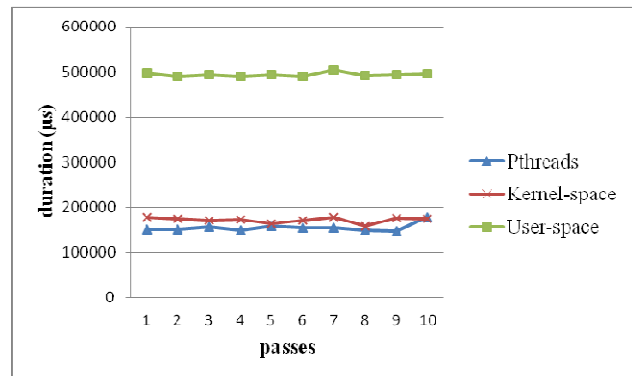

Fig.3. Results for matrix multiplication

The second test uses handling of input/output operations. There are created 2 threads, one reads from a file, and the other saves data in it. The pairs read/write operations are performed 1000000 times. It is expected that during the execution of some of the operations blocks could occur. Figure 4 shows the obtained results.
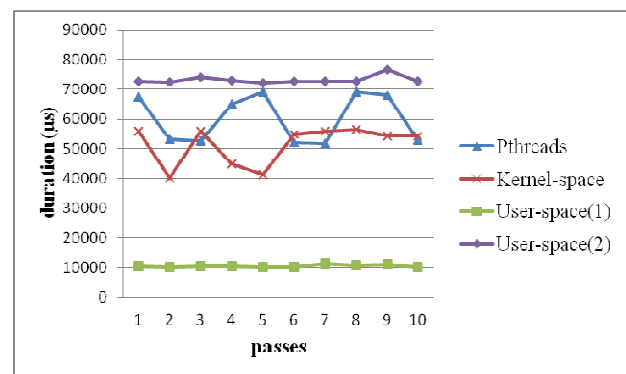

Fig.4. Results for I/O operations

The user-space library is used in two ways. With "1" is noted the use of non-blocking primitives provided by the current implementation, and with "2" - when they are replaced with standard *write()* and *read()*. As can be seen, the developed tools have better efficiency in comparison with the standard ones in Linux.

The *pthreads* and kernel-space libraries show the same performance with a slight preponderance of the kernel-space implementation.

The test results show also a greater fluctuation between different starts because it is more difficult to predict variants of execution.

The third test aims to assess the effectiveness of synchronization primitives. In the cycle of 1000000 iterations a mutex is locked and unlocked. There are created 5 separate threads, each of which performs the described action. Figure 5 shows the obtained results.

It is evident that the implementation of mutex in *pthreads* library is slower and requires an average 615154μs.

The version in the kernel-space library is about 1.5 times faster, thanks to a simple and effective structure of organization of the threads. Also the advantage is due to the fact that an active waiting is used instead immediately suspending of the execution of the thread.
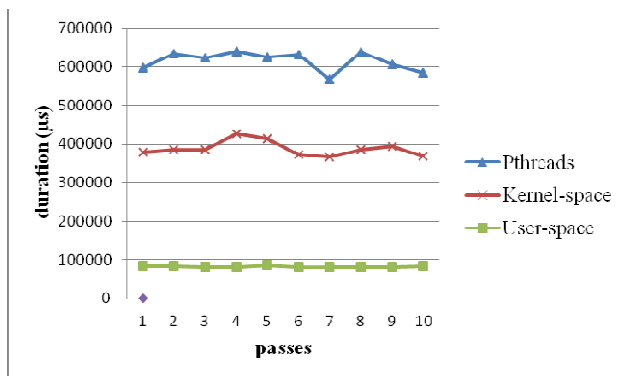

Fig.5. Results for synchronizing primitives

Quite expectedly the user-space implementation is proved fastest with an average of 82875μs. This is due to the fact that slow system calls are not used.

VII. CONCLUSIONS

This paper presents an implementation of a multithreaded library under Linux, functioning both in user-space and kernel-space mode. Some features of implementation are given. Experimental comparisons and evaluations with the famous library *pthreads* have been made.

The results show that for presented library it is achieved identically and in specific cases high performance. The developed library is relatively small, which results in faster compilation. The code is written in a simple way, which allows the use of the library for learning multithreaded programming.

Goal of future work is to improve the signal processing in user-space mode, as well as developing additional synchronization primitives, such as conditional variables and monitors.

REFERENCES

[1] Haldar S., A. Aravind. Operating Systems. Dorling Kindersley Ltd. 2010.
[2] Herlihy M, N. Shavit. The Art of Multiprocessor Programming. Morgan Kaufmann, 2008.
[3] Engelschall R. Portable Multithreading. The signal stack trick for user-space thread creation. In Proceedings of 2000 USENIX Annual Technical Conference, 2000, pp.18-23.
[4] IBM Corporation. Next Generation POSIX Threading, November 2002.
[5] Ljumovic M. C++ Multithreading Cookbook. Packt Publishing, 2014.
[6] Love R. Linux System Programming. O'Reilly. 2013.
[7] Sandem B. Design of multithreaded software. Wiley Inc. 2010.
[8] Leroy X. The Linux Threads Library. http://pauillac.inria.fr/~xleroy/.
[9] Native Posix Thread Library. http://people.redhat.com/drepper/nptl-design.pdf.
[10] Futex Semantics and Usage. http://man7.org/linux/man-pages/man7/ futex.7.html.

Analysis and Simulation of a Virtual Environment for the Organization of Competitions in Programming with Local and Remote Clients

Deyan Atanasov [1], Trifon Ruskov [2] and Vladimir Nikolov [3]

Abstract - **The aim of this paper is to analyze the services provided in a distributed system by establishing a virtual environment, which simulates the system usage, load and performance while conducting a programming contest (National Collegiate Programming Contest) where the teams are local or remote clients.**

Keywords - **Cloud computing, data center, simulation, resource management.**

I. INTRODUCTION

For the past few years there is a significant development in the distributed processing of information, based on the existence of a wide variety of computing environments and platforms. This rapid development provided the foundation for the development of new more flexible architectural models and concepts such as Grid computing and Cloud computing.

Cloud computing is based on virtualization technology and provides its customers with on-demand, reliable, secure and easy access to resources and services. Cloud computing is a dynamic model in which Cloud service providers could sell or resell services to their customers. The dynamics of the model is determined by the fact that the customer uses the service only when needed, and pays only for the period of its use.

Cloud providers offer their services according to several fundamental delivery models [1]:
- Software as a Service (SaaS);
- Platform as a Service (PaaS);
- Infrastructure as a Service (IaaS).

IaaS is the delivery model that we are simulating and describing in this paper. It is a self-service model for accessing, monitoring, and managing remote datacenter infrastructures, such as computer hardware (servers, networking technology, storage, and data center space) as a service. IaaS may also include the delivery of operating systems and virtualization technology to manage the resources.

[1]Deyan Atanasov is PhD student at the Computer Science and Engineering Department at Technical University of Varna, 1Studentska Str., Varna 9010, Bulgaria, E-mail: dido.paraskevov@tu-varna.bg.
[2]Trifon Ruskov is with the Computer Science and Engineering Department at Technical University of Varna,1Studentska Str., Varna 9010, Bulgaria, E-mail: ruskov@tu-varna.bg.
[3]Vladimir Nikolov is with the Computer Science and Engineering Department at Technical University of Varna,1Studentska Str., Varna 9010, Bulgaria, E-mail: nikolov_vn@tu-varna.bg.

This paper is organized as follows. Related problem and used instruments are described in Section II. Section III presents the experiment scenario. The results from the experiments are given in Section IV. The conclusion and future work are specified in Section V.

II. PROBLEM BACKGROUND AND USED TOOLS

Our simulation experiment aims to analyze the services provided in a distributed system by establishing a virtual environment. This virtual environment simulates the system usage, workload and performance while conducting a programming contest (National Collegiate Programming Contest), where the teams are local or remote clients. The problem arose some time ago and during the last few editions of this competition the organizers experienced some technical issues concerning load balance and system performance. Our department hosted the last contest, so we had the opportunity to examine and analyze the parameters which affected the performance. By this analysis we have obtained the values of some important parameters concerning system and network performance. Parameters such as size of tasks, amount of data to transfer, network bandwidth, etc. were taken into account and used in establishing the simulating environment.

To build such a virtual environment we used the open source simulation instrument CloudSim [2, 3]. CloudSim is an extensible simulation toolkit that enables modeling and simulation of Cloud computing systems and environments. We decided to use CloudSim for our simulation model because of its time efficiency, flexibility and applicability to different Cloud concepts. Some specific to Cloud environments functionalities that CloudSim can provide are simulating of large Cloud environments (data centers) on a single physical compute node and simulating a federation of clouds (multiple inter-networked clouds) [4].

Our simulation of distributed computing environment is also based on the use of NetBeans IDE [5] for the development and deployment of our scenario. We used a web-based instrument called CloudReports to show the results from the experiment. This is an open source tool with a graphical interface using the CloudSim kernel.

III. SIMULATION SCENARIO

A distributed system is used to support the learning process at the Computer Sciences and Engineering Department at the Technical University of Varna. This system also aims to provide various other online educational services. It consists

of three blade servers and each of them has the following parameters:

- 2 x CPU Intel XEON E5-2600v2 6 core 2 GHz;
- 32GB RAM;
- 2 x HDD 146GB 15K rpm;
- 10GbE network adapter.

These characteristics of the physical distributed infrastructure are used as parameters in our simulation environment. Thus we obtained the opportunity to analyze and compare the simulation results with the real distributed system. The simulation environment receives and processes requests and tasks that match with the tasks sent by the participating teams in the real programming contest.

For this purpose, the simulation model is based on the configuration of a datacenter, containing three servers (hosts) with the parameters of the actual blade servers. The datacenter is used by both local and remote clients. These clients are sending tasks (cloudlets) to the virtual machines running on the datacenter. The cloudlet is a class that models the Cloud-based application services (such as content delivery, social networking, and business workflow) and represent the usage Cloud-based applications and services [2, 3]. Cloudlets have predetermined parameters such as size, amount of data to transfer and length of instructions measured in millions of instructions per second (MIPS). After analyzing the results from the actual contest, these parameters were set with values that correspond to the actual tasks from the programming competition.

Cloudlets use the computing resources of the deployed virtual machines via one of the following scheduling policies. They are space-shared, time-shared and dynamic workload:

1. Space-shared scheduling policy – by implementing this policy, the virtual machines executes the tasks one by one. When a cloudlet is being executed, all the other cloudlets are waiting in standby mode;

2. Time-shared scheduling policy – by implementing this policy, each virtual machine executes several tasks simultaneously. In time-shared policy, each task is using the virtual machine for a certain period of time after which access to VM resources is given to the next task;

3. Dynamic workload policy – this policy allows dynamic resource load generation, which means dynamic usage of the virtual machines resources. Under this policy, each cloudlet is given a VM with available resources.

Other important parameters that we take into account for the purpose of the experiment are:

- Network bandwidth;

- Packets transfer delay – parameter which represents the network latency that a message can experience. We use this parameter for data sent from remote clients. In this model, there are no actual entities available for simulating network equipment, such as routers or switches. Instead, network latency that a message can experience on its path from one CloudSim entity to another is simulated;

- Resource allocation method - resource allocation (VM allocation) is the process of creating virtual machine instances on hosts that meet the characteristics and configuration requirements of the Cloud service providers in order to achieve certain levels of quality of service (QoS). The default allocation method used is First-Come, First-Served (FCFS) [6]. This is a well known allocation policy in which items are processed in order of arrival.

We deployed a total of 30 virtual machines, configured with single-core processors, performance 1000 MIPS and 1GB of RAM in the datacenter. The VMs are 30, because there were 30 teams in the actual programming contest (every competitor used 1 VM). Remote clients (competitors) used 20 of them and the other 10 were used by the local clients (competitors). The simulation test was 30 min long. The analysis of the actual competition showed that the distribution of executed tasks represents a normal (Gaussian) distribution with displacement of the maximum to the last 30 minutes of the race. Therefore, our simulation time interval is chosen to match the final 30 minutes of the contest, when the workload and the number of client requests are significant.

IV. RESULTS

The results of the simulation experiment are shown in Fig. 1, Fig. 2 and Fig. 3. They show the workload and resource utilization of the simulated infrastructure during the simulation. For the purposes of the simulation we used all three scheduling policies. The best results were obtained by using the Dynamic workload scheduling policy, therefore Fig.1-3 are representing the usage of this policy.
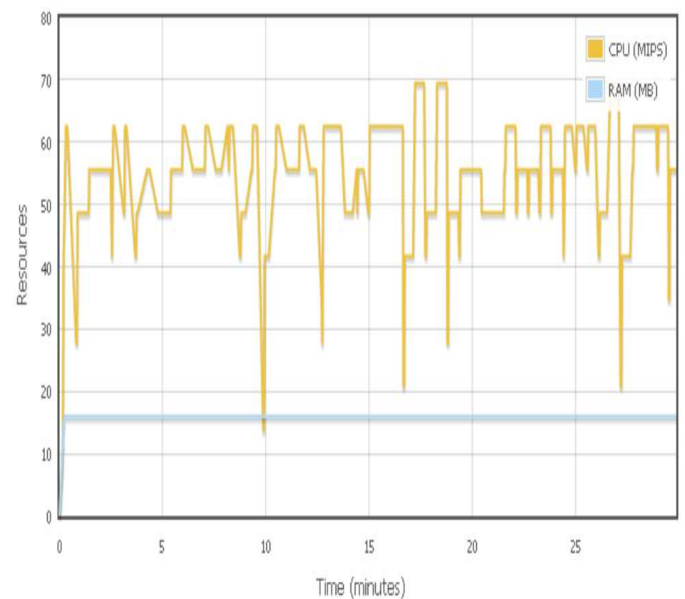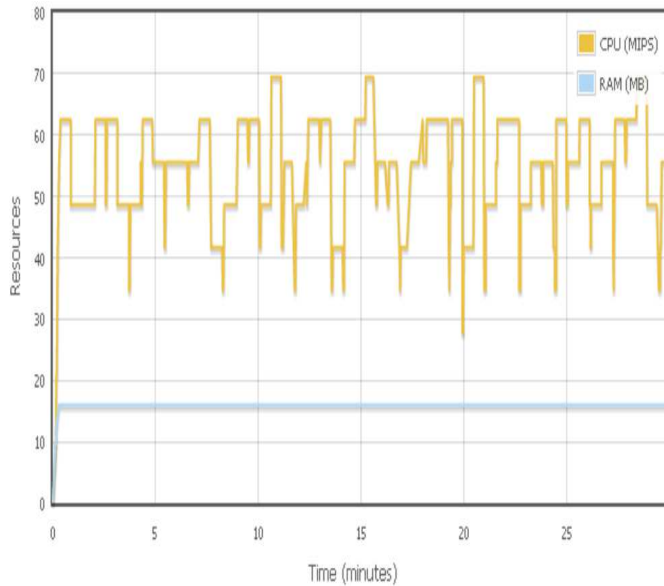

Fig.1. Used resources in Host 1.
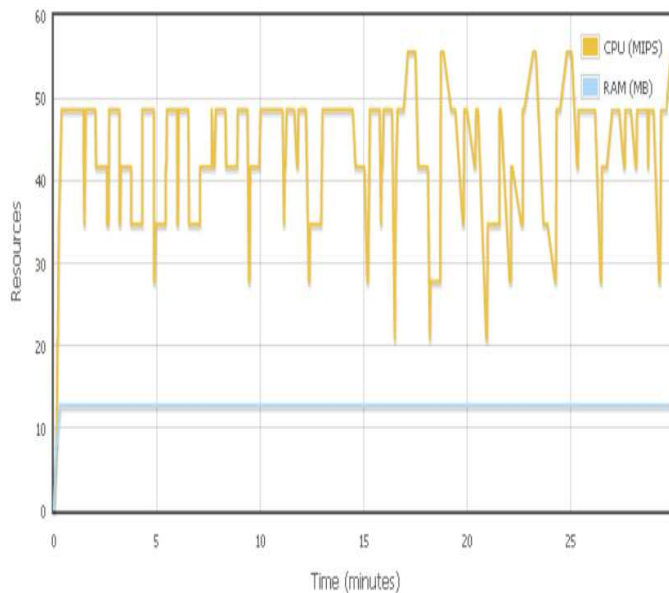
Fig.2. Used resources in Host 2.

Fig.3. Used resources in Host 3.

The obtained results indicate the following:

1. The Dynamic workload scheduling policy gives the best results in terms of uniform (balanced) load of servers and VMs;

2. Used resources in the datacenter, are distributed evenly among the servers (hosts);

3. The simulated infrastructure has enough free resources and can process applications and tasks from a much larger number of customers, both local and remote;

4. During the simulation, the average number of processed tasks by remote clients is smaller than the average number of task processed by the local clients, but the difference is negligible. Remote clients processed 660 cloudlets, which is an average of 33 cloudlets per VM. Local clients executed a total of 378 cloudlets, which is an average of 37,8

cloudlets/VM. The number of tasks processed by each virtual machine is shown in Fig.4.
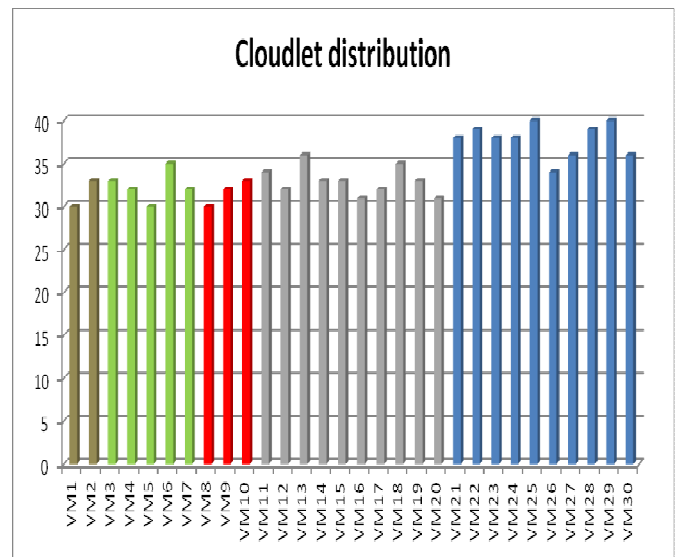
Fig.4. Distribution of executed tasks (cloudlets) by VM

Different colors in the graph are corresponding to different packet transfer delays. This delay is used to simulate different remote clients in the simulation. We have set the virtual machines (VM1-VM20) with the following packet transfer delays:

VM1 and VM2 – 15ms;
VM3 to VM7 – 20ms;
VM8 to VM10 – 18ms;
VM11 to VM20 – 10ms.

VMs with numbers 21-30 represent the local clients in the simulation. They all have been set with packet delay of 1ms.

5. The number of cloudlets executed by the VMs is much greater than the number of tasks from the real contest, which indicates that the infrastructure is able to cope with the workload and the large number of requests during the competition.

V. CONCLUSION AND FUTURE WORK

The conducted in this work experiments show the applicability of CloudSim for simulation related to the efficiency of processing applications and tasks from local and remote clients. Such a configuration of the system would handle the workload in this type of competitions.

As a guideline for future work on the simulator may be development and implementation of different allocation algorithms and scheduling policies. Such developments would give us the opportunity to compare and analyze different methods and Cloud scenarios using wide variety of policies and algorithms in order to achieve high level of QoS and performance in distributed systems.

REFERENCES

[1] P. Mell, T. Grance, "The NIST Definition of Cloud Computing", NIST Special Publication 800-145, 2011.

[2] R. Calheiros, R. Ranjan, C. A. F. DeRose, R. Buyya, "Cloudsim: A Novel Framework for Modeling and Simulation of Cloud Computing Infrastructures and Services", 2009.

[3] R. Calheiros, R. Ranjan, A. Beloglazov, C. A. F. DeRose, R. Buyya, "CloudSim: A Toolkit for Modeling and Simulation of Cloud Computing Environments and Evaluation of Resource Provisioning Algorithms", Software: Practice and Experience (SPE), Volume 41, Number 1, pp: 23-50, ISSN: 0038-0644, Wiley Press, New York, USA, 2011.

[4] R. Buyya, R. Ranjan, "InterCloud: Utility-oriented Federation of Cloud Computing Environments for Scaling of Application Services", Proceedings of the 10th International Conference on Algorithms and Architectures for Parallel Processing, pp:328-336, Busan, South Korea, 2010.

[5] NetBeans IDE, Available at: https://netbeans.org/.

[6] Dictionary of Algorithms and Data Structures, Available at: http://www.nist.gov/dads/HTML/firstcome.html.

Programming approaches for implementing web servers for static content

Hristo Nenov[1] and Sevdalin Todorov[2]

Abstract – **The paper describes the role of the web servers for static content, and what part of the overall concept of the internet they occupy. Different models and algorithms for implementation of servers are described, and their strengths and weaknesses are pointed.**

Keywords – **Servers for Static Content, Approach, Model, Thread, Process, I/O Operation.**

I. INTRODUCTION

Popularity of Internet and widespread usage of Web requires timely improvement of the technologies and software used for Internet communications and servers. The Increasing number of users of WWW, as well as its penetration in more areas makes web servers critical component in this area. The enormous number of users whose requests must be processed increases the requirements for hardware and communication resources, which in turn increases operational costs. With the dynamic development of the network consumers are becoming more demanding of the server response time when processing their requests. Most of the resources that web servers serve are static – images, CSS, client-side JavaScript, audio, video etc.. There are few popular web servers with general purpose as well as many not so common. The main goals of most of them are flexibility, lots of functionality and cross-platform, thus the performance and system resources usage remain in the background.

II. WEB SERVER IMPLEMENTATION APPROACHES

A. WEB PAGE CONTENT

The information in the World Wide Web is contained in hypertext documents or information resources, known as web pages. Usually the web page format is HTML or XHTML and allows navigating to other web pages using hyperlinks. Web pages may include other resources like multimedia, CSS and JavaScript. Content that does need extra processing prior serving it to the user is referred to as static content, if the content needs processing it is referred to as dynamic content. Web servers use HTTP protocol to

transmit the content. Often the Web server is used to provide information in other formats - most often these are images in the formats PNG, JPEG and GIF; XML documents, CSS and JavaScript files, etc.. so-called static content.
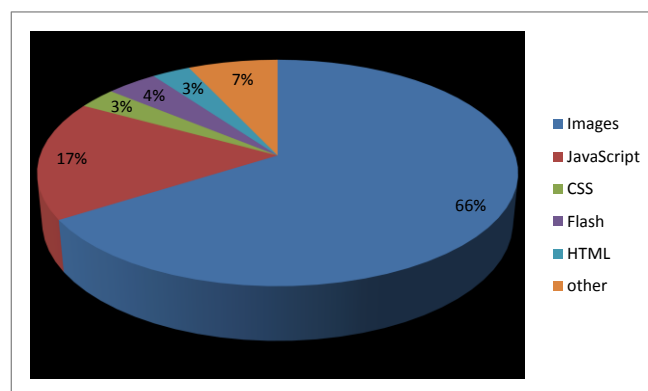

Fig.1 Average bytes per page for different type of context

Usually the Web server uses external programs to process the information before it is sent to the user using CGI scripts or application servers, however the processing module can be built-in into the server code. These scripts can be written in one of many programming languages, but the most commonly used languages are Java, PHP, Python, Ruby and Perl.
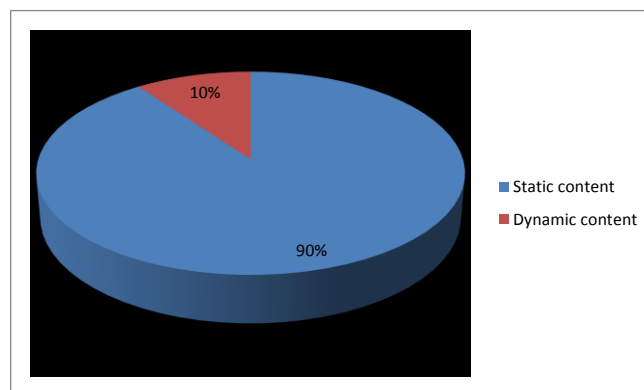

Fig.2 Average bytes per page for different type of content

B. LOADING SPEED

In September 2009 - Akamai Technologies, Inc. published a key study conducted by Forrester Consulting, which examines eCommerce web site performance and its correlation with an online shopper's behavior. The most compelling results reveal that two seconds is the new threshold in terms of an average online shopper's expectation for a web page to load and 40 percent of shoppers will wait no more than three seconds before abandoning a retail or travel

[1]Hristo Nenov assist. prof. at Faculty of Automation and Computing at Technical University Varna, 1 Studentska str.Varna 9000, Bulgaria, E-mail: h.nenov@tu-varna.bg
[2]Sevdalin Todorov graduated bachelor engineer at Faculty of Automation and Computing at Technical University Varna, 1 Studentska str.Varna 9000, Bulgaria, E-mail: dincho.todorov@gmail.com

site. Additional findings indicate that quick page loading is a key factor in a consumer's loyalty to an eCommerce site, especially for high spenders. 79 percent of online shoppers who experience a dissatisfying visit are less likely to buy from the same site again while 27 percent are less likely to buy from the same site's physical store, suggesting that the impact of a bad online experience will reach beyond the web and can result in lost store sales. In a similar study conducted in 2006, Akamai found customer expectations at four seconds or less. Considering both studies and the period of three years between them, it may be concluded that user expectations of web pages loading time were doubled. Given that, and the rapidly increasing mobile users and their significant requirements, because of the constraints of their environment (end of 2014) it can be concluded that expected loading time of web pages is less than one second. After internal research conducted in 2009 by Google, they also found that the loading times of web pages is of great matter, the greater is the delay - the less users use the service. On the one hand, from technical point of view many Internet publications argue that reducing the load time of pages dramatically reduces the load of the service equipment and therefore reduces operating costs by up to 50%. This is expected since theoretically faster request processing leads to faster resources freeing and available for use by other requests and programs, thereby increasing the ratio of efficiency of the equipment.
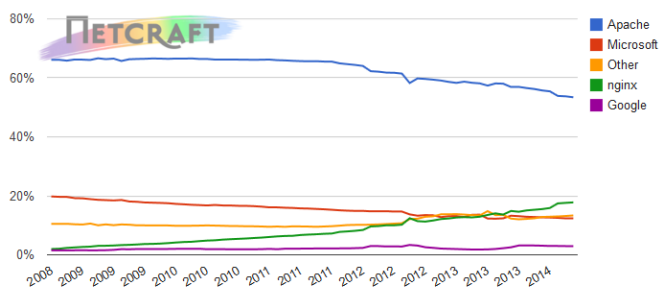
Fig.3 Web Server Survey

C. APACHE SERVER APPROACH

Apache is among the most widely used web servers. According to NetCraft -- which provides Internet research services -- Apache web servers dispatch over 50% of the overall content -- static and dynamic -- on the web.

The Apache software foundation produces two types of web servers. Apache HTTP used for static content that can also be equipped with modules to serve dynamic content (e.g. PHP, Ruby), as well as Apache Tomcat which is a web-container (i.e. application server) used for serving dynamic content written in Java. The modules add additional functionality such as CGI, SSL, virtual hosts and processing of applications written in almost any programming language. Modules can be loaded dynamically without the need for recompilation of the web server.

The main modules through which Apache processes client requests are called modules for multi-processing. These modules are responsible for client requests processing and their distribution to the worker threads. These modules interact directly with the operating system by system calls.

Multiprocess module

This module creates a lot of child processes, but with one thread per each. Therefore, a process can handle only one request at a time. It is obvious that this architecture is not scalable. The module is mainly used for compatibility with older modules that are not thread safety.

Multiprocess-multithreaded module

This module creates a lot of child processes, and many threads for each of them. Therefore, each process can serve multiple requests. This is the best option of both modules because it is scalable, but resource consumption increases dramatically with increasing of concurrent requests.

Advantages and disadvantages

The advantages of Apache web server are its modular architecture and dynamic loading of its modules. Because Apache is developed and used so for many years, there are a lot build modules for it. Another advantage is its cross-platform characteristics.

The outdated and scale-dependent architecture that uses multiprocess-multithreaded processing of requests can be pointed as a major disadvantage. The support of large number of operating systems can also be specified as disadvantage because such a realization needs a lot additional source to achieve platform independence.

D. NGINX SERVER APPROACH

Nginx is a web server, developed by Igor Sasoev. Initially used for sites with high traffic in his home country – Russia, Nginx has developed since then and is now one of the fourth most used web servers in the world according to Netcraft. Besides being able to handle static content, Nginx modules also support FastCGI - allowing processing and dynamically generated content. Nginx functionality is based on modules like Apache, however, unlike the Apache modules which can be built into the web server or loaded dynamically, Nginx modules can be added only at compile time. Nginx is designed with asynchronous event-driven architecture so you can use only one thread to processing many requests in concurrent. At high load, this architecture uses less and predictable amount of RAM- to process each request, compared to Apache model which uses "Multiprocess-multithreaded" oriented approach - depending on the selected module. Therefore Apache creates a new thread to handle new request - which requires additional memory. Nginx can handle new requests using existing threads, because of its architecture. Nginx also is cross-platform and supports many different operating systems.

Advantages and disadvantages

The biggest advantage of this web server is its architecture - asynchronous event-driven. A disadvantage is also the relatively high amount of source code needed to build: the

modular architecture, work with dynamic content and numerous of extra functionalities.

E.UNIX BASED OPERATING SYSTEMS

UNIX based systems are dominant in the sector of web services. According W3Techs their share was 67.5% percent as of May 2014. Normally, the kernel and the application layers of these operating systems are open source, which contributes to their popularity, security and stability. Also, most of these systems are free, which is an essential factor in the choice of operating system, for such applications. UNIX operating systems are proven in the years of operation, superior performance and efficiency in the use of system resources. The most common of these is Linux, which is the reason that was chosen as a platform for research in this paper.

Linux asynchronous input/output (I/O) model

Linux asynchronous I/O is a relatively recent addition to the Linux kernel. It's a standard feature of the 2.6 kernel. The basic idea behind AIO is to allow a process to initiate a number of I/O operations without having to block or wait for any to complete. At some later time, or after being notified of I/O completion, the process can retrieve the results of the I/O.

Fig.4 Simplified matrix of basic Linux I/O models

Synchronous Blocking I/O

One of the most common models is the synchronous blocking I/O model. In this model, the user-space application performs a system call that results in the application blocking. This means that the application blocks until the system call is complete (data transferred or error). The calling application is in a state where it consumes no CPU and simply awaits the response, so it is efficient from a processing perspective.

Figure 5 illustrates the traditional I/O blocking model, which is also the most common model used in applications today. Its behavior are well understood, and its usage is efficient for typical applications. When the read system call is invoked, the application blocks and the context switch to the kernel. The read is then initiated, and when the response returns (from the device from which you're reading), the data is moved to the user-space buffer. Then the application is unblocked (and the read call returns).

From the application's perspective, the read call spans a long duration. But, in fact, the application is actually blocked while the read is multiplexed with other work in the kernel.
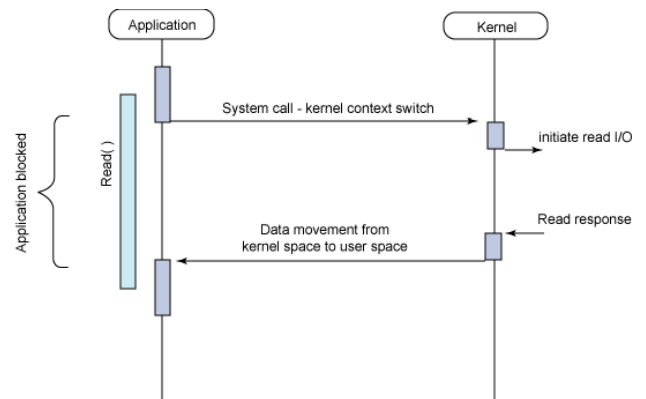

Fig.5 Typical flow of the synchronous blocking I/O model

Synchronous non-blocking I/O

A less efficient variant of synchronous blocking is synchronous non-blocking I/O. In this model, a device is opened as non-blocking. This means that instead of completing an I/O immediately, a read may return an error code indicating that the command could not be immediately satisfied (EAGAIN or EWOULDBLOCK), as shown in Figure 6.
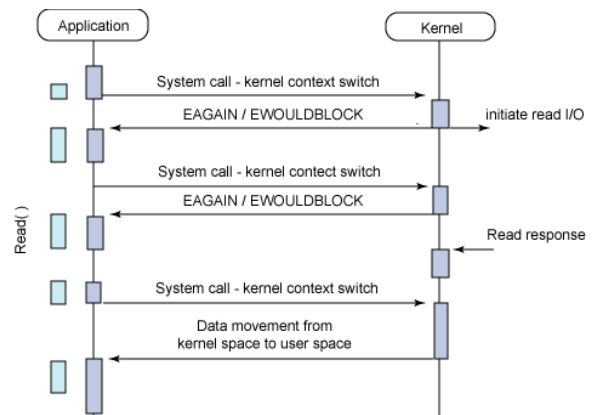

Fig.6 Typical flow of the synchronous non-blocking I/O model

The implication of non-blocking is that an I/O command may not be satisfied immediately, requiring the application to make numerous calls to await completion. This can be extremely inefficient because in many cases the application must busy-wait until the data is available or attempts to do other work while the command is performed in the kernel. As also shown in Figure 6, this method can introduce latency in the I/O because any gap between the data becoming available in the kernel and the user calling read to return it can reduce the overall data throughput.

Asynchronous blocking I/O

Another blocking paradigm is non-blocking I/O with blocking notifications. In this model, non-blocking I/O is configured, and then the blocking select system call is used to

determine when there's any activity for an I/O descriptor. What makes the selected call interesting is that it can be used to provide notification for not just one descriptor, but many. For each descriptor you can request notification of the descriptor's ability to write data, availability of read data, and also whether an error has occurred.
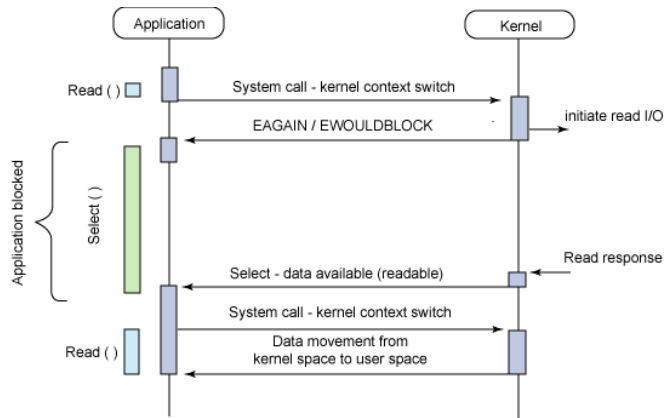

Fig7.Tipical flow of the asynchronous blocking I/O model (select)

The primary issue with the select call is that it's not very efficient. While it is a convenient model for asynchronous notification, its use for high-performance I/O is not advised.

Asynchronous non-blocking I/O (AIO)

Finally, the asynchronous non-blocking I/O model is one of the overlapping processing with I/O. The read request returns immediately, indicating that the read was successfully initiated. The application can then perform other processing while the background read operation completes. When the read response arrives, a signal or a thread-based callback can be generated to complete the I/O transaction.
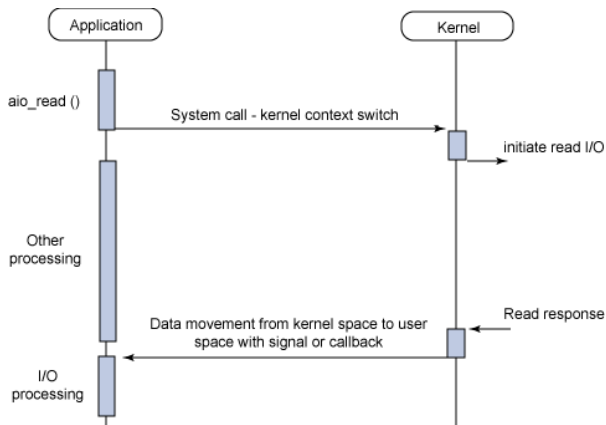
Fig.8 Typical flow of the asynchronous non-blocking I/O model

The ability to overlap computation and I/O processing in a single process for potentially multiple I/O requests exploits the gap between processing speed and I/O speed. While one or more slow I/O requests are pending, the CPU can perform other tasks or, more commonly, operate on already completed I/Os while other I/Os are initiated.

III. CONCLUSION

In this article an overview of different programming approaches and principles for creating web server for static content is shown. From the presented facts can be made the following conclusions:

- the bottle neck in job of the web servers for static content is the performance of input output operations;
- choice of the proper architecture according to the specific role of the server significantly increases productivity and efficiency.

Future point in the study will be a real implementation of a web server by choosing some of the architectures described above.

REFERENCES

[1]. S. Todorov, "Design and Implementation of web server for static content", diploma thesis 2014.
[2]. Л. Николов,. UNIX: Системно програмиране. - София: Сиела, 2009.
[3]. Л. Николов,. Операционни системи. - София: Сиела, 2012.
[4]. П. Наков, П. Добриков, Програмиране =
[5]. ++Алгоритми. - София: TopTeam Co, 2012.
[6]. Asynchronous I/O. Wikimedia, <http://en.wikipedia.org/wiki/Asynchronous_I/O> (15.09.2014)
[7]. eCommerce Web Site Performance Today. Akamai Technologies, Inc., <http://www.akamai.com/html/about/press/releases/2009/press_091409.html> (15.09.2014)
[8]. D. Rubio, "Web application performance and scalability." <http://www.webforefront.com/performance/webservers_statictier.html> (15.09.2014)
[9]. J. Brutlag, "Speed Matters for Google Web Search." Google, Inc., 22.06.2009, <http://googleresearch.blogspot.com/2009/06/speed-matters.html> (15.09.2014)
[10]. "Usage of operating systems for websites. W3Techs" <http://w3techs.com/technologies/overview/operating_system/all> (15.09.2014)

Asynchronous non-blocking IO model approach to avoid the problem C10K in web servers for static content

Hristo Nenov [1] and Sevdalin Todorov[2]

Abstract – **The paper describes one of the main problems in processing of the web servers – C10K. A proposal for implementation of the server for static content is made and test benchmark and comparison results with chosen etalon are shown.**

Keywords – **static content, asynchronous non-blocking IO, event, event handler, pool, queue, thread, C10K.**

I. INTRODUCTION

The primary function of a web server is to store, process and deliver web pages to clients. The communication between client and server takes place using the Hypertext Transfer Protocol (HTTP). Delivered pages are most frequently HTML documents which may include images, style sheets and scripts in addition to text content. A web server has defined load limits, because it can handle only a limited number of concurrent client connections (usually between 2 and 80,000, by default between 500 and 1,000) per IP address (and TCP port) and it can serve only a certain maximum number of requests per second depending on:

- its own settings;
- the HTTP request type;
- whether the content is static or dynamic;
- whether the content is cached;
- hardware and software limitations of the OS of the computer on which the web server runs.

When a web server is near to or over its limit, it becomes unresponsive.

II. *C10K* PROBLEM IN SERVER JOB

The name C10k is a numeronym for concurrently handling ten thousand connections. The C10K Problem refers to the inability of a server to scale beyond 10,000 connections or clients due to resource exhaustion. Servers that employ the thread-per-client model, for example, can be confounded when pooled threads spend too much time waiting for blocking operations-usually I/O. The native thread implementations on most OSes allocate about 1 MB of

[1]Hristo Nenov assist. prof. at Faculty of Automation and Computing at Technical University Varna, 1 Studentska str. Varna 9000, Bulgaria, E-mail: h.nenov@tu-varna.bg
[2]Sevdalin Todorov graduated bachelor engineer at Faculty of Automation and Computing at Technical University Varna, 1 Studentska str.Varna 9000, Bulgaria, E-mail: dincho.todorov@gmail.comBulgaria.

memory per thread for stack. As a result, blocking operations easily frustrate scalability by exhausting the server's memory with excessive allocations and by exhausting the server's CPU with excessive context-switching. For that reason the blocking operations should be avoided at all cost from servers that are likely to be challenged by the large number of customers at peak load.

The C10k problem is the problem of optimizing network sockets to handle a large number of clients at the same time. The problem of socket server optimization has been studied because a number of factors must be considered to allow a web server to support many clients. This can involve a combination of operating system constraints and web server software limitations. According to the scope of services to be made available and the capabilities of the operating system as well as hardware considerations such as multi-processing capabilities, a multi-threading model or a single threading model can be preferred. Concurrently with this aspect, which involves considerations regarding memory management (usually operating system related), strategies implied relate to the very diverse aspects of the I/O management.

III. DESIGN AND IMPLEMENTATION OF WEB SERVER FOR STATIC CONTENT (STATIX V.$0.1.0$)

For the implementation of the experimental server is chosen "Asynchronous non-blocking I/O" model. HTTP server listens for new connections from clients. During the accepting process of a new connection, the server reads the contents of the application submitted on the link and then makes its analysis according to the HTTP protocol. The resulting structure of the application is processed and according to it an appropriate response is generated. The response may be the require content or mistake. Finally the formatted response is sent to the appropriate client link and the connection is closed if it is not checked "keep-alive" options, otherwise the connection does not close, only resets the counters and buffers to it.
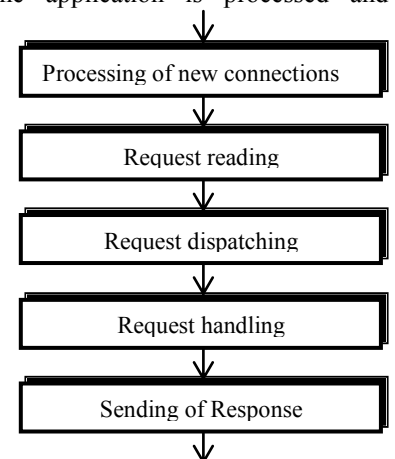

Fig.1 Generalized Work Scheme

Acceptation of new connections is processed only from one main thread. Once the connection is accepted a working

thread is selected and read request in its event queue is registered. As addition, deletion, modification and retrieval of events are done through system calls (according to event-notification system) and those system calls are serialized i.e. they are thread safe and additional synchronization is not necessary. Each working thread has its own event queue.
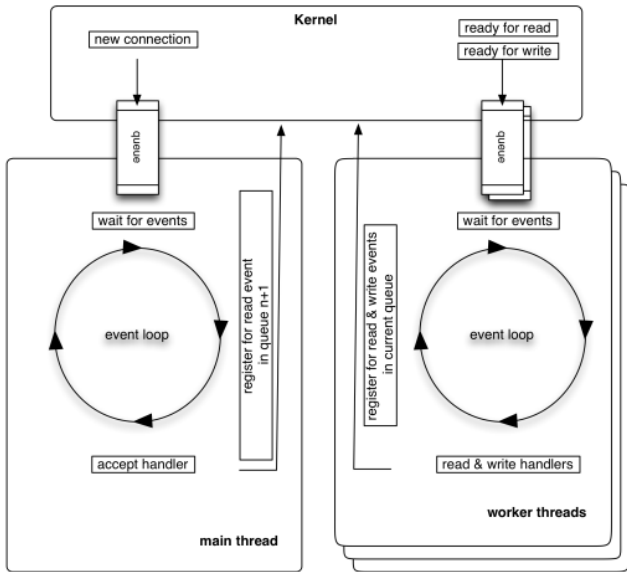

Fig.2 Generalized Work Architecture

A. Handling of new connections
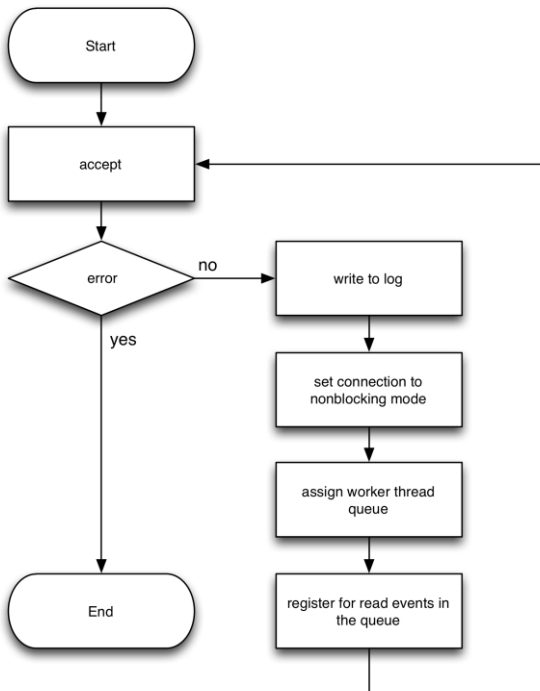
Fig.5 Processing of new connection

Upon the occurrence of new connection event, all pending connections are administering simultaneously as during the time in which the event is generated and the time in which it is processing new connections requests can be received. Once the connection is accepted, it is set

appropriately, the event is recorded in the journal and event request for queue reading is registered in one of the working threads event queue. The thread chooses simple round principle, thus achieving balancing distribution of working load between threads.

B. Reading of request

Upon receipt of the event for reading the first thing is to be determined whether the connection is still active, if it is not then it closes itself. In the first event for reading to a certain connection is verified the capacity of working thread in which the processing is executed, and if it is reached then the connection closes itself, otherwise the application is initialized. The data reading from connection can be executed repeatedly until all data is read or an error occurs. The error may be due to too much data that cannot fit in the buffer request, a transmission error if the client closed the connection or if the result from the operation will be a process of blocking. By error of blocking process, a new request event for reading is registered and the processing ends. If the client closes the connection, the server processing ends. At large volume of data or other error in the response the respective code answer is included. Control is passed to the next stage - parsing and processing application.
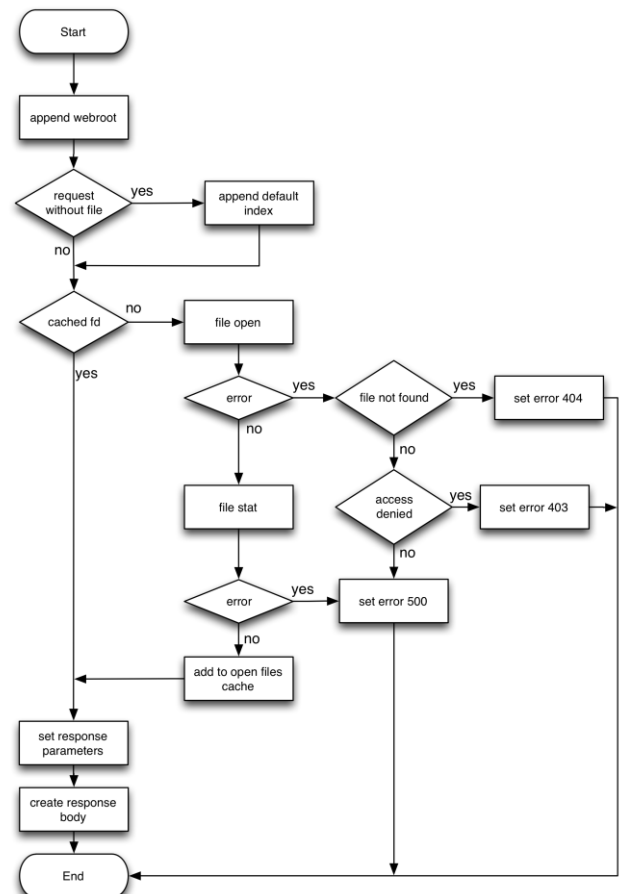

Fig.6 Request processing

C. Request handling

The Request handling consists in few steps:

- processing of the file request;
- determination of the appropriate header for different content type;
- generating of response.

In the time of file processing its absolute path on the file system server is generated. The check is if there is such a file and is it available for reading. For a high level of performance the cache of the open files is used. At successful open of the file, the respective status of the response and its size are set.

The content type of the response is determined by a static table based on the extension of the request. The construction of the response based on the formation of the text, depending on the status, type, size, and the subsequent closing of connection. If there is not an available opened file, the body of the response is formed by static table based on the status of the response. Once the request is processed, the response is send to the client.

IV. RESULTS

Tests were performed on a clean server especially designed for these tests. There are no unnecessary services installed and used otherwise. For comparatively analysis is used web server Nginx. Tool for the benchmark is Weighttp with Weighttp wrapper.

Technical part:

- HP ProLiant DL380p Gen8
- 2 x Intel(R) Xeon(R) CPU E5-2650 @ 2.00GHz
- cores: 16
- threads: 32
- RAM: 64GB DDR3 1600 MHz

Software environment:

- Ubuntu 14.04.1 LTS (64bit)
- Kernel: Linux ubuntu 3.13.0-35-generic
- Statix 0.1.0 (our project)
- Nginx 1.4.6 (etalon for benchmark)
- Weighttp 0.3
- Weighttp wrapper 5.10.7

TABLE I
STATIX V.0.1.0 RESULTS

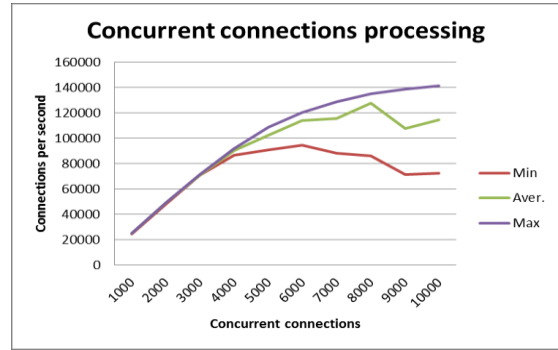| Connections	Min	Average	Max	CPU % (usage)	CPU % (kernel)	Memory (MB)
1000	24464	24880	24967	51	264	40,83
2000	48378	49051	49317	41	258	41,11
3000	70766	70960	71334	61	236	48,45
4000	86460	90592	92163	52	248	53,57
5000	90668	102468	108517	52	245	62,62
6000	94759	114013	120446	47	270	62,78
7000	88030	115415	128845	50	251	68,53
8000	86079	127605	135004	34	277	68,63
9000	71328	107752	138751	57	262	68,73
10000	72418	114732	141538	54	265	68,75

FIG.7 PROCESSING PERFORMANCE OF STATIX V0.1.0

TABLE II
NGINX 1.4.6 RESULTS

| Connections	Min	Average	Max	CPU % (usage)	CPU % (kernel)	Memory (MB)
1000	23642	23786	23932	155	170	47,49
2000	22204	30646	36092	129	146	49,69
3000	28981	38987	66015	134	199	49,91
4000	27260	32057	40562	131	146	50,41
5000	13682	31937	51481	137	164	50,04
6000	14126	35815	65161	71	115	50,28
7000	13990	29367	32551	92	140	50,53
8000	13629	25508	32106	158	185	50,36
9000	13586	23564	31398	173	215	51,79
10000	13507	25480	31678	135	208	51,47
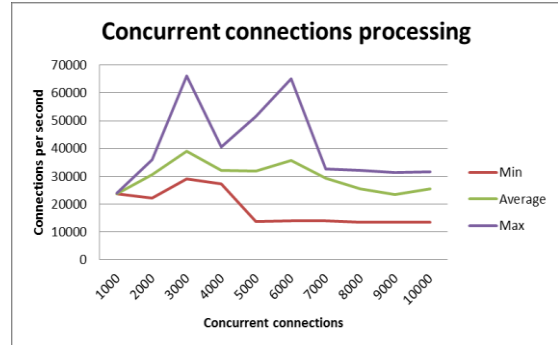
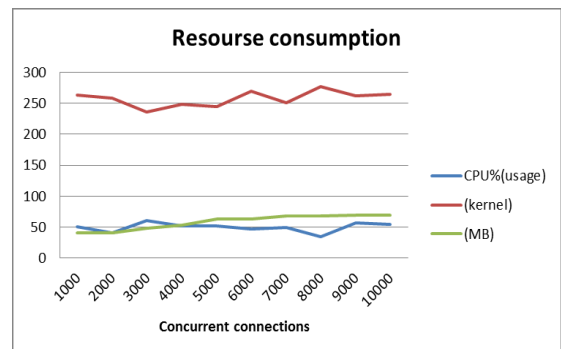Fig.8 processing performance of Nginx 1.4.6

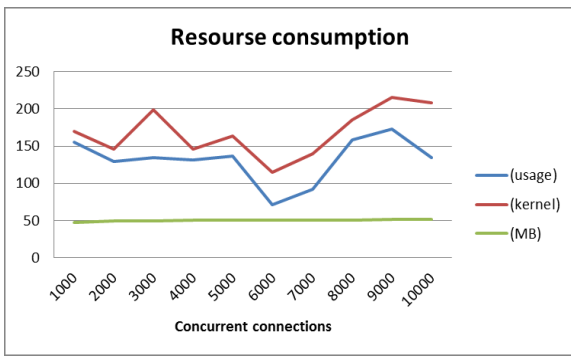Fig.9 Resource consumption of Statix 0.1.0

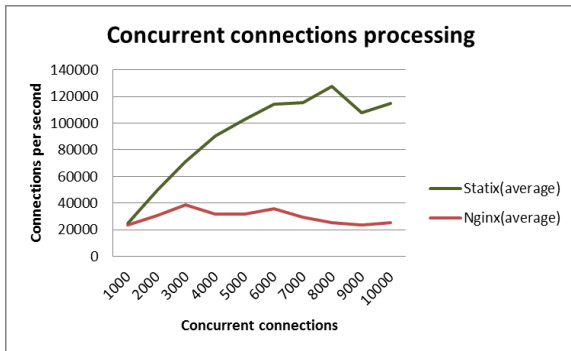Fig.10 Resource consumption of Nginx 1.4.6

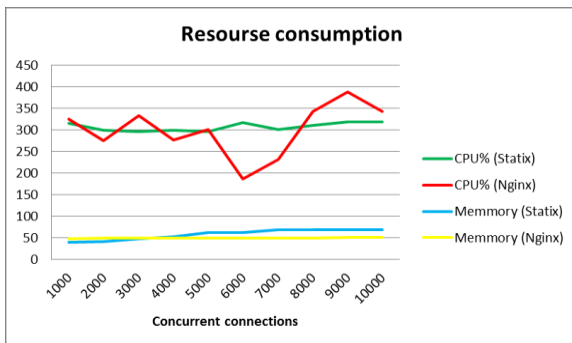Fig.11 Comparison on request processing
Statix 0.1.0 - Nginx 1.4.6

Fig.10 Comparison on resource usage
Statix 0.1.0 - Nginx 1.4.6

V. CONCLUSION

From the experiments made (shown in the preceding tables and graphs) the following can be concluded:

- at increasing of the number of client connections almost linearly increases the number of concurrent connections that are processed. At the target 10k connections, the server is doing extremely well. Moreover it reaches performance nearly five times better than the comparator etalon.
- the results are not at the expense of using a huge amount of resources – CPU, memory and etc.. Consumption of resources is in very good range. A good impression makes their stable behavior.

High productivity, efficient operation and predictability resource consumption are the results of the choice of scalable architecture for building Web server - "Asynchronous non-blocking I/O" model.

Experiments clearly show that this type of architecture is especially suitable for high load web servers.

REFERENCES

[1] S. Todorov,. Design and Implementation of web server for static content, diploma thesis 2014.
[2] D. Rubio,. Web application performance and scalability. <http://www.webforefront.com/performance/webservers_ statictier.html> (15.09.2014)
[3] "Asynchronous I/O." Wikimedia, 11.09.2014, <http://en.wikipedia.org/wiki/Asynchronous_I/O> (15.09.2014)
[4] T. Jones, "Boost application performance using asynchronous I/O. Emulex, 29.08.2006
[5] "Usage of operating systems for websites." W3Techs, 15.09.2014, <http://w3techs.com/technologies/overview/operating_system/ all> (15.09.2014)

CRYPTOGRAPHIC PROTOCOL WITH A PROPOSED CIPHER AND APERIODIC KEY REPLACEMENT

Sivo Daskalov[1] and Milena Karova[2]

Abstract – **The proposed cryptographic protocol implements a cipher with block and key length of 2^n bits. The encryption algorithm consists of several stages each of which swaps or inverts different segments of the plaintext according to the current key. Each key is used a seemingly random number of times between 0 and 15, afterwards the next generated key is encrypted and transmitted using the previous one. The initial key is established using Public key encryption.**

Keywords – **Aperiodic, block cipher, cryptography, segmentation, symmetric key**

I. INTRODUCTION

For a cryptographic protocol to be secure, it is necessary that the cryptographic algorithms used perform well under the conditions of its particular application. Developers of cryptographic algorithms and protocols take measures against an adversary's possible actions and try to ensure that the protocol's goal is achieved with regard to all possible attacks.

Cryptographic protocols are used mainly for: establishment of symmetric session secrets (for one-to-one authentication), signing message digests (for one-to-many, or broadcast authentication), schemes of user or remote workstation authentication and computerized coin tossing [3, 4].

II. CRYPTOGRAPHIC PROTOCOL

A. Description of the block cipher and processing algorithm

The data stream is divided into blocks of length equal to a power of two (2^n). For demonstration of the encryption algorithm and the majority of the experiments a length of 64 bits is chosen. The cipher uses a symmetrical key algorithm with key length equal to the block length. The encryption consists of n ($n=\log_2 blocklength$) major phases and a final inversion phase. In the example shown on Fig. 1 with block length L equal to 64 the phases are 6+1. During the six major phases the ciphertext is divided into segments of length 2^{k-1} where k is the number of the current phase. In each phase the algorithm performs $L/2^k$ operations with neighboring segments. The result of each such operation forms a segment of greater rank, having the combined length of the two original segments and preserving their location in the ciphertext.

[1]Sivo Daskalov is a student in the CST department of the Technical University of Varna, E-mail: sivodaskalov@gmail.com

[2]Milena Karova is an Associate Professor in the Computer Sciences and Technologies department of the Technical University of Varna, 1 Studentska Street, Varna 9000, Bulgaria, E-mail: mkarova@ieee.bg

The processing of each pair of segments is controlled by a predefined bit in the key and there are three possible outcomes of the operation:

- The value of the control bit of the key is 0 – The two segments preserve their location and value.
- The value of the control bit of the key is 1 and the two segments are not identical – In this case the segments preserve their value but swap their location.
- The value of the control bit of the key is 1 and the two segments are identical – In this case each bit of the two segments is inverted.

The last bit of the key is used for an inversion of the cipher text block if the bit's value is 1

TABLE I
BITS OF THE CIPHER KEY USED IN EACH PHASE OF THE ENCRYPTION

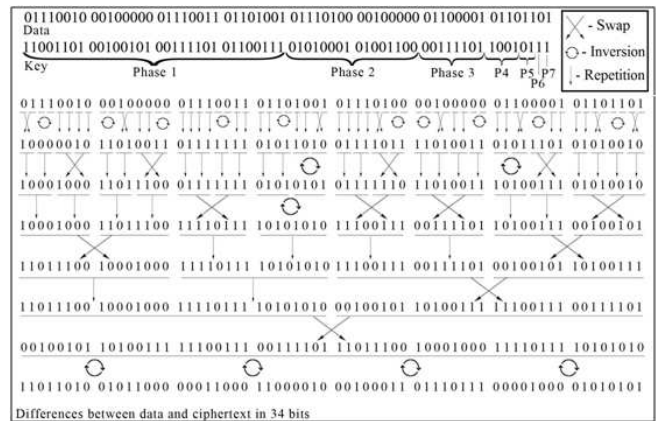| Phase number	Segment length	Number of bits of the key used	Location of the used bits
1	1	32	1-32
2	2	16	33-48
3	4	8	49-56
4	8	4	57-60
5	16	2	61-62
6	32	1	63

Fig. 1. Demonstration of the phases in the encryption of a data block

B. Initial key generation

The problem with initial key generation can be solved by the implementation of asymmetrical cryptographic approaches such as the usage of a public and private key [1, 5]. After the establishment of the initial key safe transmission of following replacement keys can be performed through the data channel as explained further in this paper.

C. Extended multi-phased cipher with doubled block length

The previously described block cipher for each phase of the encryption is applied to a half of the doubled ciphertext, using half of the key. With the unused part of the key, the XOR operator is performed on the unprocessed half of the ciphertext. This phase of ciphertext generation in the example on Fig. 2 is repeated four times to fully exhaust the combinations of semi-key and semi-block. This process is highly customizable in terms of both number of repetitions of the phase and combinations of semi-keys and semi-blocks for the encryption.
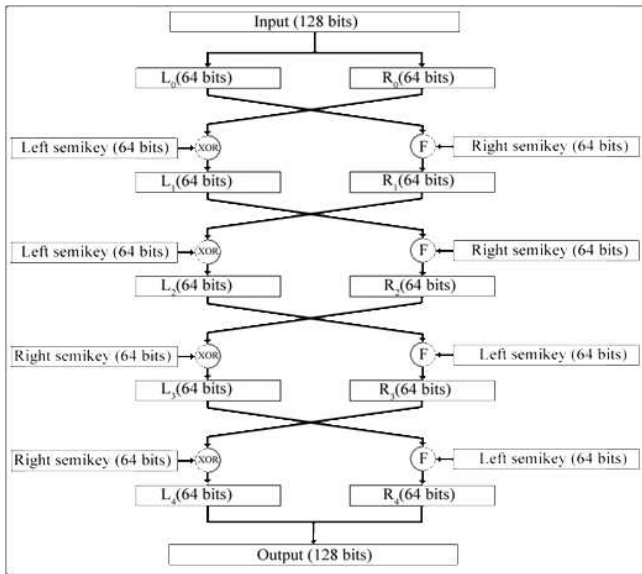
Fig. 2. Demonstration of the multi-phased cipher

D. Aperiodic key replacement

Four bits from each key are selected and form a number between 0 and 15. This number defines how many blocks of data will be encrypted with the given key. This parameter of the key will be referred to as its durability and is between 0 and 15. The position of the four bits defining the key durability is known by both sides in the communication and after the transfer of the said number of encrypted blocks through the channel a key replacement is pending. It is generated pseudo randomly by the sender and each bit of the new key has a 50% probability to be of each of the two logic values. The resulting keys are of unknown number of 'true' bits most often varying between 28 and 36 when the key length is 64 bits and the expected value of this number is half the key length.

Initially, the four bits defining the generated key's durability are positioned at the start of each quarter of the key, at positions 1, 17, 33 and 49. The position of the durability bits is shifted by the durability of the current key and hence increased unpredictability is achieved. The current key is used for the last time when encoding the generated key before its transfer through the channel. This means that before the

substitution of a key, between one and sixteen blocks of data will be encrypted, the last one being the generated replacement key. This process continues until the connection is closed or all data blocks are sent successfully. The randomness of the key durability improves the security of the transmission by increasing the unpredictability of the content of the sent blocks since the replacement keys are mixed with the sent data.

III. EXPERIMENTS

A. Realization of the algorithm and analysis setup

The simulation program has been written on the programming language C++ [2]. The encryption and decryption have been simulated on a single machine with the following characteristics:

- Processor – Intel Core i5-3570 3.4GHz Quad-Core
- Memory – GeIL EVO Corsa DDR3 2133MHz 8GB

B. Performed experiments

- Processing speed comparison with common encryption algorithms
- Analysis of the difference between plaintext and ciphertext
- Analysis of the randomness of key durability
- Performance analysis of the algorithm when using various block sizes

IV. RESULTS

A. Processing speed comparison with common encryption algorithms

The encryption and decryption speed of the proposed cipher with block size of 128 bits and its extended modification, namely Scramble and MultiScramble, has been compared to well established block ciphers such as DES, AES, RC5 and BlowFish. The results of the experiments are shown on Fig. 3 and Table II.

TABLE II
COMPARISON OF ENCRYPTION SPEED FOR VARIOUS ALGORITHMS

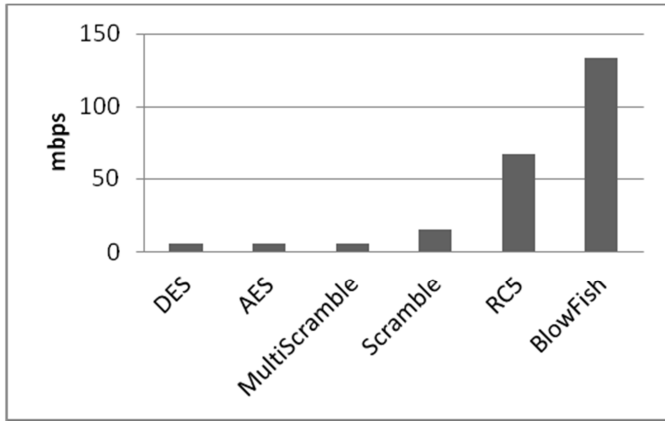|	Seconds to encrypt 1 million symbols	Encryption speed [mbps]
DES	1.41	5.67
AES	1.55	5.16
MultiScramble	1.52	5.23
Scramble	0.51	15.38
RC5	0.12	66.67
BlowFish	0.06	133.33

Fig. 3. Processing speed (encryption and decryption) in [mbps]

C. *Analysis of the difference between plaintext and ciphertext*

The differences between plain and encrypted data have been tracked for 32 768 blocks of length 64 bits. The expected value of these differences is 32, which is half of the block size. The results are presented on Fig. 4 and Table III.

TABLE III
ANALYSIS OF DIFFERENCES BETWEEN PLAINTEXT AND CIPHERTEXT

Differences	Block count	Probability
20	78	0.2%
22	315	1%
24	943	2.9%
26	2166	6.6%
28	3958	12.1%
30	5538	16.9%
32	6491	19.9%
34	5634	17.2%
36	3948	12.1%
38	2178	6.7%
40	996	32%
42	368	1.1%
44	104	0.3%
Total:	32 678	100%
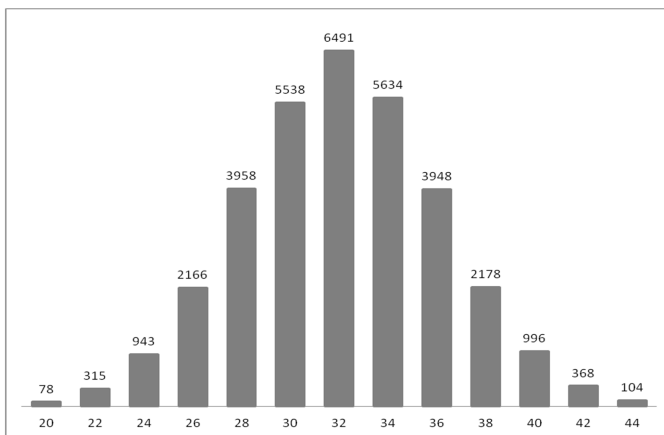
Fig. 4. Analysis of differences between plaintext and ciphertext

D. *Analysis of the randomness of key durability*

The randomness of key durability has been verified by analyzing a chain of 5000 generated keys. The sizes of the resulting key groups are nearly identical, as shown on Table IV, which proves the random nature of the parameter.

TABLE IV
ANALYSIS OF THE RANDOMNESS OF KEY DURABILITY

Durability	Key count	Probability
0	311	6.22%
1	314	6.28%
2	354	7.08%
3	335	6.7%
4	311	6.22%
5	336	6.72%
6	286	5.72%
7	329	6.58%
8	282	5.64%
9	322	6.44%
10	277	5.54%
11	300	6%
12	301	6.02%
13	300	6%
14	300	6%
15	342	6.84%
Total keys:	5000	100%

E. *Performance analysis of the algorithm when using various block sizes*

An evaluation of the time needed to encrypt and decrypt 1 million symbols has been carried out. Its results are shown on Table V and Fig. 5. The data shows that the processing speed is not strictly related to the choice of block size. However, it is clear that the lowest processing time has been reached when using blocks of size 64 bits.

TABLE V
PERFORMANCE SPEED COMPARISON WHEN PROCESSING
1 MILLION SYMBOLS FOR VARIOUS BLOCK SIZES

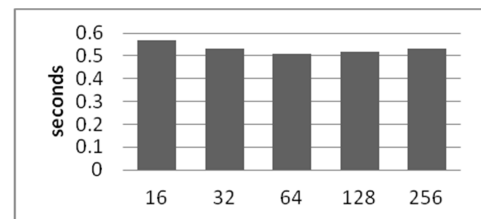| Block size	Block count	Time needed
16	567708	0.57 seconds
32	284402	0.53 seconds
64	142805	0.51 seconds
128	71463	0.52 seconds
256	36284	0.53 seconds

Fig. 5. Performance speed comparison for various block sizes

V. Conclusion

Modern block cipher protocols support several modes of operation to provide the confidentiality for the requirements of different applications. In this paper, a novel structure of multi-phased block cipher with doubled block length is proposed. The aperiodic key replacement is applied to improve the security of data transmission. The use of this method is discussed. Various analysis of the algorithm's performance have been performed, including but not limited to comparison with widely used standards. The algorithm's security and possible vulnerabilities in a running communication system are yet to be evaluated.

References

[1] Menezes A., Oorschoot P., Vanstone S., Handbook of Applied Cryptography, CRC Press, 1996, ISBN: 13-978-0849385230

[2] Schneier B., Applied Cryptography: Protocols, Algorithms and Source Code in C, Copyrighted Material, ISBN: 13-978-0471117094

[3] Mahalingam Ramkumar, Symetric Cryptographic Protocols, Springer International Publishing, Swizerland, 2014, ISBN: 978-3-319-07583-9

[4] [4]. Goots N, Izotov B., Moldovyan A., Moldovyan N., Modern Cryptography: Protect Your Data with FAST Block Ciphers, A-LIST, LLC 295, East Swedesford Rd, 2003, ISBN: 1-931769-12-5

[5] Basin D., Paterson K., Information Security and Cryptography, Springer, 2014, ISSN: 1619-7100.

Path Planning Algorithm for a Robot in a Labyrinth

Milena Karova[1], Ivaylo Penev[2], Ventsislav Nikolov[3] and Danislav Zhelyazkov[4]

Abstract – **The paper presents an algorithm for path planning for a robot in a labyrinth. The algorithm uses an image, obtained by a camera. The image is processed and converted to a matrix, presenting the labyrinth with obstacles and walls. Afterwards an algorithm, based on the Dijkstra's algorithm, is applied to find the shortest path in the labyrinth. As opposed to the classical Dijkstra's algorithm, the presented algorithm compares the size of the robot to the size of an obstacle. The implementation of the algorithm is described and the obtained results are presented.**

Keywords – **Path planning, Robot, Labyrinth, Dijkstra algorithm, Wave moving process.**

I. INTRODUCTION

In the past the orientation of a human which is in a new unknown place has been difficult. That is why the researchers have spent some time for walking, orientation and map creating for the unknown areas. Thus the geographic maps emerged to facilitate the orientation in unknown location. Later the orbital satellites have been invented and used to take pictures and automatically map the ground surface.

After the invention of computers, information technologies and programming tools the humans are trying to make a perfect electronic anthropoid called robot. Artificial intelligence is intended to be introduced in it to make it like a human. In order such intelligence to be qualitative and reliable it should be able to produce at least some thinking operations and, for example, orientation.

Many algorithms and methods have been studied for planning path of robots [5]. Great attention has been given to Genetic algorithms [2], A* algorithm [3], as well as other naturally inspired optimization algorithms [4].

In most cases the map of the environment (with obstacles and possible paths) is obtained through the sensors of the robot [1]. This approach serves well for movement of robots on terrestrial surfaces, but could be inapplicable if the robot's sensors could not be used or sensors usage is difficult. For example this is the case with movement in air or underwater areas.

This paper describes an application of a project in which a robot, labyrinth and camera are used. The camera shots and sends the pictures to the robot which is moving through the labyrinth. If the labyrinth is changed the camera informs the robot. Thus the following operations should be implemented: analyzing of the labyrinth image, finding of the optimal path, making of the path map, moving simulation. The main goal is to find the shortest path in a labyrinth given by a picture. The walls, obstructions, the robot and the exit of the labyrinth are the input data and should be marked. The output of the system is the labyrinth printed in a file and animated simulation of the robot moving.

The application goes in processing mode when a picture is selected.

II. PROCESSING MODE

A. Image processing and building of virtual labyrinth

A picture of the labyrinth is obtained by a camera, attached to the robot. Additionally on the picture the initial position of the robot is marked by a red square and the exit of the labyrinth is marked by a green square (fig. 1).
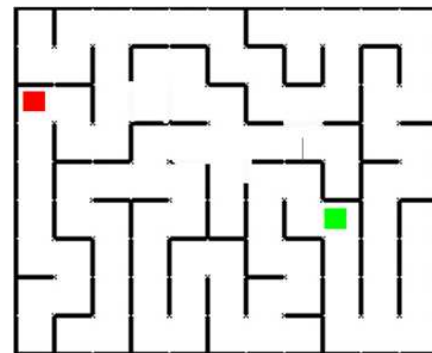

Fig.1 Picture of a labyrinth

TABLE I
SYMBOL ACCORDING

RGB	color	meaning	symbol
>200,>200,>200	light	space	' '
>200,<100,<100	nuance red	start / robot	'*'
<100,>200,<100	nuance green	end / exit	'o'
other	other	obstacle / wall	'W'

In this stage every pixel of the image is analyzed and a map is created as a two-dimensional array. Every pixel is transformed to a symbol according to Table 1. The pixels' coordinates correspond to the symbols positions in the map.

[1]Milena Karova is with the Department of Computer Science at Technical University of Varna, Bulgaria, E-mail:mkarova@ieee.bg
[2]Ivaylo Penev is with the Department of Computer Science at Technical University of Varna, Bulgaria, E-mail: ivailo.penev@tu-varna.bg
[3]Ventsislav Nikolov is with the Department of Computer Science at Technical University of Varna, Bulgaria, E-mail: v.nikolov@tu-varna.bg
[4]Danislav Zhelyazkov is a student at the Department of Computer Science at Technical University of Varna, Bulgaria, E-mail: d.zhelyazkov.7331@gmail.com

When green or red pixel occurred some additional computations are made to find the minimal and maximal values of coordinates of the robot position and the labyrinth exit, i.e.:

$(Xr_{min}, Yr_{min} : Xr_{max}, Yr_{max});(Xe_{min}, Ye_{min} : Xe_{max}, Ye_{max})$.

The width of the exit and robot are also calculated:
$k=max((Xr_{max}-Xr_{min}),(Yr_{max}-Yr_{min}),(Xe_{max}-Xe_{min}),(Ye_{max}-Ye_{min}))$.

B. Optimal path in the virtual labyrinth

Thus the algorithm guarantees that the robot will go through wide enough paths.

The previous data is integrated within the programming model of the labyrinth. It is presented as a global object, called data transfer object (DTO), accessible from any other part of the application including all interface implementations. This global object is a pure data object without any functionalities and it is able to self-validate and convert itself to plain text. The validation aims to refuse invalid or incorrect pictures.

The virtual labyrinth can be constructed either by analyzing a picture or by a scanning stream. The virtual labyrinth after its processing is shown in Fig.2.
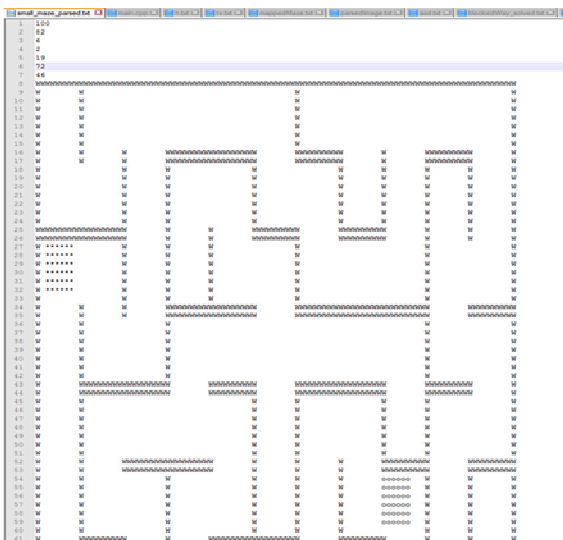
Fig.2 Virtual labyrinth shown as specially ordered symbols

III. FINDING THE OPTIMAL PATH IN THE LABYRINTH

This is the main step in the application logic. Finding the path is based on the well-known Dijkstra algorithm. The difference is, that the presented algorithm compares the size of the robot with the size of an obstacle.

A wave starting from the end of the labyrinth is observed as a final unit that is moving from one point to a next neighbor point. The wave gradually marks all points (units) directed to the final point – Fig.3. This idea is further developed in the application and is called Gasolisation. The difference is that in our case the robot, respectively the final, are with different width compared to the width of the walls and paths. The marking points are replaced with marking lines (sequences of points) with length k. The current traversing line in fact does not search neighbor points but the whole neighbor lines. If the robot is found the main wave stops it's spreading and a new small wave starts to spread trying to mark the robot. If this does not succeed then this means that the area locating the robot is too narrow and then the small wave stops and the main wave continues. Otherwise if the robot is successfully marked the algorithm is completed. If the main wave cannot continue, because the all lines are marked, then an exception is thrown saying that a path is not found. The algorithm works on the characters file, that represents the virtual labyrinth, and the marking is done by using the symbols '^', '>', 'v' and '<'. The robot follows these symbols to move to the exit, i.e. the symbols mark the back trace for the robot.
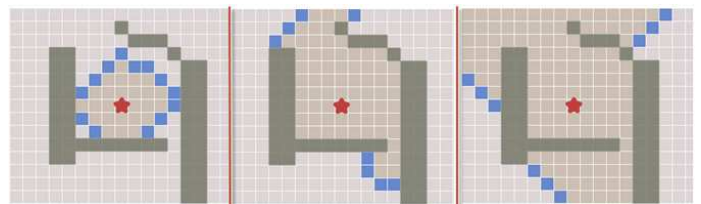

Fig. 3 The wave moving

The algorithm consists of the following basic steps:
- Creation of a queue – a set of points;
- Extraction of elements from the queue;
- For each point all neighbor points are checked. The possible neighbors are in four directions – up, right, down, left.

A neighbor point is free, if the following conditions are satisfied:
- the point is not a part of an obstacle;
- the point is not marked.
A free point is marked and added to the queue.

All free points are marked by its neighbors. The order of marking forms the shortest path.

The whole algorithm is as follows:

```
void FindPath( dot startDot, dot finalDot ) {
  Queue<dot> justAQueue;
  justAQueue.Add( finalDot );
  while (dot currentDot = justAQueue.Pop()) {
   Array<dot> neighbours = currentDot. GetNeighbours();
   foreach (dot neighbour in neighbours) {
   If( neighbor.isObstacle || neighbor.isMarked )
    next;
   currentDot.Mark( neighbor );
   if ( neighbor == startDot)
    return;
   justAQueue.Add( neighbor );
   }
  }
throw new NoPathFoundException();
}
```

IV. EXPERIMENTS - ANIMATION WITH SIMULATED MOVING

The application prints as a text file the marked virtual labyrinth. The file contains information for the size of the labyrinth and the robot as well as the coordinates of the start and end points. The walls and obstructions are shown as 'W' character, the path to the final is shown with the symbols '^', '>', 'v', '<' and the final is shown as 'o'. This file can be exported and easily used by any other software application.

The application is implemented using the Java Swing technology for user interfaces. When the process is completed a maze monitor is shown as a panel in the main user window sharing the virtual labyrinth. This monitor is responsible for the reverse transformation of the characters map to a graphical picture. A maze command executor is also used to move the robot according to the markers over which it is currently placed. If there are not markers in the robot position then this means that the target is reached or there is no path in the labyrinth. In these cases the robot cannot move any more. The executor moves the robot by a timer until a situation with an impossible move is reached (Fig. 4).
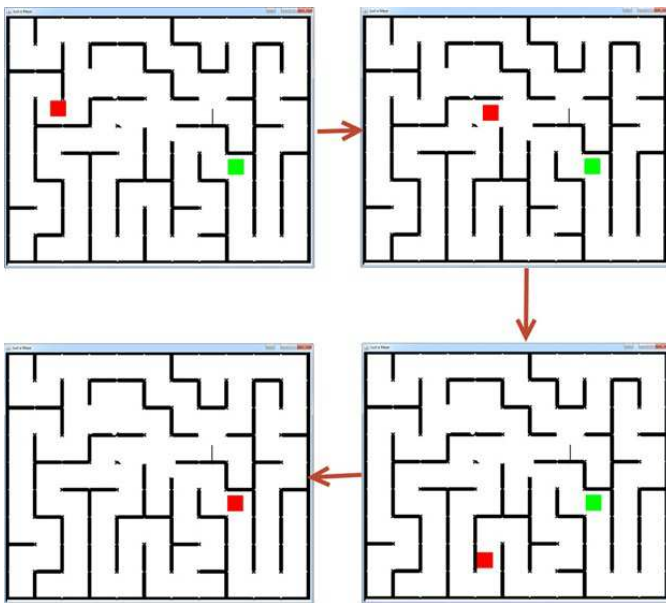
Fig. 4. Animated simulation of the robot movement

V. RESULTS

File with a model of the labyrinth with walls and obstacles and the paths found is obtained as a result of the algorithm (Fig. 5).
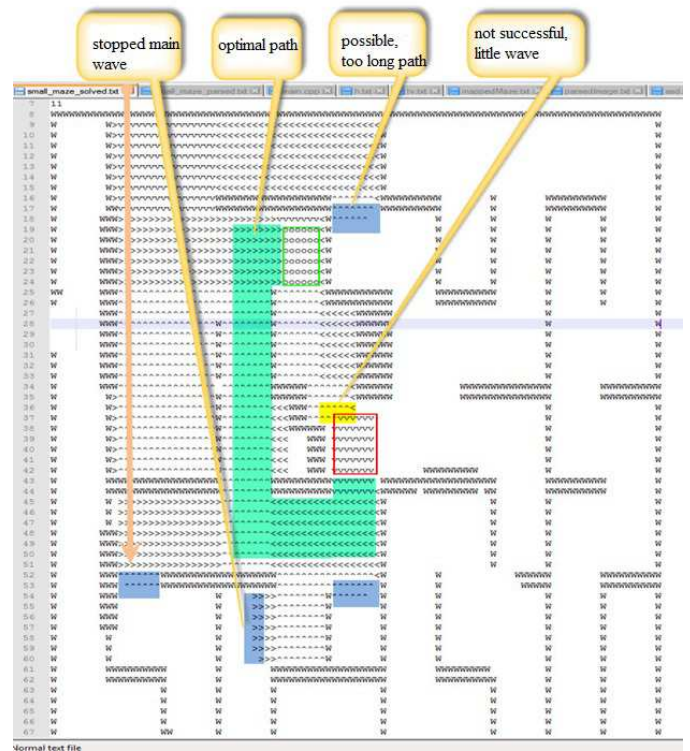
Fig. 5. Model of the labyrinth with found paths

The output file has the following structure:
- Numeric data about:
  - size of the labyrinth;
  - size of the robot and the exit of the labyrinth;
  - coordinates of the start and the exit.
- Symbol map with the following data:
  - walls and obstacles – symbol 'w' is used;
  - path to the exit – symbols '^' , '>' , 'v' , '<' are used.
  - exit – symbol 'o'.

The output, shown on Fig. 5, presents:
- several thin slices;
- unsuccessful little wave;
- successful stop of the wave;
- several possible paths.

A file with this structure could be used by other applications, using the same programming interfaces.

VI. CONCLUSION

For the present we consider significant application of the presented algorithm in the field of technology education. The algorithm has been implemented in a real robot platform (in our case LEGO EV3 robot). This way we could demonstrated to students fundamental concepts in computing and automation: path-finding and search algorithms, robot programming, device motion control, finite state machines, others. Furthermore the algorithm is a suitable basis for comparison of different path finding algorithms, for example breadth-first and depth first search with A* algorithm.

One of the most important optimizations of the proposed algorithm should be its parallel realization. This could also make possible dynamic changes like manual choice of the robot target and manual change of the labyrinth. Another future task is implementation of another searching algorithm.

REFERENCES

[1] A. Rodic, Navigation, Motion Planning and Control of Autonomous Wheeled Mobile Robots in Labyrinth Type Scenarios", Volume 8, Number 2, Intelligent Service Robotic Systems, IPSI Journal, Transactions on Internet Research, TIR, ISSN 1820 - 4503, pp. 2-9, 2012.

[2] J. Su, J. Li, Path Planning for Mobile Robots Based on Genetic Algorithms, Proceedings of Ninth International Conference on Natural Computation (ICNC), ISBN: 978-1-4673-4714-3, pp. 723-727, 2013.

[3] N. Sariff, N. Buniyamin, An Overview of Autonomous Mobile Robot Path Planning Algorithms, Proceedings of 4th Student Conference on Research and Development, ISBN: 1-4244-0527-0, pp. 183-188, 2006.

[4] S. Muldoon, L. Chaomin, F. Shen, H. Mo, Naturally Inspired Optimization Algorithms as Applied to Mobile Robotic Path Planning, IEEE Symposium on Swarm Intelligence, ISBN: 978-1-4799-4458-3, pp. 1-6, 1994.

[5] http://www.redblobgames.com/pathfinding/a-star/introduction.html

POSTER SESSION
CONTROL SYSTEMS

Acquisition of U/I characteristics of electric motors with permanent magnet using LabView software

Borivoje Milošević [1] , Slobodan Obradović[2] , Srdjan Jovković[1] and Slavenko Djukić [3]

Abstract – **This paper deals with creating user interfaces and software for measurement, recording and playback of U/I characteristics of electric motors with permanent magnet in real time, using application called LabView of the National Instruments company. Paper deals with measuring and displaying these characteristics under different engine loads, as well as measuring its speed. The realization of these tasks implied the design and the construction of an additional hardware realized by microcontroller, whose task is to convert an analog voltage with an acquisition card with the PWM signal thereby adjusting the engine design needs in order to record and display its characteristics.**

Keywords – **USB, PWM, DAQ, CASE, AD, USB-6009**

I. INTRODUCTION

The paper describes a software solution to adjust and control data acquisition and display of data within experimental methods for the determination of UI features of the electric motor over a wide load range. Apart from the software, written in a programming package LabVIEW, an example of its application is displayed together with the results in a separate chapter.

Development of the software supported system, applied to testing of instruments for monitoring and analysis of the basic UI and speed parameters is presented in this paper. This acquisition system includes two functionally connected parts. First part involves generation of the standard waveforms, including possibility for simulation of the various signals typical for electrical UI power and speed acquisitions. Software support of this procedure performs graphical presentation of the previously generated and recorded signal waveforms. Second part of this procedure includes real-time recording and presentation.

 System is supported by virtual instrumentation concept, which includes control software application in LabVIEW environment and data acquisition card USB-6009.

Data acquisition (DAQ) is the process of measuring an electrical or physical phenomenon such as voltage, current, temperature, pressure, or sound with a computer. A DAQ system consists of sensors, DAQ measurement hardware, and a computer with programmable software. Compared to traditional measurement systems, PC-based DAQ systems

exploit the processing power, productivity, display, and connectivity capabilities of industry-standard computers providing a more powerful, flexible, and cost-effective measurement solution. An example is given in Fig 1.

Fig. 1. Data acquisition system

The parts of the system are:
- Physical input/output signals
- DAQ (Data Acquisition) device/hardware
- Driver software
- Application software

Hardware component used in this paper is a USB-6009 card of the National Instruments company. A block diagram of the card is shown in Fig 2.
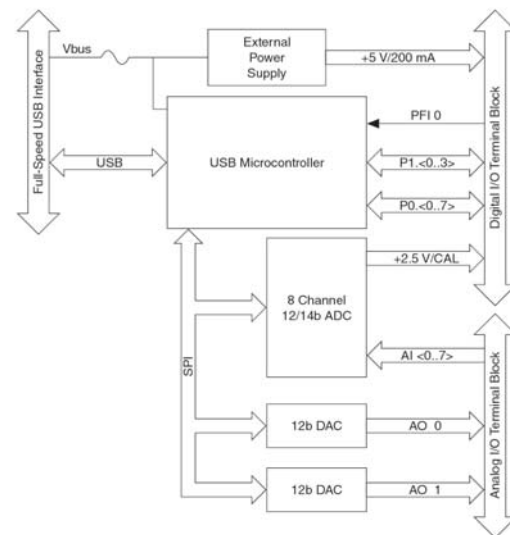

Fig. 2. Block diagram of the card USB-6009

Card USB-6009 includes:
- Four analog inputs of the differential relation (14-bit resolution, maximum sample rate 48KS/s, the input voltage at the junction of up to ± 10 V)

[1]Borivoje Milošević is with the College of Applied Technical Sciences of Niš, 20 A. Medvedeva, Niš 18000, Serbia, E-mail: borivoje.milosevic@vtsnis.edu.rs.
[2]Slobodan Obradović, VISER, 11000 Beograd, Serbia, slobo.obradovic@gmail.com
[3]Slavenko Djukić is with the College of Applied Technical Sciences of Niš, 20 A. Medvedeva, Niš 18000, Serbia.

- Two analog outputs (12-bit resolution, voltage 0-5 V, current up to 5mA, refreshing outputs with 150 Hz)
- Twelve digital inputs/outputs (there are two ports, port P0 has 8 lines that can be individually defined program as input/output marked with P0 - P7 and P1 port that has four lines that can be individually defined program as input/output, marked with P1 - P3.
- 32-bit counter with full speed USB interface (activated on the falling edge of the signal, counting only the advance, the maximum input frequency of 5 MHz).

Hardware components of the printed circuit board were created with the aim of controlling the speed of rotation of DC motors with permanent magnet. Electronics is realized using a microcontroller whose primary task is to convert an analog voltage of 0 - 5V, which is obtained from one of the two analog outputs of the USB 6009 card defined in the programme, into the corresponding PWM (Pulse Wide Modulation) signal with a ratio of 0 - 100% of the value which is proportional to the analog voltage dependent upon referrential voltage of the card. The microcontroller is programmed so that it can generate three different periods of the PWM signal while the choice available via buttons on the printed circuit board can select a range of 15.6 KHz, 3.8 KHz or 980Hz in the full range of 0 - 100% depending on the selected electric motor.

The load circuit consists of transistors and amplifiers responsive to a programmed digital outputs on the USB - 6009 card. Four relays are inserted through the load as active loads of different ohmic resistance through which electric generators indirectly charge.

A hardware block diagram of the complete part is shown in Fig 3.
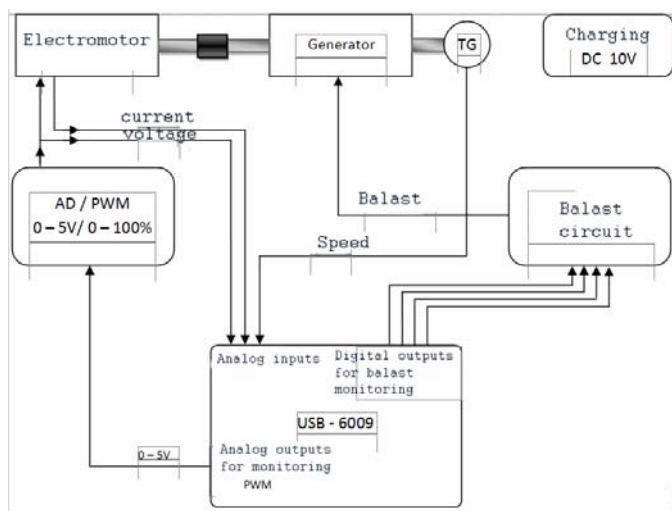
Fig. 3. Block diagram of the complete hardware part

Unlike other programming languages, LabVIEW uses a graphical programming language "G" to create programs in the form of block diagrams. LabVIEW uses terminology, icons, and ideas close to scientists and engineers and relies on graphical symbols rather than textual languages to describe the actions of the program as part of CASE tools. For lovers of C, LabVIEW has left open the possibility of programming in text mode commands. Programming of these possibilities has just been used in this paper. LabVIEW contains a huge number of functions and subroutines for most programming requirements that may occur in practice. In addition to Windows, LabVIEW supports the work of the Macintosh, Sun and Linux OS and developed its own OS. Fig 4 is showing Front Panel work tools and structures that contain Block Diagram.
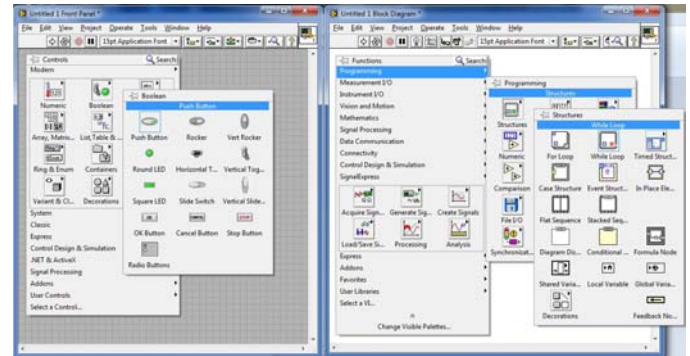

Fig. 4. Front panel

II. CREATING A PROGRAM FOR MEASURING THE U/I CHARACTERISTICS OF THE MOTOR

The program is designed as a whole in one window, where there will be two oscilloscope displays, one which will display waveform voltage to the engine Fig 5 - mark 1 and the other which will display the current flowing through the engine in real time Fig 5 - mark 2. Regulation of the output analog voltage 0 - 5V cards with USB - 6009 (exit A00) was performed programmatically by virtual vertical slider, the respective slider on the right side of the window labeled with PWM, Fig 5 - mark 5. This voltage is directly led to the input of the hardware part of the microcontroller whose task is to measure the AD voltage converter and generate a PWM signal which is in direct proportion to that voltage. Practically, the analog voltage whose value ranges from 0 - 5V has proportional PWM signal 0 - 100% and therefore the motor rotation speed control from 0 - 100%.
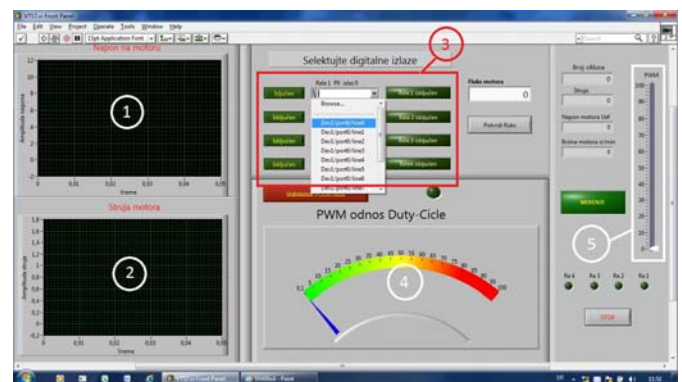

Fig. 5. The appearance of the user interface

The part of the program that deals with selecting digital outputs on the USB - 6009 card, enables that charging can be connected to any of the outputs, which is made possible by a drop-down list in the framework of Fig 5 - mark 3 in which the channels are selected thus becoming physically connected. Programming makes it possible to isolate each of these digital outputs which can be at any time switched on or off via the corresponding switch on the left drop-down list, as well as to indicate the relay activity at any time, allowing a visual overview of the current state of the relay involved, ie ohmic load. Vertical slider or slider, Fig 5 - mark 5, defines compliance with PWM signal or voltage that is entered to the motor.

A. Defining analog input channels

The procedure for defining the analog inputs is as follows: it is necessary to have a connected USB card - 6009 via a USB port that we could access or to define the analog inputs. In the block diagram, right-click on the desktop opens the pop-up menu from which we choose the icon DAQ Assist in Fig 6.

Fig. 6. Defining the analog inputs

We can define three analog channels: voltage_0, voltage_1, voltage_2, an rename them too, through which we measure speed, voltage and current of the motor, Fig 7.
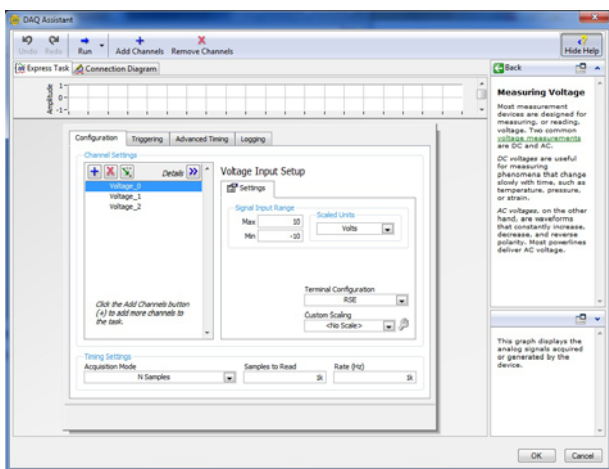
Fig. 7. Adjusting the selected channels

B. Defining analog outputs channels

Defining the analog output to generate a voltage of 0-5 V, which we use to manage PWM is performed as follows. Right click on the desktop block diagram opens the menu shown in Fig 8.
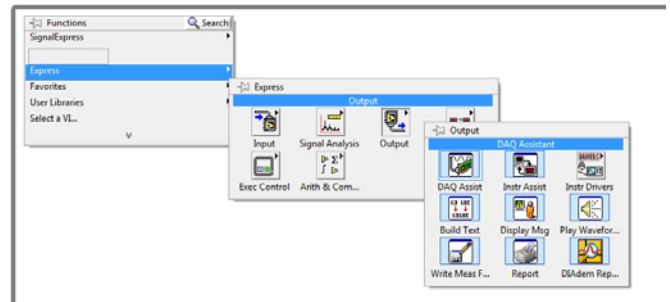

Fig. 8. Defining the analog outputs

The next activity is to define the scope of the output voltage, which is entered so that the minimum voltage value is 0V and 5V maximum value (which is the largest value of the output voltage which can be defined for this card).

The procedure of definition of the analog outputs is similar to the procedure of defining analog inputs.

C. Creating a graphic (oscilloscope) display

For the oscilloscope signal display it is needed to add Waveform graph of voltage and current in the Front panel.

To show the effective voltage value of the system, it is necessary to perform an expression that defines the effective value of the PWM signal that has a form used in this paper.

Effective voltage value is defined by the equations Eq. 1.

$$U_{ef} = \sqrt{\frac{1}{T}\int_0^T u^2(t)dt}$$

(1)

Since in our case the signal - pause can have any value, we need to define the effective value of the signal in every possible case. This is necessary in order to programme the system to display the correct values for any given voltage of PWM signal.

Fig 9 shows the PWM signal, which in general can have any signal-pause relationship, and variable is defined by variable k within the periods which can be defined at any time.

Calculation of the effective value during periods start from the equations Eq. 2.

$$U_{ef} = \sqrt{\frac{1}{T}\int_0^T u^2(t)dt} = \sqrt{\frac{1}{T}\int_0^T U^2_{max}(t)dt} = \sqrt{\frac{1}{T}\left[\int_0^{kT} U^2_{max}(t)dt + \int_{kT}^T 0(t)dt\right]}$$

$$U_{ef} = \sqrt{\frac{U^2_{max}}{T} \cdot \Big|^{kT}_0 t} = \sqrt{\frac{U^2_{max}}{T}[kT-0]} = \sqrt{\frac{U^2_{max}}{T} kT} = U_{max}\sqrt{k} \qquad (2)$$
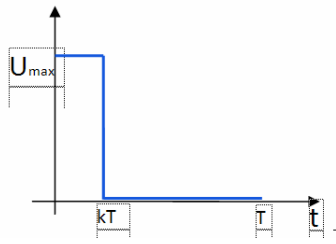

Fig. 9. PWM signal

The resulting expression is entered in the program to help that displaying the correct effective value of the supply voltage. Fig 10.
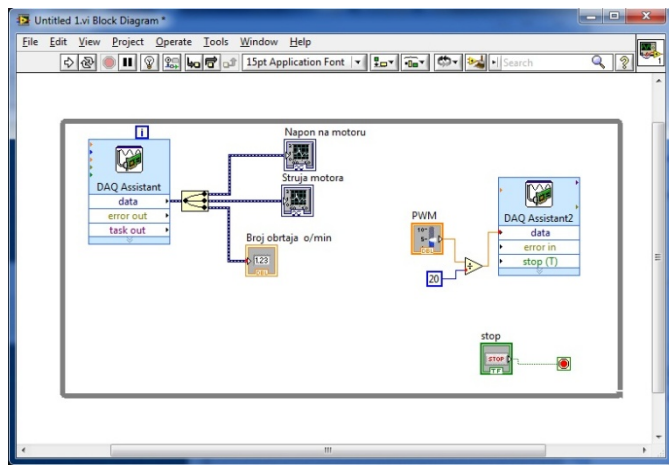

Fig. 10 System block diagram

D. Defining the digital outputs

Program is defined by CASE structure (central rectangular part of the Fig 11.) which allows the inclusion or exclusion of each digital output individually.

In the same way, the other three digital outputs that form a functional unit with the rest of the software program are defined.
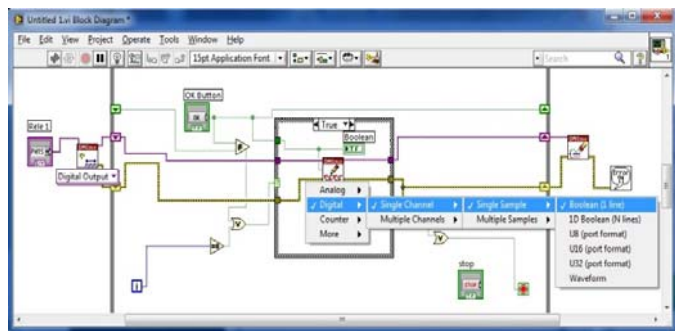
Fig. 11 Defining the digital outputs

Programming the connection and disconnection of individual relay which will include the active resistors is generated in C language.

The final look of the front panel is shown in Fig 12 where we can see all the possibilities that are built into the program and we can also record the signal voltage and current on the motor during operation under load.
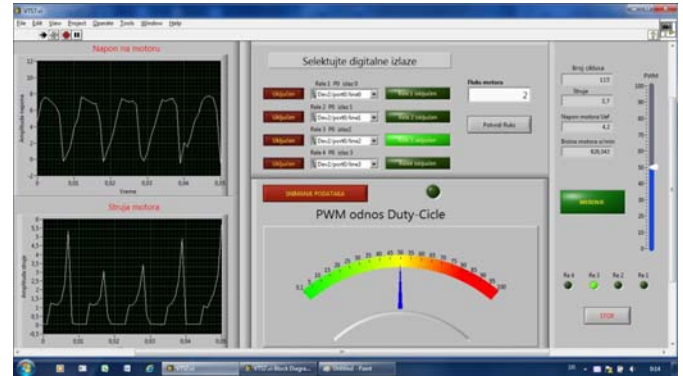

Fig. 12. The final look of the front panel

III. Conclusion

This paper describes the intelligent controller designed and developed in LabVIEW for measurement the U/I characteristics and motor speed of electric motors with permanent magnet in real time.

One benefit of LabVIEW over other development environments is the extensive support for accessing instrumentation hardware. Drivers and abstraction layers for many different types of instruments and buses are included or are available for inclusion. These present themselves as graphical nodes. The abstraction layers offer standard software interfaces to communicate with hardware devices. The provided driver interfaces save program developments.

References

[1] S. Haykin, "Neural Networks", New York, IEEE Press, 1994.
[2] Jovitha J., Aravind A. P. , Arunkumar V. Balasubramanian P. "LabVIEW based Intellegent for Speed Regulation of Electric Motor " IEEE Transactions, Instrumentation and Measurement Technology Conf., IMTC. Vol 2, Issue , Pages: 935 – 940, 16-19 May 2005.
[3] National Instruments Co. "Introduction to LabVIEW" , Austin (USA) , 2003
[4] Nihat Ozturk, and Emre Celik. "Speed control of permanent magnet synchronous motors using fuzzy controller based on genetic algorithms" Electrical Power and Energy Systems 43 (2012) 889–898
[5] Megha Jaiswal, and Mohna Phadnis. "Speed Control of DC Motor Using Genetic Algorithm Based PID Controller" Volume 3, Issue 7, July 2013 ISSN: 2277 128X International Journal of Advanced Research in Computer Science and Software Engineering
[6] HANS PETTER HALVORSEN, "Data Acquisition in LabVIEW". Faculty of Technology, Postboks 203, Kjølnes ring 56, N-3901 Porsgrunn, Norway.

Software Model for Optimization of Infill Balise Location in ETCS, Level 1 in Railway Transport

Taschko Nikolov[1] and Georgi Ganchev[2]

Abstract – **A drawback of railway ETCS, L1 is that a change of the signal aspect of a signal cannot be transmitted dynamically in the on board system. For this reason infill balises are used. Their location is a matter of discussion. A mathematical model is demonstrated here, which allows to optimize the location of infill balises, as the time gained from unnecessary halting and accelerating must be maximal, under certain limit conditions.**

Keywords – **ETCS L1, infill, braking curves calculation.**

I. INTRODUCTION

In many countries worldwide, including beyond the borders of the EU, the European standardized system for automatic train protection ETCS, Level 1, is implemented. With this system of the so-called balises, located along the railway, information about the signal aspects of the signals is transmitted to the locomotive. Usually, the signal aspect shows two velocities – the one at which the current signal should be passed and the one at which the next signal should be passed. When the signal aspect of the signal changes to more permissive after the locomotive has already passed the previous signal, there is no way for the locomotive equipment to update its information because the previous signal balise has already been passed. Thus, the locomotive is compelled to move at a lower velocity than that actually prescribed, and in some cases even halt, if the aspect before its updating has been "stop." This leads to unnecessary braking and accelerations, resulting in longer travelling time and power loss. In order to avoid such losses, the so-called infill balise is installed before some signals. The infill balise location is disputable and depends on the velocities at which the train passes the previous signal, the distance to it, and differs for the various railway administrations. In order to determine the most appropriate location of the infill balise, a software product is developed for automatic calculation of the infill balise location based on various input data. It is based on a number of analytical calculation methods from the area of mathematics, physics and safety systems in the field of railway transport. Various additional data are also calculated, which are useful for designing the section in which the infill balise is to be located. The product has a graphic interface facilitating the work with it and data visualization.

[1]Taschko Nikolov is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria.

[2]Georgi Ganchev is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: gantchev_g@yahoo.com.

II. PRINCIPLES OF BUILDING THE MATHEMATICAL MODEL

As in every mathematical model, some approximations are accepted here as well, sufficiently close to the real-life situations, allowing the use of standard mathematical and physical functions.

The basic adopted approximation is that braking and accelerating curves are parabolas whose vertex M is on the ordinate axis (Fig. 1). This type of parabola is the basis of the approximations of the given model and is preset with the following dependences:

- parabola equation : $y(x) = ax^2 + c$;   (1)

roots: x_1, x_2 as $x_1 = -x_2$

parametre a – always negative $(a < 0)$

parametre c – always positive $(c > 0)$

vertex M with coordinates $M(0,c)$;

discriminant $D = b^2 - 4ac \Rightarrow D = -4ac$ always $D > 0$

- Vieta's formulae

$$x_1 + x_2 = -\frac{b}{a}; \qquad x_1 x_2 = \frac{c}{a}$$

If the linear relocation l is taken into account, the function will appear as follows:
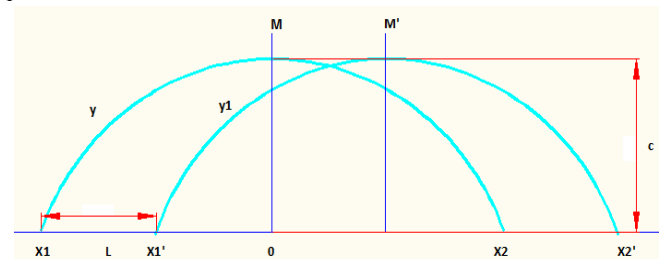
$$y_1(x) = a(x - l)^2 + c .$$

Fig. 1. Graphical view of equations

III. MODELING THE BRAKING AND ACCELERATING BEHAVIOUR

A. Braking curve

The braking curve is approximated as part of a parabola and calculated as a velocity function of the way– $V = f(s) = as^2 + c$, as V indicates the instantaneous velocity in a point of the way S. For easier perception of mathematical calculations the above formula is used $- f(x) = ax^2 + c$. The initial point of the coordinate system $O(0,0)$

coincides with the location of the so-called warning signal (*Пс*). The target point or the point where the train must stop is the entry signal (*Bx*) – point x_2 – the greater positive root in the equation. The value of the coefficient *c* is identical with the velocity at which the train enters at the warning signal. The position of the warning signal, entry signal and entry velocity is determined based on traction estimates which are not subject of this study and are accepted as preset. Due to the specificity of the subject matter, the measurement units are respectively preset as S – way [m], *V* – velocity [km/h], *t* – time [*s*].

In calculating the curve, two main problems are solved:

- to calculate the quadratic function of the parabola for preset location of the warning and entry signal and entry velocity at the warning signal;

- to calculate the time at each point of the way - $t = f(s)$ for preset location of the warning and entry signal and entry velocity at the warning signal,.

After the respective transformations and derivations, the analytical expressions for the sought curves are obtained. The graph of the time function is of logarithmic type and is shown in Fig.2. The graph of the velocity function is shown in Fig.3.
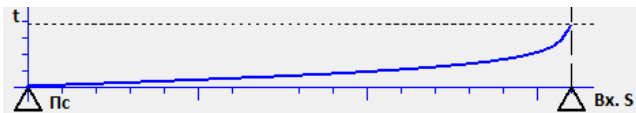
Fig. 2. Deceleration time function

Fig.3. Deceleration velocity function

B. Acceleration curve

The acceleration curve is approximated as a part of a parabola (Fig.3 quadrant II) and is calculated as a velocity function of the way – $V = f(s) = as^2 + c$ as by *V* we indicate the instantaneous velocity in a point of the way *S*. For easier perceiving of the mathematical calculations, we will use the symbols generally accepted in mathematics – $f(x) = ax^2 + c$. We perceive as an initial point of the coordinate system $O(0,0)$ as the location of the warning signal. Due to the fact that the acceleration curve (unlike the part of the primary parabola located in quadrant II) is situated

in quadrant I, and its initial point is located after the warning signal, in the curve equation a horizontal relocation is added relative to the primary parabola (Fig.1). Therefore, the equations of the acceleration curve are accordingly transformed: – $V = f(s) = a(s - l)2 + c$ and $f(x) = a(x - l)^2 + c$. The parametre *l* is always greater than 0 and greater than the distance between the warning and entry signal, i.e., $(l > x_2)$ (Fig.1).

The maximum velocity of the acceleration curve is determined by the specific situation and is accepted as preset. Due to the specificity of the subject matter, the measurement units are preset accordingly: S – way [m], V – velocity [km/h], t – time [s]. Upon calculation of the acceleration curve two main problems are solved:

- to calculate the parabola's parametres for preset location of the warning and entry signal, maximum acceleration velocity, starting acceleration point of $0\ km/h$ and preset maximum acceleration;

- to calculate the instant in each single point – $t = f(s)$ for preset location of the warning and entry signal, the maximum acceleration velocity, starting acceleration point of $0\ km/h$ and preset maximum acceleration.

After the respective transformations and derivations, the analytical expression for the sought curves and drawn graphs (Fig.4) are obtained. The graph of the time function is of logarithmic type.
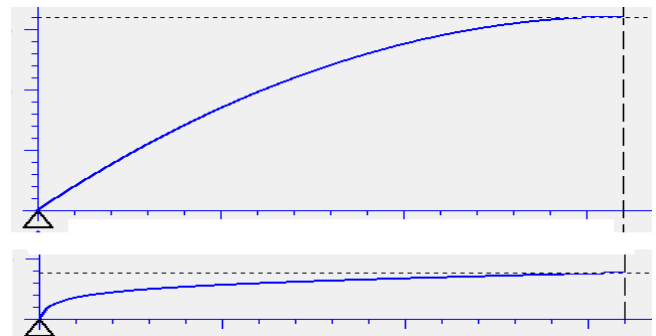

Fig.4. Acceleration, velocity and time functions

IV. DETERMINING THE INFILL BALISE LOCATION

A. Movement with constant acceleration and with constant deceleration

Here we will provide some basic formulae and graphs for movement with constant deceleration (Fig.5) and with constant acceleration (Fig.6), reworked for specific purposes, which are used in further calculations. The following symbols are used in the figures:

- initial velocity –u ;

- final velocity – v ;

- travelled way – ΔS (Δx) ;

- acceleration – a ;
- time for travelling a way with definite initial and final velocity – Δt

- average velocity $- v_{avg} = \dfrac{u + v}{2}$

where:

$u < v$ – with constant acceleration $a = const > 0$; $u > v$ with constant deceleration $a = cost < 0$;
$u = v - a = 0$;

Main equation:

$$ \text{(2)} $$

Travelled way:

$$ \Delta x = V_{avg} \times \Delta t \tag{3.a} $$

$$ \Delta x = u \times \Delta t + \frac{1}{2} a \Delta t^2 \tag{3.b} $$

$$ \Delta x = \frac{v^2 - u^2}{2a} \tag{3.c} $$

Acceleration:

$$ a = \frac{v^2 - u^2}{2 \times \Delta x} \tag{4} $$

Time:

$$ \Delta t = \frac{\Delta x}{V_{avg}} = \frac{2 \times \Delta x}{u + v} \tag{5} $$
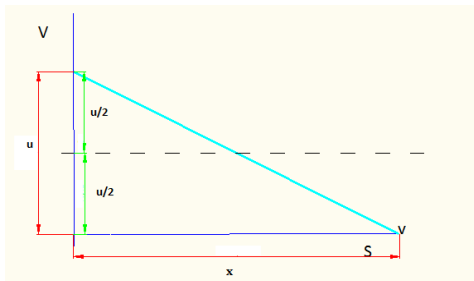
Fig.5. Average velocity calculation
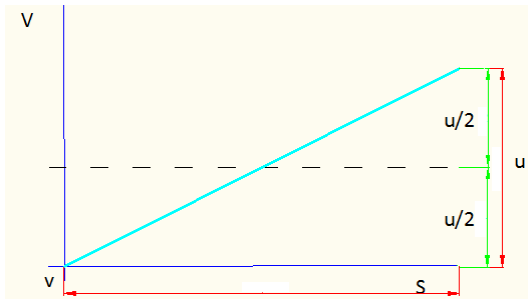
Fig. 6. Average velocity calculation

B. Determining the infill balise location

In railway transport the function of train velocity of the way – $V = f(s)$ is considered as basis for initial calculations. Time

is a determining factor in calculating the dependences between the separate elements of the law of motion: travelled distance, velocity, acceleration. That is why in creating the mathematical model of great significance is to find the dependences of the way, velocity and acceleration on time. Upon preset law (equation) $V = f(s)$ are calculated the functions (equations) $S = s(t)$, $V = v(t)$, $A = a(t)$, where S – way, V – velocity, A – acceleration, as velocity is the first derivative from $s(t)$, and the acceleration is the second derivative from $s(t)$, i.e.,

$$ V = s'(t), \qquad A = v'(t), \qquad A = s'' $$

The location of the **infill balise** is determined by two main factors, which are related in inversely proportional dependence. Both factors are based on the time t_1, for which the train travels the distance from the warning signal to the **infill balise and the time t_2, for which the train travels the distance from the infill balise to the entry signal. Significant impact is exerted by the human factor, and yet it is very difficult or impossible to take this into account mathematically, and therefore, it is neglected in calculations, save in some very specific cases. The first factor is the probability for the aspect at the entry signal to change for the time t_1 . The closer** to the entry signal an **infill balise is, the greater** the probability for the change to be read, and to respond to it appropriately, will be. The second factor is the gained time and power upon timely detecting of the aspect's change. **The closer** to the warning signal an **infill balise is located, the greater the** gain of time and energy will be. Several methods **adopted in global practice** exist for calculation of the **infill balise** position**.**
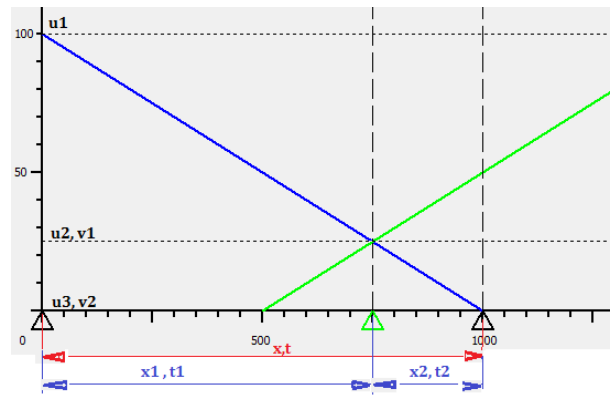

Fig.7. Theoretical probability calculation

The most widely spread method (that may as well be defined as classical) is based on the movement with constant deceleration (acceleration) (Fig.7). According to it, the probability for the aspect of the entry signal to change when the train is between the warning signal and the **infill balise and the probability** for the aspect of the entry signal to change when the train is between the **infill balise and the entry signal must be equal. This means that** $t_1 = t_2 = \dfrac{1}{2} t$.

Upon detailed development of the kinematic formulae it is proved that, irrespective of the entry velocity at the warning

signal and the distance between the warning signal and the entry signal the ratio $\dfrac{x_1}{x_2} = 3$ is preserved. The pros of the method are: easy mathematical interpretation, constant $\dfrac{x_1}{x_2}$, the movement is with constant deceleration.

In real-life conditions (the movement is parabolic) far greater error occurs. The probability for the aspect of the entry signal to change when the train is between the warning signal and the **infill balise and the** probability for the aspect of the entry signal to change is constant $\dfrac{t_1}{t} = \dfrac{t_2}{t} = \dfrac{1}{2}$.

The second method is based on the same theoretical postulates as the one above, but it uses parabolic interpretation in calculating. The mathematical justification is very complicated and is very seldom used in real calculations.

The third method (Fig.8) is based on the assumption that the most optimal position of the infill balise is at $t_1 = t_2$, i.e., the time for which the train traverses curve $v_1 - v_2$ equals the time for which it traverses curve $v_2 - v_3$. The method is comparatively new. It has a complex mathematical interpretation.
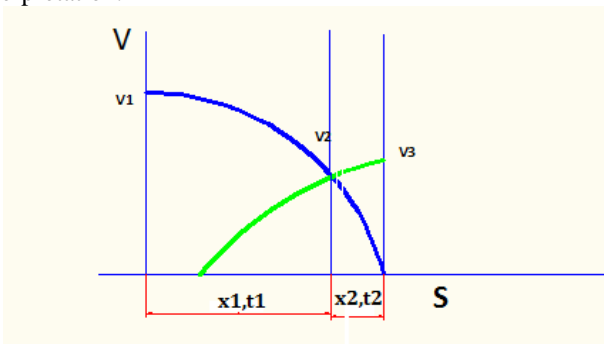

Fig.8. Real probability calculation

A **graphical – analytical** method that is very flexible has been elaborated and proposed. In this method movement is approximated as a parabola (variable acceleration). There are four variable parametres in place for determining the input parametres and an abundant "toolbox" for calculation of various output parametres which are tremendously useful to the designer in considering a specific situation. The relevant software is developed, with the help of which every change in the input parametres is visualized immediately and enables the designer to assess the optimal position of the infill balise in compliance with the requirements of the terms of reference and the allowed tolerances, both for input and for output data. This method also provides a possibility for analysis of the already existing situation and its optimization with minimum changes and expenses. As is widely known, the economic factor becomes increasingly significant in contemporary developments, which is why special attention should be given to reducing the consumption of time and power and to increasing the track carrying capacity of a particular section,

and also to reducing one-off investments in facilities on account of design optimization.

Graphical interface is developed for visualization of the method. Upon correct selection of the user data a screen appears, similar to that in Fig. 9.
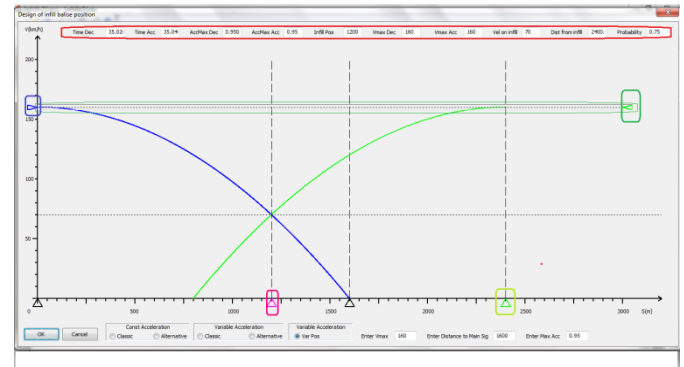

Fig.9. Flexibility calculation of infill balise position

On the screen are displayed the braking curve and the acceleration curve. The position of infill balise **Δ** is calculated via the classical parabola method. The parabolas are calculated and drawn according to the initial input data. In the upper part of the screen (red) fields with some output data are shown.

V. CONCLUSION

As a result of the performed analyses for description of the movement of trains between the warning and entry colour light signal a selection is made of the most appropriate interpretation of this movement. Mathematical and software model is developed and a method is proposed for determining the location of the infill balise, depending on the position of the infill balise **Δ**, the position at the end of the acceleration **Δ**, the entry velocity at the warning signal ▶ $(0 \div 210)$, the maximum acceleration velocity ◀ $(0 \div 210)$. With the created graphic interface the input data may be smoothly changed, as they are calculated forthwith and appear in the output data fields.

REFERENCES

[1] Vishy Shanker, The Ideal Positioning of Infill Balises, IRSE News, ISSUE 172, Nov.2011
[2] David Glendinning, Graeme Stainlay, ETCS Revealed – The Railcorp Expereince, IRSE Technical Meeting "All about ATP", Sydney, March, 2008.
[3] Tractive effort, acceleration and braking, The Mathematical Association, 2004

Experimental Investigation of Grid-connected Induction Generator's Behavior during Reconnection Transients

Milica Rašić[1], Milan Radić[2], Nenad Floranović[3], and Zoran Stajić[4]

Abstract – **This paper presents experimental results obtained by recording transient regimes of three-phase grid-connected induction generator. Transients that were considered in these experiments are induction generator's disconnection from the electrical grid and its reconnection. The induction generator, with the rated power of 2.2 kW was investigated by two experiments with the similar procedure, during which it was in two different steady operating points, i.e. generator delivered two different values of active electrical power to the grid. In both experiments Y-connected capacitor bank with the capacitance $C = 30\,\mu F$ was directly connected in parallel to the generator with the goal to reduce reactive power, taken from the electrical grid. In both situations, generator's disconnection from the grid was followed by its self-excitation, as a consequence of the mentioned capacitor bank's existing. In both experiments this state lasted a few seconds and thereafter the generator was reconnected to the grid. Obtained results showed generator's electrical parameters behavior during described transients.**

Keywords – **Experiments, Induction generator, Transient regimes, Disconnection, Reconnection**

I. INTRODUCTION

In the past, during the long history of their usage, induction machines were operating most often as electrical motors. Energy crises and introduction of renewable sources have caused more frequent use of induction machines as generators, especially for utilization of water's and wind's energy potential [1-2]. Generator's mode of induction machine occurs when the rotor is spinning by another driving machine in the direction of the magnetic field's rotation at the speed higher than synchronous [3]. Two types of induction generators are available: "grid-connected" and "self-excited". Grid-connected induction generators are connected to the electrical energy system (EES) producing active electrical energy and delivering it to the EES. At the same time, grid-connected induction generators take from the EES a certain amount of reactive energy that is used for their excitation [4-5]. Self-

excited induction generators are usually used at some specific locations with the energy potential, where generator's connection to the EES is not possible. At these locations their autonomous operation is possible under special conditions [6] (cheap solution for voltage and frequency regulation still does not exist and in this manner generated active energy should be used just for thermal consumers).

In this paper grid-connected induction generator was considered. Since it is connected to the electrical energy system, this type of generator operates at the voltage and frequency of the EES [7]. There are a lot of situations when capacitor batteries are connected in parallel to the generator in order to compensate a certain amount of reactive electrical energy taken from the electrical grid. As the demanded, by the generator, reactive power varies continuously and the capacitor batteries may be changed stepwise in order to avoid unwanted higher excitation of the generator, resulting in high voltage [4].

One of the most interesting aspects for science research in the area of induction generators is certainly its behaviour during different types of transient regimes. The main objective of the work presented in this paper is to investigate changes of induction generator's electrical parameters during its disconnection and reconnection transients. Electrical parameters that were analyzed, based on the presented experiments, are voltages, currents and frequency.

II. THEORETICAL BACKGROUND

Generally, transient regimes could be very dangerous for both electrical machine's parts – electrical and mechanical. Transients that are the subject of this paper happen during generator's disconnection from the electrical supply and thereafter during its reconnection. Grid-connected induction generator operates at the voltage and frequency of electrical system. It delivers a certain amount of active electrical power to the grid, but it takes reactive power from the electrical system which is used for its excitation. Capacitor bank is connected in parallel with the generator and reactive power consumption is reduced. Generator could operate in different steady state points, depending of its load. In the moment of generator's disconnection from the system, it continues with its operation, but his excitation is now caused by capacitor bank. Really dangerous situation could happen, if the capacitor bank is with the higher capacitance than is recommended. Higher value of capacitance causes higher excitation's current that leads to a really high value of machine stator's voltage. Meanwhile, with the increasing of generator's voltage, frequency is increasing, as well. Rotor's speed is in dependence of the frequency and its value is also higher. Controlling of rotor's speed is here really important,

[1]Milica Rašić is with the Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia
E-mail: milicarasa@gmail.com
[2]Milan Radić is with the Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia
E-mail: milan.radic@elfak.ni.ac.rs
[3]Nenad Floranović is with the Research and Development Center "ALFATEC" Ltd., Bul. Nikole Tesle 63/5, 18000 Niš, Serbia
E-mail: nenad.floranovic@alfatec.rs
[4]Zoran Stajić is with the Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia
E-mail: zoran.stajic@elfak.ni.ac.rs

because if the rotor's speed continues with increasing, mechanical equipment failures can be caused. Values of stator's currents in the moment of disconnection are lower, but with the higher voltage caused by capacitor's excitation, these values increase again. This mode of generator's operation could last a few second and it could lead to stationary state operation, if the machine could hold voltage and rotor's speed increase. All things considered, the induction generator during this period is operating as "self-excited".

Other group of transients considered in this paper are induction generator's reconnection transients. In the moment of generator's reconnection, voltage and frequency decrease to the electrical grid's values. The induction generator is "grid-connected" again. Problems that can appear in the moment of generator's reconnection can be caused by the higher values of stator's currents. These transient currents could be several times higher than the stator's currents rated values. Higher values of stator's currents are followed by higher values of Joule's heating losses that can cause stator's windings failures.

Generally, it is really important to be very careful about described transients in order to avoid generator's electrical and mechanical damages. One of factors that can have a great impact to the generator's behaviour during described transients is the capacitance of capacitor bank, connected to the electrical grid. This impact has already been explained – high values of the capacitance can cause higher excitation and stator's voltage. Another factor that can also have a great impact is the value of active power that generator delivers to the electrical grid .Experiments described in the next chapter were performed with two different values of active power and results obtained by them explain theirs impact to the generator's behaviour during described transients.

III. DESCRIPTION OF EXPERIMENTS

Experiments whose results are presented in the paper were performed using equipment connected as shown in Fig. 1.
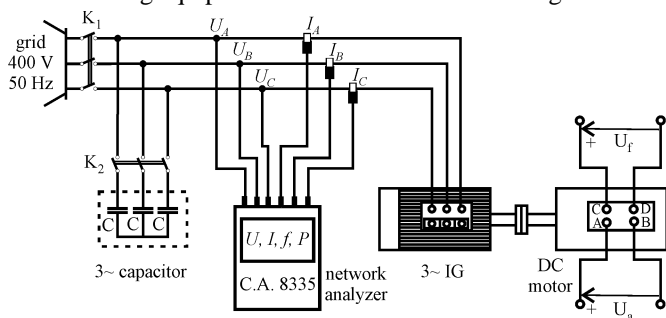

Fig. 1. Connection diagram and equipment used for experiments

An induction machine with cage rotor was used as a grid-connected induction generator. The machine's nameplate data (given for motoring mode of operation) are: $U_n = 380V$,

$f_n = 50 Hz$, $I_n = 4.9 A$, $P_n = 2.2 kW$, $\cos\varphi_n = 0.86$, $n_n = 2825 \min^{-1}$, Y-connected stator.

Generator's rotor was driven by separately excited DC motor, whose excitation current I_f was set to the rated value, while the mechanical power delivered to the generator was controlled by the applied armature voltage U_a.

At the beginning of the first experiment, both contactors K_1 and K_2 were opened, while armature voltage of the used DC motor has been set to the value that resulted with almost synchronous rotation of the system ($n_0 \approx 3000 \min^{-1}$). After that, contactor K_1 was closed and induction generator was connected to the electric grid, operating at no-load regime and consuming reactive power only. By closing the contactor K_2, symmetrical, Y-connected capacitor bank with capacitance $C = 30 \mu F$ per phase was directly connected in parallel to the generator, in order to mitigate consumption of reactive power. Note that the contactor K_2 remained closed during the rest part of the experiment. By increasing the value of armature voltage U_a, mechanical power on the generator's shaft was increased, and the generator reached steady operating point characterized by active power $P_{G1} = 1430W$ delivered to the grid. After that, contactor K_1 was opened, and closed again after a short period of time. This action had caused transient behavior whose nature has been the subject of this investigation. Relevant electrical parameters were registered using portable network analyzer C.A. 8335b, with capability of recording complete voltage and current waveforms.

The second experiment had the same procedure, with only difference that before opening and reclosing of contactor K_1, induction generator was in steady operating point characterized by greater value of delivered active power, $P_{G2} = 1720W$.

IV. RESULTS AND DISCUSSION

Diagrams of relevant electrical parameters recorded by portable network analyzer during described experiments are presented in following figures.

In Fig 2. and Fig. 3 changes of stator's line voltage U_{AB} are shown for the first and second experiments, respectively. From Fig. 2. it can be noticed that before generator's disconnection, generator's stator voltage U_{AB} was equal to the grid's voltage. From the moment of its disconnect, voltage started to grow and after about 1.7 seconds it achieved the peak value $U_{AB}=720$ V, which is a direct consequence of generator's self-excitation, caused by parallel connected capacitor bank $C = 30 \mu F$. Voltage increase settled approximately this value and a new steady state was reached. Generator continued to operate with this value of U_{AB} for 0.8 seconds more and it was reconnected to the grid. In the moment of generator's reconnection, voltage U_{AB} was again equal to the grid's voltage. Generally, during the first experiment induction generator was delivering $P_{G1}=1430$ W to the grid and while it was working as self-excited for about 2.5

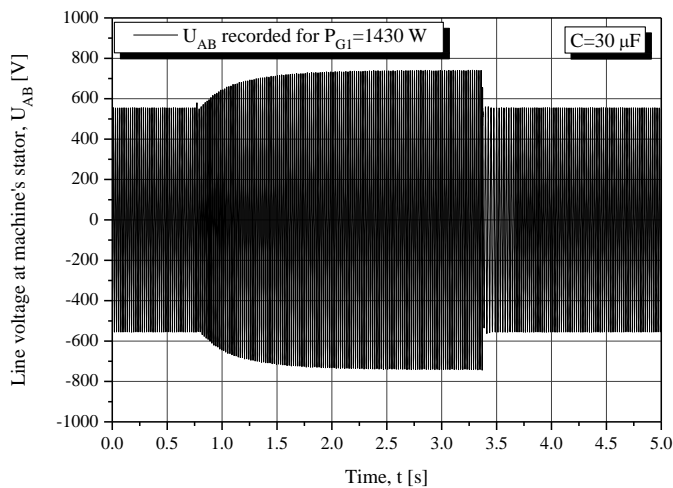seconds, it achieved peak value of stator's line voltage $U_{AB}=720$ V and reached new steady state of its operation.

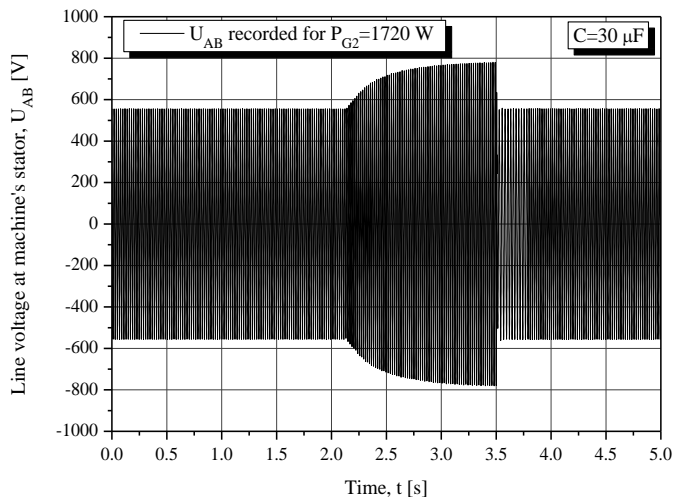Fig. 2. Diagram of line voltage U_{AB} at generator's stator for $P_{G1}=1430$ W

Fig. 3. Diagram of line voltage U_{AB} at generator's stator for $P_{G2}=1720$ W

In Fig. 3. diagram of stator's line voltage U_{AB} that induction generator had during the second experiment is presented and differences between the U_{AB} during the first experiment can be observed. In the beginning, the situation is the same − line voltage at generator's stator is equal to the grid's voltage. The generator was in steady operation point characterized by almost 30 % greater value of delivered active power than in the first experiment, $P_{G2}=1720W$. In the moment of its disconnection, voltage started to increase and after a while it achieved the peak value $U_{AB}=790$ V, which is 10 % higher value than in the first experiment. The difference between these two experiments is also in the lasting of generator's operation as self-excited and for the second experiment in this mode generator was working for about 1.4 seconds and then it was reconnected, when stator's line voltage U_{AB} was again equal to the grid's voltage. In the second experiment induction generator didn't reach new steady state as self-excited,

because it was dangerous for it to operate longer with this value of voltage taking into the account the fact that the frequency and rotor's speed were also higher.

In Fig 4. comparative diagrams of machine's frequencies are shown for two described experiments. In both situations before the disconnection frequencies at induction generator's stator were equal to the electrical grid frequency ($f=50Hz$). For both experiments in the moment of generator's disconnection frequency started to increase and in the first experiment it achieved the value of $f=64.5Hz$. In the second experiment frequency reached the value $f=67Hz$, with the possibility of continuation of its increasing in the case of longer period of generator's operation as self-excited. Approximately, steady state of generator's operation could be reach for the frequency of $f=68Hz$. Unfortunately, it wasn't safe for generator's equipment to achieve this steady state, because the frequency's rise was following with the rotor's speed increasing and rotor's overspeed could lead to mechanical damage. In general, for less than 30 % higher value of active power delivered to the grid during the disconnection transients, frequency can reach 5 % higher value. In both experiments after generator's reconnection induction generator stator's frequency again reached the value of grid's frequency.
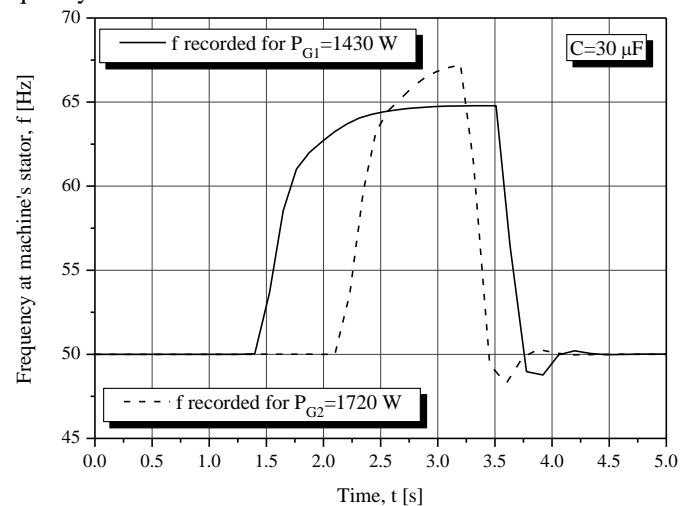

Fig. 4. Comparative diagrams of frequencies at generator's stator for $P_{G1}=1430$ W and $P_{G2}=1720$ W

Finally, in Fig. 5 and Fig. 6. changes of stator's phase current I_C are shown for the first and second experiments, respectively. This phase current is chosen, because it was useful to show what happened with the phase, which wasn't considered by line voltage U_{AB}. For both experiments in the moment of generator's disconnection the value of stator's phase current started to decrease, but after a while with the increase of stator's voltages it stopped with decreasing and started to rise. During the generator's self-excited operation, stator's phase currents didn't reach values that they had before the disconnection.

The situation that is more interesting, when currents are considered, is in the moment of generator's reconnection. In

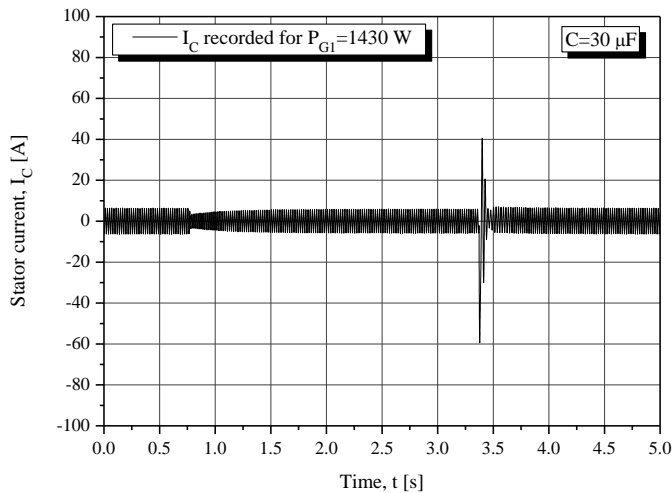that moment transient currents at generator's stator can be few times higher than their rated values.

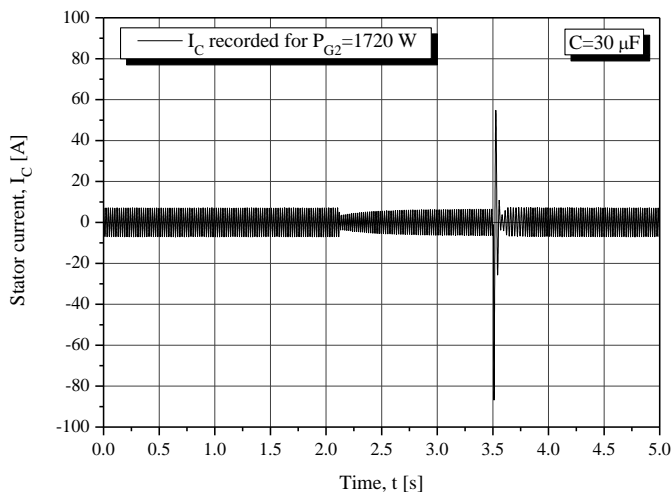Fig. 5. Diagram of phase current I_C at generator's stator for $P_{G1}=1430W$

Fig. 6. Diagram of phase current I_C at generator's stator for $P_{G2}=1720W$

In Fig. 5. it can be noticed that the signal of phase current I_C in the moment of generator's reconnection reached *40 A* in positive part of diagram and *-60 A* in its negative part. The current's signal value in negative part is higher because of the existing of aperiodic component of transient current.

In Fig. 6. changes of phase currents I_C during the second experiment are presented. The situation is similar to the first experiment with the different values that current's signal reached in the moment of generator's reconnection: *55 A* in positive and *-85 A* in the negative part of the diagram. Therefore, if the positive part of current's signal is considered (influence of aperiodic component is not taken into the account) it can be seen that the peak current I_C is almost eight times higher than the peak value of generator's rated current. In general, for less than 30 % higher value of delivered active

power, peak current during generator's reconnection transient regime can reach almost 40% higher values. Those high currents can be really dangerous, because they can produce high values of Joule's heating losses that can cause generator's windings damage.

V. CONCLUSION

Grid-connected generator's behaviour during its disconnection and reconnection transients for two different values of active power delivered to the electrical system was presented in this paper. It was shown that increasing of less than 30 % of active power that induction generator delivers to the grid can cause significant differences in diagrams of specified electrical parameters presented in previous chapter. Changes of considered electrical parameters during generator's operation as self-excited were shown and described.

Disconnection and reconnection transients were considered for only two values of electrical power that generator delivers to the grid. Future investigation in this area can include induction generator's behaviour recorded for a larger range of those values. Differences between changes of electrical parameters for that research will also be analyzed.

ACKNOWLEDGEMENT

REFERENCES

[1] M.Hadžiselimović, I. Zagradišnik, B. Štrumberger, "Induction machine: comparison of motor and generator characteristics", *Przeglad Elektrotedhniczny,* 2013, Vol. 89, pp. 103-106.

[2] N. Smith: „Motors as generators for micro – hydro power", Intermediate Technology Publications, London, UK, 1994.

[3] Z. Stajić, Đ. Vukić, M. Radić, *Asinhrone mašine,* Elektronski fakultet, Niš, 2012. (in Serbian)

[4] D. A. Gorski, J. Wisniewski, W. Koczara, "Grid connection transisents of small hydropower generator", *Proceedings of ICREPQ",* Santiago de Compostela, 2012.

[5] S. Sharma, K. S. Sandhu, "Effect of machine parameters on starting transients of grid connected induction generators used in wind farms", *International Journal of Computer Applications*, July 2011, Vol.25 No.11, pp: 34-37.

[6] C.F. Wagner, "Self-Excitation of Induction Motors with Series Capacitors" *AIEE Transactions*, February 1941, Vol. 60, pp. 1241-1247.

[7] K. S. Sandhu, S. Sharma, "Effects on major power quality issues due to incoming induction generators in power systems", *ARPN Journal of Engineering and Applied Sciences*, February 2010, Vol. 5. No. 2, pp: 57-65.

An Approach on Applying the CSM for Risk Evaluation and Assessment of Significant Changes of the Railway System

Denitsa Kireva-Mihova [1]

Abstract – **The Railway Safety Directive (2004/49/EC) requires implementation of Common Safety Methods (CSMs) to ensure that a high level of safety in railways is maintained and improved. The absence of a common approach for specifying and demonstrating compliance with safety levels and requirements of the railway system among the Member States has proved to be one of the obstacles to liberalisation of the railway market.**

The objective of this paper is to propose a common approach on applying the CSM in its part of hazards classification and risk evaluation of significant changes of the railway system.

Keywords – **Railway System, Common Safety Method (CSM), Risk Assessment (RA)**

I. INTRODUCTION

In the common European legislation, the Railway Safety Directive (2004/49/EC) requires the railway companies operating in the EU Member States to manage risks systematically. They are doing this, following amongst others the requirements described in the "Regulation (EC) No.352/2009 of 24 April 2009 on the adoption of a common safety method on risk evaluation and assessment" and repealed by the "Commission Implementing Regulation (EU) No.402/2013 of 30 April 2013 on the common safety method for risk evaluation and assessment" *(CSM RA)*. The CSM RA describes a risk management process and framework. The demonstration of compliance with this process, which is accompanied by a report of an assessment body, is a prerequisite for mutual recognition of the respective process application, [1-3].

The CSM RA has applied to all significant changes to the railway system (modernisation or renewal) since 01 July 2012, [1]. The changes may be of a technical (engineering), operational or organisational nature, where the organisational changes could have an impact on the operation of the railway. The CSM RA also applies if a risk assessment is required by a technical specification for interoperability (TSI) and is used to ensure safe integration of a structural subsystem into an existing railway system in the context of an authorisation for placing in service in accordance with the Railway Interoperability Directive 2008/57/EC, [4]. The aim of interoperability legislation is to achieve a technical and operational harmonisation of the main structural and functional railway subsystems. The structural subsystems are:

- infrastructure
- energy

- trackside control-command and signalling
- on-board control-command and signalling
- rolling stock

The functional subsystems are:

- operation and traffic management
- maintenance
- telematics applications for passenger and freight services.

The CSM RA applies to significant changes to all railway subsystems.

For the purpose of this paper the following definitions will be used [1-3]:

railway system means the totality of the subsystems for structural and operational areas, as defined in Directive 2008/57/EC, as well as the management and operation of the system as a whole;

system means any part of the railway system which is subjected to a change whereby the change may be of a technical, operational or organisational nature;

risk means the frequency of occurrence of accidents and incidents resulting in harm (caused by a hazard) and the degree of severity of that harm;

risk assessment means the overall process comprising a risk analysis and a risk evaluation;

safety means freedom from unacceptable risk of harm;

fatality means death within one year of the casual incident;

proposer means the railway undertakings (RUs) or the infrastructure managers (IMs) in the framework of the risk control measures they have to implement in accordance with Directive 2004/49/EC.

consequences means the number of fatalities, major injuries and minor injuries resulting from the occurrence of a particular hazardous event outcome;

frequency of an event is the number of times an event occurs over a specified period of time e.g. number of events/year.

II. PROBLEM DEFINITION

The CSM RA Regulation applies to the proposer (RU or IM) when making any change to the railway system in a Member State and applies only to *significant changes* of the railway system. If there is no notified national rule for defining whether a change is *significant* or not in a Member State, the proposer considers the potential impact of the change in question on the safety of the railway system using the significance criteria in the CSM RA Regulation [2]. In case the proposed change has no impact on safety, the risk management process need not be applied.

[1]Denitsa Kireva-Mihova is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: kireva@tu-sofia.bg

Whenever the railway system already in use is subject to a change, the significance of the change is to be assessed taking into account all safety-related changes affecting the same part of the system. The purpose is to assess whether or not the totality of such changes amounts to a significant change requiring the full application of the CSM RA.

The CSM RA management process comprises the following activities [3]:

➢ The proposer of a change produces a preliminary definition of that change, and the system to which it relates. It then examines it against the significance criteria in the CSM RA Regulation.

➢ The CSM risk management process starts with the system definition. This provides the key details of the system that is being changed (purpose, functions, interfaces and the existing safety measures that apply to it).

➢ All reasonably foreseeable hazards are identified and their risk is classified and analysed.

➢ Safety requirements are identified by application of one or more of the three risk acceptance principles to each hazard.

➢ A hazard record for the system that is to be changed is produced and maintained.

➢ Before acceptance, the change proposer demonstrates that the risk assessment principles have been correctly applied and that the system complies with all specified safety requirements.

➢ The assessment body provides its report to the proposer.

This risk management process is iterative and is depicted in the diagram on Fig. 1, [2]. The process ends when compliance of the system with all the safety requirements necessary to accept the risks linked to the identified hazards is demonstrated.

Hazard classification has a very particular meaning in the context of the risk management framework. It is based on an initial assessment of the risk associated with each hazard and is carried out as part of a hazard identification process.

It is very important that, at the considered level of detail, the hazard identification is complete and that hazards are neither forgotten nor wrongly classified to be associated with *broadly acceptable risk(s)*.

This is of prime importance because if hazards are not identified, they are not mitigated and are not dealt with further in the risk management, risk assessment and hazard management processes.

The classification of the identified hazards, at least into hazards associated with *broadly acceptable risk(s)* and hazards associated with risks that are not considered as broadly acceptable, enables the prioritisation of the risk assessment on those hazards that require risk management and risk control measures. The classification of hazards between these two categories is based on *expert's judgement* and does not have a quantitative measurement. Criteria are to be used to help decide whether the risk associated with a proposed change is low enough to proceed (Fig. 2).

It is the responsibility of the change proposer to evaluate whether the risk associated with each identified hazard is broadly acceptable.
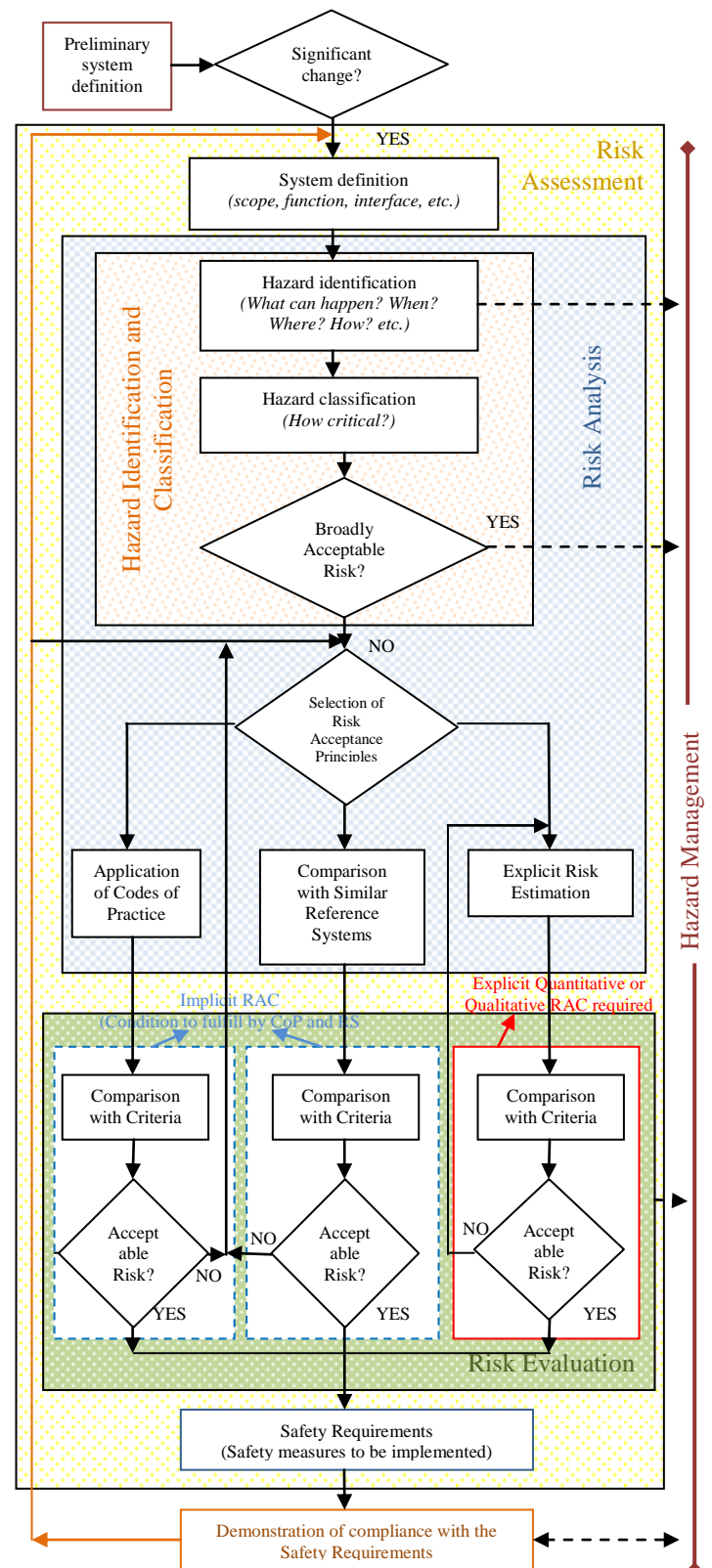

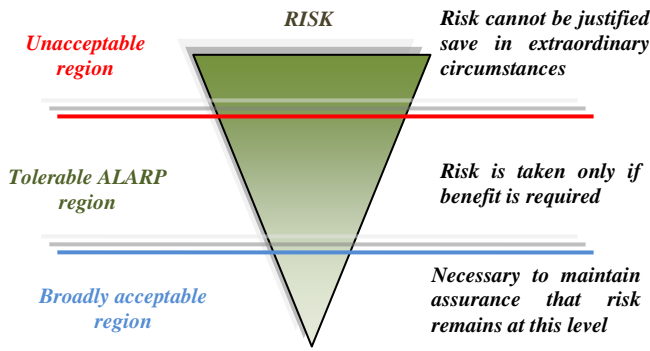Fig. 1. Risk management framework in the CSM RA Regulation

Fig. 2. Framework for tolerability of risk

The CSM RA enables the evaluation of the risk acceptability of a significant change to the railway system by using one or a combination of the following so-called *risk acceptance principles*, without giving priority to any of them:

- the application of *codes of practice (CoP)*
- the comparison with similar *reference systems (RS)*
- the use of *explicit risk estimation (ERE)*

The proposer of the change is responsible for the choice of the principle to apply, [2]. The risks, which are controlled by the application of *codes of practice* or by the safety requirements derived by a comparison with a similar *reference system*, are considered as acceptable provided that the conditions of application of these two risk acceptance principles are fulfilled and sufficiently documented as defined in the CSM RA. This means that explicit risk acceptance criteria need not be defined for the hazards controlled by these two principles. Additionally, whenever the third risk acceptance principle - the *explicit risk estimation* – is used and in order to be able to determine whether the residual risk is sufficiently low so that it is not necessary to take any immediate action to reduce it further, *risk acceptance criteria (RAC)* are used. *Explicit risk acceptance criteria* will therefore only be needed for evaluating the risk acceptability when applying the third principle - *explicit risk estimation* (Fig.3).
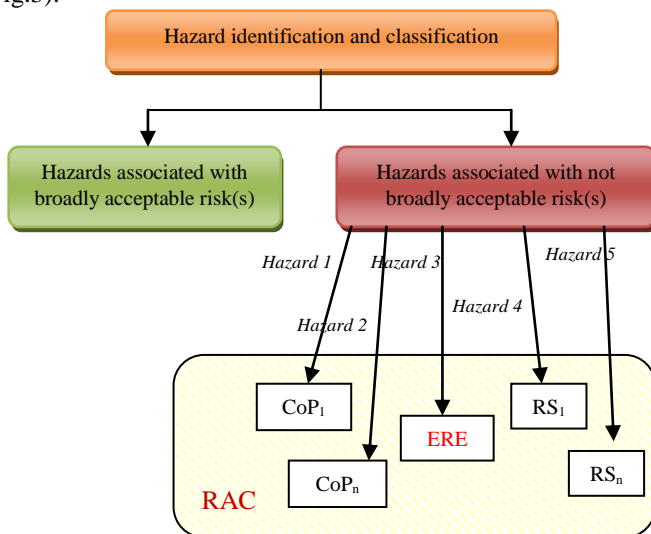

Fig. 3. Applying the risk acceptance principles

The *explicit risk estimation* is not necessarily always quantitative. The estimation of risks can be quantitative (if sufficient quantitative information is available in terms of frequency of their occurrence and severity), semi-quantitative (if such quantitative information is not sufficiently available) or even qualitative (e.g. in terms of process for management of systematic errors or failures, when quantification is not possible).

The *explicit risk acceptance criteria* that are needed to support the mutual recognition will be harmonised between the Member States by the on-going European Railway Agency work on the risk acceptance criteria. At this moment it is the responsibility of the proposer of the change to define such criteria and more frequently the inaccurate qualitative approach based on *experts' judgment* is used.

In cases the CoP and RS risk acceptance principles are not applicable an approach to *explicit quantitative risk estimation* is proposed in order to avoid the subjective disadvantages derived from the *expert's judgement*. Applying the quantitative approach the assumption is made that sufficient quantitative information for the hazards frequency and the consequences from the change of the railway system is available.

III. EXPLICIT RISK ESTIMATION AND EVALUATION

The *explicit risk estimation* principle is frequently used for complex or innovative changes of the railway system (modernisation or renewal). In order to evaluate whether the risks controlled by the application of *explicit risk estimation* is acceptable or not, *explicit risk acceptance criteria* are needed. These can be defined at different levels of a railway system. They can be seen as a *pyramid of criteria* (Fig. 4) starting from the high level risk acceptance criteria (expressed for instance as societal or individual risk), going down to subsystems and components (to cover technical systems) and including the human operators during operation and maintenance activities of the system and subsystems, [3].
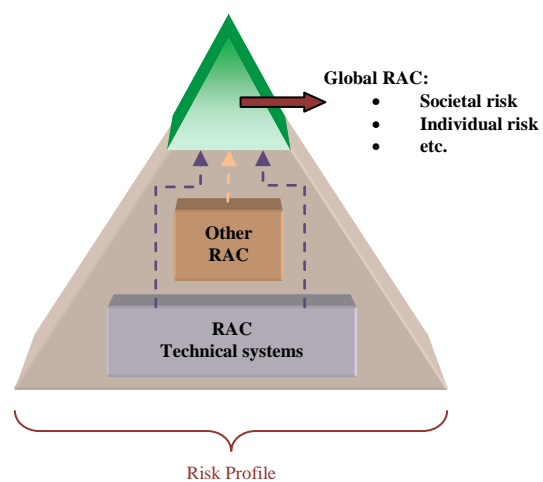

Fig. 4. Pyramid of risk acceptance criteria

Risks to people may be expressed in two main forms: *individual risk* – the risk experienced by an individual person,

and *societal risk* – the risk experienced by the whole group of people exposed to the hazard.

Low-frequency high-consequence events might represent a very small risk to an individual, but they may be seen as unacceptable when a large number of people are exposed. Railway incidents are mainly referred to such type of events which are unacceptable for a large number of people, so the *societal risk* is considered in the proposed approach.

Societal risk can be represented:

- graphically, in the form of FN-curves, or
- numerically, in the form of a risk integral.

FN-curves: Generally, risk scenarios (S_i) to be included in a risk assessment can be characterised as having a frequency (f_i) and a consequence (c_i), i.e. number of casualties (N). F is used to denote the sum of the frequencies of all the individual events that could lead to N or more fatalities, [5].

The risk is then defined by a set $\{S_i, f_i, c_i\}, i=1...n$.

Associated with each risk scenario, the risk may be defined as:

$$r_i = f_i \times c_i \qquad (1)$$

and the risk of the changed railway system defined as:

$$r = \sum_{i=1}^{n} r_i \qquad (2)$$

The risk as defined in Eq. 2 offers a measure of the risk level over all the risk scenarios. It however hides the difference between two types of incidents: one of low frequency but high impact consequences; the other one of high frequency but low impact consequences. In fact, people have different attitudes toward these two types of incidents. It is therefore of interest to include the frequency and the severity profile in the risk assessment and can be represented by the FN-curve, where N stands for the fatalities in one incident and F stands for the yearly frequency of the incidents causing N or more fatalities (Fig. 5).
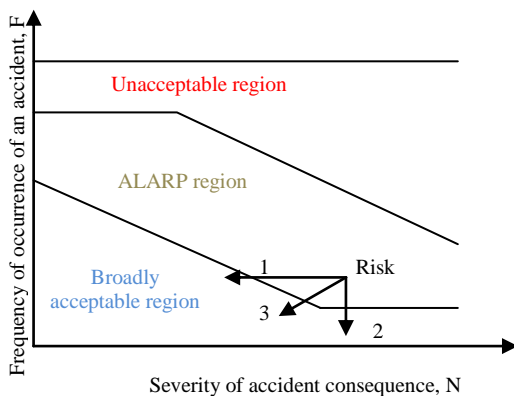

Fig. 5. Risk areas and risk reducing measures

The FN-curve can be built based on the frequencies and the consequences of the risk scenarios S_i, $i=1...n$.

First the risk scenarios are ordered to satisfy $c_{i-1} < c_i$. then plot the cumulative frequency F_i against c_i for $i=1...n$.

$$F_i = \sum_{k=1}^{n} f_k \qquad (3)$$

The *risk acceptance criteria* for societal risk are to be set up on the FN-diagram, which could be defined by the customer or authority or estimated. Two criteria lines divide the space into three regions – where risk is unacceptable, where it is broadly acceptable and where it requires further assessment and risk reduction as low as reasonably practicable (ALARP), as shown on Fig. 5. There are three options for reducing the risk: mitigate the consequences of an accident (Arrow 1), reduce the probability of occurrence of the accident by implementing additional safety barriers or by using more reliable components (Arrow 2) or both (Arrow 3).

Risk integral: A risk integral (RI) is a summary of the overall level of societal risk, taking account of the whole set of f-N pairs and is mathematically presented with Eq.4.

$$RI = \sum fN^{a} \qquad (4)$$

The calculation of a contribution to the RI can weight the value for N so as N increases, the RI contribution also increases but at a faster rate. Proposed values of α are generally in the range 1 to 2. So, the consequences of events are considered to be more important than the likelihoods.

An RI value is a summary statistic and so cannot, by itself, tell us anything about the various specific accident scenarios or their likelihood at any site. Moving from FN representation to RI sacrifices detailed information (f-N pairs) for ease of comparison and ranking - RI values are more easily and unambiguously ordered than FN-curves.

IV. CONCLUSION

CSM RA gives the responsibility to the proposer of the change to evaluate whether the risk associated with each identified hazard is broadly acceptable. If the identified hazards and the associated risks of the system under assessment cannot be controlled by the application of CoP or similar RS, an approach to *explicit quantitative risk estimation* is proposed in order to avoid the subjective disadvantages derived from the *expert's judgement*. With the criteria on the FN-curves we can judge whether more actions are needed to improve the safety of a given system subject to a change. It could be used by RUs and IMs to develop their safety management system (SMS) to manage the risks associated with their activities.

REFERENCES

[1] *Railway Safety Directive 2004/49/EC*
[2] *Commission Implementing Regulation (EU) No 402/2013 of 30 April 2013 on the common safety method for risk evaluation and assessment and repealing Regulation (EC) No 352/2009.*
[3] ERA/GUI/01-2008/SAF, *Guide for the application of the Commission Regulation on the adoption of a common safety method on risk evaluation and assessment as referred to in Article 6(3)(a) of the Railway Safety Directive,* ver. 1.1.,2009
[4] *Railway Interoperability Directive 2008/57/EC*
[5] Daosheng Cheng,*Uncertainty analysis of Large Risk Assessment Models with Application to the Railway Safety and Standards Board safety Risk Model,* 2009

POSTER SESSION
DIGITAL IMAGE PROCESSING

A Method for Colorization of the Original Grayscale Tesla Photograph

Vladan Vuckovic [1] and Sanja Spasic [2]

Abstract – **The method for picture colorization of one the Nikola Tesla's original photograph will be explained in this paper. The aim of this paper is to enliven patents which may contribute to enlarge information related to scientific knowledge. This project presents a simple method which requires precise picture colorization and its possible animation. The colorization method is based on a simple assumption of opposite pixels in space and good understanding of the period when the picture is made.**

Keywords – **Image processing, grayscale image colorization,**

I. INTRODUCTION

There is a great need for the enlivening of the picture as well as for the restoration of the picture until their movement is archived. There is more often need for the use of pictures from the distant past which are in relation to the key words for revealing details on which certain sciences are based: medicine, information technology, electrotechnics etc. One of the ways of enlivening pictures is picture colorization. Picture colorization is one of the key answers to a wide spectre of digital processing of monochrome pictures. In that way picture is enlivened and it shows the past. These are some of the reasons for appearing of big interest for the further research, improvement and the use of method for picture colorization. Picture colorization means it is necessary to animate picture when the illusion of drawings movement is achieved, by the models or lifeless things. For an animation of pictures different methods are used, the most included method is the presentation of animation of moving pictures or video animation. Animation makes the pictures become very real and there is an impression of their current, real existence.

II. PICTURE COLORIZATION

Colorization is the method which is used for adding colours to black and white picture. In order to do the successful colorization method, it is essential to prepare basic picture stem. Base improvement is done by the method of picture restauration whose main aim is to refresh some parts of the picture, to remove rustle which very often appears because of the processing. Generally speaking, the

process of picture improvement does not enriched the informative content of the picture nor is the originality of the pictures lost but it only makes easier the use of existing information in a particular picture. In order the present picture colorization in the best way we will take as an example the picture of Nikola Tesla, which is very specific and unique.

Fig 1: Nikola Tesla with his invention - original image (Nikola Tesla Museum of Belgrade).

Considering the fact that Nikola Tesla was in some way perfectionist during his life, it is a big challenge to do the colorization of his picture because it means to use the knowledge about Tesla's life. It was needed to know Tesla's habits and details of his life which means to know the life of Nikola Tesla and the start the work.

The picture, on which we will be working, is the original, which is kept from the distant past in the museum of Nikola Tesla in Belgrade (Fig. 1). Considering the fact that pictures used in the past can not be compared to the present ones, it makes this challenge greater to do the colorization of reality which was done in Grayscale technique.

As it can be seen Grayscale picture contains a great number of grey tones. In order to make the correct and exact picture colorization it is needed to define all pixels in the picture. The Gray picture is the one where every pixel represents a particular sample, which has only the information about intensity [1].

[1]Vladan Vuckovic is with the Faculty of Electronic Engineering, ul. Aleksandra Medvedeva 14, Nis, Serbia, E-mail: vladan.vuckovic@elfak.ni.ac.rs.

[2]Sanja Spasic is with the Faculty of Electronic Engineering, ul. Aleksandra Medvedeva 14, Nis, Serbia, E-mail: sanjadjor@gmail.com

This kind of picture is made of grey tones and they can vary from the one which has the weakest intensity (black) to the one with the strongest intensity (white). The intensity of pixels of grey picture is in the range from 0 to 255, where 0 is black and 255 is white, or from 0 to 1, where 0 represents black and 1 represents white, other numerical values represent different levels of grey (Fig. 2).

Fig. 2. Grey image in the range of 0 to 255.

Every pixel of grey can be expressed by formula (1):

$$W_{ij} = e \bullet \frac{\left\| F_i - F_j \right\|^2}{\sigma_I^2} \bullet \frac{\left\| X_i - X_j \right\|^2}{\sigma_X^2} \qquad (1)$$

Where is:

F_i – vector describing the characteristics most frequently pixel color

X_i - node pixels

Where the following condition (2) must be true:

$$\left\| X_i - X_j \right\| < r \qquad (2)$$

Where "i" represents special location of the pixel; While "F_i" vector describes characteristics of a pixel which are based on information related to intensity, color and texture and "r" is distances of pixels.

If this distance approaches 0 the more shades of white pixels harder. Since all the pixels are defined, it is needed to copy the color from the original patent in the most realistic way [7].

Now, the functions F_i is defined in a different way. It is based on the color used, expressed by the formula:

$$F_i = \left[v, v \cdot s \cdot \sin(h), v \cdot s \cdot \cos(h) \right]_i \qquad (3)$$

Where is:

v, s, h are the values of the color spectrum that is used for image segmentation

Important fact is that someone must subtracts function F_i and F_j between two pixels deducted all components then take into account all three components of value, which will be shown in Figure 4.
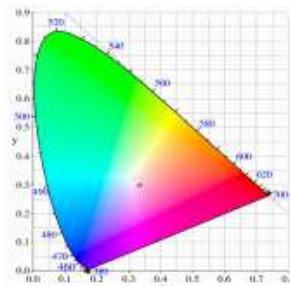

Fig 3: The color spectrum

Color [5] is related to a particular light feeling of physical quality of light, the movements of which can be noticed by a human eye (Fig. 3). Rays of light are electromagnetic movement of precisely defined wave lengths.

At different moments the human eye registers colours in different ways. The same goes for technics (cameras) which makes artificial light in order to save the color.

We can say that colours reflect in a different way from different points. In that way expression of colours in a picture can reduce the quality of picture. It is very important to give adequate colour to a definite grey.

By combination of grey tones and colorization we can make very effective and pleasant picture to our eye. In the process of colorization, after tidying of the basic picture, the separation of the edges and putting the colors are performed.

The edges represent local discontinuities of lightning (or colors) in the picture and they give a good indication of the borders of the objects in the scene [2],[6].

Fig. 4. The edge processing.

In Fig. 4, edges can be clearly seen and they should be carefully done (marked by yellow arrow) especially motor edges because the motors will be more noticeable throughout the animation. If it is not done carefully the process of

colorization animation will be quite unreal and unsuccessful because the effect of the real motor will not be achieved. All the elements should be separated and done individually in order to give qualities separately and not to lose the quality of the details of the picture.

Near the borders of the regions it is considered that measures of texture are changed from one valve to another. The same goes for the detection of edges of small slope so that degrees of operation can be used for segregation of edges and pressing of non maximum during the formation of maps of edges. In that occasion wavelet transformation is used and it presses colours on grey picture, and the result of this colorization [3],[4] is shown in Fig. 5.

The edge processing is very important especially for the region determining. Our approach is basically manual one, but we are exploring some other algorithms of the automatic image processing and region determining.

Automatic region determination is unchangeable when animation and movie processing is in focus. When we work only with single images, the manual approach is satisfied (Fig. 5).

III. PARTIAL PICTURE ANIMATION

Animation is the process of picture presentation on the computer screen which changes rapidly in a sequence. It can be made with hand-made drawings, with pictures generated by computer or three-dimensional objects or by the combination of these techniques. The position of each object in any picture is related to the position (place) of that following picture, so that objects can move indepedently one from the other without any troubles. Appliances for presenting show these pictures very fast, most frequently 24, 25 or 30 frames per second.

Colorization and animation are the methods which are used in the picture with Nikola Tesla and his motor patent (Fig. 6). This „hard" picture wasn't chosen by accident, it was because it had very visible and difficult damages and beside that we wanted to move the motor which is in the picture.

However, with this project we managed to eliminate different damages, to do the colorization, to enliven motors and to move it by animation. It was needed to put great effort to be informed about Tesla's way of life and to understand the material of the motor so that animation can show the reality of that period in the best way (Fig. 7).

The main idea connected with animation is to select objects and layers appropriate for movement, construct the layer movements, and precisely control frame rate and timings. This is very important if someone wants to achieve maximum quality of animation. Also, animation phase is sequel to colorization. The quality of animation is directly connected with a colorization phase.

For Tesla's picture, we choose two layers: Tesla induction engine and Tesla's face - so there are two animations.

Fig.5. Tesla and the induction engine in color.

Fig.6. Original image.

Fig.7. Final image.

IV. ENGINE ROTOR ANIMATION

In this section of the paper will explain the animation of the final stage of Nikola Tesla color image with a rotary engine in Adobe After Effects. Before the start of this phase of the work it is necessary to process the photo into Photoshop. First we need to make a separate Layer-s for each part of the image that we want to animate. Specifically in our case, it will be the working parts of the engine. To select a tool we use "Polygonal Lasso Tool" with the help of which we select the

first part of the photo, or part of the engine that will later be animated. After selecting the first part it is necessary to copy that, selected, part of the new layer (Fig. 8). In this way, we copy the pixel information from the selected part of the Photo. The next step is to generate a new layer in which we end up copying.
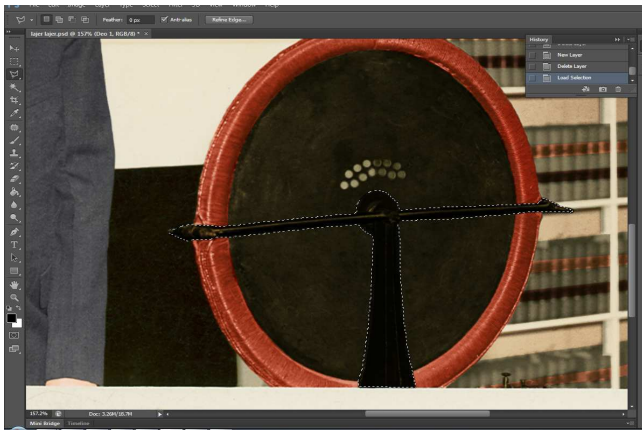

Fig. 8. Selection of the static part of the engine.

The same procedure can be applied to each portion of the image that we want to animate. In our case, that we could animate the engine, it is necessary to set aside another part of the picture. Below we use the program Adobe After Effects using the "Transform" and "Rotation". The procedure has a few steps: 1. Define layers (disk rotor and stator) 2. Define X,Y and angle parameters. 3. Define time interval. 4. Define Key-frame (the start picture). 5. Tune 1.-5. parameters and start animation. The software will generate the animation under these conditions.

Fig. 9. Selection of the moving part of the engine (rotor).

After separation of the static and dynamic part of Tesla's motor using and adjusting the rotational transformation we obtain the rotation animation (Fig. 9). This technology is extremely useful given that there are a number of photos which can be animated by a similar procedure. The result animation could be directly insert into the digital movies.

V. CONCLUSION

Picture colorization will have more and more important part in many different spheres of image processing in the future. This is the fundamental issue whose improvement opens a great number of opportunities in familiar fields of research. There are great expectations related to colorization which will enable making animations which will make easier, virtual presentation, an original patent and its way of working and its appearance. In this paper we presented a novel method for picture colorization based on mixture of manual and automatic colorization method. Also, the groundwork for possible partial picture animation is presented.

Improvement of computer technology and perfection of human ability offers wide space for practical use of sometimes too complex approach of colorization which can be later gathered and made into an animation. This offers a wide range of directions for further researches towards new possibilities and the presentation of colorization and animation. This paper has offered defined results in the sphere of science till now by showing a part of the past in a visual way. The most important aim is making different kinds of picture "re-living" in 3D films that are based on pictures similar to presented one in this paper. The results in this paper are directly implemented in project developed with Nikola Tesla Museum of Belgrade and they are installed in central Museum presentation.

ACKNOWLEDGEMENT

REFERENCES

[1] R.Yogamangalam, B.Karthikeyan, „Segmentation Techniques Comparison in Image Processing", School of Computing, SASTRA University, Thanjavur, TamilNadu, India.

[2] B.D.Venkatramana Reddy and Dr.T.Jayachandra Prasad, „Colour-Texture Image Segmentation using Hypercomplex Gabor Analysis ", Signal & Image Processing : An International Journal(SIPIJ) Vol.1, No.2, December 2010.

[3] Lilong Shi, Brian Fun, „Quaternion color texture segmentation", Computer Vision and Image Understanding 107, pp. 88–96, 2007.

[4] Junqing Chen,Thrasyvoulos N. Pappas,Aleksandra Mojsilovic and Bernice E. Rogowitz, „Adaptive Perceptual Color-Texture Image Segmentation", Submitted to IEEE transactions on image processing; revised March 4, 2004

[5] Vlatka Štimac Ljubas „Podrijetlo, sintaktička struktura i leksikografska obradba naziva za boje " , Studia lexicographica, GOD. 7 , BR. 1(12), STR. 91–115, 2013.

[6] Megumi Nishi, Takahiko Horiuchi and Hiroaki Kotera, „A Novel Picture Coding Using Colorization Technique", Graduate School of Science and Technology, Chiba University, Japan.

[7] S. Liapis, E. Sifakis, and G. Tziritas, „Colour and texture segmentation using wavelet frame analysis, deterministic relaxation, and fast marching algorithms", Department of Computer Science, University of Crete, P.O. Box 2208, Heraklion, Greece Received, 16 February 2001.

Reliable SVD-based Watermarking Algorithm Applied on Uncompressed Video

Zoran Veličković[1], Zoran Milivojević[1] and Milojko Jevtović[2]

Abstract – **The need for copyright protection of video content becomes more pronounced on the Internet. One of the ways of copyright protection is inserting invisible information - the watermark in the video content. The efficiency of the proposed algorithm is shown based on SSIM parameters obtained by an iterative algorithm for correcting the quality of the extracted watermark, ie, the occurrence of false positive problems present in conventional SVD extraction of the watermark is prevented.**

Keywords – **Digital multimedia, SVD, Reliable watermark algorithm, H.264/AVC.**

I. INTRODUCTION

Storage and sharing of digital multimedia content on the Internet is becoming the dominant form of network traffic [1]. The modern network technologies are significantly contributed to support specific communication requirements especially in the wireless environment [2]. Cross-layer communication protocols provide significant network bandwidth required for multimedia content in real time [3]. Great communication capabilities of mobile devices, as well as the easy availability of multimedia content on the Internet, have produced a range of side effects. Characteristics of digital multimedia content when copying do not lost on quality, simplify the creation of illegal copies as well as their illegal distribution. Illegal copying and distribution of multimedia contents are specifically expressed in the music and film industry. The total pirated multimedia content accounts for about 35.2% on illegal film market [4]. This paper discusses the compression algorithms for specific video content, although the results can be applied to the majority of multimedia content. Compression video algorithms are based on the imperfections of the human visual system HVS (Human Visual System). H.264/AVC is one of the most popular video coding standards for video content [5], which is discussed in this paper.

In order to prevent illegal copying and distribution of multimedia content, several methods based on hardware and/or software solutions are developed. For protection against illegal copying of multimedia content in usage are various types of cryptographic methods, as well and

techniques of inserting a watermark into the content. Some of these techniques include embedding the invisible digital information in multimedia content - watermark [6]. This concept is based on the legislation in the fight against piracy, which requires efficient and reliable proof of ownership of multimedia content. Content of the inserted watermark should unambiguously identify the owner of multimedia content and to all the time be present in it. Good inserted watermark must meet several basic criteria:

1. It should be invisible to the observer;
2. It should not cause a noticeable degradation of multimedia content;
3. It should be robust to attempt its removal;
4. It should robust to the occurrence of false positive problem;
5. It should robust to compression and transcoding video content to different bit rates.

For inserting a watermark in the video sequence in this paper, a modified SVD algorithm is used in the transformation DCT domain. Insertion of the watermark is done in uncompressed video sequence, so that the watermark is inserted into each frame. The basic idea of SVD algorithm for watermark insertion is based on the fact that most of the energy of the video frame is localized in a few singular values of the matrix that represents a frame. Small variations in the values of the singular values of video frames will not cause a noticeable degradation of quality. This characteristic of singular values has enabled the installation of singular values of the watermark in the original frame by minor modifications of singular values of video frames. However, the problem of the appearance of false positive problem with this algorithm is described in [7].

Specifically, this algorithm can extract any watermark you are looking for, so confirming his presence even though he is not inserted! To solve this problem, in this paper has been tested reliable SVD based watermarking algorithm [7], [8] which instead insertion singular values of the watermark inserts principal components of the watermark in each frame uncompressed videos. The fact that this algorithm belongs to the class of blind watermarking algorithms for extracting the watermark is necessary to possess the original image watermark. The second chapter presents the H.264/AVC encoder with their specificities and identifies potentially possible position for inserting a watermark in the algorithm for insertion. In the third chapter, the mathematical basis for the implementation of a reliable algorithm for injection and extraction of watermarks from videos based on SVD decomposition are given. The fourth chapter presents the results of watermark extraction for different watermarks, and the quality of the extracted trademarks is evaluated by objective parameter SSIM. In the fifth chapter are derived some results based on the results obtained.

[1] Zoran Veličković is with College of Applied Technical Sciences, Nis, Serbia, A. Medvedeva 20, 18000 Niš, Serbia, E-mail: zoran.velickovic@vtsnis.edu.rs

[1] Zoran Milivojević is with College of Applied Technical Sciences, Nis, Serbia, A. Medvedeva 20, 18000 Niš, Serbia, E-mail: zoran.milivojevic@vtsnis.edu.rs

[2] Milojko Jevtović is with Engineering Academy of Serbia, Beograd, Serbia, K. Miloša 9/IV, 11000 Beograd, Serbia, E-mail: vladmijev@ptt.rs
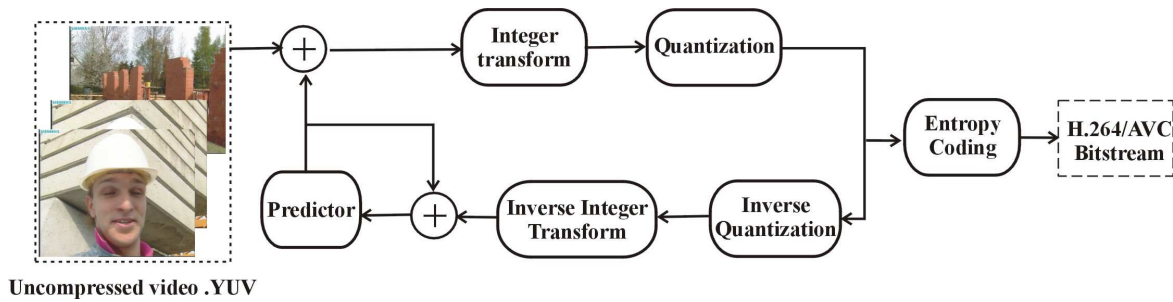
Fig. 1. Basic processing blocs of H.264/AVC encoder

II. H.264/AVC ENCODER/DECODER

Video presents a series of correlated images in the spatial and temporal domain [5]. When say H.264/AVC video, actually we refer to the video sequence, which is represented by a specific format. Video compression algorithms implemented in H.264/AVC standard are based on the elimination of additional information from the temporal, ie, the spatial domain [4]. The specific format is defined in the standard, which is implemented using the H.264/AVC syntax elements that describe different aspects of coding sequences, as well as the ways in which they presented individual elements. The syntax of H.264/AVC encoder is a hierarchical structure consisting of video sequences on top of the hierarchy to the individual frames and macro blocks on the bottom. The control parameters are stored in a special syntax section PS (Parameter Sets) or they are part of the macro blocks. Syntax elements are stored in an array of bytes RBSP (Raw Byte Sequence Payloads), which was later encapsulated in NALUs (Network Abstraction Layer Units). VCL (Video Coding Layer), which is part of the H.264/AVC standard, defines coded parts of the image (slices) in NALU. Each slice consists of a header and data are coded macro blocks. The coded sequence begins with IDR (Instantaneous Decoder Refresh) followed by the other coded parts - Data Partition Slices. Specialized NALU's Parameter Set and Supplemental Information Ehancement are not part of the VCL and carry additional information, which is not necessary for decoding H.264 stream.

Fig. 1. shows the basic blocks for processing of uncompressed video, which compose the H.264/AVC encoder. Predicting the content of the current frame is performed in the block "*Predictor*" based on one or more previous or future frames. The frames based on which the contents of the current frame are called reference frames. H.264 / AVC standard defines the frames of type I (intra), P (inter) and B (bidirectional) that uses one or more (past or future) reference frames. A powerful mechanism for exploring the similarities in the current figure or figures, which precede, respectively, below, is the basic strength of H.264 encoder. Anticipating the contents of certain parts of the image based on the perceived similarities, it is possible to form a "residual frame" with much less data. The consequence of this approach can be neglect of fine detail in the frame, which will have a negative effect on the inserted watermark. The consequence of this approach is the variable quality of the video, thus

separated watermark, especially at lower bit rates. In previous studies, we tested the survival of the inserted watermark in the video encoded H.264/AVC encoder [8].

For professional use, the standard version of the H.264/AVC encoder is extended giving new coding tools. This is the extended version is known as FRExt. FRExt version of the H.264/AVC standard is enriched with a new High profile (HP). HP support higher resolution video without changing the sampling scheme.

III. ALGORITHMS FOR WATERMARK INSERTING/EXTRACTING

A. Algorithms for watermark insertion

Insertion of watermark to video content can be realized in each of the processing blocks of H.264/AVC encoder. Depending on the position of the processing block algorithms differ inserting the watermark inside or outside reconstructive loop of H.264 encoder. Insertion of the watermark in the video before encoding is the first option that was used in this and previous papers [9]. This insertion algorithm is implemented outside of the reconstructive loop of H.264/AVC encoder. Another possibility is the insertion of watermark into the structure of video codecs (such as "motion vector" or a new syntax element Reference Index). A third possibility occurs in the process of transformation, whereas the fourth option occurs in the quantization process. It is clear that the last three algorithms implemented within reconstructive loop of H.264/AVC encoder. Last possibility of inserting a watermark in compressed video stream and it is implemented outside the loop reconstructive H.264/AVC encoder.

In this paper, watermark insertion algorithm before encoding the video content is used. The applied concept allows insertions of the watermark in each frame, which will in the extraction process to provide a range of extracted watermarks from each frame. There are the two major classes of algorithms for watermarks insertions in uncompressed video. The first class of algorithms is based on inserting a watermark in the spatial domain, while the second class of algorithms is based on modifying the coefficients in the transformation domain. In case the first class of algorithms applied to images or video, watermark is hidden in the values of luma and/or chroma components spatially distributed pixel images [7], [8], [9].
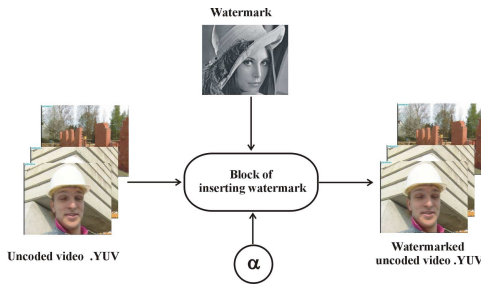
Fig. 2. Insertion of the watermark before encoding H.264 / AVC encoder

Algorithms which belonging to this class is relatively easy to implement but are not sufficiently robust when processing video signals, especially in the implementation of video compression. Another class of algorithms is based on the modification (modulating) the transformation coefficients of videos based on the transformation coefficients of the watermark. The transformation coefficients can be obtained by the DCT (Discrete Cosine Transform), FFT (Fast Fourier Transform) or SVD-a (Singular Value Decomposition). The inserted information can be extracted by inverse procedure from the modified transformation coefficients [7]. The algorithms of this class have higher performance robustness in relation to the watermark inserted in the spatial domain. Fig. 2. shows the algorithm inserting a watermark in each frame uncompressed video with inserting factor $0 < \alpha < 1$, while the encoding profile defined set of parameters H.264/AVC encoder.

B. Reliable SVD algorithm

In this paper, for inserting a watermark in the video sequence, algorithm based on SVD decomposition are used. This algorithm represents an improved version of the algorithm presented in [7], which eliminates the problem of false positives watermark. Applying this algorithm to each frame of the uncompressed video sequence, protected video are obtained. The input or the output parameters of the algorithm for the installation of the watermark are:

Input:
- A series of matrix $A_{m \times n}$ which representing uncompressed frames of video sequences.
- Matrix $W_{m \times n}$ which representing figure – watermark for embedding in video sequence.
- Inserting factor α.

Output:
- Video with embedded watermark – series of the matrix $A_{w_{m \times n}}$.

The algorithm for inserting a watermark in a video frame is displayed in four **I** steps:

Step I$_1$: SVD decomposition of the matrix, frame A:

$$A = USV^T , \qquad (1)$$

where A is the original frame, U and V are orthogonal matrices of dimensions m × n and n × n, respectively, with a diagonal matrix S of dimension m × n with elements that represent singular values. The columns of the matrix U are called the left singular vectors, while the columns of the

matrix in the right singular vectors. Singular vectors specified image geometry while singular values specified luminance (energy) images.

Step I$_2$: SVD decomposition of watermark:

$$W = U_w S_w V_w^T = A_{wa} V_w^T . \qquad (2)$$

Step I$_3$: Insertion of the principal components of the A_{wa} in diagonal matrix S with inserting factor α:

$$S_1 = S + \alpha A_{wa} . \qquad (3)$$

Step I$_4$: Creating images with watermark:

$$A_w = US_1V^T . \qquad (4)$$

The extraction of the watermark from image A_w^* that is potentially due to superimposed noise different from A_w, are implemented by algorithm, which consists of three **E** steps:

Input:
- Video with inserted watermark, series of matrices A_w^*,
- Original video, a series of matrices A,
- Inserting factor α.

Ouput:
- Extracted watermark $W*$.

Step E1: Creating differences from the Original and watermarked frame:

$$\left(A_w^* - A \right) = A_1 \qquad (5)$$

Step E$_2$: Computing *principal* component:

$$A_{wa}^* = \frac{\left(U^{-1} A_1 \left(V^T \right)^{-1} \right)}{\alpha} \qquad (6)$$
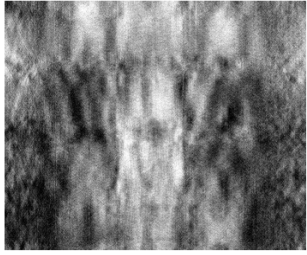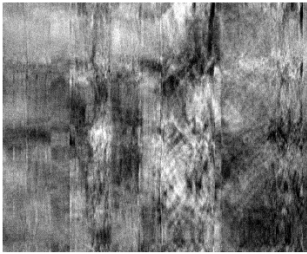
Step E$_3$: Computing watermark $W*$ as follows:

$$W^* = A_{wa}^* V_w^T \qquad (7)$$

IV. EXPERIMENT AND RESULTS

As watermarks in this paper are used central parts of the figures [10]: "*Baboon*","*Boat*" and "*Lena*" shown in the left column of Table I, respectively. Resolution of images that are used as watermarks and the resolution of the video to be protected are 352×288 pixels. In the experiments, we used the first 50 frames of uncompressed video stream "*Foreman.cif*". For inserting the watermark in all frames, the constant factor of insertion $\alpha = 0.05$ is used. The encoding and decoding of video sequences was performed JM reference software ITU (International Telecommunication Union) in version 18.4 FRExt [11]. Basic encoding parameters used in the experiment belong HIGH profile: ProfileIDC=100, IntraPeriod=12, NumberReferenceFrames=5, and NumberBFrames=7. To improve the quality of the extracted watermark was used in the presented iterative algorithm [9].

TABLE I

WATERMARKS APPLIED IN EXTRACTION ALGORITHM (LEFT COLUMN)
BABOON, BOAT AND LENA, AND EXTRACTED WATERMARKS (RIGHT
COLUMN)

Watermarks used in extraction algorithm	Extracted watermarks

To protect the video sequence "*Foreman*", watermark "*Lena*" is inserted in each frame. The watermark is inserted/extracted using a reliable SVD algorithm described in Chapter II. In the process of extracting, all three watermarks from Table I are used. To assess the quality of the extracted watermark SSIM parameter [12] was used. In the right column of Table I shows the obtained results - extracted watermarks. From Table I it can be concluded that in cases where wrong watermark ("*Baboon*" and "*Boat*") is used in the extraction process, the algorithm failed to allocate inserted watermark. Also, based on the obtained results, can be concluded that the reliable SVD algorithm is resistant to the occurrence of false positive problem. SSIM parameter was 0.122 when watermark "*Baboon*" is used in the extraction process, while, for the case of using watermark "*Boat*", SSIM parameter was 0.125, which clearly shows that the occurrence of false alarm problem solved.

In the case when the original watermark "*Lena*" is used in the process of extracting", reliable SVD algorithm has shown excellent results. SSIM parameter in this case was 0.64, which is more than five times better compared to the use of incorrect watermark. From the presented examples can be clearly concluded that the proposed reliable SVD algorithm for insertion/extraction of the watermark can be effectively applied to protect the video of code H.264/AVC encoder. The results indicate that the false positive problem solved

efficiently, which was removed a basic lack of standard SVD algorithm for insertion/extraction watermark. All the advantages of the standard SVD algorithm are preserved, and reliable SVD algorithm eliminates the drawbacks described above.

V. CONCLUSION

In order to prevent illegal copying and distribution of video content is often used technique for inserting a watermark in uncompressed video. In the process of encoding video content H.264/AVC encoder comes to video degradation, and therefore the inserted watermark, which significantly complicates the extraction. This paper, analyzes the reliable SVD algorithm for insertion/extraction of the watermark from the video that is coded H.264/AVC encoder. Results of this study show that the reliable SVD algorithm is resistant to the occurrence of false positive problems when inadequate watermarks are used in the extraction process. When used properly watermark in the extraction process, using an iterative algorithm to improve the extracted watermark, obtained satisfactory results. In this way, the authors of multimedia content can prove ownership of the video content and thus gain their copyright. Reliable SVD algorithm can be efficiently used to protect video content while preserving all the positive characteristics of classical SVD algorithm.

REFERENCES

[1] Cai, Lin, Xuemin Shen, Jon W. Mark, Multimedia services in wireless internet: modeling and analysis, Vol. 15. John Wiley & Sons, 2009.

[2] Z. Veličković, M. Jevtović, V. Pavlović, "Cross-layer Throughput Optimization in Slow Wireless Fading Channel", International Journal Elektronika Ir Electrotechnika, Vol. 19, No 6, pp. 131-137, 2013.

[3] M. Jevtović, Z. Veličković, "Protokoli prepletenih slojeva", Akademska misao, Beograd, 2012.

[4] http://www.go-gulf.com/blog/online-piracy/

[5] ITU-T, Recommendation H.264, Advanced Video Coding for Generic Audiovisual Services. Technical report, ITU-T, 2011.

[6] Z. Shahid, M. Chaumont, W. Puech, "Considering the reconstruction loop for data hiding of intra and inter frames of H.264/AVC", Signal, Image and Video Processing, Vol. 7, pp. 75-93, 2013.

[7] R. Run, S. Horng, J. Lai, T. Kao, R. Chen, "An Improved SVD-based watermarking technique for copyright protection", Expert Systems with Applications 39, pp. 673-689, 2012.

[8] C. Jain, S. Arora, P. Panigrahi, "A Reliable SVD based Watermarking Scheme", Journal CoRR, vol. abs/0808.0309, 2008.

[9] Z. Veličković, Z. Milivojević, M. Jevtović, „Iterative Algorithm for improvements of the quality of the watermark in the H.264 encoded video", Vol. 2, pp. 112-117, UNITECH, 2014.

[10] http://www.ece.rice.edu/

[11] JM reference software version 16.0, http://iphome.hhi.de/suehring/tml/, July 2009.

[12] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, Image Quality Assessment: From Error Visibility to Structural Similarity, IEEE Trans. on Image. Proc., Vol. 13, No. 4, 2004.

Comparative Analysis of the MSD and MSDM Watermarking Algorithms Based on the Schur Decomposition

Bojan Prlinčević [1], Zoran Milivojević[2], Petar Spalević[3], Darko Brodić[4]

*Abstract –***The first part of the paper describes the SDM watermarking algorithm, based on the Schur decomposition, and the MDB algorithm for detection and removal of impulse noise. The SDM algorithm has been modified in the part for watermark extraction. In the second part of the paper, proposed algorithm is tested, for resistance on superimposed impulse noise and filtering with the MDB algorithm. Testing results are presented in tables and in the graphs.**

Keywords – **Schur decomposition, watermarking, impulse noise, noise detection, filtering.**

I. INTRODUCTION

The widespread use of computer networks leads to intensive exchange of multimedia data (pictures, video, audio,...). However, copyright protection and proof of ownership is a big problem in distribution of digital images. In order to solve these problems, a principle is used of inserting visible or invisible information in the image with aim to prove ownership of the author, i.e. distributor. Inserting of hidden information in the digital image with the aim to prove ownership rights is called a digital watermark [1], [2]. The basic characteristics of digital watermark are insensitivity, robustness, capacity, non-inverse and ability to provide positive proof of ownership [3]. For inserting a watermark in the image, many transformations could be used such as DCT [3], [4] (*Discrete Cosine Transform*), DWT [5], [6] (*Discrete Wavelet Transform*), SVD (*Singular Value Decomposition*) transformation [2], [7], and Schur Decomposition. Insertion of the watermark by applying Schur decomposition can be performed for inserting the watermark into unitary matrix [8] (SD watermarking algorithm) and inserting the watermark in the upper triangular matrix [9] (SDM watermarking algorithm).

In this paper a modification of SDM algorithm for inserting a watermark has been done [9] with aim to increase efficiency in extracting the watermark. In addition, an analysis of the efficiency of the algorithm in conditions with existence of impulse noise has been done. On watermarked image, impulse noise, salt and pepper, are superimposed. From watermarked image with superimposed impulse noise a watermark has been

extracted. Algorithm for detecting and removing impulse noise is applied on watermarked image with superimposed noise. From filtered image a watermark has been extracted. Based on the obtained results which are presented in tables and graphs, a comparative analyze was made with results obtained in the paper [10]. As a measure of quality, MSE is applied (Mean-Square Error).

The paper is organized in the following way: In section II Schur decomposition was described and MSDM algorithm for inserting and extracking the watermark, based on it. In section III are presented the obtained results and analyze of results. Conclusion is given in section IV.

II. ALGORITHM

A. Shur decomposition

For apply Schur decomposition, image $A_{M \times N}$ is divided on blocs, $H_{Mb \times Nb}$. Schur decomposition is performed over each blocs:

$$A = U \times D \times U ', \qquad (1)$$

where $U_{Mb \times Nb}$ is unitary matrix and $D_{Mb \times Nb}$ is upper triangular matrix. In the paper [10] MSD watermarking algorithm has been presented, that is performed by inserting the watermark into a unitary matrix U. In the paper [9] SDM watermarking algorithm is presented, in which the insertion of one bit of the watermark is performed in the diagonal of the upper triangular matrix D. In this paper the efficiency of the algorithm in terms of SDM superimposed impulse noise has analyzed. In addition to this, over an image in which watermark is inserted and superimposed impulse noise, a DB algorithm for removing impulse noise has been applied, which was proposed in the paper [11] and which is modified from authors of this paper in order to increase the efficiency of detecting the infected pixels (MDB algorithm) [12]. The following part of the paper describes the SDM algorithm from [9] and its modification (MSDM).

B. SDM algorithm

SDM watermarking algorithm for inserting watermark is realized trough the following steps:

Input: original image $A_{M \times N}$, binary watermark $w_{Mz \times Nz}$, bloc dimension $M_b \times N_b$.

Output: image with the watermark A_w.

Step 1: The original matrix A divided on $X \times Y$ blocs $H_{Mb \times Nb}$, where $X = \lceil M / M_b \rceil$ and $Y = \lceil N / N_b \rceil$.

Step 2: Schur decomposition is applied to all blocs H:

[1]Bojan Prlinčević, Higher Technical Professional School in Zvecan, Nušićeva 6, Zvečan, Serbia, b.prlincevic@vts-zvecan.edu.rs.
[2]Zoran Milivojević, College of Applied Technical Sciences, A. Medvedeva 20, 18000 Nis, Serbia, zoran.milivojevic@vtsnis.edu.rs.
[3]Petar Spalević, Faculty of Technical Science Kosovska Mitrovica, petar.spalevic@pr.ac.rs.
[4]Darko Brodić, University of Belgrade, Tehnical Faculty Bor, V. Jugoslavije 2, 19210 Bor, Serbia, E-mail: d.brodic@tf.bor.ac.rs.

$$H_{i,j} = U_{i,j} \times D_{i,j} \times U_{i,j}^T , \qquad (2)$$

where U is unitary matrix, D is upper triangular matrix and $1 \le i \le \lceil M / M_b \rceil$ and $1 \le j \le \lceil N / N_b \rceil$.

Step 3: Insertion of one bit of watermark, $bw_{i,j}$ in $D_{i,j}$ bloc of matrix $H_{i,j}$:

$$D_{i,j}^{'} = D_{i,j} \ (I + \alpha \ I \ bw_{i,j}) , \qquad (3)$$

where I is unit matrix and $D'_{i,j}$ is upper triangular matrix with embedded watermark.

Step 4: Reconstruction of the bloc with embedded watermark:

$$H_{i,j}^{'} = U_{i,j} \times D_{i,j}^{'} \times U_{i,j}^T , \qquad (4)$$

Step 5: Watermarked image A_w is obtained from blocs H'.

SDM watermaking algorithm for extracting the watermark is realized through the following steps:

Input: Watermarked image A_w, bloc dimensions $M_b \times N_b$.
Output: Reconstructed binary watermark $w'_{Mz \times Nz}$.

Step 1: The watermarked matrix A_w is divided on $X \times Y$ blocs $H'_{Mb \times Nb}$, where $X = \lceil M / M_b \rceil$ and $Y = \lceil N / N_b \rceil$.

Step 2: Schur decomposition is applied to all blocs H':

$$H_{i,j}^{'} = U_{i,j}^{'} \times D_{i,j}^{'} \times (U_{i,j}^{'})^T , \qquad (5)$$

where U is unitary matrix, D' is upper triangular matrix and $1 \le i \le \lceil M / M_b \rceil$ and $1 \le j \le \lceil N / N_b \rceil$.

Step 3: Extraction one bit of watermark, bw' from the matrix D':

$$I_M \ w'_{i,j} = D_{i,j}^{-1} \ (D'_{i,j} - D_{i,j}) / \alpha \ \Rightarrow$$

$$bw'_{i,j} = \frac{diag(D'_{i,j} - D_{i,j}) \ diag(D'_{i,j} - D_{i,j})^T}{\sqrt{diag(D_{i,j}) \ diag(D_{i,j})^T}} . \qquad (6)$$

Step 4: Watermark w' is obtained from extracted bits bw'.

C. MSDM algorithm

In order to increase the quality of extracted watermark, authors of this paper are modified the SDM algorithm from [10] (MSDM algorithm). A modification of the following steps of SDM algorithm was done:

a) in the part for inserting the watermark:

Step 3: Diagonal of block of upper triangular matrix D has analyzed. The block is not appropriate, for insertion bit of the watermark, if the diagonal has two or more elements with the value 0.

b) in the part for extracting the watermark:

Step 3: Extraction one bit of watermark, bw' from the matrix D':

  IF   (real(D'$_{(1,1)}$) - real(D$_{m(1,1)}$)) > 0.01 | (real(D'$_{(1,1)}$) -...
    ...real(D$_{m(1,1)}$)) < -0.01
      $bw'_{i,j}$=1;
  ELSE
      $bw'_{i,j}$=0;
  END.

III. EXPERIMENTAL RESULTS AND ANALYZE

A. Experiment

For the purpose of testing the watermarking algorithm based on Schur decomposition (MSDM watermarking algorithm) the following experiment was conducted:

Step 1: Original image A is divided into blocs $M_b \times N_b$. Using algorithm based on Schur decomposition (MSDM watermarking algorithm), in blocs is inserted binary watermark W dimensions $M_w \times N_w$ with coefficient of inserting α. Watermaked image A_w is obtained from blocs with inserted watermark.

Step 2: In the watermarked image A_w impulse noise were superimposed (salt and pepper) with different percent of p.

Step 3: From the watermaked image with superimposed impulse noise, A_{ws} watermark has been extracted W_e

Step 4: Over an image A_{ws} MDB algorithm for detection and elimination of impulse noise has been applied.

Step 5: From the filtered image A_w^* a watermark has been extracted $W_e^{'}$.

As a measure of quality of filtered image and extracted watermark, the mean squared error is applied MSE:

$$MSE = \frac{\sum_{ij} \left(x_{ij} - y_{ij} \right)^2}{M \times N} \qquad (7)$$

where x_{ij} - is pixel element of original image, y_{ij} - is pixel element of reconstructed image, $M \times N$-image dimensions.

The values of the coefficient of inserted watermark are α={0.001, 0.025, 0.05, 0.075, 0.1}. Impulse noise was varied from p=10-70%.

B. Base

Images (dimensions 512x512) presented on Fig. 1 presents the image's base for the experiment: a) Lena, b) Girl, c) Baboon, d) Barbara, e) Boat and f) Peppers. Image, dimensions (128×128), presented in the Fig. 2, were used as watermark.
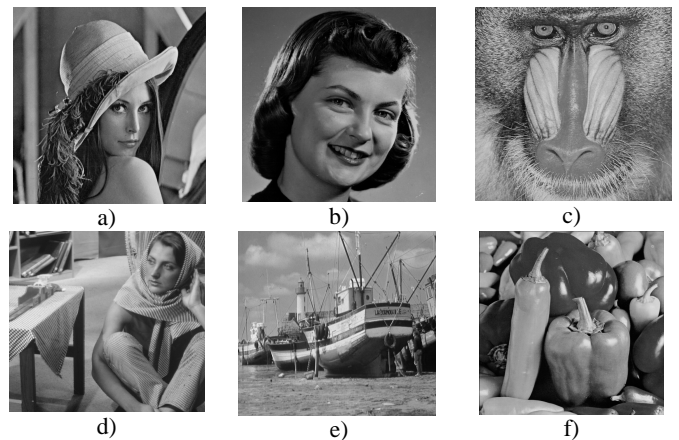

Fig. 1 Images used in the paper: a) Lena, b) Girl, c) Baboon, d) Barbara, e) Boat and f) Peppers.

POLYTECHNIC
SCHOOL FOR
VOCATIONAL
STUDIES
ZVECAN

Fig. 2. Watermark.

C. Results

The Table I presents the results of MSE for image Lena, after watermarking with MSD algorithm (MSE_{x1}), and results after applying the MDB algorithm for removing impulse noise (MSE_{xr1}). Also the Table I, presents results of MSE for Lena, after watermarking with MSDM algorithm (MSE_{x2}) and results after applying the MDB algorithm (MSE_{xr2}). The table II presents the results for the MSE for extracted watermark. Results for MSE are represented with: a) MSE_{ws1} (with superimposed impulse noise), b) MSE_{wr1} (after filtering impulse noise with MDB algorithm) for watermark inserting with MSD algorithm. Results for MSE are represented with: a) MSE_{ws2} (with superimposed impulse noise), b) MSE_{wr2} (after filtering impulse noise with MDB algorithm) for watermark inserting with MSDM algorithm. Fig. 3 shows diagrams for MSE for Lena: a) after watermarking with SDM and MSDM algorithms (MSE_x) and b) after filtering with MDB algorithm (MSE_{xr}) also shows MSE for extracted watermark from Lena c) after superimposed impulse noise and d) after applying MDB algorithm. Fig. 4 shows the appearance of extracted watermark (MSD algorithm): a) without attacks (Fig. 4.a), b) after superimposed impulse noise (p=10%) (Fig. 4.b) and c) after filtering (Fig. 4.c). Also Fig. 4 shows the appearance of extracted watermark (MSDM algorithm): a) without attacks (Fig. 4.d), b) after superimposed impulse noise (p=10%) (Fig. 4.e) and c) after filtering (Fig. 4.f). Fig. 5 shows the appearance of test images watermarked with MSDM algorithm, with the inserting coefficient α=0.05 and superimposed noise of p=30% (Figs. 4.a, 4.b and 4.c), and image appearance after applying the MDB algorithm for filtering (Figs. 4.c, 4.d and 4.e).

TABLE II
MEAN SQUARE ERROR FOR THE WATERMARK

Coef. Inserting (α)	Perc. Noise (p)	MSE_{ws1} (10^{-3})	MSE_{wr1} (10^{-3})	MSE_{ws2}	MSE_{wr2}
0.01	0	5	7.7	7693	7921
	10	300	110	9436	9396
	30	470	230	9829	9685
	50	500	320	9806	9609
	70	520	390	9585	9218
0.05	0	0.18	0.67	7694	7921
	10	270	20	9436	9404
	30	430	68	9829	9699
	50	470	130	9809	9632
	70	510	230	9586	9273
0.1	0	0.12	0.18	7694	7921
	10	240	4.5	9438	9411
	30	390	28	9829	9706
	50	450	80	9809	9651
	70	490	180	9586	9336
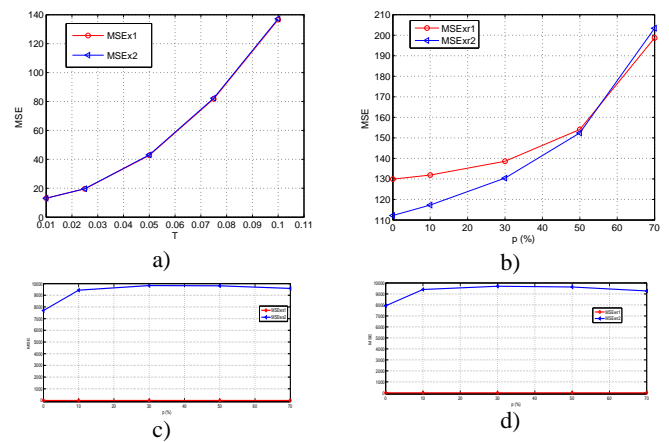
Fig. 3. MSE for: a) Lena after watermarking, b) Lena after watermarking and filtering with MDB algorithm, c) extracted watermark from Lena after superimopsed impulse noise, d) extracted watermark from Lena after filtering with MDB algorithm
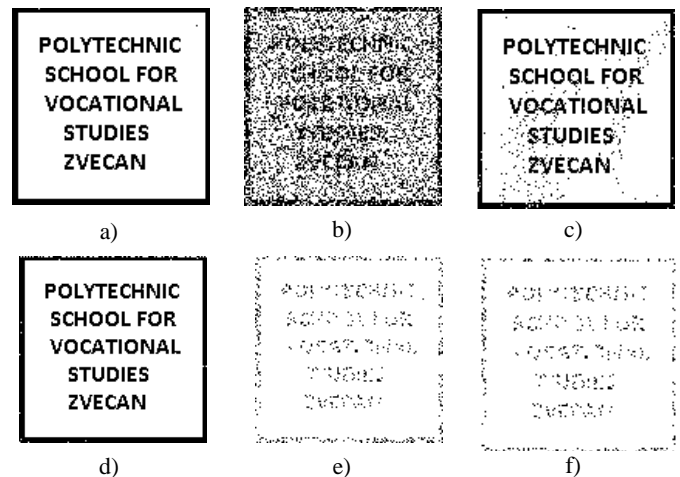
TABLE I
Mean Square Error for image Lena

Coef. Insert (α)	Perc. Noise (p)	MSE_{x1}	MSE_{xr1}	MSE_{x2}	MSE_{xr2}
0.01	0	13.1	100.2	13.1	88.2
	10		104.8		93.6
	30		116.6		107.4
	50		137.1		129.9
	70		186.7		181.6
0.05	0	42.8	129.9	43.1	112.2
	10		131.9		117.3
	30		138.6		130.4
	50		154.1		152.4
	70		198.7		203.4
0.1	0	136.6	223.3	137.2	187.4
	10		217.8		191.9
	30		209.3		203.7
	50		209.1		224.8
	70		238.2		274.1

Fig. 4. The extracted watermark (MSD algorithm): a) without attacks, b) after superimposing impulse noise (p=10%), c) after filtering with MDB algorithm; and watermark extracted (MSDM algorithm): d) without attacks, e) after superimposing impulse noise (p=10%), f) after filtering with MDB algorithm.

Fig 5. The watermarked test images, with inserting coefficient α=0.05 and superimposed noise p=30%: a) Lena, b) Girl, c) Baboon, d) Barbara, e) Boat, f) Peppers, and images after applying the MDB algorithm: g) Lena, h) Girl, i) Baboon, j) Barbara, k) Boat and Peppers.

C. Analyse

Based on the results shown in Tables I and II and in Figs 1-5 can be concluded that:

a) Applied the MSDM watermarking algorithm for inserting watermark in the image, does not lead to visual degradation.

b) The watermark, which is inserted into the image with MSDM algorithm, after extraction, visually is satisfactory quality.

c) Visual characteristics of watermarked image, inserted with MSDM algorithm, are possible to improve for very high percentage of superimposed impulse noise, even up to 70%.

d) After superimposing of 10% impulse noise to watermarked image, extracted watermark, applying MSDM algorithm, visually is not satisfactory quality.

e) MSE of extracted watermark (p=10%) after filtering with MDB algorithm, for MSD algorithm applied in the paper [11] is MSE=0.27, while for MSDM algorithm, presented in this paper, is MSE=9436, which present statistically extremely high percentage of errors.

f) MSE of extracted watermark (p=30%) after filtering with the MDB algorithm is MSE=0.068 (MSD algorithm) which is for 14×10^4 times less in relation to MSE of watermark extracted by applying the MSDM algorithm.

IV. CONCLUSION

The paper analysed the MSDM watermarking algorithm and comparing with MSD algorithm. Watermarked images are filtered with the MDB algorithm for detection and removing impulse noise. The analysis was done for the coefficient of insertion α={0.01,0.025,0.05,0.075,0.1} with variation of superimposed impulse noise p=10-70%. Measure of success of algorithms is demonstrated with quality measure MSE.

Detailed comparative analysis of the MSE parameters, and visual appearance of test images and extracted watermark, indicates that the MSD algorithm [11] is suitable for inserting a watermark and shows satisfactory resistance for superimposed impulse noise. Watermark extracted, with SDM algorithm [10] modified in this paper (MSDM algorithm), is visually satisfactory quality.

REFERENCES

[1] .J. Cox, M.L. Miller and J.A. Bloom, "Digital Watermarking", Morgan Kaufmann Publishers, 2002.

[2] J.Hernandez, M. Amado, F. Perez-Gonzalez, "DCT-domain watermarking techniques for still images: detector performance analysis and new structure", IEEE Trans. Image Process. 9 (January 2000) 55-67.

[3] R. Liu, T.Tan, "A SVD based watermarking scheme for protecting rightful ownership", IEEE Trans. Multimedia 4 (1) (march 2002) 121-128.

[4] B.Verma, S.Jain, D.P. Agarwal "Spatial Domain Robust Blind Watermarking Sheme for Color Image" Asian Journal of Information Technology 6(4):430-435, 2007.

[5] S.Baba, L.Krikor, T.Arif, Z. Shaaban, "Watermarking of digital images in Frequency Domain", International Journal of Automation and Computing, May 2009.

[6] W.Chu, "DCT-based image watermarking using subsampling", IEEE Trans. Multimedia 5 (1) (March 2003) 34-38.

[7] A.Reddy, B.Chatterji, "A new wavelet based log-watermarking scheme", Patern Recognition Lett. 26(may 2005) 1019-1027.

[8] Q.Su, Y.Niu, X.Liu, Y.Zhu, "Embedding color watermarks in color images based on Schur decomposition", Optics Communications, 285 (2012) p.p. 1792-1802.

[9] A. A. Mohamed, "A new digital image watermarking scheme based on Schur decomposition", Multimed tools appl (2012) 59: 851-883.

[10] B.Prlincevic, Z.Milivojevic, P.Spalevic, D.Brodic "Performance of the SD algorithm for inserting watermark in image based on the Schur decomposition", XIV International Scientific Symposium, Infoteh, 18-20 March, Jahorina, 2015.

[11] K.S.Srinivasan and D. Ebenezer, „A New Fast Efficient Decision-Based Algorithm for Removal of High-Density Impulse Noises", IEEE signal Processing Letters, vol.14, No.3, March 2007.

[12] B.Prlincevic, Z.Milivojevic, D.Brodic "Efficiency of MDB algorithm for filtering watermarked images", conf. IT 2014, Zabljak, 2014.

Automatic Embryonic Detection in Microscopy Images

Veska Georgieva[1]

Abstract – **In this paper is proposed an automatic embryonic stem cell detection and determination of their size and location in microscopy images. The presented approach includes pre-processing stage for obtaining better information for the number and state of the investigated cells. It includes noise reduction based on homomorphic filter and correction of illumination of the fluoroscopic microscopy images. Then an automatic method to detection and location of the embryo is applied. It is based on Hough Transform to approximate the embryo as a circle. Experimental results showed that the proposed method can detect the position of the eligible embryo accurately. The obtained result can be used to extract criteria for embryo transfer purpose.**

Keywords – **Automatic cell detection, microscopy images, homomorphic filter, illumination correction, Hough transform.**

I. INTRODUCTION

Embryonic stem cells are found in various parts of the human body at every stage of development from embryo to adult and are classified according to their potential to develop into other cell types. Using different cell markers, specialists are able to determine, by manual counting, the total number of cells, how many specialized itself into a specific mature cell and how many cells died [1].

There are methods [2, 3] to identify and quantify sections of cells cultured in suspension. However, these methods are expensive and require a trained technical specialist. Another disadvantage is that the spatial information is lost, because the cells must be separated. This information is important because the specialists are able to observe some phenomena, such as the differentiated cells are located in the colony's extremity while the specialized stem cells are located at the colony's center [4].

Although the main problem was the overlapping of the cells in the images, it was also found that the size (magnification) and the brightness also varied from one image to another [5]. Noise components, blurring and illuminations artefacts are found in the most of images.

In this paper is proposed an approach for automatic embryonic stem cell detection and determination of their size and location in microscopy images. This approach can also be applied in other groups of objects, as long as the object surface is both smooth and concave with illumination source.

The remainder of this paper is organized as follows: section 2 describes the basic steps in the proposed algorithm; section

3 presents the experimental results to show the effectiveness of the method and; section 4 presents our conclusion and some future works.

II. BASIC STAGES IN ALGORITHM FOR IMAGE PROCESSING

In the paper is proposed an effective algorithm for automatic embryonic stem cell detection and determination of their size and location in microscopy images. It consists of following basic stages:

- Noise reduction
- Correction of illumination
- Edge Detection via the Canny Operator
- Circle Detection via the Circular Hough Transform

The block diagram of the algorithm is presented in Fig.1.
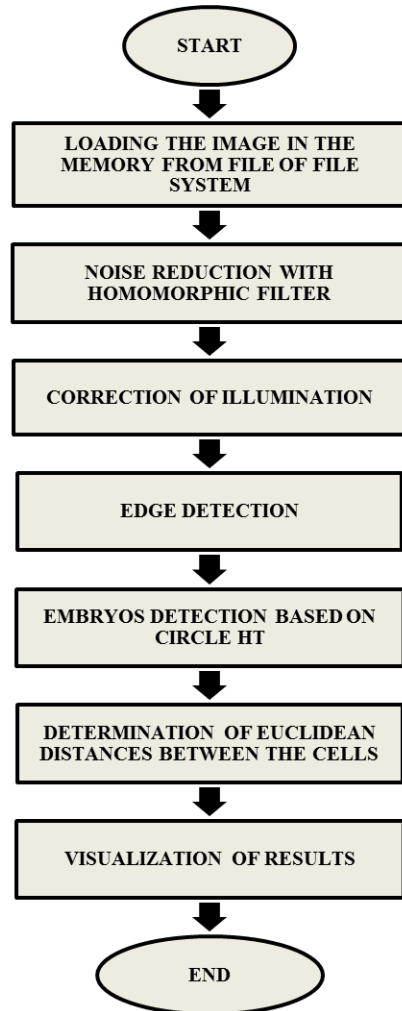
Fig. 1. Block diagram of the basic algorithm

[1] Veska Georgieva is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: vesg@tu-sofia.bg

Image pre-processing is performed in order to reduce or eliminate noise and enhance the visual quality.

Homomorphic filtering (HF) can eliminate non-uniformity luminance distribution of image, and keep its original state. The standard homomorphic filtering uses illumination-reflectance model in its operation. This model consider the image is been characterized by two primary components. The first component is the amount of source illumination incident on the scene being viewed $i(x,y)$. The second component is the reflectance component of the objects on the scene $r(x,y)$. The image $f(x,y)$ is then defined as [6-8]:

$$f(x, y) = i(x, y) r(x, y) \qquad (1)$$

In this model, the intensity of $i(x,y)$ changes slower than $r(x,y)$. Therefore, $i(x,y)$ is considered to have more low frequency components than $r(x,y)$. Using this fact, homomorphic filtering technique aims to reduce the significance of $i(x,y)$ by reducing the low frequency components of the image. However, before the transformation is taking place, logarithm function has been used to change the multiplication operation of $r(x,y)$ with $i(x,y)$ in Eq.(1) into addition operation.

$$z(x, y) = \ln f(x, y) = \ln i(x, y) + \ln r(x, y) \qquad (2)$$

Illumination correction is based on background subtraction. This type of correction assumes the scene is composed of a homogeneous background and relatively small objects brighter or darker than the background. There are two major types of background subtraction techniques depending on whether the illumination model of the images can be given as additional images or not: prospective and retrospective correction. Our approach uses retrospective correction for color images. The image is converted in this case into the HSL color space and then the correction to the lightness channel is applied. The background is estimated by mathematical morphology opening or closing. The estimated background is then subtracted from the original image. The total sequence of operations corresponds to a top-hat of the image. Top-hat removes high frequencies (considered as reflectance) and keeps low frequencies (considered as illumination). Bottom - hat is used for clear background and top-hat is used for dark background. If the background is clear, the corrected image g(x,y) is obtained using:

$$g(x, y) = T_B[f(x, y)] + mean\{ clo\,sing[f(x, y)]\} \qquad (3)$$

$$g(x, y) = [f(x, y)] - clo\,sin\,g[f(x, y)] + \\ + mean\{ clo\,sing[f(x, y)]\} \qquad (4)$$

where mean [closing (f(x,y)] is the mean value of the closed image and T_B is bottom – hat transform of the image. The bottom-hat returns an image, containing the "objects" or "elements" that: are "smaller" than the structuring element, and are darker than their surroundings. The size, or width, of the elements that are extracted by the top-hat transforms can

be controlled by the choice of the structuring element. The bigger the latter, the larger the elements extracted.

For detection of the embryos cells is proposed to perform as next edge detection via the Canny Operator. It is optimal with regards to the following criteria [9]:

1.   Detection: The probability of detecting real edge points should be maximized while the probability of falsely detecting non-edge points should be minimized. This corresponds to maximizing the signal-to-noise ratio (SNR).

2.   Localization: The detected edges should be as close as possible to the real edges.

3.   Number of responses: One real edge should not result in more than one detected edge (one can argue that this is implicitly included in the first requirement).

The Circular Hough Transform is useful for detecting circles of known radius as well to detect circles of various radii. This method is based on creating an accumulator matrix of size of the original image to be processed. The local maxima in accumulator space are obtained by voting procedure. Parameter space is defined by the parametric representation used to describe circles in the picture plane, which is given by Eq.5[10]:

$$r^2 = (x - x_0)^2 + (y - y_0)^2 \qquad (5)$$

It implies that the accumulator space is three-dimensional (for three unknown parameters x_0, y_0 and r) and defines a locus of points (x, y) centered on an origin (x_0, y_0) with radius r. Points corresponding to x_0, y_0 and r, which has more votes, are considered to be a circle with center (x_0, y_0) and radius r.

The Euclidean distances between the cells are found in addition. They are necessary by obtaining of better information about cross location of the different cells in regard to their micro-moving. This step can be applied by observation of the cells in the time.

III.   EXPERIMENTAL RESULTS

The experiments are made in MATLAB 7.14 environment by using IMAGE PROCESSING TOOLBOX. For the investigations are used 20 fluoroscopic microscopy images with size 235x170 pixels. In Fig. 2 is presented the original image and in Fig. 3 is shown its enhancement modification after homomorphic filtration.
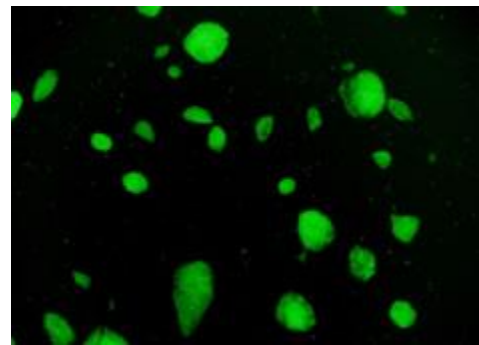

Fig. 2. Original microscopy image

The graphical presentation of background surface and effect from correction of illumination are shown in Fig.4 and Fig.5 respectively. The eligible embryo cells are colored in purple.
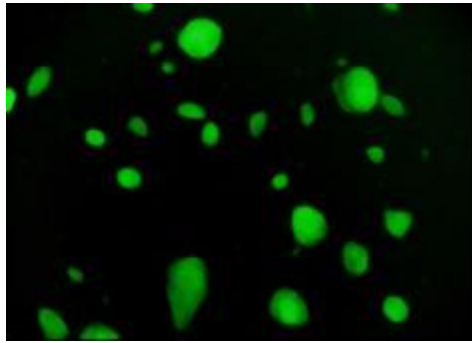

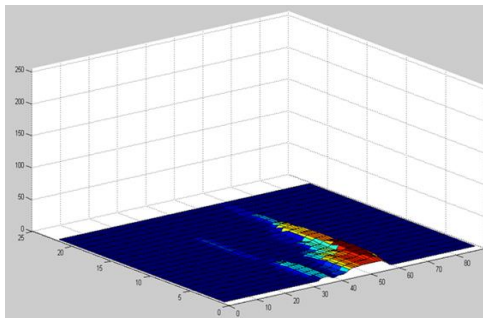Fig. 3. Microscopy image after homomorphic filtration

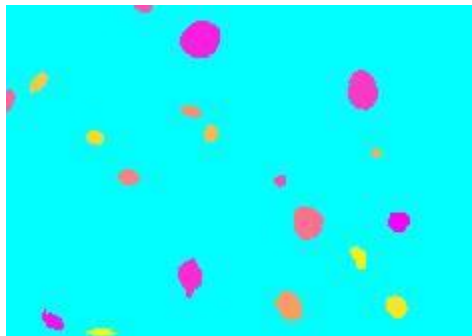

Fig. 4. Graphical presentation of background surface

Fig. 5. Microscopy image after illumination correction

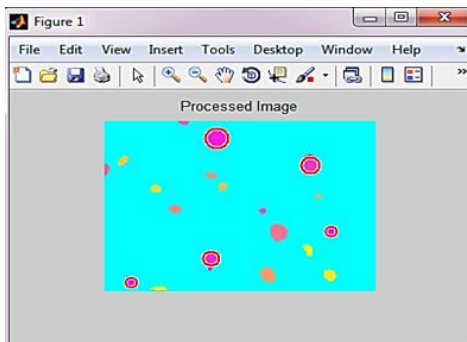The embryo cells are detected and located after application of Canny edge operator and Circle Hough transform. The obtained result is shown in Fig.6.

Fig. 6. Detected and located embryos cells

In Table 1 are given the obtained results for the radii and coordinates of the detected cells on the base of Circle Hough transform. The results are evaluated by calculating the measures precision, recall and F-measure [1].The measures were calculated for each image and, then, averaged over all images. The obtained results are given in Table 2.

TABLE I
EXPERIMENTAL RESULTS FOR DETECTED CELLS

Number of cells	X co-ordinate	Y co-ordinate	Radius R
1	97.79	18.0	10
2	178.55	45.14	8
3	92.94	138.47	7
4	196.70	11.29	5
5	24.18	161.69	5

TABLE II
EXPERIMENTAL RESULTS FOR DETECTED CELLS IN %

Precision	Recall	F-measure
97.67	95.23	96.43

The Euclidean distances between the cells are calculated as next. This information is very important for cross location of the different cells in regard to their micro-moving. The connections between the cells are presented in Fig.7 and the calculated Euclidean distances are given in Table 3.
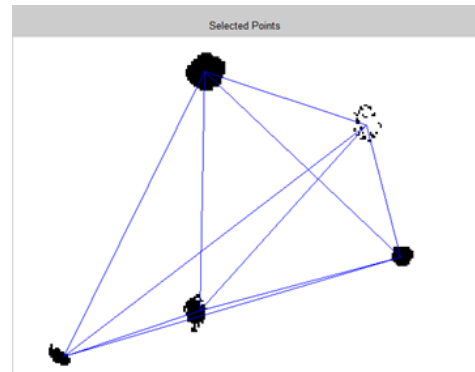
Fig. 7. Connections between the cells

TABLE III
EXPERIMENTAL RESULTS FOR EUCLIDEAN DISTANCES BETWEEN DETECTED CELLS

Number of cells	1	2	3	4	5
1	0	86.33	136.51	121.01	160.55
2	86.33	0	69.12	126.06	192.60
3	136.51	69.12	0	104.54	177.20
4	121.01	126.06	104.54	0	72.73
5	160.55	192.60	177.20	72.73	0

IV. CONCLUSION

In this paper is proposed an effective approach for automatic embryonic stem cell detection and determination of

their size and location in microscopy images. It consists noise reduction based on homomorphic filter and correction of illumination of the fluoroscopic microscopy images. Then an automatic method to detection and location of the embryo is applied. It is based on Hough Transform to approximate the embryo as a circle. Experimental results showed that the proposed method can detect the position of the eligible embryo accurately with average precision, recall and F-measure of 97.67%, 95.23% and 96.43%, respectively, which is satisfactory. Moreover, with the automatic detection, we can eliminate the subjectivity because, unlike in manual detection, is guaranteed that the same criteria are always used to detect cells. In addition, the method could be used in others applications.

Because of the image quality, which depends on a microscope type, some cells can be detected very difficult. These images have a strong noise and the cells are very small to be identified. Better results can be obtained by implementation of more effective methods for filtration. Our future works will be focused in applying of wavelet transformation for more effectiveness of filtration in cases with strong noise.

REFERENCES

[1] G.Faustino, M. Gattass, P. Carvalho, S. Rehen and C. de Lucena, *Automatic Embryonic Stem Cells Detection and Counting in Fluorescence Microscopy Images*, Monografias em Ciência da Computação, ISSN 0103-9741, No. 04, 2009.

[2] S. Sergent-Tanguy, C. Chagneau, I. Neveu, and P. Naveilhan, "Fluorescent activated cell sorting (facs): a rapid and reliable method to estimate the number of neurons in a mixed population", Journal of Neuroscience Methods, vol.129(1), pp.73 – 79, 2003.

[3] J. Logan, K. Edwards, and N. Saunders. *Real-Time PCR: Current Technology and Applications*. Caister Academic Press, 2009.

[4] F. Ambriz-Colín, M. Torres-Cisneros, J. G. Avina-Cervantes, J. E. Saavedra-Martinez, O. Debeir, and J. J. Sanchez-Mondragon. "Detection of biological cells in phase-contrast microscopy images". In MICAI, pages 68 – 77, 2006.

[5] C. Lucena, G. Faustino, M. Gatti and M. Gattass M. "A 3d multi-scale agent-based stem cell self- organization". SEAS, pp. 37 – 48, 2008.

[6] C.-N. Fan and F.-Y. Zhang, "Homomorphic filtering based illumination normalization method for face recognition," Pattern Recognition Letters, vol. 32, pp. 1468-1479, 2011.

[7] A. Holger G, "Butterworth equations for homomorphic filtering of images," Computers in Biology and Medicine, vol. 28, pp. 169-181, 1998.

[8] K. Delac, et al., "Sub-image homomorphic filtering technique for improving facial identification under difficult illumination conditions," Proc. 13th International Conference on Systems, Signals and Image Processing, pp. 95-98, 2006.

[9] R. Duda, P. Hart. "Use of the Hough transform to detect lines and curves in pictures", Communications of the ACM, Vol.15 (1), pp.11–15, 1972.

[10] S. Pei, J. Horng, "Circular detection based on Hough transform", Pattern Recognition Letters, vol.16, pp. 615-625, 1995.

POSTER SESSION
ELECTRONICS

Desing of On-Route Charging Infrastructure for EV

Nikolay Dimitrov Madzharov [1]

Abstract - **Design and development of 30kW Inductive Power Transfer system for on-route charging of Electric Vehicles is described in the present paper. There are several significant factors that make design of IPT more challenging. Its main purpose is a representation an alternative method of charge. At the same time, provides an opportunity for quick charging of electric vehicle batteries. Modes of charge, which are part of functional properties of the system during motion of electric vehicle. The basic units of the Charging converter are described, the circumstances of their design, advantages and disadvantages.**

Keywords - **IPT, coil, ON-Route, matchingcircuit, air gap.**

I.INTRODUCTION

On route charging of EVs is considered as a dynamic charging mode, with the car being charged while moving along the road. Feasibility of this technology is a big challenge in terms of effective EV integration in urban road system, but thanks to this technology the capacity and the weight of battery packs can be reduce potentially up to 60%. As a result the car is lighter and with improved driving range, so dynamic solution may become more cost effective. Because of the vehicle's motion, transfer of energy from the charging road spot to the EV's batteries can be only contactless and inductive principles are preferable.

II. STATE OF THE ARTS

Today's interest is now focused on the dynamic charging because it allows reducing the volume of the battery potentially up to 60%, and therefore the cost of the electric vehicle. Transmitting part is fixed in the roadway in the direction of the movement, and receiving part is in the EV.

The scale of power supply area is related to the charging speed, recharged volume of battery energy and running speed of vehicle and so on. Supposing it takes 30 minutes to make over 80% power for battery's energy, the length of 6 km in charging area is needed for an electric vehicle with speed of 60km/h to supply dynamically 30% of the battery capacity. If a single overall style primary track is designed for this long charging area, the energy loss is huge, with critical power requirements for every single electrical energy transducer. To avoid these problems, energy emission unit of primary side can take the form of coil array and segmented track – fig.1. The whole charging lane is segmented into several charging section and each section owns independent primary underground and side track or coil arrays. This is not only good for charging energy measurement but also for reducing power level of each power supply device.

[1] Nikolay Madzharov is with the Faculty of Electrical Engineering and Electronics at Technical University of Gabrovo, H.Dimitar str. 4, 5300 Gabrovo, Bulgaria, madjarov@tugab.bg

Moreover, the copper mass is reduced enormously, as the system only supplies power for these charging sections in which electric vehicles are being charged. What's more, an intelligent admittance mechanism could be built to insure that every EV in the system is identifiable and the system is never overloaded [1-3].

a) Coil array   b) Segmented track
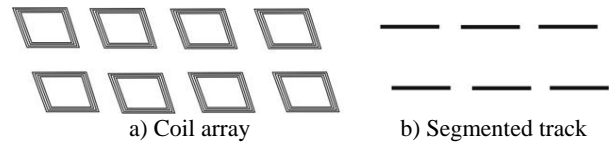
Fig. 1. The layout of energy emission unit

Optimal design should be achieved for the structure and the layout of energy emission unit, as the transmission efficiency of whole charging system, the first investment cost, operating cost and operating management are directly influenced by these design aspects.

In the on route charging system, the pick-up unit should get as close as possible to the energy emission unit, which can insure the energy pick-up is adequate and the transmission efficiency of system is high. However, in order to maintain the constant close distance, the pick-up unit should adjust its position automatically, including vertical position and horizontal position, to adapt road condition, position variation of energy emission unit and vehicle. The pick-up unit should be able to detect the position of energy emission unit in real-time, and adjust its position automatically to insure the energy pick-up reliably and safely [1-7].

Configurations of on route inductive power transfer.

A. Long Wire Loop

The long wire loop configuration is presented in Fig. 2. The roadbed loop is long compared to the vehicles and pick-up coils. Due to the short gap between long loops (assumed to be 1m), each car charges at least 95% of the time. Each large loop operates whenever a car is near, and all loops operate almost all the time with steady traffic. Reactive magnetization power is almost two orders of magnitude higher than the real-power output.
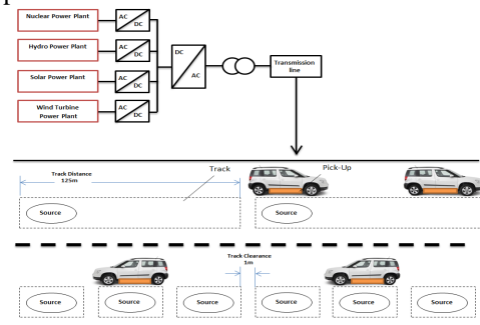

Fig. 2. Long wire loop configuration of EV on route charging.

B. Sectioned Wire Loop (Car-size Section)

The sectioned loop configuration is shown in Fig. 3a). The length of the loop is about the size of an EV. The gaps between the loops are assumed to be smaller (0.2m). The source loop in this case supplies only one car at a time, and the loop operates only when a car interacts with it. The coupling coefficient is larger than the long wire loop because the area that does not contribute directly to magnetic power transfer is reduced.
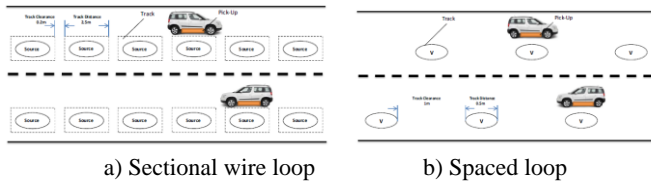

a) Sectional wire loop   b) Spaced loop
Fig. 3. Sectioned loop configuration

C. Spaced Loop (Small Loop)

The spaced loop configuration is shown in Fig. 3b). The loop (0.5 m) is much smaller than a car. Hence, the interval during a car interacts with the road coil is much smaller than for the other loop configurations. It may be possible to operate more than one loop per car, or to have a pickup coil larger than the source loop. The gaps between the loops are large (1m). However, the flux is more confined within the air gap, so external stray fields are substantially lower.

Numerous international teams [2,4,6] are currently working on the development and improvement of technology for dynamic EV charge and in particular on the efficiency of energy transfer during the movement and the possibility of transmission of more energy in a shortest distance of EV movement. This activity can be divided into two main groups - market products produced and developed by firms and companies and-experimental development of university teams.

Table 1 shows the technical parameters of the leading companies' achievements in this field. The main conclusion that arises is that the average efficiency achieved with this method is between 85 and 90%.

Table I. Leading companies achievements in the field of contactless charging

Company	Application and rated power	Type IPT	Air gap , mm	Web site
Conductix Wampfler	Industrial – $10^1 \div 10^2 kW$ Car, Bus	Static and dynamic	$10^1 \div 10^2$	conductix.com
Primove	Railway Transport – 400kW Bus – 200kW EV – 10kW	Static and dynamic	$200 \div 300$	primove.bombar-dier.com
OLEV Technologies	Bus– 100kW, 20kHz,85% eff.	Static and dynamic	200	olevtech.com

As a conclusion, it can be underline that companies' practical solutions are currently preferably oriented to the big public vehicles, like buses and trains. Dynamic on route charging solution for mini-buses and passenger cars are still subject of development and investigation. For this target group of vehicles is very important that the new developed technical products must be identical for static and dynamic charging mode and interoperable..

III. DESIGN OF ELECTRONIC MODULES FOR ON ROUTE CHARGING

The designed mock-up of charging station contains electrical and mechanical parts. The main components of electrical part are: 3 phase AC/DC rectifier, input C filter and five dual IGBT modules that compose four full bridge IGBT inverter circuits; control circuit and distribution board. The mechanical parts are heat sink and supporting sheet metal construction. In [6] detailed analysis of electromagnetic processes is made and on this basis electrical parameters of each electronic element, passive and active, are selected.

Electrical circuit of investigated modules, inverter and compensation capacitors and transformers, for on route charging mode are shown in figure 4.
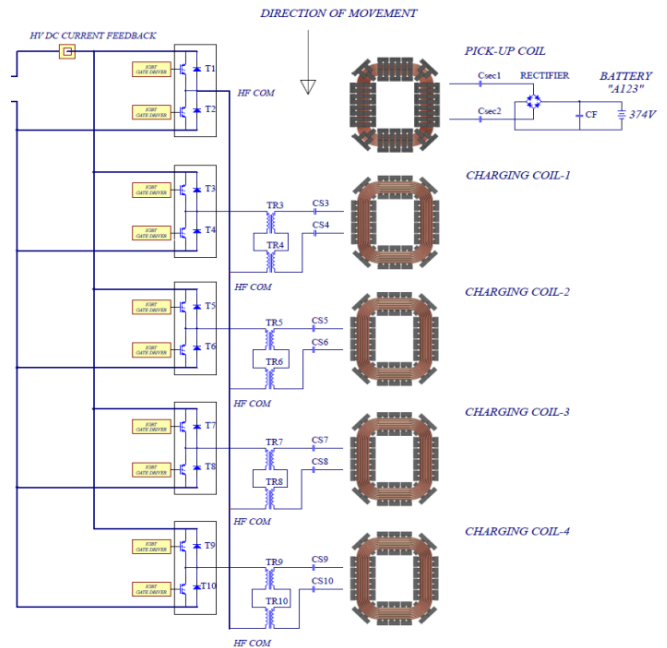
Fig. 4. Electrical circuit of investigated modules

The inductive power transfer (IPT) device is inherently a transformer, which transfers electrical energy from the primary to the secondary without direct contact between the two components. Due to inevitable air gap, parameters of such transformer are lower than for a normal transformer and the main challenge of the development was to make a system with good efficiency low sensitivity to the load parameters and air gap variation. Compared to charging system without IPT, the developed system (with IPT) has much higher reactive power and therefore is more sensitive to parameter variation. It has more reactive components as well, which can result in additional partial resonances in the range of frequency used for the generator scanning [2,4,6].The IPT primary and secondary parameters (figure 4) are shown in table 2.

Table II. IPT parameters

Element	CS3÷CS10	Csec 1,2	CF	T1÷T10
Value	2.4μF	2.4μF	550μF	200A/1200V
Transformers TR3,4÷TR9,10	Connection of primary windings - series Connection of secondary windings - series Trafo ratio - 9:5 Primary current : I1 = 60A/20kHz Secondary current : I2 = 120A/20kHz Trafo core : 4 xE100/60/28, material 3C90			
IPT primary	Weight ≈ 33kg Dimensions : 812x712x100mm Turns number: 7 LITZ cross section: 60mm^2			
IPT secondary	Weight: 25kg Dimensions: 760 x660x52mm Turns number: 7 LITZ cross section: 30mm^2			

One of the most important features for effective transfer of energy is magnetic coupling factor K. This is why influence of the misalignment between coils in direction X,Y,Z on the coupling factor has been investigated. For specified dimensions of the coils, 800mm/700mm/90mm (including winding aluminium shielding construction) the coupling factor K>0,25-0,3 is preferable. The main outputs of this investigation is that to realize on route charging is necessary to have information of misalignment between primary and secondary coils in direction X (in direction where the car is moving). The above results show that "X" misalignment up to 20 cm can be used as a threshold value for switching ON and OFF the transmitting windings - fig.5.
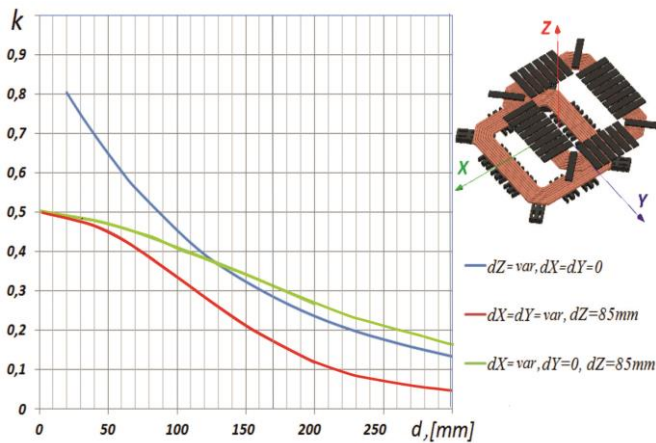
K > 0.25 - 0.3 optimal value

Fig. 5. Coupling factor K Vs. misalignment in X,Y,Z

The next step of testing and measurement of the whole power transfer electronic system with real battery load was optimization of charging station operating mode. The final outputs are presented in table 3.

Some real measurements are presented in figure 6. From these time charts is visible (see IPT primary current, IPT secondary current, IGBT voltage) that charging station operating mode is totally acceptable from electrical and safety point of view.

Table III. Electrical parameters of primary and secondary IPT coils

LOAD	Battery − A123 system
IPT ratio	$W_1 : W_2 = 7:7$
Air gap	70 ÷ 90mm
Frequency	f = 18 ÷ 22 kHz
IGBT current	I_{BRIDGE} = 60A rms @ 20kHz
Secondary matching trafos voltage TR3,4÷TR9,10	$U2_{TR3} + U2_{TR4} ≈ U_{DC} : 1.8 ≈ 300$ V
Primary IPT capacitor voltage CS3 (4,5, etc.)	U_{CS} = 600V rms @ 20kHz
IPT primary current	$I_{TX\ coil} = I_{CS} = I2_{TR} = 1.8 \times I1_{TRprim} = 110A$
IPT primary voltage	U_{TX} = 1200V rms @ 20kHz
IPT secondary current	$I_{RX\ coil} = I_{LOAD}$ = 75A rms
IPT secondary voltage	U_{TX} = 1100V rms @ 20kHz
Secondary IPT capacitor voltage Csec	U_{sec} = 550V rms @ 20kHz
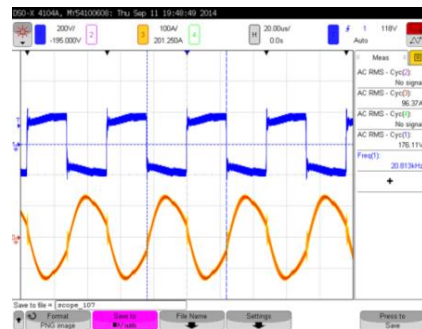
IPT primary Current (orange) 96,37 A and voltage on the secondary winding of matching trafo (blue) 176,11V at power 28.5 kW and frequency 20.81 kHz. No misalignment

a)

IPT primary current (orange) 188.32A; Matching capacitors voltage CS3-CS10 (blue) 607.3V; IPT secondary current (green) 51.42A; IGBT voltage (pink) 500V at misalignment dx=150mm, power 18.7 kW, frequency 20.54 kHz

b)

Fig. 6. Electrical tests - time charts

The optimized operating modes of all elements (currents and voltages) are in accordance with their catalogue electrical data.

IV. EVALUATION OF HFIPT LOSSES

Table 4 presents the overall IPT ferrite and Litz wire losses, which are acceptable in terms of efficiency and levels of losses. On these bases could be calculated optimal operating temperature, cooling requirements, taking account of environment temperature. Relation between volume and area of the primary and secondary coils allows operating mode of IPT module without additional cooling.

Table IV. Losses in IPT core and Litz wire

IPT		Air gap 100mm	
Output power/Output current		29,7kW / 90A	
Frequency		20kHz	
Winding	Loses	P_{loss} ,[W]	Plosses/Pout, %
Transmitting (Tx)	ferrite core	≈ 160W	0,54%
	Litz wire	$I^2R = 120^2 \cdot 0,047 = 677W$	2,28%
	Volume and surface area		$363dm^3 / 133dm^2$
Receiving (Rx)	Ferrite core	≈ 150W	0,51%
	Litz wire	$I^2R = 80^2 \cdot 0,085 = 544W$	1,83%
	Volume and surface area		$265dm^3 / 120dm^2$
Total		≈1532W	5,16%

To confirm that revised electrical parameters of electrical components guarantee optimal operating mode of each designed module, additional temperature test are implemented. On the figure 7 are shown temperature distributions of IPT module. The test condition during the temperature tests are - 3 min at 90 A and 17 min at 60 A, power 22kW.

IPT primary-max.temp. 44.9°C
a)

IPT secondary-max. temp. 40.1°C
b)
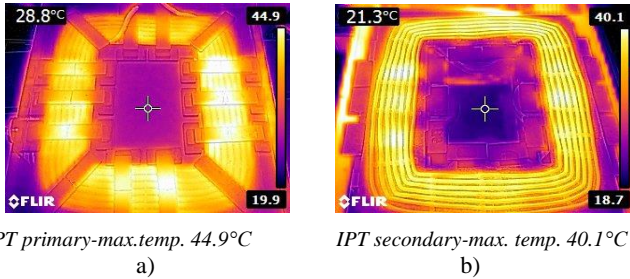
Fig. 7. IPT temperature tests

The measured maximum temperatures (in red) of the most powered modules at end of test (after 20 minutes) are minimum three times less than permissible catalogue data.

V. DESIGN OF ON ROUTE CHARGING ZONE

In Figure 8 is presented a draft of on route Charging zone with four primary coils, two guidance and one centric lines, that will help the driver to guide the vehicles in the right way with minimum Y misalignment. In front of the first coil, between next three others and after the last one are shown sensors for indication of permissible X misalignment. The green marked sensors are activated when secondary coil covers the corresponding primary coil with dX=200 mm misalignment "before" in the driving direction. Accordingly the corresponding primary coil is switched ON. Similarly, all red sensors are activated when misalignment "after" is again

200mm and they will switch OFF the corresponding primary coil.
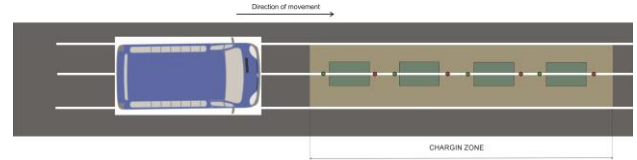
Fig. 8. On route charging zone

Taking into account this principle of activity, it is proposed to choose so called proximity sensors.Sensors will be switched in accordance with 40 mm distance between secondary coil and ground surface, where they have to be built.

VI. DISCUSSION AND CONCLUSIONS

✓ Infrastructure on route charging zone is developed.
✓ For on route charging, additional investigations of misalignments between primary and secondary coils, especially in X direction (driving direction), were made.
✓ For that reason, how the misalignment values influence the permissible magnetic coupling factor has been analyzed and final results have been presented.
✓ The currents and voltages of all elements are in accordance with their catalogue electrical data and results of implemented temperature tests have proved that.
✓ The analysis of on-route energy distribution and misalignment results has been taken into account to define the distance between primary coils built in a charging zone.
✓ It was proved that to realize effective and reliable HF generator operating mode, which correspond to permissible misalignment, the switching sensors have to be used.

VII. REFERENCES

[1] AldhaherSamer , Patrick C. K. Luk, AkramBati, " Wireless Power Transfer Using Class E Inverter with Saturable DC-Feed Inductor", IEEE Transactions on Industry Applications, Volume 50 , Issue 4, 2014, Pages 2710 - 2718

[2] AlingerDustin James, "SYSTEM ANALYSIS AND DESIGN FOR THE RESONANT INDUCTIVE NEAR-FIELD GENERATION SYSTEM ", Master of Science Thesis, University of Maryland, College Park, 2013

[3] "Arash Dabirzadeh, ""RF COIL DESIGN FOR MULTI-FREQUENCY MAGNETIC RESONANCE IMAGING AND SPECTROSCOPY"", TexasA&MUniversity,2008

[4] Asheer Sara, Amna Al-Marwani, Tamer Khattab, "Contactless power and data transfer for electric vehicle applications", International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering Vol. 2, Issue 7, July 2013, ISSN : 2320 – 3765

[5] Chopra Swagat, "Contactless Power Transfer for Electric Vehicle Charging Application", Master of Science Thesis, Delft University of Technology, 2011.

[6] Senjuti Shawon, "Design And Optimization Of Efficient Wireless Power Transfer Links For Implantable Biotelemetry Systems", Master of Science Thesis, The University of Western Ontario, Canada,2013

[7]Tomohiro Yamanaka, Yasuyoshi Kaneko, Shigeru Abe, Tomio Yasuda, "10 kW Contactless Power Transfer System for Rapid Charger of Electric Vehicle", Los Angeles, California : s.n., 2012, Vols. International Battery, Hybrid and Fuel Cell Electric Vehicle Symposium.

Hysteresis Controlled Switching-Mode Amplifier for LTE Applications

Tihomir Brusev

Abstract – **Switching-mode amplifiers are used in power supply circuits, which deliver the energy to the transmitter's power amplifier (PA) for fourth generation Long-Term Evolution (4G LTE) applications. They could be part of series or parallel hybrid envelope amplifier (EA) architectures in combination respectively with linear amplifier. In this case switching-mode amplifier supplies about 80% of necessary power to the PA. Therefore efficiency of this stage strongly affects the overall efficiency of the envelope tracking power amplifier (ET PA) system. Switching-mode converters are used also in the interleaved multiphase structures, which perform the function of EA in 4G LTE transmitter. In this paper are presented investigation results of hysteresis controlled switching-mode buck converter designed on CMOS 0.35 μm process.**

Keywords – **switching-mode converter, CMOS 0.35 μm process, Cadence, LTE wireless communication standard.**

I. INTRODUCTION

In the new wireless communications 4G LTE standard high data rate can be transferred [1], [2], [3]. This gives freedom to the costumers to watch TV programs in their mobile phones, to have fast web browsing, etc. All those advantages compare to the GSM communication standard are connected to a drawback that the battery has to be recharged more often. Efficiency improvement of the separate building blocks in the transmitter is the key to save batteries .energy of the mobile communication devices.

Increasing the system run-time of battery powered portable electronic devices is challenging task. Stringent linearity requirements for PA, which is the most energy consuming block in the transmitter, should be fulfilled. On the other hand linear PAs are not high efficient circuits. Their efficiency is unacceptable small especially when they work with low power input signal. PA of the transmitter in LTE applications have to work with high peak to average power ratio (PAPR), which lead to decreasing of efficiency values of this stage [4]. In most of the time PA have to work in the back-off mode of operation.

Envelope elimination and restoration (EER) and envelope tracking (ET) are efficiency enchantment techniques for PA [4], [5]. The both method used modulated supply voltages which are delivered to the RF PAs transistor. ET has some advantages compare to EER technique, such as lower bandwidth required to envelope amplifier and higher output gain at low power of PA [4], [6]. Those features help higher

[1]Tihomir Brusev is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: brusev@ecad.tu-sofia.bg.

power-added-efficiency (PAE) of the whole system to be achieved when ET technique is used.

Envelope amplifier in ET technique is used to supply dynamically changeable voltages to the drain or collector of PA's output transistors as a function of PA input signals. The main goal is RF PA's transistor to operate near to saturation mode in most of the working time [7]. Overall efficiency of the ET PA system is defined by formula [8] (1):

$$\eta_{ET\,PA} = \eta_{EA}.\eta_{PA}, \qquad (1)$$

where η_{EA} is the efficiency of the envelope amplifier; η_{PA} is respectively the efficiency of the PA.

Hybrid combination of parallel or series connected linear and switching amplifier are usually used for envelope amplifiers architectures. The function of switching amplifier is performed by switching-mode dc-dc converter. This stage delivers about 80% of the total output to the PA [6]. In some applications interleaved multiphase switching-mode buck dc-dc converters are used as envelope amplifiers [1]. The common thing in the all architectures of EAs is that switching-mode converters play a major role in determining the overall efficiency of the ET PA system.

In Section II are discussed different power supply circuit architectures for mobile communication applications, where switching-mode converters are used. Hysteresis controlled switching-mode amplifier is designed with Cadence on CMOS 0.35 µm process. Power losses in output MOS transistors are considered and evaluated. Efficiency of the whole converter system as function of load is investigated. The received results are presented in Section III.

II. SWITCHING-MODE CONVERTER ARCHITECTURES USED IN POWER SUPPLY CIRCUITS FOR COMMUNICATION APPLICATIONS

A. PWM controlled buck dc-dc converter

In the earlier communication standards like GSM, PWM controlled switching-mode dc-dc converters are used to deliver supply voltage to the PA of transmitter. These types of circuits are suitable for constant envelope signals, when stable dc output voltage of the converter should be ensured. In the 4G LTE standard envelope amplifiers, which have to supply voltage to the PA, should have fast tracking speed because envelope frequency is increased. The disadvantage of the PWM controlled switching converters is that they are low bandwidth circuits. The switching frequency f_s of the dc-dc converter have to be about ten times higher than the bandwidth of the LTE signal [1]. In Fig. 1 is shown block

diagram of switching-mode buck dc-dc converter system with Pulse-Width Modulation (PWM) control.
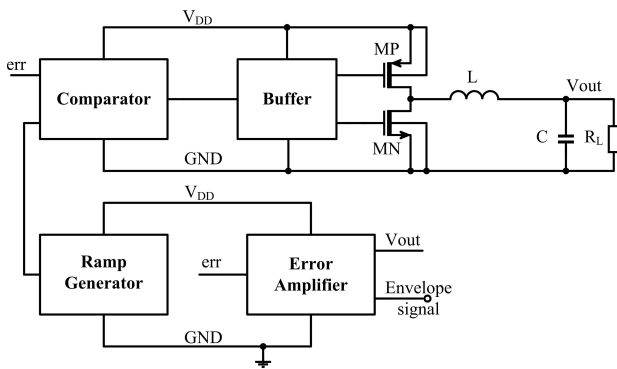

Fig. 1. Switching-mode buck dc-dc converter with PWM control system.

If PWM controlled switching-mode regulator is used as envelope amplifier in the transmitter, f_s have to be very high in order to be covered all the LTE bandwidths. Power losses in the dc-dc converter will be increased, because they are proportional to the f_s, decreasing overall efficiency in the system.

B. Hybrid envelope amplifier architectures

Hybrid combination between switching-mode converter and linear amplifier is an alternative for envelope amplifier architectures [2]. Switching-mode amplifier is used to deliver low frequency and dc voltages, while linear amplifier has to supply high frequency voltages to PA. In Fig. 2 is shown series combined hybrid architecture between switching-mode buck dc-dc converter and linear amplifier [9].
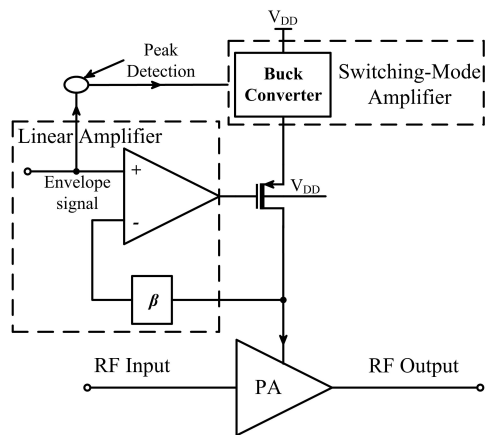
Fig. 2. Series combined switching and linear regulator topology of envelope amplifier.

In the LTE communication standard the transmitted signals have non-constant envelope waveform. Any noises coming from envelope amplifiers will be mixed and amplified with the input signal of PA, distorting output transmitted signal.

Linear amplifier in the hybrid architectures should act as filter for output ripples generated from buck dc-dc converter. The disadvantage is that for high switching frequency f_s linear amplifier has low power supply rejection ratio (PSRR) [2]. Also, the entire power deliver to the PA goes through power MOS transistor of the switching-mode amplifier. Therefore switching power losses will be increased especially for high f_s.

In Fig. 3 is illustrated block circuit of parallel combined hybrid architecture between switching-mode buck dc-dc converter and linear amplifier.
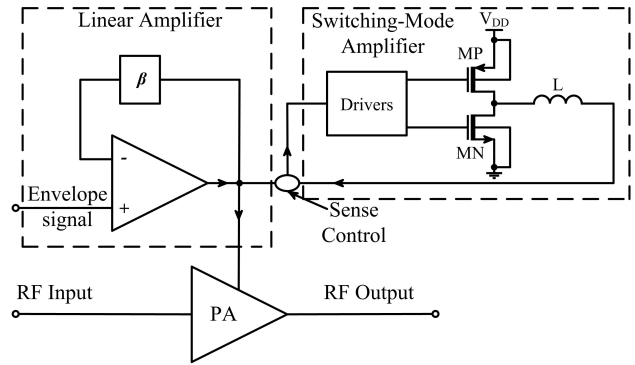
Fig. 3. Parallel combined switching and linear regulator topology of envelope amplifier.

This structures use the advantages of the both stages: high efficiency of the switching-mode converter at low frequency and wideband performance of linear amplifier. The switching-mode amplifier is used to deliver average power to PA. Linear amplifier has to filter the current ripples generated from switching type of converter and also to deliver the rest part of the power to PA, when switching regulator cannot respond quickly. The disadvantage of linear amplifier is that those circuits have low efficiency especially in the case when difference between output and input voltage is big. The parallel combined hybrid envelope amplifier architecture is widespread used for envelope amplifiers in LTE applications, because most of the envelope powers have low frequency [2], [4]. In this case high efficient switching-mode amplifier delivers most of the energy to PA, increasing overall efficiency of the system.

C. Multiphase switching-mode amplifier architectures

Multiphase envelope amplifier architecture is proposed in [1]. In this power supply circuit's topology less efficient linear stage is completely removed. Therefore overall efficiency of the envelope amplifier could be increased. The authors proposed interleaved two-phase switching converter architecture. In Fig. 4 is shown first stage of the proposed synchronized adaptive voltage tracking (SAVT) controller of the interleaved buck converter.

The second stage is phase shifted on 180°. The first advantage of the multiphase structures is that current and voltage ripples are decreased compare to standard buck converter.
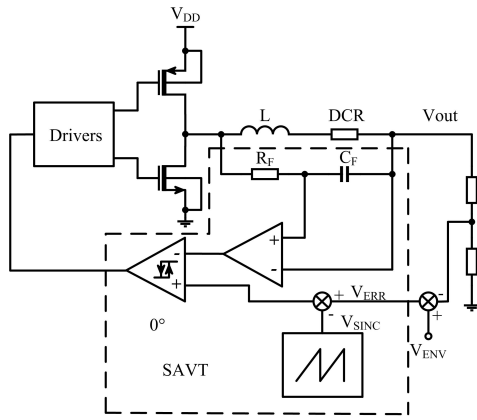
Fig. 4. Synchronized Adaptive Voltage Tracking (SAVT) control of the two-phased interleaved .converter proposed in [1].

The switching frequency f_s and the phase of the two sub-converters are synchronized in the implemented hysteresis control. Thus the spectrum of the inevitable switching noises going to power amplifier could be predicted.

III. INVESTIGATION OF HYSTERESIS CONTROLLED SWITCHING-MODE AMPLIFIER

The most used envelope amplifiers architectures for LTE applications reported in the literature and drawbacks of PWM controlled switching-mode converters have been discussed in the previous section of this paper. In the hybrid envelope amplifiers hysteresis control is used for switching-mode amplifier. This control method allows increasing of switching converter's bandwidth up to switching frequency f_s [1]. Therefore using lower f_s, compare to the PWM control technique, the LTE bandwidths could be covered.

The hysteresis controlled switching-mode amplifier is designed on CMOS 0.35 µm process. The block diagram is presented on Fig. 5.
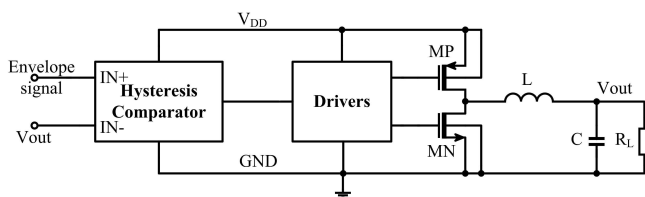
Fig. 5. Hysteresis controlled switching-mode amplifier.

Supply voltage V_{DD} of the hysteresis controlled switching-mode amplifier is chosen to be equal to 3.6 V, which a standard output voltage of lithium-ion battery. The value of the inductance of output filter inductor L is equal to 250 nH. The value of the capacitance of the output filter capacitor C is set to be 5 pF. This is equivalent value of the capacitance of parallel combination of the power transistor of linear amplifier and the filter inductor L of the parallel combined hybrid envelope amplifier architecture [6].

In Fig. 6 is presented the circuit of the comparator with hysteresis used in the investigated control. The generated output signal of this stage is applied to the drivers, which regulate the output transistors of the switching converter. The drivers provide short gap time when NMOS and PMOS transistors are both switched-off, preventing the short-circuit losses.
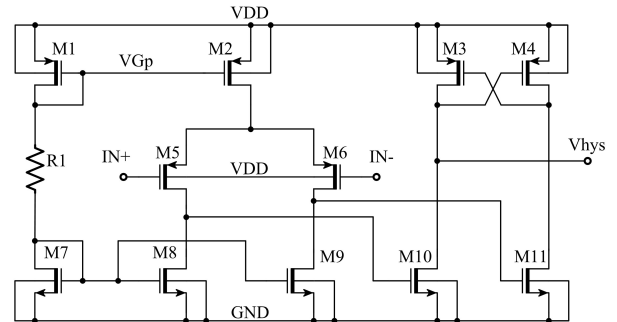
Fig. 6. Comparator with hysteresis.

The main focus of the investigation of the hysteresis controlled switching-mode amplifier is power losses in the output transistors MP and MN. They consume most of the energy of the converter determining to large extent efficiency his efficiency. Power losses in the MOS transistors are equal to [10]:

$$P_{loss,MOS} = P_{cond} + P_{sw}, \qquad (2)$$

where P_{cond} and P_{sw} are respectively conduction and switching losses in the output transistors. On the other hand P_{cond} are given by [10]:

$$P_{cond} = I_{source}^2 . r_{on,p} + I_{sink}^2 . r_{on,n}, \qquad (3)$$

where I_{source} and I_{sink} are respectively sourcing and sinking current of the MOS transistor, while $r_{on,p}$ and $r_{on,n}$ are on-resistance of PMOS and NMOS transistors. The switching losses are equal to [2]:

$$P_{sw} = f_s . C_{tot} . V_{DD}^2, \qquad (4)$$

where C_{tot} is the input total capacitance of the MOS transistor.

Power losses in PMOS and NMOS transistors of the designed converter are investigated as a function of the load R_L. The received simulation results are given in Table I.

The values of resistor R_L represent the current load of RF PA. The reported simulation results are received when the envelope input signal of comparator with hysteresis has a sinusoidal waveform. The frequency of this test signal is equal to 20 MHz. Thus fast changing LTE envelope could be emulated. The dc voltage level of the test signal is equal to 1.5 V, while the amplitude is 250 mV. The main goal of this analysis is to be illustrated the function of losses in the output transistors of hysteresis controlled switching-mode amplifier when output power is changed.

TABLE I
POWER LOSSES (OF PMOS AND NMOS TRANSISTORS) AND
EFFICIENCY AS A FUNCTION OF LOAD R_L

	R_L=10 [Ω]	R_L=15 [Ω]	R_L=20 [Ω]	R_L=25 [Ω]	R=30 [Ω]
P_{out} [mW]	83.6	95.35	113.1	144.2	217.3
P_{NMOS} [mW]	1.88	2.12	2.67	4	5.834
P_{PMOS} [mW]	38.23	45.24	58.7	90.44	198.3
Eff. [%]	50.34	58.04	62.06	63.86	64.58

In Fig. 7 are presented graphically power losses of PMOS and NMOS transistors as function of output power of the switching-mode amplifier.
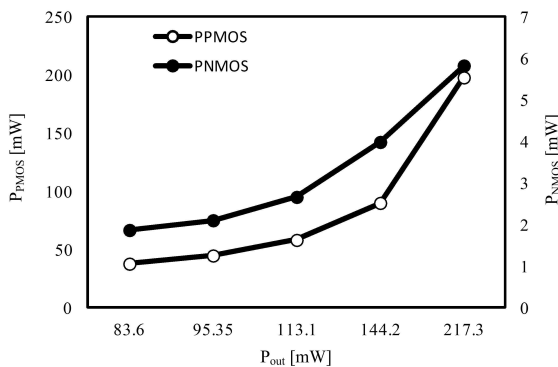
Fig. 7. Power losses of PMOS and NMOS transistors as function of output power of the switching-mode amplifier.

Simulated efficiency results of the switching-mode amplifier as a function of the load R_L are presented in Fig. 8.
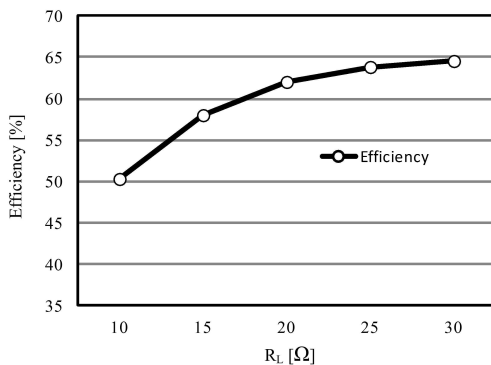
Fig. 8. Efficiency of the switching-mode amplifier as a function of the load R_L.

The values of the R_L are changed in the range between 10 Ω and 30 Ω. This range represents practical equivalent value of PA used as a load [10]. The maximum simulated efficiency of the hysteresis controlled switching-mode amplifier is 76.7 % at output power equal to 260 mW.

IV. CONCLUSION

In this paper are discussed different power supply circuit architectures suitable for portable communication devices, where switching-mode converters are used. The advantages of hysteresis instead of PWM control technique for LTE applications are considered. The output transistor's power losses of the hysteresis controlled switching-mode amplifier, designed on CMOS 0.35 µm process, are analyzed. Efficiency of the whole converter system as a function of the load is investigated.

ACKNOWLEDGEMENT

REFERENCES

[1] J. Sankman, M. K. Song, and D. Ma, "A 40-MHz 85.8%-Peak-Efficiency Switching-Converter-Only Dual-Phase Envelope Modulator for 2-W 10-MHz LTE Power Amplifier," *Symposium on VLSI Circuits Digest of Technical Papers*, June 2014.

[2] M. Hassan, L. Larson, V. Leung, and P. Asbeck, "A Combined Series-Parallel Hybrid Envelope Amplifier for Envelope Tracking Mobile Terminal RF Power Amplifier Applications," ," *IEEE Journal of Solid-State Circuits*, vol.47, no.5, pp.1185-1198, May 2012.

[3] www.3gpp.org/LTE

[4] Y. Li; Lopez, J.; Lie, D.Y.C.; Chen, K.; Wu, S.; Tzu-Yi Yang; Gin-Kou Ma, "Circuits and System Design of RF Polar Transmitters Using Envelope-Tracking and SiGe Power Amplifiers for Mobile WiMAX," *Circuits and Systems I: Regular Papers, IEEE Transactions on* , vol.58, no.5, pp.893-901, May 2011.

[5] S. Yoo, J. S. Walling, E. C. Woo, and D. J. Allstot, "A switched-capacitor power amplifier for EER/polar transmitters," in *Proc. IEEE Int. Solid-State Circuits Conf.*, San Francisco, CA, 2011, pp. 428–430.

[6] F. Wang, D. Kimball, D. Lie, P. Asbeck, L. Larson, "A Monolithic High-Efficiency 2.4-GHz 20-dBm SiGe BiCMOS Envelope-Tracking OFDM Power Amplifier," *IEEE Journal of Solid-State Circuits*, vol.42, no.6, pp.1271-1281, June 2007

[7] Tuffery, A.; Deltimple, N.; Leite, B.; Cathelin, P.; Knopik, V.; Kerherve, E., "A 27.5-dBm linear reconfigurable CMOS power amplifier for 3GPP LTE applications," *New Circuits and Systems Conference (NEWCAS), 2011 IEEE 9th International* , vol., no., pp.221,224, 26-29 June 2011.

[8] J. Lopez, Y. Li, J. D. Popp, D. Y. C. Lie, C. C. Chuang, K. Chen, S. Wu, T.-Y. Ying, and G.-K. Ma, "Design of highly efficient wideband RF polar transmitter using the envelope-tracking technique," *IEEE Journal of Solid-State Circuits*, vol. 44, no. 9, pp. 2276–2294, Sep. 2009.

[9] J. N. Kitchen, C. Chu, S. Kiaei, and B. Bakkaloglu, "Combined linear and Δ-modulated switch-mode PA supply modulator for polar transmitters," *IEEE Journal of Solid-State Circuits*, vol. 44, no. 2, pp. 404–413, Feb.2009.

[10] J. Ham, H. Jung, H. Kim, W. Lim, D. Heo and Y. Yang1, "A CMOS Envelope Tracking Power Amplifier for LTE Mobile Applications," *Journal of Semiconductor Technology and Science*, vo.14, no.2, pp. 235-245, April 2014.

Analysis and Design of a Resonant Reset GaN Forward Converter with Self-Driven Synchronous Rectifiers

Zoran Zivanovic[1] , Vladimir Smiljakovic[2] and Sinisa Jovanovic[3]

Abstract – **This paper presents the straightforward design of a resonant reset forward converter using a Gallium Nitride transistor with synchronous rectification. The prototype has been built and tested through lab measurements to verify the design. Design steps are described and well documented with measurement results.**

Keywords – **Efficiency, forward converter, GaN, resonant, self-driven, synchronous rectifier.**
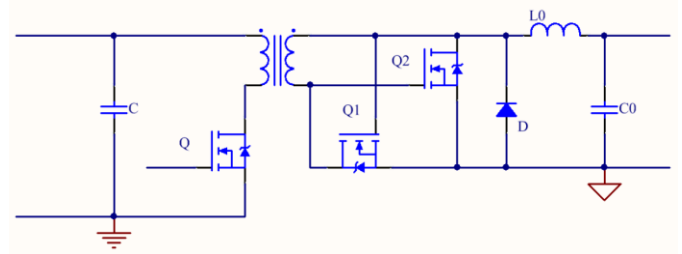
Fig. 1. Resonant reset forward converter

I. INTRODUCTION

A forward converter is a popular choice for low output voltage DC/DC converters, particularly for input voltage range of 18 to 36V or 36 to 72V. It has a single primary side switch and after each switching duty cycle, the magnetizing current in the transformer has to be reset to prevent core saturation. The active clamp approach is far the best in terms of optimal performance but not so simple. Traditional reset circuits are Third Winding, RCD Clamp, Zener Clamp, Active Clamp and Resonant Reset. As the output voltage is decreased, the efficiency of the forward converter becomes limited by the output rectifier loss. Although the Schottky diodes are widely used, synchronous rectification is needed to realize higher efficiency.

II. ABOUT THE CHOSEN TOPOLOGY

The chosen topology is the forward converter with resonant reset shown in Fig. 1. This converter differs from a basic design in three ways: it uses a resonant reset technique to reset the transformer, uses Gallium Nitride transistor Q instead a MOSFET and employs synchronous rectifiers Q_1 and Q_2 (low R_{dson} MOSFETS) on the secondary side. To keep the circuit simple and low cost, the synchronous rectifiers are self-driven, directly with the voltage from the transformer secondary. The drain of Q_1 is connected to the gate Q_2, and the drain of Q_2 is connected to the gate of Q_1. The gate-drive voltages must be high enough to ensure low FET resistance, but small enough to prevent the gate destruction. The gate-drive voltage for rectifier Q_2 is derived from the reset voltage so the resonant reset is not the most efficient solution especially at high input voltage when the dwell time is large. During the dwell time the body diodes of Q_1 and Q_2 with

relatively high voltage drops are conducting and the efficiency is reduced. This loss is minimized by connecting Schottky diode D in parallel with mosfet Q_2.

The resonant circuit consists of only parasitic elements and there is no need for additional external components. When the primary switch is turned off, a resonance is developed between magnetizing inductance of the transformer and the equivalent resonant capacitance of the circuit. The high voltage across the drain of the switch resets the transformer core. There is no need for a reset winding and diode.

Typical drain voltage of primary switch in resonant reset forward converter is given in Fig. 2.

Fig. 2. Resonant reset waveform

Resonant frequency is given by

$$f_{RES} = \frac{1}{2\pi\sqrt{L_M C_R}} \qquad (1)$$

where L_M is transformer primary magnetizing inductance and C_R is parasitic capacitance.

On the other hand the resonant half-sine wave voltage has a peak value given by

$$V_{RESET} = \frac{V_{INMIN} D_{MAX}}{L_M f_{SW}} \sqrt{\frac{L_M}{C_R}} \qquad (2)$$

The construction of the GaN transistor allows for inherent high speed operation so the switching time of 5 ns or less are possible. The GaN transistor is not avalanche rated at the breakdown of drain-source junction so it is necessary to insure

[1]Zoran Zivanovic is with the IMTEL KOMUNIKACIJE AD, Bul. Mihajla Pupina165b, 11070 Belgrade, Serbia, E-mail: zoki@insimtel.com.

[2]Vladimir Smiljakovic is with the IMTEL KOMUNIKACIJE AD, Bul. Mihajla Pupina165b, 11070 Belgrade, Serbia, E-mail: smiljac@insimtel.com.

[3]Sinisa Jovanovic is with the IMTEL KOMUNIKACIJE AD, Bul. Mihajla Pupina165b, 11070 Belgrade, Serbia, E-mail: siki@insimtel.com.

that a maximum drain-source voltage will remain within the desired limits. Available GaN transistors have a maximum gate voltage of +6V. Situation is complicated by the fact that full enhancement of the GaN is achieved at gate voltage higher then +4V. So it is advisable to have voltage regulation in the gate drive circuitry.

III. DESIGN AND ANALYSIS

The goal is to design the resonant reset forward converter using Gallium Nitride transistor and synchronous rectification. Achieving the efficiency as much as possible around 90% is the primary objective. We will choose the switching frequency to be around 320 kHz, which is a compromise between the efficiency and size. Knowing that, a good choice of a transformer and the inductor core is RM4, N49 material from TDK.

Design input parameters are given in Table I.

TABLE I
DESIGN INPUT PARAMETERS

		Min	Typ	Max	
Input voltage	V_{IN}	18	24	36	V
Output voltage	V_O		5		V
Output current	I_O	0.1	2		A
Output current limit	I_{OCL}		2.4		A
Full load efficiency	η		90		%
Switching frequency	f_{SW}		320		kHz

Starting from design input parameters we will now calculate basic parameters for the transformer and output inductor (Table II).

TABLE II
TRANSFORMER AND INDUCTOR INPUT PARAMETERS

		Max	Typ	Min	
Duty cycle	D	0.43	0.32	0.21	
Core cross sect. area			0.11		cm^2
Core effect. volume	V_E		0.29		cm^3
Prim. RMS current	I_{PRMS}	0.99	0.85	0.69	A
Sec. RMS current	I_{SRMS}	1.31	1.13	0.92	A
Flux density	B_{pp}		182		mT
Specific core losses	P_V		0.3		W/cm^3
Number of prim. turns	N_P		12		
Number of sec. turns	N_S		8		
Primary resistance	R_P		55		mΩ
Secondary resistance	R_S		29		mΩ
Output inductance	L		21		µH
Number of ind. turns	N		12		
Air gap required	g		0.1		mm
Winding resistance	R_L		27		mΩ

Now we can wind the transformer and output inductor prototype. We will use for the primary three parallel strands of 0.3mm enameled copper wire and for the secondary a bundle of seven 0.15mm wires twisted around each other, in order to

keep the AC resistance low as much as possible. For the output inductor we will use 3 parallel strands of 0.3mm enamelled copper wire.

Knowing RMS currents and specific core losses we can now calculate the losses in both magnetic components (Table III). Total transformer power loss at 24V input voltage is 164mW. This results in approximately 20°C rise above ambient temperature. The temperature rise on the inductor is 13°C. Satisfied with the results, the optimization of a transformer and inductor is not necessary.

TABLE III
TRANSFORMER AND INDUCTOR OPERATING PARAMETERS

		Max	Typ	Min	
Thermal resistance	R_{TH}		120		K/W
Core loss	P_{CORE}		87		mW
Primary loss	P_{PRI}	54	40	26	mW
Secondary loss	P_{SEC}	50	37	25	mW
Total transformer loss	P_{TOT}	191	164	138	mW
Inductor RMS current	I_{LRMS}		2		A
Inductor loss	P_{IND}		108		mW

For higher efficiency, the current sense resistor is biased, reducing the current sense amplitude, so it can be three times smaller. As a result we have smaller power loss. Power dissipated in current sense resistor is given in Table IV.

TABLE IV
CURRENT SENSE RESISTOR POWER LOSSES

		Max	Typ	Min	
Current sense resistance	R_{CS}		165		mΩ
CS resistance loss	P_{CS}	162	119	79	mW

Power switch losses can be expressed by the equation

$$P_{FET} = P_{COND} + P_{ON} + P_{OFF} + P_{QOSS} \qquad (3)$$

where

P_{COND} is the conduction loss given by

$$P_{COND} = I_{RMS}^2 R_{DS} \qquad (4)$$

P_{ON} is the turn-on switching loss given by

$$P_{ON} = \frac{V_{DS} I_P (Q_{GS} + Q_{GD}) f_{SW}}{2 I_G} \qquad (5)$$

P_{OFF} is the turn-off switching loss given by

$$P_{OFF} = \frac{V_{DS} I_P (Q_{GS} + Q_{GD}) f_{SW}}{2 I_G} \qquad (6)$$

and

P_{QOSS} is the capacitance charge loss given by

$$P_{QOSS} = \frac{Q_{OSS} V_{DS} f_{SW}}{2} \qquad (7)$$

For the primary switch Q we choose a Gallium Nitride transistor EPC2016. Using Eq. (3) to (7) we have calculate his operating parameters that are given in Table V.

TABLE V
PRIMARY GaN FET OPERATING PARAMETERS

EPC2016		Max	Typ	Min	
Drain-to source voltage	V_{DS}		100		V
ON resistance	R_{DS}		15		mΩ
Gate-to-drain charge	Q_{GD}		11		nC
Gate-to-source charge	Q_{GS}		5		nC
Output charge	Q_{OSS}		20		nC
Conduction loss	P_{COND}	23	17	11	mW
Switching loss	P_{SW}	67	90	134	mW
Total loss	P_{TOT}	85	103	143	mW

Peak stress on the forward and freewheel synchronous rectifier will be less than 30V, therefore we will use IRF7468 MOSFETS. Their operating parameters are given in Table VI.

TABLE VI
FORWARD AND FREEWHELL SYNCH MOSFET OPERATING PARAMETERS

IRF7468			Typ		
ON resistance	R_{DS}		12		mΩ
Q_1 RMS current	I_{Q1RMS}	1.31	1.13	0.92	A
Conduction loss	P_{CON}	31	23	15	mW
Q_2 RMS current	I_{Q2RMS}	1.39	1.39	1.39	A
Conduction loss	P_{CON}	35	35	35	mW

As a freewheel diode D we will use Schottky diode MBRS140 with operating parameters given in Table VII.

TABLE VII
FREEWHEEL DIODE OPERATING PARAMETERS

MBRS140			Typ		
Forward voltage drop	V_F		0.55		V
Dwell duty cycle	D_{DWELL}	0.05	0.16	0.27	
Peak current	I_{PEAK}	2	2	2	
Conduction loss	P_{CON}	55	176	297	mW

IV. REALIZATION

DC/DC converter was built on FR-4 substrate with 70μm copper. For optimal performance low side driver LM5114 is used with GaN FET. Also we have used simple gate drive circuit for synchronous rectifiers.

Using resistive load we have measured full load efficiency at various input voltages. The results are given in Table VIII.

TABLE VIII
EFFICIENCY

			Typ		
Input voltage	V_{IN}	18	24	36	V
Input current	I_{IN}	0.608	0.464	0.319	A
Input power	P_{IN}	10.944	11.136	11.484	W
Efficiency	η	91.37	89.8	87.08	%

The drain voltages of the primary power switch are recorded using Digital Phosphor Oscilloscope TDS5052, at full load and input voltages of 18, 24 and 36V are given in Figs. 3, 4 and 5 respectively.
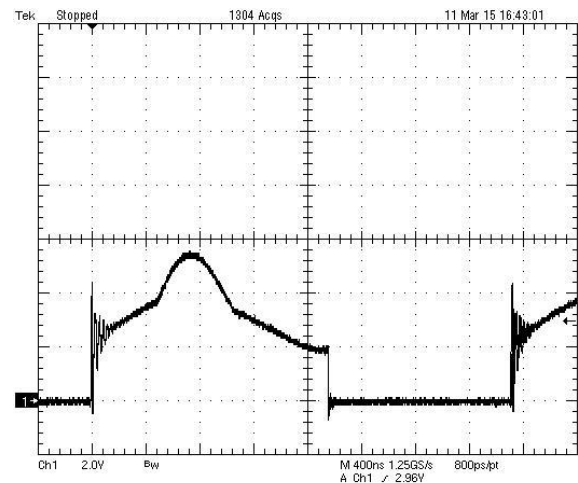
Fig. 3. Drain voltage waveform at 18V
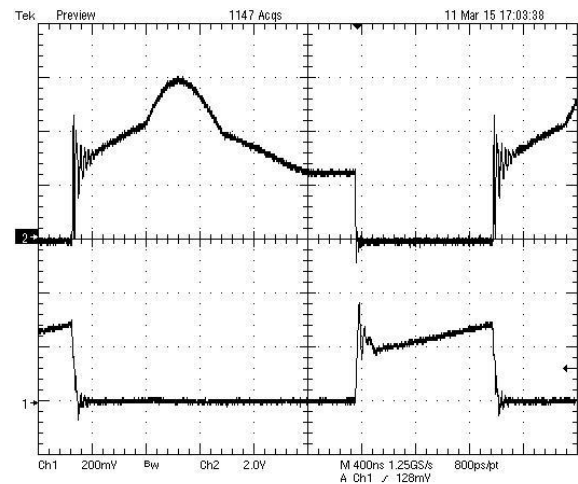
Fig. 4. Drain voltage waveform at 24V
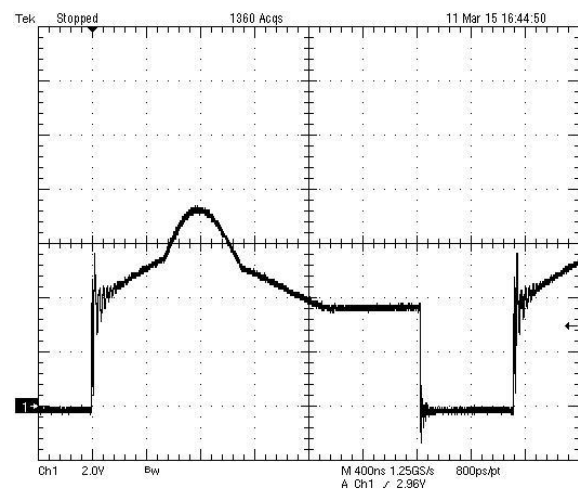
Fig. 5. Drain voltage waveform at 36V

The turn-on and turn-off waveforms for GaN FET are given in Figs. 6 and 7.
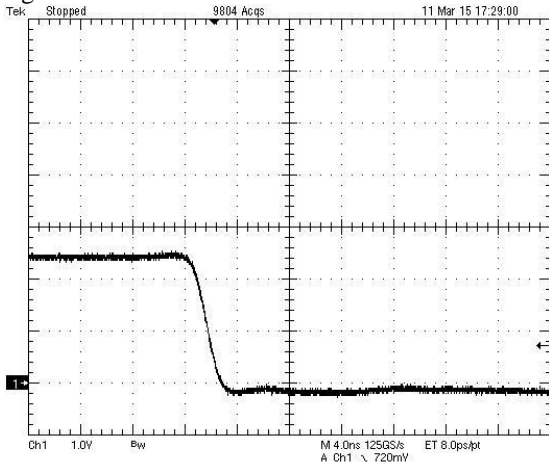
Fig. 6. GaN FET turn-on waveform at 24V
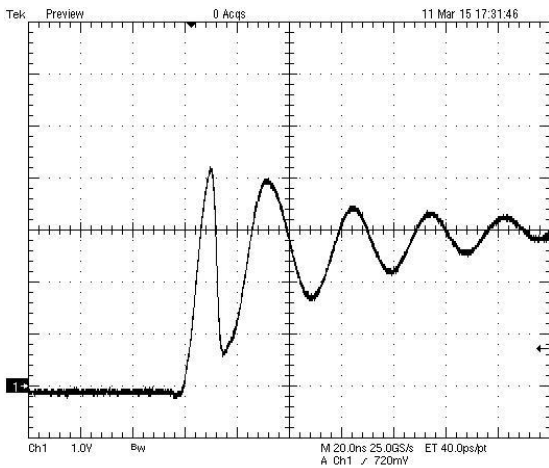
Fig. 7. GaN FET turn-off waveform at 24V

Gate and drain voltage waveforms of the forward and freewhelling rectifier are given in Figs. 8 and 9 respectively.
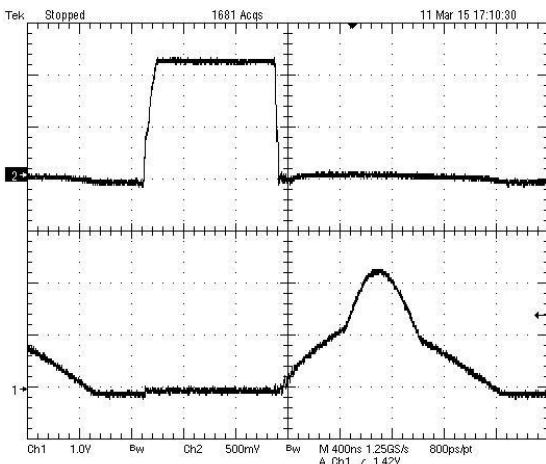
Fig. 8. Forward rectifier waveforms
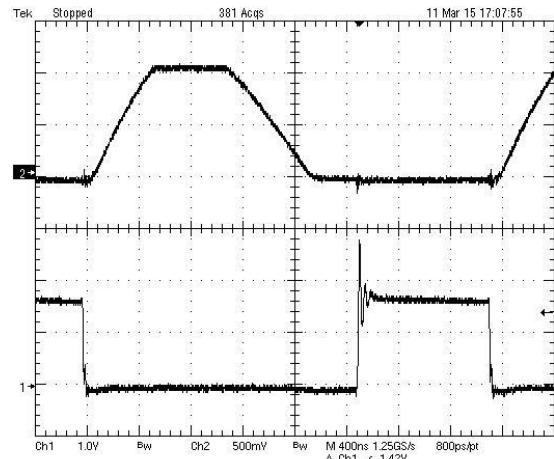
Fig. 9. Freewhelling rectifier waveforms

The picture of the converter prototype is given in Fig. 10
.
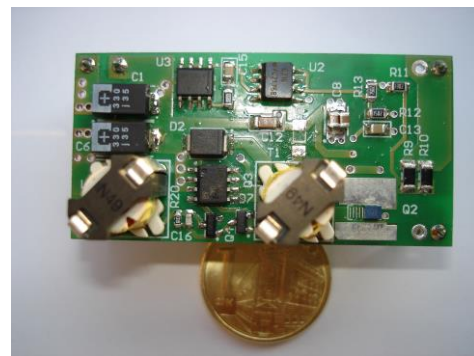
Fig. 10. Converter prototype

V. CONCLUSION

In this paper the design and analysis of resonant reset GaN forward converter with synchronous rectification is presented. The prototype was built and tested. The results verified that the full load efficiency is around 90%.

ACKNOWLEDGEMENT

REFERENCES

[1] Stephen L. Colino, Robert A. Beach, "Fundamentals of Gallium Nitride Transistors", http://epc-co.com
[2] Alex Lidow , Johan Strydom, "Gallium Nitride (GaN) Technology Overview", http://epc-co.com
[3] Zoran Zivanovic, Vladimir Smiljakovic, "Design and Realization of a small 10 Watt Forward Converter", ICEST2013 Proceedings, Volume 2, pp. 699-702
[4] Zoran Zivanovic, Vladimir Smiljakovic, "Design and Realization of a eGaN FET DCDC Converter", TELFOR2014 Proceedings, pp. 637-640

Design and Realization of a Low Noise Medical AC/DC Converter

Zoran Zivanovic[1] and Vladimir Smiljakovic[2]

Abstract – **In this paper the design of a low EMI universal input mains power supply is presented. The prototype has been built and experimental results are presented to support the theoretical analysis and to demonstrate the converter performance. The single switch flyback topology is used because of its simplicity and small number of components.**

Keywords – **CCM Flyback, EMI, Leakage current, PWM.**

I. INTRODUCTION

Power supplies designed for industrial applications are not suitable for use in medical equipment such as X-ray, CT scanners, MRI and patient monitors. Medical equipment operates with low level signals and it is more sensitive to electromagnetic interference (EMI). The Patients exposure to the smallest leakage currents may pose a threat to their life. The maximum permissible leakage current for medical equipment is no more than few hundred μA worldwide. That requirement is difficult to achieve while keeping electromagnetic interference low. Switching power supplies generate EMI and require filters to limit electrical noise. Capacitors in these filters produce leakage currents. The more effective filter produces higher leakage currents. A better approach is to minimize the amount of interference at the origin.

II. EMI REDUCTION

The power switches in switching power supplies are typically field effect transistors (FETs) and they are forced to switch as quickly as possible. Abrupt voltage and current transitions needed for high efficiency are the main causes of the noise. Because of that, input and output voltages contain low frequency ripple from 100 kHz to few MHz and high frequency content (switching spikes) with harmonics approaching 100MHz. Lowering the frequency and increasing the transition times can dramatically reduce ripple and spike amplitude.

PWM controller LT1738 from Linear Technology is a fixed frequency, single output current mode switching regulator optimized for single switch topologies such as boost and flyback, with unique circuitry to control the voltage and current slew rates of the external N-channel MOSFET.

[1]Zoran Zivanovic is with the IMTEL KOMUNIKACIJE AD, Bul. Mihajla Pupina165b, 11070 Belgrade, Serbia, E-mail: zoki@insimtel.com.

[2]Vladimir Smiljakovic is with the IMTEL KOMUNIKACIJE AD, Bul. Mihajla Pupina165b, 11070 Belgrade, Serbia, E-mail: smiljac@insimtel.com.

As seen on Fig.1 capacitor C_V provides the voltage slew rate feedback, while the current slew rate occurs by means of the sense resistor R_{SEN}. Because of the voltage slew control, MOSFET ringing is reduced and clamping circuits or snubbers in the most cases are not required. The slew rates are simply adjustable through changing the values of two resistors. The trade-off between noise and converter efficiency must be made, but fortunately the loss of efficiency is less than 10% in most cases.
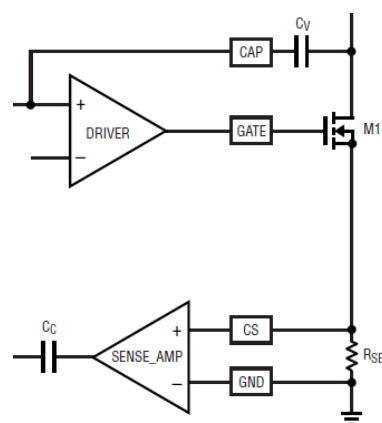
Fig.1. Slew control

DC/DC controller has all protection features including gate drive lockout for low V_{IN}, soft-start, output current limit, short circuit current limit, gate drive overvoltage clamp and input undervoltage lockout.

III. DESIGN AND ANALYSIS

The task is to design a 40W CCM flyback converter using current mode controller with careful choice of operating parameters and components. Achieving the lowest possible noise is the primary objective. The footprint size must be around 120x80mm.

First we will choose the switching frequency to be around 100 kHz, which is a compromise between the efficiency and size of the magnetics. Knowing that, a good choice of core for the transformer is PQ26/25, N95 material from TDK. Reinforced insulation is a must, but the effectiveness of which must be verified by dielectric strength testing. This means subjecting the insulation to a higher voltage than that at which it operates.

Single-section EMI filters with one stage of common mode and differential mode attenuation take the least space and have the lowest cost but require careful attention to the details. Y capacitors are planned but will not be populated.

Design specifications are given in Table I.

TABLE I
DESIGN SPECIFICATIONS

		Min	Typ	Max	
Input voltage	V_{IN}	100	220	264	V_{AC}
Output voltage 1	V_{O1}		+12		V
Output power	P		40		W
Full load efficiency	η		75		%
Switching frequency	f_{SW}		100		kHz

We will now calculate the basic parameters for the transformer, using the following equations:

$$N = \frac{N_P}{N_S} = \frac{V_{DCMIN}}{V_O + V_D} \frac{D_{MAX}}{1 - D_{MAX}} \quad (1)$$

$$I_P = \frac{I_O}{N} \frac{1}{1 - D_{MAX}} + \frac{\Delta I_P}{2} \quad (2)$$

$$L_P = V_{DCMIN} \frac{D_{MAX} T}{\Delta I_P} \quad (3)$$

$$N_P = \frac{L_P I_P}{B_{MAX} A_E} \quad (4)$$

$$I_{PRMS} = \sqrt{D_{MAX}\left(I_P^2 - \Delta I_P I_P + \frac{\Delta I_P^2}{3}\right)} \quad (5)$$

$$g = \frac{\mu_0 N_P^2 A_E}{L_P} \quad (6)$$

The results are given in Table II.

TABLE II
BASIC PARAMETERS

		Max	Typ	Min	
Duty cycle	D	0.49	0.24	0.21	
Number of prim. turns	N_P		29		
Number of sec. turns	N_S		4		
Primary peak current	I_{PPEAK}	1.28	0.98	0.96	A
Primary RMS current	I_{PRMS}	0.63	0.29	0.26	A
Secondary peak current	I_{SPEAK}	6.53	4.38	4.21	A
Secondary RMS current	I_{SRMS}	4.66	3.82	3.74	A

Now it is time to wind the transformer. We will use a bundle of 7 twisted wires - 0.2 mm enamelled copper wire for the primary and a triple insulated Litz wire with 0.28 mm wires for the secondary, in order to minimize copper losses

taking into account the skin effect and to fulfil demands for reinforced insulation.

Knowing specific core losses we can now calculate the loss in magnetic component (Table III). Total transformer power loss at $100V_{AC}$ input voltage is 1.5W. This results in approximately 43°C rise above ambient temperature. Satisfied with the results, we will keep the chosen core geometry.

TABLE III
TRANSFORMER LOSSES

		Max	Typ	Min	
Core effect. volume	V_E		6.53		cm^3
Specific core losses	P_V		0.14		W/cm^3
Core loss	P_{CORE}		130		mW
Primary resistance	R_{PRI}		137		mΩ
Primary loss	P_{PRI}	55	11	9	mW
Secondary resistance	R_{SEC}		61		mΩ
Secondary loss	P_{SEC}	1.32	0.89	0.85	W

Power switch losses can be expressed by the equation

$$P_{FET} = P_{COND} + P_{ON-OFF} + P_{QOSS} \quad (7)$$

where

P_{COND} is the conduction loss given by

$$P_{COND} = I_{RMS}^2 R_{DS} \quad (8)$$

P_{ON-OFF} is the turn-on-off switching loss given by

$$P_{ON-OFF} = \frac{V_{DS} I_P (t_{ON} + t_{OFF}) f_{SW}}{2} \quad (9)$$

and

P_{QOSS} is the capacitance charge loss given by

$$P_{QOSS} = \frac{C_{OSS} V_{DS}^2 f_{SW}}{2} \quad (10)$$

As a result of the slew rate control we can expect for t_{ON} and t_{OFF} to be around 200ns each, so the MOSFET's dissipation would be excessive. Using Eq. (7) to (10) we have calculated its dissipation given in Table IV.

TABLE IV
PRIMARY FET DISSIPATION

IRFBC40		Max	Typ	Min	
ON resistance	R_{DS}		1.2		Ω
Output capacitance	C_{OSS}		48		pF
ON time	t_{ON}		200		ns
OFF time	t_{OFF}		200		ns
Conduction loss	P_{COND}	0.71	0.15	0.12	W
Switching loss	P_{ON-OFF}	5.12	7.25	8.45	W
Capacitance charge loss	P_{QOSS}	0.09	0.33	0.47	W
Total loss	P_{TOT}	5.92	7.73	9.04	W

IV. REALIZATION

AC/DC converter was built on two layer FR-4 substrate with 35μm copper with a footprint of 120x80mm. The transformer is wounded on through hole coil former according to calculations. Output voltage is further filtered out by the added LC filter. All electrolytic capacitors are low ESR aluminum electrolytic capacitors.

Furthermore we have measured full load efficiency at various input voltages. For practical reasons we have used input voltages from 120 to 360V_{DC}. This measure gives a slightly better result than the actual value. The results are given in Table V. The efficiency is between 70 and 77% at full load.

TABLE V
EFFICIENCY

		Min	Typ	Max	
Input voltage	V_{IN}	120	300	360	V_{DC}
Input current	I_{IN}	0. 431	0.183	0.159	A
Input power	P_{IN}	51.70	55.00	57.10	W
Efficiency	η	77.37	72.72	70.05	%

Using variable isolation mains transformer and resistive load we have recorded the waveforms at the point of interest.

The drain waveforms of the primary power switch at full load and input voltages of 100 and 200V_{AC} are given in Figs. 2 and 3. As seen on the waveforms the converter is working in CCM mode. The wideband harmonic activity is entirely absent, without ringing and fast rise and fall times in the drain waveform of the power switch.

Characteristic gate voltage waveform can be seen in Fig. 4. The edges of the square wave drain waveform are slewed to trapezoidal shape (Fig. 5 and 6). As we can see the transition times are more than 200ns (in "normal" converters usually 50 to 100ns) and the net result is low harmonic content but the price is paid through efficiency reduction.
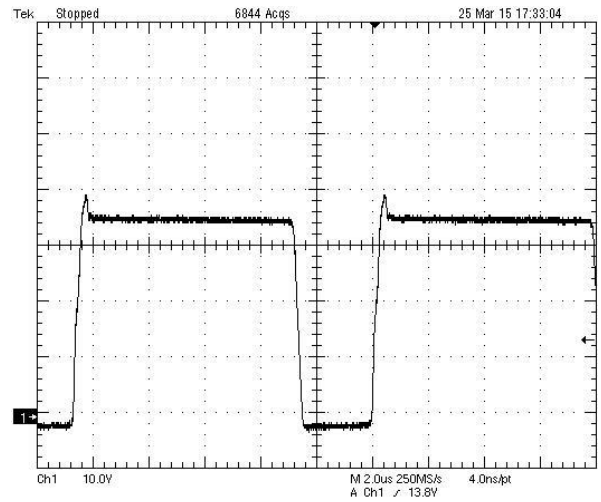
Fig.3. Drain voltage waveform at 200V_{AC}
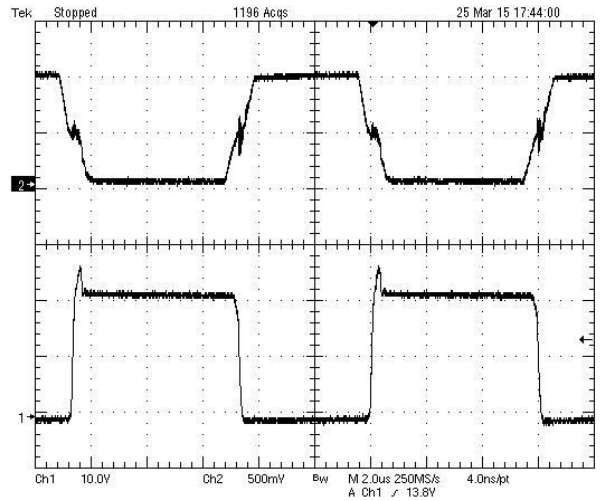
Fig.4. Drain and gate voltage waveform at 100V_{AC}
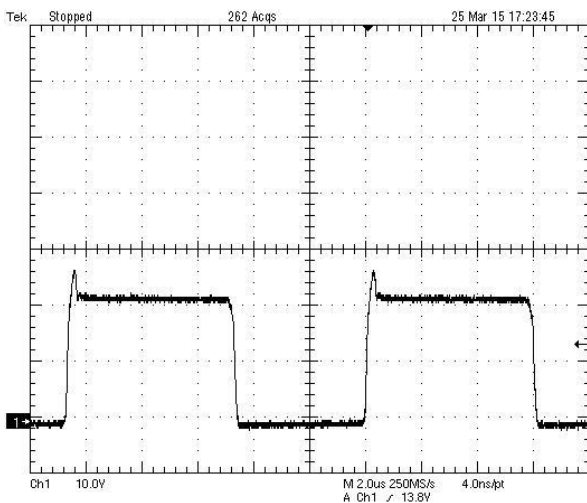
Fig.2. Drain voltage waveform at 100V_{AC}
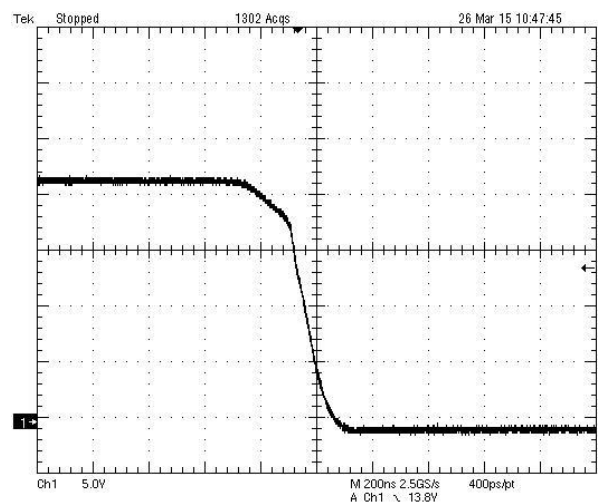
Fig.5. Drain voltage ON slew rate
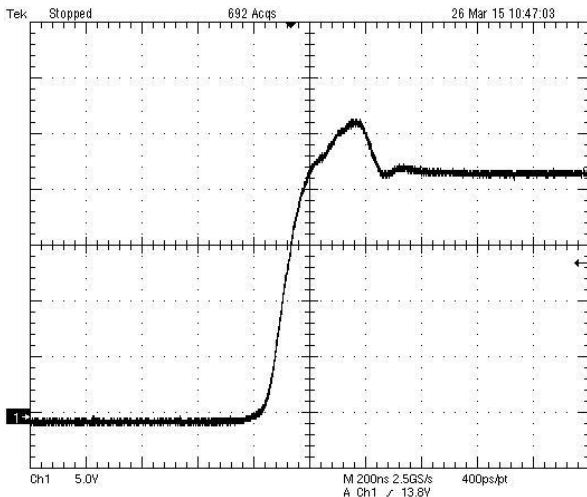
Fig.6. Drain voltage OFF slew rate
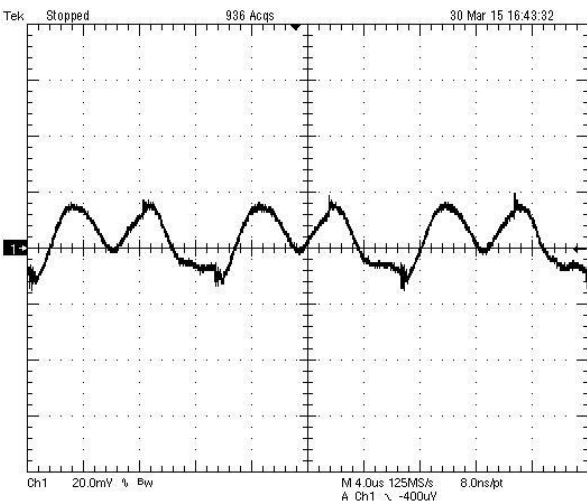
Fig.7. Input voltage ripple at full load
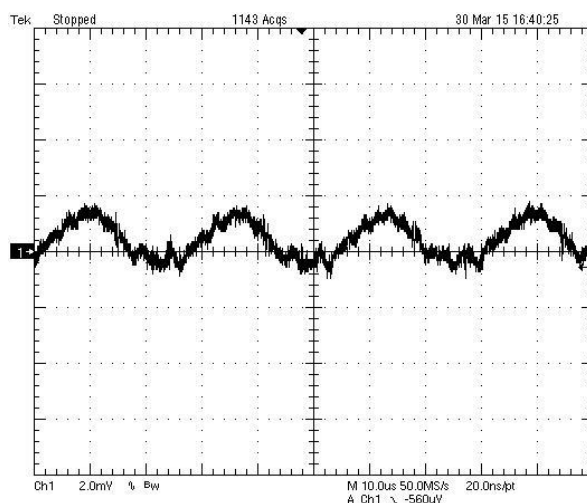
Fig.8. Output voltage ripple at full load

The input voltage ripple is around 30mV_{pp} (Fig.7) and contains only fundamental frequency without ugly spikes (measured with DC input voltage – Line Impedance Stabilization Network was not available). After the output LC filter the residual noise is 2mV_{pp} at full load and bandwidth 150 MHz (Fig.8) without high frequency components.

The picture of converter prototype is given in Fig.9.

Fig.9. Converter prototype

V. CONCLUSION

In this paper the design and analysis of 40W CCM flyback converter are presented. The prototype was built and tested. The amount of electromagnetic interference is reduced and only a small EMI filter is needed to meet the EMC requirements. With only a small amount of filtering, without Y capacitors, leakage currents are kept low, satisfying the safety requirements. The results verified that the full load efficiency is over 70%. The converter is cheap and easy to manufacture. With minor changes we can make different versions of this power supply.

ACKNOWLEDGEMENT

REFERENCES

[1] Safety Requirements for Power Supplies Intended for Medical Applications, www.artesyn.com.
[2] Zoran Zivanovic, Vladimir Smiljakovic, "Design and Realization of a Low Noise Power Converter", ICEST2014 Proceedings, Volume 1, pp. 233-236.
[3] TDK, "Ferrites and accessories PQ26/25", Data Sheet, www.tdk.eu.
[4] Power Integrations, "Flyback Transformer Construction Guide", www.power.com.
[5] Linear Technology, "Slew Rate Controlled Ultralow Noise DC/DC Controller LT1738", www.linear.com

Comparison of the piezoelectric properties of single-layer and bilayer structures with thin films of PZT and ZnO in dynamic mode

Yordanka Vucheva [1], Georgi Kolev [1], Mariya Aleksandrova [1] and Krassimir Denishev[1]

Abstract – **In this paper the piezoelectric effect was investigated for two types of samples, consisting of piezoelectric thin films PZT and ZnO with potential application as energy harvesting microdevice. The layers are deposited by RF sputtering on flexible polyethylene terephthalate (PET) substrates. Thermally evaporated aluminum (Al) films were used as bottom and top contacts. The piezoelectric reaction of the samples was measured by lab made experimental setup for dynamic mode. The generated piezoelectric charge was measured at different frequencies of the vibrations.**

Keywords – **PZT, ZnO, MEMS sensors, RF sputtering, piezoelectric thin films**

I. INTRODUCTION

Piezoelectric effect is the ability of certain materials to generate an electric charge in response to applied mechanical stress. One of the unique characteristics of the piezoelectric effect is that it is reversible, meaning that materials exhibiting the direct piezoelectric effect (the generation of electricity when stress is applied) also exhibit the converse piezoelectric effect (the generation of stress when an electric field is applied) [1]. This effect is formed in crystals that have no center of symmetry. Such crystals are the piezoelectric materials lead zirconium titanate, known as PZT and zinc oxide (ZnO).

Piezoelectric materials, such as PZT and ZnO are very important for the Micro-electro-mechanical systems (MEMS) and Nano-electro-mechanical systems (NEMS) design, as sensors and actuators. Traditionally, the PZT and ZnO are used as bulk materials with thickness higher than 100-200 μm, which are difficult to be integrated into microdevices. The requirements for miniaturization can be fulfilled by conversion of the piezoelectric materials into thin nanometric films via conventional microelectronic technology, like RF sputtering.

ZnO forms a hexagonal wurtzite structure with 6mm symmetry. It possesses the lack of center symmetry required for piezoelectric materials. The coupling coefficient for ZnO is also relatively high which makes ZnO an excellent material to use in a wide variety of piezoelectric applications [2]. PZT

[1]Yordanka Vucheva, Georgi Kolev, Mariya Aleksandrova and Krassimir Denishev are with the Department of Microelectronics, Faculty of Electronic Engineering and Technologies at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: ydv@tu-sofia.bg.

$(Pb[Zr_{(x)}Ti_{(1-x)}]O_3)$ is a ferroelectric material which has excellent piezoelectrical and pyroelectrical properties. The structure of this piezoelectric ceramic is perovskite [3]. Piezoelectric materials are important for MEMS and NEMS because of their possibility to convert one form of energy to another.

When piezoelectric material is placed under mechanical stress, a shifting of the positive and negative charge centers in the material takes place, which then results in an external electrical field. When reversed, an outer electrical field either stretches or compresses the piezoelectric material [1].

Piezoelectric actuators convert electrical energy to mechanical energy. This is why they are referred to as "motors" (often linear motors). Piezoelectric sensors convert mechanical energy into electrical energy. This is why they are referred to as "generators". In most cases, the same element can be used to perform either task [3]. Piezoelectric sensors or piezoelectric generators are very interesting for engineers for that they can be used for energy harvesting devices for MEMS. They can be single-layer, bilayer or multilayer generators depending on the number of piezoelectric layers. Single-layer structure which is longitudinal (d_{33}) generator is shown in Fig. 1. This means that mechanical stress is applied in direction parallel to polarization [4]. The case for bilayer structure is the same but there are two piezoelectric layers.
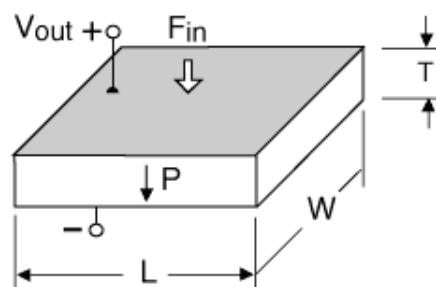
Fig. 1. Single-layer piezoelectric structure [4]

Traditionally, the PZT and ZnO are used as bulk materials with thickness higher than 100-200 μm, which are difficult to be integrated into microdevices. In this paper these materials are used like thin films with thickness around x.100 nm. A suitable approach for piezoelectric performance improvement could be combination of different piezoelectric materials, but this hypothesis should be studied. That's why in this research, the piezoelectric effect was investigated for two types of samples, consisting of piezoelectric thin films PZT and ZnO, RF sputtered. Two single-layer structures and one bilayer structure are produced. The piezoelectric reaction and

generation of the electrical energy of these structures are investigated by experimental setup for dynamic mode measurement.

II. EXPERIMENTAL WORK

A. Producing of the samples

The experimental work begins with the manufacturing of the samples. Three samples were made with two piezoelectric materials – PZT and ZnO. Two of the samples were single-layer structures with one piezoelectric layer PZT or ZnO, and the third sample was bilayer structure, consisting of two piezoelectric layers in combination PZT and ZnO. All of the samples were obtained on polyethylene therephthalate (PET) substrates. This foil type material was chosen, due to its high flexibility and ability for easy deformation, which additionally stimulates generation of piezoelectric voltage. On the PET pieces aluminum (Al) 200 nm films were deposited by thermal evaporation in vacuum chamber A400-VL Leybold Heraeus at vacuum level 10^{-5} Torr for bottom and top electrodes. The piezoelectric layers were deposited by RF sputtering at total partial pressure of 2.10^{-2} Torr. The sputtering conditions were as follows: U_{pls}=0,6kV, I_{pls}=150mA, P_{pls}=90W (specific plasma power 11,8W/cm^2). The time of the deposition was 50 min for PZT and 20 min for ZnO. The thicknesses of the PZT and ZnO layers were measured respectively 200nm and 60nm by Alfastep Tencor 100. The deposition conditions were changed, because of the fact that, due to the processing temperature the substrates started to be softer. Therefore the thicknesses of the piezoelectric layers are different than expected. The size of the top electrode was 2x1,5 mm, defined by shadow mask.

The both single-layer structures are with PZT and ZnO, separately and the other bilayer structure is consisting of the PZT and ZnO one over another. Cross-section view of the single-layer and two-layer structures is shown in Fig. 2.
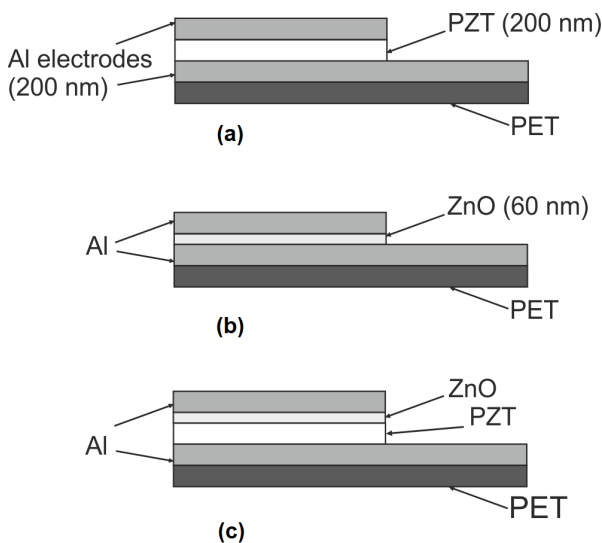
Fig. 2. Cross-section view of single-layer (a), (b) and bilayer structures (c)

B. Measuring the piezoelectric reaction in dynamic mode

The piezoelectric reaction of the samples was measured by lab made experimental setup for dynamic mode, which is shown in Fig. 3.
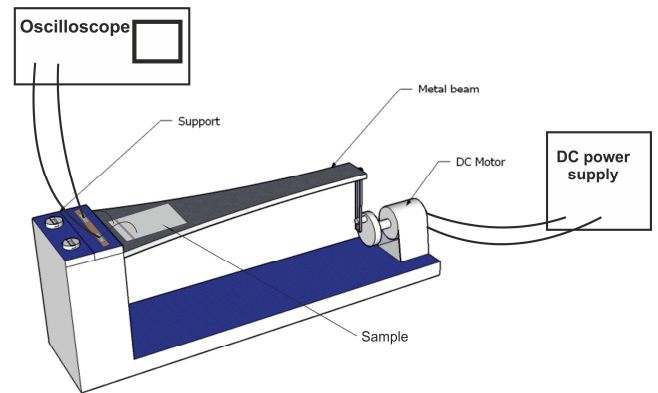
Fig. 3. Experimental setup for dynamic mode measurement of the piezoelectric reaction

A picture of the experimental setup is shown in Fig. 4.
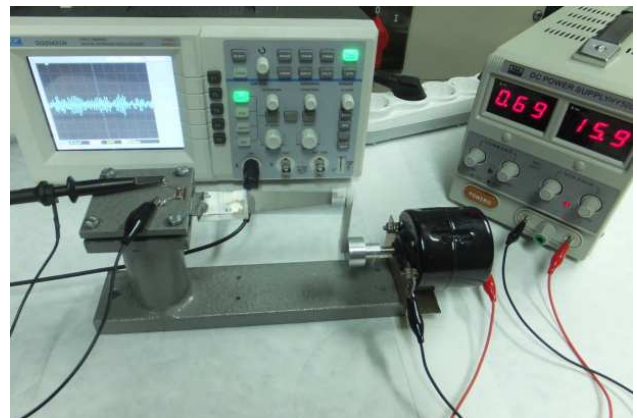
Fig. 4. The experimental setup

To measure this reaction, every sample was fixed to the beam by single component adhesive Z70 and then was bonded to the measurement leads. The generated piezoelectric voltage was measured at different frequencies vibrations of the metal beam, which are induced and defined by the DC motor. The maximum strength of the experimental stand is 1096 gr or 10,75 N. The time sweeps of the generated oscillations were observed by digital oscilloscope DQ2042CN.

The measurements for all samples were done at equal conditions for values of the voltage from the DC power supply in the range 14-20 V and vibrations of the metal beam in the range 23-55 Hz.

III. RESULTS AND DISCUSSION

The signals, for the generated piezoelectric voltage, were measured for all piezoelectric structures at the following values of the DC power supply 14 V, 16 V, 18 V and 20 V. At these values of the voltage, applied to the motor, the frequencies of the metal beam vibrations are 23 Hz, 33 Hz, 45Hz, 55Hz, respectively. Below the signals from the structures and the discussions about them are shown.

A. Single-layer structure with PZT

Single-layer structure with piezoelectric layer of PZT was subjected to vibrations with different frequencies, which induced measurable piezoelectric charge inside the layer. The generated voltage from the 200nm thick piezoelectric layer is in the range 40-60mV, as the values of 40mV are measured at 23 Hz and 33 Hz (Fig. 5 (a)) and the values of 60mV at 45 Hz and 55 Hz (Fig.5 (b)). Fig. 5. presents two signals of the mentioned values of the generated piezoelectric voltage. It should be mentioned that the value 60mV of the voltage exists as peaks in the output signal and it is not permanent. This is actually expected, because the elastic beam move frequently and it passes through certain position often so the voltage peaks with similar frequency can be monitored on the oscilloscope screen.

(a)

(b)
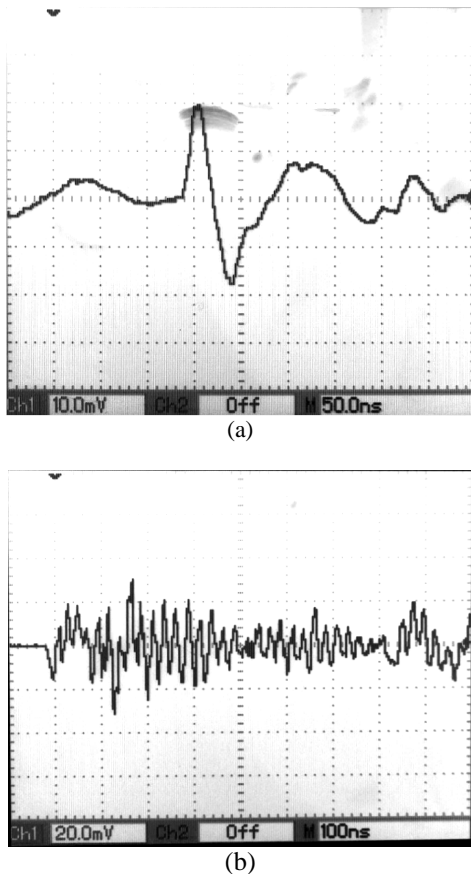
Fig. 5. Output voltage signals of the generated piezoelectric voltage from single-layer structure with PZT

With the increasing of the frequency, the peaks in the output signal of voltage are higher, but sparsely. In the range 23-33 Hz of the vibrations, the piezoelectric voltage consists of more peaks in the signal, with higher value.

B. Single-layer structure with ZnO

At the same conditions the output signals of the generated voltage were taken from the one-layer piezoelectric structure with ZnO. In this case, the generated voltage, from the 60nm thick piezoelectric layer, is in the range 80-100 mV, and like in the previous case the highest voltage is generated in the range 23-33 Hz. In Fig. 6 two output signals for 80 mV (a) and 100 mV (b) are shown. 80 mV of the generated piezoelectric voltage was measured at 23 Hz, 45 Hz and 55 Hz. The value of 100mV, was measured at 33 Hz.

(a)

(b)
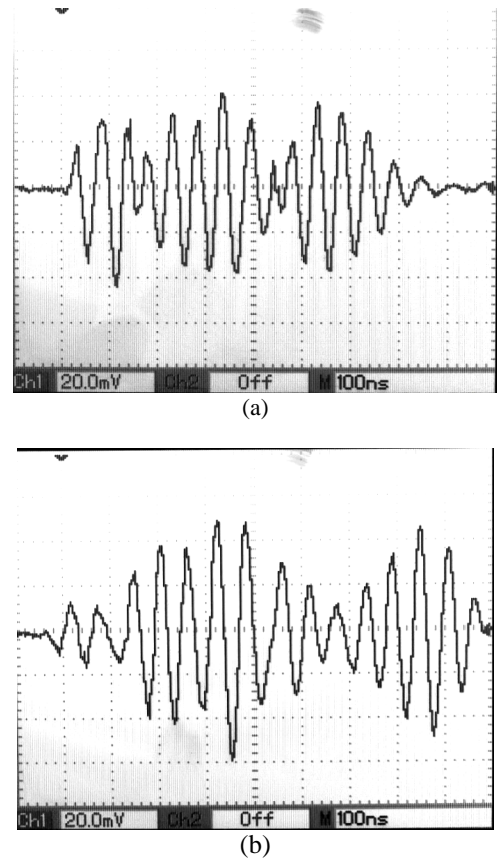
Fig. 6. Output voltage signals of the generated piezoelectric voltage from single-layer structure with ZnO

C. Bilayer structure with PZT and ZnO

For this bilayer structure PZT layer was firstly deposited, and the second one was from ZnO. In this case the generated piezoelectric voltage is almost half of this in the previous one-layer structures. Its value is in the range 30-40 mV for all frequencies and values of 80mV were monitored rarely for 45 Hz and 55 Hz measuring points, due to depolarization or inability of the dipoles of both materials to follow the change of the mechanical load with one and the same rate, so the

slowest one (PZT) cannot produce a charge. This behavior is probably due to the increased inner capacitance of the bilayer structure. Fig. 7. presents some of the captured signals of the generated piezoelectric voltage from the bilayer structure.
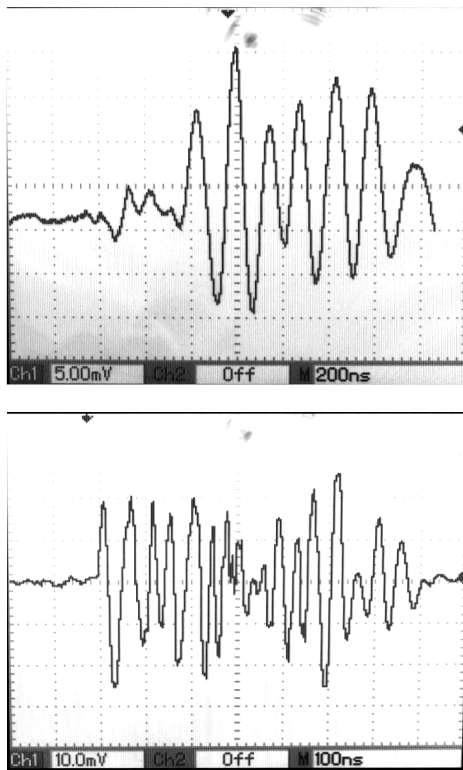
Fig. 7. Output voltage signals of the two-layer structure with PZT and ZnO

From the comparison of the bilayer structure with the single-layered ones, it can be seen that the voltage values, generated by the bilayer structure, are almost half of these from the single-layer structures.

IV. CONCLUSION

After considering the generated piezoelectric voltage signals of the all structures, it could be concluded that, bilayer structure have a lower output voltage, comparing the single-layer piezoelectric structures. The voltage values, generated by the bilayer structure, are almost half of these from the single-layer structures. This can be explained by the electrical charge of the structures. Bilayer structure generates a higher charge than single-layer structures, but the inner capacitance is bigger and, as a result, the generated voltage is lower. That should be studied in the future work.

For the single-layer structures, more interesting is the fact that the 60nm ZnO layer has almost the same voltage values as the voltage, obtained by the 200nm PZT layer. This could be as a consequence of the piezoelectric coefficients of the layers, but the future work should be done for confirmation the reason of that.

Another fact is that the value of the output voltage, at 23 Hz and 33 Hz, for single-layer samples, is lower but for bilayer sample is higher. At high frequencies, is just in opposite, the value of the voltage, for single-layer samples, is higher but for bilayer samples is lower.

ACKNOWLEDGEMENT

REFERENCES

[1] http://www.nanomotion.com/piezo-ceramic-motor-technology/piezoelectric-effect/
[2] Johnson, R. L., "Characterization of piezoelectric ZnO thin films and the fabrication of piezoelectric micro-cantilevers", Iowa State University, 2005.
[3] Kolev, G.D., "Investigation of piezoelectric effect in thin layers, for application in harvesting devices and MEMS sensors", Annual Journal of ELECTRONICS, ISSN 1314-0078, 2012.
[4] http://www.piezo.com/tech2intropiezotrans.html
[5] Gretarsson D., "Energy Harvesting using Piezoelectric Generators", Copenhagen, 2007.
[6] B. Jaffe, W. Cook, H. Jaffe, "Piezolectric Ceramic", Academic Press, London and New York, 1971.
[7] T.L. Jordan, Z. Ounaies, "Piezoelectric Ceramics Characterization", ICASE Report No. 2001-28.

POSTER SESSION
ENERGY SYSTEMS AND EFFICIENCY

Experimental Investigation of the Electrical Parameters of the Soil for the Purpose of the Grounding System Design

Rositsa Dimitrova[1], Margreta Vasileva[2], Marinela Yordanova[3]

Abstract – **The aim of the paper is to propose a mathematical equation, describing the influence of the electromagnetic field' frequency on the specific volume resistivity and the dielectric permittivity of multilayer soil. To obtain the dependence experimental studies have been done in the electrical power substation construction area.**

Keywords – **Specific volume resistivity, Relative dielectric permittivity, Multilayer soil**

I. INTRODUCTION

There are not many data for the specific volume resistivity ρ_v and the relative dielectric permittivity ε_r of the soil at different frequencies f of the electromagnetic field, arising due to a lightning current. Data for frequency of 50 *Hz* can be found more frequently. Since the specific volume resistance directly affects the grounding resistance, it is necessary to determine ρ_v depending on the frequency of the flowing through the grounding rod current. Such research has been done and published by the authors [1], but for the significantly narrower frequency range from 12 *Hz* to 100 *kHz*. A direct lightning stroke over an object leads an impulse current to flow through the grounding elements and high frequency processes develop in the soil area around them. The specific volume resistivity and the relative dielectric permittivity of the soil depend on those processes. The precise determination of the electrical characteristics of the soil ρ_v and ε_r is necessary to establish a correct model of the grounding system for the study of the wave processes.

In Section II the experiment equipment is described and the mathematical equations for data processing are given. In Section III, the experimental results and the analyses are presented. The concluding remarks are given in Section IV.

II. EXPERIMENT DESCRIPTION

Samples of the soil from the area where an electrical power substation will be constructed have been taken. The samples were taken from depths 0.4 *m* and 0.8 *m*, as well as from the soil surface. The values of the specific volume resistivity and the dielectric permittivity have been measured for wet and dry soils, as well as for the loose and the pressed soils.

[1] Department of Electrical Power Engineering, Technical University of Varna, Varna, Bulgaria E-mail: r.dimitrova@tu-varna.bg

[2] Department of Electrical Power Engineering, Technical University of Varna, Varna, Bulgaria E-mail: m.vasileva@tu-varna.bg

[3] Department of Electrical Power Engineering, Technical University of Varna, Varna, Bulgaria E-mail: m.yordanova@tu-varna.bg

The measurements of the electrical characteristics R, C, C_0 for each of the samples, having different moisture contents, are made with the device Precision Impedance Analyzers 6500B Series, shown in Fig.1. The analyzer can measure directly the values of the resistance and the capacitance of the samples and after that the specific volume resistivity ρ_v and the relative dielectric permittivity ε_r can be calculated using formulas.
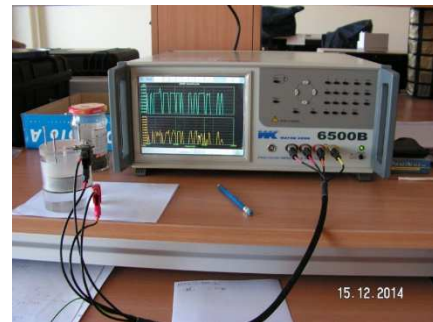
Fig. 1. Precision Impedance Analyzers 6500B Series

The device can measure the following parameters: capacitance C; inductance L; resistance R; reactance X; conductance G; dissipation factor D; quality factor Q, impedance Z; admittance Y; phase angle θ. The total frequency range is from $20Hz$ up to $5MHz$. The speed of measurement for the used frequency from 1 *kHz* to 1000 *kHz* is from 60 *ms* to 250 *ms* (in the Meter Mode) and from 60 *ms* to 190 *ms* (in the Analysis Mode). The measurement accuracy is up to 0, 05%, and the dissipation coefficient has accuracy up to 0, 0005%.

For each sample the analyzer allows rapid and accurate measurements of a certain number of values of R, C, C_0, Z and D at different frequencies.

The samples are placed in a suitable sensor (Fig. 2), which is a flat capacitor with circular copper plates with a diameter of 37,78 *mm* and distance between them 49,2 *mm* (other sensors with different dimensions can be used). The surrounding surface is made from plexiglass, because of its great resistance. On the top of the sensor there is a spring so that the sample is under a constant mechanical pressure. This provides a minimum presence of air in the test soil.

Fig. 2. General view of the measuring sensor

The analyzer measures the resistance R_v of the soil sample. Using the measured value of R_v the specific volume resistivity can be calculated:

$$\rho_v = \frac{R_v S}{h}, \Omega.m \qquad (1)$$

where R_v – volume resistance of the soil sample, Ω; S – surface of the electrode, m^2; h – thickness of the soil sample.

The specific volume resistivity can be calculated after measurement of the parameters C and D and using following formula [2, 5]:

$$D = \omega C R_v \qquad (2)$$

where $\omega = 2.\pi.f$

$$R_v = \frac{D}{2\pi f C} \qquad (3)$$

The dielectric permittivity of the soil can be determined using the dielectric measurement method [3], which is the most appropriate for granulated materials. This method measures the capacity C of the capacitor with a dielectric between the plates (the soil sample) and the capacity of the air capacitor C_0 (without soil between the plates). The alteration of values of the dielectric (the soil sample) parameters at different frequencies of the electric field is observed. Under the influence of the electric field the dielectric polarizes itself and its dielectric permittivity can be measured by measuring the capacitance at different frequencies of the alternating current.

The capacitance of the flat vacuum capacitor is:

$$C_0 = \varepsilon_0 \varepsilon_r \frac{S}{h} \qquad (4)$$

where $\varepsilon_r=1$.

The capacitance of the same capacitor when there is a dielectric (the soil sample) with relative dielectric permittivity ε_r between the electrodes is:

$$C = \varepsilon_r \varepsilon_0 \frac{S}{h} \qquad (5)$$

where $\varepsilon_0= 8,85 .10^{-12}$ F/m; $S=2374,6.10^{-6} m^2$.

Then the relative dielectric permittivity of the soil sample is:

$$\varepsilon_r = \frac{C}{C_0} \qquad (6)$$

The formula (6) allows, after measurement of the two capacitances, to determine the relative dielectric permittivity of the soil sample between the electrodes of the sensor. The relative dielectric permittivity is determined at different soil humidity.

The humidity of the samples ψ is evaluated by the relative content of the water into them [2, 3]:

$$\psi = \frac{m_1 - m_2}{m_2}.100\% \qquad (7)$$

where m_1 is the mass of the wet sample, m_2 is the mass of the same dry sample.

The difference $m_1 - m_2$ is exactly the mass of the water into the sample.

The drying at the samples is carried out by a special methodology - Gravimetric method [3]. The mass m of the wet material is measured. Then the sample is placed in an oven at about $120^0 C$ to dry for an hour or two. After that they are placed into a desiccator for about an hour to reach the thermal equilibrium. The mass m_0 of the dried sample is measured. By formula (7) is calculated the humidity. For greater accuracy should be used electronic scales for measuring of the mass.

The measurements are carried out at different densities of the soil, which are calculated using the following formulae:

$$\gamma_n = \frac{m[kg]}{V[m^3]} \qquad (8)$$

$$V = Sh \qquad (9)$$

$$S = \frac{\pi d^2}{4} \qquad (10)$$

where m – mass of the soil sample; V – volume of the soil in the sensor; S – surface of the electrode; d – diameter of the electrode; h – thickness of the sample.

The obtained from measurements results are processed by the computer program Grafer [4]. This is a mathematical program for processing of the experimental results. The program has an option to draw a theoretical curve by the experimental results and to give the best mathematical function describing the experimental curve as well as the coefficient of determination, which shows how the approximated (theoretical) curve is close to the experimental.

The resultant mathematical expression of the dependencies of the specific volume resistivity and the relative permittivity from the frequency (from 1 kHz to 1000 kHz) of the electric field can be used for different humidity and density of the samples. This equation can be used in the calculation of the specific volume resistance of the soil and the grounding resistance value at a specific frequency.

III. EXPERIMENTAL RESULTS

Measurements for investigation of the dependence of the specific volume resistivity and the dielectric permittivity from the frequency in the range 1 kHz to 1000 kHz at different humidity and density of the samples were made.

A. *Dependence of the specific volume resistivity of wet pressed ($\gamma_p = 1239,8$ kg / m^3) soil samples from the frequency of the electric field at a depth of 0.4 m.*

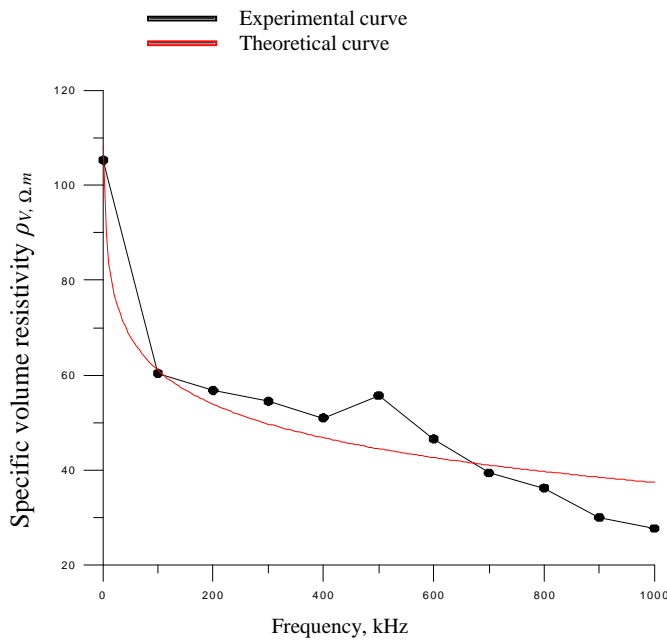Figure 3 presents the results from the fulfilled test measurements of wet pressed soil samples.

Fig.3. Specific volume resistivity ρ_V of wet pressed soil from point №1 at a depth of 0.4 m versus the frequency

For the results the following equation is processed by Grafer:

$$Y = -10{,}25159429\ ln(X) + 108{,}2316067$$

The coefficient of determination, R-squared = 0,914924.

The results present that the alteration of the specific volume resistivity ρ_v is according to a power decreasing function which is confirmed and by other authors [1, 2, 3, 6, 7] for granulated materials for different frequency ranges. The high coefficient of determination 0,914924 shows that the experimental curve can be described very precisely by the function as:

$$\rho = \rho_0\left(\frac{f}{f_0}\right)^{-n}$$

B. *Dependence of the relative permittivity of wet pressed ($\gamma_p = 1207{,}9\ kg/m^3$) soil samples from the frequency of the electric field at a depth of 0.4 m.*

The results are shown in Fig. 4

For the results the following equation is processed by Grafer:

$$Y = -20{,}17964574\ ln(X) + 145{,}8162477$$

Coef of determination, R-squared = 0,855494.

In most cases, experimental curves, describing the permittivity dependence on the frequency, get a complex character and the Grapher program processes them with less accuracy.
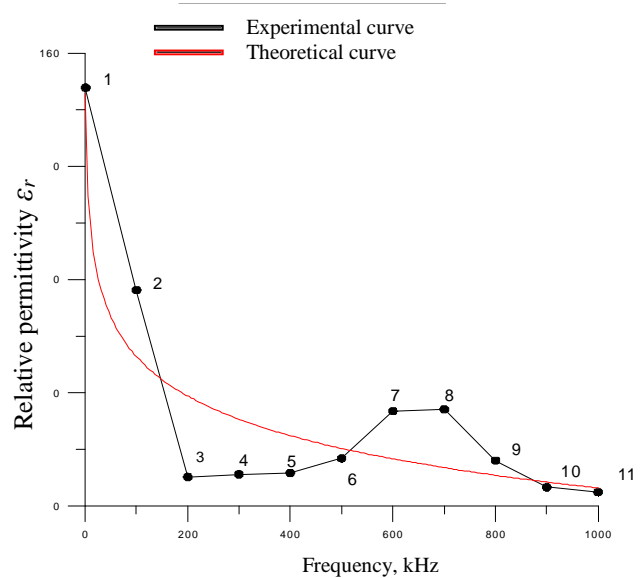
Fig.4. Relative permittivity ε_r of wet pressed soil from point №2 at a depth of 0.4 m versus the frequency

In this case the curve is examined in several sections – Fig.5÷7.
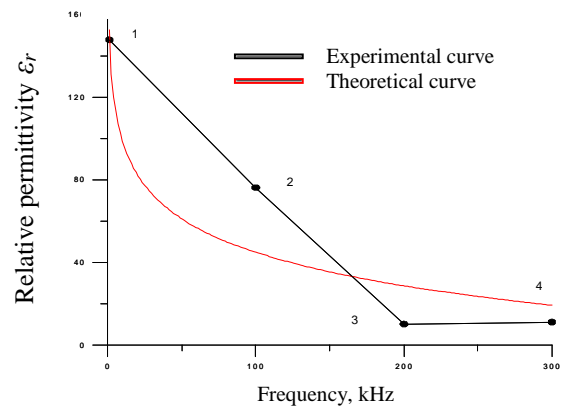
Fig.5. Relative permittivity ε_r of wet pressed soil from point №2 at a depth of 0.4 m for the sector №1 versus the frequency

Equation: $Y = -23{,}32579912\ ln(X) + 152{,}4130114$

Coef of determination, R-squared = 0,88972
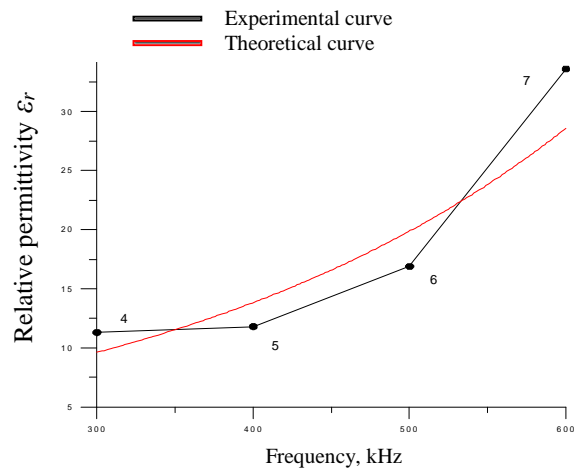
Fig.6. Relative permittivity ε_r of wet pressed soil from point №2 at a depth of 0.4 m for the sector №2 versus the frequency

Equation: *ln(Y) = 0,003628384114 X + 1,175912635*

Coef of determination, R-squared = 0,863944
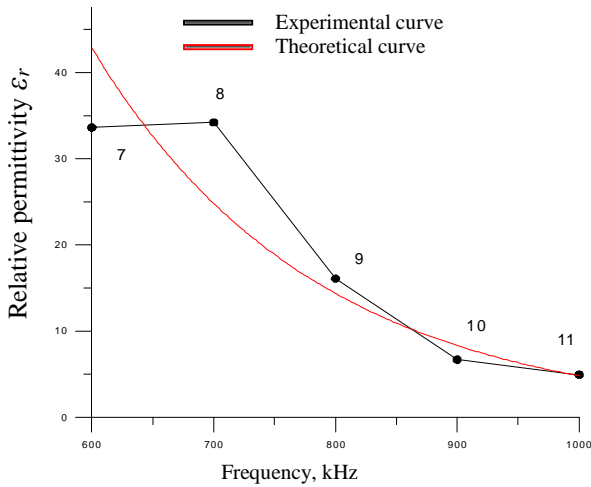
Fig.7. Relative permittivity ε_r of wet pressed soil from point №2 at a depth of 0.4 m for the sector №3 versus the

Equation: *ln(Y) = -0,005454867346 X + 7,030009011*

Coef of determination, R-squared = 0,930534

For the three section:

$$Y = \begin{cases} -23{,}32579912\,ln(X) + 152{,}4130114; & 0 \prec X \prec 300kHz \\ ln(Y) = 0{,}003628384114\,X + 1{,}175912635; & 300kHz \prec X \prec 600kHz \\ -0{,}005454867346\,X + 7{,}030009011; & 600kHz \prec X \prec 1000kHz \end{cases}$$

C. Dependence of the specific volume resistivity and the relative permittivity of the soil samples from the depth at which the sample was taken.

The results are shown in Fig. 8 and Fig. 9.
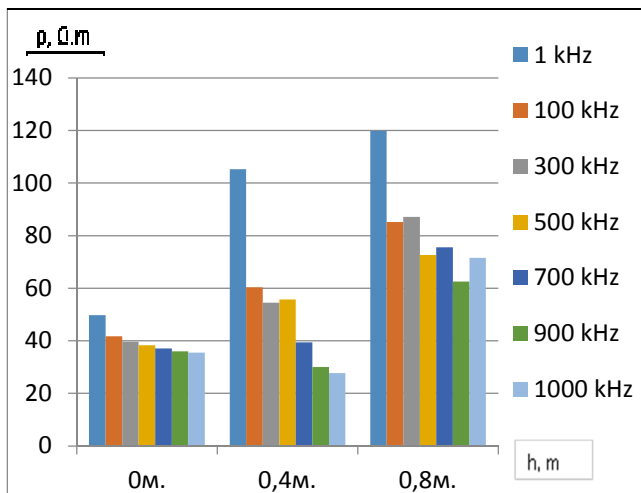
Fig. 8. Specific volume resistivity of the wet pressed soil versus the depth at which the sample was taken.
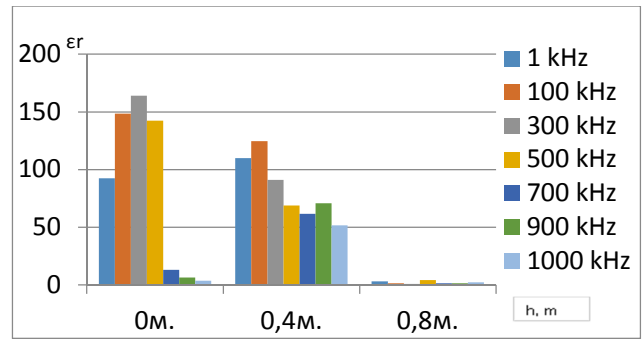
Fig. 9. Dielectric permittivity versus the depth at which the sample was taken.

IV. CONCLUSION

The dielectric method for measuring the electrical characteristics of the soil and the received experimental results, which are confirmed in other studies, give reason to continue the investigation for other samples - multilayer soils.

The resultant mathematical expression of the dependence of the specific volume resistivity from the frequency can be used in the calculation of the resistance of the grounding system for a specific frequency value.

The resulting mathematical expressions of the dependence of the dielectric permittivity from the frequency can be used in calculating the capacitance of the grounding system, which is necessary to create a model scheme of the installation for the study of the wave processes in it.

ACKNOWLEDGEMENT

REFERENCES

[1] Dimitrova R., M. Vasileva, K. Kardjilova, Influence of Frequency on Resistivity and Dielectric Permittivity of Multilayer Soil, SIELA 2014, XVIIIth International Sumposium on Electrical Apparatus and Technologies, 29 – 31 May 2014, Bourgas, Bulgaria, ISBN 978-1-4799-5817-7, pp. 37-40.

[2] Barudov S., V. Iliev, B. Nikov, Materials and components in electronics, TU-Varna, 2005, ISBN – 954-20-0294-7 .

[3] Kardjilova K., Specific methods for measurement of physical properties of biological materials., Varna 2014, ISBN 978-954-760-316-5, PP.224-234.

[4] Manual for the Grafer software.

[5] Precision Impedance Analyzers 6500B Series – User Manual, Issue A4, Part № 9H6500B

[6] Margolin N. F., Currents into the ground, Gosudarstvennoe energeticheskoe izdatelystvo, Moscow, 1947.

[7] Dolin P. A., Fundamentals of technical safety for electrical equipment, Energoatomizdat, Moscow,1984.

Study of the Energy Characteristics and the Temperature Influence on the Liquid Medium Resistance during the Formation of a High Voltage Discharge

Milena D. Ivanova[1]

*Abstract –***Under the formation of high voltage discharges in a liquid medium many physical and chemical processes occur, which depend on many different factors (characteristic of the electric field, composition of the liquid medium, external parameters such as temperature, pressure, etc.). Considering the dynamics of these processes, the purpose of the present work is experimental study and analysis of the influence of the temperature and the energy characteristics on the liquid medium resistance change at the formation of a high-voltage discharge pulse. The obtained results and dependencies can be used for design of devices for formation of such discharges in liquids for different technological applications.**

*Keywords – **High voltage discharge in liquid, Capacitive energy accumulation, Periodically attenuating discharge pulse.***

I. INTRODUCTION

High voltage pulse discharges in liquids are an object of many scientific researches. The specifics of the processes which occur in the liquid medium are related to the appearance of shock waves, UV radiation, generation of chemically active components, ozone and others. Such electrical discharges are widely used for water treatment aiming biological disinfection, cleaning of scaling and deposits on pipes, removal of chemicals in liquids, etc. [1,3]

The resistance of the formed plasma channel at a high voltage discharge in liquid depends on the alteration of the free current carriers concentration in it. Some of the factors which impact on the process are:

- recombination of current carriers;
- drift movement of current carriers;
- scattering of current carriers outside the plasma channel at the expense of a thermal diffusion and appearance of a hydraulic wave.

The processes develop with high dynamics and different time constants, which most commonly depend on the structure of the medium, external factors (temperature, parameters of the electric field, elements of the discharge circuit and spatial geometry of plasma channel [6]. Their mutual influence is difficult to be covered analytically and this requires conducting of parametric experimental studies. The results from them can be used for the design of devices for formation of high voltage discharge pulses in liquids for certain technological applications.

The present work is dedicated on the experimental study and analysis of the influence of the liquid medium temperature and the energy parameters of the formed discharge pulse on the liquid resistance alteration.

In Section 2, the experimental system and the obtained results from the conducted experiments are presented. The analyses are presented in Section 3. The concluding remarks are given in Section 4.

II. EXPERIMENT

A. Prototype

The experiments are conducted by using the prototype system, showed in Fig.1 [2].
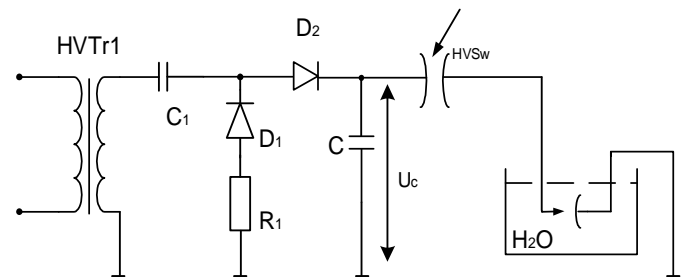
Fig.1 Device for formation of high voltage discharges in water

The charging circuit consists of a high voltage transformer (HVTr1) and a voltage doubler (C_1, D_1, D_2, C), which charges the work capacitor battery C. The discharging circuit includes a high voltage switch (HVSw) – trigatron – Fig.2 and the discharge gap in the water.

The trigatron is a three – electrode air discharger with two separated discharge circuits and two spatially divided gas-discharge channels, which evolve consistently in time. The first channel ensures the appearance of a discharge between the control electrode and the hemispherical electrode adjacent to it. The other channel forms a high voltage pulse discharge on the base of the energy accumulated in the capacitor battery between the two hemispherical electrodes. The commutation capabilities of the trigatron has been investigated in the literature [2].

Determining the alteration of the liquid resistance is important with respect to the sizing of the charging and discharging circuits.
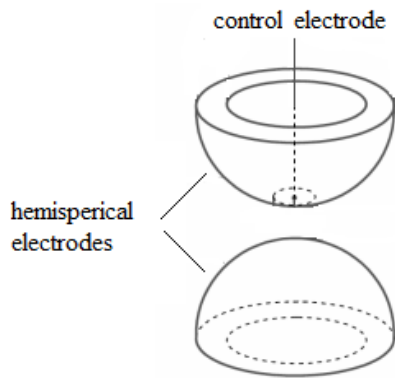
[1]Milena Ivanova is with the Department of Electrical Power Engineering, Electrical Engineering Faculty at Technical University of Varna, 1, Studentska Str., Varna, 9010, Bulgaria,
e-mail: m.dicheva@tu-varna.bg.

Fig.2 Trigatron

The experimental study of a high voltage discharge pulse in water is conducted with the following parameters of the system from Fig.1:
- Capacitance of the work capacitor battery C=1÷3μF;
- Voltage, to which it is charged – U_c=8,5÷12kV;
- Temperature of the water – T=16÷90°C.

B. Experimental results

The energy characteristics of the discharge pulse have been recorded for all values of the described parameters above.

The formed high voltage discharge pulse has a periodically attenuating character. The time diagrams of the discharge current and discharge voltage for some values of the parameters are shown in Fig.3 ÷ Fig.5. On the graphs – the upper curve presents the discharge voltage and the lower one – the discharge current.

The discharge current can be described by Eq.1 [5]:

$$\frac{d^2i}{dt^2} + 2\delta\frac{di}{dt} + \omega_0^2 t = 0 \qquad (1)$$

where:
- $i(t_0) = 0$ – initial condition;
- $\delta = \frac{R}{2L}$ is damping ratio;
- $\omega_0^2 = \frac{1}{LC}$ - resonant frequency of the circuit.

When δ<<ω, the process is periodically attenuating.

For the circuit resistance R is valid R = R_{HVSw} + R_L, where the R_{HVSw} is the resistance of the trigatron and R_L is the resistance of the liquid medium.

Two consecutive periods from the discharge formation can be observed on the time diagrams:
- Predischarge period – the time between the moment of arising a discharge in the trigatron, followed by accumulation of current carriers in the discharge gap and appearance of the discharge in the liquid medium;
- Discharge period (duration of the high voltage discharge pulses) – from the moment of the discharge appearance in the liquid medium until the attenuation of the discharge process.
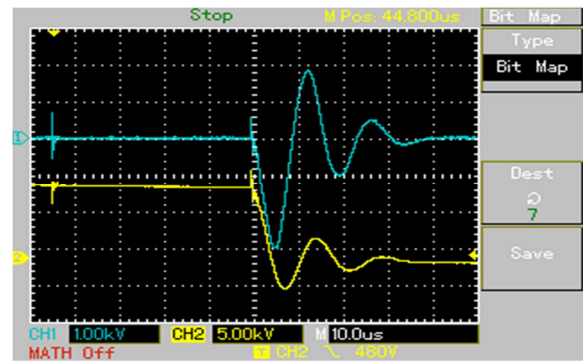
Fig.3 Discharge current and discharge voltage at C=1μF, T=16°C, Uc=11kV
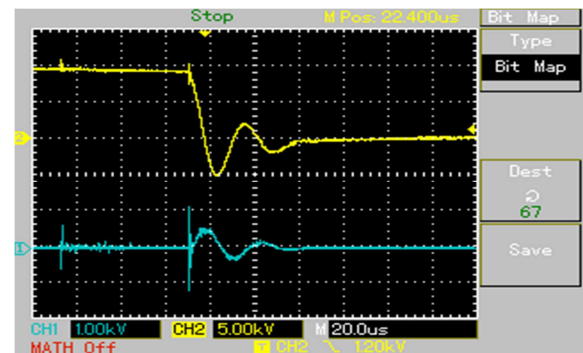
Fig.4 Discharge current and discharge voltage at C=3μF, T=60°C, Uc=11kV
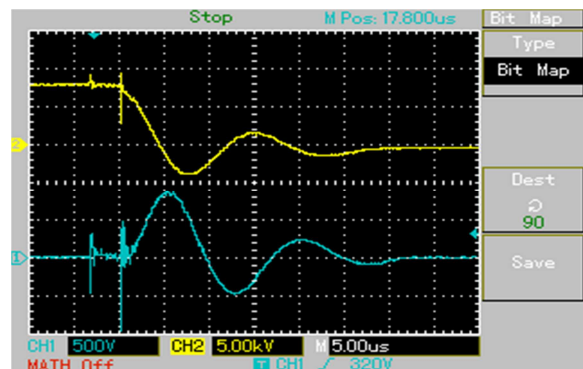
Fig.5 Discharge current and discharge voltage at C=1μF, T=90°C, Uc=9,5kV

From the conducted experiments it can be noticed that the change of the predischarge period has a random character and does not depend on the change of the liquid medium temperature. This fact is due to the different parameters of the liquid medium before the appearance of each separate discharge pulse, which depend on different factors. The predischarge time does not depend on the capacitance of the capacitor battery and the voltage U_c [4].

III. ANALYSIS

The experimental results have been analyzed by using the following methodology:

- Defining the discharge process duration;
- Recording the values of the discharge current and discharge voltage for all maximums of the positive half-wave and the moment of their appearance from the high voltage periodically attenuating oscillations;
- Drawing the wrap curves from the amplitude values of the discharge current and voltage;
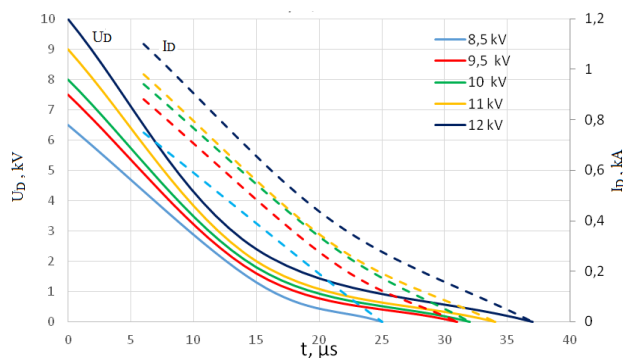- Calculation of the liquid medium resistance R_L on the basis of the obtained curves.

In Fig.6 are presented the wrap curves of the discharge current and discharge voltage for some values of the parameters C, U_c and temperature of the liquid medium.

a)   C=1µF and T=16°C

b) C=3µF and T=60°C.
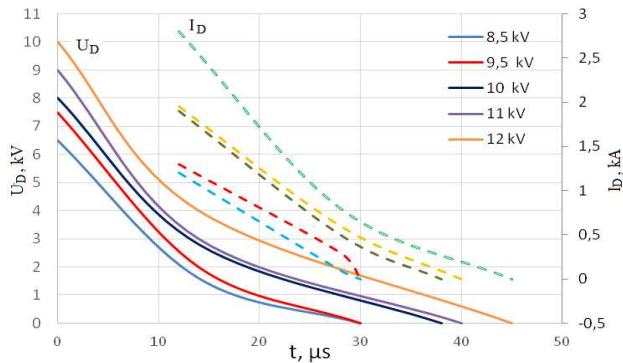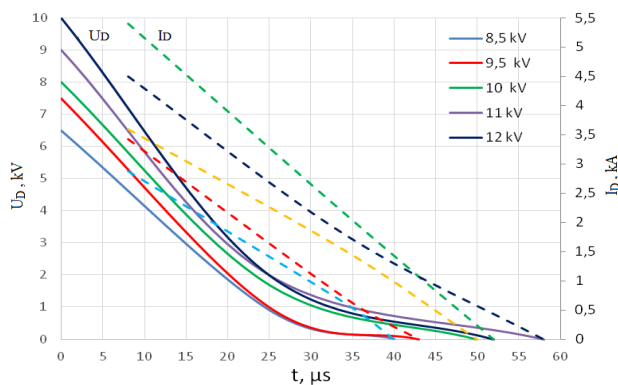
c) C=1µF and T=90°C.

Fig.6 Wrap curves of the amplitude values of the discharge current and voltage

When increasing the capacitance of the work capacitor battery and the voltage, to which it is charged, the values of the discharge current and discharge voltage increase, too. That is due to the higher energy, accumulated in the capacitor battery.

The discharge process duration shortens with increasing the temperature and decreasing of C, because of the higher electrical conductivity of the liquid medium.

The alteration of liquid resistance at the arising of a high voltage periodically attenuating discharge pulse is defined on the basis of the wrap curves – Fig.7. For comparison, the results are presented for the same values of C, Uc and the temperature T as in Fig.6 a,b,c.

The values of the resistance vary within the range $0,39 \div 10,5\Omega$.

a) C=1µF and T=16°C

b) C=3µF and T=60°C.
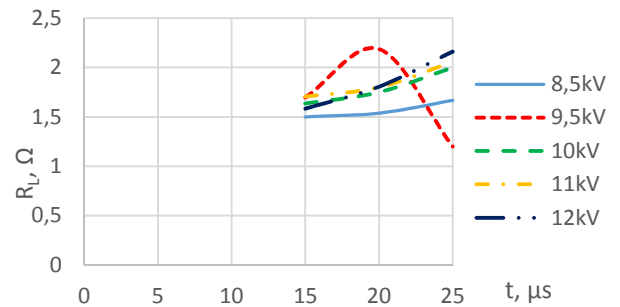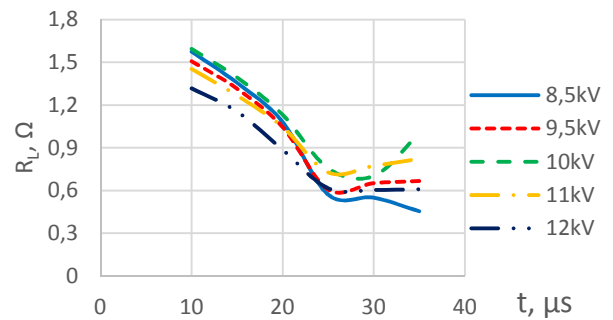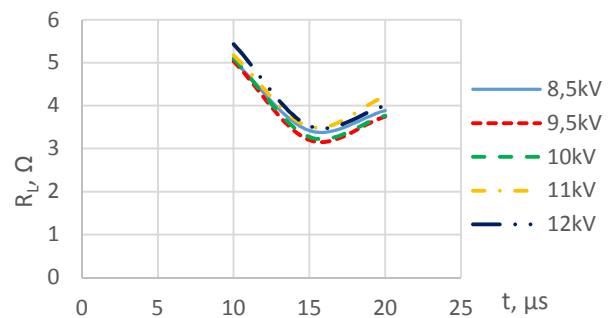
c) C=1µF and T=90°C.

Fig.7 Resistance of the liquid medium during the discharge process

At higher capacitance of the work capacitor battery it is observed decreasing the resistance and changing of the curve character – Fig.8.
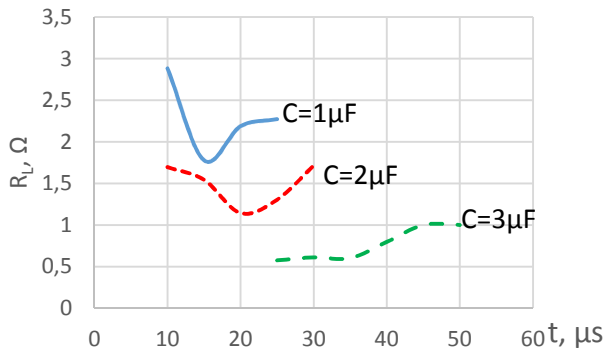
Fig.8 Resistance of the liquid medium at T=34ºC and Uc=10kV

For a certain type of the characteristic of the resistance alteration at arising of the high voltage periodically attenuating discharge pulse a model for analysis by approximation with a suitable mathematical function is proposed [5].

During high voltage periodically attenuating discharge in liquid, the main part of the energy is released in the first half-wave of the process. The resistance alteration depending on the temperature during this period for Uc=9,5kV is presented in Fig.9.
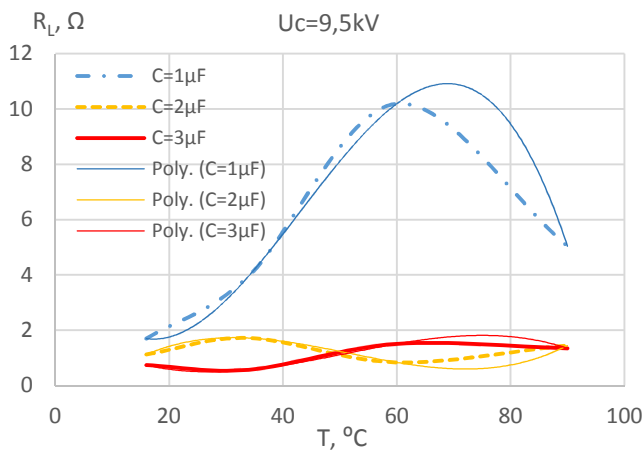
Fig.9 Resistance of the liquid medium for the first half-period of the discharge process

The same curves are observed and for the other values of the voltage U_c.

The curves from Fig.9 could be approximated with sufficient accuracy with third order polynomials (shown with thinner lines next to each of the experimental curves). The respective equations are:

- for C=1μF

$$y = -1{,}34*10^{-4}x^3 + 0{,}017314x^2 - 0{,}477706x + 5{,}456026 \quad (2)$$

- for C=2μF

$$y = 34*10^{-6}x^3 - 0{,}005214x^2 + 0{,}227871x - 1{,}323895 \quad (3)$$

- for C=3μF

$$y = -24*10^{-6}x^3 + 0{,}003660x^2 - 0{,}145547x + 2{,}228896 \quad (4)$$

The relationship between the polynomials coefficients and the values of C and U_c should be subjected to future analysis.

The approximation with a mathematical function gives a possibility for preliminary assessment of the liquid medium resistance. This is necessary for the design of devices for control of high voltage discharge pulses in liquids and for sizing of the charging and discharging circuits.

IV. CONCLUSION

A high voltage pulse discharge in liquid (water) has been experimentally studied as:

- the energy characteristics of the periodically attenuating discharge process (discharge current and voltage) are taken down at parameters: capacitance of the work capacitor battery; voltage, to which it is charged to and the temperature of the liquid medium;
- the liquid resistance alteration for different temperatures of the water during the discharge process is defined.
- it is proposed a description of the liquid resistance alteration with a mathematical function - third order polynomial.

The received results and dependencies can be used for sizing of systems for generation of high voltage discharge pulses for different technological applications.

ACKNOWLEDGEMENT

REFERENCES

[1] Anpilov A., E. Barkhudarov, Y. Bark, Y. Zadiraka, N. Chrstofi, Y. Kozlov, I. Kossyi, V. Kop'ev, V. Silakov, M. Taktakishvili, S. Temchin, 2001 Electric discharge in water as a source of UV radiation, ozone and hydrogen peroxide. J. Phys. D: Appl. Phys. 34, pp.993–999.

[2] Barudov S., M. Ivanova, R. Dimitrova, Study of the commutation capabilities of a controllable air discharger in circuit of high-voltage pulse discharge. Annual proceedings of the Technical University of Varna, 2011, Bulgaria, vol. I, ISSN: 1311-896X, pp.15-20.

[3] Bogomaz A., V. Goryachev, A. Remmenui, Ph. Rutberg, Efficiency of pulsed electric discharges in water disinfection. Lett. J. Theor. Phys. (JTP) 17, 1991, pp.65–68.

[4] Dankov L., S. Barudov, Study of high electrical discharges in water for use in a device for cleaning scale and scallops, Marine scientific forum 16-17th May 2013, Naval Academy "N.Y.Vaptsarov", Bulgaria, ISSN 1310-9278, pp.124-129.

[5] Ivanova M., S. Barudov, Modelling of high voltage periodically attenuating discharge in liquid medium with controllable high voltage switch trigatron, Marine scientific forum 16-17th May 2013, Naval Academy "N.Y.Vaptsarov", Bulgaria, ISSN 1310-9278, pp.118-123.

[6] Ushakov V., Impulse Breakdown of Liquids, 2007, ISBN: 978-3-540-72759-0.

Limiting of lightning overvoltages in the electrical Substations 110 kV

Margreta Vasileva[1] Neli Velikova[2] Yordan Ivanov[3] and Danail Stanchev[4]

Abstract – **When a lightning strikes on a power line surges are generated, which are being transferred by it. In the actual circuitry the wave is meeting discontinuities, which are based on various combinations of elements, resulting in occurring processes of multiple refraction and reflection of the wave that alter its shape and amplitude. In the report are presented the results of the model study of lightning stroke on overhead line and its impact on the equipment of a 110 kV substation switchgear including the MOSA.**

Keywords – **metal-oxide surge arrester; lightning protection, high-voltage substation.**

I. INTRODUCTION

Electrical equipment in electrical substations (ES) with voltage 110 kV, which are connected to overhead lines are protected from atmospheric surges by metal-oxide surge arresters (MOSA). The MOSA's electrical and mechanical properties must comply with the operational conditions of the switchyard. Surge arresters are placed at the inlet of each power line; each bus system of PS, as well as in front of the power transformers on the supply side. It is also necessary to choose the place of mounting of the surge arresters to make additional calculations related to their selection. [1,2,6].

The aim of the report is to present the results of the model study of lightning stroke on overhead line and its impact on the equipment of a 110 kV substation switchgear and the protective action of MOSA.

In the MATLAB system a simulation model is created of an electric system with six air accessions, which uses real data of substation 110 kV. Through the model is being reviewed the influence of lightning overvoltages on the electrical equipment in the substation. The cases are considered when atmospheric surges enter the substation through the longest power line in normal operation when protected with MOSA and the same case, but without protection equipment. A case of emergency operation of the substation is considered, in which three of its power lines are dropped out.

[1]Margreta Vasileva is with the Electrical Engineering Faculty of the Technical University of Varna, 1 Studentska St, Varna 9010, Bulgaria, E-mail: m.vasileva@tu-varna. bg.

[2]Neli Velikova is with the Electrical Engineering Faculty of the Technical University of Varna, 1 Studentska St, Varna 9010, Bulgaria, E-mail: nely.velikova@gmail.coom

[3]Yordan Ivanov is with the NOD Varna ESO, Varna 9010, Bulgaria, E-mail:yivanov@abv.bg

[4]Danail Stanchev is with the Electrical Engineering Faculty of the Technical University of Varna, 1 Studentska St, Varna 9010, Bulgaria, E-mail:dstanchev1990@gmail.com.

II. MODEL STUDY OF LIGHTNING PROTECTION OF 110 kV SUBSTATION

A substation is being viewed with six air accessions, namely: "Beton" - 2.51 km; "Avrora" - 13.22 km; "Dobrina" - 14.15 km; "Marciana" - 7 km; "Gigant" - 4.44 km; "Gitnica" - 11.03 km. The substation is whit a single section bus system and has two 110/20 kV power transformers with power 25 MVA. The protection of the facilities is carried out by two metal oxide surge arresters with a protective level U_p = 265 kV, which are being attached to the bus system.
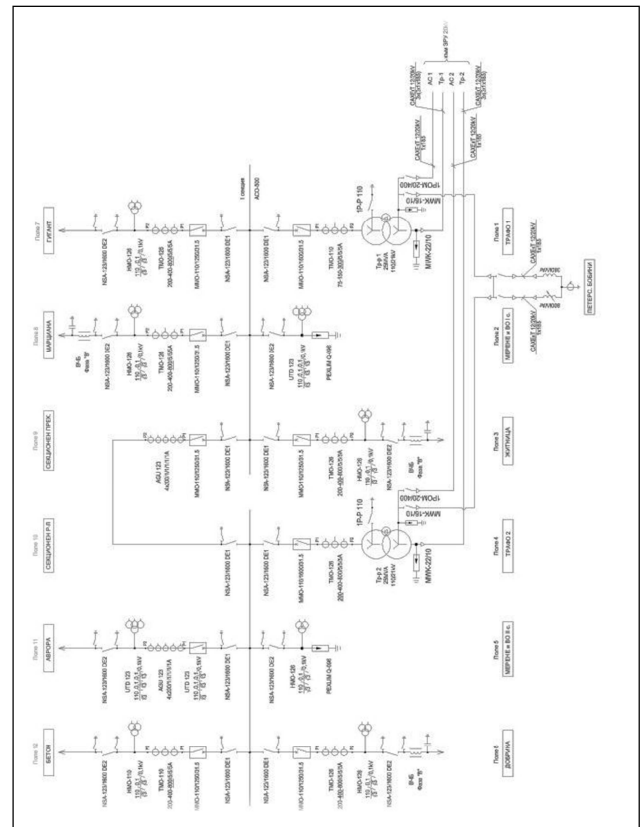

Fig.1 Single-line diagram of the studied 110 kV substatio

The mathematical equations for the definition of the parameters of the substitute circuit of the power transformer (R_T, X_T, z_T, G_T, B_T and others) are described in the literature [4,5].
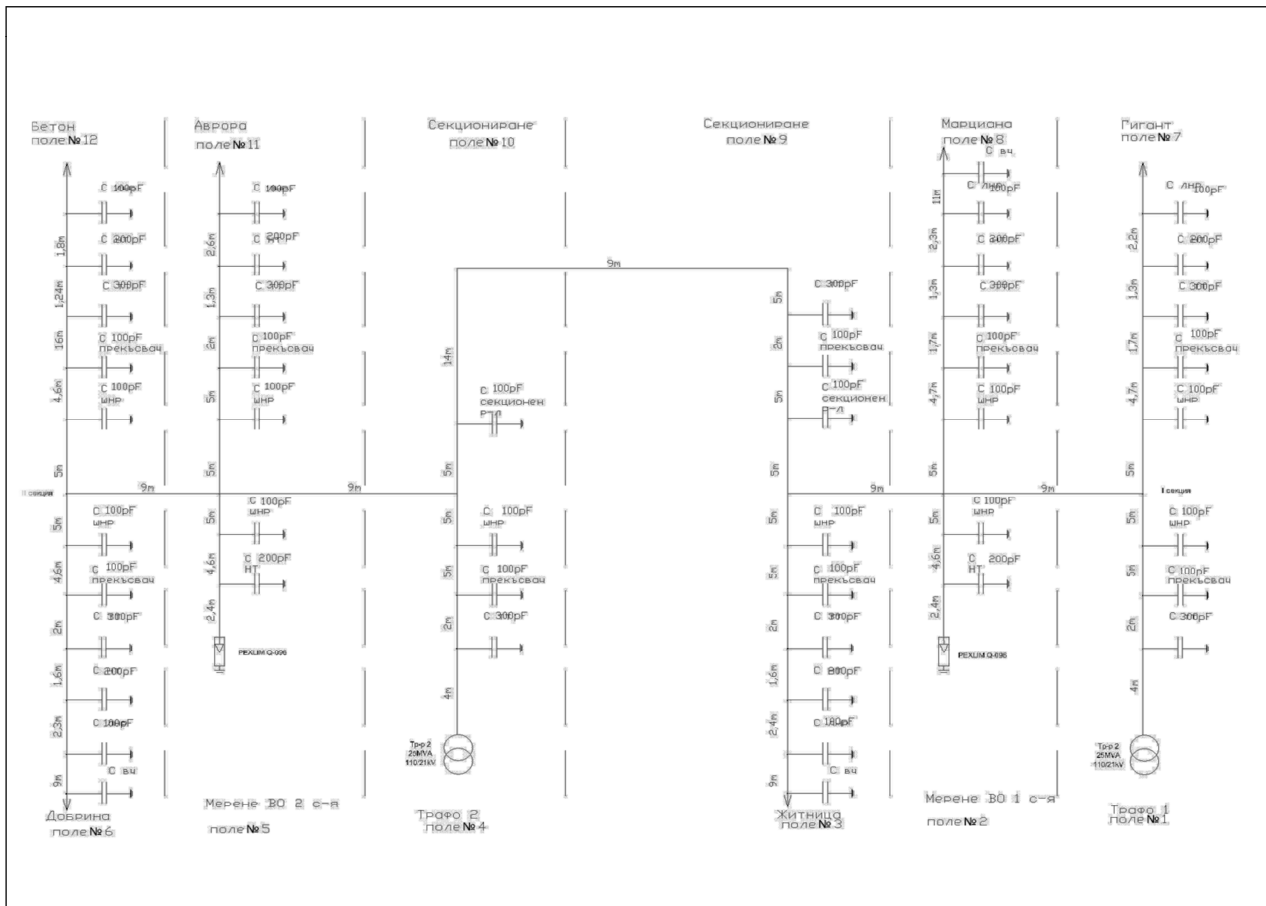
Fig. 2. Equivalent electrical circuit of the 110 kV substation

Fig. 2 is a full replacement scheme of a 110 kV switchyard. It is shows the distance in meters between devices and nodes, and the numbers to the capacitors are the values of their capacity in pF.

Established three-phase model scheme of 110 kV substation in the program MATLAB, in which the equipment is being presented as capacitors with a specific capacitance [3], and power lines and substation buses are modelled by equivalent electrical circuit of line with distributed parameters. Lightning is modelled as a current source with an amplitude of 80 kA and a shape 1/10 μs. The parameters of the replacement schemes are calculated according to [4].
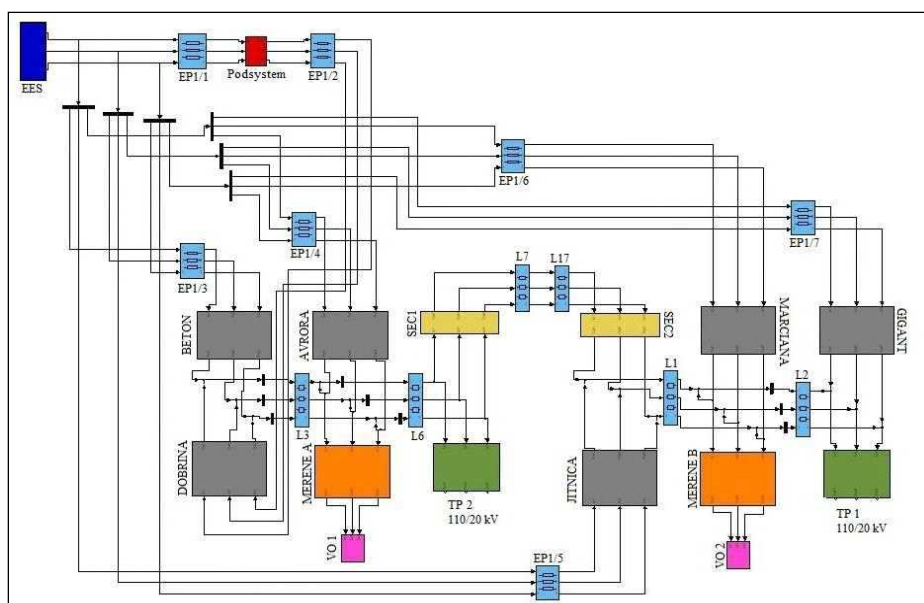
Fig. 3. Simulation model of the Substation 110 kV in a Matlab

III.   SIMULATION RESULTS

III.1. Study on impact o a lightning stroke arising in power lines, on the electrical equipment in the substation 110 kV without MOSA

The voltages in the three phases are being controlled - in the place of the lightning stroke and on the power transformers.

A. Damage from lightning on a phase conductor of power line "Dobrina".

The case of a lightning strike in phase B of overhead power line "Dobrina" is being examined. The affected area is on a distance of 100 m away from the substation. The results are shown in fig. 4. and fig. 5.
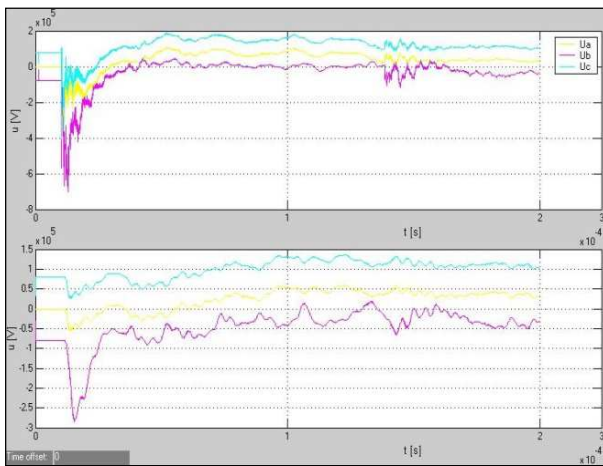
Fig. 4. Voltages in case of lightning stroke on phase B of "Dobrina" without MOSA in the ES (a) the point of the stroke and (b) on the first transformer terminals
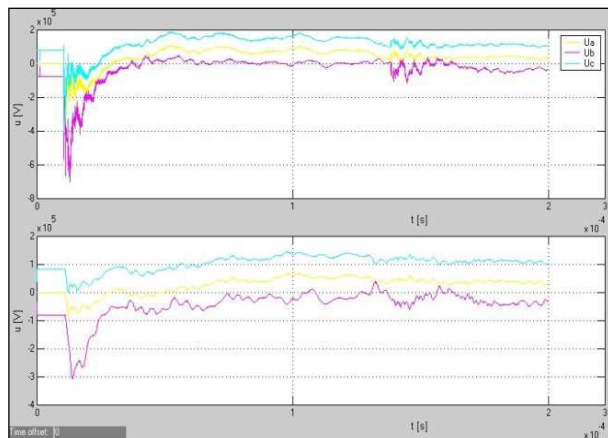
Fig. 5. Voltages in case of lightning stroke on phase B of "Dobrina" without MOSA in the ES (a) the point of the stroke and (b) on the second transformer terminals

B. Damage from lightning on the phase conductor of overhead power line "Beton".

The case of a lightning strike in phase B of the overhead powerline "Beton" is being examined. The affected area is on a distance of 100 m away from the substation. The results are shown in fig. 6. and fig.7.
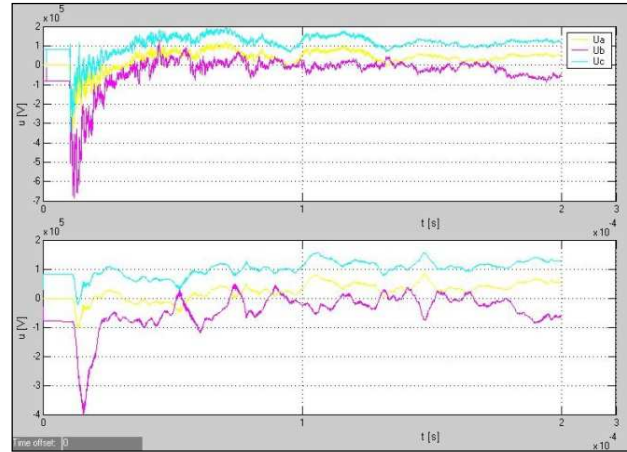

Fig. 6. Voltages in case of lightning stroke on phase B of "Beton" without MOSA in the ES (a) the point of the stroke and (b) on the first transformer terminals

Fig. 7. Voltages in case of lightning stroke on phase B of "Beton" without MOSA in the ES (a) the point of the stroke and (b) on the second transformer terminals

The overvoltages affecting the insulation of the power transformers, if they are not being protected, may exceed their insulation level [5]. It is therefore necessary to be restricted.

III.2. Study on the protective effect of MOSA in case of impact of lightning stroke arising in power lines , on the electrical equipment in the substation 110 kV with MOSA

The situations A and B of p. III.1., are being discussed, but in the case if there is MOSA available in the ES. The results are shown on fig. 8, fig. 9, fig. 10 and fig. 11.

The voltages in the three phases are being controlled - in the place of the lightning stroke, on the power transformers and on the MOSA.
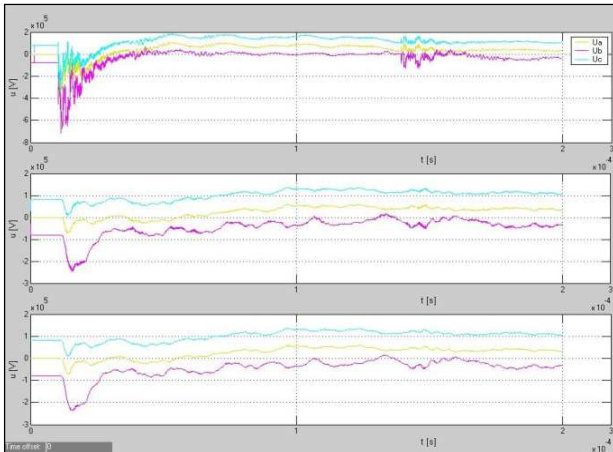

Fig.7. Voltages in case of lightning stroke on phase B of "Dobrina" with MOSA in the ES (a) the point of the stroke; (b) on the first transformer terminals (c) on the surge arrester
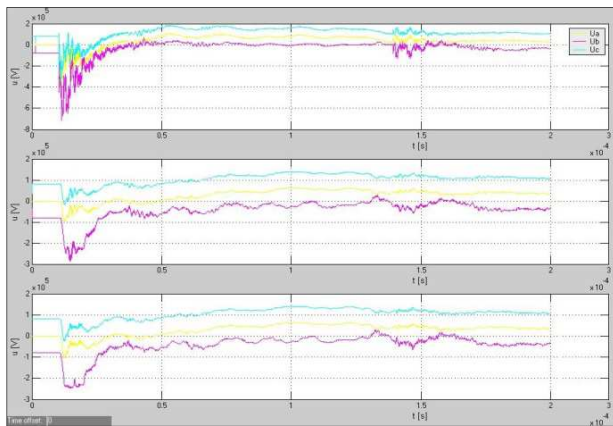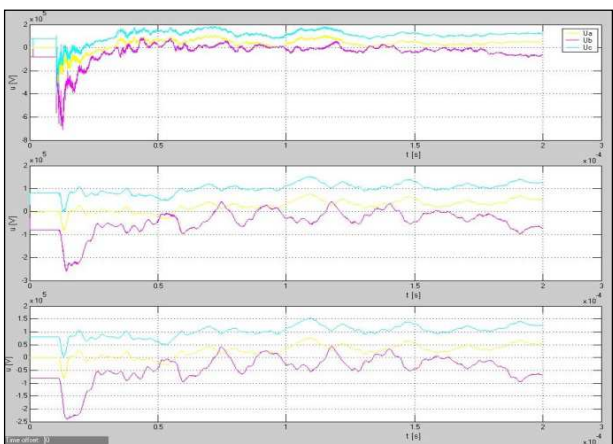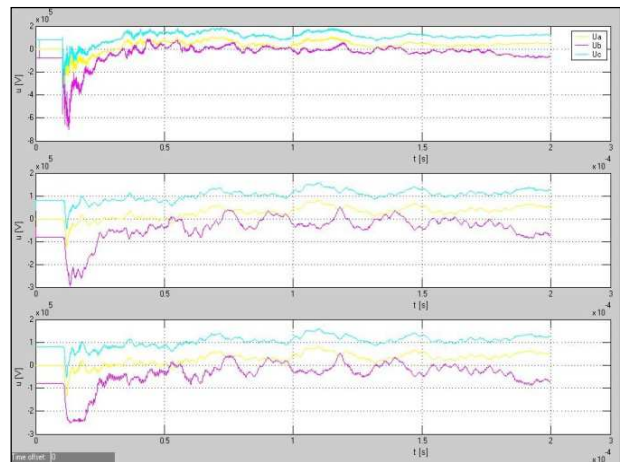

Fig.9. Voltages in case of lightning stroke on phase B of "Dobrina" with MOSA in the ES (a) the point of the stroke; (b) on the second transformer terminals (c) on the surge arrester

Fig.10. Voltages in case of lightning stroke on phase B of "Beton" with MOSA in the ES (a) the point of the stroke; (b) on the first transformer terminals (c) on the surge arrester

Fig.11. Voltages in case of lightning stroke on phase B of "Beton" with MOSA in the ES (a) the point of the stroke; (b) on the second transformer terminals (c) on the surge arrester

From the results it is apparent, that MOSA can ensure the reliable operation of the equipment and the energy system as a whole. For this purpose it is necessary to make the correct choice of the technical parameters and the place of their installation.

IV. CONCLUSIONS

In operation of the substation with six accessions the received surges on the facilities are under their insulation level.

Risk of damage to the equipment in the substation on impact of an atmospheric surge is present in the cases of a smaller number of accessions connected to the bus bar system.

The simulation model of the electric system, presented in the report can be used to study wave processes in electrical substations.

The developed model of the study on the protective effect of metal oxide surge arresters in electrical substation can be used for a more precise choice, taking into account the configuration of the scheme and the participating elements on the effect of the atmospheric overvoltage.

REFERENCES

[1] M. Vasileva, "High Voltage technique", Varna, TU-Varna, 2007.
[2] Regulation 3 of Structure of electrical installations and power lines, 2004.
[3] M. Goldstajn, J. Korovin,   A. Prokudin "Lightning and switching overvoltages in electrical power systems", 2010.
[4] K. Gerasimov, J. Kamenov, "Modeling in power systems", Sofia, 2007.
[5] IEC 60071-1, 2006. Insulation co-ordination in highest voltage systems. Definitions, principles and rules.
[6] S. Barudov, E. Bardov „System for management of discharges", Academy of Sciences   Azerbaijan,„"Problems of power systems" №2-3, pp. 96÷100, ISSN 1302-6461,2008 .

Design and Construction of a Laboratory SCADA System

Yulian Rangelov[1], Aleksandar Avramov[2] and Nikolay Nikolaev[3]

Abstract – **During the last decade, the industry has become very "digitalized". Industrial, commercial and domestic electricity consumers are becoming part of the digital revolution. In the near future even the light bulb is expected be able to control and monitor itself. The electric power engineering is also subject to digitalization on a global scale. Every electric power facility is either fully or partly equipped with automated controls, numerical relay protections, multifunctional digital measurement devices and other intelligent electronic devices. The energy industry has been looking for years to find a universal approach to deal with the unification of substation automation and communication devices manufactured by different companies. The standardization institutes in Europe and America have developed the common standard IEC61850, which allows the integration of the communication, the information, the control and all other subsystems.**

This paper is dedicated to the design and development of electronic system for supervision, control and data acquisition (SCADA) of an operational switchgear. It briefly presents basic aspects regarding the development of a SCADA system for a realistic and fully-operational switchgear in the laboratory "Power Plants and Substations" at Technical University – Varna.

Keywords – **Electric Power System, substation, SCADA, standard.**

I. INTRODUCTION

Department "Electric Power Engineering" in Technical University of Varna has a unique laboratory named "Electrical Power Plants and Substations", which is built just like real and operational switchgear [1]. The busbar system is single and is divided into three sections (fig. 1), located in two neighboring rooms (area approx. $120 \, m^2$) [2]. Two remotely-controlled synchronous generators are connected to the first section. They can be synchronously connected to the power grid or to each other. On the control panel, there are control buttons, signaling lights and various measurement equipment to supervise the operational variables. The turbines are replaced by variable-speed DC motors, which are powered by another motor-generator aggregate, installed in the same room.

Two-power transformers are connected to the second section and represent a physical implementation of a substation.

The substation feeds different consumers – two rotary fans with AC motors, a complex *R-C* load and a complex computer-controlled active power load.

Section 3 is a physical implementation of a classical indoor 10 kV switchgear which practically operates at 0.4 kV. There are three separate cells – laboratory switchgears (LSG) – which are powered by this section. A circuit-breaker with motor-spring mechanism, model SCI4, along with a disconnector are mounted in LSG1. A discrete active power load is fed by a vacuum circuit-breaker (model Tavrida), located in LSG2. LSG3 is a measurement field and has different types of voltage transformers connected to it. All three sections are interconnected by circuit-breakers and disconnectors.

The switchgears of section 1 and 2 are designed to provide operational flexibility, i.e. the generators can be connected directly to the power grid or operate in island mode, while energizing complex loads in the substation.

The laboratory is connected to the external grid by two cable feeder (W1 and W2) which originate from the near-by transformer station of the faculty. There are several digital multifunctional power meters ABB model DMTME, connected to main nodes of the lab's switchgear. There is a total number of 23 remotely-controlled circuit-breakers, 24 disconnectors and 5 power switches.

The following report presents the design and construction of an automation SCADA system, which allows to control and supervise the switching devices and to observe important electrical variables related to the operation of the switchgear (fig. 1). Considering the available resources, the SCADA system should provide graphical user interface (GUI) which allows local or remote control via the web. The control and signaling functions are performed by a local controller based on the Arduino Uno [3] open electronic environment and specifically designed input-output interface circuit boards. The operational data, measured by the ABB DMTME devices is collected via RS485 communication and Modbus protocol, and the converted to LAN interface by an ABB CUS device are sent to the computer station. The input-output data of the local controller can be transmitted over the internet outside the laboratory. The computer interface is developed in Matlab.

[1]Yulian Rangelov is with Department "Electric Power Engineering" at Technical University of Varna, Studentska str. 1, Varna 9010, Bulgaria, E-mail: kkgerasimov@tu-varna.bg

[2]Aleksandar Avramov is with Department "Electric Power Engineering" at Technical University of Varna, Studentska str. 1, Varna 9010, Bulgaria, E-mail: y.rangelov@tu-varna.bg

[3]Nikolay Nikolaev is with Department "Electric Power Engineering" at Technical University of Varna, Studentska str. 1, Varna 9010, Bulgaria, E-mail: n.nikolaev@tu-varna.bg
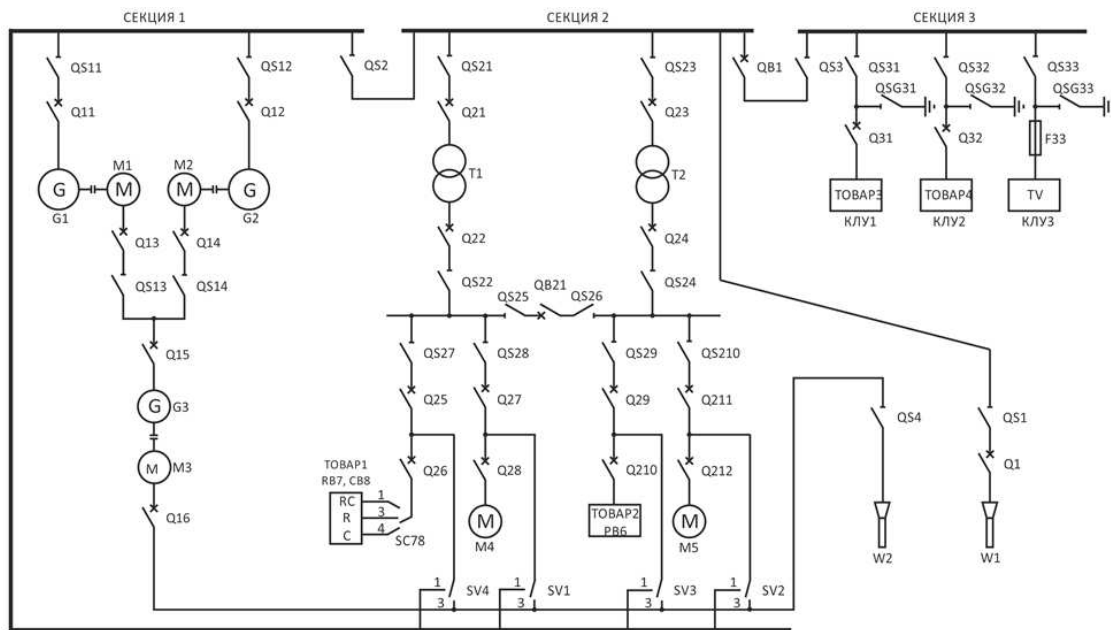
Fig. 1. Single-line diagram of laboratory "Electrical Power Plants and Substations

II. DESIGN OF AN AUTOMATED SYSTEM FOR CONTROL AND SUPERVISION OF LABORATORY "ELECTRICAL POWER PLANTS AND SUBSTATIONS"

A. General Concepts

The approach to build an automated control system of the laboratory is based on the Bulgarian regulations and standards [4, 5]. What is built is the so-called informational control complex with a computer system for real-time telecontrol, telesignalling and telemeasurement. This is basically a system for data acquisition, supervision, data processing and control of switching devices – SCADA [4].

A telecontrol system is developed for switching devices which are technically available for remote control. The switch-on and –off commands are released via individual general-purpose relays for each individual switching device.

The telesignalling provides information about the state of the switching devices and thus gives valuable information about the system's topology. The device state is observed be means of its own auxiliary contact. If there is no available auxiliary contact, additional relay is connected to provide more contacts.

The telemeasurement system transmits information about the important electrical quantities.

The project documentation complies with the standards and regulations regarding the symbolic and numerical designations [6].

To provide for clear presentation of the paper's ideas, it is necessary to depict all electrical diagrams for control and signaling of every switchgear component. However, these diagrams are too many and cannot be presented due to the restricted length of this report.

The local controller is based on Arduino Uno board, extended with an Ethernet shield, four custom-designed boards with 16-inputs and four boards with 16-outputs, main board that controls the I/O of the other boards. All measurement units are connected in a RS485 network, which is accessed from the computer via TCP/IP gateway ABB CUS. All the boards are located in a dedicated SCADA panel.

The Arduino board and the I/O boards have independent 5 V power supplies. The TCP/IP gateway is powered with 220 V via a miniature circuit-breaker.

In the SCADA panel there is additional 24 V source which supplies the general purpose relays.

Fig. 2 depicts the main structure diagram and the connections between the local controller and the other boards.
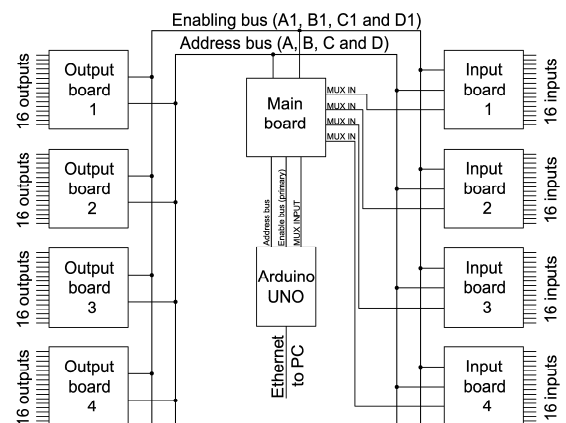

Fig. 2. Main structure diagram and connections between the boards

B. Main Board Design

Several of the Arduino outputs form an address bus (A, B, C and D), an enable bus (A1, B1, C1 and D2) and an input bus (MUX IN) which transmits the state of the switching devices to the Arduino board. The address bus sets a binary code to

choose among the inputs or outputs of the demultiplexors (DEMUX) or multiplexors (MUX), chosen by the enable bus.

C. Input Boards Design

Every input circuit-board has a TTL DEMUX chip SN74150J and an EL817 optocoupler. The DEMUX collects the binary signals from the optocoupler. The optocouplers have independent power supply in order to avoid the penetration of high potentials in the electronic part of the SCADA system.

D. Output Boards Design

Each output boar consists of one SN74154N MUX chip, three SN7404N inverter chips, 16 relays HM4100F and two ULN2803A relay drivers. Via the address and enable bus, the Arduino board chooses which MUX output it will activate and the respective relay will either switch ON or OFF the device. The activation command holds for 0.5 seconds to make sure that the tripping or closing action will be successful.

A dedicated software is written for the Arduino platform. The Arduino itself is a bridge between the I/O boards and the Matlab computer software. The diagram of the Arduino's algorithm is depicted in fig. 3. The code starts with initialization of the Arduino's inputs, outputs and the TCP/IP communication with Matlab. Then the program checks whether Matlab is connected. If yes it reads the variable *matlabMessage* and if it is equal to 255 then the Arduino reads all the inputs and sends information about the state of the switching devices to Matlab. If the message is less than 255, then, depending on the number code, the Arduino decides which output relay it activates for 0.5 seconds.
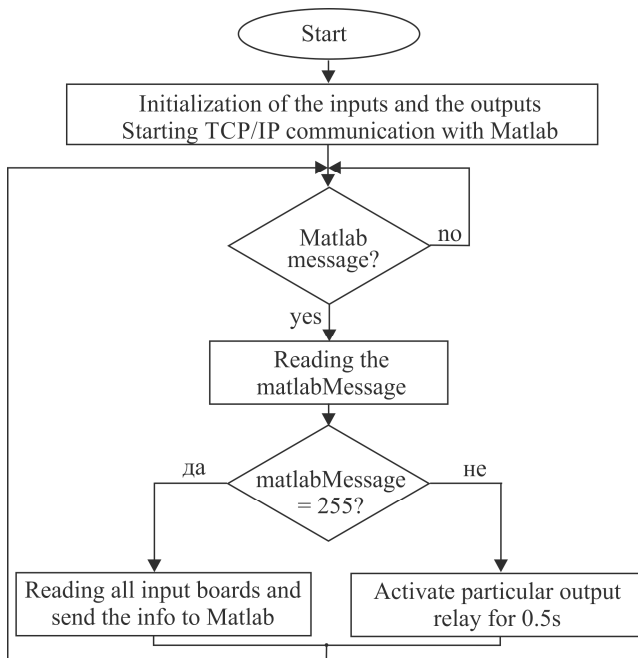

Fig. 3. Algorithm diagram of the Arduino board

Fig. 4 depicts the algorithm diagram of the Matlab program. It starts by loading the graphical user interface. At the same time the configuration file is loaded and the objects from classes *DMTME* and *SWITCH* are created. Then follows the setup of the communication with the Arduino board and the measurement devices. If there is a command to stop the SCADA software, the TCP/IP ports are closed and the connection is terminated. Next, the program checks if the user wants to switch a device. If *yes*, then a number code less than 255 is sent to the Arduino. After that, code 255 is sent to tell the Arduino to read all input boards and to return information about the status of the switching devices and the GUI is refreshed. Finally, the DMTME measurement devices are read and the GUI refreshed again.
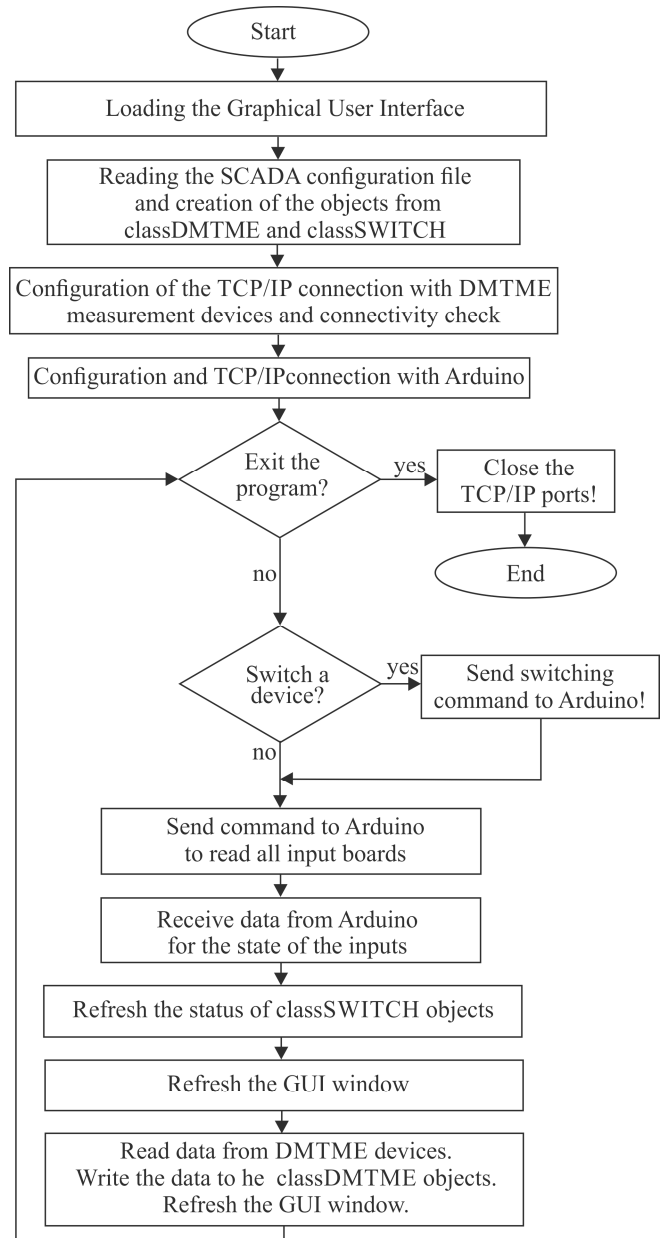

Fig. 4. Algorithm diagram of the Matlab computer program

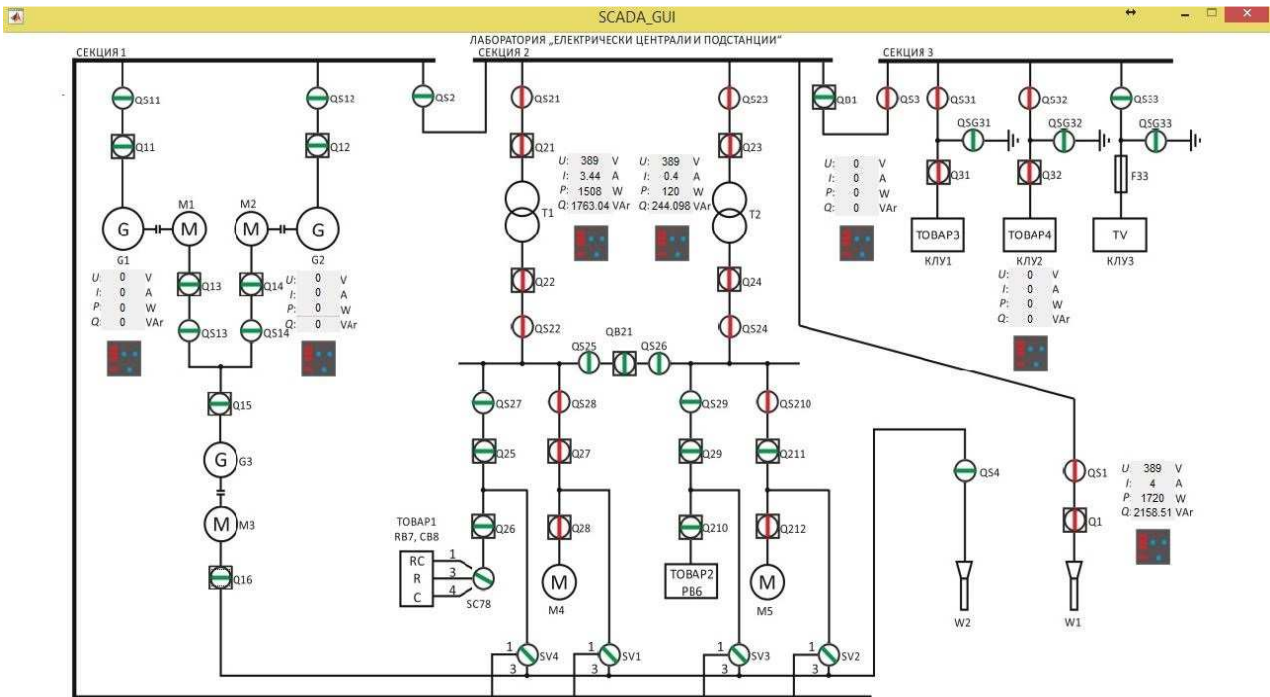Fig. 5 depicts the local controller and the final look of the SCADA panel.
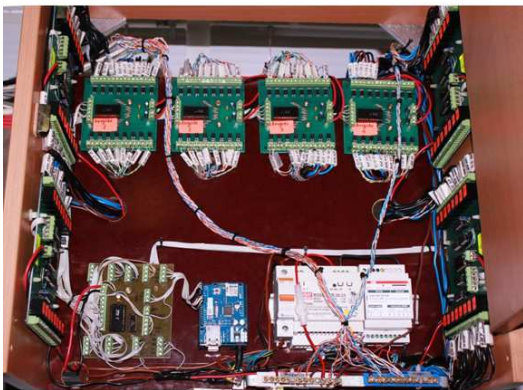
Fig. 5. Graphical User Interface of the SCADA

Fig. 6. Local controller in the SCADA panel

The GUI of the SCADA system is presented in fig. 6. As a first step, it is developed just enough to test all control, signaling and measurement functions. The object-oriented programming approach is used to make the software easily maintainable and extendable. The device symbol turn red when it's ON and green when it is OFF to make the visualization perceivable. The entire single-line drawing is depicted in single GUI window. Display panels are placed at the measurement points to allow the user to keep track of the voltage, the current and the power.

III. Conclusions

Here we draw the main conclusions of the paper:

1. Based on thorough study on existing internet reports, national regulations and standards an intelligent electronic platform is created, which we call a local controller. It implements all the inputs and output to interface the switching and measurement devices which are part of existing switchgear and makes connection to the operator's computer station.

3. Approximately 1000 meters of cables are used to connect every component of the switchgear (83 devices) to the local controller in laboratory "Electrical Power Plants and Substations".

4. A human-machine interface is created to observe and control the operation of the laboratory switchgear.

5. The paper shows that it is possible to build a low-cost fully-functional SCADA system with existing open-source microcontrollers, such as Arduino. The required devices and equipment is worth approximately 500 euro.

References

[1] (2015, 15 03.). Department "Electric Power Engineering" in Technical University of Varna. http://ee.tu-varna.bg

[2] Добрев, Г. Проектиране на лаборатория по електрическа част на електрически централи и подстанции. Дипломна работа. ТУ-Варна, 1982.

[3] (2015, 15 03.). Arduino, http://arduino.cc/en/Main/Products

[4] Наредба №3 за устройството на електрическите уредби и електропроводни линии. Министерство на енергетиката и енергийните ресурси, обн., ДВ, бр. 90 и 91 от 2004 г., в сила от 15 януари 2005 г.

[5] Leupp, P., Cl. Rytoft. Special Report IEC 61850. The corporate technical journal.

[6] Rangelov, Y. Overview of the methods for symbolic and alphanumeric designation of the power and the control circuits in the Bulgarian power system. Сборник с научни доклади, ISBN 978-954-760-316-5. Проект: BG051PO001-3.3.06.0005. pp 83-91.

Algorithms for Precise Anticipation of Reactive Energy Savings in Compensated Power Systems

Milica Rašić [1], Milan Radić [1], Nikola Milosavljević [2], and Zoran Stajić [1]

Abstract – **This paper presents some problems that can appear during the precise determination of exact values of reactive energy savings in compensated power systems, as well as specific methods used for their solutions. Final results of the research, contained in this paper are obtained from a comparative analysis of results of reactive energy consumption monitoring, registered by two different measuring devices: one commercial energy meter and one smart measuring device that works as a controller of the accuracy of mentioned energy meter. Those measurements are performed at one customer of electrical energy, where compensation system has already been implemented. Knowledge of the exact values of energy savings is an important aspect of the ESCO concept and therefore algorithms presented in this paper, could have high importance in energy efficiency.**

Keywords – **Algorithms, reactive power, energy savings, compensated system, energy efficiency.**

I. INTRODUCTION

The transfer of reactive power through transmission and distribution lines and transformers has many disadvantages with respect to the power system construction and operation. Active losses are increased and a higher cross-section of lines is sometimes required. Reactive losses are increased, as well. Furthermore, the voltage drops in the system are increased, resulting to the need to choose a higher regulation range of tap-changers for the transformers. Therefore, consumed reactive energy is charged. Power distribution companies use different billing methods: pricing based on reactive energy, based on apparent energy and different amounts for low and high tariffs.

Energy efficiency and power quality over past decades became one of the most important global issues. In power distribution networks one of the most common improvement measure is reactive power compensation. There are several methods and types of compensation in practice, but this paper considers use and effects of the simplest shunt type compensation.

For ESCO model financing of energy efficiency projects, the main stuff based in every project is predicting of total

[1]Milica Rašić, Milan Radić and Zoran Stajić are with Faculty of Electronic Engineering, University of Nis, Aleksandra Medvedeva 14, 18000 Niš, Serbia, e-mail: milicarasa@gmail.com, milan.radic@elfak.ni.ac.rs, zoran.stajic@elfak.ni.ac.rs
[2]Nikola Milosavljević is with Public Utility Company "ED Centar d.o.o.", Slobode 7, 34000 Kragujevac, Serbia, e-mail: nikola.milosavljevic@edcentar.com

available budget in correlation with potentially realized savings [1].

Power utility companies take registered values from the power utility meters, as elements in the preparation of bills for electricity. However, during the anticipation of realized savings, according to ESCO concept, it is necessary to have information about monthly consumption of reactive energy in the certain consumer system, as well as values of uncompensated reactive energy in the same system, after the implementation of system for compensation. The great importance of this information is in the evaluation and subsequent precise determination of realized savings and expressing of their economic effects.

II. FORMULATION OF THE PROBLEM

The largest number of analysis was based on collected data, recorded by systems for the electrical energy consumption monitoring. These systems have already been placed in several industrial facilities, which are in possession of one customer and each of those systems consists two parts used for different purposes. One part of smart measuring device is used for registering all current values of electrical parameters at the place of implementation: phase and line currents and voltages, frequency, active, reactive and rated power, consumed active and reactive electrical energy, as well as calculation of their effective values. Another part of this device is used to record measured results and to control electrical energy consumption [2, 3]. Mentioned industrial facilities can be classified to the bakery industry.

In this paper, measurements in two different facilities are going to be analysed. At both measuring locations, systems for shunt compensation have been implemented. The main goal in this analysis is precise calculation of reactive electrical energy savings at both places, as well as the calculation of total reactive electrical energy, that would be taken from the electrical energy system (EES), if the compensation system wasn't implemented. The main difference between these measuring points is the order of placing utility energy meter, smart control measuring device and capacitor battery to the customer's electrical supplying. In both cases, some problems can appear during specified calculations. In order to skip these problems certain algorithms were used.

At the first measuring point utility energy meter, smart measuring device and shunt compensation system are placed in order, shown at Fig. 1. In this case utility energy meter measures the value of uncompensated reactive energy while control measuring device registers the value of total consumed reactive energy.
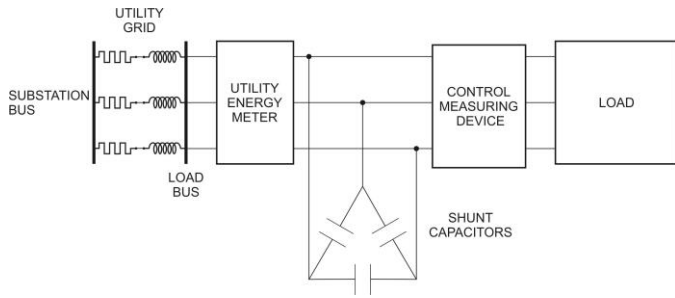
Fig. 1. System for measuring and reactive power compensation implemented at the first location

Time period for which calculations are going to be done is one month in accordance with period for which electricity supply companies deliver bills for consumed energy. Reactive electrical energy that can be saved using the shunt type compensation is:

$$Ec = Q_{bat_n} \cdot t \qquad (1)$$

where Q_{bat_n} is rated value of battery capacity, and t is period during which compensation system works. For this calculation, this period is one month (720 hours). If the proper operation of utility energy meter should be checked, it is possible to do it by simply adding the value that is registered by utility power meter and reactive energy saved by compensation system and compare with the value registered by smart measuring device. If the utility meter operates properly, these two values will be approximates.

This method can be applied only for the rough estimate where the only important thing is proper operation of utility energy meter. In all other calculations, it is necessary to take into the account dependence of reactive power compensated by capacitor battery on quality of voltage in electrical grid. This dependence is given as:

$$Q_{bat} = Q_{bat,n} \cdot \left(\frac{U_{mean}}{U_{bat,n}} \right)^2 \qquad (2)$$

where U_{mean} is mean value of the line voltage for each individual measurement, $U_{bat,n}$ is voltage for battery's optimal operation and Q_{bat} is real value of reactive power saved by compensation system for each individual measurement, as well. In attention of precise calculation of reactive energy savings and real value of consumed reactive energy before compensation, specified algorithms are used.

Measuring of electrical parameters by smart measuring device was performed at intervals of ten seconds during the period of six months. Results of those measurements were stored as .b files. In the beginning, it is necessary to import all measuring results to one of programmable packages. Parameters that have the special importance for following calculation are line voltages and total value of reactive power, given as a sum of phase reactive powers. The next step is to calculate mean value of the line voltage that affects calculation of real value of reactive power saved by

compensation system. Then, real value of capacitor battery should be determined by Eq. (2). This process should be done for each individual measurement performed during the considered period. Now, it is possible to calculate uncompensated reactive power by subtracting real value of compensated power from the value of reactive power registered by smart measuring device. This step should be done for each individual measurement, as well. The next step is to calculate total uncompensated reactive power during the period of one month as a sum of all real values of uncompensated reactive power for the same month, determined in the previous step. Finally, it is possible to calculate total uncompensated energy taken from EES as a product of total uncompensated power in one month and number of hours in the consider month (720). Analogously, total reactive energy taken from the grid during the same month can be determined as a product of total reactive power (sum of all individual values of reactive power registered by smart measuring device during the period of one month) and number of hours in the month. Real value of energy savings can be determined here by subtracting the value of uncompensated reactive energy from the value of total reactive energy taken from the EES.

In order to check utility energy meter's proper operation, the value of uncompensated reactive energy, calculated by described algorithm, should be compared with the value registered by utility power meter.

One of the initial ideas, on which is based ESCO concept is knowing the real value of reactive energy savings. Other advantages obtained from this algorithm are conversance of all electrical parameters, which are describing the quality of voltage and electrical energy delivered to the customer.

At the second measuring point utility energy meter, smart measuring device and shunt compensation system are placed in order, shown at Fig. 2.
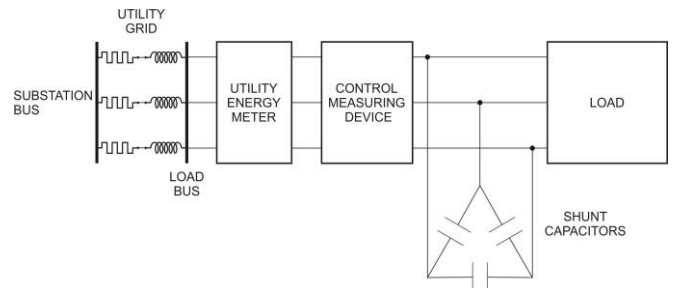

Fig. 2. System for measuring and reactive power compensation implemented at the second location

In this case both measuring devices register the value of uncompensated reactive energy. The parameter that should be determined is value of compensated reactive power, as a function of the line voltage. It is also important to calculate the value of reactive energy that would be taken from the electrical energy system in a case that compensation system wasn't installed.

As in the previous case, it is necessary to take into the account the dependence of reactive power on a quality of

voltage, given by Eq. (2). As at the first measuring location, measuring of electrical parameters by smart measuring device was performed at intervals of ten seconds during the period of six months. Results are also stored as .b files. The process of importing the recorded data and procedure for determining mean value of line voltage and real value of reactive power saved by compensation system is absolutely the same as in the first model. Parameters that should be determined here are values of total reactive power that customer's system would take from the grid in the case of absence of system for shunt compensation, for each individual measurement during the considered period. It can be done by adding the real value of compensated power to the value registered by smart measuring device. Therefore, total reactive power taken from the EES during the one month should be determined as a sum of values of total reactive power for each individual measurements which are performed during the same month. Finally, the overall reactive energy that could be taken from the EES, in the period of one month at this measuring location could be determined as a product of total reactive power and number of hours in the month.

This calculation is applicable for customers, where it is not possible to set smart measuring device close to the place of customer's supplying. Electrical parameters calculated by this algorithm have a great importance for electrical energy consumption planning that represents the basis for "smart grid" systems development [4].

III. RESULTS AND DISCUSSION

Algorithms presented at previous chapter were applied to data gained from monitoring by smart measuring devices installed at two compensated industrial systems. At both measuring locations parameters of electrical energy consumption were being monitored during the period of six months. Algorithm used for measurements at first measuring location gave results shown in Table I.

TABLE I

Months	REBC* [kvarh]	REAC – SMD** [kvarh]	REAC – UEM*** [kvarh]
January	3184.7	-3110.7	1262
February	3220.7	-2488.5	969
March	3592	-2720.0	1055
April	3465.9	-2721.4	1187
May	3840.8	-2529.9	1034
June	4343.9	-1795.2	990

* Reactive energy before compensation registered by smart measuring device
** Reactive energy after compensation – calculated by first algorithm
*** Reactive energy after compensation registered by utility energy meter

After comparison of values of uncompensated reactive energy registered by utility energy meter and obtained by presented algorithm, large differences between them can be observed. Calculated values of uncompensated reactive energy by the first algorithm are smaller than zero and indicate deep system's overcompensation. On the other side, utility energy meter registered large positive values of

uncompensated reactive energy. Those results indicate to incorrect operation of utility energy meter and its replacement.

Diagram of the power supply line voltage for one of considered months is shown in Fig. 3. Diagrams of reactive power savings and reactive power before and after using shunt compensation are shown in Fig. 4.
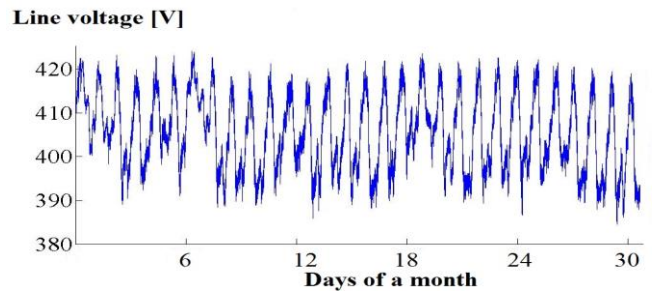

Fig. 3. Monthly diagram of power supply line voltage changes at first measuring location
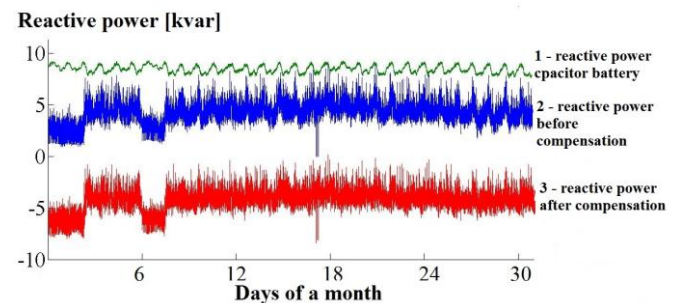

Fig. 4. Monthly diagrams of reactive power at first measuring location

Diagrams in Fig 3. and Fig. 4 show how in practice power of capacitor battery depends on the line voltage changes: with higher mean value of line voltage, compensated reactive power is also higher. It is also possible to notice that during the whole month, power of the capacitor battery is higher that the reactive power taken from the EES that is a proof of great overcompensation at this location.

Algorithm applied to measurements done at second customer's system obtained results shown in Table II.

TABLE II

Months	REBC* [kvarh]	REAC – SMD** [kvarh]	REAC– UEM*** [kvarh]
January	2062	-2479.4	0
February	1924	-2246.2	0
March	2100	-2524.4	0
April	2265	-2223.8	0
May	2509	-2121.6	0
June	2735	-1762.0	0

* Reactive energy before compensation – calculated by second algorithm
** Reactive energy after compensation registered by smart measuring device
*** Reactive energy after compensation registered by utility energy meter

After comparison of values of uncompensated reactive energy measured by smart measuring device and utility energy meter it can be concluded their deviation. Values registered by smart measuring device are negative as opposed to values registered by utility energy meter (zeros for all six months). This difference is result of utility energy meter's constructional impossibility of measuring in all four quadrants – it can't register values of reactive power smaller than zero. Instead of negative values it records zeros. Anyway, utility energy meter's proper operation is not brought to the suspicion at this measuring location.

On the other side, at both measuring locations great overcompensation is present. This is really positive impact to EES, because a large amount of reactive electrical energy was generated to the grid. However, this situation is not fair to the customer, because he doesn't receive any fee for mentioned reactive energy's generation, while his costs for compensation system could be lower if he replaced this system by another one with the capacitor battery of less rated power.

Anyway, main goal achieved by application of this algorithm to results of measurements at this location is overall electrical energy that could be taken from EES in a case shunt compensation's absence. This data is really important for consumption planning that represents one of main ideas of ESCO concept.

Diagram of power supply line voltage for one of considered months at second measuring point is shown in Fig. 5. Diagrams of reactive power savings, reactive power before and after using system of compensation at the same point are shown in Fig. 6.
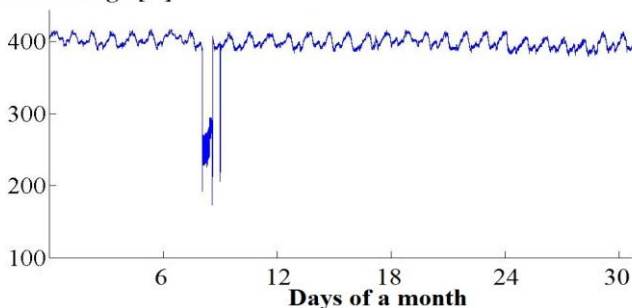

Fig. 5. Monthly diagram of power supply line voltage changes at second measuring location
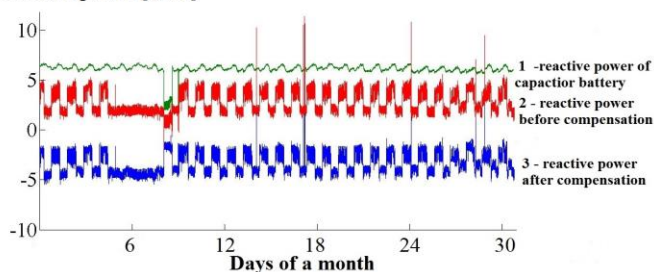

Fig. 6. Monthly diagrams of reactive power at second measuring location

Voltage drop shown in a diagram in Fig. 5. is followed with the lower value of reactive power that is compensated by shunt capacitors, presented in Fig 6. Greater reactive power of capacitor battery than a reactive power taken from EES represents overcompensation at customer's system. A fact that desired results could be achieved with the battery of less rated power had been really important for the customer. At both measuring locations, existing capacitor batteries were replaced with the batteries of less nominal power and in this way customer costs were reduced and these industrial systems became more energy efficient.

IV. CONCLUSION

Precise anticipation of reactive energy savings in shunt compensated power systems can be performed by using of presented algorithms. Naturally, the main condition for their applying is presence of measuring results, gained by smart measuring device for electrical parameters. Knowledge of energy savings is base for energy efficiency improvement and knowledge of electrical parameters in EES represents the first step in successful monitoring and controlling of power distribution grid.

Described algorithms were applied to the simplest systems for reactive power compensation. Further researches in this area could include other types of compensation systems, such as dynamic system, where automatic control for its optimal operation is installed. In these installations optimal capacitors batteries switching on and off should be included, achieving its longer lifetime and reducing of the customer's costs.

ACKNOWLEDGEMENT

REFERENCES

[1] S. Bobbino, H.Galvan, M. Gonzalez-Eguino, "Budget-Neutral Financing to Unlock energy Savings Potential: An Analysis of the ESCO Model in Barcelona", BC3 Working papers series, Barcelona, January 2013.

[2] M. Rašić, B. Kojičić, N. Simić, "The application of smart measuring devices for control of utility power meter's correct operation", *Proceedings of the 7th Student Projects Conference, IEEESTEC,* Niš, Serbia, 2014, pp. 105-108.

[3] N. Floranović, M. Radić, Z. Stajić, "Smart metering and Power Loggers as key element of effective energy efficiency improvement", *Proceedings of the XIII International Conference on SAUM,* Niš, Serbia, 2014, pp. 373-376.

[4] Z.P. Stajić, A. Janjić and Z. Smendić, "Power quality and energy losses as a key driver for smart grid platform development", *WSEAS/NAUN international Conference Corfu Island*, pp: 417-422, Greece, July 2011.

POSTER SESSION
ENGINEERING EDUCATION

Teaching with Technology: Promoting Knowledge Work Practices in Technology School Education

Stela Stefanova [1] and Tania Vasileva [2]

Abstract – **The paper considers challenges that face educational institutions to meet requirements of knowledge society. Based on brief introduction of trialogical learning approach, the educational practices to promote necessary competences are outlined. Against this background the effort done at Technology School "Electronic Systems" in changing pedagogical practices to promote collaborative knowledge creation in technology-rich environment is considered. The compulsory course re-designing from traditional face-to-face to project oriented applying modern online learning platform, cloud collaboration and social software is discussed. The results from pilots conducted in using collaboratively development of shared project are also highlighted.**

Keywords – **trialogical approach, project oriented online learning, cloud collaboration.**

I. INTRODUCTION

Knowledge has always been a factor of production, and a driver of economic and social development. The digitization of information and widespread of the Internet facilitate a new intensity in the application of knowledge to economic activity, to the extent that it has become the predominant factor in the creation of economic growth. The emergence of the knowledge society, building on the extensive influence of modern information and communication technologies (ICTs), create conditions for a fundamental reshaping of the global economy [1].

Education is of vital importance in the knowledge society, as a source of basic skills, as a foundation for development of new knowledge and innovation, and as an engine for socio-economic development. Workers at all levels in the 21st century knowledge society will need to be lifelong learners, adapting continuously to changed opportunities, work practices, business models and forms of economic and social organisation [2]. Education is, therefore, a critical requirement in creating knowledge societies that can stimulate development, economic growth, and prosperity.

Thus, effective education in a knowledge society must also deal with sharing information, knowledge, and other resources. In this context, the link between ICT, education and development appears obvious.

A current challenge for education is to prepare learners for the emergent knowledge society through appropriate pedagogical practices that promote competencies for sharing, creating and working with knowledge and knowledge artifacts in an innovative way. Formal education is expected to develop methods that support students in acquiring versatile competencies for knowledge work.

The current approaches of working with knowledge in educational are still focused on individuals' skills and knowledge structures (knowledge acquisition) on the one hand, or on social and cultural interaction (participation) on the other hand. The problem is that they do not provide sufficient models for facilitating processes of knowledge creation with related technological, practical and organizational means.

The key features of 21st century pedagogy include [3]:
- Building technological, information and media fluencies;
- Developing thinking skills and making use of high order thinking skills;
- Making use of project based learning;
- Fostering problem solving as a teaching tool;
- Using assessment with timely, appropriate and detailed feedback and reflection;
- Encouraging collaboration using enabling and empowering technologies;
- Fostering contextual learning bridging the disciplines and curriculum areas.

To teach using 21st century pedagogy educators must be student centric. The curricula and assessments must be inclusive, interdisciplinary and contextual based on real world examples. Educators must develop, in students, key fluencies and make use of higher order thinking skills. They should make use of collaborative, project based learning, using enabling tools and technologies to facilitate this. Teachers must establish a safe environment for students to collaborate in but also to discuss, reflect and provide and receive feedback in.

Educational analysts and industry representatives report that students leave higher education with an underdeveloped ability to solve open-ended problems. Pedagogical methods are still largely based on well-defined problems with known solutions. Facing with real-world complexity and collaboratively solving complex problems are ways of bringing educational practices closer to the requirements of the surrounding society. This is a same challenge for secondary and higher education. Secondary schools should prepare students for the practices of higher education, including demands for more independent studying and managing complex, open-ended tasks, which are not well supported in the rather teacher-centered practices in secondary schools today.

To answer these challenges the KNORK (Promoting Knowledge Work Practices in Education) project [4] aims at developing pedagogical models and technology to support collaborative practices in technology-rich environment. The focus is on supporting secondary and tertiary level teachers to promote students' knowledge-related practices where digital

[1]Stela Stefanova is with the Technology School "Electronic Systems" associated with Technical University of Sofia, 2 Vl. Pashov Str, Sofia 1750, Bulgaria, E-mail: sstefanova@elsys-bg.org.

[2]Tania Vasileva is with the Faculty of Electronics at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: tkv@tu-sofia.bg.

competence is crucial. KNORK is an EU-funded integrated project with 9 partners from 4 countries. The partners represent the synergies between high education institutions and secondary schools in each country.

The paper discusses efforts done at the Technology School "Electronic Systems" associated with the Technical University of Sofia to include new technologies and creative pedagogies in the curriculum to address knowledge work practices and digital competence. Such key competencies are most effectively acquired through collaborative learning approaches around shared advancement of knowledge objects. A case study of restructuring Computer Aided Design (CAD) in Electronics course to promote knowledge work practices in cloud environment is considered. The result from pilots conducted with 11 grade students (17 years old) are also highlighted.

II. NEW PEDAGOGICAL DESIGN

Present-day students will be employed in positions representing modern knowledge work. These involve abilities of group work, collaborative learning, networking, working in multidisciplinary and multicultural teams, complex problems, and dealing with uncertainty and confusion.

The development of schools through the introduction of information and communication technology (ICT) is a central focus of educational discussion. New strategies are needed because research results indicate that pedagogical change in schools through ICT has not actualized as expected. Teachers often reproduce practices focusing on content learning rather than use technologies to foster higher-order innovation skills.

There is an emergent trend to highlight knowledge creation practices as a basis for understanding modern knowledge work, but fewer pedagogical approaches for promoting related competencies. Often the focus is still on the individual acquisition of domain knowledge, although increasing amounts of study programs are based on modules including distance learning, group work or web-based discourse. Besides these monological (acquisition of knowledge by individuals) or dialogical (participation in social interaction) models, there is a need for practical examples of supporting "trialogical" learning (collaborative knowledge creation): processes where the aim is to develop new products and solutions as in real knowledge work [5].

Thrialogical approach builds on the assumption that learning is not just individual knowledge acquisition (monological) or social interaction (dialogical), but activity is organized around transforming, or creating shared knowledge objects. While the acquisition and participation approaches provide valuable resources, respectively, for understanding individual and social aspects of learning, these metaphors do not appear to provide tools for understanding deliberate processes of advancing and creating knowledge typical of knowledge-intensive work in the present age. The trialogical approach is intended to elicit innovative practices of working with knowledge within educational and professional communities. In order to achieve this aim, it is essential to increase cross-fertilization between schools, higher education institutions and professional organizations. Students on all levels need to apply their expertise in interdisciplinary, goal-oriented projects that go beyond separate, single study units.

III. NEW PEDAGOGICAL SOLUTION – COLLABORATIVE PROJECT BASED COURSE IN THE SECONDARY SCHOOL

Our primary goal was to investigate and develop pedagogical practices that support students' knowledge work competences using trialogical design principles.

The problem was how to re-design our courses to better promote students' knowledge work competencies and how to implement the trialogical design principles in own teaching, using modern computer and communication technologies. The CAD in Electronics course in the Technology School "Electronic Systems" was re-designed to be project oriented.

The goal of course restructuring was:

- To increase the commitment and motivation of students, engaging them through 21st Century Learning.
- To meet the requirements of business for better practical training, team work on common task, shared responsibility for the quality of the overall product, distribution of tasks in line with the specified deadline.

In order to achieve these objectives new educational approaches are introduced, which use cloud computing technologies, up-to-date communication tools for student-teacher connection, continuous monitoring and assistance of students' activities. The compulsory course plan for education in "CAD in Electronics" was re-designed to implement the trialogical design principles, as shown in Table 1.

We decided to reconstruct the whole course to give students opportunity to work collaboratively in group with clear role of each participant in common work. Instead of giving students many separate or loosely connected tasks we provide them with a large task (a three month long project), continuous working process, shared research plan and final presentation in groups. All group activities are organized around shared objects – collaboratively development of common project, and preparation of shared report.

Working in teams of 2 persons, the students are required to design and simulate digital and analog circuits. During the long term projects teams have to gather information, discuss the given problem in collaborative environment, analyze and simulate the digital or analog circuit using dedicated CAD software – Cadence Orcad Capture and Pspice. Project development in such practice permits for self-selected time and place allocation of the participants and teachers. Guidance is provided through systematic instructions and group work rules. Assessment includes process and product assessments, group's self-assessment, and contribution evaluation of each participant to the collaborative project development.

IV. COLLABORATIVE WORKSPACE FOR PROJECT WORK

For the collaborative project development the environment consists of public cloud based services, combined in a way that supports collaborative electronic design projects

development (see Fig. 1). Tools for collaborative development of a common shared object in the cloud are used, which makes the participation of each member independently of the others in any place and at any time – Google Drive, Google Docs, Google Sheets.

Google calendar aims to set deadlines and to monitor progress – assignments, intermediate stages reporting, deadline for submission of project.

Table I. IMPLEMENTATION OF TRIALOGICAL DESIGN PRINCIPLES IN "CAD IN ELECTRONICS" COURSE

| Design principle	Implementation in own teaching
DP1: **Organizing activities around shared objects**	Collaboratively development of common project, and preparation of shared report Teams formation (team members will choose the partners they want to work with). Task distribution between the members of a team Teaching activities: regular meetings for discussion project tasks and preliminary review of the used tools and the progress of the project development
DP2: Supporting integration of personal and collective agency and work	Coordinating participants – team members to choose an appropriate project they want to develop (offering lists of possible projects themes). Motivating students to distribute tasks between team members having respect to project deadline. Collective responsibility – all members in a group should contribute to the group solution. Each member of the group takes the responsibility for his/her project task. They should decide on their own how each one does this. There is freedom to choose on which parts of the solution each member will contribute.
DP3: Emphasizing development and creativity through knowledge transformations and reflection	Discussion and analysis of problems the teams faced during their collective work on the common project. Support versatile use of various kinds of knowledge: theoretical or literary sources; practical examples and cases. Students comment on each other's / other groups work throughout the course Practice already gained knowledge and skills in using dedicated CAD software to solve the tasks of the project. If students need help they can always send messages to the teachers and get the supervision needed. Reflection: Students are expected to reflect on their individual report regarding their collaboration in the group
DP4: Fostering long-term processes of	Prolonged working process with iterative circuits simulations – performing number of analysis of the designed circuit to refine the circuit
knowledge advancement	parameters and characteristics. Planning and writing reports, sharing the drafts, asking the teacher and other students for feedback, improving the project and project documentation, submitting respective report and presenting the obtained design and simulation results. Using forums, blogs and social media for discussing problems and talk about their points of view and opinions.
DP5: Promoting cross-fertilization of knowledge practices and artifacts across communities	Students collaborate with specialists from the CAD industry. Students are provided with professional project work models, working templates, and good examples. Industry professionals, teachers and students discuss and analyze collaborative experience. Students use modern professional tools to plan, organize, and execute the project tasks and to write project documentation.
DP6: Providing flexible tools for developing artifacts and practices	Skype for face to face and virtual meetings. Google Apps for easy sharing materials or/and comment. Google Docs for collaborative editing and commenting. Google Drive for file sharing. Google+ for discussions. As alternative a CMS based site might be a suitable solution. Project management –Google Apps The Google Calendar is very useful for project scheduling – related events by sending RSVP invitations (request for a response from the invited person or people)
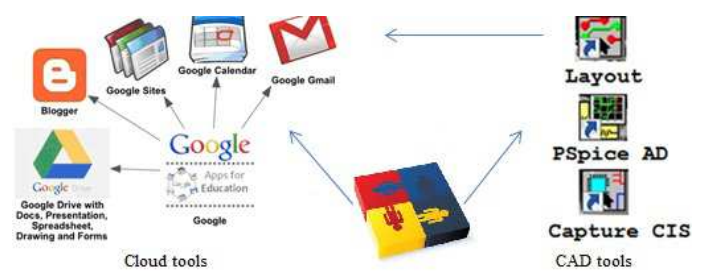
Fig. 1. Collaborative workspace structure used in CAD course

Most of developments take place outside the regular classes. For their intra-team communication, the students are free to choose whatever tools they prefer (chat, conferencing, email). For student-teacher communications we decide to use the Google tools (see Fig. 1): Groups, Gmail, Docs, Talk, Calendar, Drive and Google+. Students were encouraged to submit their questions as emails instead of chat messages.

In addition to long term project development the students have to prepare several weekly assignments, in the field of Analog and Digital circuits design and simulation. These

homework activities are presented, discussed and analyzed in the class and uploaded in Google Drive and Google Site of the CAD Course (See Figs 2 and 3).

V. PILOTS ACTIVITIES

To experiment trialogical approach we have conducted pilot courses using collaboratively development of shared project with two classes of 52 students – 11 grade classes (17 year's old students) within 15 weeks. Each team had to choose a project subject from a list provided by the teacher. In addition to the project work, students were required to submit several homework assignments. All participants had to register individual Google accounts. The teacher was responsible for creating a Google Docs document for each project report and sharing it with the team.
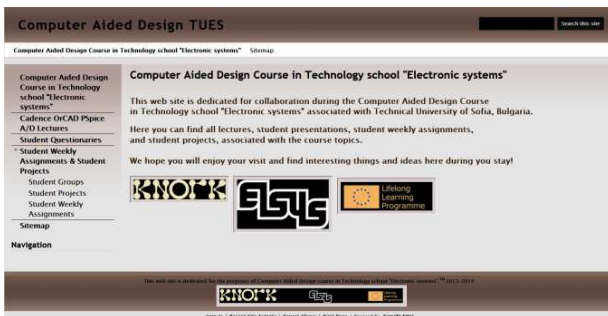
Fig. 2 Home Page of Google Site of CAD Course
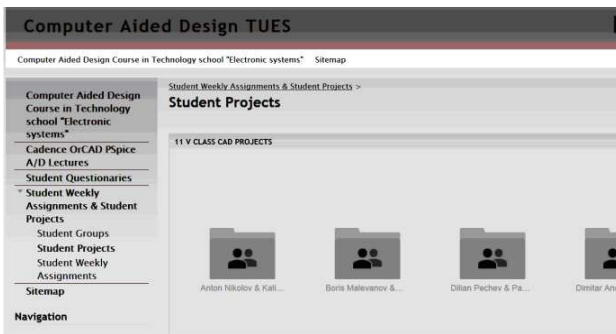
Fig. 3 Students Projects in CAD Course

In the beginning of the course a pre-survey was done. A questionnaire was sent to the students in order to find out more about their background and study skills. After the course the students were asked several questions to evaluate students' self-reflections concerning knowledge work practices related to their experiences in the implemented CAD course and to observe their progress. The 35 students' answers to the seven statements after the course are reported in Fig. 4 together with their answers to the statements before the course.
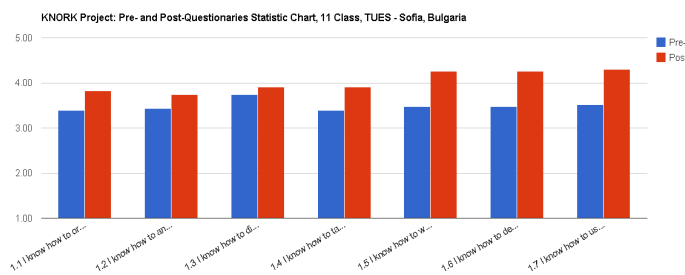
Fig. 4. Average of the students' answers

VI. CONCLUSION

The paper discusses a course redesign to promote new pedagogical practices, which were successfully used for introducing knowledge work in compulsory course "CAD in Electronics" at secondary education. These practices include longitudinal work supporting in-depth focusing and students' collaboration for a shared outcome. Students learned knowledge work practices (e.g., information processing, analysis and presentation, sharing, versioning, as well as commenting, using cloud digital tools and group work).

Introducing new technologies and paradigms in established courses is always challenging. In addition to the core subject matter, students had to learn new tools. Besides their knowledge on the subject they acquire skills to work in a team and to use advanced tools for collaboration and communication in the network. The restructured course gives the students opportunity to obtain better:

- Practical knowledge in electronic circuits design;
- Ability to use up-to-date professional tools for circuit design;
- Skills in teamwork;
- Knowledge how to manage their work in terms of tasks and time distribution for fulfilling deadlines;
- Ability to present and report their work, considering the problems they face, how they are resolved, or why these problems cannot be resolved.

In a whole it has been rewarding experience for both students and teachers. Greater interest and involvement of the majority of students were observed. Playing (and learning) with new technologies is a fun. The transformed CAD course has room for improvement. We need to find ways to promote even further the collaboration between the students and monitor their group process and individual progress.

ACKNOWLEDGEMENT

REFERENCES

[1] N. Butcher et al. (2011, December, 1). ICT, Education, Development, and the Knowledge Society. [Online]. Available: http://www.gesci.org/assets/files/ICT,%20Education,%20Devel opment,%20and%20the%20Knowledge%20Society.pdf

[2] M.Diaconescu, EU Higher Education in the Knowledge Society, Proceedings of the 11-th International Conference of ISSEI, Helsinki, Finland, August, 2008

[3] 21st Century Pedagogy. (n.d.). Retrieved July 1, 2014, from Educational Origami website, [Online]. Available: https://edorigami.wikispaces.com/21st+Century+Pedagogy

[4] Promoting knowledge practices in education (KNORK) Lifelong Learning EU Project http://knork.metropolia.fi/website/

[5]. Paavola, S., Lipponen, L., & Hakkarainen, K., Models of Innovative Knowledge Communities and Three Metaphors of Learning. Review of Educational Research 74 (4) 2004, 557-576, 2004.

Trends in the Education of the Modern Power Electronics and Motor Drives

Georgi Yordanov[1] and Tsvetana Grigorova[2]

Abstract – *The paper shows the technical aspects in the education of the modern power electronics and motor drives. In the master degree course, as a part of the student laboratory work is used the medium Voltage Digital Motor Control (DMC) kit, DRV8412-C2, from Texas Instruments. The software available with the kit is completely open source, which allows students to create their own projects and thus to better understand the basics of applying digital signal processor (DSP) in power electronics and digital motor drive control. The Master's thesis results are reported in the paper, which consist of a new software realization for controlling DC motors with bipolar PWM and unipolar PWM in the same mode of the driver DRV8412 operation. This removes the need for hardware changes in the kit when training students the various methods of modulation. The experimental results are presented.*

Keywords – **Digital motor control, DC motor .**

I. INTRODUCTION

The paper examines some trends in the education of the modern power electronics and motor drives. Universal, brushed DC and stepper motors comprise the majority of motor applications given their low cost and simplicity of control. The use of the digital motor control allows more complex and more intelligent types of motor control [4]. The students in the field of power electronics and motor drives are expected to be good with hardware and software. They need to understand not only the principles of motor control, but also digital signal processing, as well as analog and digital circuitry. The students need to graduate with skills that make them easily marketable to potential employers [3].

To meet the high requirements and competencies that are placed on the students of the power electronics, in the master degree course, is used the Medium Voltage Digital Motor Control (DMC) kit, DRV8412-C2, from Texas Instruments (TI) as a part of the student laboratory work [4, 5].

The software available with the kit is completely open source, which allows students to create their own projects and thus to better understand the basics of applying digital signal processor (DSP) in power electronics and digital motor drive control as well as to prepare their Master's thesis.

The motor control techniques are realized using the TMS320F28035 microcontrollers and the DRV8412 Dual Full Bridge PWM Motor Driver. TMS320F28035 devices are part of the family of C2000 microcontrollers which enable cost-effective design of intelligent motor controllers by reducing

[1]Georgi Yordanov is student (M.Sc) with the Faculty of Electronics and Automation at TU- Sofia, Branch Plovdiv, 63 Sankt Petersburg Blvd, Plovdiv 4000, Bulgaria, E-mail: georgi.pld@abv.bg

[2]Tsvetana Grigorova is with the Faculty of Electronics and Automation at TU- Sofia, Branch Plovdiv, 63 Sankt Petersburg Blvd, Plovdiv 4000, Bulgaria, E-mail: c_gr@tu-plovdiv.bg

the system components and increasing efficiency. The DRV8412 are high performance, integrated dual full bridge motor drivers with an advanced protection system [5]. In the DRV8412 DMC Kit are included two brushed DC motors and 8-wire bi-polar stepper motor (Fig. 1).
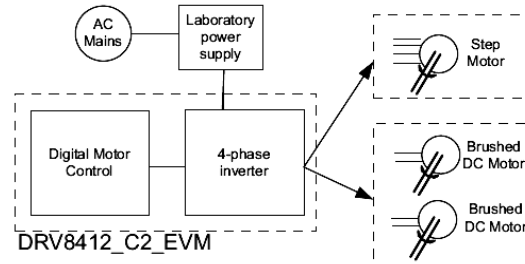

Fig. 1. Block Diagram for a Typical Motor Drive System Using

II. PWM MODULATION TECHNIQUE FOR BRUSHED DC MOTOR CONTROL

The overall system implemention of a 2-axis brushed DC motor is depicted in Fig 1. In this system, current control of the brushed DC Motor will be demonstrated. The brushed DC motors are driven by the conventional H-bridge configuration (Fig. 2). With the flexibility of four switches, a number of different control methods (bipolar or unipolar) can be used to produce fourquadrant output voltage and current [1, 2]. The waveforms for bipolar PWM voltage switching are shown in Fig. 3.
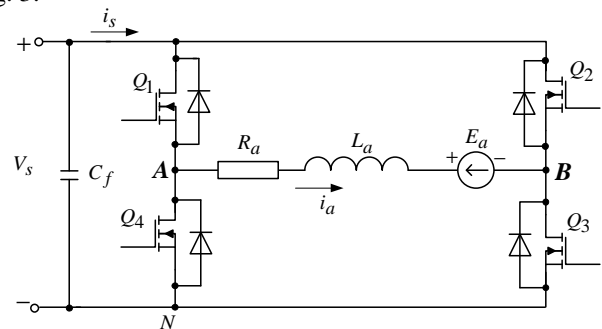

Fig. 2. Four-quadrant H-bridge converter

During the on-period for Q1 and Q3 ($0 \le t \le T_{on}$), when $v_a(t) = V_s$ (Fig. 2d)

$$V_S = R_a i_a + L_a \frac{di_a}{dt} + E_a \qquad (1)$$

which yields

$$i_a(t) = \frac{V_S - E_a}{R_a}\left(1 - e^{\frac{-t}{\tau_a}}\right) + I_{a_{\min}} e^{\frac{-t}{\tau_a}} \qquad (2)$$
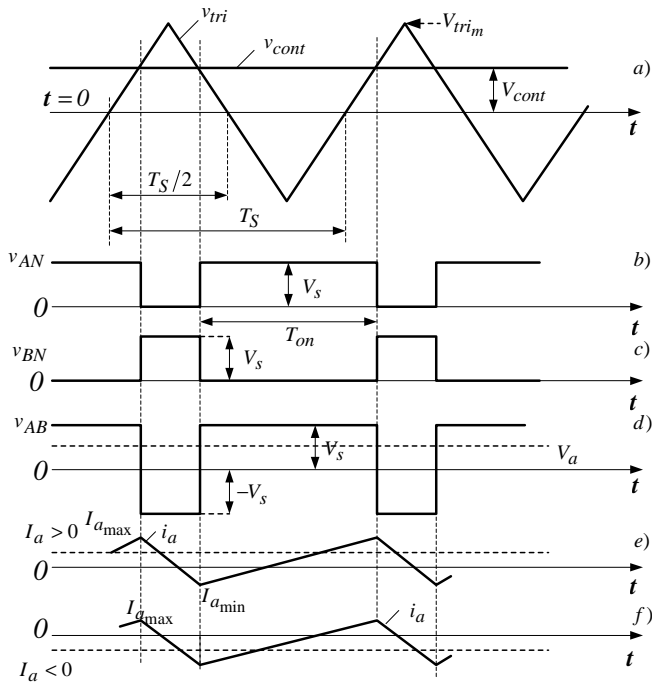
Fig. 3. PWM with bipolar voltage switching

During the on-period for Q2 and Q4 ($T_{on} \leq t \leq T_S$), when $v_a(t) = -V_S$

$$-V_S = R_a i_a + L_a \frac{di_a}{dt} + E_a \qquad (3)$$

which, gives

$$i_a(t) = -\frac{V_S + E_a}{R_a}\left(1 - e^{\frac{-t}{\tau_a}}\right) + I_{a_{max}} e^{\frac{-t}{\tau_a}} \qquad (4)$$

The initial conditions $I_{a_{max}}$ and $I_{a_{min}}$ are determined by using the steady-state boundary conditions:

$$I_{a_{max}} = \frac{V_S}{R_a} \cdot \frac{1 - 2e^{-DT_S/\tau_a} + e^{-T_S/\tau_a}}{1 - e^{-T_S/\tau_a}} - \frac{E_a}{R_a} \qquad (5)$$

$$I_{a_{min}} = -\frac{V_S}{R_a} \cdot \frac{1 - 2e^{-(1-D)T_S/\tau_a} + e^{-T_S/\tau_a}}{1 - e^{-T_S/\tau_a}} - \frac{E_a}{R_a}, \qquad (6)$$

where D – duty ratio ($D = T_{on}/T_S$); τ_a - load time constant; E_a - back EMF.

The average output voltage V_a is

$$V_a = (2D - 1)V_S . \qquad (7)$$

The F28035 is being used to generate the four pulse width modulation (PWM) signals needed to drive the DRV8412. Two input currents of each motor are measured from the H-bridge and they are sent to the F28035 via four analog-to-digital converters (ADCs).

Using the included project in the kit is generated one PWM sequence that appears on the output PWM-A or the output PWM-B depending on the motors rotation. With the

resulting signals unipolar and bipolar PWM technique for motor control can be accomplished, which is determined by the operating mode of the driver. A disadvantage is that in order to use one or the other method of PWM hardware switching has to be performed.

As mention above, the software available with the kit is completely open source and in the paper is presented the releasing of the different modulation strategies (4-Quadrant Bipolar and 4-Quadrant Unipolar technique) for brushed DC motors. The software for controlling one DC motor with bipolar PWM and the other DC motor with unipolar PWM in the same mode of the driver DRV8412 is developed. This removes the need for hardware changes in the kit when training the various methods of modulation.

It is assumed that two motors are connected to the device (Fig. 4). This build verifies the multi-axis concept which is two DC motors running simultaneously with current control. For the operation mode is selected mode 1 of DRV8412 (M1 = M2 = M3 = 0), therefore the control algorithm should be done so as to generate control signals for each side of the bridge circuits. The control method of the bipolar PWM is applied to the motor connected between the outputs A and B of the driver DRV8412, i.e. motor 1. The control method of the unipolar PWM is applied to the motor connected between the outputs C and D of the driver DRV8412, i.e. motor 2.
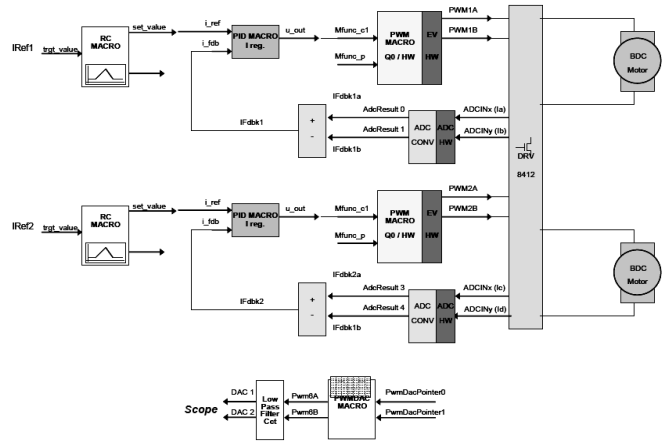

Fig. 4. Incremental System Build Block Diagram

The triangle carrier signal which is used to perform PWM sequences digitally represents the increase of the value of the counter TBCTR from 0 (0x0000) to a given value for a time equal to a half-period of the modulated signal (TBPRD), which is set by the function **PWM_Macro**, and depends on the output of the PID controller, and therefore the reference current IRef. Since ePWM module works with the frequency of the processor core, so in the span of time TBPRD the value, which the TBCTR counter reaches, is the value recorded in the register TBPRD.

A. PWM with unipolar voltage switching

PWM with unipolar voltage switching, in classic view, is realized as follows (Fig. 5): a triangular waveform is compared with the control voltage v_{cont} and $-v_{cont}$ for determining the switching signals for leg A and leg B,

respectively. A comparison of v_{cont} with v_{tri} controls leg A transistors, whereas leg B transistors are controlled by comparing $-v_{cont}$ with v_{tri}.
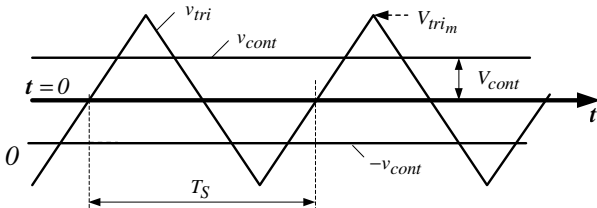

Fig. 5.

The triangle carrier signal in our project is generated as follows(Fig. 6): the equivalent of a zero will be considered half of the maximum value that the counter TBCTR reaches, i.e. one half from TBPRD. Compared to this level there should be two symmetric control signal v_{cont} and $-v_{cont}$, which are generated ePWM2A and ePWM2B, respectively. The values of these control signals are recorded in the registers CMPA and CMPB.
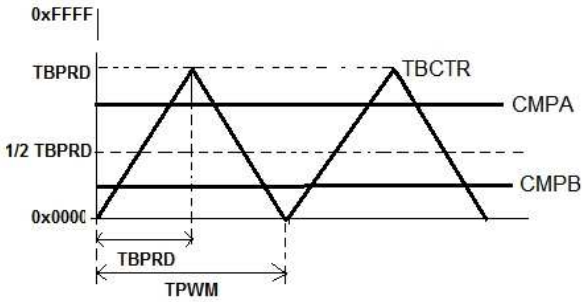
Fig. 6.

As we will use the value of CMPA register for the generation of signal ePWM2A and CMPB register for the generation of signal ePWM2B, it is necessary in the initialization of ePWM module to make appropriate initialization of the action qualifier control registers (AQCTLA and AQCTLB), determining what operation will be carried out with the output signals, thus reaching the value of the counter TBCTR with one of the two registers.

The value of these registers is set as follows:

EPwm2Regs.AQCTLA.all = (CAU_SET + CAD_CLEAR);
EPwm2Regs.AQCTLB.all = (CBU_SET + CBD_CLEAR);

Thus if CMPA > TBCTR ePWM2A signal will be in a high level, and similarly if CMPB > TBCTR ePWM2B signal will be in the high level. If CMPA < TBCTR and CMPB < TBCTR output signals will be in the lower level. The problem here is getting the values CPMA and CMPB, so that they are symmetric towards the half TBPRD value. The value recorded in CMPA register is actually a signal which is obtained based on the current setpoint value and the current through the motor, resulting in the feedback control. To obtain an accurate value for the CMPB register, so that it is symmetrical to the CMPA the following equation is proposed:

$$CMPB = TBCTRmax - CMPA = TBPRD - CMPA \quad (8)$$

In Fig. 7 the switching signals generation for this modulation is proposed. The two control signals are equating continuously.
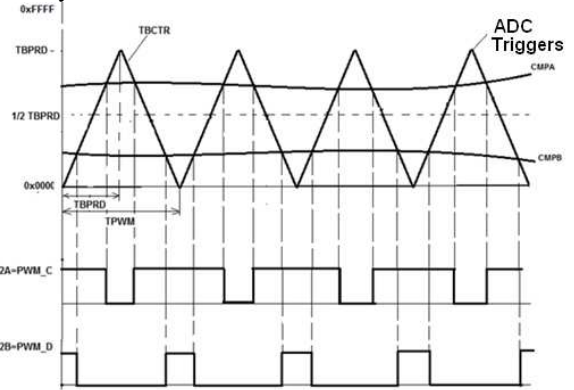

Fig. 7. Unipolar PWM control signals

B. PWM with bipolar voltage switching

For the realization of bipolar PWM necessary ePWM1 processor module to simultaneously generate two PWM sequences (ePWM1A and ePWM1B). These two signals are connected to the inputs and PWM_A PWM_B of DRV8412. Both control signals must be opposite, so as to switch both pairs in the bridge circuit, i.e. they are inverted relative to one another. In the classic bipolar PWM signals for generating ePWM1A and ePWM1B used triangular signal from the register TBCTR of ePWM1 module and the value of the register CMPA. For the proper generation of the control signals it is necessary for both register and AQCTLA AQCTLB to be initialized as follows:

EPwm2Regs.AQCTLA.all = (CAU_SET + CAD_CLEAR);
EPwm2Regs.AQCTLB.all = (CAU_ CLEAR + CAD_ SET);

Thus, if CMPA > TBCTR ePWM1A signal will be in the high level and ePWM1B signal will be in a low level. If CMPA < TBCTR ePWM1A signal will turn in low and ePWM1B in high level, respectively.
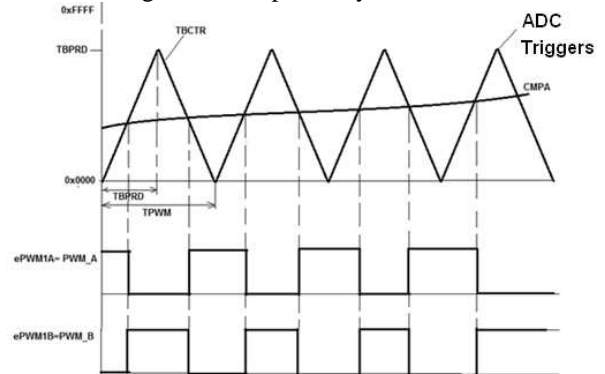

Fig. 8. Bipolar PWM control signals

III. EXPERIMENTAL RESULTS

The experimental results are made by the default PWM frequency 10kHz.

Fig. 9 shows ePWM2A (trace1) and ePWM2B (trace2) signals, which are generated with unipolar PWM for D=0.75.
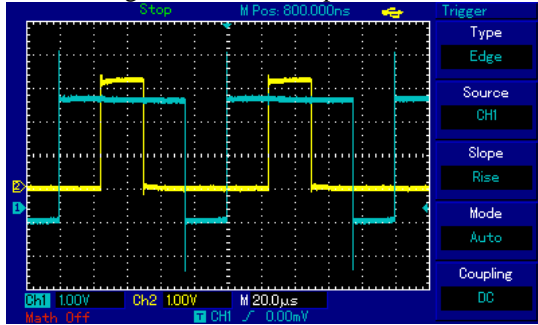

Fig. 9. ePWM2A (trace1) and ePWM2B (trace2) signals

Fig. 10 illustrates the output voltage (trace1) and the output current (trace2) with unipolar PWM for same D=0.75.
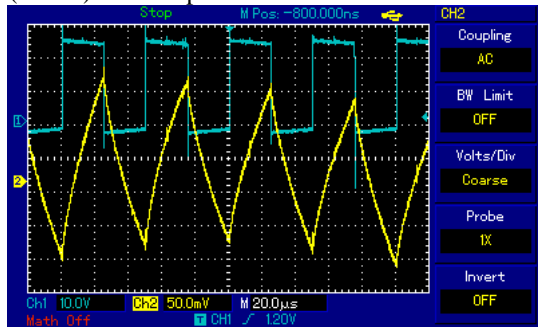

Fig. 10. Output voltage (trace1) and output current (trace2)

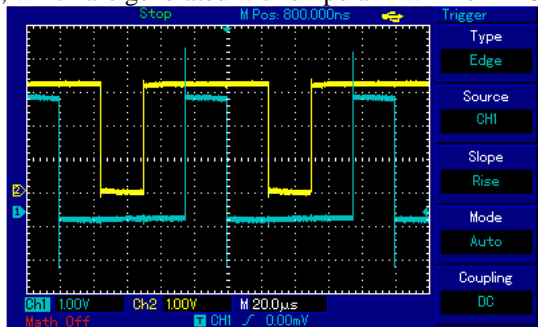Fig. 11 shows ePWM2A (trace1) and ePWM2B (trace2) signals, which are generated with unipolar PWM for D=0.25.

Fig. 11. ePWM2A (trace1) and ePWM2B (trace2) signals

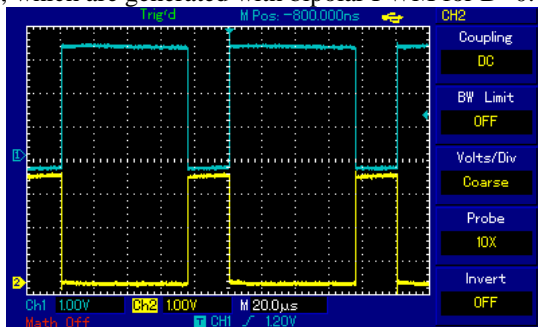Fig. 12 shows ePWM1A (trace1) and ePWM1B (trace2) signals, which are generated with bipolar PWM for D=0.75.

Fig. 12. ePWM1A (trace1) and ePWM1B (trace2) signals

Fig. 13 shows output voltage (trace1) and output current (trace2) with bipolar PWM for D=0.75.
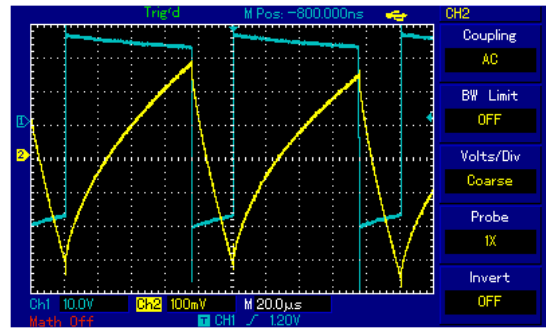

Fig. 13. Output voltage (trace1) and output current (trace2)

If the switching frequency of the switches is the same in these two PWM strategies, then the unipolar voltage switching results in better output voltage waveform and in a better frequency response, since the output voltage frequency is doubled and the ripple is reduced (Fig. 10 and Fig. 13).

In experimental testing of the control algorithms, it was found that when changing the current reference the response of the digital control system is 6,65μsec. When changing the conditions of motor operation with the proposed algorithm, the system will react to timing functions for PID, PWM modules, number of cycles for updating the condition and ADC conversion. This response time is 5,15μsec.

IV. CONCLUSION

The paper presents the realization of different modulation strategies (4-Quadrant Bipolar and 4-Quadrant Unipolar technique) for brushed DC motors, using DMC kit, DRV8412-C2, from Texas Instruments as a part of the student laboratory work. The Master's thesis results are reported in the paper, which consist of a new software realization for controlling DC motors with bipolar PWM and unipolar PWM in the same mode of the driver DRV8412 operation. This removes the need for hardware changes in the kit when training students the various methods of modulation. The experimental results are presented.

REFERENCES

[1] B. Williams, *Power Electronics*: *Devices, Drivers Applications and Passive Components*, 2006
[2] N. Mohan, T. Undeland and W. Robbins *Power Electronics*: *Converters, Applications, and Design*, New York, J. Wiley & Sons, Inc., 2003
[3] Y. Alsmadi, K. Tsai, M. Scott and L. Xu, "New Trends and Technologies in Power Electronics and Motor Drives Education", 121st ASEE Annual conf. and Exposition, Paper ID #10584, 2014, Indianapolis, Indiana.
[4] DRV8412-C2-KIT Hardware Reference Guide, Texas Instruments, www.ti.com/cn/lit/pdf/tidu314.pdf
[5] Dual-Axis Current Control of Brushed DC Motors Application Reports, Texas Instruments, www.ti.com/lit/pdf/tida012.pdf

Teaching Basic Skills in Embedded Systems Using Open-source Platforms

Peter Yakimov

Abstract – **In this paper a practical approach for teaching the basics of embedded systems hardware and software is proposed. It is emphasizing on the programme control of inputs and outputs to convince the students in the flexibility and universality of the programmable devices. Some examples are presented.**

Keywords – **Embedded systems, Open-source platforms, Arduino.**

I. INTRODUCTION

Embedded systems are everywhere. In recent years, they are becoming increasingly more important due to their widespread utilization in every aspect of people's lives [1]. Learning to design and program embedded systems is a critical skill that is necessary for many industry and scientific jobs [2]. At the level of today's technology consumer, there appears to be an increasing desire to interface the technological power-machines to the real physical world. This desire to connect may have been always present, but there appears to be more of a push towards closing the gaps between human and technology, by leveraging technology in a more personal, private and autonomous manner, under control of the user [3]. In order to prepare the students for the challenges of their future job the Faculty of Electronic Engineering and Technologies at Technical University of Sofia accepted a new curriculum for Bachelor degree in Electronics which was created after many iterations and discussions with the partners companies and employers organisations. The goal was to give the students theoretical knowledge about basic electronic circuits and devices and practical skills for their programme control. Also it was accepted that the practical training has to begin in the first semester in order to prepare the students for the specialised courses in the next years. As a development environment was chosen the Arduino platform which is open-source hardware, designed to make the process of using electronics in multidisciplinary projects more accessible. The hardware consists of a simple open hardware design for the Arduino board with an Atmel AVR processor and on-board I/O support. The software consists of a standard programming language and the boot loader that runs on the board. Arduino hardware is programmed using a Wiring-based language (syntax + libraries), similar to C++ with some simplifications and modifications, and a Processing-based IDE [4]. Another advantages of Arduino are that it can run on Windows,

Peter Yakimov is with the Faculty of Electronic Engineering and Technologies at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: pij@tu-sofia.bg.

Macintosh and Linux, and the active community of users [5].

II. NEW COURSE OBJECTIVES

According to the philosophy of the new curriculum a brand new course entitled "Practice on open source platforms programming" with two hours laboratory work per week was included. One major part of it is based on the usage of Arduino. The development board OLIMEXINO-328 based on the microcontroller ATmega328P is the hardware [4]. The aim of the course is with appropriate examples the students to understand the relationship between computer devices and the surrounding world. The experimental work includes an application software writing and debugging and measuring the response of controlled peripheral circuits. Thus, students acquire practical skills and obtain knowledge about basic electronic circuits and devices, and the possibilities for their programme control. During the laboratory work experiments on development boards are carried out. Initially the given task is analysed and the algorithm is drawn. Then a programme is written and run. Thus the students individually find possible errors and after analysis the results conclusions are made and mistakes are corrected. The topics are directly related to the field of the next courses.

III. LABORATORY SET-UP

Except the development board the necessary equipment includes a personal computer with an installed integrated development environment (IDE), a breadboard, some simple components as resistors and LEDs, and USB cable.
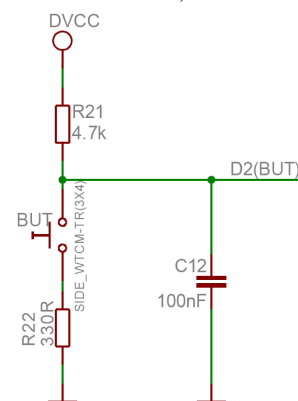

Fig. 1. User button with name BUT connected to ATmega328P pin 32 (digital signal D2)

The on-board user buttons and LEDs are used too as it is shown on Figs. 1 and 2 [4]. They are enough for introductory studying of the basic operations for control of digital inputs

and outputs. To use them the students have to know preliminary that the digital input reads logic "0" from the pressed button and to light the LED a high level must be set from the digital output.

Fig. 2. Light emitting diodes connected to ATmega328P pins 17 (LED1 - digital signal D13) and 17 (LED2 - digital signal D9)

To study the operation of the analog inputs and outputs some additional tools are needed. To set the analog input voltage a potentiometer has to be connected to the chosen analog input from Analog 0 to Analog 5 (pins from 23 to 28). Digital outputs with numbers 3, 5, 6, 9, 10 and 11 can be used as analog outputs because of their ability to set a pseudo analog voltage using pulse width modulation (PWM). To observe the response of the analog outputs a digital multimeter is necessary. Also off-board LEDs are used for this purpose. This gives the students an additional knowledge – to understand the analog operation of the LED and the principle of colours mixing.
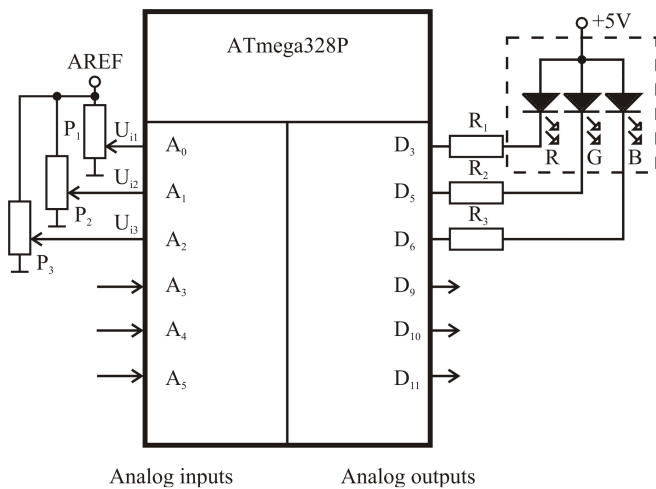

Fig. 3. Analog inputs and outputs connections for RGB-LED control

The laboratory set-up is shown on Fig. 3. The analog inputs A0, A1 and A2 are connected to the wipers of three potentiometers with 10k value. The voltages that are derived are set by the reference voltage of the built-in analog-to-digital converter in order to be stable and also that the maximal code will correspond to the maximal value of the input voltage. Outputs D3, D5 and D6 are connected to the cathodes of the RGB-LED with common anode organisation. This laboratory set-up is used for studying the basics of the

analog-to-digital and the digital-to-analog conversion. Also using the RGB-LED the students will understand the difference between the digital and the analog control of the brightness. They will realise the principle of colours mixing too.

IV. EXPERIMENTS

Here will be given some examples from the laboratory work that represent the practical approach for introductory teaching in embedded systems programming. Arduino development board and a software IDE are used. The goal is that using simple tasks the students will master the basic statements of C language which is the most often used language in the field of embedded systems. Also the students have to learn the structure of the programme and the possible approaches in the software design. The exercises start with very easy projects where the result is obvious and the complexity slightly increases with the progress of the course. Arduino IDE gives a lot of examples of code that can be used for preparation and self study. Learning the examples from the preloaded in the IDE the students study the basic statements of the programming language for control of digital inputs and outputs. After understanding the given programmes the students have to modify the code in order to master the knowledge. The next code examples which will be presented here are designed by the teaching team and they can be used in the process of learning.

The first task is to control the alternative blinking of the two LEDs on the development boards shown on Fig. 2. It is intended to study the statements for digital outputs control. The code is as follows:

```
int GreenLed = 13; // green LED is connected to pin D13
int YellowLed = 9; // yellow LED is connected to D9
void setup()
{
    pinMode(13, OUTPUT); // D13 is set as output
    pinMode(9, OUTPUT); // D9 is set as output
}
void loop()
{
    digitalWrite(GreenLed, HIGH); // LED1 is on
    digitalWrite(YellowLed, LOW); // LED2 is off
    delay(1000); // delay one second
    digitalWrite(GreenLed, LOW); // LED1 is off
    digitalWrite(YellowLed, HIGH); // LED2 is on
    delay(1000); // delay one second
}
```

The next step in programming is to study the statements for reading the state of a digital input and making decisions according to it. The following example of code illustrates this:

```
int Button = 2; // button is connected to pin D2
int GreenLed = 13;
int YellowLed = 9;
int State = 0; // variable for the read value
void setup() {
  pinMode(GreenLed, OUTPUT);
  pinMode(YellowLed, OUTPUT);
  pinMode(Button, INPUT); // D2 is set as input
}
void loop() {
  // read the state of the pushbutton value:
  State = digitalRead(Button);
  // check if the pushbutton is pressed.
```

```
// if it is, the State is LOW:
if (State == LOW) {
  // turn LED1 on and LED2 off:
  digitalWrite(GreenLed, HIGH);
  digitalWrite(YellowLed, LOW);
}
else {
  // turn LED1 off and LED2 on:
  digitalWrite(GreenLed, LOW);
  digitalWrite(YellowLed, HIGH);
}
}
```

The built-in analog-to-digital converter has 10 bits resolution. So the code returned after the conversion is in the range from 0 to 1023. The digital-to-analog conversion equals effectively to 8 bits resolution. To control analog inputs and outputs the students have to learn the proper statements. They are similar to those used for the digital ones. The following example contains these statements. The programme reads the value of an analog input A0 and using this value controls the brightness of a LED connected to D9 which can be used as an analog output too.

```
int YellowLed = 9;
int AnalogInput = 0; // potentiometer connected to A0
int Value; // Analog value
void setup () {} // there is no need of setup
void loop ()
{
  // reading the analog value
  Value = analogRead (AnalogInput);
  // correspondence between input and output ranges
  Value /= 4;
  //setting analog output value
  analogWrite (YellowLed, Value);
}
```

The above programme can be made slightly harder if the task is to control alternatively the brightness of two LEDs according to the analog input value set by a potentiometer. This will be useful for the students to master the skills for analog inputs and outputs control. The code example is:

```
int LED1 = 3; // LED1 connected to A3
int LED2 = 5; // LED2 connected to A5
int AnalogInput = 0; // potentiometer connected to A0
int Value; // Analog value
void setup () {} // there is no need of setup
void loop ()
{
  Value = analogRead (AnalogInput);
  Value = Value / 4;
  analogWrite (LED1, Value);  // LED1 control
  analogWrite (LED2, 255 - Value);  // LED2 control
}
```

To attract the attention of the students and give them an understanding about the modern information technologies is useful to show them the basic principles of colours mixing. For this purpose a RGB LED can be used as it is shown on Fig. 3. The organisation of the LED is with common anode thus the needed level to turn on the chosen LED is low. The three built-in LEDs can be controlled using digital or analog outputs. The students will recognise the difference and will understand the two level digital control and multilevel analog control. The following code example illustrates the colours mixing using digital control.

```
int RedLed = 3;
int GreenLed = 5;
int BlueLed = 6;
```

```
int Button = 2;
int state = 0;
void setup()
{
pinMode(RedLed, OUTPUT);
pinMode(GreenLed, OUTPUT);
pinMode(BlueLed, OUTPUT);
pinMode(Button, INPUT);
}
void loop()
{
if (digitalRead(Button))
{
if (state == 0)
{
digitalWrite(RedLed, LOW); //red colour
digitalWrite(GreenLed, HIGH);
digitalWrite(BlueLed, HIGH);
state = 1;
}
else if (state == 1)
{
digitalWrite(RedLed, HIGH);
digitalWrite(GreenLed, LOW); //green colour
digitalWrite(BlueLed, HIGH);
state = 2;
}
else if (state == 2)
{
digitalWrite(RedLed, HIGH);
digitalWrite(GreenLed, HIGH);
digitalWrite(BlueLed, LOW); //blue colour
state = 3;
}
else if (state == 3)
{
digitalWrite(RedLed, LOW); //yellow colour
digitalWrite(GreenLed, LOW);
digitalWrite(BlueLed, HIGH);
state = 4;
}
else if (state == 4)
{
digitalWrite(RedLed, LOW); //magenta colour
digitalWrite(GreenLed, HIGH);
digitalWrite(BlueLed, LOW);
state = 5;
}
else if (state == 5)
{
digitalWrite(RedLed, HIGH); //cyan colour
digitalWrite(GreenLed, LOW);
digitalWrite(BlueLed, LOW);
state = 6;
}
else if (state == 6)
{
digitalWrite(RedLed, LOW); //white colour
digitalWrite(GreenLed, LOW);
digitalWrite(BlueLed, LOW);
state = 7;
}
else if (state == 7)
{
digitalWrite(RedLed, HIGH); //black colour
digitalWrite(GreenLed, HIGH);
digitalWrite(BlueLed, HIGH);
state = 0;
}
delay(1000);
}
}
```
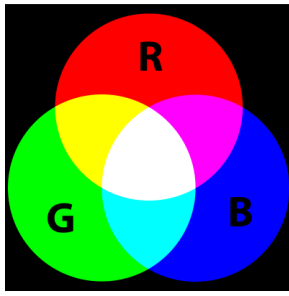
Fig. 4. Colours mixing using RGB-LED with digital control

The possible colours are red, green, blue, yellow, magenta, cyan, white and black. They are changed after pressing the button connected to D2. The mixing of the colours is shown on Fig. 4.

To apply analog control for the colours mixing every LED from the RGB must be controlled with output voltage that can be set with up to 255 levels as it is shown on Fig. 5.
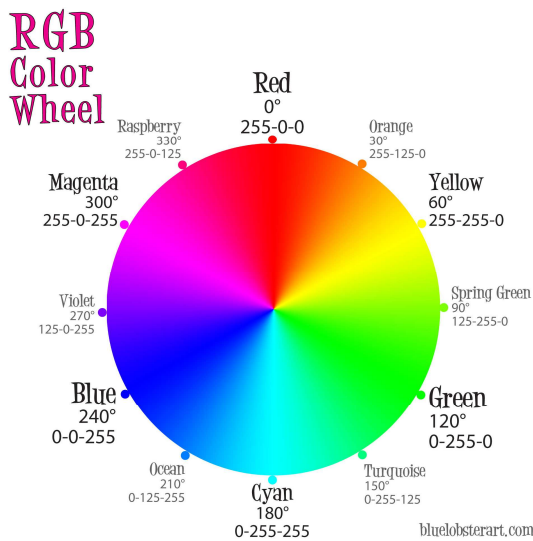

Fig. 5. Colours mixing using RGB-LED with analog control

The following code example illustrates the control. The laboratory set-up is shown on Fig. 3.

```
int RedLed = 3;
int GreenLed = 5;
int BlueLed = 6;
int AnalogInput1 = 0; // potentiometer connected to A0
int AnalogInput2 = 0; // potentiometer connected to A1
int AnalogInput3 = 0; // potentiometer connected to A2
int Value1; // Analog value1
int Value2; // Analog value1
int Value3; // Analog value1
void setup () {} //there is no need of setup
void loop ()
{
  Value1 = analogRead (AnalogInput1);
  Value1 = Value1 / 4;
  Value2 = analogRead (AnalogInput2);
  Value2 = Value2 / 4;
  Value3 = analogRead (AnalogInput3);
  Value3 = Value3 / 4;
  analogWrite (RedLed, 255 - Value1);
  analogWrite (GreenLed, 255 - Value2);
  analogWrite (BlueLed, 255 - Value3);
}
```

V. RESULTS

After completing the laboratory work dedicated to programming using Arduino board the students obtain experience in working with programmable devices. Basic skills in C programming language are mastered. The students realise the philosophy of the embedded systems and their application in all fields of the human activities and especially in the technical area. Also they are given knowledge about the most popular indicator elements and basic skills to work with them. The initial explanation of the principles of data conversion and colours mixing introduce the students in the world of the modern information technologies.

VI. CONCLUSION

In this paper a part of the course programme in "Practice on open source platforms programming" from the new curriculum for Bachelor degree in Electronics at the Technical university of Sofia has been presented. The course gives the students theoretical knowledge about embedded systems programming and practical skills in this field. The practical approach is useful for the beginners. With emphasizing on the practice rather than on the academic theory they accept the material easier. The open source hardware and software - Arduino development board and the software IDE are very useful for the introductory education. They offer friendly environment for beginners and give them experience for the further courses.

ACKNOWLEDGEMENT

REFERENCES

[1] S. Wong, and S. Cotofana, "On Teaching Embedded Systems Design to Electrical Engineering Students". Available: http://www.researchgate.net/publication/228408026_ On_Teaching_Embedded_Systems_Design_to_Electrical_ Engineering_Students.

[2] http://cse.unl.edu/~carrick/courses/2012/236/236_2012_spring_ embedded_systems.pdf.

[3] M. H. Lamers, F. J. Verbeek, and P. W.H. van der Putten, "Tinkering in scientific education", Proceedings of 10th International Conference on Advances in Computer Entertainment Technology (ACE 2013), LNCS 8253, pp. 568–571, 2013.

[4] OLIMEXINO-328 development board. Users Manual. OLIMEX Ltd, 2011.

[5] Banzi, M. Getting Started with Arduino. O'Reilly Media, Inc., 2011, ISBN: 978-1-449-309879.

Comparative Study of Three-phase, Two-phase and One-phase Impedance Tests for Induction Machines

Milan Radić[1], Milica Rašić[2] and Zoran Stajić[3]

Abstract – **This paper presents experimental results obtained using different types of impedance tests, performed on three-phase induction machines. Four different induction machines, with rated power of 0.18 kW, 1.5 kW, 3.7 kW and 7.5 kW were tested. The first type of applied tests was the standard three-phase impedance test, based on utilization of symmetrical three-phase power supply. Two other types of tests can be considered as modifications of standard test method, and they are based on utilization of single-phase power supply. With that type of power supply, two phases of stator winding were energized (two-phase impedance test), or just one phase was energized (one-phase impedance test). Obtained results have shown that the equivalent circuit parameters, calculated using measurements from three-phase and two-phase impedance test, are in excellent agreement for all tested machines. On the other side, parameters calculated using data from one-phase impedance test have always been substantially lower.**

Keywords – **Induction machine, Impedance test**

I. INTRODUCTION

Identification of three-phase induction machine equivalent circuit parameters is of crucial importance for any type of performance analysis, or for implementation of some control system. The IEEE standard 112 defines exact experimental procedures and calculations that should be performed in order to obtain accurate parameters of the machine's equivalent circuit. Mentioned procedures require combined usage of results recorded during no-load test and impedance test ([1]). Detailed analysis of formulae given in [1] suggests that, no matter which one of four standardized impedance test methods has been exploited, precise calculation of parameters in the stator and the rotor branch of the equivalent circuit should include the influence of the magnetizing branch. On the other hand, in majority of relevant textbooks, some simplified calculation approaches are given ([2-4]). Most authors suggest that the influence of the magnetizing branch can be neglected during the analysis of data obtained from the locked rotor test (which is „impedance test", if terminology from [1] is being used). Regardless of the actual calculation method, default premise is that a three-phase induction

machine is powered from a balanced, three-phase variable voltage source. Such situation requires that rotor of the machine has to be mechanically blocked, because the starting torque developed due to the three-phase supply tends to cause rotation.

In order to avoid rotation during the impedance test, some authors propose utilization of single-phase power supply, using the term „single-phase test". No starting torque is then produced, while the electric behavior of the machine is expected to be the same as in the case of three-phase excitation ([5, 6]). However, careful analysis of mentioned references reveals that different authors have different consideration of single-phase excitation. In [5], single-phase test means that single-phase supply is applied on two phase windings connected in series, while in [6] only one phase winding is energized.

The goal of the work presented in this paper was to investigate the influence that specific configuration of energized windings could have to obtained numerical results. For this purpose, single-phase supplying of two stator windings connected in series has been regarded as the „two-phase" impedance test, while single-phase supplying of the single stator winding has been regarded as the „one-phase" impedance test. Obtained results have been compared with results originating from the standard three-phase impedance test at the rated frequency, and conclusions are given.

II. MATERIALS AND METHODS

A. Machines used for experimental work

Three-phase, two-phase and one-phase impedance test were performed on each of four induction machines, whose nameplate data are given in this subsection:

Machine A:

$P_n = 0.18\,kW$, $U_n = 380\,V$, $f_n = 50\,Hz$, $I_n = 0.65\,A$,

$n_n = 885\,min^{-1}$, $\cos\varphi_n = 0.75$, stator Y, cage rotor.

Machine B:

$P_n = 1.5\,kW$, $U_n = 380\,V$, $f_n = 50\,Hz$, $I_n = 3.2\,A$,

$n_n = 2860\,min^{-1}$, $\cos\varphi_n = 0.86$, stator Y, cage rotor.

Machine C:

$P_n = 3.7\,kW$, $U_n = 380\,V$, $f_n = 50\,Hz$, $I_n = 8.4\,A$,

$n_n = 1400\,min^{-1}$, $\cos\varphi_n = 0.79$, stator Y, wound rotor.

Machine D:

$P_n = 7.5\,kW$, $U_n = 380\,V$, $f_n = 50\,Hz$, $I_n = 15\,A$,

$n_n = 1460\,min^{-1}$, $\cos\varphi_n = 0.87$, stator Y, cage rotor.

[1]Milan Radić is with the Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia E-mail: milan.radic@elfak.ni.ac.rs

[2]Milica Rašić is with the Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia E-mail: milicarasa@gmail.com

[3]Zoran Stajić is with the Faculty of Electronic Engineering, University of Niš, Aleksandra Medvedeva 14, 18000 Niš, Serbia E-mail: zoran.stajic@elfak.ni.ac.rs

B. Connection diagrams

During all experiments, tested machines were supplied with a voltage U, lower than the rated value U_n, using regulated three-phase autotransformer. Frequency of the applied voltage was equal to the frequency in the electric network, $f = 50\,Hz$. Voltages, currents, active and reactive power were measured using three-phase digital laboratory network analyzer. Used measuring device also enables direct reading of the actual power factor value.

Connection diagram for the three-phase impedance test is given in Fig. 1.
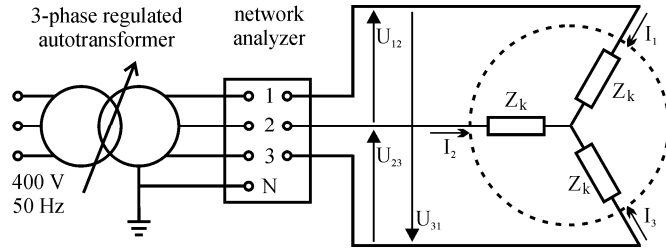
Fig. 1. Connection diagram for the 3-phase impedance test

For the two phase impedance test, connection diagram is given in Fig. 2. Note that two phase windings are connected in series and they are supplied using just one phase of the regulated autotransformer.
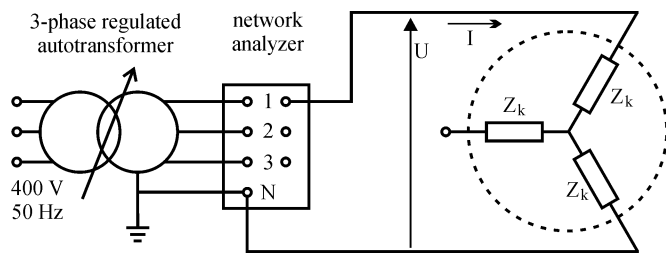

Fig. 2. Connection diagram for the 2-phase impedance test

Finally, connection diagram for the one-phase impedance test is shown in Fig. 3. In this case, current flows only through one phase winding, while other two windings are not energized. Since all tested machines have Y connected stator, neutral point of the regulated autotransformer has been directly connected to induction machine's neutral point.
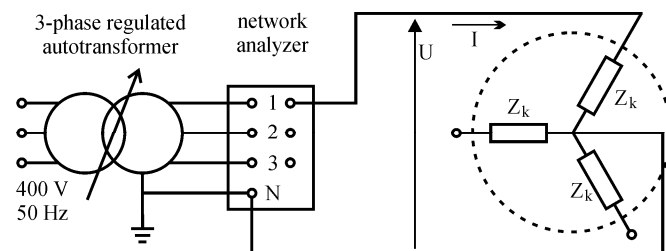

Fig. 3. Connection diagram for the 1-phase impedance test

C. Calculation of parameters Z_k, R_k and X_k

If the standard 3-phase impedance test is performed, while the rotor is mechanically blocked (see Fig. 1), total impedance per phase can be calculated as

$$Z_k = U / \sqrt{3}I \qquad (1)$$

where U is the average value of voltages between phases, calculated as

$$U = (U_{12} + U_{23} + U_{31}) / 3 \qquad (2).$$

and I is the average value of measured currents,

$$I = (I_1 + I_2 + I_3) / 3 \qquad (3)$$

Note that all four tested machines have Y connected stator, which means that the current calculated using Eq. (3) is also the average value of actual currents flowing through the stator windings.

If the 2-phase impedance test is exploited, according to the connection diagram shown in Fig. 2, total impedance per phase is calculated as

$$Z_k = U / 2I \qquad (4)$$

where U represents effective value of the applied single-phase voltage, while I is effective current flowing through the series connection of two stator phase windings.

Finally, for the case of the 1-phase impedance test, from Fig. 3, it is clear that the total impedance per phase is given by

$$Z_k = U / I \qquad (5)$$

Regardless of the exploited type of impedance test, total resistance per phase R_k, and total reactance per phase X_k have been calculated using simple formulae

$$R_k = Z_k \cdot \cos\varphi_k \qquad (6)$$

$$X_k = Z_k \cdot \sqrt{1 - \cos^2\varphi_k} \qquad (7)$$

where $\cos\varphi_k$ is the actual value of power factor, measured by laboratory network analyzer.

III. EXPERIMENTAL RESULTS

Measured electrical quantities are presented in Tables I, II, III and IV, along with calculated values of total impedance per phase Z_k, total resistance per phase R_k and total reactance per phase X_k. Distinct segregation of R_k and X_k on stator and rotor parameters was not performed, since such action was not necessary at this stage of investigation.

Calculated values were then expressed in p.u., using a relevant parameter (Z_k, R_k or X_k) obtained from the standard 3-phase impedance test of the analyzed machine, as a normalization base. Results are shown in Figs. 4, 5 and 6.

TABLE I
MACHINE A: 0.18 KW, CAGE ROTOR

parameters		3-phase test	2-phase test	1-phase test
measured	U [V]	123.55	141.38	60.79
	I [A]	0.653	0.645	0.649
	$\cos\varphi_k$	0.763	0.764	0.78
calculated	Z_k [Ω]	109.24	109.54	93.67
	R_k [Ω]	83.32	83.71	73.11
	X_k [Ω]	70.65	70.66	58.56

TABLE II
MACHINE B: 1.5 KW, CAGE ROTOR

parameters		3-phase test	2-phase test	1-phase test
measured	U [V]	53.51	63.06	25.41
	I [A]	3.206	3.245	3.146
	$\cos\varphi_k$	0.642	0.638	0.699
calculated	Z_k [Ω]	9.64	9.72	8.07
	R_k [Ω]	6.19	6.20	5.65
	X_k [Ω]	7.39	7.48	5.77

TABLE III
MACHINE C: 3.7 KW, WOUND ROTOR

parameters		3-phase test	2-phase test	1-phase test
measured	U [V]	81.06	93.59	39.23
	I [A]	8.436	8.449	8.429
	$\cos\varphi_k$	0.463	0.469	0.468
calculated	Z_k [Ω]	5.55	5.54	4.65
	R_k [Ω]	2.57	2.6	2.18
	X_k [Ω]	4.92	4.89	4.11

TABLE IV
MACHINE D: 7.5 KW, CAGE ROTOR

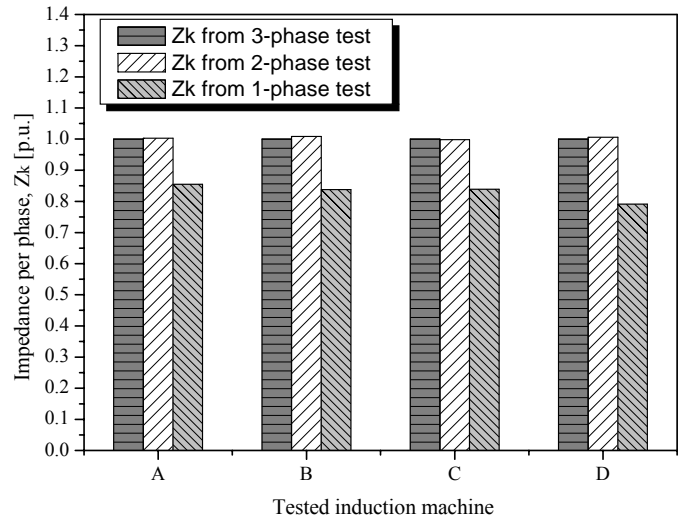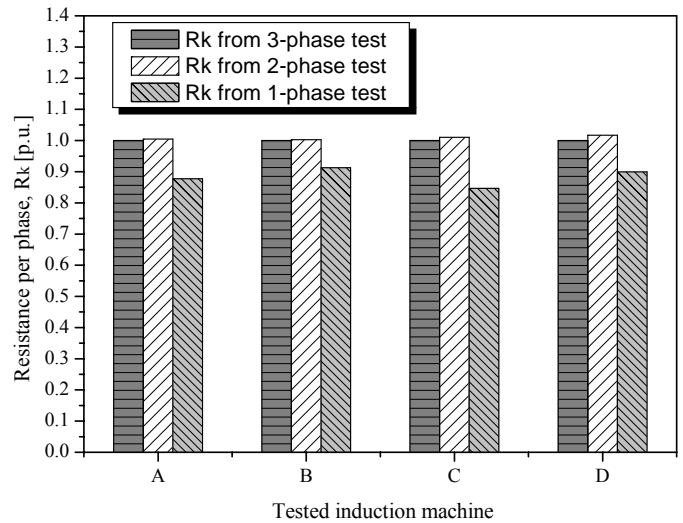| parameters		3-phase test	2-phase test	1-phase test
measured	U [V]	82.35	93.57	37.43
	I [A]	14.278	13.967	14.21
	$\cos\varphi_k$	0.337	0.341	0.383
calculated	Z_k [Ω]	3.33	3.35	2.63
	R_k [Ω]	1.12	1.14	1.01
	X_k [Ω]	3.14	3.15	2.43

Fig. 4. Relative values of total impedances

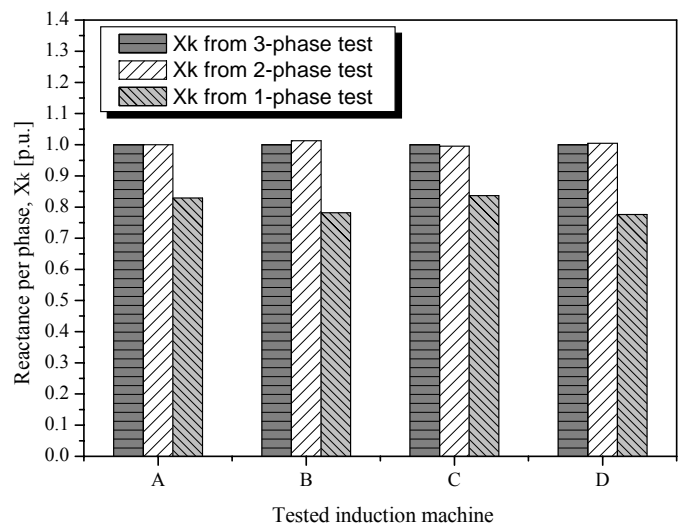Fig. 5. Relative values of total resistances

Fig. 6. Relative values of total reactances

IV. Discussion

Observing relative values of total impedances per phase shown in Fig. 4, one can conclude that results obtained from the standard 3-phase test and from the 2-phase test are in excellent agreement. For two of tested machines (A with cage rotor, and C with wound rotor), it is almost impossible to notice any difference. For two other machines (B and D, both with cage rotor), a slight disagreement between total impedances obtained from 3-phase and 2-phase test can be seen. However, it is not higher than 1%, and can be neglected. On the contrary, results for total impedance per phase obtained from 1-phase test have always been significantly lower compared to those from 3-phase and 2-phase test. As it can be seen in Fig. 4, for machines A, B and C deviation is about 15%, while for the machine D it reaches almost 20%.

Considering values of total resistances per phase, shown in Fig. 5, similar conclusions can be made. The standard 3-phase test and the 2-phase test have given almost identical results. The highest deviation has been noticed for the machine D, but it is still less than 2%. Results for total resistance obtained using data measured during the 1-phase test have been notably lower, with deviation greater than 10%.

Looking at Fig. 6, almost the same conclusion can be made when values of total reactance per phase are taken into consideration. All results for total reactances per phase are in a good agreement when the 3-phase test is compared to the 2-phase test, while results from the 1-phase test are significantly lower.

Presented results could be explained by the fact that space distribution of the magnetic field inside an induction machine is not identical during standard 3-phase, 2-phase and 1-phase impedance test. However, it seems that differences in the magnetic field distribution are smaller, or at least have less influence to the final results, when 3-phase and 2-phase impedance tests are being compared. For the case of the 1-phase impedance test, space distribution is significantly different (only one third of the complete stator winding is energized), thus leading to notable deviation in obtained results.

These results are valid for the rated value of frequency applied to the stator of the machine. Knowing that frequency can affect values for rotor winding resistance obtained during an impedance test (due to deep-bar effect), it will be of interest to investigate if the value of the applied frequency has some significant influence when 3-phase, 2-phase and 1-phase impedance tests are compared.

V. Conclusion

According to results presented in this paper, standard 3-phase impedance test (which demands mechanical blocking of machine's rotor), can be substituted by the modified, 2-phase test, without any important data being lost. It has been shown that results obtained using 3-phase and 2-phase impedance tests are in excellent agreement. For the case of the 1-phase impedance test, where only one phase of the stator winding is being energized, measured data do not give accurate results after calculations, and this type of the test should be avoided.

In the future work, it could be investigated if 3-phase and 2-phase impedance tests, performed at a frequency of the applied voltage different than the rated frequency, will still have similar output, or obtained results will be in an unacceptable disagreement.

Acknowledgement

References

[1] IEEE Standard 112 Test Procedure for Polyphase Induction Motors and Generators, IEEE Power Engineering Society , New York, 2004.

[2] I. Boldea, S. Nasar, *The Induction Machine Handbook*, London - New York - Washington D.C., CRC Press, 2002.

[3] M. G. Say, *Alternating Current Machines*, Pitman Publishing, London, 1977.

[4] R. Wolf, *Ispitivanje električnih strojeva*, Sveučilište u Zagrebu, Zagreb, 1964. (in Croatian)

[5] A. Gastli, "Identification of Induction Motor Equivalent Circuit Parameters Using the Single-phase Test", IEEE Trans. on Energy Conversion, vol. 14, no. 1, pp. 51-56, March 1999.

[6] M.O. Sonnaillon, G. Bisheimer, C. De Angelo and G.O. Garcıa "Automatic Induction Machine Parameters Measurement Using Standstill Frequency-domain tests", IET Electr. Power Appl., vol. 1, no. 5, September 2007.

On Cloud Computing in Engineering Education

Valentina Rankovska [1]

Abstract – **In this paper the opportunities offered by cloud computing in education are presented. Applying it together with innovative digital devices and modern methods for creating educational courses leads to increased effectiveness of teaching and learning. A brief overview of cloud services and their functions is provided.**

Keywords – **Cloud computing, Cloud Storage, Engineering education.**

I. INTRODUCTION

The trends, methods and tools, the problems of teaching in the higher education and particularly in the engineering education are a hot topic and are widely discussed.

Various aspects in teaching Microprocessor Systems for the students in the bachelor degree of Electronics are presented in numerous publications, very few of which are the basis of the examination, presented in [1].

Microcontrollers have a very complicated architecture that is too abstract for the students as it implies in advance well acquired knowledge and skills in the field of impulse and digital circuits. On the other hand, the development of an embedded system includes the design of both hardware and software.

Along with the widespread use of the microcontrollers in the embedded systems a variety of "intellectual", programmable integrated circuits (digital sensor circuits, memories, schemes for access control, real-time clocks and others) are used. They usually communicate with the microcontroller by various serial interfaces.

All this is related to the adoption and use of large amounts of information and design automation software.

At the same time, the presence of the Internet and the development of computer, multimedia equipment and mobile devices have changed the thinking, attitudes, and the ways of perceiving, processing and application of the information needed by students.

Along with the known advantages of the modern digital technologies, however, the peculiarities of their application in the field of learning must be taken into account [2]: the huge amount of "multichannel" information surrounding the students leads to a surface "scanning" without mastering by the students; the students expect immediate results, they are rarely inclined to an analytical thinking; young people more easily and quickly perceive visually presented information;

[1]Valentina Rankovska is with the Faculty of Electrical Engineering and Electronics at Technical University of Gabrovo, 4 H. Dimitar str., Gabrovo 5300, Bulgaria, E-mail: rankovska@tugab.bg.

the suffer of a lack of concentration, etc.

There are a variety of modern tools applicable in e-learning: universal digital devices, such as smartphones, laptops, tablets, etc., and also such with special functions for the field of learning - multimedia projectors, interactive presentation systems, document cameras, interactive classrooms, etc. [2], [3].

Their main drawback however - a high price when combining the "ideal" range of computers, communication and multimedia, and other digital devices, limits their application. Hence the new technologies in the Bulgarian universities grow significantly slower than necessary and behind the interests of the students, which are surrounded by digital devices and take their existence for granted and necessary.

Along with the "smart" digital e-learning devices, Internet applications and virtual environments are developed and implemented, expanding the opportunities for the e-learning.

Various approaches for effective learning in the field of the engineering education (and in particular in the digital and microprocessor circuits) are presented in [1] and many other articles.

This work focuses on the application of Internet and remote software platforms - the innovative cloud technology, for successfully engaging the attention of the students and increasing the effectiveness of training.

The learning in the engineering courses include various types of activities, such as attending lectures, working out of the laboratory and seminars, development of coursework, course and diploma projects, etc., This requires an individual searching for additional sources of information, such as datasheets, user guides, articles, textbooks and others. The cloud computing facilitates and supports the work of the students and the teachers at their collaboration, using the blended model [4], [5].

II. CLOUD COMPUTING IN E-LEARNING

Cloud computing means that shared computing resources and software are provided to the user through a browser. To satisfy computing needs, the necessary software and user data are stored on remote servers. It is implemented on several levels, the main ones which are [6] (Fig. 1 [7]):

- *Software as a service, SaaS* - the users have access to licensed software applications that are installed on the cloud, such as for word processing, image processing, e-mail, etc.
- *Infrastructure as a service, IaaS* - computing resources, operational and permanent memory, network access and other resources are provided necessary for the development of consumer applications (for instance the Google's App Engine).
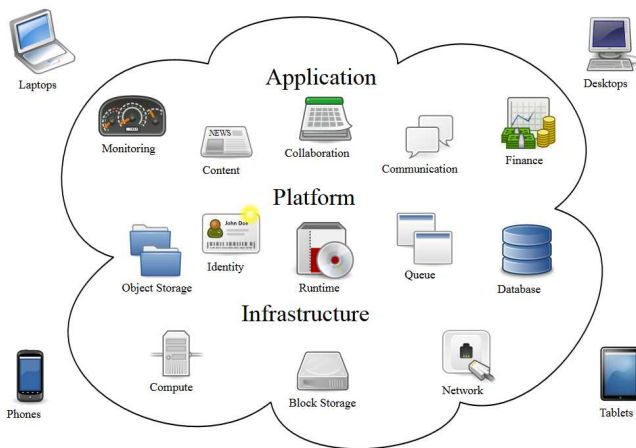
Fig. 1.

- *Platform as a service, PaaS* - The users have access to use programming languages and software to develop their own applications, which are then available in the cloud. Such a service is Google's Compute Engine.

Other cloud services are Windows Azure, Amazon Elastic Compute Cloud (EC2) and others.

The main characteristics of the cloud computing according to [8] are:

- *On-demand self-service* - The users receive and use the necessary computing resources and occupy only capacity and time necessary to accomplish the task.
- *Ubiquitous network access* - The users have access to computing resources at any time providing they have relevant digital devices to connect to the global network.
- *Location independent resource pooling* - The users access necessary resources regardless of their location.
- *Rapid elasticity* - The users can use different amounts of resources as necessary for them as they appear to be inexhaustible volume.
- *Pay per use* - Only a volume of computing power capacity used for storing user information is paid, etc.

There are several types of cloud computing, depending on the holder of the provided resources:

- *Public Cloud*, where the resources are available to the user from an external source, such as Google and Amazon;
- *Private Cloud* - the resources are provided by the organization, whose member is the user;
- *Hybrid Cloud*, as a combination of the above two options.

The basic prerequisites that make cloud computing comfortable for use in the modern education are:

- Availability (accessibility) of personal computers and other digital devices with similar features; fast Internet connections; the tremendous amount of information in the global network.

This makes the educational resources available to students anywhere in an interactive process - they can obtain information and make contact with other students for collaboration and mutual aid and with the teachers for

consultation, receiving and giving tasks, problems, educational materials, etc. in any place and at any time.

- Reducing the quantity of necessary local resources and enlarging the opportunities of the educational environment;

The saved money, time and other resources can be used for creating quality educational content.

Reducing the funds necessary to maintain an e-learning is determined by the need of advanced IT technologies and tools that develop rapidly, thereby pay a higher price; by electricity consumption of supported devices (PC, servers, peripheral devices, etc.), the necessity of cooling the hardware, and more. While Cloud computing provides a similar opportunity on the principle pay-as-you-go to pay only for the resources used for definite time [9].

- The advantages of e-learning and distance learning are combined.

There are some unresolved issues for now, as the security of the "entrusted" to the cloud information, the lack of a guarantee for the availability of the information at any time in the event of a failure in a server or the Internet connection and more.

III. CLOUD-BASED SERVICES IN E-LEARNING

There are various groups of paid and free web-based services. Most of the free services provide a limited volume of resources, which can be increased if it is necessary and the user pays for this. Some of the most common used are:

- **Online Desktop (Webtop) Applications**

These are applications that make the illusion that the user is working alone on their own PC. Many people could access the resources however, using a password. This allows the users to work in a team, for instance in the same project; the teacher has access to monitor and assess their students work and others.

There are various platforms with similar functions. Among them are free Ulteo Online Desktop [10], Zeropc [11], Oodesk [12], Jumptoit [13], etc.

Ulteo Online Desktop provides access to a number of free applications, such as Firefox web browser, OpenOffice.org office suite, KPdf to manage pdf documents, Kopete messaging software, Skype, Gimp and Digikam for managing pictures, Thunderbird + Enigmail email, Inkscape and Scribus for creating graphics, newspapers, etc. They plan also adding Windows applications. The user can pay for instance when needs more capacity, the more frequent synchronization, etc.

Zeropc is another free application that "simulates" a common personal computer for the students and teachers, allowing courses to be organized for them (Figs. 2, 3 and 4).

Hence it is possible access to cloud resources to local resources (Fig. 3) other sites, etc.

- **Personal sites**

Personal sites, like Protopage, Igoogle (Fig. 5), etc., allow the user to organize an own place with their own selection of news and blogs, keep bookmarks, to-do lists, sticky notes, calendar of events. The information can be organized in different tabs, for example for each course, diploma project, other sites (such as the widespread educational platform moodle - Fig. 6), etc.
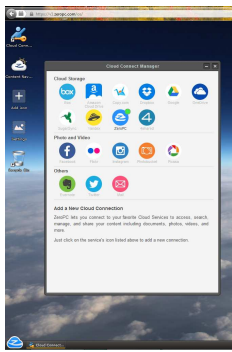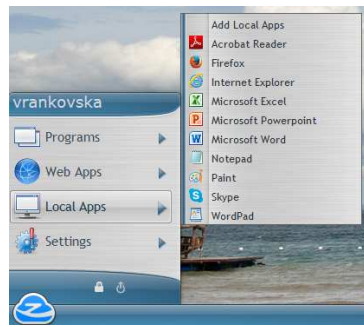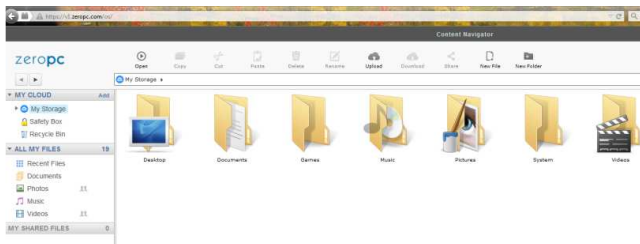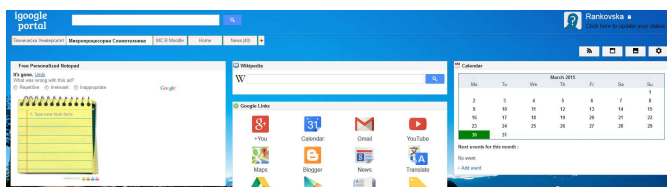
Fig. 2.
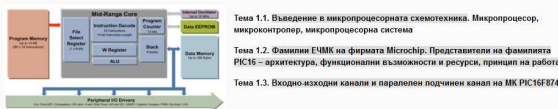

Fig. 3.

Fig. 4.

Fig. 5.

Fig. 6.

- **On-line document creating**

Applications such as *Live Documents* allows creating, storing, sharing on-line various types of documents such as spreadsheets, presentations, word processing, etc. And here, like other cloud applications, collaborative work on documents is possible without the limitation of a specific hardware, location, time.

- **On-line document and data storage** – Google Docs, Evernote, Dropbox, googledrive, skydrive, IDrive, etc. [14] provide space to store documents accessible to all the users of a course.

Opportunities provided by the cloud computing for the purpose of education are presented in a series of publications [15], [16], [9].

An infrastructure and application model is provided in [16] which to be used in the educational process and the administrative departments of the university.

Cloud computing together with innovative digital devices for learning allows organizing a virtual classroom [5]. In this case, cloud technologies complement the traditional classroom, i.e. a blended educational model is realized.

In [9] the conditions and experiences of The University of Westminster are summarized for the introduction and use of the free Google Apps (Education Edition) in relation to the activities concerning the students. For example, it has replaced the outdated students e-mail system (which forced them to start using personal e-mail accounts), providing them with 7,3 GB of free disk space per student. In addition this reduces the use of unreliable USB sticks, the spread of viruses and spam, prevents putting important emails from the University to the spam folder and others. At the same time the university system that stores and serves important resources such as documentation related to teaching and research is maintained.

In [17] a model of hybrid cloud is suggested, thereby increasing the flexibility and the reliability of the stored information and increasing the range of services provided to students, the authors of the courses, other teachers and administrative staff.

IV. CONCLUSION

Cloud computing is an innovative and useful approach to implement in the education. It provides many useful opportunities making the educational process easier for the teachers and attractive to the learners. It reduces the limitations of the place (teaching only in the classroom) and the time. The resources and the services are used on the principle pay-as-you-go. This allows reducing the total cost of the IT services used, thus increasing the quality of educational courses and applying other types of innovative digital devices in teaching.

Of course cloud computing will not replace the traditional approaches in the education, but will help to increase its effectiveness.

ACKNOWLEDGEMENT

REFERENCES

[1] Rankovska, V. A Case-Study Approach in Microcontroller Education. ICEST 2014, Conference Proceedings, vol. 2, Nish, Serbia, ISBN: 978-86-6125-109-2. pp. 299-302.

[2] Stoikova, V., A. Ivanova, A. Smrikarov. Good Practices in E-learning of Students of the Digital Generation. Rousse

University Research Papers, vol. 51, Issue 3.2., 2012, pp. 149-154. (in Bulgarian)

[3] Stoikova, V. et al. Interactive Tools in the Training of Students of the Digital Generation - Need or Extra? V National Conference on E-learning in the High Schools, Conference Proceedings, Rousse, Bulgaria, 2014, pp. 197-207, ISBN-978-954-712-611-4. (in Bulgarian)

[4] Porumb, S. et al. Cloud Computing and its Application to Blended Learning in Engineering. The Second International Conference on Cloud Computing, GRIDs, and Virtualization. Rome, Italy, 2011. pp.173-180.

[5] Rajaei, Hassan, and Eman A. Aldakheel. "Cloud computing in computer science and engineering education." American Society for Engineering Education. American Society for Engineering Education, 2012.

[6] Piskov, I. Cloud Technology Models for Providing Services. 2012. http://blog.icn.bg/%D0%BD%D0%BE%D0%B2%D0%B8%D0%BD%D0%B8-%D0%BE%D1%82-icn-bg/%D0%BE%D0%B1%D0%BB%D0%B0%D1%87%D0%BD%D0%B8-%D1%82%D0%B5%D1%85%D0%BD%D0%BE%D0%BB%D0%BE%D0%B3%D0%B8%D1%87%D0%BD%D0%B8-%D0%BC%D0%BE%D0%B4%D0%B5%D0%BB%D0%B8-%D0%B7%D0%B0-%D0%BF%D1%80%D0%B5%D0%B4%D0%BE/ (accessed in May, 2015) (in Bulgarian)

[7] Cloud Computing. http://bg.wikipedia.org/wiki/%D0%98%D0%B7%D1%87%D0%B8%D1%81%D0%BB%D0%B5%D0%BD%D0%B8%D1%8F_%D0%B2_%D0%BE%D0%B1%D0%BB%D0%B0%D0%BA (accessed in May, 2015) (in Bulgarian)

[8] Maria Spínola. The Five Characteristics of Cloud Computing. Cloud Computing Journal 2009. http://cloudcomputing.sys-con.com/node/1087426 (accessed in March, 2015)

[9] Nabil Sultan. Cloud computing for education: A new dawn? International Journal of Information ManagementVolume 30, Issue 2, 2010, Pages 109–116. (accessed in March, 2015) http://www.sciencedirect.com/science/article/pii/S0268401209001170

[10] http://www.ulteo.com/home/en/onlinedesktop

[11] http://www.zeropc.com/

[12] http://www.oodesk.com

[13] https://www.jumptuit.com/

[14] Free Cloud Storage Tools for Teachers. http://www.educatorstechnology.com/2012/06/free-cloud-storage-tools-for-teachers.html (accessed in March, 2015)

[15] K., Praveena and T. Betsy, Application of Cloud Computing in Academia (August 27, 2009). The IUP Journal of Systems Management, Vol. VII, No. 3, pp. 50-54, August 2009.

[16] Tuncay Ercan. Effective use of cloud computing in educational institutions. Procedia - Social and Behavioral Sciences. Volume 2, Issue 2, 2010, Pages 938–942. (accessed in March, 2015) http://www.sciencedirect.com/science/article/pii/S1877042810001709

[17] Md. Anwar Hossain Masud, Xiaodi Huang. An E-learning System Architecture based on Cloud Computing. World Academy of Science, Engineering and Technology, Vol. 6, Issue 2, 2012, Pages 736–740.

Short circuit investigation in telecommunication equipment through the theoretical electrical engineering

Svetlin Antonov

Abstract – **This paper shows an analysis of the features of the short circuits in technics which are precondition for ignition in telecommunication equipment. Keeping an non insulated electrodes of currency source may cause to a short circuit and processes of a huge energy transfer in a moment of time. This can lead to inflaming of some materials which ate the systems built by. The currency potentials, internal resistance and an external resistance in the electrical circle and the quantity of released and consumed energy of the outer element are investigated.**

Keywords – **Short circuit, Telecommunication systems, Theoretical electrical engineering, Ignition, Fire, Resistance.**

I. INTRODUCTION

Ignition in telecommunication equipment is an everyday occurrence. Some of the reasons are chemical reactions, old currency sources [1], short circuits in the electrical section, some external influences etc. When the fire affects a telecommunication technics, this leads to a huge costs to die out the flames and then to repair or to change the equipment. If the engineers know how the short circuit let the electrical energy to flow so intensively even for a second the will be able to prevent situations of a damage.

The current in case of a short circuit means a huge amount of energy transfer throw a low resistance medium (so small resistance as it lean to a 0 ohm). This process is described as Ohm's law, which is deeply investigated in the present paper.

Solving the problem of currency source and rising flames in telecommunication equipment is an analysis problem. It can be admitted: the construction and the other parameter of the source are known, the structure of the electrical circle and the included electrical parts with their parameters are known. Then the solution of the described problem is to find the currency value and the power in shorted circle which are enough to produce too much heat and a flame in the equipment.

II. MATHEMATICAL MODEL

An ideal voltage source is shown on Fig. 1. For the purposes of this investigation we assume the realistic case of availability of an internal resistance caused by the properties of the incorporated constructive materials in the technics (Fig.

Svetlin Antonov is with the Faculty of Automatics at Technical University Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: svantonov@yahoo.com.
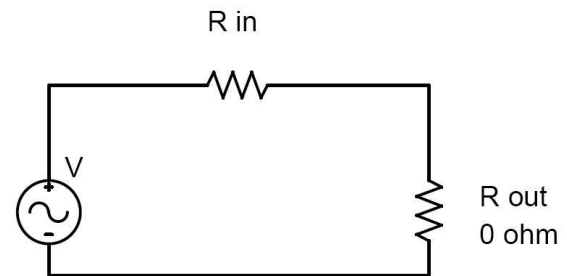
2).

Fig. 1. Figure example

Fig. 2. Figure example

Current flowing through the circuit is determined by the voltage between the source terminals, its internal resistance and consumers (external) one. In case of short circuit current will only depend on the internal resistance Rin and the potential difference between the electrodes Eq. (1) [2]. The accuracy of calculation of the potential in the point is determined by the integration constant.

$$V_1 - V_2 = \int\limits_1^{M_0} \vec{E}\vec{dl} - \int\limits_2^{M_0} \vec{E}\vec{dl} \qquad (1)$$

where: V1 and V2- value of the potential in the point measured in volts; M_0- point zero potential; $\vec{E}$ - Intensity of the electric field; $\vec{dl}$ - Tangential vector at the center of the element.

After the unification of the two integral obtained Eq. (2):

$$V_1 - V_2 = \int\limits_1^{M_0} \vec{E}\vec{dl} + \int\limits_{M_0}^{2} \vec{E}\vec{dl} = \int\limits_1^{2} \vec{E}\vec{dl} \qquad (2)$$

It follows that the potential difference coincides with the voltage between the two points in an electric field, i.e., between the source terminals.

When introduced in Fig. 2 internal resistance is correct to record inequality $|e(t)| \ge |u(t)|$. It becomes equality only in the absence of flowing current. If it is taken into account the internal resistance of the source:

$$u(t) = e(t) - R_{in} i(t) \qquad (3)$$

The determination of the resistance of the consumer is done with assumptions as: assumed to take place a short circuit between the source terminals by conductor made of homogeneous material; electric field is evenly $\vec{E} = const$; the conductive material has a conductivity γ, the conductive medium is linear $\gamma = const$; the cross section of the connecting wire is constant $s = const$. Based on the foregoing, it can save the differential form of Ohm's law (determining the density of the current) [2]:

$$\vec{J} = \gamma \vec{E} \qquad (4)$$

$$J = \frac{i}{s} = \gamma E \qquad (5)$$

$$u = \int_l \vec{E}\vec{dl} = \int_l E \, dl = E \int_l dl = El \qquad (6)$$

After substituting $E = \dfrac{u}{l}$ for i is prepared:

$$i = \frac{\gamma s}{l} u \qquad (7)$$

If the power source is adopted as the current source, the equivalent circuit (Fig. 3) includes a parallel resistor. This changes the ratio of the currents in this part of the circuit, and then obtained the inequality $|je(t)| \ge |i(t)|$. It becomes equality if implemented short circuit as a result of which the current through the internal resistance will be zero and will run exclusively through the external circuit.
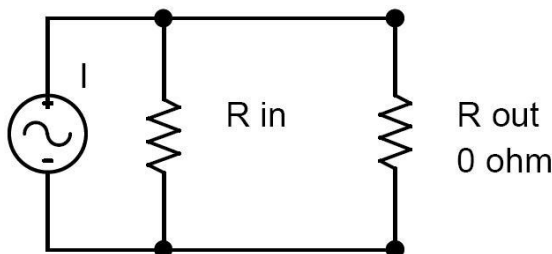

Fig. 2. Figure example

The resistance of the conductor is determined by the geometric dimensions and the specific conductivity of dependency:

$$R = \frac{l}{\gamma s} \qquad (8)$$

Released power from this element to the flow of current is determined by:

$$p = Ri^2 \qquad (9)$$

and power consumption is determined by:

$$p = \frac{1}{R} u^2 \qquad (10)$$

The development of Eq. (10) reference goes here to determine the instantaneous power, depending on the voltage on the conductive element and the current flow:

$$p(t) = u(t) i(t) \qquad (11)$$

For the purposes of the study overlooked the probability of the existence of a capacitive or inductive nature of the conducting part in the implementation of short-circuiting.

III. CONCLUSION

Taking into account the described parameters and their relationships a voltage/ currency source must be taken to isolate the electrodes in order to avoid short circuits. Isolation may be accomplished by any insulating, adhesive tape, it has to be tightly covered and the two electrodes of the battery and the strip are tightly glued edges. Other kind of isolation is to remove the source from the power nest in the corpus of the telecommunication equipment, which is the most secured option in the listed below.

Power consumption of the currency flow in short circuit leads to a huge power spending transformed from currency to heat. This is a precondition to ignite not so the conductor as the dialectical elements in the telecommunication equipment.

REFERENCES

[1] Fire and Arson Investigator, p20-28, ISSN: 1059-7298, January, 2010.
[2] Farhi, S, S, Papazov. Theoretical electrical engineering 1, Tehnika, 1987.

POSTER SESSION
INFORMATICS AND COMPUTER SCIENCE

Modified Quiescence Procedure in Axon Chess Engine

Vladan Vučković [1]

Abstract – **This paper presents the modified Quiescence algorithm. As the framework for this improvement basic Quiescence procedure is defined. Some original modifications and illustrative positions are presented in detail. All of these theoretical results and novelties are successfully implemented and verified in authors' chess applications Axon and Achilles.**

Keywords – **Theory of Logic Games; Computer Chess; Algorithms; Search procedures.**

I. INTRODUCTION

Computer chess is one of the very important subject for researching in the field of artificial intelligence. The large amount of theoretical and practical algorithms and applications are developed and proved their value in practice.

The modern PC chess engines are reached the grandmaster strength. There is a set of standard algorithms including *Alfa-Beta, Null Move, Transposition tables, Opening and Endgame* tablebases for constructing such a chess engines. Also, there are many of other supporting algorithms responsible to maximize the engine play strength. One of them is *Quiescence procedure.*

By definition, the evaluation function is applied to the terminal nodes in the tree [1],[2]. Using algorithms to search that contain implemented minimax procedure, the value is prolonged in reverse direction through the tree to the root node. Even in the earliest stages of computer chess researchers observed a significant negative effect that may occur in the final stage of evaluation nodes. This effect is called *horizon effect* because of fatal errors that could be generated trying to evaluate terminal dynamic positions. The *Quiescence* procedure intends to solve that problem introducing the extra tree selective searching in terminal nodes, instead of simply to evaluate all of them. By removing the horizon effect, the engines significantly increases their strength [3],[4],[5].

This paper has following structure. After Introduction, we will define *Quiescence* procedure and present basic algorithm. In the next section we will show the original improvement of the function including a few support procedures. Tuning and balance of these new procedures is very important, so we deal with it in following of the paper. We will conclude the paper with implementation of these novelties in stand-alone test application and after that directly into the author's chess engines *Axon* and *Achilles*.

[1] Vladan Vuckovic is with Faculty of Electronic Engineering, ul.Aleksandra Medvedeva 14, Niš, Serbia, E-mail: vladan.vuckovic@elfak.ni.ac.rs
vladanvuckovic24@gmail.com

II. QUIESCENCE PROCEDURE

To illustrate the horizon effect problem we will first postulate the basic variant of Quiescence procedure [6]. As an example, let's try to evaluate the position given in the following diagram (Figure 1):
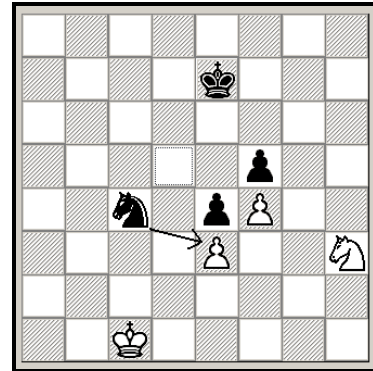

Fig. 1. The diagram shows a dynamic position that has to be evaluated.

Applying static evaluator in this position, where the material is completely identical for white and black, the evaluation would be very close to draw. However, if we take into account the dynamic parameters whether it is on the move white or black, it is clear that the white pawn on E3 cannot be saved, and that after the loss of that pawn, white will be forced to very difficult position: material evaluation is -1.00, black receives strong defended free of promoters by E line, the white pawn on F4 becomes very vulnerable. So, with the other parameters of evaluation positions created after taking a white piece could be freely approximated in interval -2.00 to -2.50 .

It is clear that attempting to evaluate a position that is not a clear static type led to significant errors, which could be crucial for the final evaluation and selection of the best moves [7]. In practical chess game, in every stage of the game, different positions generate a large number of unstable, dynamic position so that the application of the presented algorithms by definition would not lead to the realization of successful chess program. The solution for the horizon effect problem is that we must use specific procedures instead of pure evaluators to process the terminal nodes. These class of procedure are *Quiescence* searchers. This algorithm is a classic alpha-beta searcher with a slightly different logic to find the best continuation.

The procedure works as follows:

- In each generated terminal node the *Quiescence* searcher performs the evaluation [4]. If the evaluation value is greater or equal of the input variable is BETA, immediately exits the procedure. In each node, the heuristic generator determines the list of moves that can improve the existing evaluation.

These are most often moves that are taking the opponent's pieces (capture), some checks and promotions. Each of these moves is considered calling recursively *Quiescence* procedure. If the return value is greater than the variable BETA exit then the process is repeated but with improved value for ALFA. Because the moves in which take opposing figures represent in most cases the source of instability positions, their influence to these dynamic characteristics are successfully resolved. If we are in a position with no sharp bits from the group moves we mentioned, we directly do pure evaluation. In this way, in each node, the side that has the action provides the ability to maintain or improve its evaluation. Final positions are evaluated by static evaluator. In this way, this function overrides the horizon effect problem with introducing additional heuristics, so level of the game program improves significantly. Unfortunately, the introduction of *Quiescence* procedure, which in each node calls the dispute evaluative function, leads to a significant decrease in the overall speed *(PPS - position per second rate)* to the average level. It is necessary to very precisely adjust the balance among the amount of knowledge, speed and number of moves that are discussed in this procedure in order to reach the best performance [5]. In the following listing, we represent a basic form of *Quiescence* procedure:

```
procedure Quiescence (position, alpha, BETA):integer;
var …
  { Position – Current position }
  { Alpha – Upper limit }
  { BETA – Lower limit }

begin
    value:=evaluator(position)   {Static evaluation }

    if (value >= BETA) then   { If dynamic value is greater than
upper limit, exit }
      begin
        Quiescence:=BETA;
            exit;
      end;

    if (value > alpha) then alpha = value; { New value for alpha
variable }
    GenerateCaptures();

    while (CapturesLeft) do    { Generate search branch for
every element of list }
      begin
       GenerateNextCapture;
       value := -Quiescence(-BETA, -alpha);       {    Quiescence
procedure recursive call }
        UnmakeMove;

      if (val >= BETA) then begin Quiescence:=BETA; exit; end;
{ Exit }
        if (val > alpha) then alpha = value;
```

```
    end;
   Quiescence:=alpha;
 end;
```

Displayed procedure only considers the actions that take other figures (captures). Significant favorable factor in *Quiescence* procedure that considers only exchange pieces is auto-regulation of combinatorial explosion. Specifically, any recapture reduces the number of pieces on the board for one. After a series of piece recaptures, positions derives to a quasi-static position where there is no more active exchange of figures, and over which the procedure can be applied static evaluator.

III. MODIFIED QUIESCENCE PROCEDURE

However, it is important to consider this procedure and extensions in which one party in chess, as in many variants comes to mate [3]. The procedure considers the checks incurred after taking some figures and reveals mates shown in the following listing:

```
procedure QuiescenceCheck (position, alpha, BETA):integer;
var …
  { Position – Current position }
  { Alpha – Upper limit }
  { BETA – Lower limit }

begin
    if (NodeIsInCheck) then  { If terminal node king is in check
… }
      begin
       Quiescence:=Quiescence(1, alpha, BETA);
       exit;
      end;

    if (value >= BETA) then   { If dynamic value is greater than
upper limit, exit }
      begin
       NULLMOVE:=BETA;
           exit;
      end;

    if (value > alpha) then alpha = value;  { New value for
alpha variable }
    GenerateCaptures();

    while (CapturesLeft) do    { Generate a branch for every
element in a list }
      begin
       GenerateNextCapture;
       value       :=      -QuiescenceCheck(-BETA,       -alpha);{
QuiescenceCheck procedure recursive check }
        UnmakeMove;
```

```
    if (val >= BETA) then begin QuiescenceCheck:=BETA; exit;
end; { Exit }
    if (val > alpha) then alpha = value;
  end;
  QuiescenceCheck:=alpha;
 end;
```

So it is the modified form of *QuiescenceCheck* recursive procedure. According to many authors, with the implementation of the checks in *Quiescence* procedure should be very careful, because a large percentage of a completely useless moves, which cannot improve evaluation leading to a significant increase in the number of processed nodes.

IV. MVV/LVA AND SEE PROCEDURES

A very important issue that must be addressed is how to sort the list of moves that need to be processed at each node. The two main techniques used are: MVV/LVA *(Most Valuable Victim/Least Valuable Attacker)* and SEE *(Static Exchange Evaluation)*. Both of these two procedure have found their place in modified *Quiescence* procedure [8]. MVV/LVA technique is used for brief screening of list moves toward the expected material gain. It will always move that takes the opponent queen that has a higher value of the moves that takes the opposing pawn. In the event that there are more pieces that can take the same opponent piece, the first deals with the suffix that means minimal material costs.

For example, if we have a move PxQ *(Pawn takes Queen)* it will be in the list moves ahead moves QxQ *(Queen takes a Queen)*. This technique does not take into account the fact that the opponent's pieces can be defended but is easily deployed and working very quickly.

As an illustration, we present an example of forming a list of moves in the position on the diagram using the second MVV/LVA techniques.
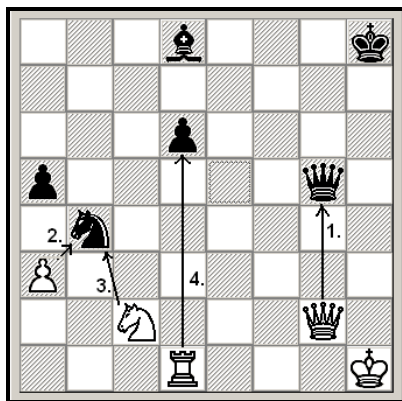

Fig. 2. The sequence of moves using the MVV/LVA procedure.

The sequence of moves is: (QxQ, PxN, NxN, RxP).

SEE technique is more complex than MVV/LVA and involves predicting the outcome of a series of trade figures. The procedure approximates the result of a series of amendments figure without having to call *Quiescence* procedures. Unlike the previously exposed techniques, if the opposing figure that is taken, defended by some other opponents figures, this fact is taken into account.

On the way to the much more realistically predicts the outcome of a series of trade figures and thus significantly improves the sequence of moves that are under consideration. Figure 2 shows the identical position to that of Figure 3, but is now used in processing SEE instead of MVV/LVA techniques:
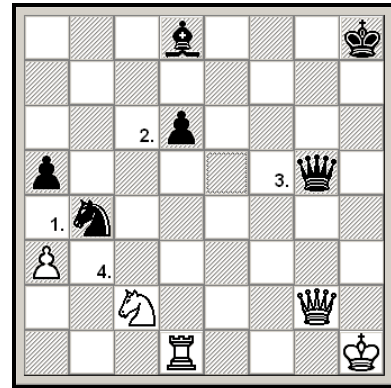

Fig. 3. The sequence of moves using the SEE technique.

The sequence of moves using the SEE techniques is slightly different: (PxN, RxP, QxQ, NxN). The sequence of moves that generates and order processing sub-tree is different in relation to the MVV/LVA technique. Bearing in mind the anticipated losses on each side, sort of moves is with much higher quality in the later stage generates a significantly higher number of cutting trees in *Quiescence* procedure.

In some older programs, SEE is used to approximate the dynamic evaluation of the terminal nodes of *Quiescence* without calling procedure, but in later versions of this it was changed in the way to eliminate too much of the tactical oversights. On the other hand, SEE procedure is considerably slower than the MVV/LVA procedure which reflects the overall NPS *(node per second)* feature of the program. However, in most modern programs including *Axon* [9], we use almost exclusively SEE.

V. APPLICATION AND TESTING OF THE QUIESCENCE PROCEDURE

After begin development of the new 32 bit version of the *Axon* chess engine, the author was faced with the problem of efficient development and testing of *Quiescence* procedure in the independent application [10].

This application is only for research purpose, it is open and intended only for the construction of some key procedures of the program [11]. Graphical environment for testing these functions is shown in Figure 4:
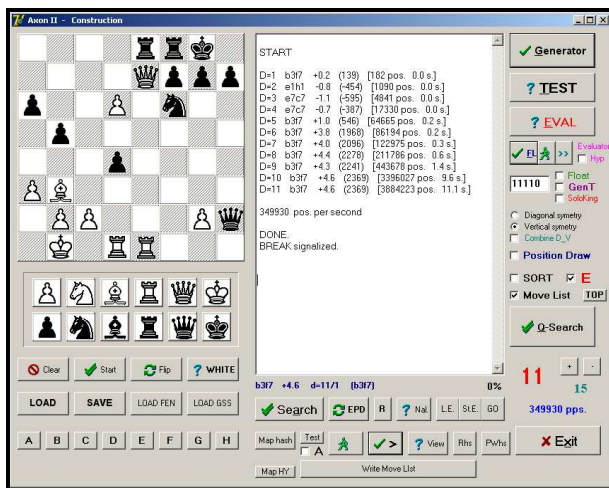
Fig. 4. Graphical environment of the program for testing of the Quiescence function.

The graphical environment contains all the necessary functions that are used for development of the *Quiescence* function. In addition to the standard options for creating and storing of the position, in relation to the visual display, various parameters and arguments and special options are embedded [12]. There is automatic test of the *Quiescence* function based on the random assignment of a set of N pieces on the chessboard, taking into account the legality of the position. Testing positions on various parameters including the symmetry is very fast, which significantly simplifies the programmer implementation and change certain aspects of the *Quiescence* function, that is necessary in a dynamic phase of development applications.

Connection to the main application is very simple and is done via UNIT common files used by the main application. After translation to machine language, the application automaticly create ENGINE.DCU file that contains all of the relevant data structures, procedures and functions for the realization of the chess machine [13].

At the stage of compilation of the basic program, *Axon II* simply lists used directive states in engine. dcu unit, so the connection is made in a very efficient manner. Control of individual processes in the unit ENGINE.DCU is done through a series of linked procedures. Every upgraded version of application then use new reference in ENGINE.DCU simply generating the new version of *Quiescence* function [14].

VI. CONCLUSION

This paper presents some original improvements embedded into the basic general Quiescence procedure with direct implementation in author's computer chess program *Axon.* Modified procedure with some extra procedures added (MVV/LVA,SEE) improved basic function and increase the total playing strength of the chess engine.

There are many possible tuning of these procedures. We will concentrate our algorithm to optimal solution balancing between efficiency and complexity needed for master computer chess game.

This solution has been already proved in direct implementation in *Axon's* tournament practice. Also, all other improvements will be checked in this way. The future prospect will be to increase amount of embedded chess knowledge into the *Quiescence/Evaluation* procedure using the constant rise of CPU processing power.

ACKNOWLEDGEMENT

REFERENCES

[1] Althöfer, I., Sören W. Perrey: Mathematische Methoden der Künstlichen Intelligenz: zur Quiescence-Suche in Spielbäumen. *ICCA Journal*, Vol. 14, No. 2, p. 84. ISSN 0920-234X, 1991.

[2] Beal, D. F. A Generalized Quiescence Search Algorithm. *Artificial Intelligence*, Vol. 43, No. 1, pp. 85-98. ISSN 0004-3702, 1990.

[3] Beal, D. F. Mating Sequences in the Quiescence Search. *ICCA Journal*, Vol. 7, No. 3, pp. 133-137. ISSN 0920-234X, 1984.

[4] Kaindl, H. Dynamic Control of the Quiescence Search in Computer Chess. *Cybernetics and Systems Research* (ed. R. Trappl), pp. 973-977. North-Holland, Amsterdam, 1982.

[5] Schrüfer, G. A Strategic Quiescence Search. *ICCA Journal*, Vol. 12, No. 1, pp. 3-9. ISSN 0920-234X, 1989.

[6] Slate D. J. , Atkin. L. R. CHESS 4. 5 – "The Northwestern University Chess Program", *Chess Skill in Man and Machine* (ed. P. W. Frey), pp. 82-118. Springer-Verlag, New York, N. Y. 2nd ed. 1983. ISBN 0-387-90815-3. , 1977.

[7] Vučković, V. , Vidanović, Đ. "An Algorithm for the Detection of Move Repetition Without the use of Hash-Keys", *Yugoslav Journal of Operations Research (YUJOR)*, Volume 17, Number 2, pp. 257- 274. Belgrade, Serbia. ISSN 0354-0243. , 2007.

[8] Vučković, V. "The Theoretical and Practical Application of the Advanced Chess Algorithms", *PhD Theses*, The Faculty of Electronic Engineering, The University of Nis, Serbia, 2006.

[9] Vučković, V. *xon/Achilles* experimental chess engines information could be find at: http://axon. elfak. ni. ac. rs , http://chess. elfak. ni. ac. rs, 2007.

[10] Vučković, V. , "The Compact Chessboard Representation", ICGA Journal, Volume31, Number 3, Tilburg, The Netherlands, ISSN 1389-6911. pp. 157- 164., 2008.

[11] Vučković, V. , "The Method of the Chess Search Algorithms Parallelization using Two-Processor Distributed System", The Scientific Journal Facta Universitatis, Series Mathematics and Informatics, Volume 22, Number 2, Niš , ISSN 0352-9665. pp. 175-188., 2007.

[12] Šolak, R. , and Šolak, Vučković, V. , "Time Management During a Chess Game", *ICGA Journal*, Volume 32, Number 4, Tilburg, The Netherlands, ISSN 1389-6911. pp. 206- 220., 2010.

[13] Vučković, V., "Advanced Chess Algorithms and Systems", monography, Zadužbina Andrejević, Biblioteka Dissertatio, ISSN 0354-7671, Belgrade, 2011

[14] Владан Вучковић, "Специјални елементи евалуационе функције", Зборник радова са 53. Конференције ЕТРАН-а, CD ROM Proceedings, Секција Вештачка интелигенција, рад VI1. 3, Врњачка Бања, (ISBN 978-86-80509-64-8), 2009.

The Estimation of Acoustic Suitability of the Amphitheater for Lecturing

Violeta Stojanović [1] and Zoran Milivojević [2]

Abstract – **This paper presents the acoustic analysis of the amphitheater provided for lecturing. In the first part of paper we describe the experiment which measures the acoustic impulse response and show the calculated acoustic parameters. The results are shown both graphically and tabular. In the second part of this paper we analyse the parameters for the empty amphitheater and their accordance to ISO 3382 standard.**

Keywords –**Room impulse response (RIR), reverberation, acoustic parameters.**

I. INTRODUCTION

The analysis of the acoustic parameters of a room can determine its suitability for certain purposes. Namely, for auditoriums, i.e., amphitheaters, the important parameters are Reverberation Time, *RT,* and Speech Intelligibility. The empirical equation for estimation of *RT* was given by C. W. Sabine in 1922 [1]. The estimation of *RT* is based on the analysis of the volume, the dimensions and wall absorption coefficients for a certain room. Further, there were several empirical equations where the estimated and measured *RT* showed consistency within the smaller reverberation rooms in relation to Sabin's prediction. The most significant one is Norris - Eyring equation which was presented in 1930 [2]. When a wall absorption coefficients are considerably different, the best results are obtained by Milingtone Sette equation which was presented in 1933 [3]. In 1965 Schreder proposed the algorithm estimations of *RT* using RT_{60} acoustic parameter, which is calculated by analysing of the acoustic Room Impulse Response, RIR [4]. Early Decay Time, *EDT*, was defined by Jordan in 1970 [5]. Centre Time, T_C, was defined by Cremer and Curer in 1971 [6]. In 1971, Peutz suggested the parameter referring to Articulation Loss of Consonants, Al_{cons} [7]. Abdel Alim and Reichard suggested the parameter Clarity C_{50} and C_{80} for speech and music in 1974 [8]. In 1985 Houtgast and Steeneken suggested the acoustic parameter Speech Transmission Index, *STI*, within the room [9].

This paper analyses the acoustic suitability of the amphitheater by analysing some of the acoustic parameters. For the purpose of the analysis, measuring of RIR was done at 9 measuring points. After that, using RIR analysis, some of the acoustic parameters were calculated along with their mean values and standard deviations as follows *EDT*, T_{R10}, T_{R20}, and

[1] Violeta Stojanović is with the College of Applied Technical Sciences of Niš, 20. Aleksandra Medvedeva, St, 18000 Niš, Serbia, e-mail: violeta.stojanovic@vtsnis.edu.rs
[2] Zoran Milivojević is with the College of Applied Technical Sciences of Niš, 20. Aleksandra Medvedeva, St, 18000 Niš, Serbia.

T_{R30}, T_C, C_{50}, D_{50}, AL_{cons}, *STI*, Just Noticable Difference in acoustic parameters' values detected by listeners, *JND*, was also calculated. The obtained values of these parameters were analysed in relation to the values defined by ISO 3382 standard [10]. Finally, we drew the conclusions concerning the acoustic suitability of the amphitheater. Recording of RIR and the analysis of the acoustic parameters was done using software package EASERA and Matlab.

The organization of this paper is as follows. In the section II the experiment is explained and the results are shown. In the section III, the analysis of the results of the analysed amphitheater is shown. The section IV is the conclusion.

II. EXPERIMENTS

The estimation of the acoustic suitability of the amphitheater at The College of Applied Technical Sciences in Niš (Serbia) is shown in Figure 1. The dimensions of the amphitheater are $X = 11.5$ m, $Y = 11.25$ m, $H = 5$ and the volume $V = 646.875$ m^3. The amphitheater capacity is $9 \times 17 = 153$ students. The reflexive coefficients of walls, ceilings and floor are 0.95, 0.85, 0.88, 0.88, 0.85 and 0.88 respectively. The air temperature is $t = 20$ ºC. Measuring of the impulse response is carried out using incentive log sweep signal with the duration of 6 s sampling frequency is $f_S = 44.1$ kHz. The equippment used for the experiment as follows: (a) an omnidirectional microphone (PCB 130D20), having a diaphragm diameter of 7mm; (b) a B&K omnidirectional sound source type 4295 (dodecahedron loudspeaker); (c) a B&K audio power amplifier, rated at 100W RMS, stereo, type 2716-C; (d) a laptop, incorporating a Soundmax Integrated Digital Audio sound card from Analog Devices.

A. The Basis

The data base is consisted of wav files obtained by measuring of RIR at 9 measuring points, marked with circles with ordinal numbers (Fig. 1.). There were 10 measurements for each measuring points which is total of 90 files.

B. The Results

The mean values and standard deviations of the acoustic parameters for 9 measuring points are shown Table I, Table II, Table IV and Table V. The mean values and standard deviations of these parameters for the amphitheater (for all measuring points) are shown in Table III and Table VI respectively. The mean values of the acoustic parameters and their standard deviations at frequencies $f = 500$ Hz, 1000 Hz are shown in Table VII and Table VIII respectively. In Fig. 2. the mean values and standard deviations for RT_{30}, C_{50}, *STI* and

AL_{cons} at measuring points are shown at all measuring points. The Fig. 3. shows the regression line for STI and RT_{10}.
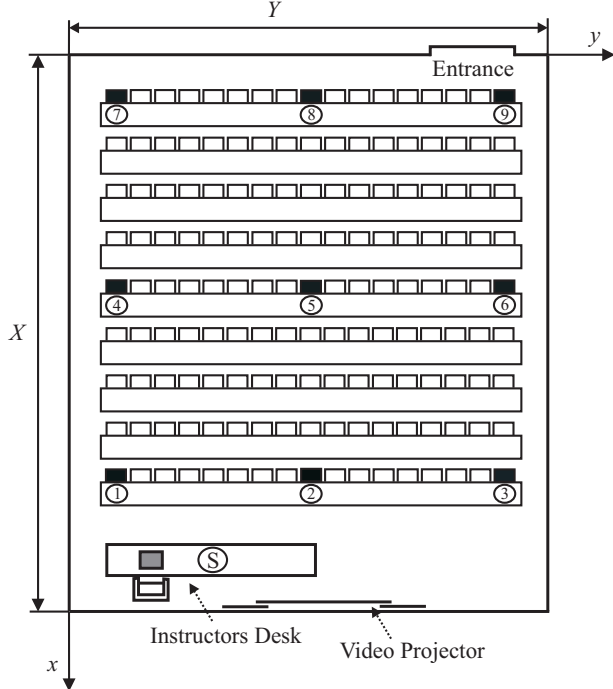

Fig. 1. The show of the room where the impulse response is measured: S - the location of the sound source, 1 - 9 - the location of the receiver.

TABLE I
THE MEAN VALUES OF THE ACOUSTIC PARAMETERS FOR THE MEASURED POINTS.

MP	$\overline{EDT}(s)$	$\overline{RT_{10}}(s)$	$\overline{RT_{20}}(s)$	$\overline{RT_{30}}(s)$
1	1.393	1.747	2.262	2.613
2	1.637	1.828	2.311	2.807
3	1.737	1.966	1.984	1.783
4	1.61	1.906	2.06	1.894
5	1.689	1.956	1.957	1.832
6	1.83	2.144	2.071	1.912
7	1.856	2.023	1.991	1.808
8	1.837	2.1	2.128	1.98
9	1.83	2.058	2.16	1.975

TABLE II
THE MEAN VALUES OF ACOUSTIC PARAMETERS FOR THE MEASURED POINTS.

MP	$\overline{T_C}(ms)$	$\overline{C_{50}}(dB)$	$\overline{STI}$	$\overline{AL_{cons}}(\%)$
1	56.944	4.14	0.551	8.61
2	71.21	3.13	0.53	9.665
3	131.261	-2.86	0.436	16.167
4	80.518	2.14	0.51	10.716
5	91.25	1.22	0.488	12.109
6	129.542	-2.07	0.422	17.337
7	109.148	-0.01	0.47	13.341
8	113.187	-0.43	0.449	14.921
9	129.381	-2.11	0.429	16.658

TABLE III
THE MEAN VALUES OF THE ACOUSTIC PARAMETERS FOR ALL MEASUREMENT POINTS.

	$\overline{EDT}(s)$	$\overline{RT_{10}}(s)$	$\overline{RT_{20}}(s)$	$\overline{RT_{30}}(s)$
MP	1.713	1.97	2.103	2.067
	$\overline{T_C}(ms)$	$\overline{C_{50}}(dB)$	$\overline{STI}$	$\overline{AL_{cons}}(\%)$
MP	101.382	0.35	0.476	13.28

TABLE IV
THE STANDARD DEVIATION OF ACOUSTIC PARAMETERS FOR THE MEASURED POINTS.

MP	$\sigma_{EDT}(s)$	$\sigma_{RT_{10}}(s)$	$\sigma_{RT_{20}}(s)$	$\sigma_{RT_{30}}(s)$
1	0.013	0.008	0.008	0.018
2	0.005	0.008	0.012	0.031
3	0.046	0.109	0.205	0.183
4	0.013	0.023	0.046	0.061
5	0.08	0.178	0.275	0.269
6	0.011	0.025	0.042	0.041
7	0.076	0.0166	0.25	0.236
8	0.08	0.0163	0.041	0.064
9	0.009	0.0193	0.031	0.086

TABLE V
THE STANDARD DEVIATION OF ACOUSTIC PARAMETERS FOR THE MEASURED POINTS.

MP	$\sigma_{T_C}(ms)$	$\sigma_{C_{50}}(dB)$	σ_{STI}	$\sigma_{AL_{cons}}(\%)$
1	0.3	0.052	0.0006	0.031
2	10.202	0.067	0.0008	0.032
3	1.882	0.334	0.0009	0.074
4	1.78	0.165	0.0011	0.066
5	3.912	0.042	0.0006	0.03
6	0.451	0.048	0.001	0.077
7	3.494	0.032	0.0007	0.05
8	0.574	0.082	0.0008	0.064
9	0.684	0.057	0.001	0.082

TABLE VI
THE STANDARD DEVIATION OF ACOUSTIC PARAMETERS FOR ALL MEASUREMENT POINTS.

	$\sigma_{EDT}(s)$	$\sigma_{RT_{10}}(s)$	$\sigma_{RT_{20}}(s)$	$\sigma_{RT_{30}}(s)$
MP	0.149	0.149	0.181	0.379
	$\sigma_{T_C}(ms)$	$\sigma_{C_{50}}(dB)$	σ_{STI}	$\sigma_{AL_{cons}}(\%)$
MP	26.432	2.35	0.0442	3.031

TABLE VII
THE MEAN VALUES OF THE ACOUSTIC PARAMETERS FOR ALL MEASUREMENT POINTS AT CERTAIN FREQUENCIES.

f(Hz)	$\overline{EDT}(s)$	$\overline{RT_{10}}(s)$	$\overline{RT_{20}}(s)$	$\overline{RT_{30}}(s)$
500	2.496	2.357	2.134	2.002
1000	2.242	2.214	2.001	1.857

f(Hz)	$\overline{C_{50}}(dB)$	$\overline{STI}$	$\overline{T_C}(ms)$	
500	-3.586	0.475	182.157	
1000	-3.212	0.475	163.285	

TABLE VIII
THE STANDARD DEVIATION OF ACOUSTIC PARAMETERS FOR ALL
MEASUREMENT POINTS AT CERTAIN FREQUENCIES.

f(Hz)	σ_{EDT} (s)	$\sigma_{RT_{10}}$ (s)	$\sigma_{RT_{20}}$ (s)	$\sigma_{RT_{30}}$ (s)
500	0.1166	0.18	0.238	0.334
1000	0.094	0.145	0.262	0.351

f(Hz)	$\sigma_{C_{50}}$ (dB)	σ_{STI}	σ_{T_C} (ms)
500	2.059	0.046	28.593
1000	1.515	0.046	17.414
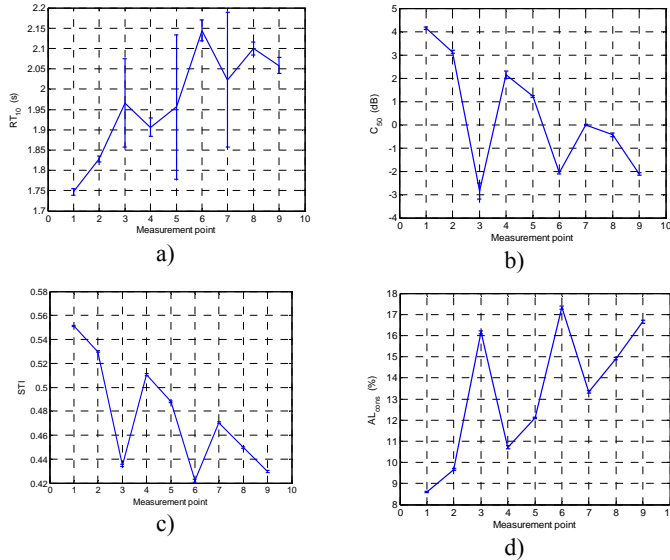
Fig. 2. The mean values and standard deviations at all measuring points for: a) RT_{10}, b) C_{50}, c) STI and d) AL_{cons}.
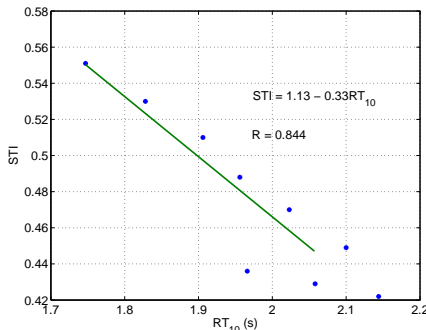
Fig. 3. The regression line for STI and RT_{10}.

III. RESULTS ANALYSIS

Based on the results shown in the Tables (I-VIII) and the Figures 2.and 3. the following can be concluded:

a) the values $\overline{RT_{10}}$ for the measuring points and for all measuring points within the amphitheater ($\overline{RT_{10}}$ = 1.747 – 2.144 s and $\overline{RT_{10}}$ = 1.97 s) in relation to the values $\overline{RT_{20}}$ and $\overline{RT_{30}}$, have the least standard deviations: $\sigma_{RT_{10}}$ = 0.08 – 0.109 s and $\sigma_{RT_{10}}$ = 0.149 s. $\overline{RT_{10}}$ is higher than the value of the optimal reverberation time RT_{opt} = 1.55 s which is proposed by ISO 12001 standard for auditoriums of certain volumes.

The permitted range of values RT for the auditoriums for general purposes (both speech and music) determined by standard is 1.4 – 1.9 s. The value $\overline{RT_{10}}$ = 1.97 s for all points within the amphitheater is the most proximate to the upper limit of this range with the least standard deviation ($\overline{RT_{20}}$ = 2.103 s, $\sigma_{RT_{20}}$ =0.181 s and $\overline{RT_{30}}$ = 2.067 s, $\sigma_{RT_{30}}$ = 0.379 s). The value $\overline{RT_{30}}$ (1000 Hz) = 1.857 s is also at the upper limit of this range, though $\sigma_{RT_{30}}$ (1000 Hz) = 3.77 JND. For $\overline{RT_{20}}$ (500 Hz) = 2.134 s and $\overline{RT_{20}}$ (1000 Hz) = 2.001 s are $\sigma_{RT_{20}}$ (500 Hz) = 2.23 JND and $\sigma_{RT_{20}}$ (1000 Hz) = 2.62 JND. The least standard deviations have $\overline{RT_{10}}$ (500 Hz) = 2.357 s and $\overline{RT_{10}}$ = 2.214 s: $\sigma_{RT_{10}}$ (500 Hz) = 1.53 JND and $\sigma_{RT_{10}}$ (1000 Hz) = 1.31 JND.

b) the values of $\overline{EDT}$ for measuring points are $\overline{EDT}$ = 1.61 – 1.856 s; their standard deviations have the values σ_{EDT} = 0.005 – 0.08 s. For all measuring points $\overline{EDT}$ = 1.713 s and σ_{EDT} = 0.149 s. These values significantly deviate from the defined range for both music and speech. However, the values of $\overline{EDT}$ at the medium frequencies belong to the range for music: 2.1 – 4.2s: $\overline{EDT}$ (500 Hz) = 2.496 s and $\overline{EDT}$ (1000 Hz) = 2.242 s. Their standard deviations are lower than 1 JND: σ_{EDT} (500 Hz) = 0.93 JND and σ_{EDT} (1000 Hz) = 0.84 JND.

c) the values $\overline{T_C}$ for the measuring points and their standard deviations are: $\overline{T_C}$ = 56.944 – 131.261 ms, σ_{T_C} = 0.3 – 10.202 ms. For all measuring points is $\overline{T_C}$ = 101.382 ms and σ_{T_C} = 26.432 ms. At measuring points 1,2 and 3, the values for $\overline{T_C}$ correspond to determined values for speech (60 – 80 ms) and at measuring points 4, 5, 6, 7, 8 and 9 are consistent to determined values for music (70 – 150 ms). At medium frequencies only the lower limit is $\overline{T_C}$ (1000 Hz) $\approx$ 150 ms, but other values are higher. The standard deviations don't exceed 1 JND (σ_{T_C} (500 Hz) = 0.2 JND, σ_{T_C} (1000 Hz) = 0.15 JND)

d) for the measuring points $\overline{C_{50}}$ and their standard deviations are $\overline{C_{50}}$ = -2.86 – 4.14 dB, $\sigma_{C_{50}}$ = 0.052 – 0.334 dB. For all measuring points are $\overline{C_{50}}$ = 0.35 dB and $\sigma_{C_{50}}$ = 2.35 dB. These values are consistent to determined values for speech ($C_{50} \geq$ -2 dB) but also meet the requirements determined for music (C_{80} = 0 ± 1.6 dB, for singers: C_{80} =1 – 5 dB). At medium frequencies the values of this parameter exclusively characterize the speech: $\overline{C_{50}}$ (500 Hz) = -3.586 dB, $\overline{C_{50}}$ (1000 Hz) = -3.212 dB. These values are consistent to standard deviation higher than 1 JND: $\sigma_{C_{50}}$ (500 Hz) = 2 JND, $\sigma_{C_{50}}$ (1000 Hz) = 1.5 JND.

e) the obtained values for STI are equable and confirm the acceptable speech intelligibility (0.45 – 0.6). For the measuring points of medium values standard deviations of this parameter are: $\overline{STI}$ = 0.422 – 0.55, σ_{STI} = 0.0006 – 0.0011 whereas for all measuring points $\overline{STI}$ = 0.476, σ_{STI} = 0.0442. At medium frequencies the values of this parameter are the same $\overline{STI}$ = 0.475 and σ_{STI} =0.046. Strong negative correlation between $\overline{RT_{10}}$ and $\overline{STI}$ (r = - 0.9189, R = 0.844), shown in Figure 3., provides as for the RT_{opt} estimated value for $\overline{STI}$ = 0.618, which confirms good speech intelligibility.

f) the obtained values for the parameter AL_{cons} for all measuring points are within the range of permitted values which show good speech intelligibility (7 – 15 %): $\overline{AL_{cons}}$ = 13.28%, $\sigma_{AL_{cons}}$ = 3.031%. Such standard deviation indicates bad speech intelligibility at certain measuring points. Those points are 3, 6 and 9. The best intelligibility is shown at the measuring point 1. The mean values of this parameter and their standard deviations are $\overline{AL_{cons}}$ = 8.61 – 17.337 %, $\sigma_{AL_{cons}}$ = 0.03 – 0.082.

Based on the performed analysis of the empty amphitheater it is concluded that $\overline{RT_{10}}$, $\overline{RT_{20}}$ and $\overline{RT_{30}}$ higher than RT_{opt}. The estimated values of the reverberation time for the full amphitheater can be obtained using reverberation time for the empty amphitheater and Schulc's diffusion time [11]. At medium frequencies they are: $\overline{RT_{30occ}}$ (500 Hz) = 1.59 s, $\overline{RT_{30occ}}$ (1000 Hz) = 1.5 s, $\overline{RT_{20occ}}$ (500 Hz) = 1.64 s, $\overline{RT_{20occ}}$ (1000 Hz) = 1.54 s, $\overline{RT_{10occ}}$ (500 Hz) = 1.68 s, $\overline{RT_{10occ}}$ (1000 Hz) = 1.61 s. All these values are within the range of values for auditoriums for general purposes and $\overline{RT_{20occ}}$ (1000 Hz) is the most proximate to the optimal reverberation time RT_{opt} = 1.55 s which is recommended by ISO 12001 standard [12] for the auditoriums of the volume which the analysed auditorium has. Strong negative correlation between $\overline{RT_{10}}$ i $\overline{STI}$ (r = - 0.9189), shown in Figure 3, gives the estimated values for $\overline{STI_{occ}}$ (500 Hz) = 0.576 $\approx$ 0.6 and $\overline{STI_{occ}}$ (1000 Hz) = 0.6 which are now limiting values of acceptable good speech intelligibility.

The acoustic analysis of the full amphitheater will be the issue of further studies.

IV. CONCLUSION

This study presents the acoustic parameters of the amphitheater calculated according to measured impulse responses. Measuring of the impulse responses was performed in the empty amphitheater.

Based on both measured and calculated acoustic parameters for the empty amphitheater, it is concluded that $\overline{RT_{10}}$ = 1.97 s (with the lowest standard deviation $\sigma_{RT_{10}}$ = 0.149 s) is the most proximate to the upper limit of the values recommended by ISO 12001 standard for the auditoriums for general purposes (both speech and music). The mean values of the acoustic parameters which are characteristic for speech transmission in an empty amphitheater confirm acceptable good speech intelligibility: $\overline{STI}$ = 0.476, $\overline{AL_{cons}}$ = 13.28%. Based on the estimation for the full ampfitheater, it is concluded that $\overline{RT_{occ}}$ (500 Hz) = 1.59 – 1.68 s i $\overline{RT_{occ}}$ (1000 Hz) = 1.5 – 1.61 s also include the optimal reverberation time which is recommended by ISO 12001 standard for the auditoriums of certain volume. The reduction of the reverberation time value resulted in expected increase of the value of the STI: $\overline{STI_{occ}}$ (500 Hz) $\approx$ 0.6 and $\overline{STI_{occ}}$ (1000 Hz) = 0.6 that confirmed acceptably - good speech intelligibility of the full amphitheater.

Based on the previous analysis, it is concluded that the analysed amphitheater for lecturing is suitable for its purpose, but for the sake of acoustic improvement, additional acoustic treatment can be performed.

REFERENCES

[1] Sabine, W. Collected Papers on Acoustics (1922). (Harvard University Press.Reimpresión Dover, 1964).
[2] Eyring, C.F. "Reverberation Time in "Dead" Rooms", J. ACOUST. SOC. AM., Vol. 1 217-241, 1930.
[3] Sette, W. H. "A New Reverberation Time Formula", J. ACOUST. SOC. AM., Vol. 4, pp 193-210, 1933.
[4] Schroeder, M.R.: New method of measuring reverberation time JASA 38 (1965), S. 329 and 40, S. 549.
[5] Jordan, V.L., 1970. "Acoustical Criteria for Auditoriums and Their Relation to Model Techniques", - JASA, Vol. 47, No. 2 (Part 1), pp. 408-412.
[6] Kürer, R.: Einfaches Messverfahren zur Bestimmung der Schwerpunktzeit"raumakustischer Impulsantworten (A simple measuring procedure for determining the "center time" of room acoustical impulse responses), 7th INTER. CONGRESS ON ACOUSTICS, Budapest 1971.
[7] Peutz, V.M.A- Articulation Loss of Consonants as a criterion for Speech Transmission in a Room - J. AUDIO – ENG. SOC. Vol. 19, p. 915-919, 1971.
[8] Abdel Alim, O.: Abhängigkeit der Zeit - und Register- durchsichtigkeit von raumakustischen Parametern bei Musikdarbietungen (Dependence of time and register definition of room acoustical parameters with music performances) Dissertation, TU Dresden 1973.
[9] Steeneken, H. J. M. – A Physical Method for measuring Speech Transmission Quality - J. ACOUST. SOC. AM., Vol. 19, 1980.
[10] ISO 12001:1998, Acoustics Noise emitted by machinery and equipment, Rules for the drafting and presentation of a noise test code, 1998.
[11] M. Topa, N. Toma, B. Kirei, I. Saracut and A. Farina, Experimental Acoustic Evaluation of an Auditorium, HINDAWI PUBLISHING CORPORATION ADVANCES IN ACOUSTICS AND VIBRATION, Volume 2012.
[12] ISO 12001:1998, Acoustics Noise emitted by machinery and equipment, Rules for the drafting and presentation of a noise test code, 1998.

RBF Neural Network and Filter Methods for Feature Selection in Medical Classification Problems

Jasmina Đ. Novakovic [1] Alempije Veljovic [2] and Sinisa S. Ilic [3]

Abstract – **In this paper we evaluate classification accuracy of radial basis function (RBF) neural network and filter methods for feature selection in medical data sets. To improve diagnostic process in everyday routine and avoid misdiagnosis, machine learning methods can be employed. Diagnosis of tumor, heart diseases, hepatitis, liver and Parkinson diseases are some of the medical problems in which we used artificial neural networks. The main objective of this paper is to show that it is possible to improve the performance of the system for inductive learning rules with RBF neural network for medical classification problems, using the filter methods for feature selections. The goal of this research is also to present and compare different algorithmic approaches for constructing and evaluating systems that learn from experience to make the decisions and predictions and minimize the expected number or proportion of mistakes.**

Keywords – **Classification problems, Feature selection, Filter methods, RBF.**

I. INTRODUCTION

A lot of information is available to medical specialists, ranging from details of clinical symptoms to various types of biochemical data and outputs of imaging devices. To modernize the diagnostic process in daily routine and avoid misdiagnosis, machine learning methods can be employed. These algorithms can handle diverse types of medical data and integrate them into categorized outputs. In machine learning, a wide range of algorithms for classification is available, each with their own strengths and weaknesses. One of the most commonly used algorithms for diagnosing the disease is artificial neural networks [1, 2, 3, 4]. Coronary heart disease diagnosis [5, 6], tumor diagnosis [7], diagnosis of hypoglycemic episodes [8], prognosis of chronic myeloid leukemia [9], tuberculosis disease diagnosis [10], predict thyroid bending protein diagnosis [11], prediction approach of carcinoma patients [12], control of blood glucose [13], colorectal cancer metastases [14], detection of ovarian cancer [15] are some of the medical problems in which the used artificial neural networks.

In this paper, we used RBF neural network as artificial

[1]Jasmina Đ. Novakovic, professor of applied studies, Belgrade business school, Higher education institution for applied science, Kraljice Marije 73, 11000 Belgrade, Serbia, jnovakovic@sbb.rs.

[2]Alempije Veljovic, full professor, Faculty of technical science Cacak, University of Kragujevac, Svetog Save 65, 32000 Cacak, Serbia.

[3]Sinisa S. Ilic, associate professor, Faculty of technical science in Kosovska Mitrovica, University of Pristina, Kneza Milosa 7, 38220 Kosovska Mitrovica, Serbia.

neural networks. RBF neural network offers a number of advantages, including requiring less formal statistical training, the ability to implicitly detect complex nonlinear relationship between the dependent and independent variables, the ability to detect all possible interactions between predictor variables and the availability of multiple algorithms for training.

The main aim of this paper was to experimentally verify the impact of filter methods as one of the feature selection technique on classification accuracy with RBF neural network. A process that chooses a minimum subset of M features from the original set of N features, so that the feature space is optimally reduced according to a certain evaluation criterion can be defined as feature selection. Finding the best feature subset is usually intractable and many problems related to feature selection have been shown to be NP-hard.

For this purpose we have organized the paper in the following way. In the second part of this paper we present a model of RBF neural network, in the third part of the paper we present description of data sets. The fourth part of the paper describes the methodology of the experiment and also we present the results of experimental study. In the fifth part of the paper, we discuss the obtained results and give directions for further research.

II. RBF NEURAL NETWORK

The classification of neural networks proved to be very good just for serious classification problems, problems where is difficult or impossible to use the classical technique. Besides, neural networks are well suited to work in conditions of noise in the data. The above listed reasons are conditioned to use neural networks.

From a structural point of view, depending on the model used to build neural networks, neural networks can be divided into static and dynamic. From the point of view of a layered mode of organization of neurons in the network, the network can be divided into single-and multi-layered. The first layer is called the input layer, the last layer output, and all other layers are called hidden layers. As a rule, each layer receives inputs from the previous layer and sends their outputs on a next layer.

The main characteristic of static neural networks is that the neurons are organized beforehand, so that neurons connected in a way with no form of feedback. These networks can not contain dynamic members, making them structurally stable. Since there are no dynamic members, static response of the neural network depends only on the current state and input values of the network parameters. Static neural networks are commonly used in the identification process, process management, and signal processing and pattern recognition. The most common types of static neural networks are MLP

and RBF neural networks. In this paper we are used the RBF neural network.

Dynamic neural networks are used much less frequently than static neural networks because their stability is not guaranteed. The learning process of the dynamic neural networks is much more demanding than learning static. Another shortcoming is that the output dynamic neural network depends on its initial conditions making it difficult to study because of poorly selected initial conditions may give poor results.

In the case where the application of static neural network does not give good results it is reasonable to use the dynamic neural network. The most common types of dynamic neural networks are the Hopfield, Elman and NARX dynamic neural network. Pseudo code for RBF training [16] is shown in Fig. 1.

```
trainRBF (in, out, width, MaxError, data) {
hidden = 0;
net = initRBFNetwork (in, out, hidden);
do {
        / / find the vector data that produces the highest error
        i = findMaxNetworkError (data, net); / / i = index
        vectors
            / / add a layer of RBF neurons in the same
            place where the vector data addRBFNeuron
            (net, width, data (i)); / / data (i) = midpoint
        / / find the overall network error
        NetError = trainOutputWeights (net, data);
} while (NetError> MaxError);
}
```

Figure 1. Pseudocode for RBF training

III. DESCRIPTION OF DATA SETS

Nine real medical data sets were used for tests, taken from the UCI repository of machine learning databases [17]. We used these data sets to compare results of classification accuracy with RBF network. In the following, we provide the details for the benchmark data sets we have used from UCI repository of machine learning databases.

Breast cancer (bc): The task of this data set is to predict whether or not there is recurrence of breast cancer. In this data set there is a single instance of the class 201 (no recurrence of breast cancer), and 85 instances of other classes (a recurrence of breast cancer).

Cardiography (ct): This set of data consists of attribute measurement of fetal heart rate and uterine contractions attributes on ultrasound that are classified doctors [18]. This data set contains 2126 instances and 23 attributes.

Hepatitis (he): The main aim of this data set is to predict whether hepatitis patients will die or not. In this data set, there are two classes: live (123 instances) and die (32 instances).

Liver (li): In this data set, the first five variables are all blood tests, which are thought to be sensitive to liver disorders that might arise from excessive alcohol consumption. Each row in this data set constitutes the record of a single male individual.

Lung cancer (lc): A set of data for the cancer of the lung contains data describing the three kinds of pathological forms of lung cancer. There are 32 instances and 56 attributes.

Mammographic mass (ma): The task is to predict the severity (benign or malignant) of a mammographic mass lesion from BI-RADS features and the patient's age [19].

Parkinson (pa): This data set consists of a range of biomedical voice measurements in 31 persons, 23 of them suffering from Parkinson's disease [20]. Each column in the table is a distinctive feature of a person's voice, and each row corresponds to one of the 195 recordings of person's voice. The main goal of this dataset is to separate healthy people from those people who are suffering from Parkinson's.

Pima Indians diabetes (pi): In this data set the diagnostic is whether the patient shows signs of diabetes according to World Health Organization criteria (i.e., if the 2 hour post-load plasma glucose was at least 200 mg/dl at any survey examination or if found during routine medical care).

Statlog Heart (sh): The task is to predict absence or presence of heart disease. This data set contains 13 features (which have been extracted from a larger set of 74).

TABLE I
REPRESENTATION OF DATA SETS

Data set	Attributes			Number of class	Size of training	Default accuracy
	all	categorical	numerical			
bc	9	9	0	2	286	70.30
ct	23	0	23	3	2126	95.00
he	19	13	6	2	155	78.10
li	6	0	6	2	345	58.10
lc	56	0	56	3	32	26.80
ma	5	0	5	2	961	84.00
pa	23	0	23	2	195	76.00
pi	8	0	8	2	768	65.10
sh	13	3	10	2	270	55.00

Table I shows the comparative characteristics of the observed data sets. There are 9 real data sets relating to diagnosis and prognosis of medical diseases. They were obtained by collecting data from real systems that exist. The default accuracy is the accuracy of always predicting the majority class on the whole dataset.

In addition, the observed total number of attributes in each set of data, as well as the number of attributes which belong to the category of categorical and numerical attributes. Three sets of data has more than 20 attributes, *lc* with 56 attributes, *ct* and *pa* with 23 attributes. Data sets with minimum attributes are *ma* with 5 attributes and *li* with 6 attributes. We conclude that the observed data sets are the sets with a very large number of attributes, as well as those sets that have a small number of attributes, which is good from the perspective of research. The observed data sets are balanced because there are sets that contain only or categorical or numerical attributes, as well as datasets containing both categorical and numerical data.

Regarding the number of classes in the analyzed data sets, 7 sets of data has two classes; only two data sets have three classes, namely *ct* and *lc*. The reason for this is the fact that in most cases in the classification problems, instances classify in two, possibly three classes, and rarely in a larger number of classes.

In Table I, shows that the number of instances provided for training varies from a small number of collected instances, which is the case with *lc* which has only 32 instances, to events that have a much greater number of instances such as *ct* which has 2126 training instance. Regarding the size of the testing set, initially in all real datasets, we had prepared a set of data for testing using 10-fold cross-validation.

IV. RESULTS OF EXPERIMENTAL RESEARCH

The experiment was performed using WEKA (Waikato Environment for Knowledge Analysis) tools for data preparation and research developed at the University of Waikato in New Zealand. Weka contains tools for data pre-processing, classification, regression, clustering, association rules, and visualization. It is also well-suited for developing new machine learning schemes.

When searching for the model that best approximates the target function, it is necessary to provide measures of quality models and learning. Different measures can be used depending on the problem, in our experimental studies; we used the classification accuracy as a measure of the quality of the model. In experimental research we used filter method to reduce the dimensionality of data. In all experiments is selected the solution with the number of attributes that will be used further in the study, which gives the highest classification accuracy. Our results provide the accuracy that is obtained as the average of ten repetitions each time with a 10-fold cross-validation.

In our experimental research, we used Paired *t*-test, where the level of significance was set to a value of 0.05. We use Paired *t*-test to determine whether the value obtained by different methods differs significantly. Paired *t*-test test the significance of the mean differences pairs d according to the following equation: $t = \frac{\bar{d}\sqrt{N}}{s_d}$, where s_d is standard deviation of the obtained differences. If the calculated value of the parameter t is greater than tabular (critical value), the null hypothesis is rejected and it is said that d is significantly different from zero, or that the difference in pairs statistically significant.

In the table of classification accuracy "+" indicates a significantly higher value for classification accuracy, while "-" indicates a significantly lower value for classification accuracy. In our experimental research, whenever we compare two or more algorithms, we give a table of classification accuracy. Comparison is such that the second algorithm is an algorithm in which was performed pre-selection attributes, and the first algorithm is a standard algorithm without pre-selection of attributes.

In this paper, we use the following filtering methods for ranking attributes that are statistically and entropy-based, and show good performance in various domains: IG, GR, SU, RF, OR and CS. All methods of filtering, IG, GR, SU, RF, OR and CS have done the ranking of attributes for each data set. Considering that the method of ranking lists all the attributes in the order as to their importance for classification problem, this method does not perform automatic reduction of the number of attributes. In order to realize the reduction of the number of attributes, there are two possibilities: (1) the use of a threshold, or (2) the use of an appropriate number of attributes for each data set and each of the filtering methods. In our experiment we use second possibility.

TABLE II
THE NUMBER OF ATTRIBUTES IN THE ORIGINAL DATASET AND NUMBER OF ATTRIBUTES SELECTED BY THE FILTER METHOD

Data set	Original data set	IG	GR	SU	RF	OR	CS
bc	9	8	3	3	2	8	3
ct	23	18	12	8	21	19	6
he	19	1	1	5	6	2	1
li	6	4	4	4	4	5	4
lc	56	5	17	2	4	9	4
ma	5	3	2	2	4	2	2
pa	23	21	21	21	13	22	21
pi	8	1	4	1	1	4	1
sh	13	3	9	3	6	3	3

Table II shows the optimal number of attributes for the purposes of classification, after searching the set of all possible solutions for each method. The table shows the original size of the set, in order to compare the effects of the reduction of the dimensionality of data. Using filter methods, the six data sets, from 9 observed, reduce the number of attributes exactly half or more than half compare with the original data set. These data sets are *bc*, *he*, *lc*, *ma*, *pi* and *sh*.

The greatest benefit of reducing the dimensionality of the data set has *lc*, with 56 attributes. These methods for *lc* data set selected a small number of attributes, even less than one-sixth, for each method, except GR method. For *he* data set, which originally has 19 attributes, all filtering methods show that the most important attribute for this studied phenomenon is 6. Filtering methods for data collection *pi*, show that up to 4 attributes are important for the classification problem.

TABLE III
CLASSIFICATION ACCURACY OF RBF ALGORITHM FOR ORIGINAL AND REDUCED DATA SET

Data set	RBF	RBF_IG	RBF_GR	RBF_SU	RBF_RF	RBF_OR	RBF_CS
bc	71.41	71.34	74.32	74.46	71.00	71.20	73.62
ct	97.93	98.35	98.41	97.65	98.13	96.90	96.27 -
he	85.29	81.31	83.45	83.05	80.49	82.69	81.25
li	65.06	58.16 -	58.16 -	58.16 -	57.33 -	60.96	58.16 -
lc	76.00	73.58	79.75	79.00	76.75	72.92	74.92
ma	77.31	77.66	79.67	79.24	77.07	77.51	79.16
pa	81.22	80.92	80.92	80.92	83.39	81.98	80.67
pi	74.04	73.84	76.28	73.84	73.84	75.32	73.84
sh	83.11	78.44	83.44	78.15	81.56	78.44	78.52

Further experimental research, the optimal number of selected attributes for each data set and filtering method, checked the accuracy of the classification algorithm using RBF network.

In none data set, we have significantly worse data for all filtering methods, which means that we can always choose the

method for a given set of data that has better results or results that approximate the original dataset (table III). Using RBF classifier, we can conclude that the GR method of filtering in most cases led to better results in the observed data sets (in six cases have better results). Only one data set (ma), with all implemented filter methods, achieves less than the default accuracy for this data set (84.0%). In all other data sets, values of classification accuracy with at least one filter method are greater than the default accuracy.

TABLE IV

THE STANDARD DEVIATION FOR THE ACCURACY OF RBF ALGORITHM WITH ORIGINAL AND REDUCED DATA SET

Data set	RBF	RBF_IG	RBF_GR	RBF_SU	RBF_RF	RBF_OR	RBF_CS
bc	7.88	7.66	6.41	6.16	6.81	8.34	6.28
ct	1.02	1.02	0.94	1.21	1.00	2.19	1.29
he	8.29	7.38	8.54	7.90	9.44	8.25	7.51
li	8.80	8.10	8.10	8.10	7.66	9.62	8.10
lc	22.91	22.91	21.10	20.70	23.31	22.17	22.52
ma	3.31	3.67	4.14	4.51	3.83	4.35	4.50
pa	7.37	7.49	7.42	7.53	7.39	7.24	7.35
pi	4.91	4.65	5.18	4.65	4.65	5.31	4.65
sh	6.50	7.28	6.44	7.25	7.29	7.13	7.28

Table IV shows the standard deviation for the classification accuracy of RBF algorithm with original and reduced data set using filter methods. From the table it can be seen that the standard deviations generally do not differ much between the standard algorithm and algorithms that use reduced data set.

V. DISCUSSION OF RESULTS AND FUTURE RESEARCH

According to the obtained results, we can conclude that the basic hypothesis was proved - it is possible to improve the system performance of inductive learning rules in the medical classification problems, using the filter method for reducing the dimensionality of the data. To prove the hypothesis, have been implemented and empirically tested filter methods for reducing the dimensionality of the data.

In further research it would be interesting to apply other techniques to solve the problem of dimensionality reduction of data, such as wrapper methods and extraction of attributes and analyze and compare the effects of their implementation. These techniques could also improve the performance of classification learning algorithms.

ACKNOWLEDGEMENT

REFERENCES

[1] F. Amato, A. López, P. Vaňhara, A. Hampl, J. Havel, "Artificial neural networks in medical diagnosis", Jornal of Applied Biomedicine 11: 47–58, 2013.

[2] E. Alkim, E. Gürbüz, E. Kiliç, "A fast and adaptive automated disease diagnosis method with an innovative neural network model", Neur Networks. 33: 88–96, 2012.

[3] D. Brougham, G. Ivanova, M. Gottschalk, D. Collins, A. Eustace, R. O'Connor, J. Havel, "Artificial neural networks for classification in metabolomic studies of whole cells using 1H nuclear magnetic resonance", J Biomed Biotechnol. 2011: 158094, 2011.

[4] M. Catalogna, E. Cohen, S. Fishman, Z. Halpern, U. Nevo, E. Ben-Jacob, "Artificial neural networks based controller for glucose monitoring during clamp test", PloS One. 7: e44587, 2012.

[5] O. Atkov, S. Gorokhova, A. Sboev, E. Generozov, E. Muraseyeva, S. Moroshkina S, N. Cherniy, "Coronary heart disease diagnosis by artificial neural networks including genetic polymorphisms and clinical parameters", J Cardiol. 59: 190–194, 2012.

[6] H. Uğuz, "A biomedical system based on artificial neural network and principal component analysis for diagnosis of the heart valve diseases", J Med Syst. 36: 61–72, 2012.

[7] D. Barbosa, D. Roupar, J. Ramos, A. Tavares, C. Lima, "Automatic small bowel tumor diagnosis by using multi-scale wavelet-based analysis in wireless capsule endoscopy images", Biomed Eng Online. 11:3, 2012.

[8] K. Chan, S. Ling, T. Dillon, H. Nguyen, "Diagnosis of hypoglycemic episodes using a neural network based rule discovery system", Expert Syst Appl. 38: 9799–9808, 2011.

[9] P. Dey, A. Lamba, S. Kumari, N. Marwaha, "Application of an artificial neural network in the prognosis of chronic myeloid leukemia", Anal Quant Cytol Histol. 33: 335–339, 2012.

[10] E. Elveren, N. Yumuşak, "Tuberculosis disease diagnosis using artificial neural network trained with genetic algorithm", J Med Syst. 35: 329–332, 2011.

[11] AS. Gannous, YR. Elhaddad, "Improving an Artificial Neural Network Model to Predict Thyroid Bending Protein Diagnosis Using Preprocessing Techniques", WASET. 50: 124–128, 2011.

[12] Ho W-H, Lee K-T, Chen H-Y, Ho T-W, Chiu H-C., "Disease-free survival after hepatic resection in hepatocellular carcinoma patients: a prediction approach using artificial neural network", PloS One. 7: e29179, 2012.

[13] BS. Leon, AY. Alanis, E. Sanchez, F. Ornelas-Tellez, E. Ruiz-Velazquez, "Inverse optimal neural control of blood glucose level for type 1 diabetes mellitus patients", J Franklin I. 349: 1851–1870, 2012.

[14] L. Spelt, B. Andersson, J. Nilsson, R. Andersson, "Prognostic models for outcome following liver resection for colorectal cancer metastases: A systematic review", Eur J Surg Oncol. 38: 16–24, 2012.

[15] A. Thakur, V. Mishra, S. Jain, "Feed forward artificial neural network: tool for early detection of ovarian cancer", Sci Pharm. 79: 493–505, 2011.

[16] S. Russell, P. Norvig, Artificial Intelligence: A Modern Approach, Second edition, Prentice Hall, 2003.

[17] A. Frank, A. Asuncion, UCI Machine Learning Repository, Irvine, CA: University of California, School of Information and Computer Science, 2010, [http://archive.ics.uci.edu/ml].

[18] D. Ayres de Campos, et al., "SisPorto 2.0 A Program for automated analysis of cardiotocograms", J Matern Fetal Med 5:311-318, 2000.

[19] M. Elter, R. Schulz-Wendtland, T. Wittenberg, "The prediction of breast cancer biopsy outcomes using two CAD approaches that both emphasize an intelligible decision process", Medical Physics 34(11), pp. 4164-4172, 2007.

[20] M.A. Little, P.E. McSharry, S.J. Roberts, D.A.E. Costello, I.M. Moroz, "Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection", BioMedical Engineering OnLine, 6:23, 2007.

POSTER SESSION
MEASUREMENT SCIENCE AND TECHNOLOGY

High Voltage Indicator with a Fiber-optic System

Emil Barudov[1]

Abstract – **Providing management of the power distribution grids and the equipment in the electric power system, in order to reduce the time of power outages in emergency situations, requires continuous monitoring for the presence of the respective voltage in certain points of the system.**

The work is dedicated to the development and research of a voltage indicator, designed for medium voltage grids, which is based on a sensor – voltage-to-light pulse frequency converter. The sensor gives a possibility for galvanic separation from the receiver on the basis of a fiber-optic system. This allows operation in a medium with strong electromagnetic interference, remoteness of the sensor from the receiver and additional including of a system for processing of the received result from the measurement.

Keywords – **high-voltage indication, fiber-optic systems.**

I. INTRODUCTION

The purpose of the present work is to be synthesized a device for high-voltage indication for work in grids over 1000V with a transmitter located next to the live parts under voltage and operating without the need of operational supply voltage and with a communication channel between the transmitting and the receiving parts based on a fiber optic line.

Different devices are known in the literature [1, 3, 4, 5], for indication of high voltage, which include capacitive divider, converter and indicator - a glyme lamp. Each of them has some advantages and disadvantages. In some cases disadvantages of these devices are the absence of galvanic isolation between the high voltage source and the indicator, low light indicator intensity and low selectivity at the presence of high level electromagnetic interference.

An example structure of such device for indication is shown in Fig.1, where:
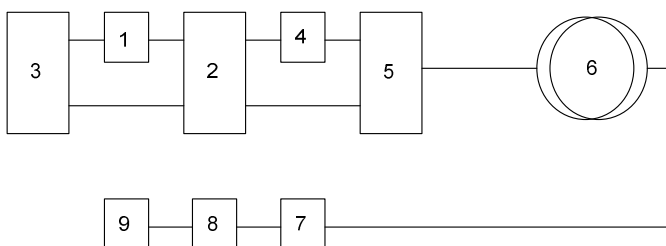
Fig.1 Schematic circuit of a device for measurement of high voltage on the base of a fiber-optic system

- 1 – ballast;

[1]Emil Barudov is with the Engineering Faculty of Naval Academy "N. Vaptsarov", 73 Vasil Drumev Str, Varna 9026, Bulgaria, E-mail: ugl@abv.bg.

- 2 – rectifying section with accumulating capacitor;
- 3 – high voltage source;
- 4 – switching element with a fixed threshold of activation;
- 5 – emitting LED;
- 6 – fiber-optic line;
- 7 – photodiode converter;
- 8 – trigger expander;
- 9 – indicator.

The principle scheme of the transmitting part is shown in Fig.2:
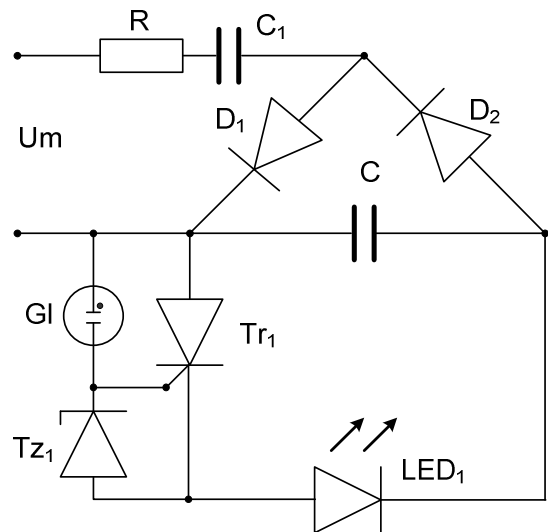
Fig.2. Principle scheme of the transmitter in a device for indication of high voltages

The transmitting part of high voltage indicator is a voltage to light pulse frequency converter. The frequency of the light pulses depends on the charging of the capacitor C to a voltage at which occurs a discharge in the glyme lamp Gl, triggering on the thyristor Tr_1 and forming a light pulse by the LED_1.

II. EXPERIMENTS

In Fig.3 is shown the principle scheme of the voltage doubler.
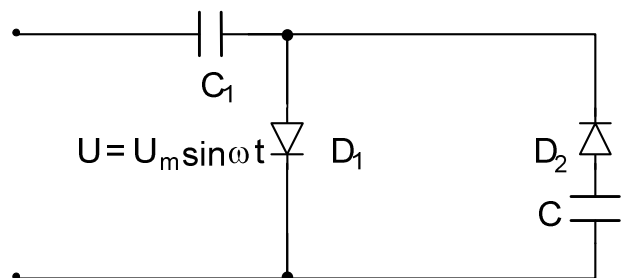
Fig.3 Principle scheme of a voltage doubler

In Fig.4 is presented the time diagram of the voltage change in the capacitor C from Fig.3.

For a random moment in the n-th pulse from the charging of C is valid Eq.1, [2]:

$$u_{Cn} + u_{C1n} = u = U_m \sin(\omega t + \varphi_n) \qquad (1)$$

where: φ_n is dephasing of the moment of triggering on the diode.

After differentiation of Eq.1 is obtained Eq.2:

$$i_{3n} = C \frac{du_{Cn}}{dt} = C_1 \frac{du_{C1n}}{dt} \qquad (2)$$
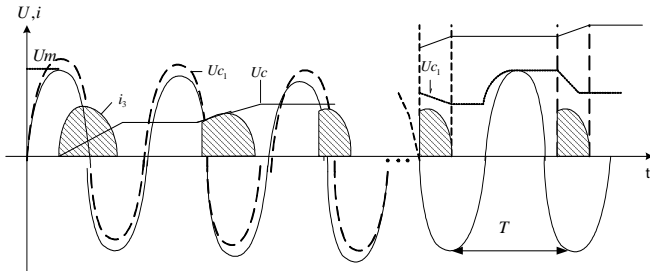
Fig.4. Time diagram of the voltage change of the capacitor C from Fig.3

Then, for the charge current through a triggered on diode D_2 is valid Eq.3:

$$i_{CHn} = U_m \omega \frac{C_1.C}{C_1 + C} \cos(\omega t + \phi_n) \qquad (3)$$

At the moment of triggering on the diode D_2 is valid:

$$|u_{Cn}(0)| - |u_{C1}(0)| = U_m \sin \varphi_n \qquad (4)$$

At this moment $u_{C1n}(0) = U_m$, therefore:

$$u_{Cn}(0) - U_m = U_m \sin \varphi_n \qquad (5)$$

from where:

$$\sin \varphi_n = \frac{u_{Cn}(0)}{U_m} - 1 \qquad (6)$$

The voltage u_{Cn} at the moment of passing of the n-th charge current pulse can be defined by Eq.7:

$$u_{Cn} = \frac{1}{C} \int i_{3n} dt = \frac{U_m C_1}{C_1 + C} \sin(\omega t + \varphi_n) + K \qquad (7)$$

Here, K is an integrating constant and is defined by the border conditions.

The charge current appears at $\omega t = 0$ and vanishes at $\omega t + \varphi_n = \pi/2$.

$$\left. \begin{aligned} u_{Cn}(0) &= \frac{U_m C_1}{C_1 + C} \sin \varphi_n + K \\ u_{Cn+1}(0) &= \frac{U_m C_1}{C_1 + C} + K \end{aligned} \right\} \qquad (8)$$

Replacing $\sin\varphi_n$ from Eq.6 in Eq.8, it can be found the increase of the voltage, to which C is charge in the n-th interval of the charging process – Eq.9:

$$\Delta u_{Cn} = u_{Cn+1}(0) - u_{Cn}(0) = \frac{C_1}{C + C_1}[2U_m - u_{Cn}(0)] \qquad (9)$$

After transformation of Eq.9 is obtained:

$$\frac{u_{Cn+1}(0)}{2U_m} = 1 - A\left[1 - \frac{u_{Cn}(0)}{2U_m}\right] \qquad (10)$$

where: $A = \dfrac{C_1}{C}$. In this case:

$$\frac{u_{C2}(0)}{2U_m} = 1 - (A)^1$$

$$\frac{u_{C3}(0)}{2U_m} = 1 - (A)^2 \qquad (11)$$

$$\frac{u_{Cn}(0)}{2U_m} = 1 - (A)^{n-1}$$

At a ration $C_1 : C \approx 1:1000$ the increase step of the voltage over the capacitor C - Δu_{Cn} for each period of the input supply voltage is constant and can be presented with Eq.12:

$$\Delta u_{Cn} = 2U_m \frac{C_1}{C} \qquad (12)$$

At $U_m = 14\ 100V$, the dependency $u_{Cn}(0) = f_1(t)$ is shown in Fig.5.

The amplitude of the charge current pulse can be defined by Eq.13:

$$I_{CHm} = U_m \omega \frac{C_1.C}{C_1 + C} \qquad (13)$$
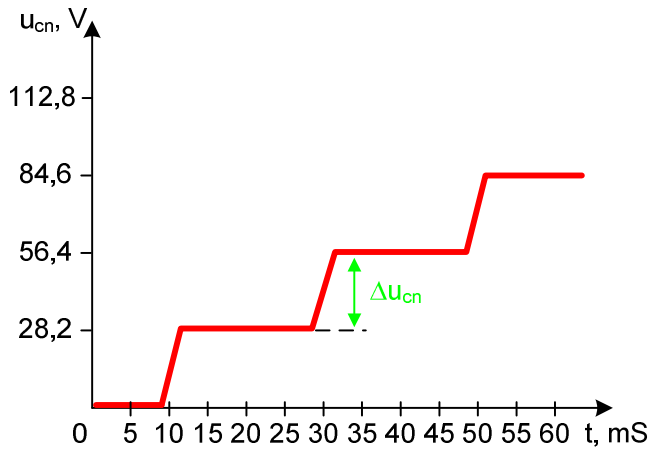
Fig.5. Change of $u_{Cn} = f_1(t)$ at $\Delta u_{Cn} = const$.
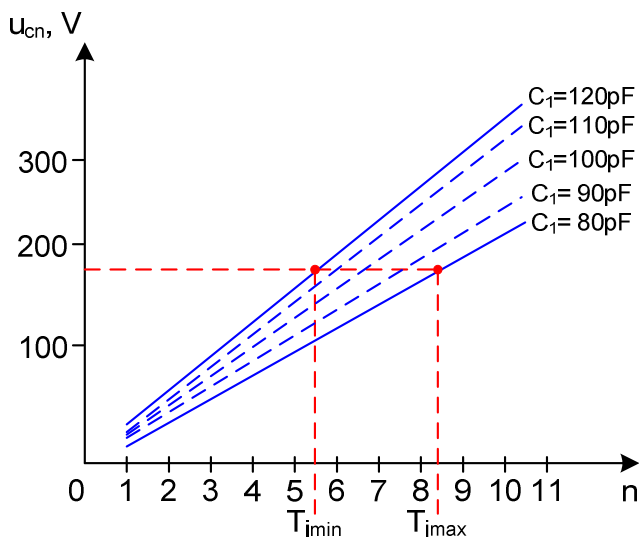
Fig.6. Change of $u_{Cn} = f_2(n)$ for C=100nF and U=10kV at a parameter C_1.
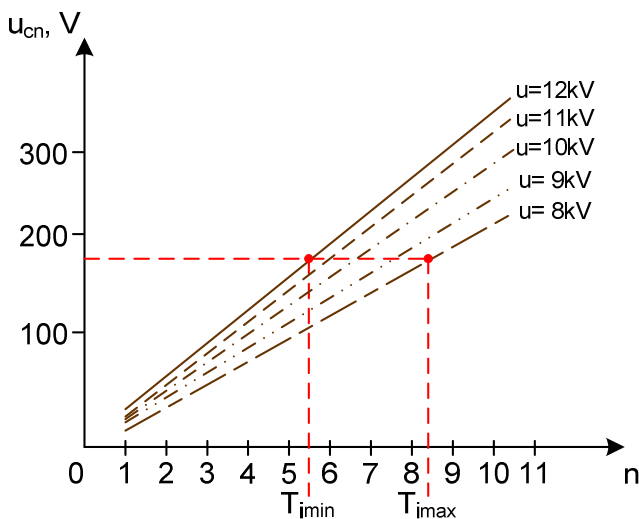
Fig.7. Change of $u_{Cn} = f_3(n)$ for C=100nF and C_1=100pF at a parameter U.

In the suggested analytical model [2] are not covered the differences in the charging process depending on which of the diodes D_1 or D_2 is triggered on during the first half wave of the charging process. It is also not considered the issue about the non-zero initial conditions for the charging process. For the proposed converter, representing the transmitting part of the high voltage indicator, the pointed limitations are not of significant matter, because at the formation of the light pulse the capacitor C (Fig.2) is discharged practically to 0V.

In Fig.8 is shown the connection of a high voltage transmitter and fiber optic line to a 20kV electric pole. This allows operation in a medium with strong electromagnetic interference, remoteness of the sensor from the receiver and safety for the operators. Here the transmitter operates without need of operational supply voltage, which is one of the advantages of the presented device.
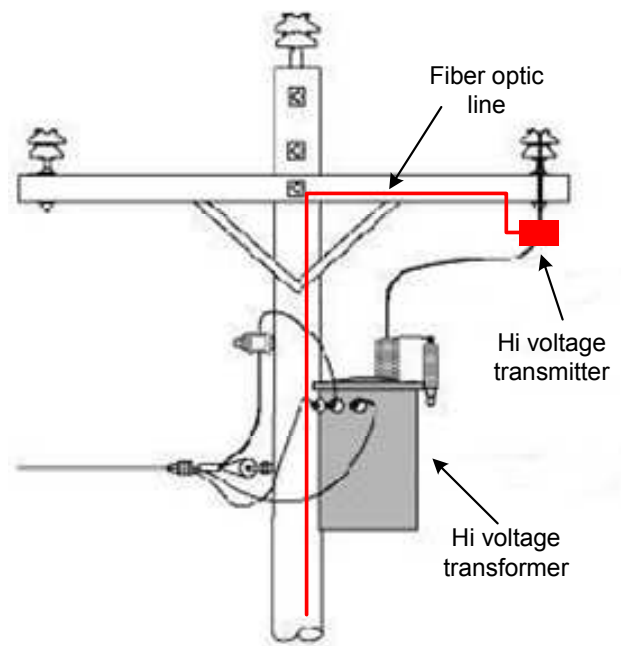
Fig.8. Connection of a high voltage transmitter and fiber optic line to a 20kV electric pole

III. CONCLUSION

From the conducted analysis can be summarized:

1. On the basis of the voltage-to-light pulses frequency converter from Fig.2 is possible the realization of devices for high voltage indication with a fiber-optic system with the following advantages:
   - A possibility for movement of the indicating part at a considerable distance;
   - Insensitivity to electromagnetic interference;
   - High level of the output signal.

2. Independently from the linear connection "amplitude of the input supply voltage – charge voltage of the

accumulating capacitor u_{cn}", the application of the proposed converter for measurement of high voltages could be difficult because of the sudden increase of u_{Cn} for few milliseconds in the range of one period of the supply voltage - 20mS.

REFERENCES

[1] Барудов С., „Теория и практика построения аппаратуры управления электрическим разрядом", Санкт – Петербург, ПЭИПК, 2011, ISBN 978-5-905042-23-2.

[2] Волков И., Вакуленко А., „Источники электропитания лазеров" Киев, Техника, 1976.

[3] Lee S., „Electrooptic voltage sensor: birefringence effects and compensation methods", App. Optics, vol. 29, no. 30, 1990.

[4] Niewczas P., Dziuda L., Fusiek G., McDomale J., "Design and evaluation of a pre-prototype hybrid fiber-optic voltage sensor for a remotely interrogated condition monitoring system", IEEE Trans Instrum Meas 54 (2005), 1560÷1564.

[5] Touminen M., Olsen R., "Electrical Design Parameters of All-Dielectric-Self-Supporting Fiber Optic Cable", IEEE Transactions on Power Delivery, Vol. 15, No. 3, July 2000.

Analysis and evaluation of the tram vibrations to human body

Emil Iontchev[1], Rosen Miletiev[2], Vladimir Bashev[3] and Ilia Mryankov[4]

Abstract – **The current paper discusses the impact of the rail motor carriage vibrations over the standing human comfort in the tram. The three axis accelerometer is installed on the tram floor at the wheel set center. The acceleration data are frequency weighted to model the human reaction to the different frequencies. The most commonly used weightings are Wk, Wd, Wb. The main characteristic which is used to evaluate the vibration impact is an effective acceleration value. Sometimes the sudden shock and the high amplitude accelerations are also included in the evaluation with the maximum transient vibration value (MTVV) and the vibration dose value (VDV) according to ISO 2631-1:1997 standard and the human health risk is valuated according to Directive 2002/44/EC. Some criteria are defined to interpret the results. According to these criteria the test shows that the tram accelerations are comfortable for the standing human.**

Keywords – **Human body, comfort, accelerations, vibrations**

I. INTRODUCTION

The rail transport hold out the more comfortable travelling towards other transport types. The travel comfort depends from lots of factors such as temperature, noise, seats, etc., but the accelerations have the main impact over the comfort. Their main parameters are recognized as an amplitude and frequency distribution and they depend from the vehicle construction, its technical condition, railway status, driver qualification, etc. The acceleration impact is valuated according to the ISO 2631-1:1997 standard - Mechanical vibration and shock - evaluation of human exposure to whole body vibration [1], national standards such as BS 6841 in UK [2], EU Directive 2002/44/EC about the safe and health conditions [3] and international standard ISO 8041:2005 [4] Human response to vibration — measuring instrumentation. A summary of the standards is made in the book [5]. The specificity of the vibrations in the rail transport is described in the UIC513 standard and the valuation methods are developed as Sperling's ride index.

The current paper describes a method to specify the

[1] Emil Iontchev is with the Higher School of Transport "T. Kableshkov" 158 Geo Milev Street, Sofia 1574, Bulgaria, E-mail: e_iontchev@yahoo.com
[2] Rosen Miletiev is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria. E-mail: miletiev@tu-sofia.bg
[3] Vladimir Bashev is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria. E-mail: v.bashev@gmail.com
[4] Ilia Mryankov is with the Higher School of Transport "T. Kableshkov" 158 Geo Milev Street, Sofia 1574, Bulgaria

vibration level on the tram floor and to evaluate the standing human comfort.

II. SYSTEM DESCRIPTION

The 3D acceleration measurement consists of EK3LV02DL evaluation system from STMicroelectronics [6] and personal computer. The system includes the 3D MEMS linear accelerometer LIS3LV02D. It is able to measure accelerations in three axes in the range ± 2g, ± 6g, with a frequency from 40 to 2560 Hz and the sensitivity is equal to 1mg. The measurement system is installed on the rail motor carriage floor at the wheel set center, which axes are orientated according to the requirements of ISO 2631-1:1997 standard (Figure 1).

Figure 1. Basic axes of the standing human body

III. DATA ANALYSIS BASICS

The measured accelerations are processed according to the algorithm shown at Figure 2. The vibrations with equal amplitudes and different frequencies are perceived in a different way by humans, because the human body is very sensitive to the frequency equal to 5Hz. The equalized sensitivity is obtained by the filtration of the data by specially designed filters [1]. This standard defines the filter curves for different human body positions and different frequencies. The Figure 3 shows three of these curves. The Wk filter is used for filtration of the z data direction and for vertical recumbent direction (except head) while Wd filter is used for the x and y directions and for horizontal recumbent direction.

The main characteristic which is used for the valuation of the human comfort is recognized as an effective value of the

weighted acceleration. Its value is estimated according to the following equation for each axis [1]:

$$a_{wj}rms = \sqrt{\frac{1}{N}\sum_{n=1}^{N}\left(a_{wj}(n)\right)^2}$$

(1)

where N is sample number and j is the acceleration of the corresponding axis.

When some sudden shock appears then the crest factor is used to evaluate the human comfort. Its value is calculated according to the equation:

$$CF = \left|\frac{max(a_w)}{a_{wj}rms}\right|$$

(2)
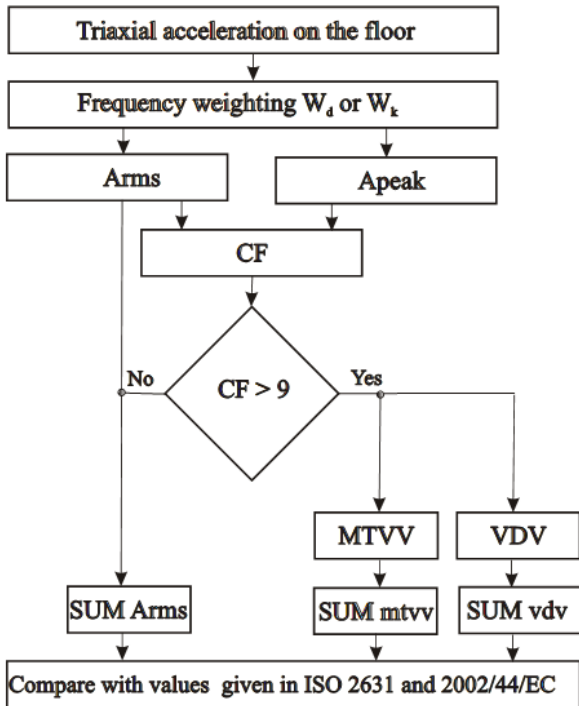
Figure 2. Algorithm of evaluation and assessment of whole-body vibration
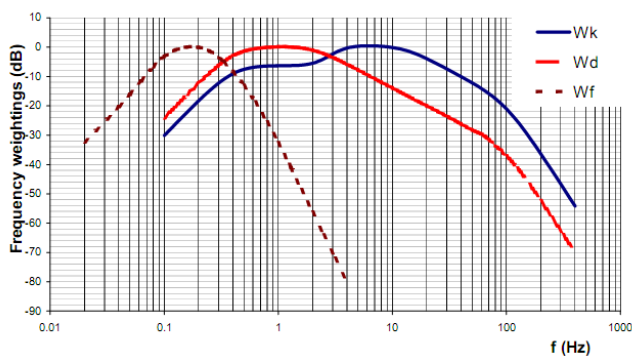
Figure 3. Frequency weighting curves for principal weightings

The crest factor is equal to 1 for the square wave signal, 1.4 for the sinusoidal signal and 1.7 for the Gaussian distributed vibrations. Its value is changed significantly when the sudden shock appears while the effective value remains unchanged. When the crest factor value exceeds 9 it is recommended to use the additional characteristics to evaluate the human

comfort. The first characteristic is recognized as a maximum value of the vibrations. The moment vibrations and shocks are obtained by the data integration for a short time. The standard recommendations at [1] set this time to 1s. The maximum transient vibration value is calculated as:

$$MTVV_j = max\left(\sqrt{\frac{1}{fs}\sum_{n=n_0}^{n_0+fs}\left(a_{wj}(n)\right)^2}\right)$$

(3)

where fs is the sampling frequency, $n_0=0,1,2\ldots.N-fs-1$.

The second characteristic is recognized as a vibration dose value (VDV). This characteristic is more sensitive to the peak values than the MTVV value. The vibration dose value is calculated as:

$$VTV_j = \sqrt[4]{\frac{1}{fs}\sum_{n=0}^{N}\left(a_{wj}(n)\right)^4}$$

(4)

The additional characteristics are important for the final valuation when the following thresholds are exceeded:

$$\frac{MTVV_j}{a_{wj}rms} = 1.5$$

(5)

and

$$\frac{VDV_j}{a_{wj}rmsT^{1/4}} = 1.75$$

(6)

The total effective value is calculated from the effective values of each axis as follows:

$$a_v = \sqrt{\left(k_x^2 a_{wx}rms\right)^2 + \left(k_y^2 a_{wy}rms\right)^2 + \left(k_z^2 a_{wz}rms\right)^2}$$

(7)

where $k_x, \_k_y \_k_z$ are coefficients which values depends from the human position and vibration direction.

When the human body is standing then the coefficients are equal to 1. The weighted vibration magnitude (total of three axes) is compared with the standard values (Table 1) and the final valuation is established.

TABLE 1 [1]

Weighted vibration magnitude (total of three axes) [m/s²]	Likely reaction in public transport
Less than 0.315	Not uncomfortable
0.315 to 0.63	Little uncomfortable
0.5 to 1	Fairly uncomfortable
0.8 to 1.6	Uncomfortable
1.25 to 2.5	Very uncomfortable
Greater than 2	Extremely uncomfortable

The Directive 2002/44/EC defines the top values of these characteristics regarding the exposure of workers to the risks arising from physical agents (vibrations).. These values are shown at Table 2.

TABLE 2

Exposure action value	$a_w$rms [m/s²]	VDV [m/s$^{1.75}$]
A daily exposure action value	0.5	9.1
A daily exposure limit value	1.15	21

The daily vibration dose which is obtained during the 8 hour working day is calculated as follows:

$$A_{wj}(8) = k_j a_{wj} rms \sqrt{\frac{T_{exp}}{8}}$$

(8)

where k_j is equal to 1.4 for x and y axes and 1 for z axis and T_{exp} is the exposing time.

In this case the maximum exposing time may be calculated to reach the maximum values from Table 2. The equation also may be used to calculate the vibration dose as follows:

$$VDV_{exp j} = k_j VDV_j \sqrt[4]{\frac{T_{exp}}{T_{meas}}}$$

(9)

where T_{meas} is a recording time, VDV_j is measured value for the time period equal to T_{meas}.

IV. EXPERIMENTAL RESULTS

The experimental data are measured by two transitions of the tram on the same road. The sampling frequency is equal to 640Hz and the accelerometer dynamic range is set to ±6g. The recorded time is equal to 25 minutes so the data are statistically representative for the railway section. The recorded data are processed according to the algorithm shown at Figure 2 and it is integrated in the MATLAB environment. The weighted value of the Z acceleration, the maximum transient vibration value (MTVV) and the vibration dose value (VDV) are shown at Figure 4. The Figure 5 represents the data during the time interval between 840 and 1200 s.
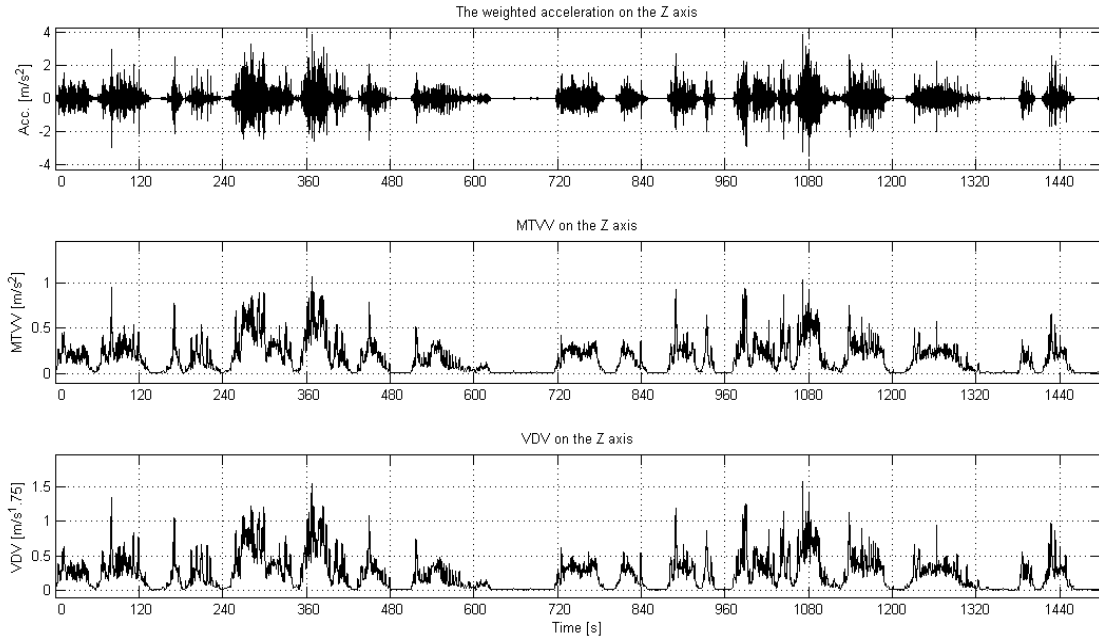
Figure 4. A_{wz}rms, MTVV and VDV
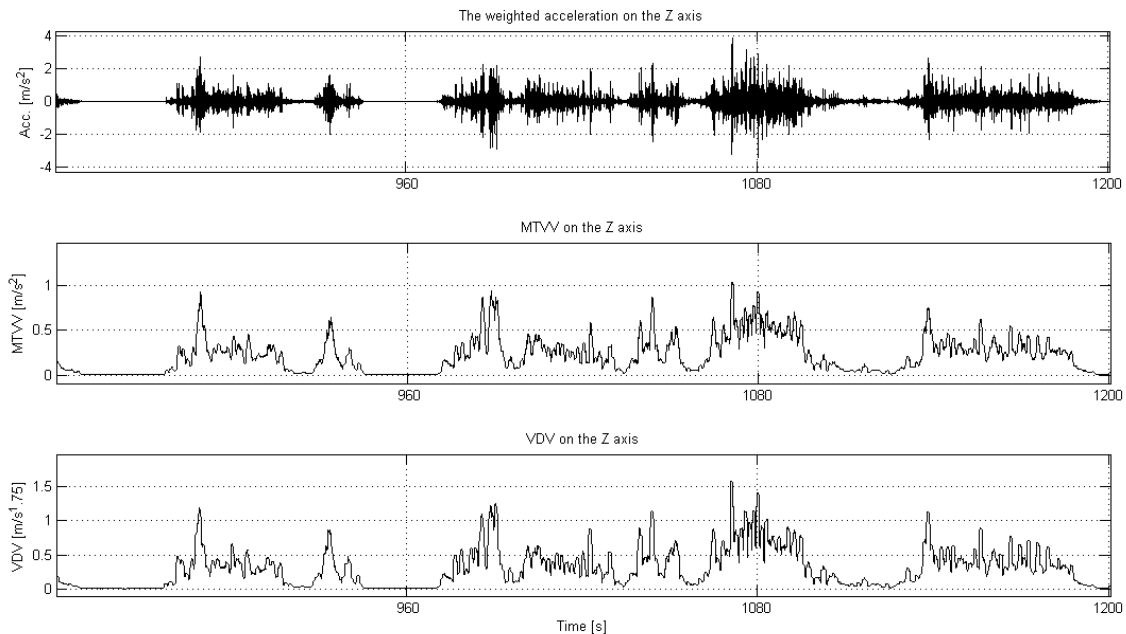
Figure 5. A_{wz}rms, MTVV and VDV during the time interval 840-1200s
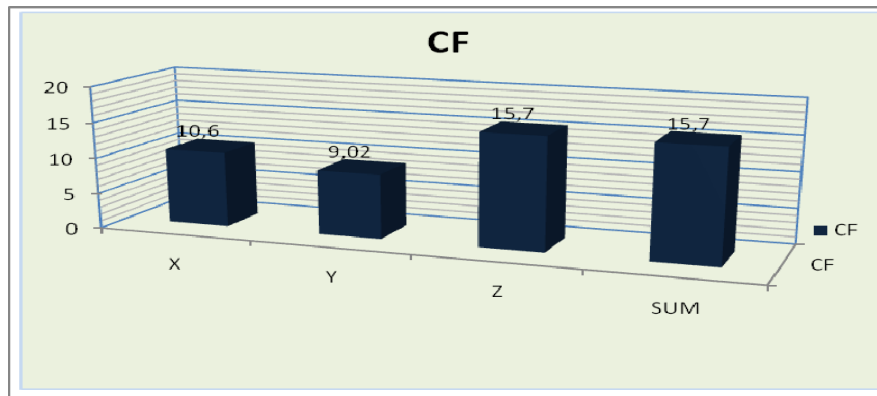
Figure 6. Crest factor
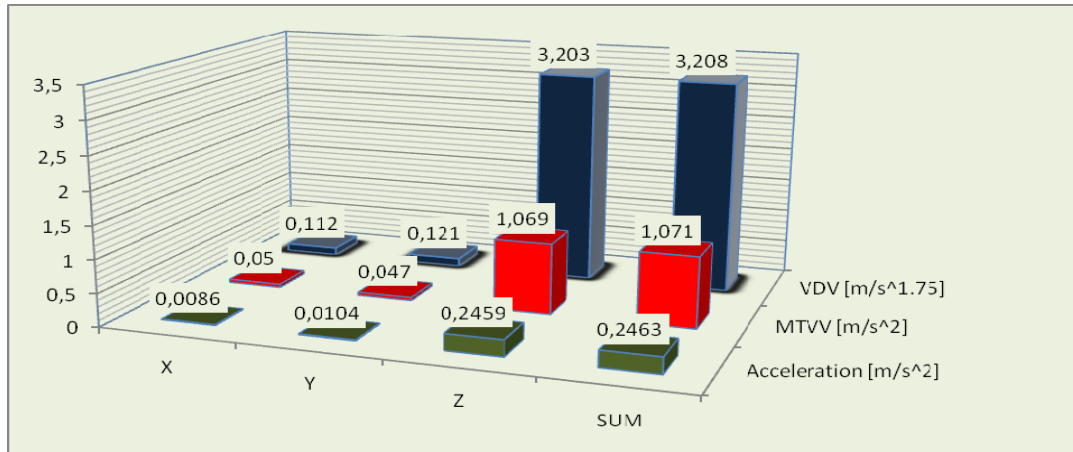
Figure 7. Values of the A_{wj}rms, MTVV and VDV

The calculated values of the crest factor exceed 9 for all three axes (Figure 6). The conditions (5) and (6) are also fulfilled which require the calculation of the additional characteristics. The measurements results for the effective value of the weighted acceleration, the maximum transient vibration value (MTVV) and the vibration dose value (VDV) for each axis are shown at Figure 7.

The calculated values do not exceed the threshold values described at Table 1 and Table 2 which defines the valuation of the human comfort as not uncomfortable and the vibration values as safe for the human health.

V. CONCLUSION

The current paper discusses a method for standing human comfort valuation according to standards ISO 2631-1:1997 and EU Directive 2002/44/EC. Also the algorithm for calculation of the main characteristics is developed and studied which may be also used for valuation of the vibration influence to the sitting human. The obtained results show that the vibrations are not uncomfortable and the vibration values are safe for the human health.

REFERENCES

[1]. ISO 2631-1, 1997. Mechanical vibration and shock Evaluation of human exposure to whole-body vibration. Part 1: General requirements. International Organization for Standardization (ISO).

[2]. BS 6841, 1987. Guide to measurement and evaluation of human exposure to whole-body mechanical vibration and repeated shock. BSI

[3]. EUROPEAN COMMISSION, 2002. Directive 2002/44/EC of the European Parliament and of the Council of 25 June 2002 on the minimum health and safety requirements regarding exposure of workers to the risks arising from physical agents. Official Journal of the European Communities L177

[4]. ISO 8041, 2005. Human response to vibration - Measuring instrumentation. International Organization for Standardization (ISO).

[5]. Mansfield, N.J., Human response to vibration. CRC Press, ISBN 0-203-57102-9, 2005

[6]. www.st.com, UM0285, User guide, EK3LV02DL Evaluation Kit, September, 2006

Process Control System of the Converters Plant in the RTB Bor Company, Serbia

Viša Tasić[1], Marijana Pavlov-Kagadejev[2], Vladimir Despotović[3], Darko Brodić[4],

Marko Anđelić[5], Ivana Stojković[6]

Abstract – **The article describes improvements of the process control system in the converters plant in RTB Bor Company, Serbia. The main task of converting process is to obtain metallic copper as a result of the oxidation reaction of iron, sulphur and other components, which can generate oxygen from the air that is blown into the melt under certain pressure. Monitoring and control of the converting process are provided using the process control system that is a part of complex distributed system for control of the copper smelting and refining process in RTB Bor Company. Some hardware and software solutions developed for this particular control system, as well as configuration and topology of industrial computer network, are emphasized.**

Keywords – **control system, network, communication**

I. INTRODUCTION

The Municipality of Bor is located in the southeastern part of the Republic of Serbia, close to the Bulgarian and Romanian borders. The town of Bor has been the major centre for mining and processing of copper and other precious metals since 1903. Air pollution is perceived as the main environmental problem in Bor. The main source of air pollution with SO_2 gas, heavy metals in particulate matter and aero sediments is the Copper Mining and Smelting Complex Bor (RTB Bor Company) which has been in operation for more than 100 years [1].

[1] Viša Tasić is with the Institute of Mining and Metallurgy, Department of Science, Zeleni bulevar 35, 19210 Bor, Serbia, e-mail: visa.tasic@irmbor.co.rs,

[2] Marijana Pavlov is with the Institute of Mining and Metallurgy, Department of Industrial Informatics, Zeleni bulevar 35, 19210 Bor, Serbia, e-mail: marijana.pavlov@irmbor.co.rs,

[3] Vladimir Despotović is with the University of Belgrade, Technical Faculty in Bor, Vojske Jugoslavije 12, 19210 Bor, Serbia, e-mail: vdespotovic@tf.bor.ac.rs,

[4] Darko Brodić is with the University of Belgrade, Technical Faculty in Bor, Vojske Jugoslavije 12, 19210 Bor, Serbia, e-mail: dbrodic@tf.bor.ac.rs,

[5] Marko Anđelić is with the Institute of Mining and Metallurgy, Department of Science, Zeleni bulevar 35, 19210 Bor, Serbia, e-mail: dmarko.andjelic@hotmail.com,

[6] Ivana Stojković is with the University of Niš, Faculty of Electronic Engineering, Aleksandra Medvedeva 14, 18000 Niš, Serbia, e-mail: ivana.stojkovic@elfak.ni.ac.rs

A typical pyrometallurgical copper smelting process includes four steps: roasting, smelting, concentrating, and fire refining. Copper ore concentrate is roasted to reduce impurities, including sulfur, antimony, arsenic, and lead. Smelting of roasted ore concentrate produces matte, a molten mixture of copper sulfide (Cu_2S), iron sulfide (FeS) and some heavy metals. By converting of matte, the high-grade 'blister' copper of 98.5 to 99.5 % is recovered. Typically, blister copper is then fire-refined in an anode furnace, cast into 'anodes', and sent to the electrolytic refinery for further impurity elimination.

The main activity of the converters plant is to obtain metallic copper as a result of the oxidation reaction of iron, sulfur and other components, which can generate oxygen from the air that is blown into the melt under certain pressure. In the standard Pierce-Smith converter (as in the copper smelter in Bor), the off-gases are treated in electrostatic precipitator system to remove particulate matter, and in the sulfuric acid plant to remove SO_2. Thus, reliable and timely information about process parameters are of great importance for the process itself, as well as for the preservation of the air quality in surroundings of the copper smelter.

Department of Industrial Informatics, at the Mining and Metallurgy Bor, has a long tradition in designing of the real-time systems for monitoring and control of the industrial processes. Three generations of UMS (Universal Measuring Station) have been developed since 1990. UMS is industrial PLC (Programmable Logical Controller) and it is a core of process control system. Main objectives of such systems are real-time data processing, automatic process control, real-time presentation and visualization of results, and forming the databases [2].

The existing process control system in the converters plant includes measurements of the key process parameters (temperature, pressure, speed, level, etc.). Also, the constant and reliable communications between the production plants in the copper smelter are of the critical importance. Because of the fact that the copper smelter production plants are located in the several halls, at the distance of several hundreds of meters, the industrial computer network consists of several dislocated segments. In order to include all the required locations into the industrial computer network and to transfer all actual information to the specific locations, building of the network nodes and segments were carefully planned and realized.

In order to improve the existing control system in the converters plant, a new industrial PLC has been installed. Manual data collection using instrumentation on command tables and panels is replaced by microcontroller based real-time control system. Appropriate software application has

been developed as well. The main objectives of a new control system are real-time data processing, data presentation (in the form of dynamic synoptic schemes, real-time graphs and tables) and database management.

II. CONTROL SYSTEM HARDWARE

Programmable logic controllers (PLC) are controlled by programmed code created on a PC and downloaded to the memory of the controller. Depending on the state of the inputs and results of programmed functions, the controller will activate proper outputs. Flexible solutions for data acquisition and control are also an important tool for measurements in scientific and research laboratories. PLCs are modular devices and this modularity is often used to provide additional functions, e.g. increase the number of available inputs and outputs. The UMS standard configuration includes CPU module (one-board computer based on MC68HC11E microcontroller with internal eight channels, 8-bit A/D converter), analog input modules (up to 3 modules, each with 20 channels), digital input and output modules (64 + 64 channels), RS232 communication port, LC display and keyboard. Block diagram of the UMS is shown in Fig. 1.
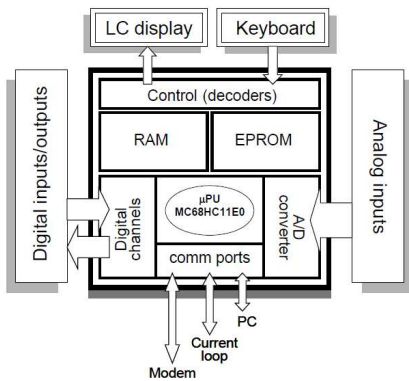
Fig. 1. Functional block diagram of UMS
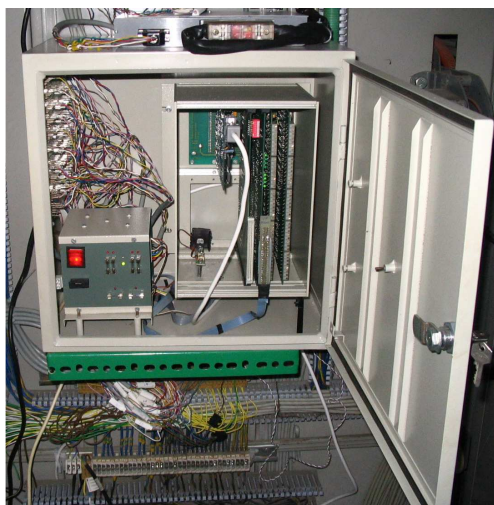
Fig. 2. UMS installed in the converters plant

Fig. 3. PC workstation in the converters plant control room

Improvement of the control system demanded the installation of new sensors, transmitters and controllers. All output signals from the different types of transmitters (temperature, pressure, flow, speed, electric power) were connected to the UMS inputs.

Since the most of sensing elements were dislocated throughout the facility hall, the output signals from sensors and transmitters had to be concentrated at one place. The control room of the converters plant was chosen as the appropriate location for this purpose (shown in Fig. 2).

Process parameters are imported to UMS as standard current (4-20 mA) or voltage signals (0-5 V DC). UMS performs measurements and upon request, transmits results to PC workstation (shown in Fig. 3). UMS together with the associated PC workstation (server) operate as a node of the industrial computer network. Industrial computer network consists of several dislocated segments. In order to include all the required nodes into industrial computer network design, all network nodes and segments were carefully planned and realized.

III. CONTROL SYSTEM SOFTWARE

UMS can operate independently and control the process parameters itself (local control mode). Also, it can operate as a data logger, and store more than 3000 data messages in RAM. The EPROM of the UMS holds residential software. It consists of executable versions of test, control, operational and communication software modules. The source code is written in the symbolic Motorola language. For that purpose a special environment of ∏-assembler [3] was used.

Supervising Control and Data Acquisition (SCADA) real-time application, named Process Control Program (PCP), is developed in order to support communications with the UMS and other real-time control system operations [4]. PCP is based on a client/server architecture running on both master and remote workstations. In that way, PCP enables integration of the network nodes in a complex distributed control system network.

PCP is developed using Microsoft Visual C++ [5]. It has a complex structure and consists of several modules. Main program modules are designed for communication with UMS,

data acquisition, real-time data processing and presentation, interaction with process according to the appropriate algorithm, reporting, data archiving, off-line data processing and database management.

All the required information, such as: measuring ranges, operating ranges, operation curves, working regimes, etc., can be modified in PCP. Interactions with the process are performed according to the control algorithms considering the actual values of measured parameters.

The PCP presents data using dynamic screens, graphs or tables, as shown in Fig. 4 and Fig. 5. The results of the measurements are stored in a database. Thus, information about the process parameters can be retrieved and displayed in the same manner as in the real-time. All data can be easily exported into the applications suitable for reports preparation and further analyses (e.g. Microsoft Excel).
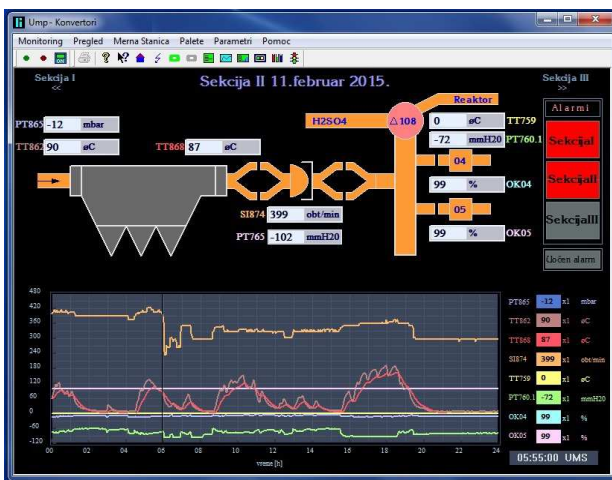
Fig. 4. Synoptic representation of the process parameters

Fig. 5. Graphical representation of the measurements results

IV. DISTRIBUTED CONTROL SYSTEM

The PC workstation is connected to UMS which executes the PCP. It acts as a data server in the industrial network

configuration. The other workstations execute passive version of PCP, called Remote Control Program (RCP) that is used for the remote monitoring only. RCP can access databases and retrieve the data according to the assigned privileges. PCP generates very short text files (ASCII), with the size of only few KB, which contains the actual information about monitored process parameters.

The workstations that execute RCP simply download these files over LAN. From a technical point of view, this is a low-cost solution having in mind that distant LAN segments are mostly connected via conventional telephone copper wires (RTB Company owns private telephone lines between the most of its facilities). Each production plant has a server PC workstation, which is interconnected with a number of client workstations. The realized industrial network consists of several sub-networks with a number of clients for remote monitoring as shown in Fig. 6.

Data analysis, processing, and presentation are performed locally, on client workstations [6, 7]. Client workstation is able to run as many RCP programs as needed at the same time (by accessing servers in the different production plants).
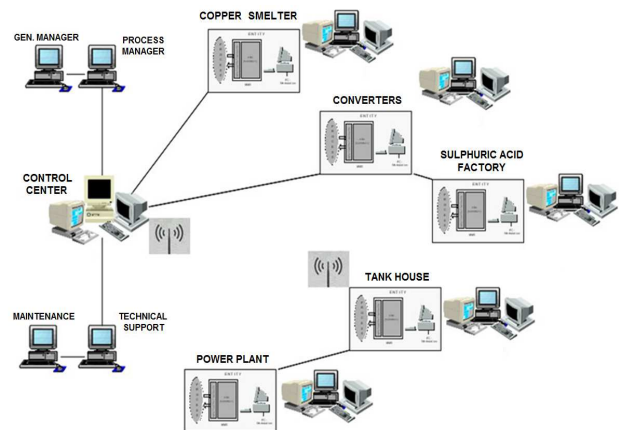
Fig. 6. Block diagram of the industrial computer network in the copper smelter in Bor

CONCLUSION

The realized control system showed good stability and resistance to external influences. The presented system can be used for local control, or as integral parts of the distributed control system. The main advantage of the given solutions is their modularity, ease of use, low price, and its applicability in the different production processes. The system response time is usually less than few seconds. This can be considered as satisfactory response time for this kind of the production processes. The possibility of distant monitoring is very important for making the different business and production strategy decisions on time. Implemented network enabling managers to monitor the production process in real-time. Appropriate data analyzes and creation of reports can be also performed on the client side at any time.

ACKNOWLEDGEMENT

REFERENCES

[1] V.Tasić, N. Milošević, R. Kovačević, N.Petrović, "The analysis of air pollution caused by particle matter emission from the copper smelter complex Bor (Serbia)," Chemical Industry & Chemical Engineering Quarterly, vol. 16, no. 3, pp. 219-228, 2010.

[2] D.R.Milivojevic, V.Tasic, "MMS in Real Industrial Network," Information Technology and control, vol. 36, no. 3, pp. 318-322, 2007.

[3] Z. Radonjic, ∏-assembler user guide, Niš, Serbia, 1996. (in Serbian)

[4] D.R.Milivojevic, V. Despotovic, V.Tasic, M. Pavlov, "Process Control Program as an Element of Distributed Control System," Information Technology and control, vol. 39, no. 2, pp. 152-158, 2010.

[5] B.Stroustrup, *The C++ Programming Language*, Addison-Wesley Pub Co, 3rd edition, 2000.

[6] B. Bamieh and P. G. Voulgaris, "Optimal Distributed Control with Distributed Delayed Measurement," In Proceedings of the IFAC World Congress, 2002.

[7] D.Milivojević, V.Tasić, M.Pavlov and V.Despotović, "Synthesis of DCS in Copper Metallurgy," ICEST 2007, Conference Proceedings, Book 2, pp.629-631, Ohrid, FYR Macedonia, 2007.

Time and frequency analysis of bogie-railway dynamics

Rosen Miletiev[1], Rumen Yordanov[2], Vladimir Bashev[3] and Emil Iontchev[4]

Abstract – The paper discusses the application of the electronic measurement system based on MEMS inertial sensors and GPS receiver which is capable to measure and record the dynamic parameters of the interaction of the bogie-railway system. The inertial data are combined with the GPS receiver and SD card to record all data which may be processed later to find the places with highest linear and angular accelerations.

Keywords – **MEMS, inertial sensor, bogie-railway dynamic**

I. INTRODUCTION

The safety and comfort motion of the urban trams is defined by the condition of the tramway and the dynamic interaction between the rails and vehicle. The tram operations influence over the rail parameters, so the rail deviation is increased from their initial position. The rail faults which depend from the geometry may be defined as periodic, non-periodic or custom ones. The rail faults also conduct to the vibration generation jointly with the tram wheels. The measurement of the railway current condition may be accomplished with specialized motor cars [1] and the railway geometry may be analyzed according to the measurements. This procedure requires expensive motor cars and often the railway condition is evaluated according to the manual railway measurements. The correlation between the obtained valuation and the generated forces is inadequate, but the passenger comfort depends from the dynamic forces. This is the reason to develop and evaluate the measurement systems, which are capable to measure the dynamic reactions between the railway and the tram. The main parameter of the dynamic reaction is the movement noise. The measurement of the generated noise is based on the microphone dynamics and this method is often used to determinate the wave wear of the rails [2]. The European standard EN3095 [3] defines the permissible levels of the generated noise from the railway roughness.

The measurement of the dynamic forces is accomplished by the strain gauges, which are stick on the wheel [4] or translation sensors and tensometers on the rails [5]. The force

[1] Rosen Miletiev is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria. E-mail: miletiev@tu-sofia.bg

[2] Rumen Yordanov is with the Faculty of Electronic Engineering and Technology at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria. E-mail: rsyordanov@yahoo.com

[3] Vladimir Bashev is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria. E-mail: v.bashev@gmail.com

[4] Emil Iontchev is with the Higher School of Transport "T. Kableshkov" 158 Geo Milev Street, Sofia 1574, Bulgaria, E-mail: e_iontchev@yahoo.com
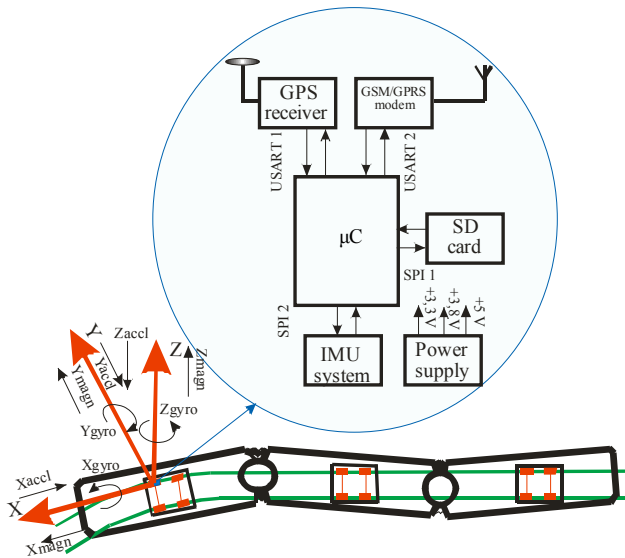
amplitude calculation is described in the European standard EN14363 [6]. Their amplitude is used as a criterion for valuation of the railway safety and its loading capacity.

Another main parameter is defined as vertical and horizontal accelerations and tram angular rate. The acceleration amplitudes are used as an index to control the rail failures. The acceleration measurements are accomplished by linear or angular rate accelerometers, which are installed on the tram or on the rails [7]. This method gives the valuation of the combined roughness of the rails and the wheels. In these systems [8] the roll value is also measured to obtain the accelerations which are caused by the varied rail level.

The paper discusses the time and the frequency analysis of the linear and angular accelerations of the bogie-railway dynamics.

II. SYSTEM DESCRIPTION

The MEMS based inertial system is built to measure the linear accelerations and to calculate the dynamic characteristics especially the rotation angle of the bogie. The system is recognized as 9DoF (Degrees of Freedom) and it is based on 3D linear accelerometer, 3D digital gyroscope and 3D digital magnetometer. The localization of the vehicle is obtained by build-in GPS receiver with a refresh rate up to 10Hz. The refresh rate of the navigation data is set to 4Hz.

The data acquisition system is based on the MEMS three axes digital output linear accelerometer and magnetometer LSM303DLHC and 3D ultra-stable MEMS angular rate accelerometer L3G4200D, both produced by *ST*. These sensors include a sensingelement and an IC interface able to take the information from the sensing element and to provide the measured acceleration signals to the external world through an I^2C serial interface.

The system reads the inertial and magnetic data 40 times per second and stores the navigation and inertial data in the internal FLASH memory (SD card) with a capacity up to 4GB. The navigation data are transmitted also to the remote server via GPRS connection. The block diagram of the system and the installation place are shown at Figure 1.

The measurement system is installed on the terminal clam and is firmly fixed to ensure proper transfer function between the rails and the system. The integrated GSM module transmits the GPS data to the server to realize real-time monitoring of the tram position. The acceleration data are not transmitted via GSM network due to their high volume but they could be analyzed directly in the device to detect the critical points. The detection of the critical point may be transferred to the server to point the railway fault positions.

Fig. 1. The measurement system and its orientation relative to the vehicle axes

Fig.2. Geographic coordinates of the experimental track

III. DATA ANALYSIS BASICS

The navigation and the inertial data are recorded in a binary format on SD card and analyzed later by MATLAB routine. The spectrum analysis is realized on the basis of a Short Fast Fourier Transform. STFTs as well as standard Fourier transforms are frequently used to locate the frequencies of specific vibrations (especially when used with greater frequency resolution) or to find frequencies which may be more or less resonant in the space where the signal was recorded.

There are two main parameters which have to be determined to obtain an optimal resolution. The first parameter is directed to the analyzed window size N and the second one – the number of the overlapped samples. Because the data are noisy, we have to choose between the better noise suppression windows and the higher spectral resolution windows. So we have to analyze the different windows and find the best one for the application. The proper selection of the ST-FFT window length and overlap processing produces excellent and accurate time-frequency resolution to show the frequency at any given time of the original waveform.

IV. EXPERIMENTAL RESULTS

The experimental data are collected using the described measurement system, which is installed on the 6-axes tram motor car T6M 400. Its suspension consists from three carts: power ones (I and III-rd) T65 and supporting one (II-nd cart) type T_{sp}65. The carts consist of two-stage spring suspension (cylindrical springs) and the H-shaped open type cart frame.

The time analysis of the experimental data is shown at Figure 2 (Geographic coordinates of the experimental track), Figure 3 (3D Time analysis of Z linear acceleration) and Figure 4(3D Time analysis of Z angular acceleration). Their amplitude is used as a criterion for valuation of the railway safety.
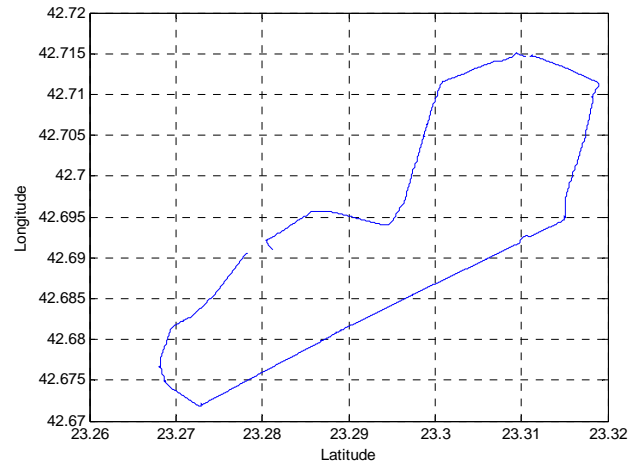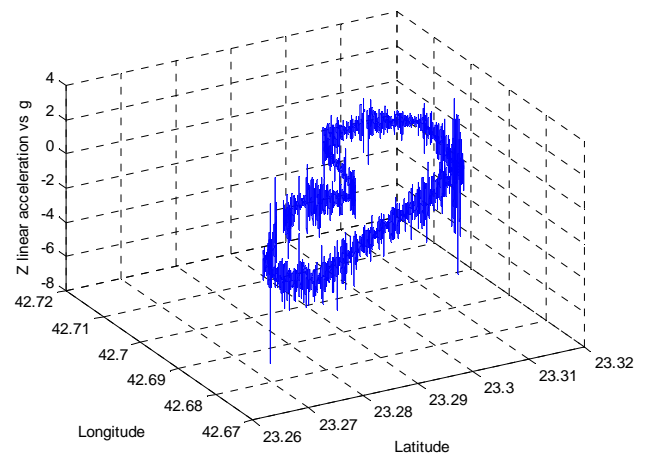
Fig.3. 3D Time analysis of Z linear acceleration
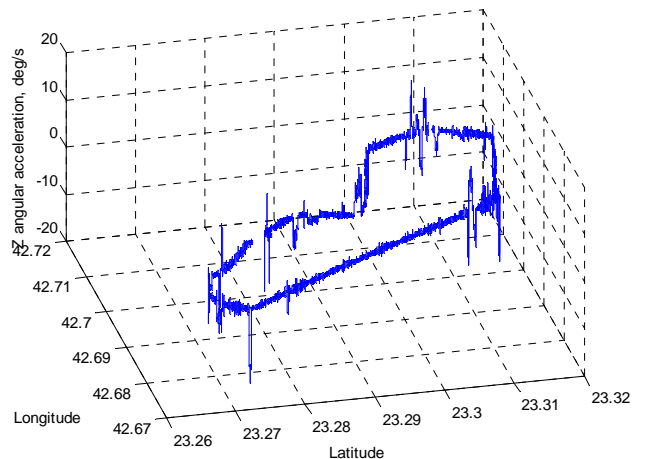
Fig.4. 3D Time analysis of Z angular acceleration

The frequency analysis is realized on the basis of a Short Time Fast Fourier Transform. The window has a rectangular shape with different size (from N=64 to N=256 samples) and 25% overlapped samples. The frequency bits may be calculated according to the equation f_i=i/N*Fs, where Fs = 40Hz – sampling frequency.

The results of the frequency analysis of the Z linear acceleration using 256, 512 and 1024 point window are shown at Figure 5.

signal to achieve the maximum value of the spectrum peak. The Figure 6 represents the zoomed picture of the spectrum peaks at the selected region of the time-frequency distribution.
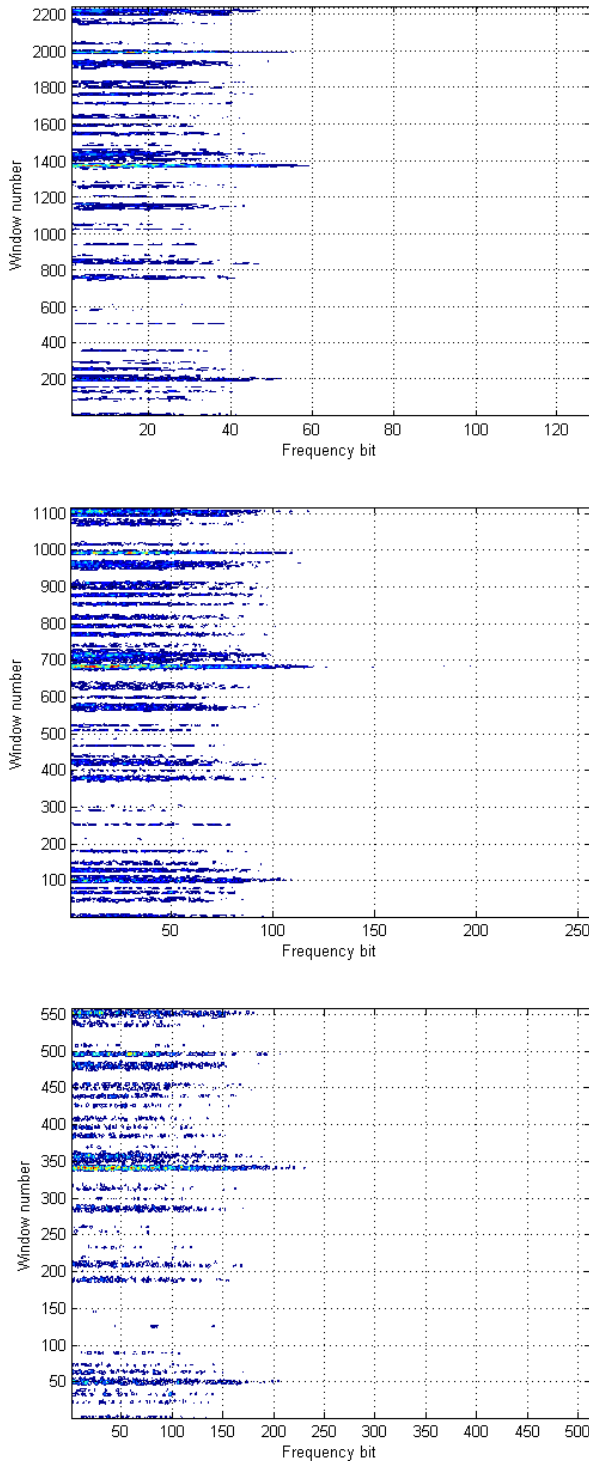
Fig.5. ST-FFT analysis of the Z linear acceleration using 256, 512 and 1024 point window respectively

Fig.6. Zoomed picture of the spectrum peaks using 256, 512 and 1024 point window respectively

The results show that the maximum magnitude of the frequency peaks appears at the window number equal to approximately to 1370 when the window size is equal to 256 points. The right choice of the window size is determined from the circumstance to include the maximum points of the

It is clearly visible that the choice of the window size is essential to obtain the maximum value of the spectrum peaks. When the window size is equal to 512 point the maximum sampling points of the vibrations are included in the analyzed

window. As the window size is higher than this optimum value the analyzed window includes not only all signal samples but also a lots of noise samples which blurred the spectrum peaks. If the window size is lower than the optimum value then the window includes a part of the signal samples and the spectrum peaks are not clearly defined.

The same situation appears when the Z angular acceleration is analyzed, but due to the fact that the spectrum peaks are situated at very low frequencies it is very important to obtain excellent frequency resolution by increasing of the window size. The results are shown at Figure 7.

V. CONCLUSION

The current paper represents the inertial measurement system based on MEMS inertial sensors which is capable to measure and analyze the linear and angular accelerations while their amplitude is used as a criterion for valuation of the railway safety. It is shown that the proper selection of the ST-FFT window length and overlap processing may produce excellent and accurate time-frequency resolution to show the frequency at any given time of the original waveform.

REFERENCES

[1]. E. Sackl, The EM-SAT 120 Track Survey Car, an integrated part of the track geometry data base of the Austrian Federal Railways ÖBB, Railway Technical Review, 2/2004, pp.39-43

[2]. Y. Hayashi, T. Kojima, H. Tsunashima and Y. Marumo , Real time fault detection of rail way vehicles and tracks, Railway Condition Monitoring 2006, pp. 20-25

[3]. European Standard EN ISO 3095:2005: Railway Applications—Acoustics—Measurement of Noise Emitted by Railbound Vehicles, August, 2005.

[4]. Per Gullers, Lars Andersson, Roger Lunden, High-frequency vertical wheel–rail contact forces—Field measurements and influence of track irregularities, Elsevier Wear 265, 2008, 1472–1478

[5]. Yoshihiro Suda, Hisanao Komine, Kosuke Matsumoto, Yasunobu Endo, Takuji Nakai, Masuhisa Tanimoto, Feedback friction control between wheel and rail by detecting yaw moment of wheelset, Elsevier, Wear 265 (2008) 1512–1517

[6]. EN 14363, Railway applications – Testing for the acceptance of running characteristics of railway vehicles – Testing of running behaviour and stationary tests, 2005.

[7]. E. Verheijen, A survey on roughness measurements, Journal of Sound and Vibration 293 (2006), pp 784–794

[8]. H. Tsunashima, T. Kojima, Y. Marumo, A. Matsumoto and T. Mizuma, Condition monitoring of railway track and driver using in-service vehicle
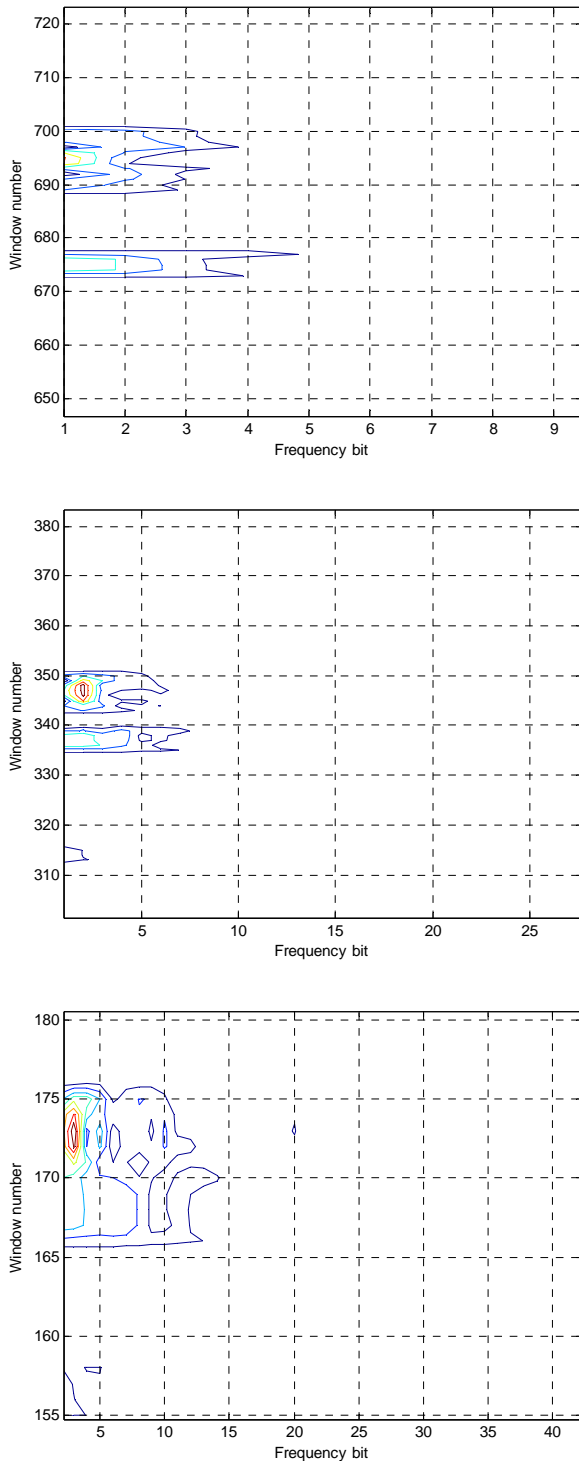
Fig.7. Zoomed picture of the Z angular acceleration spectrum peaks using 512, 1024 and 2048 point window respectively

Development of Data Acquisition Module Based on PIC18F4550 Microcontroller

Marko Anđelić[1], Viša Tasić[2], Marijana Pavlov-Kagadejev[3], Darko Brodić[4], Ivana Stojković[5]

Abstract – **This paper describes the development of low-cost data acquisition module which should be implemented in the existing process control systems. The main goal of such implementation is to improve the A/D conversion resolution of the realized and applied control systems. The new data acquisition module is based on PIC18F4550 microcontroller. Some specific hardware and software details of the realized module are specially emphasised.**

Keywords – **data acquisition module, A/D conversion, microcontroller**

I. INTRODUCTION

Department of Industrial Informatics, at the Mining and Metallurgy Institute Bor, has been designing real-time systems for monitoring and control of industrial processes since 1990. Three generations of Microprocessor Measuring Station (MMS) have been developed. MMS was used as a programmable logic controller (PLC) in the control systems mainly applied in the metallurgy. It is based on Motorola MC68HC11 microcontroller with integrated 8-bit A/D converter, 64 analog inputs, 64+64 digital state signals (input +output), RS232 communication port, 48 kB RAM and16 kB EPROM [1].

In order to improve the characteristics of A/D conversion, a new data acquisition module is developed. It is based on PIC18F4550 microcontroller with 10 bit A/D converter (13 analog input channels), integrated USB module, 32 kB of flash (program) memory, 2 kB RAM and 256 bytes of EEPROM memory [2].

Block diagram of PIC18xx microcontroller family

[1] Marko Anđelić is with the Mining and Metallurgy Institute Bor, Department of Science, Zeleni bulevar 35, 19210 Bor, Serbia, e-mail: marko.andjelic@hotmail.com,

[2] Viša Tasić is with the Mining and Metallurgy Institute Bor, Department of Science, Zeleni bulevar 35, 19210 Bor, Serbia, e-mail: visa.tasic@irmbor.co.rs,

[3] Marijana Pavlov-Kagadejev is with the Mining and Metallurgy Institute Bor, Department of Industrial Informatics, Zeleni bulevar 35, 19210 Bor, Serbia, e-mail: marijana.pavlov@irmbor.co.rs,

[4] Darko Brodić is with the University of Belgrade, Technical Faculty in Bor, Vojske Jugoslavije 12, 19210 Bor, Serbia, e-mail: dbrodic@tf.bor.ac.rs,

[5] Ivana Stojković is with the University of Niš, Faculty of Electronic Engineering, Aleksandra Medvedeva 14, 18000 Niš, Serbia, e-mail: ivana.stojkovic@elfak.ni.ac.rs

architecture is shown in Fig. 1. Layout of the realized data acquisition module is shown in Fig. 2.
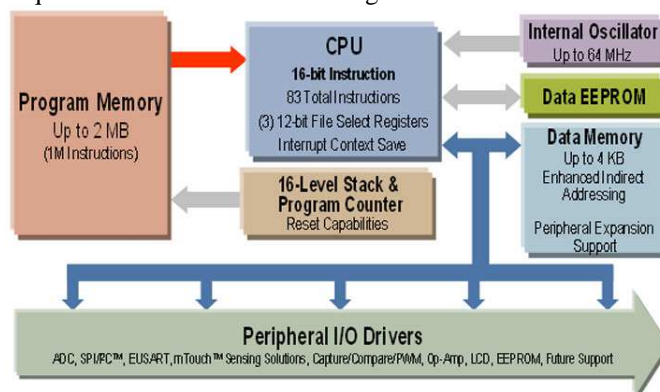
Fig. 1. PIC18xx Architecture Block Diagram [3]

During the realization of this project, some firmware and hardware tools for the development of PIC18F4550 microcontroller based systems were used:

- PICkit 2 In-Circuit Debugger/Programmer [4]
- PICBASIC PRO™ Compiler 3.0 [5]
- MPASM Microchip assembler [6]

For testing purposes, a special application for the PC workstation was developed by using Borland Delphi and mcHID.dll library [7].

In the paper, a brief overview of PIC18F4550 microcontroller A/D conversion process is presented. After that, an example of A/D conversion programming is explained in details.

Fig. 2. Layout of the realized data acquisition module based on PIC18F4550 microcontroller

II. A/D CONVERSION PROCESS IN REALIZED MODULE

The A/D converter module integrated in PIC18F4550 has 13 inputs and allows conversion of an analog input signal to corresponding 10-bit digital number as shown in Fig. 3.
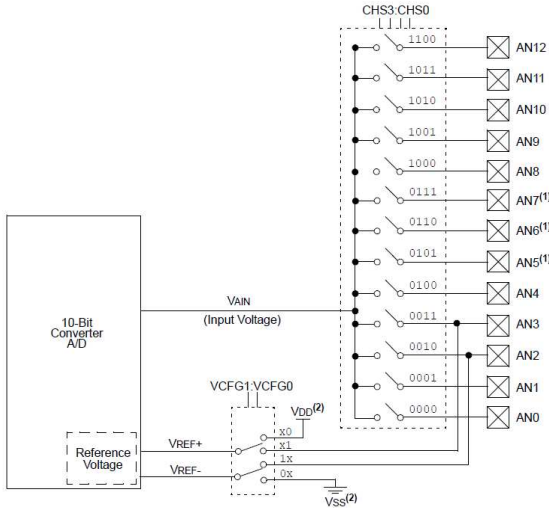
Fig. 3. A/D Block diagram [2]

The A/D conversion (ADC) module has five registers (2 data registers and 3 control registers) [2]:

- A/D Result High Register (ADRESH)
- A/D Result Low Register (ADRESL)
- A/D Control Register 0 (ADCON0)
- A/D Control Register 1 (ADCON1)
- A/D Control Register 2 (ADCON2)

ADRESH and ADRESL registers serve as data registers for storing the A/D conversion results. ADCON0 control register configures ADC operations (start of A/D conversion and input channel selecting). ADCON1 register configures voltage reference and whether the input pins are analog or digital. ADCON2 register configures A/D acquisition time, ADC clock and A/D result format (justification).

In the realized data acquisition module, reference voltage for A/D conversion is set to device's positive and negative supply voltage (V_{DD} = 5 V and V_{SS} = 0 V). The output of the sample and hold is the input into the converter, which generates the result via successive approximation. When the A/D conversion is complete, the result is loaded into the ADRESH:ADRESL register pair, the GO/DONE bit (ADCON0 register) is cleared and A/D Interrupt Flag bit, ADIF, is set.

After the configuration of A/D module, and elapsing of the acquisition time, the A/D conversion of the selected channel can be started. The next steps should be followed to perform an A/D conversion:

- Configure the A/D module,
- Configure A/D interrupt,
- Wait the required acquisition time,
- Start conversion (ADCON0 register),
- Wait for A/D conversion to complete, by waiting for the A/D interrupt,
- Read A/D result registers (ADRESH:ADRESL),

When the analog input channel is selected, it must be sampled for at least the minimum acquisition time before starting a conversion. Calculation of the minimum required acquisition time T_{ACQ} is done with following equation:

$$T_{ACQ} = T_{AMP} + T_C + T_{COFF} \qquad (1)$$

Where T_{AMP} is amplifier settling time, T_C is holding capacitor charging time and T_{COFF} is temperature coefficient. Refering to datasheet specification of the microcontroller, calculation (1) is based on folowing system preferences: C_{HOLD} = 25 pF, Rs = 2.5 kΩ, Conversion Error ≤ 1/2 LSb, V_{DD} = 5V → R_{SS} = 2 kΩ, Temperature = 85°C (system max.), and is given by:

$$T_{ACQ} = 0.2\ \mu s + 1.05\ \mu s + 1.2\ \mu s = 2.45\ \mu s \qquad (2)$$

The A/D conversion time per bit is defined as T_{AD} and per 10-bit conversion 11 T_{AD} is required. The acquisition time parameter is the amount of T_{AD} cycles to delay before the actual A/D conversion starts. The source of the A/D conversion clock is software selectable by bits ADCS2:ADCS0 in the ADCON2, in terms of the microcontroller clock (T_{OSC}). T_{AD} have to be selected as short as possible, but greater than T_{ADmin} (0.8 μs for PIC18F4550 [2]).

Since developed module has main crystal of 20 MHz, T_{OSC} is 0.05 μs. This means that lowest selectable value of T_{AD} is 16 T_{OSC} = 0.8μs, and the ADCS2:ADCS0 bits should be set to 101 B (binary). It is possible to either keep track of T_{ACQ} delay in the program code or to simply set the A/D acquisition time select bits to an even multiple of T_{AD}.

Based on the previous calculations, the minimum A/D processing time per channel in our case is given by:

$$T_{ACQ} + 11\ T_{AD} = 11.25\ \mu s \qquad (3)$$
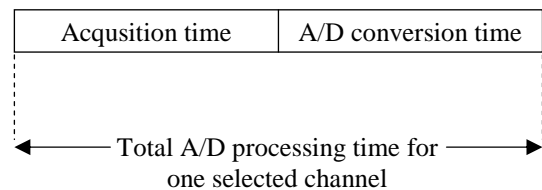
Fig. 4. A/D conversion processing time

III. THE A/D CONVERSION MODULE PROGRAMMING EXAMPLE

PIC assembler program is written using the PICBASIC PRO™ Compiler 3.0 and MPASM universal assembler.
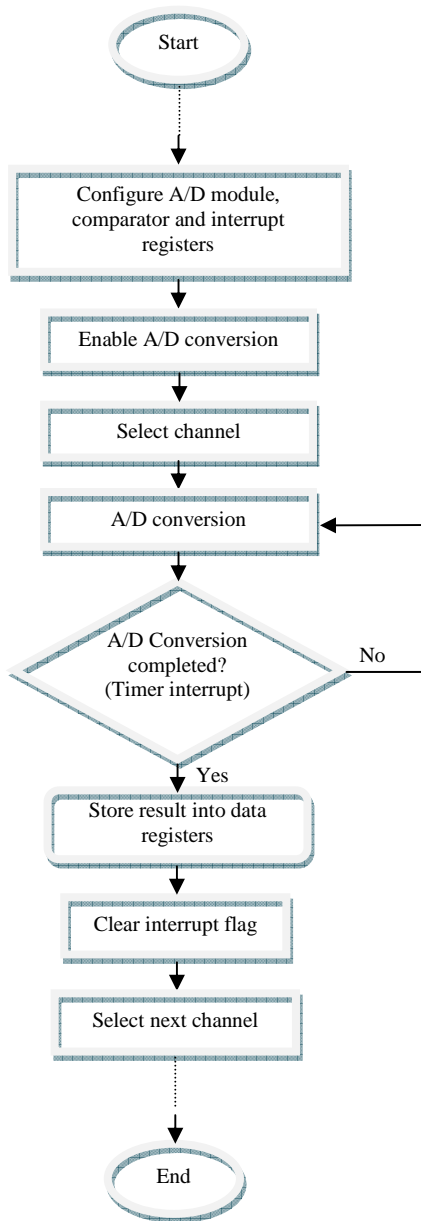
Fig. 5. Data acquisition flowchart (one channel)

Generally speaking two basic principles are used: polling and interrupt mechanism. Polling mechanism is used for sampling, and the handling of different tasks execution is achieved using an interrupt mechanism. Data acquisition flowchart is shown in Fig. 5. Assembler program for ADC includes all steps as described in the previous section that is required for the ADC operation. It was downloaded into the microcontroller using the PICkit2 In-Circuit Debugger/Programmer.

Assembler code sequence, which is used to configure and enable A/D module, comparator, data direction and interrupt registers is shown in Fig. 6. Also, code sequence for the selection of input channel, A/D conversion loop and storing the ADC results is shown in Fig. 7.

Timer predefined for ADC has to be reset to automatically repeat the A/D acquisition period with the minimal software overhead. The ADC module has the ability to generate an interrupt upon completion of an A/D conversion.

Fig. 6. ADC module configuration sequence

Fig. 7. A/D input channel selection and conversion

IV. TESTING THE FUNCTIONALITY OF THE DATA ACQUISITION MODULE

Once the program code has been built and checked from the syntax point of view, it needs to be tested. In order to test the code, appropriate software for supporting the data acquisition module operation from the PC workstation is developed. It is developed using Borland Delphi programming environment (application window is shown in Fig. 8.). API (Application Program Interface) calls are transferred to USB port and to connected module with built-in system mcHID.dll library that is dynamically linked with the application. Since the used A/D converter has 10 bit resolution each analog input can be represented as a number ranged from 0 to 1023.
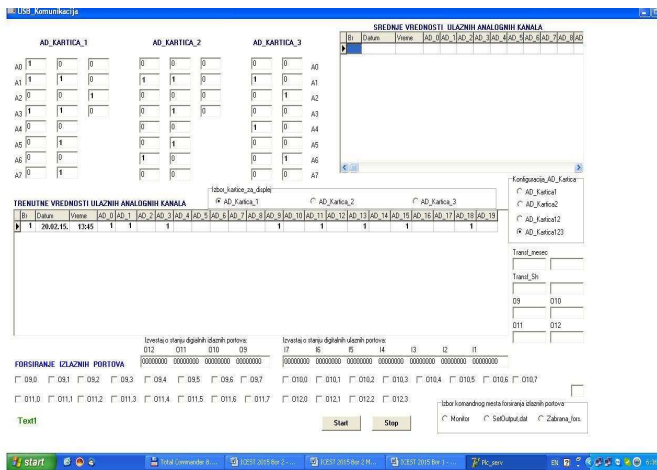
Fig. 8. Test application screenshot

Testing has revealed that for the input signal in the range of 0-3.75 V A/D conversion error of the new acquisition module, measured under laboratory conditions, is not greater than 1 LSB (0.0036 V). A/D conversion error of the old ADC module [8], for the same input range, is also 1 LSB (0.0146 V). The old ADC module has the 8-bit A/D converter so the accuracy of the new ADC module is actually four times higher.

V. CONCLUSION

From the need for a more precise A/D conversion, a new data acquisition module was realized. The appropriate programs for testing the operation of data acquisition module were downloaded into the microcontroller and installed on a personal computer. By testing the module in laboratory it was obtained that the A/D conversion accuracy is improved (four times better accuracy in comparation with the old ADC module).

Also, the data acquisition module works reliable, has USB interface, it is low-cost and enables further development. In the coming period it is necessary to develop adequate software on the PC workstation which will suit industrial requirements, or simply allow the monitoring of specific industrial processes. Minor corrections in the PIC program are also planned to allow some additional functionalities.

ACKNOWLEDGEMENT

REFERENCES

[1] D.R.Milivojevic, V.Tasic, "MMS in Real Industrial Network," Information Technology and control, vol. 36, no. 3, pp. 318-322, 2007.

[2] Microchip Technology Inc., "PIC18F2455/2550/4455/4550 Data Sheet", 2009.

[3] http://www.microchip.com/pagehandler/en-us/family/8bit/architecture/pic18.html (accessed 10 February 2015)

[4] Microchip Technology Inc., PICkit™ 2 Programmer/Debugger User's Guide, 2008.

[5] MicroEngineering Labs, Inc., Pic Basic Pro™ Compiler Reference Manual, 2011

[6] http://ww1.microchip.com/downloads/en/DeviceDoc/33014L.pdf (accessed 10 February 2015)

[7] http://delphiandpic.jimdo.com/delphi-usb-y-herramientas/usb-hid/la-mchid-dll/ (accessed 10 February 2015)

[8] V.Tasić, D.Milivojević, M.Pavlov, V.Despotović, D. Brodić, "Microcontroller Based Systems for Peak Load Reduction," Proceedings of 35th International Convention, MIPRO 2012, pp. 919-923, Opatija, Croatia, 21.5.-25.5.2012.

POSTER SESSION
RADIO COMMUNICATIONS, MICROWAVES, ANTENNAS

Application of Microwave Hyperthermia in Oncology

Hristo Gochev [1], Boncho Bonev [2] and Peter Petkov [3]

Abstract – **In this article the basic concepts of microwave hyperthermia are given. The general types of hyperthermia are discussed and results of clinical studies on hyperthermia are given. Types of microwave hyperthermia applicators are presented – single and multielement array applicators. Methods for temperature measurement of body tissues are also given. In the final paragraph there is a short overview of the problems that have to be solved in order to improve the efficiency of the method.**

Keywords – **Microwave hyperthermia, Cancer treatment, Applicators, Temperature measurement.**

I. INTRODUCTION

The research of the effects of Electromagnetic Field on biological systems is of growing interest. High levels of electromagnetic radiation are considered harmful to human health that's why many countries have introduced Specific Absorption Rate (SAR) limitations. SAR is a measure of the rate at which radio frequency (RF) energy is absorbed by the human body. ANSI standards introduce safe exposure levels of SAR of 4W/kg. A safety factor of ten is incorporated so the final recommended protection guidelines of a SAR level is 0.4W/kg. However there is a disagreement over exactly what levels of absorbed RF energy is harmless. For instance in eastern European countries SAR thresholds are much lower that the thresholds introduced in western European countries.

Apart from the harmful aspects, Electromagnetic Field has some positive effects when used in medicine in the fields of oncology, physiotherapy, urology, cardiology, surgery, ophthalmology.

II. MICROWAVE HYPERTHERMIA TO TREAT CANCER

Microwave hyperthermia is procedure of increasing the temperature of the human body or some part of it using microwaves for the purpose of treating medical conditions such as cancer. There are three broad categories of hyperthermia- localized hyperthermia, regional hyperthermia and whole body hyperthermia [1][2]. Using microwaves body temperature is increased in the interval 41 and 45℃.

In localized hyperthermia heat is applied to a small area

[1]Hristo Gochev is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: gochev.h@gmail.com.

[2]Boncho Bonev is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: bbonev@tu-sofia.bg..

[3]Peter Petkov is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: pjpetkov@tu-sofia.bg..

where the tumor is located using different methods of delivering the energy [3]. Depending on the tumor location there are three approaches to introduce microwave heating:

- External method is used to treat tumors that are located on just below the skin. This approach uses external applicators that are positioned around the appropriate region and microwave energy is focused to raise the temperature of the desired part of the body.
- Intraluminal or endocavitary methods may be used to treat tumors within or near body cavities. This approach is based on inserting probe in the cavity in question for directly heating the tumor.
- Interstitial method is used to treat tumors located deep within the body, such as brain tumors. Imaging techniques such as ultrasound are used to determine the exact location of the tumor. After that, under anesthesia, probes or needles are inserted onto the tumor with high precision. Microwave radiation is then applied to the tumor to increase its temperature with the purpose to kill cancer cells.

Regional hyperthermia is used to treat larger areas of the human body such as a limb, body cavity or an organ [3].

- Deep tissue approaches may be used to treat cancers within the body, such as cervical or bladder cancer. External applicators are positioned around the body cavity or organ to be treated, and microwave or radiofrequency energy is focused on the area to raise its temperature.
- Regional perfusion techniques can be used to treat cancers in the arms and legs, such as melanoma, or cancer in some organs, such as the liver or lung. In this procedure, some of the patient's blood is removed, heated, and then pumped (perfused) back into the limb or organ. Anticancer drugs are commonly given during this treatment.

Whole-body hyperthermia is used to treat metastatic cancer that has spread throughout the body. This can be accomplished by several techniques that raise the body temperature, including the use of thermal chambers (similar to large incubators) or hot water blankets [3].

The effects of hyperthermia on the host and cancer tissue are pleiotropic and depend mainly on the temperature and the physical techniques applied. The biological and molecular mechanisms of these effects are changes in the membrane, the cytoskeleton, the ion-gradient and membrane potential, synthesis of macromolecules and DNA-replication, intra- and extracellular pH and decrease in intracellular ATP. Genes can be up-regulated or down-regulated by heat, for example the heat-shock proteins.

Synergistic effects by interactions with antineoplastic agents, radiation and heat can be several powers of ten even at moderate temperatures. In addition, reduced chemotherapy resistancy, possibly due to increased tissue penetration,

increased membrane permeability, and activated metabolism, has been observed.

Immunological effects of hyperthermia may play an additional role in cancer therapy such as immunological effects on cellular effector cells (emigration, migration and activation), induction of cytokines, chemokines and heat shock proteins (chaperones), and modulation of cell adhesion molecules. The induction of heat-shock proteins might increase specific immune responses to cancer cells.

There are three most common ways to use hyperthermia-heat alone, combination of heat and radiation, combination of heat and drugs [1].

Heat alone

This method is based on directly killing the malignant cells using high temperatures in the interval 41- 42 ℃. However the thermal response of the cells depends on micro environmental factors such as pH. Main phenomenon justifying this technique is protein denaturation which is happening in the temperature range between 39 and 45 ℃.

Heat and radiation

Aggregation of nuclear proteins damage is thought to be the central event by which heat makes cells more sensitive to radiation. The synergy between heat and radiation, often expressed as thermal enhancement ratios (TERs), is highest when the two modalities are given simultaneously. When heat precedes radiation, the synergy is lost when the time interval between the two modalities increases; this loss of TERs nicely parallels the decline in protein aggregation.

Heat and drugs

A lot of physiology-related features make a combination of heat and drugs very attractive. Moreover, heat can cause supra-additive killing when combined with alkylating agents, nitrosureas, platinum drugs, and some antibiotics, although for some drugs only additive effects or even less than additive effects on cell death are found. The most impressive results in this regard are for heat and cisplatin treatments. Synergistic killing is already found at rather mild heat treatments.

When cells are exposed at elevated temperatures to drugs, their response is frequently very different from that seen at 37°C. Drugs whose rate-limiting reaction is primarily chemical (i.e., not involving enzymes) would, on thermodynamic grounds, be expected to be more efficient at higher temperatures. The rates of alkylation of DNA, or of conversion of a nonreactive species to a reactive one, can be expected to increase as the temperature increases. Tissue culture studies have shown this to be true for the nitrosoureas and cisplatin. For other drugs, there appears to be a threshold at or near 43°C. Below that temperature, drug activity is only mildly enhanced. At higher temperatures, however, cell killing proceeds at a greatly enhanced rate. The combination of chemotherapy with hyperthermia still deserves attention and has high potential.

III. CLINICAL RESULTS

A study in the Netherlands that used three different systems to deliver hyperthermia found that hyperthermia in addition to radiation may be especially useful in locally advanced cervical tumors [4]. The Dutch study involved 358 patients with bladder, cervical, or rectal cancer, randomly assigned to radiotherapy or radiotherapy plus hyperthermia. For patients with cervical cancer, 3-year overall survival was 27% after radiotherapy alone vs 51% after radiotherapy plus hyperthermia.

A study conducted at nine centers in Europe and North America randomly assigned 341 patients with localized, high-risk soft-tissue sarcoma to neoadjuvant chemotherapy alone or with regional hyperthermia. Patients receiving the combination therapy had significantly better local progression-free and disease-free survival [5].

The results from a study in Czech Republic are also promising. More than 500 patients were treated by combination of hyperthermia and radiotherapy. Complete response of the tumor was observed in 52.4% of the cases, partial response at 31.7% and no response at 15.9% [6].

IV. MICROWAVE APPLICATORS USED FOR HYPERTHERMIA

Single applicators

Early hyperthermia trials were conducted with single-aperture devices having no ability to steer or focus energy other than shifting patient position relative to the applicator. These trials included 27 MHz ridged waveguide, 82 MHz helix, 70 MHz coaxial TEM applicator, and 27 to 70 MHz evanescent-mode waveguide excited below the cutoff frequency by entering a resonant circuit (lumped capacity and inductance) with a wave impedance build-up band-pass filter for the operating frequency. Most of the microwave equipment includes a water bolus for surface cooling. Low-profile, light-weight microstrip applicators, which are easier to use clinically, are also used. The type of applicator selected depends on the production of sufficient thermal field distributions at different depths of the tumor in a variety of anatomical sites. Single-element applicators can safely deliver optimum thermal doses to relatively small superficial tumors. Over the years, several types of applicators for external local hyperthermia have been investigated by many researchers based on the principle of a dielectric filled waveguide or horn antenna [1].

Multielement Array Applicators

To increase the value of the SAR at depth relative to the surface SAR in hyperthermia therapy, we must geometrically focus energy deposition from multiple E fields generated by an array of applicators [1]. A basic array for external deep heating will likely consist of an annular ring of radiating apertures. The parameters of interest are the external E field

within an array at the surface of the patient's body, the SAR pattern within the target volume, and the radiation leakage levels of the scattered fields around the applicator. Several different RF electrode arrays have been investigated. Two arrays of needle electrodes arranged in two planes, with a bipolar RF current between the arrays were examined. In the bipolar system, RF current is passed between two electrodes instead of between a single electrode and a ground path, so two electrodes heat the tissue instead of one, resulting in a larger ablation zone. Other groups investigated different array configurations, and segmented needle electrodes have been suggested to allow for better control of tissue heating. An array of applicators with variations in phase, frequency, amplitude, and orientation of the applied fields can add more dimensions to controlling the heating patterns during hyperthermia cancer therapy. Because of the constructive interference of E fields at the intended focus and destructive interference of E fields away from the focus, multichannel coherent phased-array applicators can theoretically provide deeper tissue penetration and improved localization of the absorbed energy in deep-seated tumor regions without overheating the skin and superficial healthy tissues, compared to single or incoherent array applicators. When comparing array applicators with a single applicator, array applicators provide deeper tissue penetration, reduce undesired heating of normal surrounding tissues between the applicator and tumor, and improve local control of the tumor temperature distribution. Heat generated by RF devices is delivered regionally across a much larger area. However, a microwave array system requires target compression because of the shallow penetration of the higher microwave frequencies. RF array applicators surrounding the body are used in attempting to heat deep tumors. However, studies in external RF array thermotherapy have shown the difficulty of localizing RF energy in malignant tissue deep within the human body without damaging superficial healthy tissue due to hot spots. Improvements in RF energy deposition are achieved when the RF phased array is controlled by an adaptive algorithm to focus the RF energy in the tumor and tumor margins, while the superficial RF fields are nullified. Clinically, the use of phased arrays as heating applicators has several advantages. Phased arrays can easily compensate for the effects of inhomogeneities of the treatment volume (which includes the tumor and the surrounding tissues). The heating pattern can be controlled electronically, thus eliminating the need for mechanical movement of the applicator head. This simplifies the machine-patient interface and allows for better use of the available power. Also, electronic switching can be performed rapidly, thus enabling swift response to changes in the tumor environment. However, clinicians cannot always accurately predetermine or manually adjust the optimum settings for output power and phase of each antenna to focus heat reliably into deep-seated tumors. Two outstanding challenges in EM phased-array hyperthermia are (1) to selectively elevate the temperature in the cancerous tissue without excessively elevating the temperature of the surrounding healthy tissues in the presence of electrical and thermodynamic inhomogeneities, and (2) to react to unexpected changes in the patient positioning and physiology (e.g., sudden change in

blood flow in the tumor) that can significantly impact the quality of the delivered treatment. Significant research progress has been obtained recently in heating devices appropriate for deep hyperthermia including ultrasonic arrays, RF arrays, and microwave arrays.

V. TEMPERATURE MEASUREMENT

The choice of a particular thermometry system should be made following careful consideration of its intended application. Any measuring instrument can be considered acceptable for a given purpose if the error contributed by the instrument is small compared to other errors or uncertainties in the measurement. Hence, in local hyperthermia where large temperature fluctuations are encountered, an extremely accurate thermometer may not be required. An overall error not exceeding $0.3°C$ is probably acceptable, but achieving this accuracy in a clinical environment may require that the thermometer agree with an institution's standard to within 0.1 or 0.2%. In whole body hyperthermia, on the other hand, where temperatures are only a few tenths of one degree below the lethal limit and homogeneity is excellent, instrument errors of more than +/-0.05% may be unacceptable [7].

Temperature control by microwave radiometry

Human tissues spontaneously emit electromagnetic radiations of thermal origin which can be measured by a very sensitive receiver called "a radiometer". When this measurement is carried out in the microwave frequency range, it is possible to evaluate the tissues temperature. Microwave radiometry is used to detect thermal anomalies inside the human body but also to evaluate, noninvasively the temperature of the body. In the microwave domain, the thermal noise power emitted by the body is directly proportional to the temperature and can be obtained by integrating the spectrum brightness B(f). The temperature of a dissipative body can thus be determined by a measurement of the electromagnetic power radiated in a given frequency bandwidth. This measurement is achieved by systems which use an antenna as an electromagnetic power captor in the microwave region [8].

Temperature measurement using thermocouples

Thermocouples are used as invasive thermometers in clinical hyperthermia. The probe is inserted in the area of interest and the temperature is measured directly [8].

For measuring the surface body temperature an infrared camera may be used.

VI. FUTURE CHALLENGES

Although a lot has been done in the field of microwave hyperthermia, there are still many obstacles to be overcome. Some of the major technological challenges that make hyperthermia therapy complicated are the inability to achieve

uniform temperature distribution in a tumor, inability to precisely monitor the temperatures of the tumor and the healthy tissue and wavelength limitations.

To achieve uniform temperature in the tumor and to focus the electromagnetic energy in the tumor alone and not in the surrounding tissues, more sophisticated applicator designs are required. To face this challenge more work has to be conducted by examining how electromagnetic energy propagates through the human body and determining optimal radiation patterns for the applicators.

In order to have more concentrated beam of energy a higher frequency wave must be used. From the other hand it is known that the penetration ability on the waves decreases for shorter wavelengths. This leads to limitations on the frequencies to be used. For superficial tumors higher frequency wave can be used. If we want to achieve deeper penetration however a radio wave with lower frequency must be used.

Very important part of hyperthermia is to monitor the temperature of the tissues during the treatment. There is a limitation to monitor the temperature at the same time the tissue is being subjected to RF radiation because the RF field might interfere with the temperature measurement equipment. This prevents real time temperature measurements and is a drawback for temperature control. Improvements in this area are also necessary to increase the control over the process of heating.

As a conclusion it's evident that a lot of research must be conducted before a fully functional equipment for microwave hyperthermia is created and released for general usage.

REFERENCES

[1] Riadh W. Y. Habash, Rajeev Bansal, Daniel Krewski, Hafid T. Alhafid, "Thermal Therapy, Part 2: Hyperthermia Techniques"

[2] van der Zee J., "Heating the patient: a promising approach?" Annals of Oncology 2002;

[3] Wust P, Hildebrandt B, Sreenivasa G, et al. "Hyperthermia in combined treatment of cancer." *The Lancet Oncology* 2002; 3(8):487–497.

[4] van der Zee J, Gonzalez DG, van Rhoon GC, et al, "Comparison of radiotherapy alone with radiotherapy plus hyperthermia in locally advanced pelvic tumours: A prospective, randomised, multicenter trial." Lancet 355:1119-1125, 2010.

[5] Issels RD, Lindner LH, Verweij J, et al, "Neo-adjuvant chemotherapy alone or with regional hyperthermia for localised high-risk soft-tissue sarcoma: A randomised phase 3 multicenter study." Lancet Oncol 11:561-570, 2010.

[6] Jaroslav Vorlíček, Barbora Vrbova, Jan Vrba, "Prospective Applications of Microwaves in Medicine", Microwave Techniques, 2008. COMITE 2008. 14th Conference, 1-4.

[7] Geoffrey S. lbbott,Ivan Brezovich, Peter Fessenden, Yakov Pipman, Taljit Sandhu, V. Sathiaseelan, Paul Stauffer, Adrianne Galdi, Tillman Saylor, "PERFORMANCE EVALUATION OF HYPERTHERMIA EQUIPMENT, REPORT OF TASK GROUP NO. 1 HYPERTHERMIA COMMITTEE"

[8] Mark Haynes, John Stang, Mahta Moghaddam, "Real-time Microwave Imaging of Differential Temperature for Thermal Therapy Monitoring", J. Phys. D: Appl. Phys. 38 (2005) 2633–2639

System Approach to Ka-Band Earth Station Beam Pointing Accuracy Verification

Peter Petkov [1]

Abstract – **In this paper an experimental setup for verification of Earth station antenna pointing accuracy is described. Collected experimental data for two cases is presented. Conclusions are delivered about the functionality of the experimental setup and the antenna pointing system.**

Keywords – **Ka band tracking, Earth station antenna, RF data collection, pointing accuracy.**

I. INTRODUCTION

Recent development in broadband satellite communications brought to use the very promising Ka frequency band. However due to exceptional requirements to Earth Station Antennas pointing, the successful tracking of the spacecraft is impaired by variety of problems of mechanical and electrical matter. In order to verify the consistency of the tracking system, a reference setup, described in the paper, has been developed and successfully tested.

One of the most important features of the medium (3-9m) and large (9-18m) Earth Stations is the pointing accuracy. It is widely recognized that pointing accuracy within 0.5 dB is considered as good. However it puts challenges to the pointing mechanics and also to the instrumentation to prove that accuracy [1-4].

II. THE MEASUREMENT SETUP

The measurement setup is depicted on Fig, 1. It consists of a Large Earth Station (9m Antenna Under Test), Small size (3ft) reference antenna, RF Switch with integrated LNA, Spectrum analyzer and PC With Labview platform to collect the measurement data and control the switch. Similar setups are defined by other authors [2]

The size of the reference antenna is selected in order that Antenna Pattern is broad enough to accommodate the diurnal satellite drift within 0.05 dB down from the antenna beam peak. Therefore the signal coming from this antenna can be used as reference for the signal from AUT.

The other parameters are as follow: Data acquisition rate - 2 sec., Duration of full data cycle - 24h (86 400 sec.), Spectrum Analyzer settings: 1.5 MHz span, 801 points per sweep, LNA Noise temperature: 105K@20GHz. LNA Gain 53dB, AUT Gain 62dBi. On Fig.2 two approaches of data collections are

[1]Peter Petkov is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: pjpetkov@tu-sofia.bg.

displayed. The Blue line represents Beacon peak value over time, the red one represents integrated channel power (beacon included) over time. Due to scintillations reduction, the second approach was selected to perform the measurement.

The reference antenna is used to provide a stable reference beacon level, independent of the satellite movement. The antenna has small aperture size and therefore broad main beam. The satellite movement in the box will not cause signal variation due to pattern slope more than 0.05dB.

Fig. 1. Data Acuisition system (antenna photos are exemplary)
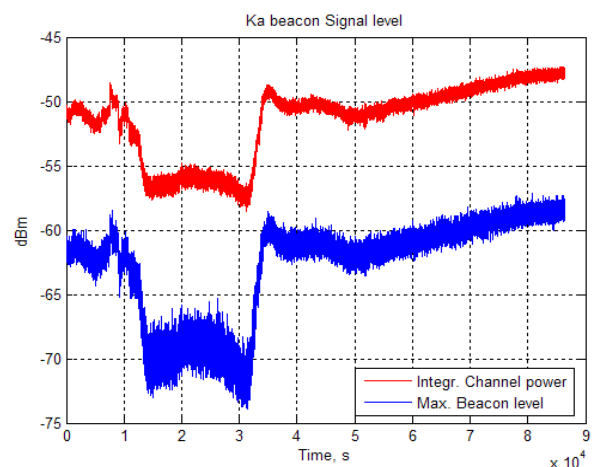
Fig. 2. Satellite beacon reception level

III. THE EXPERIMENTAL DATA

The real conditions measurement data is displayed on Fig. 3. Every second the switch flips back and forth and spectrum analyzer reads AUT channel or Reference Antenna channel either. One full data cycle takes 86 400sec (24h).
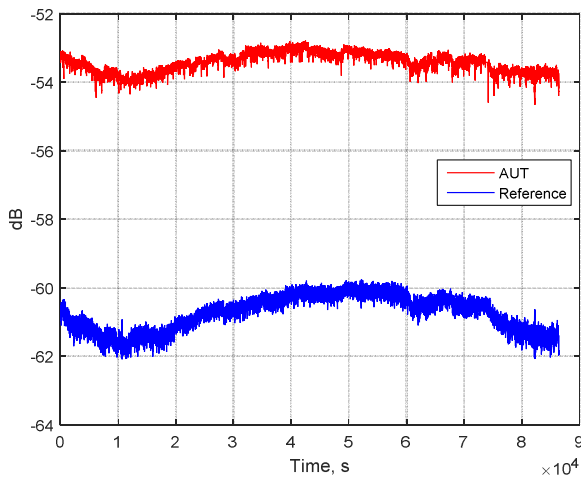
Fig. 3. Measurement Data (Both channels)

The difference (pointing accuracy) is displayed on Fig. 4. For verification purposes, the tracking function of the AUT was switched off. It is clearly evident that despite the exceptional 3D stabilization, the satellite shows a diurnal drift of about 1dB in signal level for the selected AUT.
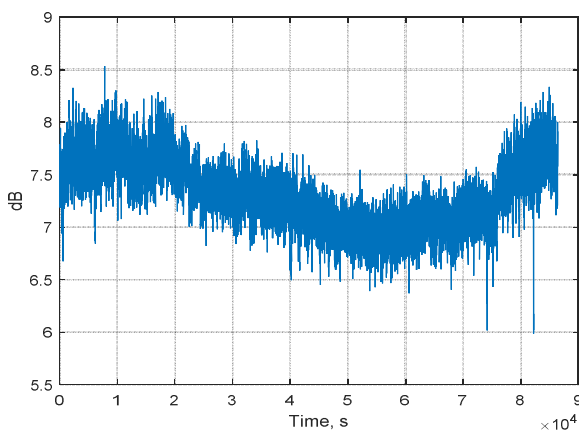
Fig. 4. The pointing accuracy (AUT tracking system turned off)

On the next set of test cycle the AUT tracking system was turned on (at approx. 15 000-th second on Fig. 5). The Mispointing error immediately fell below 0.5dB which is
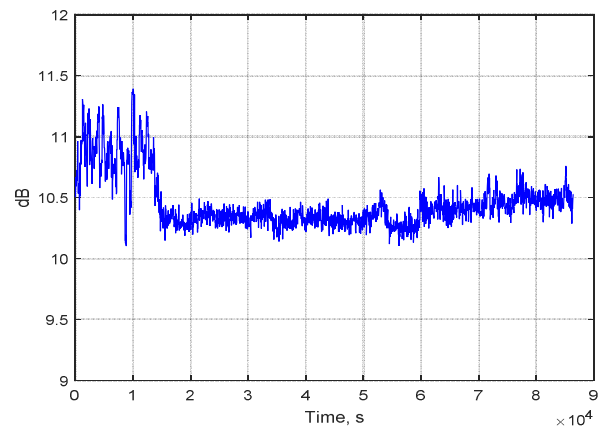
Fig. 5. The pointing accuracy (AUT tracking system turned on)

within the requirements for Earth station pointing accuracy and keep the same level until the end of the test cycle. Despite the rainy weather conditions the tracking system kept the AUT well pointed toward the satellite. It must be mentioned that the original setup was provided with complex weather station, measuring ambient temperature, wind speed and solar radiation for the purposes of in-depth structural analysis. However such analysis is not the purpose of this paper and therefore not included here.

The measurement data is collected and stored with a Labview application and eventually processed in Matlab. Simple moving averaging is applied for the cases presented here. However it is clearly evident that there is a need of more sophisticated data processing.

IV. CONCLUSION

From the measurement results and data presented above it can be concluded that:

1. Data collecting and verification system is feasible enough to deliver continuous and consistent tracking data and completely fulfils measurement accuracy requirements.

2. Earth station tracking system completely fulfills the requirements for pointing accuracy.

REFERENCES

[1] Brain, D.J.; Dang, N.D.; Davis, I.M.; Granville-George, D.A., "Ka-band tracking antenna for inter-orbit communications," *Antennas and Propagation Society International Symposium, 1989. AP-S. Digest* , vol., no., pp.1606,1609 vol.3, 26-30 June 1989

[2] Holzman, E.L.; Eisenhart, R.L.; Robertson, R.S.; Kurtz, L.A.; Kihm, R.T., "A low-cost Ka-band Cassegrain polarimetric tracking antenna," *Antennas and Propagation Society International Symposium, 1992. AP-S. 1992*

[3] Gawronski, Wodek; Gudim, M.A., "Design and performance of the monopulse control system," *Antennas and Propagation Magazine, IEEE* , vol.41, no.6, pp.40,50, Dec 1999

[4] Bayer, H.; Krauss, A.; Stephan, R.; Hein, M.A., "Multimode monopulse tracking feed with dual-band potential for land-mobile satellite communications in Ka-band," *Antennas and Propagation (EUCAP), Proceedings of the 5th European Conference on* , vol., no., pp.1169,1172, 11-15 April 2011

Passive Optical Sensor Network with Energy Harvesting

Jovan Shikoski[1] and prof. Rumen Arnaudov[2]

Abstract **- We have considered an optical communication system for data transfer from sensor node powered by an energy harvesting system. The advantages of this data transfer method, compared to RF modules, are the greater distance of transferred data and the resistance to electromagnetic interference.**

Keywords – **Passive optical sensor network, energy harvesting, sensor node.**

I. INTRODUCTION

In the principles of energy harvesting systems, as a means of communication are accepted for use RF modules characterized by low consumption, which is a primary and basic requirement for the design of these sensor systems. The use of RF modules for communication in the conditions of strong electromagnetic fields or radiowaves, proves to be a tall order, having in mind the interference as a basic phenomenon in such environments. When the disturbances are greater, communication errors occur. This necessitates the use of a greater transmit power, respectively a higher consumption, which contradicts the basic principle of energy harvesting: maximum energy economy and effectiveness [1].

II. ENERGY HARVESTING FOR SENSOR NODES

The sensor node is the basic energy consumer in the energy harvesting process. The power consumed varies from tens of milliwatts to ones of microwatts, depending on the working mode. When the supply is provided by the environment, the converter of energy, for example photovoltaic, transforms the solar energy into electric with a certain effectiveness. For the normal functioning of a system in the energy harvesting process, it is necessary for the power gained from the environment to be greater than the average consumed power of the sensor node. The energy is stored in a rechargeable battery or a supercondenser, and the inconsistency of the energy derived from the environment needs to be taken into account. The use of a DC/DC convertor makes it possible for the photovoltaic to work at a maximum power point (MPP) and thus to increase the gained energy [2]. We have used switching DC/DC convertors with hysteresis control of the output voltage in order to achieve low consumption of the convertor with high effectiveness.

[1]Jovan Shikoski is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: icest@tu-sofia.bg.
[2]Rumen Arnaudov is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria.

III. PASSIVE OPTICAL SENSOR NETWORK (POSN)

The building of sensor networks allows flexible solutions for monitoring, control and management of different parameters of physical environments. The value which are observed, are: temperature, humidity, lighting, pressure, vibrations, noise level, etc. The presence of powerful electromagnetic interference strongly restricts the possibility of normal RF communication. As a solution to this problem is suggested the building of a passive optical sensor network which consists of: base station, optical fibers, splitters and smart sensors.

A. Fiber-optic communication for sensor node

A block scheme of an optical communication system with optical fiber, suitable for sensor nodes with energy harvesting is shown in fig.1. It consists of a transmitter and a receiver. In order for communication to take place along a fiber, a passive optical splitter 1x2 is used. At the two ends, the receiver and the transmitter are plugged, and at the input of the splitter is the optical fiber along which the information is transmitted to the base station.
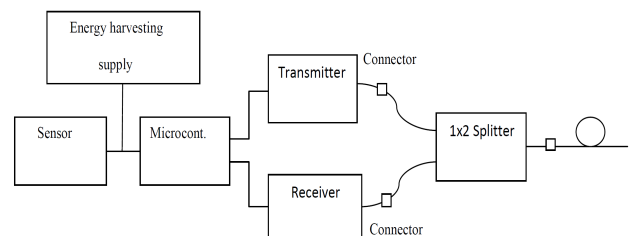
Fig.1 Sensor node

B. Transmitter

In order to transmit the optical signal, is used a highly effective infrared laser OPV314YAT, with wavelength 850 nm and an in-built photodiode for precise control of the optical power. The high speed of data transmission (2,5 Gbps), the great optical power 600μW (-2,2 dBm) and the low consumption 7 mA at a working voltage of 1,6 V to 2,2V makes it perfect for application in the energy harvesting systems. In order to manage the impulse mode and control of the output optical power of the laser, is used a driver of the type iC-NZN. In fig. 2 is given the model scheme of the optical transmitter. The Driver iC-NZN is designed to maintain the control of the optical power of all laser configurations (N, P or M).
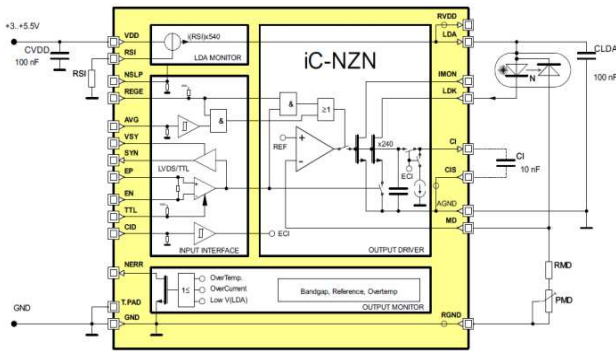
Fig. 2 Transmitter

The optical power level of the laser is chosen with the aid of the resistor $R_{MON}=R_{MD}+P_{MD}$. In order to calculate R_{MON}, one should know the value of the monitor electricity IM of the laser diode. It is actually the reverse current through the photodiode of the laser and depends on the optical power. R_{MON} needs to be chosen in a way that the monitor electricity generated by the desired optical power, creates a lowering of voltage of 500 mV along the whole length of R_{MON}. The monitor electricity I_M of the laser OPV314YAT is 30μA.

$$R_{MON} = \frac{V_{MP}}{I_{MP}} = \frac{0.5}{0.03} = 16k\Omega \qquad (1)$$

A higher standard value can be chosen. iC-NZN also controls the average laser electricity running from the output of LDA. The restriction of the electricity through the laser is set through the resistor of pin R_{SI}. Sizing of the resistor R_{SI}:

$$I_{\max(LDA)} = 540.\frac{0.5}{R_{SI}} \qquad (2)$$

$$R_{SI} = 540.\frac{0.5}{I_{\max(LDA)}} = \frac{270}{0.007} = 38.57k\Omega \qquad (3)$$

The lower standard value is chosen.When it is in working mode, iC-NZN is characterized by lower consumption (10÷15 mA), sleep mode - 5μA, with supply voltage 3÷5 V.

C. Receiver

The receiver HFBR-2412TCZ in fig.3 includes a photodetector, DC amplifier and Schottky output transistor. The receiver is designed to work with supply voltage of 3÷5 V
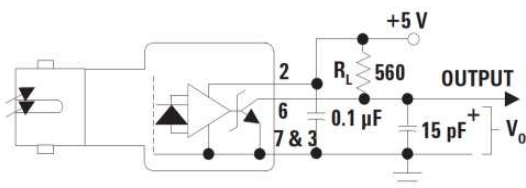
fig. 3 Receiver

and maximum electricity up to 10 mA. The input optical power varies from 0,1μW (-40 dBm) at a high level and 120 μW (-9,2 dBm) at a low level.

D. Base station

The base station is a device controlling the data transmission from different smart sensors in the network, as well as their re-programming. Through a serial interface, the data received from the sensors are transmitted to a computer, where monitoring is done. In fig. 4 is shown a block scheme of the base station.
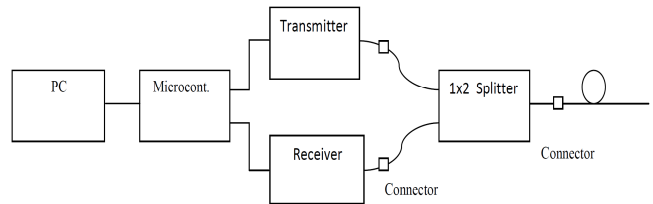
Fig. 4 Base station

The microcontroller of the base station controls and manages the processes in the sensor network. It makes a query to each sensor for data availability and transmits the received data to the computer. Through the microcontroller to the computer can be carried out a distance re-programming of each smart sensor separately. With the aim of a greater compatibility, the communication part of the base station is similar to the one of the intelligent sensor. It consists of an optical receiver HFBR-2412TCZ, and the transmitter is an infrared laser OPV314YAT, controlled by the driver iC-NZN.

F. Working principle of passive optical sensor network

The base station is connected, through a multi-mode optical fiber 62,5μm, to the input of the passive optical splitter [3]. On the other side, the 32 outputs of the splitter are connected to 32 sensors – one at each sensor is an autonomous product, independent on the other sensors and most of the time it is in sleep mode. When the measured value changes, the sensor is activated, it records the respective change in the memory and switches to economical (sleep) mode again. Each sensor has its own address to which it responds. The base station queries for data availability to the first sensor with address S1, the optical signal is received by all sensors, but only the one to which the query was sent, responds, in this case, S1. When there are data in the memory, the sensor sends them and switches to sleep mode again. The query continues to the next sensor with address S2, the sensor is activated and if there are no data in the memory, it responds "NO". If a sensor does not respond to the query sent to its address, the base station alarms for the absence of this sensor. The queries continue until the last address, in this case S32. After that, this query cycle finishes.
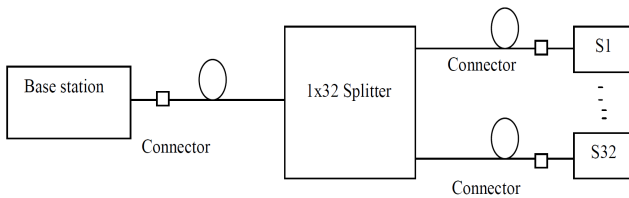
Fig. 4 Passive optical sensor network

A new query cycle begins in a certain time, which is defined during the design of the sensor network, depending on the change frequency of the measured value. The number of sensors in the network depends on the number of outputs in the optical splitter. Except for splitters with 32 outputs, such with 8, 16 and 64 outputs are also produced.

G. Design of passive optical sensor network (POSN)

When designing a passive optical sensor network (POSN) [4] one needs to calculate the total energy losses as a sum of the losses in all network components:

$$B = n_{connector}B_{connector} + n_{splice}B_{splice} + (z_1 + z_2)b_f +$$
$$B_{32\,splitter} + 2B_{2\,splitter} + B_{re}$$
$$(4)$$

where: $n_{connector}= 4$ is the number of connectors from the sensor to the base station, $B_{connector} = 0,5$dB are the losses in the connectors, $n_{splice} = 5$ – number of splices in the optical fiber, $B_{splice}= 0,1$ dB are the losses in the splices of the optical fiber, $b_f = 2,3$ dB are the losses through a multi-mode 62,5μm optical fiber when the light wavelength is 850 nm [5]. The total length of the optical fiber in the system is $z=(z_1+z_2)$, where z_1 is the length of the optical fiber from the base station to the 1x32 optical splitter, and z_2 is the length from the splitter to the respective sensor. The losses in the optical splitters are $B_{Nspliter}$ and they are calculated by the formula:

$$B_{Nsplitter} = IL + 10\lg N$$
$$(5)$$

where - N is the number of outputs from the splitter, Insertion Loss are given by the manufacturer IL= 3[dB]. B32spliter= 18 [dB] – the losses in the 1x32 optical splitter, B2spliter= 6 are the losses in the 1x2 splitters (one in the base station and one in the communication system of the sensor). Bres is provided in case of additional losses (unexpected bends, pollution of the connectors, extra splices, etc.). The energy design amounts to the verification of the condition:

$$B \leq P_{owerBudget}$$
$$(6)$$

$$P_{owerBudget} = P_{TX} - P_{RX\,min}$$
$$(7)$$

On the basis of the calculations above, we can calculate the maximum length of the optical fiber with the given parameters:

$$z = (z_1 + z_2) = \frac{B - n_{connector}\,B_{connector} + n_{splice}\,B_{splice}}{b_f} +$$
$$+ \frac{B_{32\,splitter} + 2B_{2\,splitter} + B_{res}}{b_f}$$
$$(8)$$

During the design of POSN with the devices described above: PTX = -2,2 dBm, PRXmin= -40 dBm, for the energy budget from expression (7) is derived PowerBudget = 37,8 dB. When this is substituted in condition (6), the total losses should be B ≤ 37,8 dB. If we accept B= 37,8 [dB], it means that the energy use is at the upper limit for data transfer and each loss increase means interruption of the communication along the optical fiber. When the results in expression (8) are substituted, the maximum admissible length of the optical fiber zmax=1 km is derived. For stable communication, we need to choose an optical fiber with a length smaller than zmax, for example up to 950m, which is a distance 8 times greater than the RF module. When the fiber length is 950m from expression (4) for the total losses we derive B=37,7dB, which means that condition (6) is fulfilled. The low power consumption of the proposed optical communication systems meet the basic criteria for energy harvesting. The total consumption in data transmission is 22mA, 10mA in receive mode, while in sleep mode is 5μA at a supply voltage of 3÷5V. A comparison of the power consumption of the examined optical communication system and a RF module MRF24J40MA, both in transmitting and receiving mode is given in the following table.

TABLE I
Comparison of consumption

	Optical communication system		RF module MRF24J40MA	
	TX	RX	TX	RX
I_{CC} (mA)	22	10	23	19
V_{CC} (V)	3÷5	3÷5	2.4÷3.6	2.4÷3.6

H. Network protocol

The block algorithm of Network protocol, by which the communication in POSN is carried out [6], is given in fig. 6.
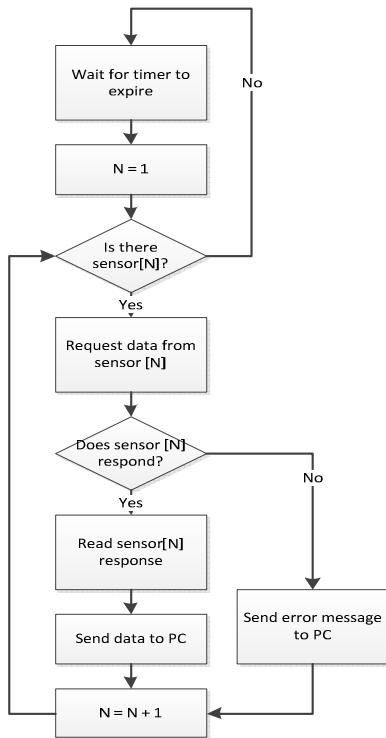

Fig. 6. Working principle of the base station
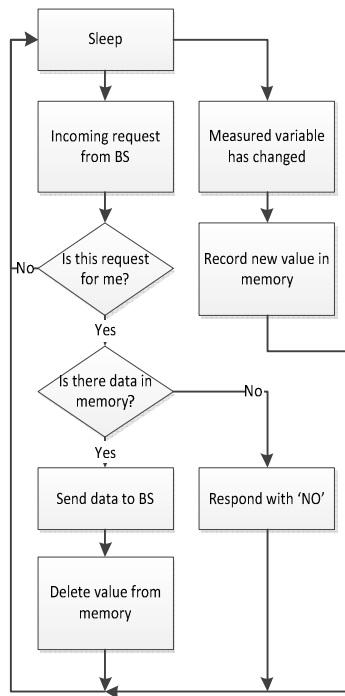
I. Working principle of sensor

Fig. 7. Working principle of each sensor

Most of the time, the sensor nodes are in economical sleep mode. When the physical phenomenon changes, the sensor wakes up, measures the variable quantity, records the measured data and switches to sleep mode. When there is a call from the base station with its address, it is activated, sends the data (if there are no data from new measurements, it responds NO) and goes back to sleep mode. The block algorithm of the working principle of each sensor is given in fig. 7.

IV. CONCLUSION

The low consumption of electric energy is the main parameter on which is focused the design of the suggested energy harvesting sensor system. An order is proposed for the design and calculation of the maximum fiber length, considering the energy budget with guaranteed reliability of communication in conditions of strong electromagnetic interference. Except in conditions of strong electromagnetic interference, the optical communication system can replace the RF module, when required a greater distance between the sensor node and the base station. We have developed algorithms for managing the network and sensor activity.

REFERENCES

[1] Zahariev P., Hristov G., Tsvetkova I.,. An Approach towards Balanced Energy Consumption in Hierarchical Cluster-based Wireless Sensor Networks.// Journal of Computing and Information Technology Vol.20, 2012, No 3, pp. 159-165, ISSN 1330-1136.
[2] Oscar López-Lapeña, Maria Teresa Penella, A New MPPT Method for Low-Power Solar Energy Harvesting Student Member, IEEE, and Manel Gasulla, Member, IEEE,Vol. 57, No. 9, September 2010 p.p. 3129-3138
[3] Calvin C. K. Chan, Passive Optical Networks Principles and Practice, The Chinese University of Hong Kong, 2007,
[4] Mitzev C., Dimitrov K., Optical Communications seminar exercises guide, Technical University, Sofia, ISBN:978-954-438-77-8, p.p. 10-27, 2013.
[5] Corning Incorporated, Corning Clear Curve Multimode Optical Fiber Product Information, , p.p.2-4, 2011.
[6] Michael P. McGarry, M. Reisslein, M. Maier, Ethernet Passive Optical Network Architectures and Dynamic Bandwidth Allocation Algorithms, University of Akron Arizona State University, IEEE CS&T, Vol.10, October 2008, p.p. 46-60.

About some problems in the transmission of signals in HFC/CATV networks

Oleg Borisov Panagiev[1]

Abstract – **In this paper are presented the results of a research into the influence of the change in the parameters of elements of the coaxial part of a HFC/CATV network, as a result of different unfavorable factors. This on its part leads to worsening in the quality of receiving or inability to receive, as well analogue (CSO, CTB < 60 dB) as digital (BER increases, MER and C/N decreases) signals. Thus it is necessary to research and analyze the influence of the parameters of the network and signals, transmitted in it, onto BER, MER, C/N, CSO and CTB and to make the corresponding recommendations to insure a promising and qualitative signal transmission in the HFC/CATV network.**

Keywords – **HFC/CATV, distortions, BER, MER, C/N.**

I. INTRODUCTION

Cable television systems are traditionally one-way and distribute mostly television and radio signals to the subscribers. Since the development of Internet and HDTV, cable TV distributors have been interested in distributing those services, too. A good solution is to transmit analog and digital signals together in hybrid communication networks by optical fibers and coaxial cables. It is a fact that the hybrid transmission of the existing AM-VSB and the new digital (M-QAM) signals is essential to the expansion and the increase in efficiency of cable television systems in the future [1], [2].

The architecture of the HFC network is bidirectional communication infrastructure (Fig. 1). Each group of 500 to 2000 subscribers is serviced by two single mode fibers, one for "Downstream" and another for "Upstream" channel, which connects the Head End and the fiber node. Furthermore, signals from the fiber node to the subscribers are transmitted via coaxial cable through few amplifiers and taps/splitters by a frequency division: from 5 MHz to 65 MHz ("Upstream" channel) or from 85 MHz to 862 MHz ("Downstream" channel).

HFC/CATV systems are one of the most attractive, because the M-QAM signals provide high spectral efficiency and purity considering the noise and the nonlinear distortions. However from "clipping" in the laser diode (DFB – distributed feedback) is generated a noise, which causes deterioration (increase) of BER (Bit-Error-Rate) when transmitting simultaneously analog and digital signals [3], [4]. This paper examines the experimental results from the investigation into the influence of the variation of devices' parameters on the coaxial part of the network: increase/decrease of the levels of the analog (AM-VSB, FM) and digital (M-QAM) signals.
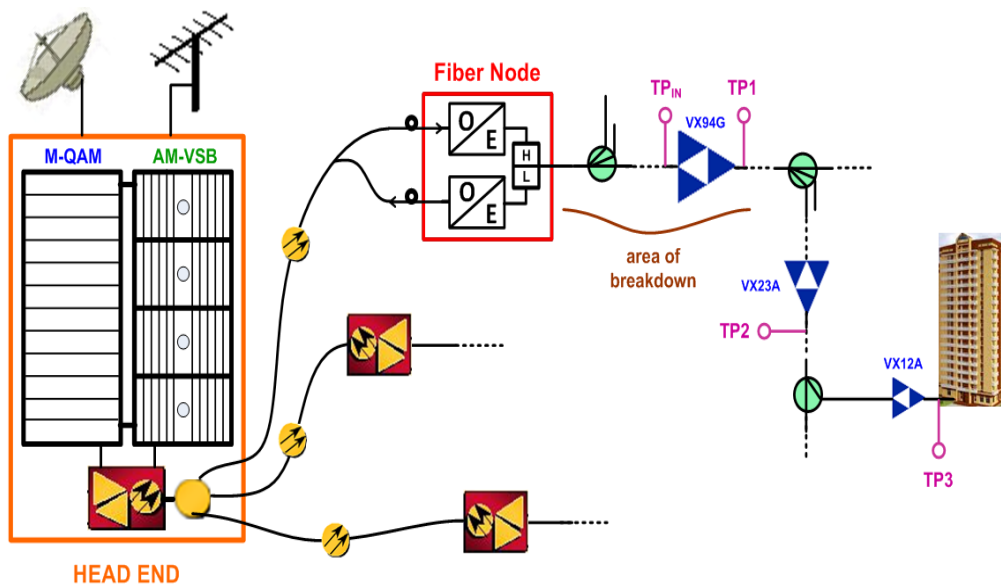
Fig.1. Experimental setup of part of a HFC/CATV network

[1]Oleg B. Panagiev is with the Technical University of Sofia, Bulgaria, E-mail: olcomol@yahoo.com.

Due to disturbances in the integrity of the coaxial cable (trunk, subtrunk, drop), the nominal levels of the signals change. Reasons for that could be: soil consolidation, rodents, construction works, cracking of the outer insulation due to manufacturing defects or such, caused during the installment process into the underground ducts (Fig. 2), [5]. Other causes for the variation in the nominal level of the signals could be: change in the parameters of the taps and splitters; degradation of the contact resistance, in the level regulators, and of the slope of the amplitude-frequency response of the electronic amplifiers; defects in the amplifying elements (monolithic and hybrid integrated circuits, transistors); change of the supply voltage (problems with SMPS); problems in AGC/ALC, NMS modules and others. Effects from such problems (in the present or next element/device) are at first a level change of the signals (linear distortions), but after the signals have entered the input of the next amplifier they can cause non-

Fig.2. Cable fault - two cables

linear distortions (increased level of the signals compared to the nominal) or degradation of the noiseimmunity (decreased level of the signals compared to the nominal).

Since the amplifiers are connected in series (Fig. 1), the unfavorable effects multiply, which leads to decrease in the quality of receiving or to the impossibility of receiving both the analog (CSO, CTB < 60 dB) and digital (BER increases, MER and C/N decrease) signals. In this sense it is required to measure and analyze the influence of the parameters of the network and the signals, transmitted through it, over BER, MER, C/N, CSO and CTB.

II. EXPERIMENTAL SETUP AND RESULTS

A. Initial conditions

In this section are presented the results of the research made in a part of real functioning HFC/CATV system/network, which carries 75 channels (Fig. 3). Part of the analog channels are located between the digital ones (in the range above 470 MHz), and some of them coincide by frequency with broadcasted terrestrial channels (DVB-T, LTE and others) at the place of the experiment.

The studies were carried out for the worst case: when the studied channels coincide by frequency with channels from DVB-T (channels 40 and 52) and when parts of the trunk cable and the subtrunk cable are under the ground – PVC conduit in trench and pit/cable chambers (Fig. 4), and on the power poles (Fig. 5), [6].

Fig.4. Cable chamber

The carrier frequencies of the channels are synchronized in the Head End (Fig. 1), by a PLL, and their levels (Table1) are equalized according to [7], [8], the parameters of the amplifiers and to the ΔU_n (Table2). The equalization of the signals' levels in the composite/group signal is done for channels SR1 and 45 (standard D/K). The difference between the levels of the analog (AM-VSB) and digital (M-QAM) channels is 10÷13dB. The composite nonlinear distortions (CSO, CTB) are measured in channel 22. M=256 for channel 40 and M=64 for channel 52. The electronic amplifiers that were used are by the WISI company: VX94G, VX23A and VX12A (Table3), [9].
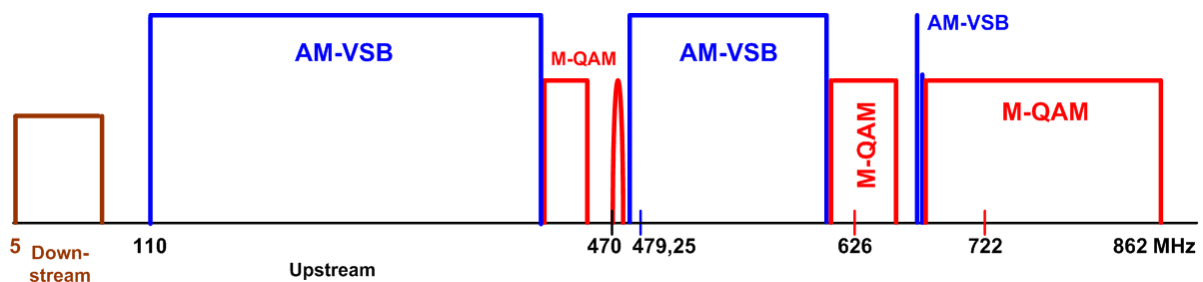
Fig.3. Spectrum of HFC/CATV system

TABLE 1

TP	TPin			TP1			TP2			TP3		
Channel	22	40	52	22	40	52	22	40	52	22	40	52
f, MHz / Parameter	479,25	626	722	479,25	626	722	479,25	626	722	479,25	626	722
CSO, dB	64			67			67			67		
CTB, dB	>63			64			64			62		
preBER		1,5E-7	6,2E-8		2,8E-6	1,5E-8		1,8E-9	3,6E-9		1,2E-7	1,3E-8
postBER		9E-10	1,2E-9		9E-10	1,2E-9		9E-10	1,2E-9		9E-10	1,2E-9
MER, dB		37	29		36	29		40	30		37	29
C/N, dB	44	30	31	43	29	30	43	30	33	43	29	30
U, dBμV	69	66	66	91	90	90	102	101	101	89	87	88

Fig.5. Power pole

The measurements are done for 4 test points (TPin, TP1, TP2, TP3) according to Fig.1, with the area of breakdown being close to the fiber node (around the trunk amplifier). The initial state is at normal operating mode of the HFC/CATV system (Table1), after that the levels of the signal increase/de-

TABLE 2

Number of channels n	42 CENELEC	55	65	75	85
ΔU_n dB	12,1	12,99	13,55	14,02	14,43

TABLE 3

Parameter / Model amplifier	$U_{out,2}$ dBμV	Gain dB	Attenuator dB	TILT dB
VX94G	121	28	10	20
	109*			
VX23A	122	35	<10	>10
	110*			
VX12A	114	38	18	18
	98*			

Note: * 42 channels CENELEC EN50083

crease to 10 dB. In the first case, the influence of the nonlinear distortion's increase over BER, MER, CSO and CTB is investigated, as well as the alternation of C/N. The second case studies the influence of the decrease in the noiseimmunity due to the impaired shielding and grounding. This leads to the growth of the interferences from other radiocommunication devices and systems/networks (DVB-T, LTE, radio transceivers – used by amateurs, police, fire brigade and etc.) in air mounted coaxial cables and devices.

B. Results

The results from the measurements are presented graphically (Fig.6, Fig.7, Fig.8, Fig.9, Fig.10) for three channels - one analog (AM-VSB) 22 and two digital (QAM) - 40 and 52 by D/K standard. The constellation diagrams are also given for the test point TP3 in normal operating mode and in breakdown mode. From the graphics, it is seen that for the analog signals: C/N worsens with 5-10 dB (TP3) in the edges of the investigated range of alternation of U_{out}; CTB with $\Delta U_{out} \geq 6$ dB drops to 50 dB for TP2 and to 48 dB for TP3 and with $\Delta U_{out} \leq -4$ dB its values drop under 60 dB and reach 53 dB for TP3. The change of the nominal level substantially influences the digital signals, as follows: channel 40 MER it worsens with 13-15 dB (TP2, TP3) at the ends of the studied variation range of U_{out} and improves with 4 dB for TP1 with $\Delta U_{out} \geq 6$ dB; for channel 52 MER worsens with 5-8 dB (TP3) in the ends of the studied variation range of U_{out}; C/N_{QAM} drops by around 8 dB (TP2, TP3) with -6 dB $> \Delta U_{out} > 6$ dB - for the two channels; BER increases in the ends of the investigated range of U_{out} up to 3×10^{-3}, but for values of -4dB $\leq \Delta U_{out} < 0$ dB (channel 52) and -2 dB $\leq \Delta U_{out} < 0$ dB (channel 40) BER decreases with around 0,02 and reaches up to $1,8 \times 10^{-9}$ (40-TP2).
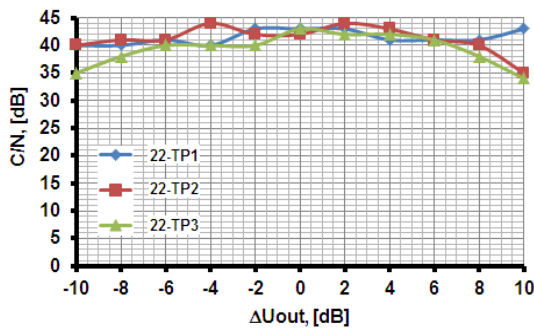
Fig.6. *C/N=func(ΔU_out) – AM/VSB*
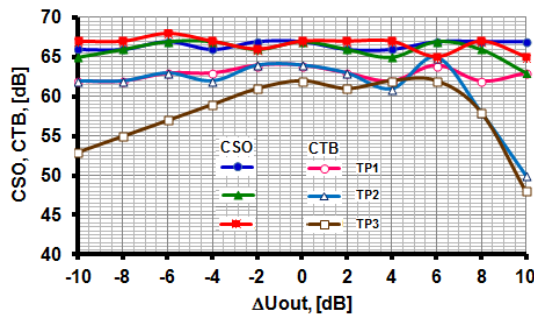
Fig.7. *CSO= func(ΔU_out) and CTB= func(ΔU_out)*
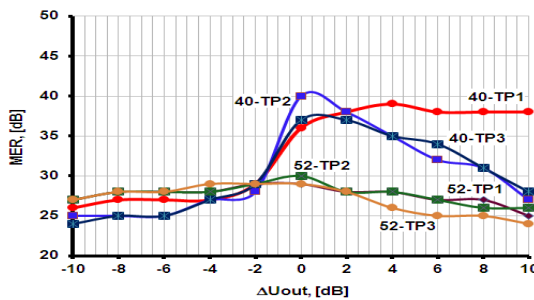
Fig.8. *MER= func(ΔU_out)*
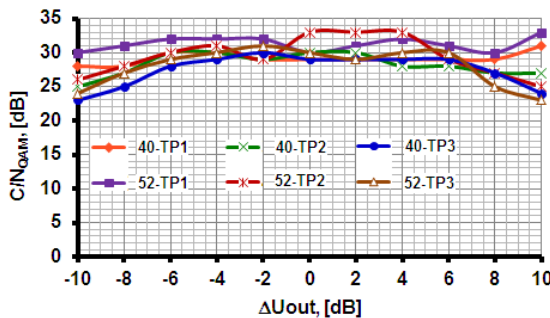
Fig.9. *C/N_QAM=func(ΔU_out)*
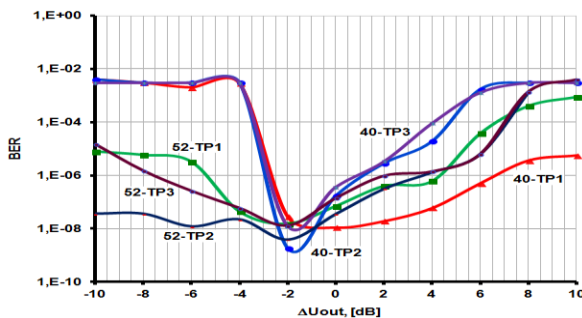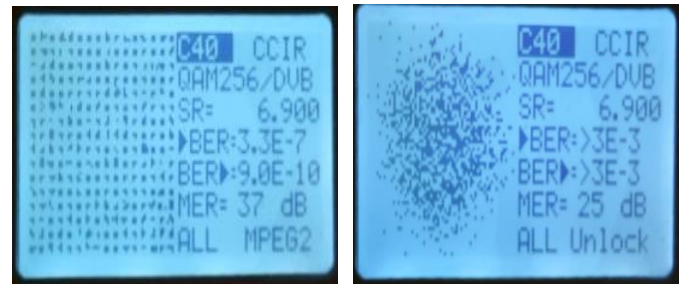

Fig.10. *BER= func(ΔU_out*

a) $\Delta U_{out} = 0dB\_(TP3)$ *b)* $\Delta U_{out} = \pm 10dB\_(TP3)$

Fig.11. Constellations: *a)* without and *b)* with breakdown

III. CONCLUSION

The research, the results and their analysis let us make the following conclusions: a variation with even so much as 2-4 dB of the nominal level in the cable distributing network of the HFC/CATV system can lead to a worsening in the quality of the transmitted signals (QoS) and even to prevent them from being (of M-QAM) received ($BER > 10^{-6}$). Furthermore, in the tests it was found that in the adjacent channels (39, 53) BER, MER, CSO, CTB and C / N do not deteriorate as much as for the channels 40 and 52, which is a result of the lack of influence of the DVB-T channels' (40, 52) broadcast in the area in which the study was conducted. Thus the recommendations, which can be given regarding the attenuation in the nominal level with ± 10 dB are as follow: to be used amplifiers with AGC on the long distances (>2 km); the maximal gain of every amplifier should be ca. 35-40 dB, but the working gain about 5-10 dB lower; all equipment and cables should be placed underground, be well isolated, as well as dampness and rodent proved et al. Taking the necessary measures would considerably improve the reliability and quality of the offered services, including those of the "triple play" – television, telephone and Internet [10].

REFERENCES

[1] L. Harte, *Introduction to Cable Television*, 3rd ed., Althos, 2010.
[2] K. Koitchev, K. Angelov and S. Sadinov, *Design of interactive cable television networks*, Ex-Press, Gabrovo, 2010.
[3] O. Panagiev, "Study of the parameters' influence of HFC communication network onto BER", NCIP "Telecom' 2009", 8-9 oct. 2009, Varna, pp 274-279.
[4] G. Morthier, P. Vankwikelberge, *Handbook of Distributed Feedback Laser Diodes*, 2nd ed., Artech house, 2013.
[5] http://electrical-engineeringportal.com
[6] http://www.moore.org.au
[7] O. Panagiev, "Algorithm for defining the limits possibilities of a system for cable television with the aim of minimal nonlinear distortions", ICEST, Proc. of Papers, vol.1, Bitola, 16-19 June 2004, pp. 217-219.
[8] European standard CENELEC "EN 50083-3", 2006.
[9] WISI, A link to the future. Catalogue'2004.
[10] V. Hristov, M. Kanchev, "Investigation of Triple Play Services", 3rd International Conference on Communications, Electro-magnetics and Medical Applications (CEMA'08) Athens, Greece, Nov. 06th-08th, 2008, pp.59-63.

AUTHOR INDEX