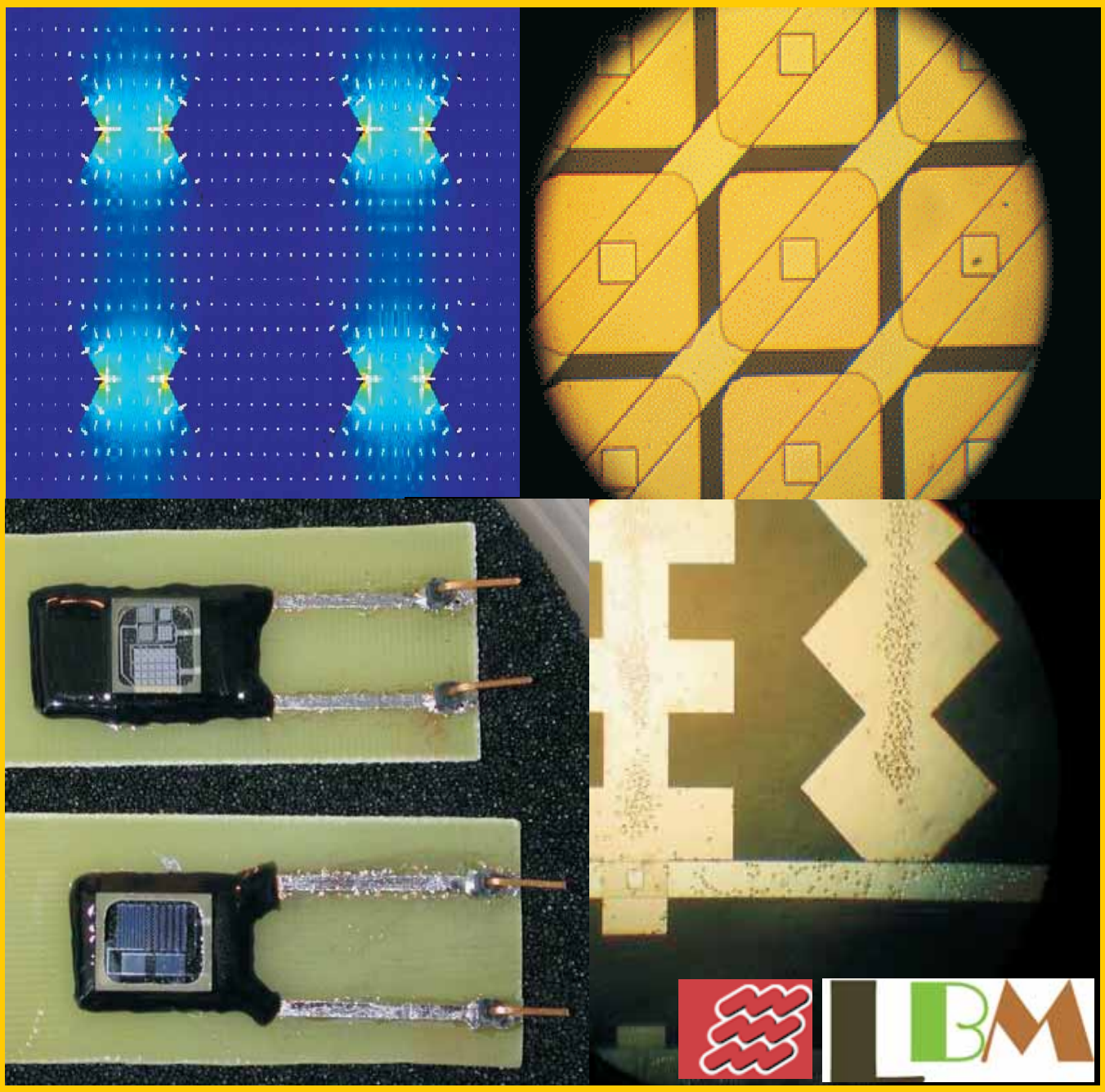
UDK 621.3:(53+54+621+66)(05)(497.1)=00 ISSN 0352-9045

INFORMACIJE MIDEM 2°2008

Strokovno društvo za mikroelektroniko elektronske sestavne dele in materiale

Strokovna revija za mikroelektroniko, elektronske sestavne dele in materiale
Journal of Microelectronics, Electronic Components and Materials

INFORMACIJE MIDEM, LETNIK 38, ŠT. 2(126), LJUBLJANA, junij 2008



UDK 621.3:(53+54+621+66)(05)(497.1)=00 ISSN 0352-9045

INFORMACIJE MIDEM 1 ◦ 2008

INFORMACIJE MIDEM LETNIK 38, ŠT. 1(125), LJUBLJANA, MAREC 2008
INFORMACIJE MIDEM VOLUME 38, NO. 1(125), LJUBLJANA, MARCH 2008

Revija izhaja trimesečno (marec, junij, september, december). Izdaja strokovno društvo za mikroelektroniko, elektronske sestavne dele in materiale - MIDEM.
Published quarterly (march, june, september, december) by Society for Microelectronics, Electronic Components and Materials - MIDEM.

Naslov uredništva
Headquarters
Uredništvo Informacije MIDEM
MIDEM pri MIKROIKS
Stegne 11, 1521 Ljubljana, Slovenija
tel.: + 386 (0)1 51 33 768
faks: + 386 (0)1 51 33 771
e-pošta: Iztok.Sorli@guest.arnes.si
http://www.midem-drustvo.si/

Letna naročnina je 100 EUR, cena posamezne številke pa 25 EUR. Člani in sponzorji MIDEM prejemajo Informacije MIDEM brezplačno.
Annual subscription rate is EUR 100, separate issue is EUR 25. MIDEM members and Society sponsors receive Informacije MIDEM for free.

Znanstveni svet za tehnične vede je podal pozitivno mnenje o reviji kot znanstveno-strokovni reviji za mikroelektroniko, elektronske sestavne dele in materiale. Izdajo revije sofinancirajo ARRS in sponzorji društva.
Scientific Council for Technical Sciences of Slovene Research Agency has recognized Informacije MIDEM as scientific Journal for microelectronics, electronic components and materials.
Publishing of the Journal is financed by Slovene Research Agency and by Society sponsors.

Znanstveno-strokovne prispevke objavljene v Informacijah MIDEM zajemamo v podatkovne baze COBISS in INSPEC.

Prispevke iz revije zajema ISI® v naslednje svoje produkte: Sci Search®, Research Alert® in Materials Science Citation Index™

Scientific and professional papers published in Informacije MIDEM are assessed into COBISS and INSPEC databases.

The Journal is indexed by ISI® for Sci Search®, Research Alert® and Material Science Citation Index™

Po mnenju Ministrstva za informiranje št.23/300-92 šteje glasilo Informacije MIDEM med proizvode informativnega značaja.

Grafična priprava in tisk
Printed by
BIRO M, Ljubljana

Naklada
Circulation
1000 izvodov
1000 issues

Poštnina plačana pri pošti 1102 Ljubljana
Slovenia Taxe Percue

UDK621.3:(53+54+621+66), ISSN0352-9045 Informacije MIDEM 38(2008)2, Ljubljana

Slika na naslovnici:
LBM, Laboratorij za bioelektromagnetiko, se ukvarja z razvojem in raziskavami vpliva električnega polja na biološke sisteme. Na naslovnici je prikazana numerična simulacija, načrtovanje ter izdelava polprevodniških mikrostruktur za proučevanje dielektroforezne sile na biološke celice.

Front page:
LBM, Laboratory for Bioelectromagnetics, studies influences of electrical fields on biological systems. Front page shows simulation, design and realization of semiconductor microstructures used in research of dielectroforesis force acting on biological cells.

VSEBINA **CONTENT**

Obnovitev članstva v strokovnem društvu MIDEM in iz tega izhajajoče ugodnosti in obveznosti

Spoštovani,

V svojem več desetletij dolgem obstoju in delovanju smo si prizadevali narediti društvo privlačno in koristno vsem članom.Z delovanjem društva ste se srečali tudi vi in se odločili, da se v društvo včlanite. Življenske poti, zaposlitev in strokovno zanimanje pa se z leti spreminjajo, najrazličnejši dogodki, izzivi in odločitve so vas morda usmerili v povsem druga področja in vaš interes za delovanje ali članstvo v društvu se je z leti močno spremenil, morda izginil. Morda pa vas aktivnosti društva kljub temu še vedno zanimajo, če ne drugače, kot spomin na prijetne čase, ki smo jih skupaj preživeli. Spremenili so se tudi naslovi in način komuniciranja.

Ker je seznam članstva postal dolg, očitno pa je, da mnogi nekdanji člani nimajo več interesa za sodelovanje v društvu, se je Izvršilni odbor društva odločil, da stanje članstva uredi in **vas zato prosi, da izpolnite in nam pošljete obrazec priložen na koncu revije.**

Naj vas ponovno spomnimo na ugodnosti, ki izhajajo iz vašega članstva. Kot član strokovnega društva prejemate revijo »Informacije MIDEM«, povabljeni ste na strokovne konference, kjer lahko predstavite svoje raziskovalne in razvojne dosežke ali srečate stare znance in nove, povabljene predavatelje s področja, ki vas zanima. O svojih dosežkih in problemih lahko poročate v strokovni reviji, ki ima ugleden IMPACT faktor.S svojimi predlogi lahko usmerjate delovanje društva.

Vaša obveza je plačilo članarine 25 EUR na leto. Članarino lahko plačate na transakcijski račun društva pri A-banki : 051008010631192. Pri nakazilu ne pozabite navesti svojega imena!

Upamo, da vas delovanje društva še vedno zanima in da boste članstvo obnovili. Žal pa bomo morali dosedanje člane, ki članstva ne boste obnovili do konca leta 2008, brisati iz seznama članstva.

Prijavnice pošljite na naslov:

MIDEM pri MIKROIKS

Stegne 11

1521 Ljubljana

Ljubljana, junij 2008

Izvršilni odbor društva

DEVELOPMENT OF OUTDOOR PHOTOVOLTAIC MODULE MONITORING SYSTEM

Jurij Kurnik, Marko Jankovec, Kristijan Brecl, Marko Topič

Faculty of Electrical Engineering, University of Ljubljana, Ljubljana, Slovenia

Key words: PV module, *I-V* curve measurement, outdoor monitoring

Abstract: Performance of photovoltaic (PV) modules is usually specified under standard test conditions (STC). But the performance of the modules under real field conditions can differ from the expectations derived from the results of STC tests due to variety of continuously different conditions. Therefore it is important to monitor PV modules' outdoor performance.

An automated computer-controlled monitoring system that is able to measure *I-V* curves of 16 PV modules in real-time has been developed and put in operation on the roof of Faculty of Electrical Engineering in Ljubljana. Beside the *I-V* curves the presented system also measures total and diffused solar irradiance, module temperature and other meteorological parameters. Selected measurement results are presented and discussed.

Razvoj sistema za merjenje fotonapetostnih modulov pri realnih vremenskih pogojih

Kjučne besede: PV modul, merjenje *I-U* karakteristik, meritve pri realnih vremenskih pogojih

Izvleček: Zmogljivost fotonapetostnih (photovoltaic - PV) modulov je običajno izmerjeno in deklarirano pri standardnih testnih pogojih (STC). Vendar so pogoji, pri katerih PV moduli običajno obratujejo, lahko zelo različni od pogojev pri STC. Zaradi tega lahko pride tudi do večjih razlik med pričakovanim energijskim donosom PV modulov glede na deklarirane parametre modulov ter dejanskim pridobljenim energijskim donosom. Zaradi tega je za natančnejše predvidevanje delovanja modulov potrebno meriti njihovo delovanje pri realnih zunanjih pogojih obratovanja.

V članku je prikazana zasnova in razvoj sistema za merjenje PV modulov pri realnih vremenskih pogojih, ki je postavljen na strehi Fakultete za elektrotehniko v Ljubljani. Razvit sistem omogoča avtomatično meritev *I-U* krivulj 16 PV modulov. Poleg *I-U* krivulj se sočasno merijo tudi globalno in difuzno sončno sevanje, temperature modulov, zraka ter ostali meteorološki parametri. Predstavljeni in ovrednoteni so značilni merilni rezultati.

1. Introduction

Performance of photovoltaic (PV) modules is determined by conversion efficiency of solar energy into electrical energy. The most important PV modules' performance parameters are STC efficiency (η_{STC}) and effective efficiency (η_{EFF}). The η_{STC} is determined as the maximal output power (P_{MPP_STC}) under standard test conditions (STC - solar irradiance 1000 W/m^2, AM 1.5 spectrum and module temperature 25 °C, wind speed 1 m/s) normalized to incident solar irradiance (G) of 1000 W/m^2 /1/ and η_{EFF} is calculated as a ratio of total available annual energy generated by a PV module (E_M), divided by annual solar energy the module receives (E_{SOL}):

$$\eta_{STC} = \frac{P_{MPP\_STC}}{1000\ W/m^2} \tag{1}$$

$$\eta_{EFF} = \frac{E_M}{E_{SOL}} = \frac{\int_{year} \eta\,(G, T_{PV}) \cdot G \cdot dt}{\int_{year} G dt} \tag{2}$$

Several procedures have been developed /2, 3, 4/ to asses PV modules' annual performance (effective conversion efficiency or energy yield) at a given location and installation by means of combining location specific data (irradiance, solar energy and air temperature) and PV module installation specific data (installation azimuth, inclination and mounting method) with PV module specific data (Figure 1).

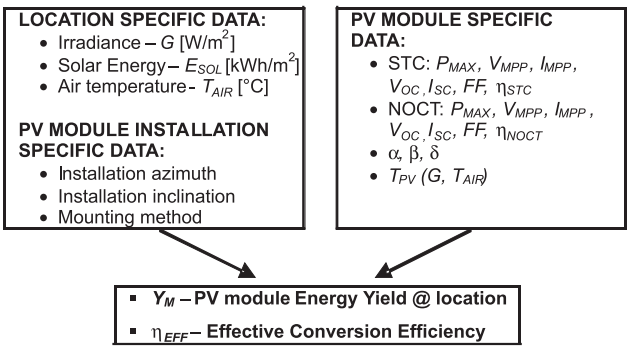


Fig. 1: *Input data needed to assess the PV module annual performance at specific location.*

The location specific data are normally measured by the local meteorological institute, whereas to obtain PV module specific data different approaches are possible. In the simplest approach one can use standard manufacturers' PV module parameters measured at the STC. Some PV module manufacturers also give module parameters measured at nominal operating conditions (NOCT - solar irradi-

ance 800 W/m^2, AM 1.5 spectrum and 20 °C air temperature) /1/. Since performance of PV modules varies with solar irradiance, module temperature, solar spectrum, incident angle, etc., predicting PV module annual performance on the basis of module parameters measured at STC or NOCT can not guarantee accurate results. To be able to give a better daily or annual PV module performance prediction, its parameters have to be defined by outdoor monitoring at different times of the year, different weather conditions and at different mounting options (open rack mounting, building integrated).

For this purpose we have developed an outdoor PV monitoring system (Figure 2). The system has been in operation since 1st of January 2007. At the moment we are monitoring PV modules of the following technologies: poly-Si, mono-Si, back contact mono-Si, a-Si, CIGS, HIT, flexible single junction a-Si and flexible triple junction a-Si. The modules are either mounted on an open rack or integrated on a metal roof. The monitoring site is located on the roof of the Faculty of Electrical Engineering in Ljubljana (Slovenia), oriented south with an inclination angle of 30° (optimal for Ljubljana). To monitor integration of PV modules on building walls one module is also mounted on a metal wall oriented south and 90° inclined.

Fig. 2: *Outdoor PV module monitoring site located on the roof of the Faculty of Electrical Engineering in Ljubljana.*

First results acquired by the outdoor PV monitoring system have been reported elsewhere /5, 6/. We have found relatively large difference between outdoor monitoring results and the modules' performance at STC or NOCT. This confirms the need of continuous outdoor measuring of PV modules' at different conditions to accurately predict their long term performance. Similar conclusions have also been drawn by other authors /7, 8, 9, 10, 11, 12/.

2. Development of the outdoor PV monitoring system

The most important requirement of an outdoor PV monitoring system is that it is able to simultaneously measure several PV modules of different technologies with wide range of nominal powers. If new modules are added to the system it should be easily expandable. Each module should be completely characterized by *I-V* curve scanning from short to open-circuit conditions and its temperature needs to be measured as well. The *I-V* scan should be as fast as possible to assure stable weather conditions during the scan since sample and hold technique can not be applied. In addition, meteorological data (solar irradiances, air temperature, wind speed, etc.) need to be monitored. To minimize the measurement error due to long connection cables, the monitoring system should be located as close as possible to the monitored PV modules. This usually means that measurement instruments have to be located outdoors, which can lower measurement accuracy due to temperature and humidity variations. If a personal computer (PC), which controls the system, is situated indoors a long and reliable communication bus has to be implemented.

On the basis of those requirements we designed and built a microcontroller based outdoor PV monitoring system that is comprised of a PV measurement unit, a module switching unit, different irradiance and several temperature sensors and a commercially available weather station (Figure 3). PV measurement and switching units are controlled by a PC via RS-485 bus. Another RS-485 bus is used to connect remote weather sensors to the weather station display, which is connected to the PC with a RS-232 bus. A separated switching unit allows us to easily expand the system capacity by simply adding additional switching units, which can be connected to the same RS-485 bus. The monitoring system measures global and diffused irradiances on a horizontal surface, air temperature, wind speed, rain-fall, air pressure and humidity at chosen time intervals (every two minutes). The measurement time interval for *I-V* curves, PV modules' temperatures and irradiance in the plane of PV array (G_{POA}) is ten minutes. Entire monitoring system is controlled by monitoring software, developed in LabVIEW programming language, which runs on a PC. All the measured data is stored in an SQL database.

2.1 PV measurement unit

The main part of the whole PV monitoring system is the PV measurement unit (Figure 4), which enables 4-wire connection of the measured PV module /13/. The electronic load is a power MOS-FET controlled by a 14 bit D/A converter. Current of the module is measured as a voltage drop on a precise shunt resistor with a 24 bit A/D converter, which also acquires the module voltage, scaled by a precision voltage divider. The input current and voltage limits are 12 A and 100 V, respectively. The limits are sufficient to measure all commercially available PV modules on the market at the moment. The unit also has six inputs for different types of irradiance sensors. Four of them are voltage inputs (two with 24 mV and two with 120 mV input limits), while the rest are current inputs with 24 mA maximal current input. Specifications of the PV measurement unit are shown in Table 1. In addition to the current and voltage inputs, the measurement unit also features one-wire digital communication bus for connection of digital temperature sensors. The unit can be powered from the grid, in that situation it is connected to the measurement compu-

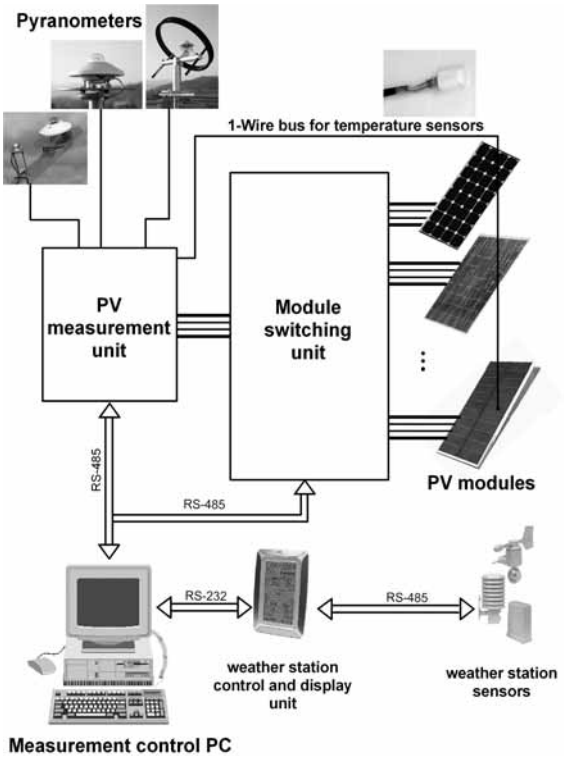


Fig. 3: *Block diagram of the PV monitoring system.*

ter via a RS-485 bus (as it is done at our monitoring system), or via USB connection, when using it as a portable PV module *I-V* curve measurement device.

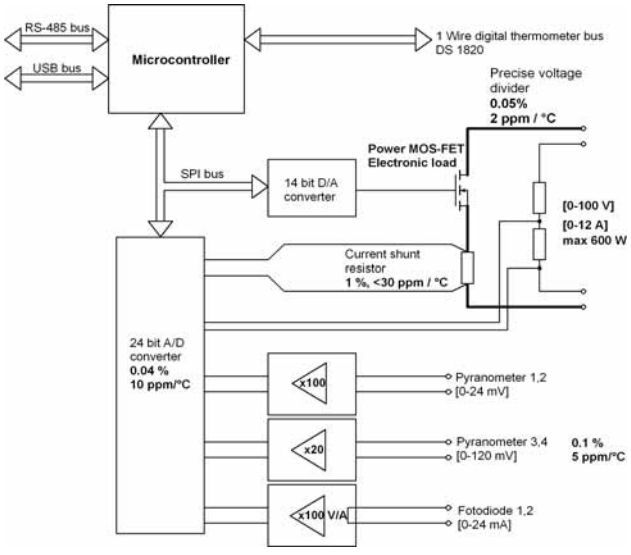


Fig. 4: *Block diagram of the PV measurement unit.*

2.2 Module switching unit

The purpose of the module switching unit in the designed PV measurement system is to enable arbitrary 4-wire connection of multiple PV modules to a single *I-V* measurement unit. In our case, the switching unit allows connecting up to 16 PV modules to the measurement unit. Similar to the PV measurement unit, the module switching unit can also be powered from the grid or from the USB bus power

Table 1: *Specifications of the developed PV measurement unit.*

Input Channel	Input range	Max. offset [0-60°C]	Max. error [0-60°C]
PV current input	0 ÷ 12 A	±0.5 mA	0.25%
PV voltage input	0 ÷ 100 V	±0.75 mV	0.1%
Pyranometer voltage input	0 ÷ 24 mV	±15 µV	0.25%
Pyranometer voltage input	0 ÷ 100 mV	±15 µV	0.3%
Reference cell current input	0 ÷ 24 mA	±0.15 µA	0.1%

supply (5 V) in case of using RS-485 or USB connection, respectively. Block diagram of the four point module switching unit is shown in Figure 5.

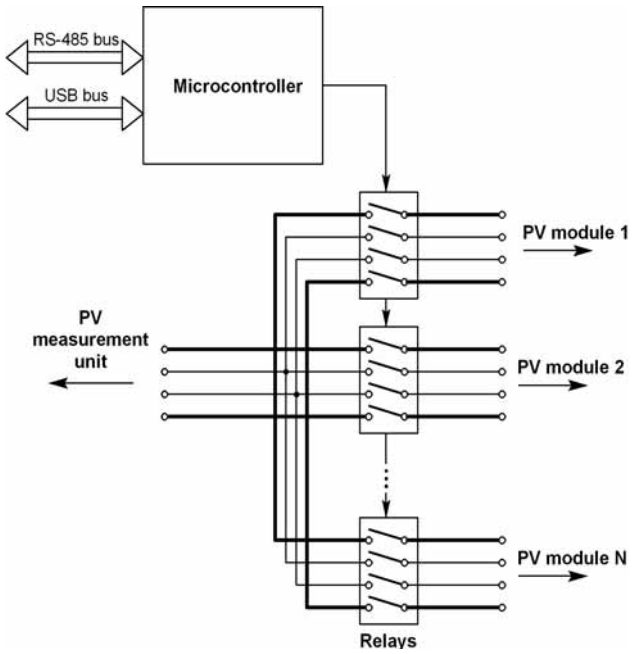


Fig. 5: *Block diagram of the 4-point module switching unit.*

2.3 Sensors

As local specific data we measure instant total solar irradiance in the plane of array, total and diffused irradiance on flat surface, module and air temperature, wind speed and direction, relative humidity, air pressure and rain-fall.

The total solar irradiance at the plane of array (G_{POA}) is measured with two different types of irradiance sensors. One is a thermopile based pyranometer (*CMP 6*) /14/ and the other one is a pyranometer with a photodiode (*SP-Lite*)

/14/. The first one is more accurate but relatively slow and the second one has a fast response time, but it is less accurate. However, the measurements show that sudden irradiance changes do not contribute to the irradiation, so the difference in dynamic properties is not important. It is their spectral range that differs. While the *CMP 6* covers a broad spectral range, the *SP-Lite* is spectrally adopted to fit crystalline-Si PV modules. Such a combination allows us to accurately evaluate not only crystalline - Si PV module, but also PV modules from other technologies. Total and diffused irradiance on a horizontal surface is measured with two *CMP 6* pyranometers. For the diffused irradiance measurement a shadow ring /14/ is used in combination with the *CMP 6*.

We also monitor temperatures of each module *(T_{PV})* by digital temperature sensors (*DS18B20*) /15/ glued on the back sheet of the PV modules. The temperature sensors are thermally connected to the back sheet of the module using a thermal conducting paste and shielded from the ambient temperature influence on the measurement of module temperature by polystyrene and silicon sealant.

For monitoring air temperature, wind speed and direction, relative humidity, air pressure and rain-fall a weather station WS3600 is used /16/.

Specifications for the pyranometers and the digital temperature sensors are shown in Table 2.

Table 2: *Specifications of the pyranometers and digital temperature sensors.*

Pyranometer with thermopile (Kipp&Zonen *CMP 6*)	
Measurement range	0 ÷ 2000 W/m^2
Spectral range	310 ÷ 2800 nm
Response time	18 s
Temperature range	-40 ÷ +80 °C
Directional error (at 80°)	2%
Pyranometer with photodiode (Kipp&Zonen *SP-Lite*)	
Measurement range	0 ÷ 2000 W/m^2
Spectral range	400 ÷ 1100 nm
Response time	< 1 s
Temperature range	-30 ÷ +70 °C
Directional error (at 80°)	5%
Digital temperature sensor (*DS18B20*)	
Measurement range	-55 ÷ +125 °C
Conversion time	0.75 s
Accuracy [-10 ÷ +85 °C]	±0.5 °C
Resolution	1/16 °C

2.4 Measurement control

The PV monitoring system is controlled by a LabVIEW based software, running on a standard PC, located indoors. The measurement sequence, which is shown in Figure 6, is executed every ten minutes. First the air temperature, wind speed, air pressure, humidity and PV modules' temperatures are measured. Then the solar irradiances are acquired. After that the system switches the first PV module to the PV measurement unit, which acquires its *I-V* curve. Then the system disconnects the PV module and calculates parameters of the measured PV module, which are together with the whole *I-V* curve and weather parameters stored to an SQL database. System switches to the next PV module and restarts the procedure with solar irradiance measurement until all PV modules are measured. The whole measurement procedure for 16 modules takes less than a minute. Between the measurement sequences the modules are kept in open-circuit conditions.

The user can compare G_{POA} irradiances measured by *SP-Lite* sensor prior and after *I-V* scanning and see whether the solar irradiance changed during the *I-V* scan. If the difference is too high, indicating unstable weather conditions, the *I-V* curve may be regarded as uncertain. The system also enables the user to manually select and measure individual PV modules at any time.

2.5 I-V curve measurement

The procedure of the PV module *I-V* curve acquisition starts with the measurements of the short and open circuit values of the curve. Using this information the optimal increment of the electronic load is set to measure the *I-V* curve with equidistant steps from shortcircuit to opencircuit conditions. Since the MOS-FET is highly non-linear, the step is adjusted from point to point during the *I-V* scan. Due to capacitive effects of the PV modules, the accuracy of the measured *I-V* curve strongly depends on the speed of the scan. *I-V* curves of the Sanyo HIT PV module /17/, measured at various acquisition times (t_{AP}) of each *I-V* point, are shown in Figure 7. From the study of different modules' behavior we conclude that t_{AP} should be at least 1 ms.

The total number of points per one *I-V* scan is adjustable up to 350 points, but we found a trade-off between total scan time and accuracy with around 70 points per *I-V* scan. To accurately determine the module's power at the maximal power point (P_{MPP}), the voltage (V_{MPP}) and the current (I_{MPP}) at the maximal power point, a fourth order polynomial interpolation is used. Since the electronic load is passive, the ideal opencircuit and shortcircuit circuit conditions cannot be achieved. Therefore, linear extrapolation is used to determine module's shortcircuit current (I_{SC}), opencircuit voltage (V_{OC}) as also the series and shunt resistances. From that parameters module's fill-factor (*FF*) and efficiency (η) is calculated by equations 3 and 4, respectively. G_{POA} represents measured total irradiance in plane of the measured PV module and *A* module area.

$$FF = \frac{I_{MPP} \cdot V_{MPP}}{I_{SC} \cdot V_{OC}} = \frac{P_{MPP}}{I_{SC} \cdot V_{OC}} \quad (3)$$

$$\eta = \frac{P_{MPP}}{G_{POA} \cdot A} \quad (4)$$

A screen shoot of the main measurement screen with a measured selected PV module's *I-V* curve, calculated *PV*

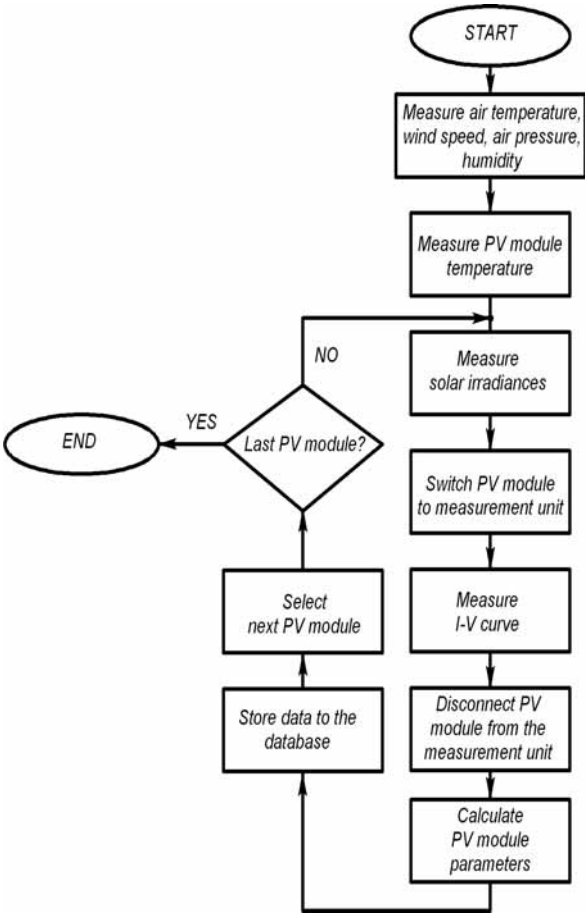


Fig. 6: *I-V measurement sequence block diagram.*

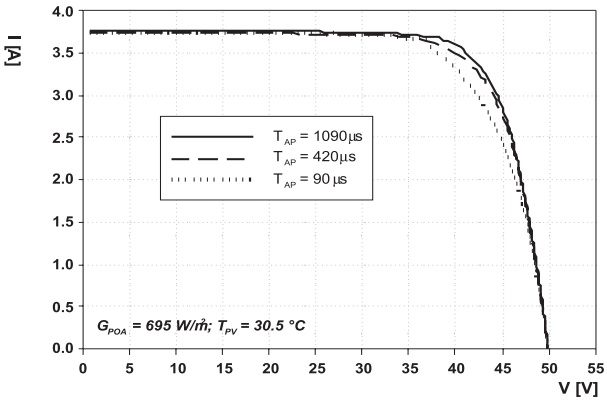


Fig. 7: *Measured I-V curves of the Sanyo HIT PV module at different t_{AP} (t_{AP} = 1090 µs – solid line, t_{AP} = 420 µs – dashed line and t_{AP} = 90 µs – dotted line).*

curve, irradiance and temperature data and from the *I-V* curve calculated PV module parameters is shown in Figure 8.

3. Measurement results

At our monitoring site we constantly monitor 16 different PV modules. Measured *I-V* curves of the Sanyo HIT PV module are shown as an example in Figure 9. The curves were measured in the afternoon of the 30th of March 2008

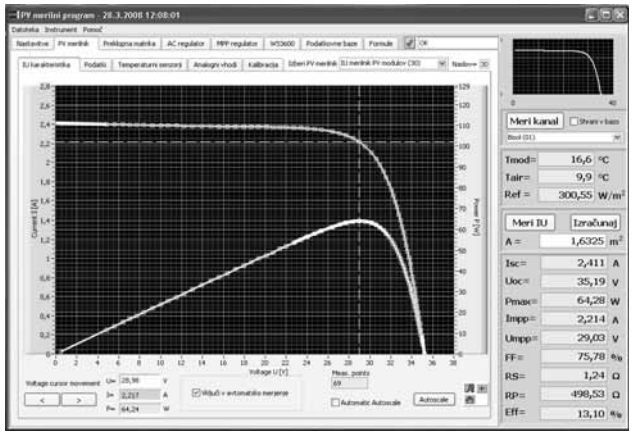


Fig. 8: *Screen shoot of the measurement software developed in program language LabVIEW.*

from noon till five o'clock under clear sky conditions with solar irradiance values ranging from 1015 W/m^2 to 265 W/m^2 and module temperature values ranging from 48 °C to 21 °C during the measurement period (see current in Figure 9).

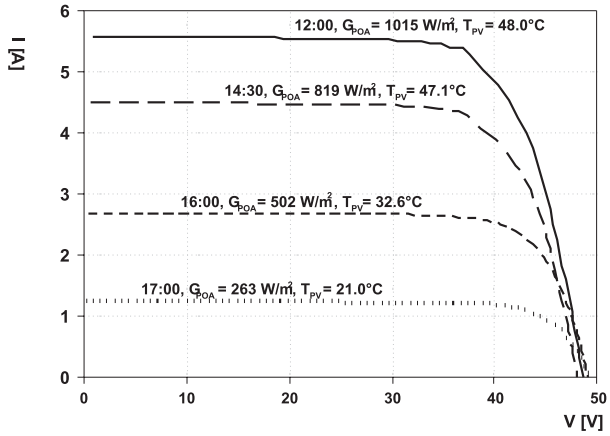


Fig. 9: *Measured I-V curves of a Sanyo HIT PV module on the 30th of March 2008.*

Measurements of solar irradiance on horizontal surface (Figure 10) for the selected day show cloudy and foggy conditions in the morning. That is noticed by almost the same values of total and diffused irradiance until 10:30. Afterwards the fog and clouds cleared out and it became a nice sunny afternoon, depicted by a large difference between the total and the diffused irradiance and a smooth curve of measured solar irradiances in the afternoon hours. Together with *I-V* curves and irradiances we measure module and air temperatures and wind speed. The measured parameters are shown in Figure 11. The 30th of March 2008 was a pretty calm day with wind speeds under 2 m/s and air temperatures up to 18 °C. Because of clear sky conditions in the afternoon, with G_{POA} up to 1000 W/m^2 and low wind speeds the HIT module heated up to almost 50 °C. In hot and clear-sky summer days module temperatures of up to 75 °C have been measured.

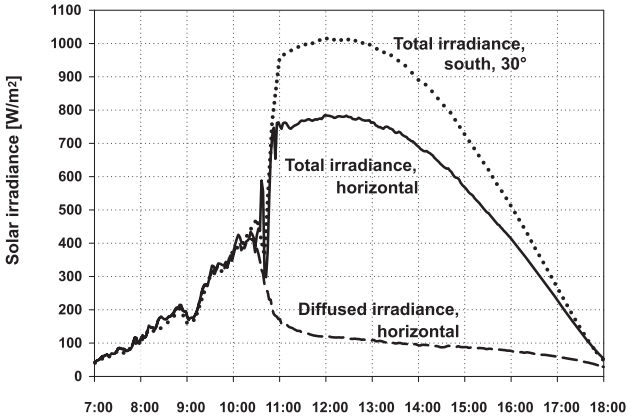


Fig. 10: *Total (solid line) and diffused solar irradiance (dashed line) on flat surface and total solar irradiance on plane of array (dotted line) on the 30th of March 2008.*

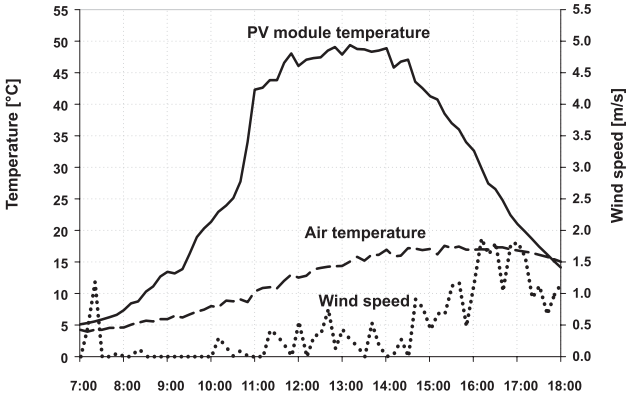


Fig. 11: *HIT PV module temperature (solid line), air temperature (dashed line) and wind speed (dotted line) on the 30th of March 2008.*

4. Conclusion

An outdoor PV monitoring system has been successfully developed and put into operation on the roof of Faculty of Electrical Engineering in Ljubljana. The monitoring system offers real-time measurement of PV modules' *I-V* curves together with their temperatures. Additionally we monitor the total and diffused irradiances on the horizontal plane and the total irradiance in the inclined plane of the PV modules. Meteorological parameters are monitored as well enabling us to accurately determine different short and long term module parameters. Longterm monitoring results are the basis for building and validating mathematical models of PV modules of different technologies and also the basis for study of PV modules' longterm stability.

5. Rreferences

/1/ Standard IEC 60904-3, Measurements Principles for Terrestrial PV Solar Devices with Reference Irradiance Data, International Electrotechnical Commission - IEC, Geneva, Switzerland

/2/ S.R. Williams, M. Strobel, T.R. Betts, R. Gottschalg, D.G. Infield, W. Kolodenny, M. Prorok, T. Zdanowicz, N. van der Borg, H. de Moor, G. Friesen, A. Guerin de Montgareuil, *"Accuracy of European energy modeling approaches"*, Proceedings of 21-EU-PVSEC, 2006, pp. 24522455.

/3/ M. Topič, K. Brecl, J. Sites, *"Effective efficiency of PV modules under field conditions"*, Progress in Photovoltaic: Research and Applications, 2007, Vol. 15, pp. 19-26.

/4/ M. Topič, K. Brecl, J. Kurnik, J. Sites, *"Effective efficiency and performance ratio as energy rating system for PV modules"*, Proceedings of 21-EU-PVSEC, 2006, pp. 25072510.

/5/ J. Kurnik, K. Brecl, M. Jankovec, M. Topič, *"Comparison of fixed, 1axis and 2axis tracking PV system performance"*, Proceedings of 43rd International Conference on Microelectronics, Devices and Materials - MIDEM, 2007, pp. 101104.

/6/ J. Kurnik, K. Brecl, M. Jankovec, M. Topič, *"First measurement results of effective efficiency of different PV modules under field conditions"*, Proceedings of 22EUPVSEC, 2007, pp. 27312734.

/7/ T. Zdanowicz, T. Rodziewicz, M. Zabkowska-Waclawek, *"Evaluation of actual PV modules performance at low insolation conditions"*, Opto-Electronics Review, Vol. 8, 2001, pp. 361366.

/8/ E. Bura, N. Cereghetti, D. Chianese, A. Realini, S. Rezzonico, *"PV Module Behaviour in Real Conditions: Emphasis on Thin Film Modules"*, Proceedings of 17EUPVSEC, 2001, pp. 714-717.

/9/ R. Gottschalg, T.R. Betts, S.R. Williams, D. Sauter, D.G. Infield, M.J. Kearney, "*A critical appraisal of the factors affecting energy production from amorphous silicon photovoltaic arrays in a maritime climate"*, Solar Energy, Vol. 77, 2004, pp. 909916.

/10/ R.P. Kenny, A. Ioannides, H. Mullejans, W. Zaaiman, E.D. Dunlop, *"Performance of thin film PV modules"*, Thin Solid Films, Vol. 511-512, 2006, pp. 663-672.

/11/ B. Zinßer, G. Makrides, W. Schmitt, G. E. Georghiou, J. H. Werner, *"Annual energy yield of 13 photovoltaic technologies in Germany and in Cyprus"*, Proceedings of 22EUPVSEC, 2007, pp. 31143117.

/12/ E. Rustu, O. Sener, *"Comparison of 18-month kWh/kWp energy output of four photovoltaic systems with four different module technologies"*, Proceedings of 22EUPVSEC, 2007, pp. 31143117.

/13/ J. Krč, M. Jankovec, M. Topič, *"Electronics on the way from a detector to the system unit"*, Informacije MIDEM, Vol. 32 , 2002, pp. 298302.

/14/ *http://www.kippzonen.com*

/15/ *http://datasheets.maxim-ic.com/en/ds/DS18B20.pdf*

/16/ *http://www.heavyweather.info/new_english_us/index.html*

/17/ *http://www.sanyocomponent.com/fileadmin/mc/products/photovoltaics2/ datasheets/ NHE5/HIP_215_210_205NHE5_e.pdf*

Jurij Kurnik, univ. dipl. ing. el.
Dr. Marko Jankovec, univ. dipl. ing. el.
Dr. Kristijan Brecl, univ. dipl. ing. el.
Prof. Dr. Marko Topič, univ. dipl. ing. el.

University of Ljubljana,
Faculty of Electrical Engineering,
Laboratory of Photovoltaics and Optoelectronics,
Tržaška cesta 25, SI-1000 Ljubljana, Slovenia

E-mail: jurij.kurnik@fe.uni-lj.si

Prispelo (Arrived): 24.04.07 *Sprejeto (Accepted):* 28.5.08

CANTILEVER METHOD FOR DETERMINATION OF d_{31} COEFFICIENT IN THIN PIEZOELECTRIC FILMS

Samo Penič, Uroš Aljančič, Drago Resnik, Danilo Vrtačnik, Matej Možek, Slavko Amon

Laboratory of Microsensor Structures and Electronics (LMSE), Faculty of Electrical Engineering, University of Ljubljana, Ljubljana, Slovenia

Key words: piezoelectric, characterization, d_{31} coefficient, PZT, simulation, FEM, ANSYS

Abstract: A cantilever method for characterization of thin piezoelectric films is proposed. Using the proposed cantilever method, piezoelectric coefficient d_{31} of thin film piezoelectric material on various samples was determined. Cantilever based characterization method provides a fast comparison of different piezoelectric material samples, since multiple samples can be mounted simultaneously on the testing structure. It is shown how, when combined with numerical simulation, piezoelectric coefficient d_{31} can be determined from fitting measured voltage response with simulated response.

Exact knowledge of geometry and material properties of cantilever and samples proved to be important in order to determine piezoelectric coefficients with sufficient accuracy. Stainless steel cantilever was adequately characterized by measuring its Young's modulus. Silicon properties are adequately determined by published data. Mechanical properties of PZT layers are on the other hand more difficult to acquire, since they are rather dependent on the actual PZT preparation procedure and composition. Nevertheless, we expect that error here introduced is small due to very thin PZT layer compared to stainless steel cantilever and silicon substrate. To improve the proposed method, based on numerical simulation results, guard chips were mounted at the side of the cantilever to reduce stress variation over samples.

Determined values of piezoelectric coefficients d_{31} for PZT layers under test were in reasonable agreement with results available in the literature.

Metoda za določanje koeficienta d_{31} tankih piezoelektričnih filmov

Kjučne besede: piezoelektrik, karakterizacija, d_{31} koeficient, PZT, simulacija, silicij, FEM, ANSYS

Izvleček: V članku je predstavljena metoda za karakterizacijo tankih piezoelektričnih plasti. Z uporabo ročice smo določili piezoelektrični koeficient d_{31} tankih piezoelektričnih filmov. Metoda omogoča hitro primerjavo lastnosti različnih materialov, ter s pomočjo numerične simulacije hkratno karakterizacijo parametra d_{31} večih vzorcev.

Poznavanje geometrije in materialnih lastnosti ročice in vzorjev je ključno za natančno določitev piezoelektričnih koeficientov. Mehanske lastnosti jeklene ročice smo določili z meritvijo Youngovega modula, za mehanske lastnosti silicijevega substrata pa smo uporabili podatke v literaturi. Mehanske lastnosti tankih PZT plasti so težje dostopne, saj se razlikujejo zaradi same zgradbe PZT keramike ter njene priprave. Zaradi tanke plasti PZT materiala, ocenjujemo, da je napaka pri uporabi vrednosti za debele materiale zanemarljiva. Na osnovi simulacij smo predstavljeno metodo izboljšali z dodatnimi stranskimi čipi, ki izboljšajo homogenost stresa na vzorcih.

Vrednosti za piezoelektrični koeficient d_{31}, ki smo jih določili s predlagano metodo, se ujemajo s podatki iz literature.

1. Introduction

When designing a new product or device, proper material selection is of basic importance. Material properties are also used in numerical analysis, when predicting device behavior. In case of piezoelectric microstructures, the properties of thin film piezoelectrics are influenced by chemical composition and other parameters of piezoelectric manufacturing process. It is thus important to have means for analyzing specific samples of piezoelectric thin films.

Due to unique properties of piezoelectric effect, piezoelectrics are important materials in micro-electromechanical system (MEMS) technology, used for actuation or sensing, energy harvesting etc. Characteristics of piezoelectrics, especially piezoelectric coefficients d, play important role in device design, simulation and behavior prediction. In general, thin film materials used in microengineering behave differently than bulk, thus requiring an adequate characterization of their core properties. Several methods are in use /1/ and new ones are being developed. Selection of the appropriate material for a certain application requires comparison of different materials using datasheet specifying core information about these materials.

Properties of piezoelectric materials vary with chemical composition, preparation technique e.g. sintering temperature and other effects. These influences present difficulties for comparison of different materials prepared by different methods, of different thicknesses and possibly from different producers, usually taking plain datasheet information from catalogue as a starting point.

To overcome this obstacle a comparative method for characterization of different thin film piezoelectric samples bonded to a stainless steel cantilever is proposed. The relative

response of different piezoelectric samples to the same mechanical stress gives immediate comparison of their basic properties such as sensitivity and linearity. Furthermore, coupling the measured results with numerical simulation based on finite element method (FEM) enables determination of absolute value for piezoelectric coefficient d_{31}.

The paper presents in detail the proposed technique for thin film piezoelectrics characterization and introduces a comparative method for simultaneous evaluation of multiple piezoelectric samples based on numerical simulation in combination with measured results. The result of this characterization is the absolute value of d_{31} coefficient for multiple samples and comparison of piezoelectric response to mechanical stimulus. The method is practically tested on different thin film Lead Zirconate Titanate (PZT) chip samples prepared on silicon substrates. Measured results are matched with numerical simulation and piezoelectric coefficients are determined using ANSYS finite element analysis software.

2. Basic properties of piezoelectrics

Piezoelectrics are materials that respond to the applied mechanical stress with electric voltage on the electrodes. This is called the direct piezoelectric effect, which serves as a basis for sensors and generators. The effect can be reversed and it is then called converse or inverse piezoelectric effect. Here mechanical strain is induced when voltage is applied. The response is dependent on the polarity of applied voltage and can therefore vary between elongation and contraction.

Equations that describe electromechanical relations in a piezoelectric material are given in Voight notation with relations /2/

$$\begin{aligned} \{T\} &= [c]\ \{S\} - [e]\{E\} \\ \{D\} &= [e]^{\mathrm{T}}\ \{S\} - [\check{n}]\{E\} \end{aligned} \qquad (1)$$

Where *{T}* is stress tensor, *{S}* strain tensor, *{E}* electric field vector and *{D}* electric displacement vector. Material properties are described with stiffness matrix [c] which includes information about Young's modulus Y and Poisson ratio σ of the material, with piezoelectric stress matrix [e] (superscript T denotes matrix transpose) related to piezoelectric strain matrix [d] and with permittivity matrix [ň]. Piezoelectric strain coefficients d_{ij} and piezoelectric stress coefficients e_{ij} are related with stiffness coefficients c_{ij} by matrix equation [e] = [c] [d].

Piezoelectrics can be used for sensing or actuation, depending on whether the applied input load is mechanical or electrical, respectively. The two modes of operation can also be used interchangeably which makes piezoelectrics extremely versatile electromechanical materials since the same structure can act as a sensor or an actuator. Though the effect is reversible, certain considerations must be taken into account during the design of the structure /3/.

3. Piezoelectrics characterization

3.1 Bulk piezoelectrics characterization

A complete characterization process of bulk piezoelectric material includes determination of stiffness coefficients c_{ij} (including Young's modulus *Y* and Poisson ratio σ), permittivity ($\check{n}_{ii}$) and piezoelectric coefficients (d_{ij}).

Most widely used method adopted as IEEE standard for piezoelectric characterization is the resonance method /4/. For such characterization, piezoelectric material is prepared as a flat rectangular plate between two electrodes, forming a capacitor. The capacitor impedance Z is measured at different frequencies. From *Z(f)* diagram, the resonant (f_r) and anti-resonant (f_a) frequencies are found. Then, the elastic compliance (inverse stiffness matrix) and piezoelectric coefficients for practical purposes usually d_{31} and d_{33} can be derived /1/.

Direct methods for determining piezoelectric coefficients d_{ij} include deformation measurements when voltage is applied to the electrodes. These methods are used to quantify the direct and converse piezoelectric effect. Direct methods are also used to investigate the behavior of the piezoelectric material in terms of hysteresis and nonlinearity, thermal behavior and aging. Mechanical deformation measurement of piezoelectric sample vs. applied voltage is used to determine piezoelectric coefficients d_{ij}, calculated from relation in Voight notation $S_i = d_{ij} E_i$ /1/.

A different method for measuring piezoelectric coefficients d_{ij} is based on direct piezoelectric effect. Here, sample is mechanically loaded, therefore the bounded electric charge becomes free, ready to flow out from the electrodes /5/. Electrodes are short circuited and electric displacement *D* is measured. Piezoelectric coefficient d_{ij} is here calculated from equation in Voight notation $D_i = d_{ij} T_j$ /1/.

In order to determine the relative permittivity $\check{n}_r$, capacitance measurements are carried out at low frequency, usually 1 kHz and for low AC voltage excitation levels, ranging few mV /1/. The relative dielectric constant is then calculated as

$$\check{n}_r = \frac{Ct}{\check{n}_0 A} \qquad (2)$$

where *t* is thickness of piezoelectric layer, *A* electrode area, *C* measured capacitance and $\check{n}_o$ permittivity of free space.

3.2 Thin film piezoelectrics characterization

In general, the properties of thin film materials can differ significantly from its bulk counterparts. Therefore, adequate characterization of piezoelectric thin film properties is essential. Thin film characterization methods are usually based on similar principles as for bulk. The prevailing methods use converse piezoelectric effect where electrically excited thin piezoelectric film results in mechanical displace-

ment, which is typically in the order of a few angstroms /6/. Sometimes the direct piezoelectric effect is used. Thin film piezoelectric together with electrodes are deposited on a substrate wafer and fixed in a rigid frame above pneumatic pressure cavity /7/. Pressure in the cavity is varied thus applying different mechanical stress to the piezoelectric layer. The charge integrator is used to measure the induced charge which is used in combination with excitation pressure to determine piezoelectric coefficients d_{ij}.

For determining Young's modulus of thin piezoelectric films, several approaches exist. One of the possibilities to characterize mechanical thin film properties is presented in /8/. The experiment consists of loading a membrane with a line load applied to the middle of the span using nanoindenter. A Mireau microscope interferometer is used to observe fringes that are formed on the loaded sample. Using a CCD camera these fringes are recorded and strains determined. From known stresses and strains in the material, Young's modulus can be determined.

3.3 Cantilever method for characterization of thin piezoelectric films

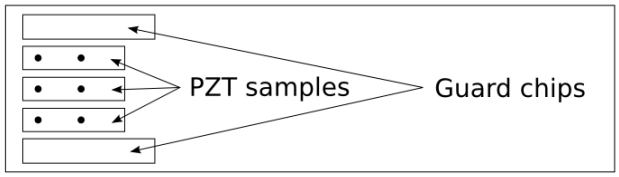
In this case, characterization method is focused on piezoelectric coefficient d_{31} using direct piezoelectric effect.

In the proposed characterization method we introduce a cantilever with mounted piezoelectric samples on silicon substrate, with exact control of deflection. Mounting several samples simultaneously to the same cantilever provides us a comparison of piezoelectric responses of various piezoelectric materials to the same stimulus. This provides fast and accurate comparison of different piezoelectric materials appropriate for R&D work. When comparing responses of different materials, relative comparative method is usually sufficient and sometimes preferred to comparing absolute values due to its simplicity. However, determination of absolute values of piezoelectric coefficients is also possible, upgrading the proposed method with analysis of mechanical setup using appropriate numerical simulation as shown later. For this purpose, finite element analysis (FEA) software ANSYS was used.

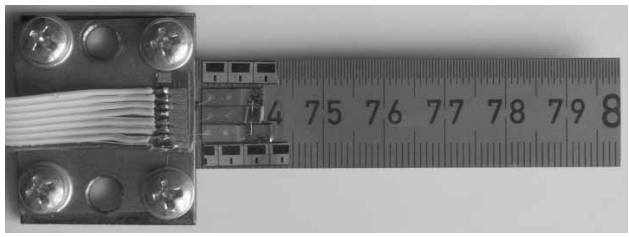
Mechanical properties of piezoelectric and silicon were taken from literature /9,10/. Permittivity was determined from capacitance measurements.

4. Experimental setup

Experimental setup consisting of rectangular cross-section cantilever with mounted samples is shown in Fig. 1. Due to the simplicity of cantilever with rectangular cross-section, also analytical expressions for stress distribution exist, enabling comparison with numerical results. Proposed characterization method uses samples with thin film piezoelectric capacitor structure on silicon substrate, mounted on stainless steel cantilever. The selection of optimal samples placement is essential, usually selected for high sensitivity as the region of maximum stress distribution in the beam still having sufficient uniformity. Stress decreases in cantilever longitudinal direction towards the cantilever free end where it reaches zero. Therefore, the samples are mounted in the region of maximum stress being at the root of the cantilever. Following our simulation results, care must be taken not to induce an excessive error in the placement of samples.



(a)



(b)

Fig. 1: Top view of the cantilever with mounted samples and side guards: (a) schematic, (b) photograph

For adequate characterization of piezoelectric thin film samples, high repeatability of sample loading is essential. The testing cantilever setup, together with bonded samples represents such a test structure. Stainless steel was selected as the material for cantilever, providing possibility of high repeatable deflections. Furthermore, stainless steel cantilever is mechanically resistant and can be reused after replacing samples.

During characterization, samples are often exposed to higher mechanical stresses as during the normal sensor or actuator operation. To achieve such a wide measurement range, cold rolled austenitic stainless steel (1.4310) was selected for the cantilever realization. This material has an extended elastic range due to a special treatment during the fabrication. In this case, the cantilever returns to its initial position even after extremely large deflections.

To achieve large measured range of stresses for samples under test, the mechanical part of testing system has to provide adaptability. Therefore, 10 cm long and 18 mm wide stainless steel strips (cantilevers) of thickness 0.5 mm were cut by milling and then pressed between two rigid stainless steel plates acting as a fixed support. In this approach, the cantilever length is adjustable, resulting in increased measured range with high repeatability and accuracy.

To illustrate the characterization of piezoelectric samples with described experimental setup, various thin PZT layers

were deposited by sol-gel method on silicon chips covered by Pt/Ti as reported elsewhere /11/. Gold electrodes were placed on top of PZT layer by sputtering and shaped by shadow mask method. Thin Ti and Pt layers with thicknesses of 10 and 100 nm respectively are not significant for the overall mechanical properties of the relatively thick samples and were thus neglected in numerical simulations. As an example of the proposed characterization procedure, three samples with two different thicknesses of PZT layer were introduced, marked as samples PZT1a, PZT1b and PZT2. Due to our numerical simulations, two dummy guard chips were added at cantilever sides to achieve better stress uniformity over the samples (Fig. 1).

To assure a reliable transfer of induced mechanical stress from the cantilever to the PZT samples, a strong and stable bond between the cantilever and the samples has to be achieved. Therefore, an epoxy adhesive (UHU endfest 3000) with high bonding strength of 3000 N/cm^2 was used for PZT samples bonding. The extended elastic range of the selected stainless steel, in the combination with the mentioned adhesive enable highly reliable loads on testing samples, up to the silicon tensile strength. In addition, samples fixed with the selected adhesive can be easily removed at relativly low temperatures what makes the testing cantilever reusable /11/.

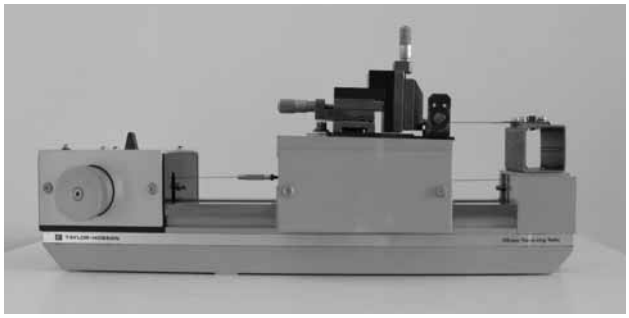


Fig. 2: Experimental setup: Taylor-Hobson traversing table and micromanipulator are used to achieve high deflection repeatability.

To achieve highly repeatable stresses, testing cantilever with bonded samples is mounted on the fixed part of modified Taylor-Hobson 150mm Traversing Table, as shown in Fig. 2. The computer controlled worktable is motor driven in both directions, but can also be moved manually. Straightness accuracy of the worktable is within ±1 µm over the full 150mm range. In order to assure deflection repeatibility, a micromanipulator with 8 mm tall pointed pin is mounted at the top of the worktable, as described in detail elsewhere /12/.

Voltage response of PZT samples is measured by Semiconductor Parameter Analyzer HP4155A, including SMU and PMU Generator Expander HP41501A.

For determination of piezoelectrics permittivity, capacitance on test capacitors is measured with HP4284A Precision LCR Meter at various frequencies, at excitation amplitude 1 V and DC bias 0 V.

5. Numerical modeling

For the purpose of simulation, commercial FEM modeling and simulation software ANSYS was used. Simulator input for cantilever test structure with samples is built using ANSYS proprietary scripting language APDL. Meshing is done using built-in automatic mesh generator. The resulting hexahedral mesh of simulated test structure is shown in Fig. 3. Local improvement of the mesh was done manually to refine mesh in structure critical regions such as thin PZT layer and to avoid badly shaped elements.

The test structure basically consists of several different layers – stainless steel (SS) cantilever, silicon (Si) substrate chip, metal and PZT layer. Electrodes and interface layers were neglected at mechanical simulation due to their small thicknesses. For modeling SS and Si materials, three-dimensional SOLID95 elements were used. PZT layer was modeled with SOLID226 elements with capability to couple mechanical and electrical quantities using piezoelectric effect.

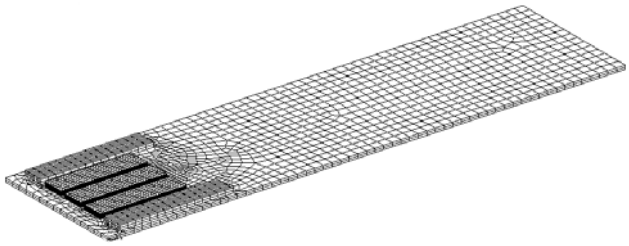


Fig. 3: Generated mesh of cantilever with 3 bonded samples and two side guards.

When we take into account material symmetry, general form of stiffness matrix [*c*] for ceramics, permittivity matrix [*ǐ*] and piezoelectric coefficients matrix [*d*] can be simplified /1/.

$$[c]=\begin{bmatrix} c_{11} & c_{12} & c_{13} & 0 & 0 & 0 \\ c_{12} & c_{11} & c_{13} & 0 & 0 & 0 \\ c_{13} & c_{13} & c_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & c_{44} & 0 & 0 \\ 000 & & & 0 & c_{44} & 0 \\ 00000 & & & & & (c_{11}-c_{12})/2 \end{bmatrix} \quad (3)$$

$$[\check{n}]=\begin{bmatrix} \check{n}_1 & 00 & \\ 0 & \check{n}_1 & 0 \\ 00 & & \check{n}_3 \end{bmatrix} \quad (4)$$

$$[d]=\begin{bmatrix} 0 & 0 & 0 & 0 & d_{15} & 0 \\ 0 & 0 & 0 & d_{15} & 0 & 0 \\ d_{31} & d_{31} & d_{33} & 000 & & \end{bmatrix} \quad (5)$$

Due to the lack of exact information in the literature, mechanical properties of thin PZT layer were approximated by bulk values. Therefore, values c_{11} = 13,9x10^{10} Pa,

c_{33} = 11,5x10^{10} Pa, c_{44} = 2,56x10^{10} Pa, c_{13} = 7,43x10^{10} Pa, c_{12} = 7,78x10^{10} Pa, were taken from literature /10/. Due to the small thickness of PZT compared to the cantilever and Si substrate, the error introduced is negligible.

SS material is usually considered isotropic. The Young's modulus of SS material was measured using nanoindentation method /13/. The measured value of SS Young's modulus is Y = 167.56 GPa.

Silicon is very well known material. Due to Si crystal symmetry, it is described by 3 stiffness coefficients c_{11}, c_{12} and c_{44}. In our case Si was modeled using anisotropic symmetric matrix with coefficients c_{11} = 0,1657x10^6 Pa, c_{12} = 0,0639x10^6 Pa, c_{44} = 0,0796x10^6 Pa/14/.

Due to the longitudinal stress dominating in our case as confirmed by our numerical simulation, only piezoelectric coefficient d_{31} was taken into account.

Boundary conditions for cantilever at FEM simulation were fixed support on the cantilever left side (deflection and its derivative equal to 0) and free deflection on the right side. To allow simple load variation, the deflection was described in the program as a parameter. Electrical ground boundary condition was set on the bottom electrode.

Standard sparse direct linear solver was used for solving the model having 85000 elements with 4 basic variables (degrees of freedom) of the problem: electric potential and displacements in x, y, z direction. Sparse direct solver is a robust and fast solver for linear and nonlinear analysis, appropriate when poorly shaped elements are present in the model, such as the high aspect ratio (thickness vs. width) elements in the model of PZT layer. The sparse direct solver is based on a direct solution of equations by elimination, as opposed to iterative solvers where the solution is obtained through an iterative process that successively refines an initial guess to the final solution that is within a prescribed tolerance of the final solution. Direct elimination requires the factorization of an initial very sparse linear system of equations into a lower triangular matrix followed by forward and backward substitution. Drawback of this solver is that it requires a significant amount of memory, thus it is not suitable for larger scale models with more than a half million variables. Because sparse direct solver is based on direct elimination, poorly conditioned matrices do not pose difficulty in producing the solution /15/. Direct solver was chosen for our simulated approach since it does not exceed the recommended number of equations and there was enough computer memory available to perform computation.

Simulations were performed on Intel Core Duo 6600 64-bit processor architecture with 4GB RAM memory, running at 2.4GHz. A single simulation run with chosen solver required typically 6 minutes.

6. Procedure for determination of piezoelectric coefficient d_{31}

The described experimental setup was used to deflect cantilever. Corresponding voltage response of the mechanically loaded PZT samples was measured with parametric analyzer as described previously. Numerical simulator was configured as discussed in the previous section, to translate the test structure into numerical model. The characterization of piezoelectric effect and related d_{31} parameter was performed by fitting the measured voltage response with simulated response: d_{31} parameter value was varied in the simulator until a good match between measured and calculated voltage response was found. The value of d_{31} that provided best fit throughout all deflections between calculated and measured voltage response was selected as the final result for the piezoelectric coefficient d_{31}.

7. Results and Discussion

The PZT samples capacitance was measured using LCR meter at frequencies ranging from 20 Hz to 10 kHz, at excitation voltage of 10 mV. A relatively small dependence of capacitance vs. frequency was detected (Fig. 4). Measured capacitance value at 1 kHz was taken, as stated in /1/. Top electrode area was measured under the microscope. PZT layer thickness was measured after the fabrication of the layer. From data given in Table 1 the relative permittivity $\check{n}_r$ for samples was calculated. Samples PZT1a and PZT1b are built on the same PZT layer differing only in their electrode position, regarding to the cantilever support (Fig. 1). Electrode of PZT1a was located 2.7 mm from the support, while the electrode of PZT1b was located 5.4 mm from the cantilever support. Sample PZT2 was prepared with modified processing for double thickness of PZT1, resulting in changed value of PZT permittivity. The electrode location for PZT2 was the same as for PZT1a, 2.7 mm from the support.

Simulated stress profile in PZT layer is shown in Fig. 5 (simulation path is shown in the inset). From Fig. 5 can be concluded that electrode exact position is important when performing characterization of multiple samples. Following our numerical simulations results, to minimize the difference of stress profile in neighbor samples, two longer guard chips are added at the sides, as shown in Fig. 1. Calculated stress distribution in the cantilever and samples is shown in Fig. 6a.

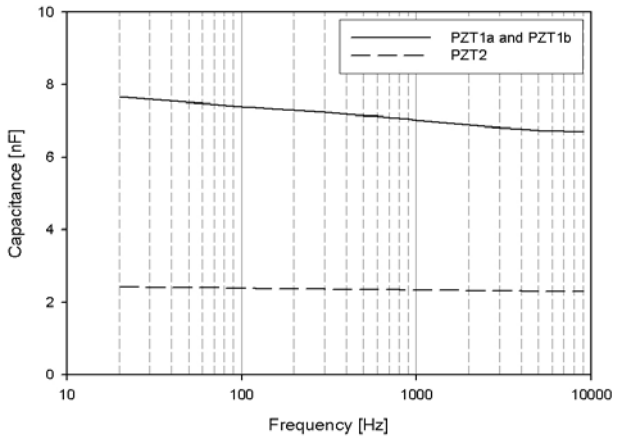


Fig. 4: Measured samples capacitance vs. frequency.

Table 1: Measured sample parameters and calculatedrelative permittivity of PZT layers

Sample	PZT Thickness [nm]	Electrodes [mm^2]	Capacitance [nF]	Rel. permittivity $\check{n}_r$
PZT1a	740	0.87	7.67	737
PZT1b	740	0.87	7.67	737
PZT2	1554	0.87	2.40	484

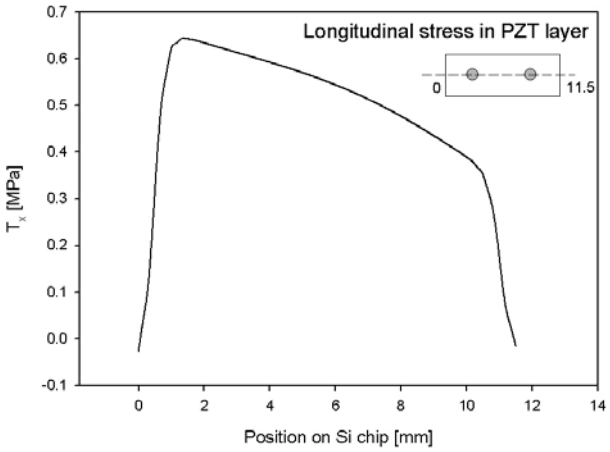


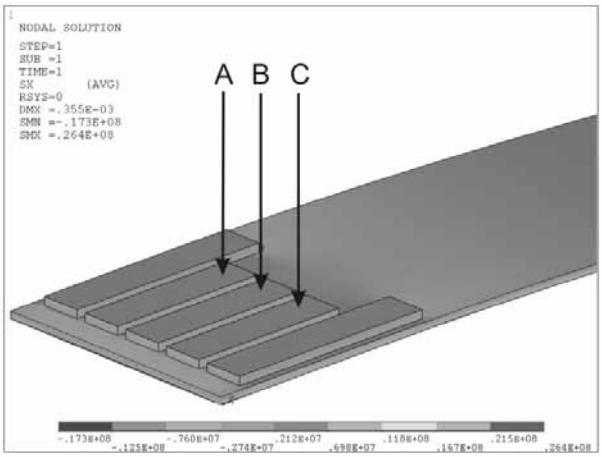
Fig. 5: Simulated longitudinal stress profile in PZT layer vs. position on Si chip.

The effect of guard chips is quantified in Fig. 6 and Table 2. As shown, the absolute stress in samples is decreased when guard chips are present. However, stress uniformity over the samples improves significantly. The relative difference in stress in both cases, without and with guard chips, was calculated between central and side samples. Guard chips thus provide more homogenous stress conditions on all samples.

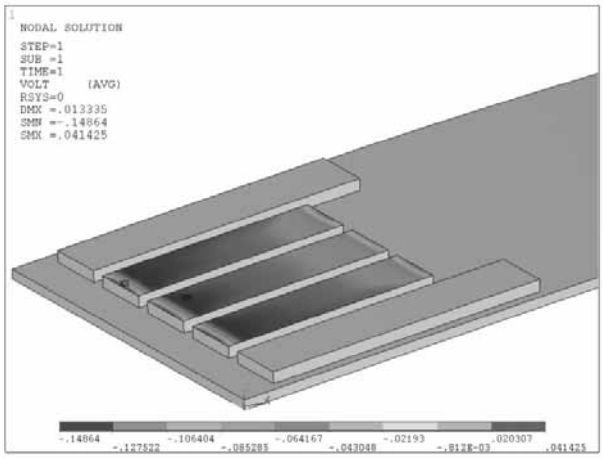
According to the piezoelectric effect, voltage response is proportional to the stress, what is described by piezoelectric coefficients. Calculated voltage response of PZT samples due to calculated stress is given in Fig. 6b.

Measured time dependent voltage response of PZT samples during testing is shown in Fig. 7. Here, the cantilever was deflected to predefined values using the micromanipulator as previously described. At start, the cantilever was first manually deflected over the desired deflection value, and then after this it was released to rest in final position determined by micromanipulator. Similar procedure was applied also during the end of testing. Consequently, voltage spikes always occurred at the start and at the end of loading.

As also seen in Fig. 7, the response for constantly deflected cantilever slowly decreases with time, probably due to piezoelectric internal effects such as leakage and recombinations, and due to external effects such as input impedance of HP4155A connected to the sample. Therefore, measurement of the response was done after the spike settled down, typically after 10 seconds.



(a)



(b)

Fig. 6: Simulated longitudinal stress distribution in stainless steel cantilever and silicon chips (a) and corresponding voltage on top of PZT layer due to accumulated charge (b). Positions on the chips A, B and C show where stresses were compared.

Table 2: Improvement of the stress uniformity over samples when guard chips are used.

Stress in samples **without guard chips** [MPa]			Stress in samples **with guard chips** [MPa]		
Position B	Positions A, C	Rel. difference	Position B	Positions A, C	Rel. difference
13.19	14.74	10.5 %	10.04	10.08	3.7 %

Table 3: Measured voltage response and simulated values for different deflections of cantilever with best fit value for d_{31} parameter.

	PZT1a		PZT1b		PZT2	
Deflection	Meas[mV]	Sim [mV]	Meas[mV]	Sim [mV]	Meas[mV]	Sim [mV]
3.175 mm	35.5	35.24	24.9	27.80	41	37.38
5.715 mm	60.0	63.40	50.8	50.10	63.7	67.30
8.255 mm	95.4	91.62	76.9	72.43	92.5	97.19
10.795 mm	121	119.8	93.9	94.70	125	127.1
13.335 mm	144	148.0	114	117.0	158	157.1
15.875 mm	173	176.2	141	139.1	187	186.9

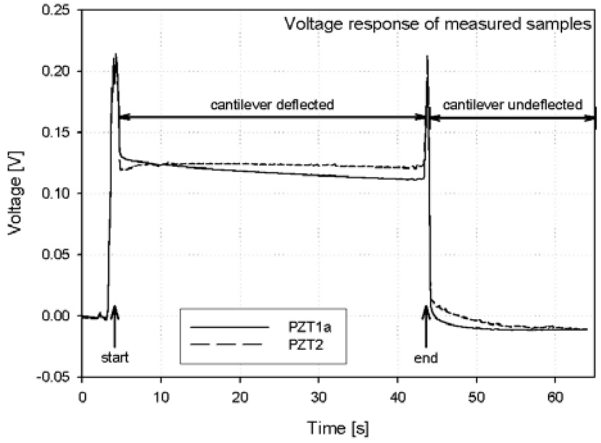


Fig. 7: Measured voltage response vs. time during testing

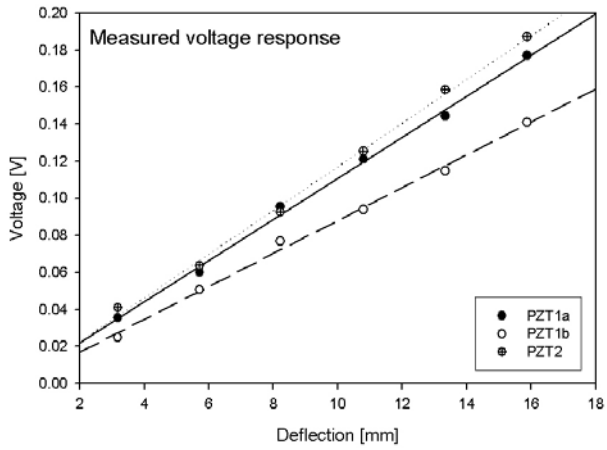


Fig. 8: Measured voltage response of PZT samples vs. deflection.

Measured voltage response results are graphically displayed in Fig. 8. The response amplitude is dependent on electrode distance from the cantilever support and is in correlation with simulated stress profile in PZT layer shown in Fig. 5. The voltage on PZT1a is thus considerably higher than voltage on PZT1b. PZT2 that differs in thickness and material properties produces response that is slightly higher than with PZT1a.

Determination of piezoelectric coefficient d_{31} was done by using numerical simulation as described previously. Successive simulations were performed for various values of coefficient d_{31} until close agreement between simulated and measured voltage response was obtained. Measured and simulated responses at various deflections for best values of piezoelectric coefficient d_{31} are given in Table 3. The summary of measured values for relative permittivity n_r and piezoelectric coefficient d_{31} for PZT materials under test is given in Table 4. Results obtained are in reasonable agreement with available values from the literature /7/.

Table 4: Measured properties of PZT layer.

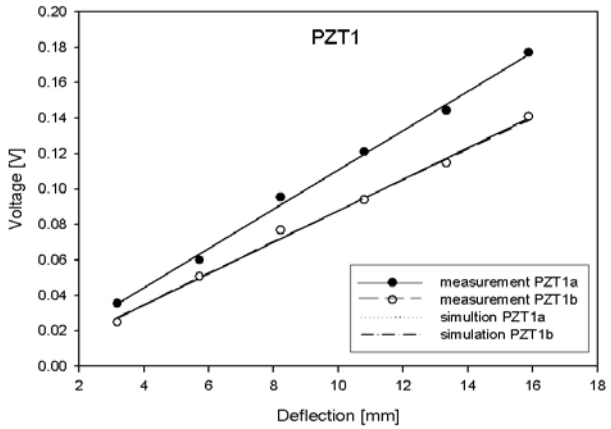
Sample	Relative permittivity - ň$_r$	Piezoelectric coef. - d_{31} [pC/N]
PZT1a	737	-66.1
PZT1b	737	-66.1
PZT2	484	-20.7

Graphical representation of measured and simulated voltage responses vs. deflection for all three PZT samples are shown in Fig. 9. In the range of measured deflections, the simulated response displays linearity while it is slightly distorted for measured values, probably due to measurement error.

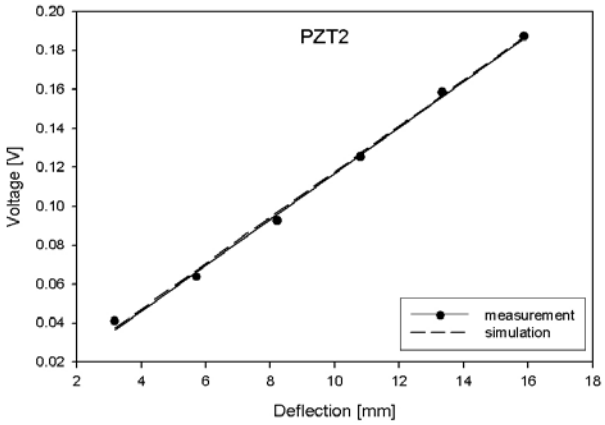
8. Conclusion

Using the proposed cantilever method, piezoelectric coefficients d_{31} for various thin film piezoelectrics were determined. Cantilever based characterization method provides a fast comparison of different piezoelectric material samples, since multiple samples can be mounted simultaneously on the testing structure. Furthermore, when combining experimental data with numerical simulation, piezoelectric coefficient d_{31} can be determined by matching simulated results with voltage response measurements.

Exact knowledge of geometry and material properties of cantilever and samples proved to be important in order to measure piezoelectric coefficients with sufficient accuracy. Stainless steel cantilever was adequately characterized by measuring its Young's modulus. Silicon properties are adequately determined by published data in the literature. Mechanical properties of PZT layers are on the other hand more difficult to acquire, since they are rather dependent on the actual PZT preparation procedure and composition. Nevertheless, we expect that error here introduced



(a)



(b)

Fig. 9: Graphical representation of measured and simulated voltage response of PZT samples vs. deflection.

is small due to very thin PZT layer compared to stainless steel cantilever and silicon substrate. To improve the presented method, based on numerical simulation results guard chips were mounted at the side of the cantilever to reduce stress variation over the samples. Determined values of piezoelectric coefficients d_{31} for PZT layers under test were in reasonable agreement with results available in the literature.

Acknowledgment

Authors would like to acknowledge Electronic Ceramics Department – K5, Jožef Stefan Institute, Slovenia for PZT samples preparation.

This work was supported by Ministry of Higher Education, Science and Technology and Slovenian Research Agency.

References

/1/ T.L. Jordan, Z. Ounaies, "Piezoelectric Ceramics Characterization". NASA/ CR-2001-211225 ICASE report to NASA Langley Research Center. Report No. 2001-28, September 2001.

/2/ Ansys Inc. "Ansys Inc. Theory reference". Ansys Inc., 2005.

/3/ S. Penič, U. Aljančič, D. Vrtačnik, D. Resnik, M. Možek and S. Amon "Numerical modeling of PZT/SiO_2 microcantilever with interdigitated electrodes", Proc. 43rd International Conference on Microelectronics, Devices and Materials and the Workshop on Electronic Testing, Bled, Slovenia, September 2007, pp. 63-68.

/4/ "IEEE Standard on Piezoelectricity", (IEEE Standard 176-1987), Institute of Electrical and Electronic Engineers, 345 East 47th St, New York, NY 10017.

/5/ K. C. Kao, "Dielectric phenomena in solids", Elsevier Academic Press, San Diego, California, 2004.

/6/ J.T. Dawley, G. Teowee, B.J.J. Zelinski and D.R. Uhlmann "Piezoelectric Characterization of Bulk and Thin Film Ferroelectric Materials using Fiber Optics". MTI Instruments application note, http://www.mtiinstruments.com/

/7/ J.F. Shepard Jr., P.J. Moses and S. Trolier-McKinstry "The wafer flexure technique for the determination of the transverse piezoelectric coefficient (d_{31}) of PZT thin films". Sensors and Actuators A, vol. 71, 1998, pp. 133-138.

/8/ H.D. Espinosa , B.C. Prorok, M. Fischer "A methodology for determining mechanical properties of freestanding thin films and MEMS materials". Journal of the Mechanics and Physics of Solids, vol. 51, 2003, pp. 47-67.

/9/ Efunda, http://www.efunda.com/

/10/ X.J. Zheng, Y. C. Zhou and J.Y. Li "Nano-indentation fracture test of $Pb(Zr_{0.52}Ti_{0.48})O_3$ ferroelectric thin films". Acta materialia, vol. 51, 2003, pp. 3985-3997.

/11/ U. Aljančič, B. Malič, M. Mandeljc, M. Vukadinović, D. Vrtačnik, D. Resnik, M. Možek, M.Kosec, S. Amon "Cantilever as Testing Structure for Characterization of PZT Thin Films on Pt/Si Substrates". Proc. 42nd International Conference on Microelectronics, Devices and Materials and the Workshop on MEMS and NEMS, Strunjan, Slovenia, September 2006, pp. 271-276.

/12/ U. Aljančič, M. Vukadinović, D. Resnik, D. Vrtačnik, M. Možek, S. Penič, S. Amon "Cantilever Characterization Method for Static Behavior of PZT Thin Films". Proc. 43rd International Conference on Microelectronics, Devices and Materials and the Workshop on Electronic Testing, Bled, Slovenia, September 2007, pp. 115-120.

/13/ S. Penič, U. Aljančič, D. Vrtačnik, D. Resnik, M. Možek, M. Makovec, R. Bošnjak and S. Amon "FEM modeling of piezoresistive force sensor for medical retractor and design verification". Proc. 6th EUROSIM Congress on Modelling and Simulation, Ljubljana, Slovenia, September 2007, p. 158.

/14/ A. M. Fitzgerald "Practical Issues in Finite Element Analysis of MEMS". Ansys Workshop, March 2006.

/15/ Ansys Inc. "Ansys Inc. Basic Analysis Guide". Ansys Inc., 2005.

Samo Penič, univ. dipl. inž. el.
mag. Uroš Aljančič
doc.dr. Drago Resnik
doc.dr. Danilo Vrtačnik
mag. Matej Možek
prof.dr. Slavko Amon

University of Ljubljana,
Faculty of Electrical Engineering,
Laboratory of Microsensor Structures and Electronics
Trzaska 25, Ljubljana 1000, SLOVENIA
e-mail: matej.mozek@fe.uni-lj.si
Telefon: 01 4768 303, Telefax: 01 4264 630

Prispelo (Arrived): 08.05.07 *Sprejeto (Accepted):* 28.5.08

NAČRTOVANJE PRENOSNEGA MERILNEGA SISTEMA ZA MERJENJE POSPEŠKOV

Ciril Močnik, Dejan Križaj

Laboratorij za bioelektromagnetiko, Univerza v Ljubljani, Fakulteta za elektrotehniko, Ljubljana, Slovenija

Kjučne besede: senzor pospeškov, zapisovalnik podatkov, mikroprocesor, SD kartica, FAT datotečni sistem

Izvleček: Miniaturni merilniki pospeška pridobivajo vedno večji pomen v medicini in športu. V določenih primerih je potrebno zajemati podatke pospeškometrov s prenosnim merilnim sistemom. V prispevku je predstavljeno načrtovanje in izvedba prenosne naprave za zajem podatkov merilnikov pospeška in njihovo shranjevanje na spominsko kartico. Izdelan sistem omogoča zapisovanje osmih analognih signalov frekvence 1,7 kHz.

Design of Portable Data Logger System for Accelerometer Sensors

Key words: accelerometer sensor, data logger, microprocessor, SD card, FAT file system

Abstract: Accelerometer sensors have gained wide use in medicine and sport in recent years. Our goal is to upgrade our recent studies of muscle response with accelerometer sensor data logging instrument. As instruments already present on the market either didn't satisfy our needs or are very expensive, we decided to develop a dedicated device for specific needs. This paper presents a design of a lightweight, portable data logger system capable of measuring and storing eight signals from accelerometer sensors. The design is based on a low cost 8 bit microprocessor supported by a 12 bit eight channel A/D converter and an SD media storage card. For ease of operation an LCD character display and four keys are added. The instrument can operate either with a wall DC adapter, two AA size batteries or NiCd rechargeable batteries. Collected data is stored in data files on an SD card formatted with a FAT file system, which makes them compatible with any PC for further analysis. Final tests fulfill all our expectations. The instrument is very light, with eight analog channels and up to 512 MB of storage space on an SD card. It is capable of recording a wide spectrum of different tests in sport and medicine research with maximal frequency of 1,7 kHz that can be even increased when measuring short bursts.

1 Uvod

Merilnike pospeška najdemo v avtomobilih, računalnikih, navigacijskih napravah, športnih rekvizitih itd. Ker njihova nabavna cena pada, hkrati pa se veča zmogljivost, se v bodočnosti obeta še bolj pogosta uporaba teh elementov. Posebno se je razširila uporaba merilnikov pospeška z njihovo miniaturizacijo, kar je omogočila mikroelektronska tehnologija z dodatnimi znanji in tehnologijami iz načrtovanja in izdelave mikromehanskih struktur. Tehnologija izdelave in princip delovanja MEMS (Micro-Electro-Mechanical Systems) naprav sta natančneje opisani v (1). Ti miniaturni pospeškometri omogočajo določanje pospeška iz premikov, ki jih lahko beležijo s pomočjo piezoelektričnega pojava ali pa s pomočjo spremembe kapacitivnosti. Piezoelektrični senzorji imajo boljše šumne lastnosti, senzorji na podlagi spremenljive kapacitivnosti pa so cenejši in enostavnejši za izdelavo in s tem tudi bolj razširjeni.

Prvotno zanimanje za uporabo miniaturnih pospeškometrov kot merilnik mišičnega odziva smo želeli nadgraditi z napravo, ki bi omogočala zajemanje večje količine podatkov pospeškometrov v situacijskem okolju - s prenosljivo napravo za zajem in shranjevaje podatkov. Določene meritve je sicer možno izvesti v laboratorijskem okolju, kjer je na voljo potrebna oprema za zajemanje in shranjevanje podatkov in njihovo kasnejšo obdelavo kot v primeru treninga veslačev (2). Če pa želimo izvajati take meritve v športu, hitro naletimo na omejitev tako prenosljivosti kot priročnosti opreme. Obstaja kar nekaj naprav, ki omogočajo zajemanje podatkov pospeškometrov z različnimi področji uporabe. V industriji in terenskih meritva so v rabi robustnejše naprave za stalno namestitev, ki lahko delujejo samostojno dlje časa. Dve sta predstavljeni v (3) in (4). Prenosne izvedba kot je (5) bi bile zelo primerne za meritve v športu, vendar imajo premajhno hitrost zajemanja podatkov in pomnilnik, tudi cena je glede na zmogljivost previsoka. Bolj ustreza (6), ki pa obstaja samo kot razvojni prototip ali pa (7), ki je univerzalna in visoko zmogljiva naprava, vendar brez uporabniškega vmesnika na napravi, in žal tudi z zelo visoko ceno.

Večina naprav, ki so dostopnie na tržišču in smo si jih ogledali ne ustreza popolnoma našim zahtevam (so prevelike, niso prenosne, premalo zmogljive, so v fazi razvoja ali pa so preprosto predrage). Zato je bila sprejeta odločitev, da izdelamo lastno napravo, ki bo lahka (160 g z akumulatorji), cenovno ugodna, ter obenem dovolj zmogljiva.

V tem prispevku je predstavljena izvedba merilnega sistema z lastnim napajanjem, osmimi A/D kanali in shranjevanjem podatkov na izmenljiv pomnilniški medij.

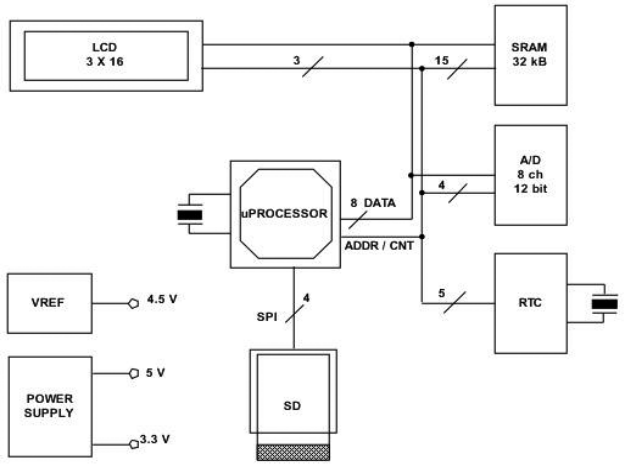
2 Načrtovanje

2.1 Zahteve

Za željeno funkcionalnost mora prenosni merilnik pospeškov izpolnjevati sledeče zahteve:

- lahka, prenosna izvedba, primerna za nošnjo na telesu ali v oblačilih
- upravljanje z napravo preko LCD zaslona in miniaturnih tipk, brez uporabe računalnika
- osem A/D vhodov z 12 bitno ločljivostjo
- nastavljiva frekvenca vzorčenja, z najvišjo vrednostjo 1 kHz ali več
- zapis posamezne meritve v obliki ASCII datoteke na SD kartico (8) v datotečnem sistemu FAT
- ura realnega časa s pomožnim napajnjem
- baterijsko, akumulatorsko ali zunanje napajanje

Glede na zahteve je načrtovanje naprave razdeljeno na zaključene funkcionalne sklope: mikroprocesor s periferno opremo, A/D pretvornik z izvorom referenčne napetosti, zunanji pomnilnik, ura in napajalnik, uporabniški vmesnik in programska oprema.



Slika 1: Shematski prikaz merilnega sistema

2.2 Mikroprocesor

Naprava ne rabi velike računske moči procesorja, ker se podatki samo zbirajo in zapisujejo na pomnilniški medij, nadaljna obdelava pa se vrši šele po prenosu na PC računalnik. Zato je bil za procesor izbran sicer manj zmogljiv osem bitni procesor, ki pa ima zaradi svoje priljubljenosti na voljo veliko uporabnih knjižnic v C jeziku (9) in cenena razvojna orodja. 16 Mhz Atmelov procesor (10) s 128 kB programskega flash pomnilnika, SPI vodilom in vmesnikom JTAG, ki omogoča enostavno programiranje in razhroščevanje se je pokazal kot primerna izbira. Edina pomanjkljivost je vgrajeni A/D pretvornik, ki ima samo 10 bitno ločljivost, zato je bilo dodano zunanje A/D vezje. Procesor lahko deluje pri napetostih 3,3 ali 5 V, vendar je pri nižji napetosti frekvenca delovanja omejena na 8 Mhz, zato je bila izbrana višja napetost, kar pomeni, da je pri komunikaciji s SD kartico, ki deluje samo pri napetosti 3,3 V, nujna uporaba pretvornika nivojev. Ker smo želeli izločiti kakršenkoli vpliv na stabilnost delovanja vezja zaradi slabo definiranih logičnih nivojev, so bila izbrana namenska integrirana vezija 74LVC4245, čeprav so dražja in težje dobavljiva.

2.3 Analogno digitalni pretvornik

Izdelovalcev A/D pretvornikov je več, še mnogo več pa je raznih modelov in izvedb, različnih tako po zmogljivosti kot ceni. Kljub skoraj nepregledni množici pa lahko hitro omejimo ustrezne izvedbe na sprejemljivo število. Prva pogoja sta bila osem analognih vhodov in 12-bitna ločljivost, ki sta precej zožila izbiro, še bolj pa zahtevi po paralelnem vmesniku in zunanji napetostni referenci velikosti 4,5V. Dodatni zahtevi sta bili še majhna poraba in hitro zajemanje podatkov.

Osem analognih vhodov potrebujemo zaradi možnosti priključitve najmanj štirih senzorjev pospeška, od katerih ima vsak dva (X, Y) analogna izhoda. 12 bitna ločljivost pa naj bi omogočala registriranje tudi majhnih sprememb pospeška, ko je lasten šum senzorja še manjši od ločljivosti.

Paralelni vmesnik zagotavlja veliko hitrost prenosa podatkov med A/D pretvornikom in procesorjem in hkrati omogoča preprosto naslavljanje A/D pretvornika kot periferne naprave in s tem poenostavljen dostop do podatkov in krmilnih registrov. Nekateri A/D pretvorniki sicer omogočajo priključitev na dovolj zmogljivo vodilo SPI, vendar je v našem primeru to že zasedeno s prenosom podatkov na pomnilniško kartico.

Le manjši del A/D pretvornikov na tržišču omogoča priključitev zunanje reference, ki je višja od 3 V, ker pa smo želeli doseči univerzalnost naprave tudi z možnostjo priključitve drugih analognih izvorov, je bil cilj, da je referenčna napetost čim bližja napajalni napetosti oziroma petim voltom. Z izvorom referenčne napetosti se napajajo tudi merilniki pospeška. S tem izločimo vpliv nestabilne napajalne napetosti. Izhodna napetost merilnikov je namreč odvisna ne samo od pospeška ampak je tudi sorazmerna napajalni napetosti.

2.4 Zunanji pomnilnik

Zunanja pomnilnika sta dva, eden je že omenjena SD kartica, povezana na procesor preko serijske SPI povezave in služi za shranjevanje podatkov osmih A/D kanalov. Drugi pa je 32 kB SRAM pomnilnik z 8 bitno vzporedno povezavo, ki ima vlogo izravnalnega FIFO medpomnilnika za podatke iz A/D pretvornika . Ta pomnilnik je nujen zaradi uporabe FAT sistema na SD kartici.

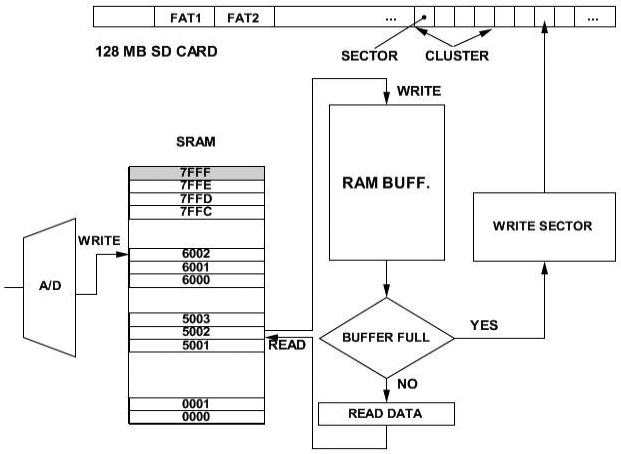
Sistem FAT16 razdeli pomnilnik na največ 65.535 gruč, ki vsebujejo ustrezno število sektorjev po 512 bajtov. V primeru 128 MB kartice je vsaka gruča velika štiri sektorje, torej 2kB. Datoteka se pri vpisovanju v pomnilnik zapiše v več gruč, redosled pa se hrani v FAT tabeli. Podatki se zapisujejo v sektorje velikosti 512 B in to cel sektor naenkrat. Tako lahko izračunamo, da se bo pri frekvenci vzorčenja 1 kHz in osmih kanalih zgodil zapis sektorja vsake:

512/16000 = 32 ms. (16000 = osem kanalov X dva Bajta X 1000 vzorcev/s)

Knjižnica za delo s FAT sistemom rezervira del RAM pomnilnika v procesorju za hranjenje tega sektorja in ko je sek-

tor poln se prepiše v spominsko kartico, v tem času pa v pomnilnik na procesorju ne moremo shranjevati podatkov iz A/D pretvornika.

Rutine v knjižnici za delo s FAT sistemom zapišejo sektor v 1 - 2 ms, kar pomeni od 16 do 32 bajtov podatkov, ki jih moramo medtem shraniti v začasni register. Večji problem se pojavi, ko tabela FAT preseže okvir ene gruče. Pri 128 MB pomnilniški kartici se to zgodi pri datotekah večjih od 500 kB. Takrat je lahko zakasnitev dolga do 250 ms kar predstavlja 250 X 16 = 4 kB podatkov. Ker ima procesor na razpolago le 4 kB RAM pomnilnika, od katerega je del že zaseden s podatki FAT sektorja, del pa s programskimi spremenljivkami, je potrebno realizirati zunanji FIFO register s SRAM pomnilnikom. Podatki iz A/D pretvornika se vpišejo najprej v SRAM pomnilnik, iz katerega jih program prepisuje v medpomnilnik sektorja v procesorju. Ko je ta poln se prepiše na SD kartico. V najslabšem primeru bo v SRAM 4 kB podatkov, ki jih, potem ko se FAT tabela dopolni, program prepiše v spominsko kartico v času cca. 16 ms.



Slika 2: Izravnalni medpomnilnik

Ker je naslavljanje pomnilnika 15 bitno, je za povezavo med pomnilnikom in procesorjem potrebno vstaviti 8 bitna D vrata (74AHC573), ki v fazi naslavljanja pomnilnika zaklenejo spodnjih osem bitov naslova. Proizvajalec procesorja opozarja, da v primeru uporabe pri najvišji frekvenci HC izvedba vrat ne zadošča, ker je prepočasna in lahko pride do napak pri dostopanju do pomnilnika, zato priporočajo uporabo izvedbo AHC. Žal so slednja težje dosegljiva in obenem precej dražja.

2.5 Ura realnega časa in napajalnik

Naprava je namenjena predvsem za terensko delo kjer lahko izvajamo veliko število meritev v različnih časovnih presledkih, zato je uporaba ure realnega časa nujna. Ura in datum se tako zapišeta v glavo datoteke s podatki, ki se generira ob začetku vsake nove meritve. Visokokapacitiven kondenzator služi kot vir napetosti za delovanje ure tudi, ko je naprava izklopljena.

Napajalnik je izveden z DC/DC pretvornikom navzgor, ki deluje do vhodne napetosti 0,7 V. Dodaten linearni regulator zagotavlja 3,3 V napetost za napajanje SD kartice.

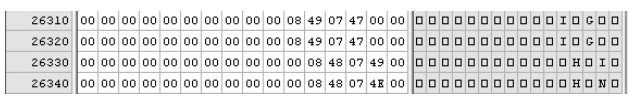
2.6 Uporabniški vmesnik

Uporabnik upravlja z napravo preko štirih tipk in trivrstičnega LCD prikazovalnika. Zaradi majhnih dimenzij (višina) je bil izbran LCD prikazovalnik z integriranim krmilnikom na steklu prikazovalnika (COG). Štiri miniaturne tipke zadostujejo, da lahko preko sistema menijev v napravo vnašamo parametre in prožimo delovanje. Spreminjamo lahko frekvenco vzorčenja, trajanje meritve, aktivne merilne kanale, način merjenja in nastavimo uro realnega časa. Na začetek in konec meritve opozori zvočni signal, kar je zelo uporabno pri kratkotrajnih meritvah.

2.7 Programska oprema

Programska oprema, ki teče v mikroprocesorju je v celoti napisana v jeziku C, zbirnika ni bilo potrebno uporabiti niti v časovno kritičnih delih programa. Uporabljeni sta bili knjižnica AVR libc in knjižnica za delo s FAT sistemom. Posebna pozornost je bila posvečena delu za uporabo zunanjega izravnalnega pomnilnika. Za programsko realizacijo registra FIFO v zunanjem pomnilniku uporabljamo dva kazalca. Prvi kazalec (RAM_WRITE) kaže na prvo zaporedno prosto mesto, kamor lahko program piše podatke, drugi (RAM_READ) pa na prvo zaporedno neprebrano vrednost v pomnilniku. RAM_WRITE lahko prehiti kazalec RAM_READ, medtem ko je slednji lahko največ enak prvemu.

V testnem primeru tako vzorčimo štiri analogne vhode, dva brez napetosti na vhodu in dva s konstantno ali spremenljivo napetostjo. Vrednosti v celicah v SRAM si sledijo z zamikom enega bajta ob vsakem resetu kazalca (to je posledica velikosti pomnilnika, ki je deljiva z osem, vendar ene pomnilne celice ne uporabljamo). V datoteko na pomnilniški kartici pa se podatki vpisujejo sekvenčno, tako, da tega zamika ne opazimo. Opazimo pa lahko zamik podatkov v primeru, da kazalec pisanja v SRAM dohiti in prehiti kazalec čitanja iz pomnilnika SRAM, to se zgodi takrat, ko je hitrost vpisovanja podatkov v zunanji SRAM večja od hitrosti, s katero se podatki prenašajo v pomnilniško kartico. To lahko najlažje vidimo v heksadecimalnem prikazu, ko se podatki z istega senzorja ne pojavljajo več na istem mestu, ampak se premaknejo.



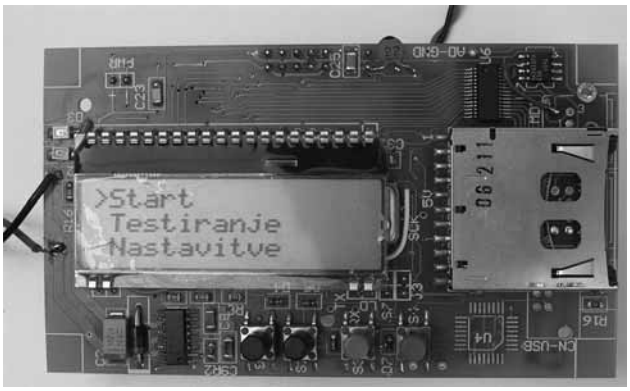
Slika 3: Primer premika podatkov v ASCII datoteki zaradi prevelike frekvence vzorčenja

Prvič se to pojavi pri frekvenci vzorčenja ~1,9 kHz. Ker se pri prvi nižji frekvenci ~1,7 kHz, to ne pojavi, lahko sklepamo, da je najvišja frekvenca, s katero še lahko zajemamo podatke brez izgube ~1,7kHz.

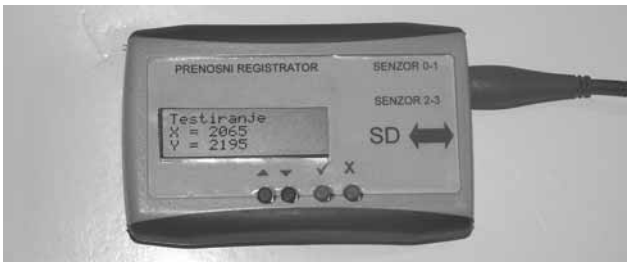
3. Realizacija in testiranje

Elektronsko vezje je izvedeno na dvostranskem tiskanem vezju dimenzij 100 mm x 57 mm in vgrajeno v priročno plastično ohišje z ločenim prostorom za dve AA bateriji ali NiCd akumulatorja.

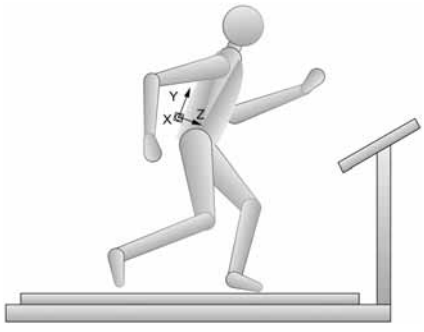
Zaradi omejenega prostora v ohišju je bilo potrebno poiskati dimenzijsko ustrezne komponente (LCD, konektor SD kartice,…), in preizkusiti več kombinacij pri postavitvi elementov na tiskanem vezju. V končni verziji je dodana še USB povezava, ki pa v programski opremi še ni podprta. Da bi še povečali uporabnost naprave, je na voljo tudi možnost shranjevanja podatkov samo v zunanji SRAM pomnilnik in prepis v pomnilniško kartico šele po končani meritvi. Tako se izognemo ozkemu grlu, ki se je pojavljalo pri vpisu v pomnilniško kartico, smo pa seveda s tem omejili število vzorcev na 16.000 vzorcev velikosti dveh bajtov.



Slika 4: Sestavljeno vezje v fazi razvoja programske opreme

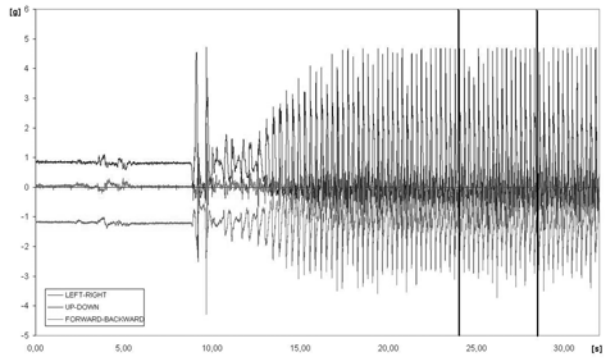


Slika 5: Izgotovljen merilni sistem

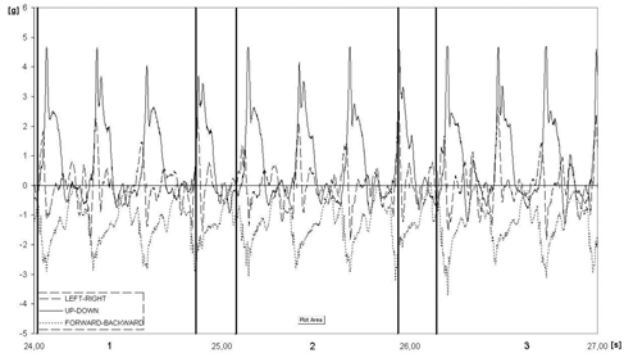


Slika 6: Položaj senzorja na poizkusni osebi med testiranjem. X os v smeri levo-desno, Y os v smeri gor-dol, in Z os v smeri naprej-nazaj.

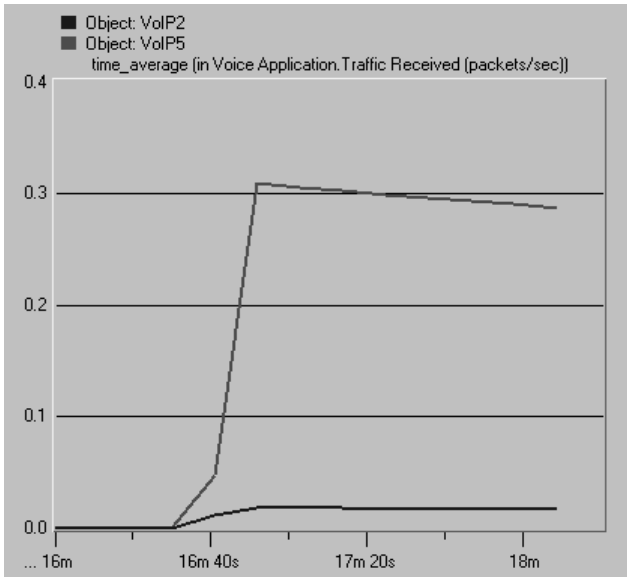
Za test delovanja naprave je bila izbrana poizkusna oseba ki je izvajala enakomerni tek na neskončnem traku. XYZ senzor je bil pritrjen na ledvenem delu ob hrbtenici. Usmeritev osi je razvidna s slike 6.



Slika 7: Posnetek prvih 30 sekund odziva senzorja na tek poizkusne osebe. Z navpičnima črtama je označen izsek prikazan na sliki 8.



Slika 8: Povečan izsek gibanja, kjer so vidni pospeški v vseh treh smereh. S številkami 1 do 3 so označeni odseki odziva na gibanje gor-dol, ki so poravnano prikazani na sliki 9.



Slika 9: Prikaz rezultata merjenja pospeška gibanja gor-dol za tri zaporedne eno sekundne odseke.

4 Zaključek

Merjenje pospeškov s prenosno in miniaturno napravo omogoča spremljanje pomembnih kinematičnih parametrov v športu. V ta namen smo specificirali, načrtali, izdelali in testirali prenosno merilno napravo, ki omogoča hitro zajemanje merilnih podatkov osmih merilnikov pospeškov oz. 8. signalov z zajemom v analogni obliki. Naprava shranjuje izmerjene pospeške s frekvenco do 1700 Hz, kar omogoča zaznavanje hitrih sprememb pospeškov. Podatki se shranjujejo na standardno SD kartico (Secure Digital), ki omogočajo zapise velike količine podatkov. Izdelava je miniaturna, kar je bistveno za nemoteče izvajanje meritev v situacijskem okolju. Testiranje je pokazalo, da smo izdelali napravo, ki izpolnjuje vse naše zastavljene cilje, poleg izpolnjenih tehničnih zahtev je tudi majhna, lahka in enostavna za uporabo. V primerjavi s cenejšimi izdelki na tržišču je mnogo zmogljivejša, presega pa tudi mnogo dražjih naprav.

Nekoliko moteče je le ožičenje, saj je potrebno pospeškometre fizično povezati z merilno napravo, kar v določenih primerih uporabe lahko deluje moteče ali celo onemogoča uporabo. V takih primerih bi bila koristna miniaturna izvedba senzorskega vezja z lastnim napajanjem in radijskim prenosom podatkov v računalnik.

5 Literatura

/1./ Drago Strle in Volker Kempe. "MEMS-based inertial systems". Informacije MIDEM, 4/2007, 199-209.

/2. / "Application of Accelerometers in Sports Training". Analog devices. (www.analog.com)

/3. / "Tri-axial shock data logger". Magdetech inc. (www.magdetech.com)

/4. / "G-Logger. Acceleration acquisition system". Silicondesigns inc. (www.silicondesigns.com)

/5. / "GT1M activity monitor". Actigraph inc. (www.theactigraph.com)

/6. / Tom Ahola, Pekka Korpinen, Juha Rakkola, Teemu Ramo, Jukka Salminen in Jari Savolainen. "Wearable FPGA Based Wireless Sensor Platform". Proceedings of the 29th Annual International Conference of the IEEE EMBS Cité Internationale, Lyon, France August 23-26, 2007.

/7. / "MIE Data logger", **MIE Medical Research Ltd. (www.mie-uk.com)**

/8. / SanDisk corporation. "SanDisk secure digital card - product manual". Ver.: 1.9, december 2003.

/9. / "AVR-libc reference manual". Ver.: 1.4.3, 2006.

/10. / ATMEL, "ATMega 128 user manual", november 2004.

Ciril Močnik, univ. dipl. inž. el.
dr. Dejan Križaj, univ. dipl. inž. el.

Laboratorij za bioelektromagnetiko (LBM)
Univerza v Ljubljani, Fakulteta za elektrotehniko
Tržaška 25, 1000 Ljubljana, Slovenija

e-mail:
ciril.mocnik@siol.net
dejan.krizaj@fe.uni-lj.si

tel.: +386 1 4768 720, fax.: +386 1 4264 658

Prispelo (Arrived): 03.05.07 *Sprejeto (Accepted):* 28.5.08

HYBRID FUNCTIONAL VERIFICATION OF A USB HOST CONTROLLER

Primož Puhar[1], Andrej Žemva[2]

[1]LEA, d.o.o., Lesce, Slovenia

[2]University of Ljubljana, Faculty of Electrical Engineering, Slovenia

Key words: functional verification, SystemC, TLM, ABV, SCV, simulation, USB

Abstract: With everyday growing demands, complexity of electronic devices has been constantly increasing. Functional verification has become the major bottleneck in the design and verification flow. In order to respond to modern demands, new devices are made of standard pre-verified reusable IP blocks created by using abstract TLM. The paper proposes a three-step design and verification flow based on a reusable test bench. It enables a short design time for a fast-simulating functionally-verified TL model, to be used in early SW development, and a functionally-verified RTL model, ready for HW implementation. The approach is demonstrated on a USB host controller design.

Hibridno funkcionalno preverjanje USB gostitelj krmilnika

Kjučne besede: funkcionalno preverjanje, SystemC, TLM, ABV, SCV, simulacija, USB

Izvleček: Zaradi vsak dan večjih zahtev postajajo elektronske naprave vse bolj kompleksne. Funkcionalno preverjanje je zato postalo najožje grlo v postopku načrtovanja in verifikacije le-teh. Da bi se prilagodili modernim zahtevam, moramo nove naprave sestavljati iz standardnih pred-preverjenih IP blokov, ki smo jih ustvarili s pomočjo abstraktnega TLM. Članek predlaga tristopenjski postopek za načrtovanje in preverjanje, ki temelji na večkratno uporabnem testu. Postopek omogoča načrtovanje hitrega funkcionalno preverjenega TL modela, uporabnega za zgodnji začetek načrtovanja programske opreme. Dodatno omogoča načrtovanje funkcionalno preverjenega RTL modela, pripravljenega za implementacijo. Pristop je prikazan na primeru USB gostitelj krmilnika.

1 Introduction

Though electronic devices, such as smartphones, multimedia players and others, already combine a lot of different functions, the market incessantly demands new functionalities like new audio and video decoders, new accessibility features and support for new interfaces. In future, functionality of any single device shall have to be improved, meaning that its complexity will be drastically increased. A higher level of complexity requires more effort in a device design and verification. Considering also the extreme time-to-market pressures, there is no doubt that new advanced solutions shall have to be provided.

Functions of electronic devices can be assured either by software (SW) or hardware (HW) components or a combination of both. Though there have been new verification techniques developed, simulation is still the most used approach to functional verification. By using slow-simulating Register Transfer Level (RTL) HW models verification times have increased to the level when they now take up to 70% of the device design time and cost /1 – 3/.

In order to respond to the constantly growing new demands, several approaches have been proposed. By using standard, already verified building blocks in new designs, the verification time can be considerably reduced. This applies to both the HW and SW blocks. Another approach to reducing the device design time is the Transaction Level (TL) Modeling (TLM) /1/. It can be used for HW/SW modeling, co-design and co-verification. Compared to the HW model described at the RTL, TLM uses abstract communication with approximate timing thus enabling faster simulation and verification /4/. Another reason to use TLM is availability of the executable models early in the development cycle. They confirm the expected functionality and can be used for design space exploration and early SW design.

Though the TL models can be easily debugged due to their abstractness, the Assertion-Based Verification (ABV) is added to the Simulation-Based Verification (SBV) to allow for faster and easier debugging. The use of assertions helps pinpointing the bug in the design more than SBV itself.

The focus of the proposed Design and Verification Flow (DVF) is on the HW-supported functionality. It captures all the above improvements, i.e. TLM, SBV and ABV for new Intellectual Property (IP) block design.

There have been several design and verification approaches proposed in literature. Vaumorin et al. /5/ favourise the design flow from a specification through SystemC and RTL to the FPGA implementation. Verification takes place at several stages while RTL design described in the HW Description Language (HDL) is verified by using co-simulation with RTL described in SystemC. No assertions are used.

Wei et al. /2/ describe the design and verification scheme for IEEE 802.15.3 MAC and Carbognani et al. /6/ for ARM AMBA. Both groups of authors present their model in TLM

and RTL with support for the reusable test bench; none uses ABV.

Habibi et al. /7/ propose a verification approach with the SystemC model translated to the AsmL language. Only ABV is used. Assertions are described with the PSL language.

Our focus in this paper will be on a hybrid DVF of an IP block. Our approach was verified on a Universal Serial Bus (USB) host controller design constructed according to /11/.

2 Design and verification flow

In this section, we present a three-step DVF (Fig. 2). It starts with a non-formal specification in a written or verbal form of the model behavior and operating conditions. The model behavior specification is used for model description and test-bench construction and the operating condition specification is used for the functional-coverage metric (FCM) definition.

In the first step, the Verification Environment (VE) is constructed enabling efficient Test-Bench (TB) generation to cover the FCM. Therefore, the first FCM is defined according to the model operating conditions. The impact of well-defined operating conditions on the verification quality and time consumed by it is considerable. If some operating conditions are not captured by the specification, the model is not verified for them. If operating conditions are set wider than necessary, VE performs verification for the non-existent conditions and unnecessarily wastes time for that (Fig. 1).

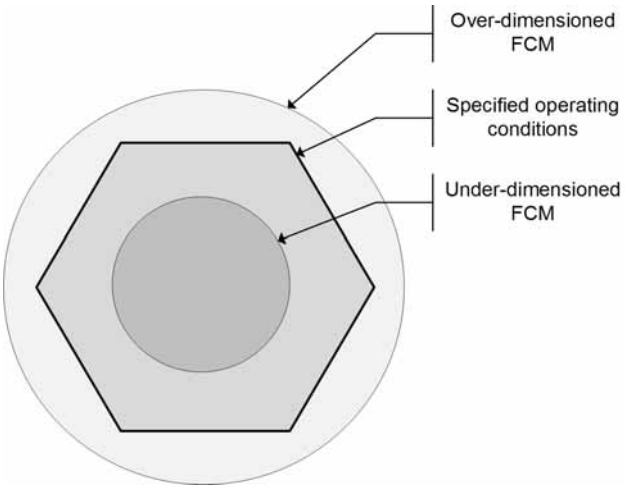


Fig. 1: Model operating conditions space

In the same step, the TB that satisfies the defined FCM is constructed. TB is responsible for generating test vectors that are later applied to the model inputs.

In the following step, the model is translated from the non-formal specification to an executable specification. It is then placed into VE as a Device Under Verification (DUV). DUV and VE are described using TLM which can be employed on many abstraction levels. TLM with an approximate timing (also known as "Programmer's View with Timing" - PVT) was selected for DVF. TLM is therefore approximate-timed and uses abstract transactions for communication. Due to its abstractness, TLM enables a shorter development time and faster simulation.

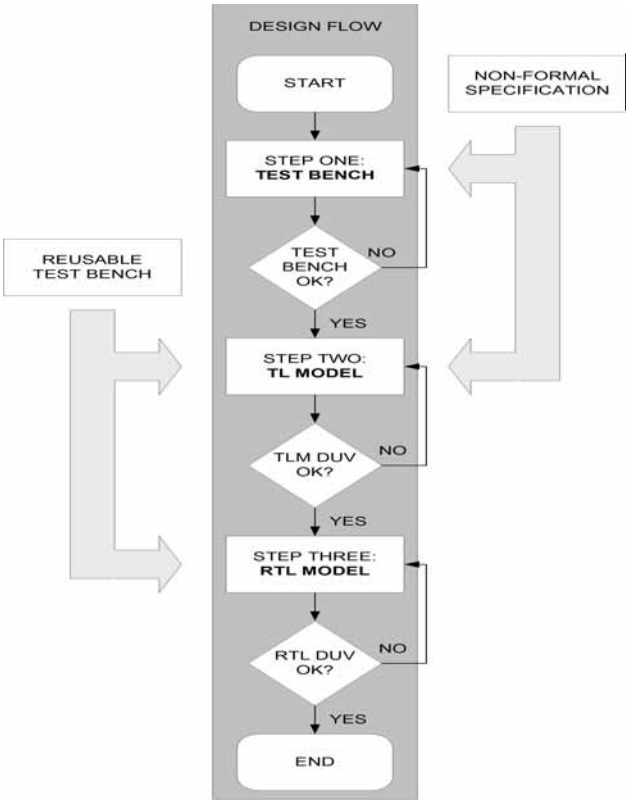


Fig. 2: Design and verification flow

After DUV is placed into VE, the TB composed in the previous step is applied. TL DUV is corrected until it positively passes the test bench.

At the end of the second step, a functionally-verified HW model is available for design-space exploration and early SW development. The TB responses are used as a reference for RTL refinement (golden model).

In order to implement HW with standard tools, the DUV model has to be refined to RTL. This is achieved in the third step. The RTL model is cycle-timed and uses four-state signal logic ('0', 1', Z' or 'X') for communication. In order to reuse TB and the VE on the RTL model, an interface referred to as transactor has to be developed. The transactor adapts the RTL model to the TL environment by translating abstract communication to signal-level communication and provides synchronization. The DUV RTL model is confirmed OK when it positively passes TB.

After using the proposed DVF, the functionally verified RTL model is ready for implementation with standard tools.

2.1 Functional verification

In the next chapters, each step will be explained in detail. The first we shall deal with will be the FCM definition step.

The functional coverage definitions are extracted from the operating conditions specification. FCM states which functions have been exercised. When there are several variables in a function, the problem becomes more complex. The question is whether to test each function with any possible combination of variables or to group the variable values first? An exhaustive verification procedure is much time consuming and therefore expensive. Our solution is to reduce the function coverage space and speed-up the verification process. This allows us to detect the major errors in the design and to reduce the verification time by a few decades, though some of the minor errors might be left undetected. Under the current time-to-market pressures, manufacturers find it hard to perform exhaustive verification (Fig. 3).

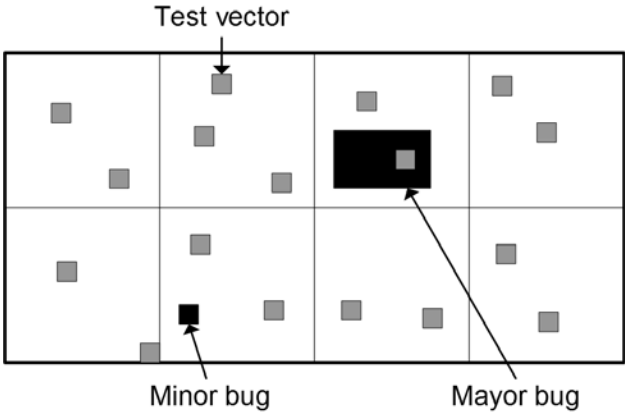


Fig. 3: Verification space

To make it clearer, let's assume the function F(A,B) variables, where A and B can be assigned any value from 0 to 31. The value range is linearly divided into four testing ranges (bags) defined as (0, 7), (8, 15), (16, 23) and (24, 31). Two examples are shown in Fig. 4. If the function parameter is tested within a certain bag, the bag is said to have a hit.

Ranges and numbers of bags of each variable may vary. The number of bags per variable defines the variable resolution. When all the defined bags have at least the predefined number of hits, a full functional coverage is being achieved /9/.

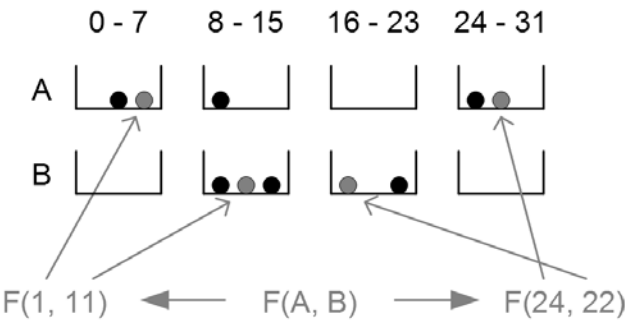


Fig. 4: Function bags

The variable resolutions and the minimum number of hits for a full functional coverage define the verification effectiveness (Fig. 5). When more exhaustive verification is required, higher resolution is selected and more combinations of test vectors are run. On the other hand, when fast and consequently less exhaustive verification is required, lower resolution is selected and fewer combinations of test vectors are run.

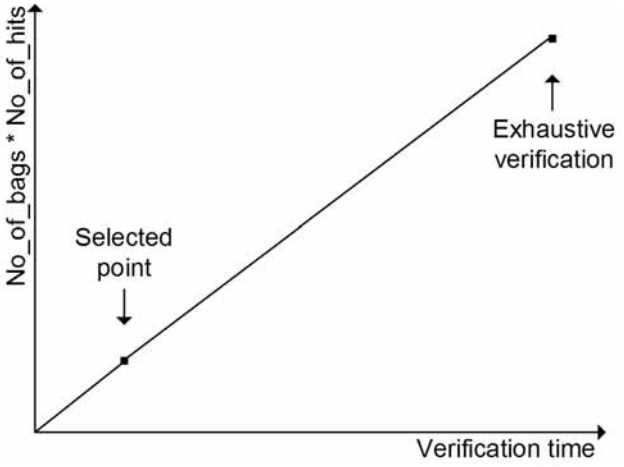


Fig. 5: Parameter selection graph

Since modern devices consist mostly of programmable logic, bug-fixes can be issued in case of missed bugs during first verification.

Verification of DUV runs at TL. The VE in Fig. 6 consists of a test controller, simulator, monitor, evaluator, master (DUV) and multiple slaves when required. Verification runs as follows. The stimulator sends a request to DUV. DUV processes the given request by sending new requests to the slave. After the slave returns response, DUV returns it to the stimulator (Fig. 7). The monitor monitors the stimulator and DUV transactions and forwards them to the evaluator. The evaluator checks correctness of these requests and measures the verification coverage. The test controller controls the environment according to the TB /4/.

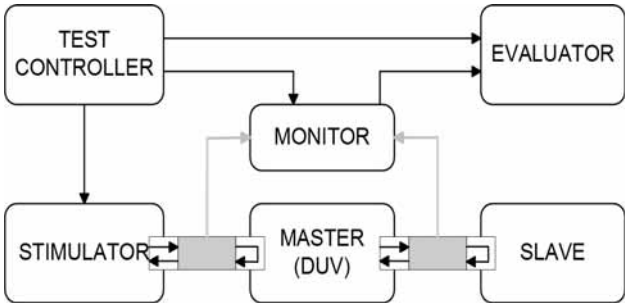


Fig. 6: Verification environment

Test vectors are generated by using the SystemC Verification (SCV) library which enables Constrained Randomization (CR). It shortens the verification time when the pro-

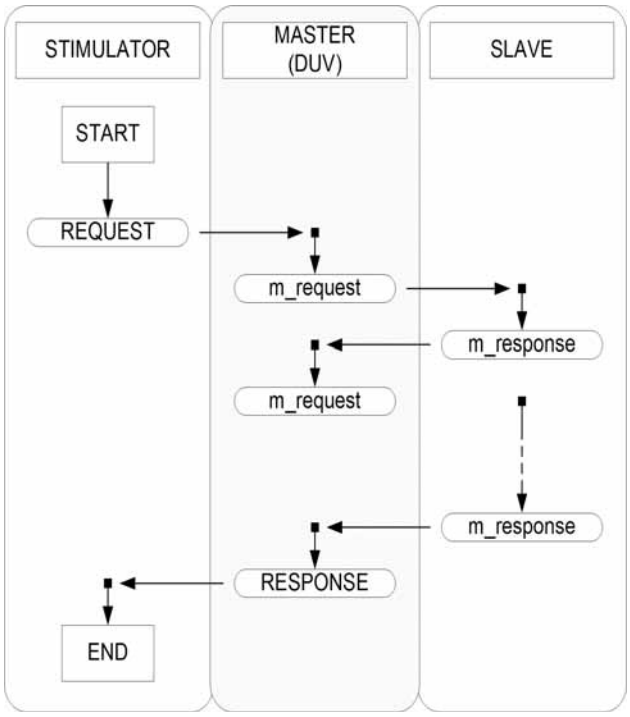


Fig. 7: Transaction-time diagram

posed FCM needs to be used /10/. CR makes it possible to limit test vector generation to a certain range and also to define inter-variable dependencies.

DVF verification consists of SBV and ABV. SBV is responsible for FCM definition, TB construction and storage of DUV requests. It is up to the designer to detect bugs in these requests. ABV is on the other hand responsible for assertion definition which enables automatic bug detection.

Since test vector generation is random, the number of test vectors for full functional coverage (test vector set size) varies. Different test vector sets lead to a full functional coverage. Test vector sets can be combined in the distribution graph.

2.2 TLM & SystemC

SystemC is a C++ library which enables HW modeling. Since it originates from the SW world and supports the HW description, it represents the basis for HW/SW co-design and co-verification.

Similar to other HW developing languages, SystemC also supports architectural design. Modules can be described and connected via input and output ports. The base for the module description is class SC_MODULE. This class consists of several ports, processes and other modules.

SystemC enables separate modeling for computation and communication. Computation is modeled by processes described using either SC_THREAD or SC_METHOD. The computer thread-like processes are described using SC_THREAD. They run all the time and can only be delayed. The RTL-like processes are described using SC_METHOD and are executed only on a trigger event. The trigger events like clock signal are listed in the sensitivity list.

Communication is modeled by channels connected between two ports. The data transmitted over the channels can be very simple like a Boolean value or complex like a second video sample /1, 7/.

Separate modeling of computation and communication is the basis for TLM. The only process running inside the TL model is described using SC_THREAD. From this process, supporting functions are called. The process is used for communication to the VE, while the supporting functions describe functionality and transact with slaves.

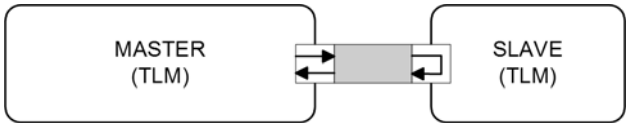


Fig. 8: TLM block scheme

TLM uses request - response transactions for communication between modules. Transactions are abstract data types described by using the custom C++ class. Communication runs as follows. An initiator (master) sends a request to a target and the target (slave) returns the response (Fig. 8).

The process time can be measured in real time (seconds) or in clock cycles. The more the types used in modeling are abstract, the less time is required to simulate the design.

As any other C++ class, the TL model too, is instantiated as an object on the top level and connected to a slave.

2.3 RTL refinement

The RTL model can be described using SystemC or standard HDL (VHDL, Verilog).

The SystemC RTL model is refined from the TL model as shown in Fig. 10. Since the architecture has already been defined at TL, the RTL refinement requires less effort than the RTL design from the start. During refinement, supporting functions are extracted and modeled as SC_METHOD processes. In contrast to the TL model, where transactions are used for communication, the RTL model communicates by 4-level signal buses. When describing the RTL model, each register or the finite-state machine (FSM) has to be described using its own SC_METHOD process.

The RTL model in HDL is identical to the RTL model in SystemC. There are few differences though; when the SystemC model is simulated using any standard compiler, a mixed-language environment is required for HDL model simulation within the SystemC VE (Aldec Riviera, Matlab).

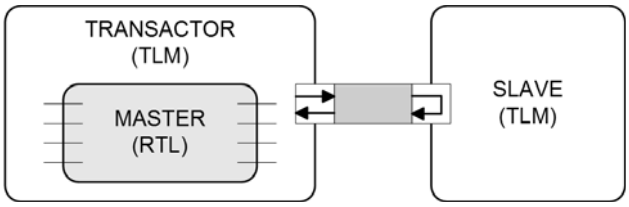


Fig. 9: RTL modeling with transactor

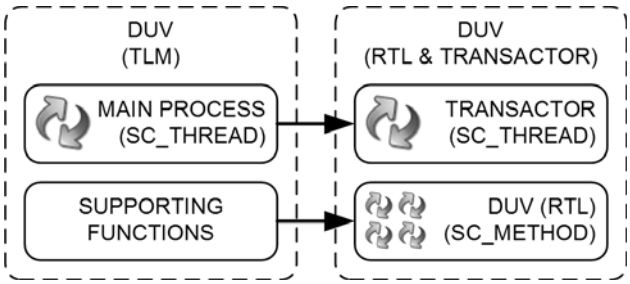


Fig. 10: TL to RTL refinement

In order to connect the less abstract cycle-accurate RTL model to the TL TB, a transactor should be constructed (Fig. 9). The transactor is a process, modeled as an SC_THREAD, originally used as the process in the TL model (Fig. 10). Its duty is to translate requests from the VE into the signal level, signals into requests for a slave, responses from the slave into the signal level, and finally signals into responses for the VE.

3 Universal Serial Bus

USB is a serial-bus standard to interface devices. Some of its features are: fast data transfer, plug&play capabilities and providing power to low-power devices. It is intended to help retire all legacy varieties of serial and parallel ports. USB connects computer peripherals such as mouse devices, keyboards, PDAs, scanners, digital cameras, printers, personal media players, and flash drives. For many of those devices USB has become the standard connection method. USB was originally designed for personal computers, but it has become commonplace on other devices such as PDAs and video game consoles. In 2004, there were about 1 billion USB devices in the world.

USB specification 1.1 was released in 1996. It includes "low-speed" and "full-speed" data rates. A new specification 2.0 was released in 2000 with some additional features. Among them is the "high-speed" data transfer.

3.1 USB host controller

The USB system has a star topology with a host in the center. The host connects up to 127 functions (devices), with the hub acting as a host for the next level (Fig. 11). Up to five levels are supported. There are several types of devices: hub, human-interface device, mass-storage device, printer, audio, video devices and others. The USB system has a master-slave organization where the master initiates the communication to which the slave can respond.

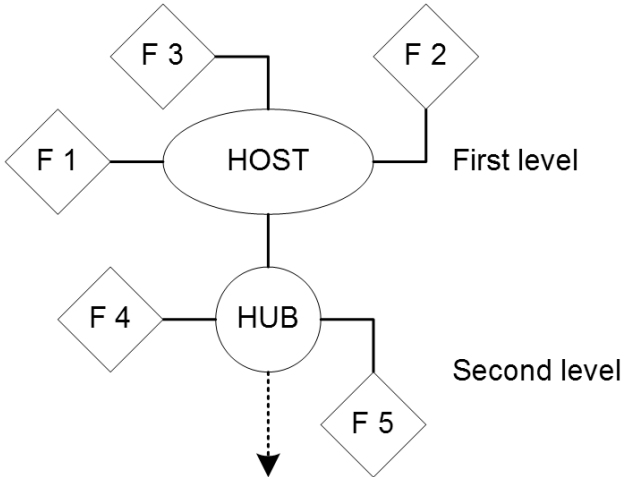


Fig. 11: USB system topography

In the startup operation the host first performs the enumeration of the connected devices and assigns addresses. It then reads the device properties like the device type and number of endpoints. Thereupon, it is ready to transfer the data to or from the devices.

The means to transfer the data are the device endpoints. To enable communication, the host controller creates virtual pipes to these endpoints.

3.2 USB host controller layers

The USB host controller consists of several layers of functionality. They can be implemented either in SW or HW or in a combination of both.

The layers of the USB host controller can be presented in many different ways. We will define layers as shown in Fig. 12. The first layer takes care of the USB electrical part like serialization and new device detection. The second layer composes the packet where the data is added the synchronization field (Sync) and "end of packet" (EOP). The third layer generates the packet data and makes the cyclic redundancy check (CRC) when required. The fourth layer provides different types of transfers. The fifth layer is the functionality layer for connection to the endpoints. The sixth and upper layers are application layers.

When some data from a certain function is needed for an application, a pipe is established. The Pipe layer then calls a specific transfer type, the Packet layer prepares the data required for the operation and adds CRC. The packet is completed with Sync and EOP in the Frame layer. The Base layer serializes the complete data and sends it to the slaves.

3.3 USB host controller model

Our model doesn't include all the features described in the specification, but only certain features from the USB

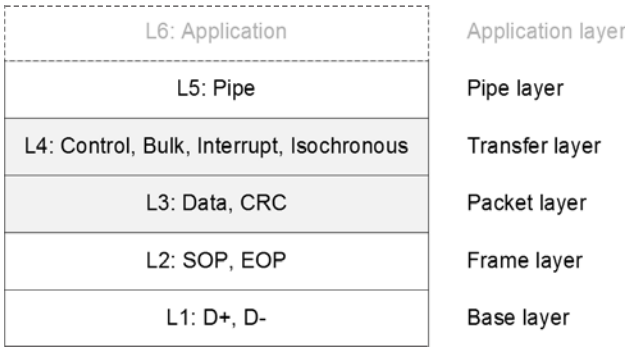


Fig. 12: USB host controller layers

specification 1.1. The model was developed with intention to show how only some selected features of certain layers can be modeled. It allows for an easy upgrading with other features and includes parts of the Packet and Transfer layers (L3 and L4) given in Chapter 8 ("Protocol Layer") of the specification /12/.

First, let's take a look at the common USB packet fields, i.e. the Sync field, Packet ID (PID) field, address (ADDR) field, endpoint (ENDP) field, data (DATA) field, CRC field and EOP field. The Sync field is responsible for synchronization, the PID field identifies the packet, the ADDR field indicates the device the packet is designated for, the ENDP field specifies the device endpoint, the DATA field contains the data, the CRC field detects errors and the EOP field defines the end of the packet.

The Packet layer distinguishes between the Token, Data and Handshake packets (Fig. 13). There are some other packet types defined in the USB specification which are beyond the frame of this paper. Each packet type has a different structure. The Token packet consists of the Sync, PID, ADDR, ENDP, CRC and EOP fields, the Data packet of the Sync, PID, DATA, CRC and EOP fields and the Handshake packet of the Sync, PID and EOP fields. Since the Sync and EOP fields occupy every packet, we extracted their assembly from the Packet to a separate Frame layer and will not model them.

Two types of CRC are required for the USB packet. The first is the CRC5 type and is used for CRC calculation of the Token packet. The second one is the standard CRC16 type and is used in CRC calculation of the Data packet. CRC will not be calculated by our model. The abstract model will specify only the CRC type.

As already mentioned above, PID is used to identify the packet. It first identifies whether the packet is of the Token, Data or the Handshake type. It then distinguishes between the SETUP, IN, OUT and the SOF Token packet type, the DATA0, DATA1, DATA2 and the MDATA Data packet type, and the ACK, NAK, STALL and the NYET Handshake packet type. PID occupies eight bits of each packet. PID will be modeled on the abstract level with the PID type only. The reader can find out more on CRC calculation and PID in /11/.

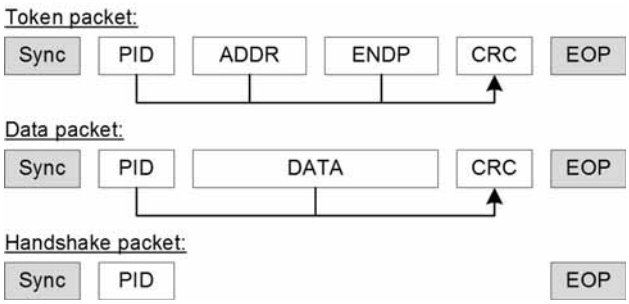


Fig. 13: USB packets and fields

The Transfer layer distinguishes between the four data-flow types: Control, Bulk, Interrupt and Isochronous transfer. The Control transfer is used for properties and status recognition, the Bulk transfer for large data, the Interrupt transfer for interrupts, and the Isochronous transfer for audio and video stream.

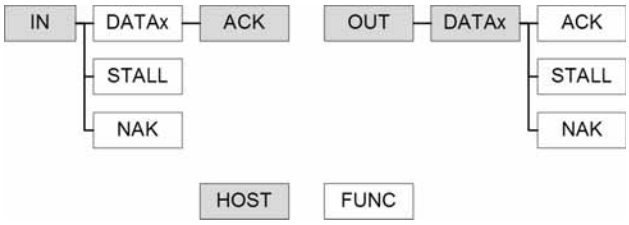


Fig. 14: USB Bulk transfers

Each of the data-flow types has a defined data flow. Since we only modeled the Bulk transfer, we will focus on this data-flow type alone. Bulk transfer can be the IN (read) or OUT (write) transfer. The IN Bulk transfer starts with the host issuing the IN Token packet. The addressed function responds with the Data packet containing the data from the target endpoint. In this case, the host responds with the ACK Handshake packet. In case the addressed function is busy, it responds with the STALL or NAK Handshake packet. The OUT Bulk transfer starts with the host issuing the OUT Token packet and a following Data packet. The function responds with the ACK Handshake packet on a successful reception and with the STALL or NAK one on an unsuccessful reception (Fig. 14). To enable better presentation, our model will only support the DATA0 and the DATA1 Data packet type and the ACK Handshake type. As said above, further functionality can be added later.

3.4 USB function model

The USB function model combines 128 functions with 16 endpoints each. In order to verify the complete specter of functions, the function model responds to all addresses and all endpoints.

3.5 Verification flow

During DVF, the stimulator is used to generate test vectors which model requests from a higher (Pipe) layer. They con-

sist of variables fAddr, fData, fLength, fEndPoint, fDirection and fDataSlot (Fig. 15). The fAddr variable targets the 7-bit ADDR packet field and varies from 0 to 127. The fEndPoint variable targets the 4-bit ENDP packet field and varies from 0 to 15. The fData variable targets the DATA field whose size can at the "full-speed" USB Bulk transfer vary from 0 to 1023 bytes of data. The fData variable is modeled with 1024 bytes of data while the actual size of the transmitted data is defined with fLength. The fDirection and the fDataSlot variables can hold values 0 or 1. For the first variable, value 0 specifies the IN transfer and value 1 the OUT transfer. For the second variable, value 0 specifies the use of the DATA0 transfer and value 1 the use of the DATA1 transfer.

```
class Function {
public:
    unsigned fAddr;
    unsigned char fData[1024];
    unsigned fLength;
    unsigned fEndPoint;
    bool fDirection;
    bool fDataSlot;

    // Constructors
    Function() {}
};
```

Fig. 15: Test vector model

In chapter 3.1, we showed that FCM can have different variable resolutions and different amounts of the required hits per bag for a full functional coverage. We stated that the verification time depends upon the selection of these parameters. Therefore, we ran four different combinations of these two parameters. We selected two different variable resolutions and ran them with a minimum of 100 and 200 hits per bag. The different resolutions are presented in Table 1. The different FCMs are labeled FCM1A (1 for 100 hits, A for resolution A), FCM2A, FCM1B and FCM2B.

Each of the four defined FCMs has many different test vector sets that lead to a full functional coverage. From these test vector sets, the average test vector set size can be calculated. Also, distribution of the test vector set sizes for each FCM can be presented.

Table 1: Resolutions of variables

Variable	Range	Resolution A	Resolution B
fAddr	0 – 127	8	16
fLength	0 – 1023	8	16
fEndPoint	0 – 15	4	8
fDirection	0, 1	2	2
fDataSlot	0, 1	2	2

```
class Packet {
public:
    enum PIDType {
        USB_NULL, USB_OUT, USB_IN,
        USB_SOF, USB_SETUP, USB_DATA0,
        USB_DATA1, USB_DATA2, USB_MDATA,
        USB_ACK, USB_NAK, USB_STALL,
        USB_NYET};
    enum CRCType {CRC5, CRC16};

    unsigned pMask [5];
    PIDType pPID;
    unsigned pAddr;
    unsigned char pData[1024];
    unsigned pEndP;
    CRCType pCRC;
    unsigned pLength;

    // Constructors
    Packet() : pPID(USB_NULL) {}
};
```

Fig. 16: Abstract packet model

The USB host and USB function communicate via the abstract packet. The packet fields are modeled using the following variables: pMask, pPID, pAddr, pData, pEndP and pCRC and pLength (Fig. 16). The pMask variable masks the presence of the major five variables and the pLength variable defines the actual size of the pData variable. The abstract packets are stored to a file and represent a golden reference of SBV. Due to the large amount of the overall data, the data is modeled using only its length (pLength). The complete data (pData) is verified using ABV.

For ABV, five assertions were implemented in the VE. The assertion monitor checks for correct packet contents, sequence of packets, Address, Data, Slot & Length and Direction & Destination. Assertions are checked on the packet sent from the host and are logged for further analysis. Assertions that have not been met report an error.

3.6 Results

Using the proposed methodology, the design of the USB host controller models was efficient. We spent 16 hours for the VE description, 30 hours for the TL model description and 12 hours for the RTL model refinement.

The minimum amount of test vectors for full functional coverage (minimum test vector set size) can be calculated by multiplication of the required number of hits per bag and the maximum number of bags per variable of a defined FCM. The numbers are 800 for FCM1A, 1600 for FCM1B and FCM2A and 3200 for FCM2B. After exercising 1000 random test vector sets, we calculated the average test vector set sizes. The results are presented in Table 2. We

also compared deviations of the average test vector set sizes from the minimum test vector set sizes in percents.

Table 2: Average test vector set sizes

FCM	Average size	Deviation
FCM1A	941	17.6%
FCM2A	1796	12.2%
FCM1B	1941	21.3%
FCM2A	3684	15.1%

Distribution for FCM1A is shown in Fig. 17. The values were grouped for better presentation. The step for the x axis is 2% of the minimum test vector set size; for FCM1A the number is 16. The y axis shows the number of test vectors within the group. As seen from this particular case, there were 129 out of 1000 test vector set sizes between 896 and 911 (labeled "904").

To have results from different FCMs better analyzed, we normalized the results to the minimum test vector set size. The normalized results are shown in Fig. 18. From these results and those from Table 2, FCMs with resolution A converge faster than FCMs with resolution B. Also, FCMs with the required 200 hits per bag converge faster than FCMs with the required 100 hits per bag. We can conclude that FCM2A finishes faster than FCM1B, although it might miss more bugs than the latter. For better interpretation of these results, a more in-depth study would be needed since they depend upon SCV CR.

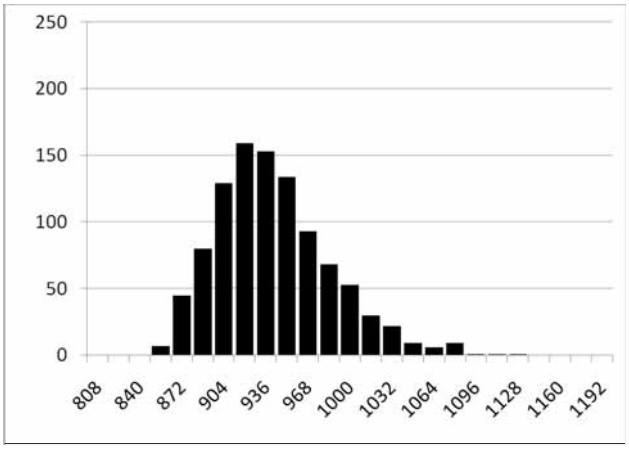


Fig. 17: FCM1A distribution chart

By using DVF, we created fast-simulating TL and RTL models of the USB host controller with the described features. The RTL models are described by using SystemC and VHDL. Table 3 shows that the TL model in average simulates six times faster than the RTL models. Since the design used in our verification was simple, the absolute time difference is small. If we added all the features from the USB specification 2.0, the verification times would be much longer and the difference would be more obvious.

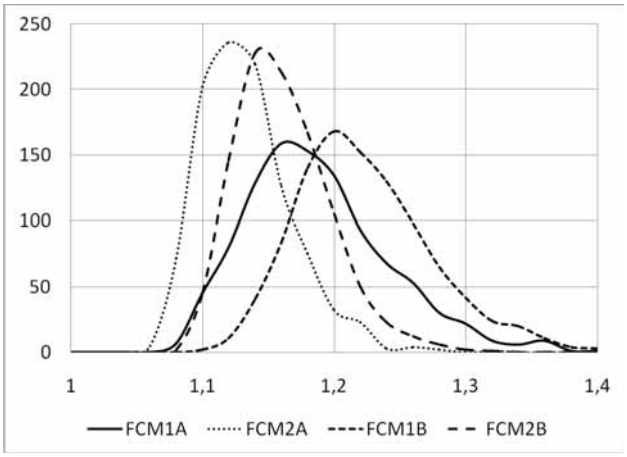


Fig. 18: Normalized distribution chart

Table 3: Average verification times of FCM1A

TLM	RTL	
	SystemC	VHDL
1,2s	6,7s	7,8s

Compared to /13/, the USB host controller shows several differences. Its verification converges much faster than that of /13/. The reason for it can be found in the inter-independent variables of the USB host controller, while those of /13/ are inter-dependent. This is a great challenge for SCV CR. Another difference is in the scope of the design verified by using SBV compared to the share verified by using ABV. The USB host controller design is a good example showing how both SBV and ABV can be efficiently used for verification, whereas using ABV for verification of /13/ is needless.

For simulation purposes, we used a computer with the Intel Pentium M 740 processor running at 1730 MHz with 1 GB of memory. The wall time was used for time measurement.

4 Conclusion and future work

In this paper we propose a hybrid three-step design and verification flow. We prove correctness of our approach on a simplified USB host controller. Further functionality can be added and verified at any time. It can be realized by either SW or HW or a combination of both. For instance, the Transfer layer was implemented in SW, and the Packet layer in HW. For realization in SW, the RTL refinement is not required.

One of the reasons to adopt the proposed approach is to shorten the design time for the verified TL model, which in our case study was 30 hours. Another 16 hours were required for the VE and another 12 hours for the RTL refinement.

We showed the free choice of FCM at the expense of possible missed bugs. Also, ABV was very helpful at pinpointing the bugs by means of which the design took us less time.

The use of the proposed DVF requires no special tools. TL and SystemC RTL models description and simulation are made only by a free C++ compiler. Since no SystemC implementation tool is currently available, the VHDL RTL model in combination with mixed-language simulators is also required.

In our case study we show that the TL model simulates six times faster than the RTL model. The simulation speed becomes more relevant at more complex models that take more time to simulate.

Future work will focus on improving the proposed coverage metric and convergence towards full coverage. The USB host controller model will be added further functionality in order to explore the described FCMs with more complex models. Other nonlinear FCMs will be studied so as to optimize the verification.

5 References

/1/ S. Swan, "SystemC Transaction Level Models and RTL Verification", 43rd ACM/IEEE Design Automation Conference, pp. 90-92, 2006.

/2/ Y. Wei, H. Guanghui, X. Ningyi, Z. Zucheng, "SystemC Transaction Level Modeling and Verification of IEEE 802.15.3 MAC", International Conference on Communications, Circuits and Systems Proceedings, pp. 2554-2558, 2006.

/3/ S. Tasiran, K. Keutzer, "Coverage metrics for functional validation of hardware designs", IEEE Design & Test of Computers, vol. 18, no. 4, pp. 36-45, 2001.

/4/ F. Ghenassia, "Transaction Level Modeling with SystemC", Springer, 2005.

/5/ E. Vaumorin, T. Romanteau, "From Behavioral to RTL Design Flow in SystemC", electronic file available at https://www.systemc.org/, 2004.

/6/ F. Carbognani, C. K. Lennard, C. N. Ip, et al, "Qualifying precision of abstract SystemC models using the SystemC Verification Standard", Design, Automation and Test in Europe Conference and Exhibition, pp. 88-94, 2003.

/7/ A. Habibi, S. Tahar, "Design and Verification of SystemC Transaction-Level Models", Very Large Scale Integration (VLSI) Systems, IEEE Transactions on, vol. 14, no. 1, pp. 57 - 68, 2005.

/8/ T. Grotker, S. Liao, G. Martin, S. Swan, "System Design with SystemC", Kluwer Academic Publishers, 2002.

/9/ R. Siegmund, U. Hensel, A. Herrholz, et al, "A Functional Coverage Prototype for SystemC-based Verification of Chipset Designs", 9th European SystemC Users Group Meeting, electronic file available at http://www-ti.informatik.uni-tuebingen.de/~systemc/, 2004.

/10/ Members of the SystemC Verification Working Group, "SystemC Verification Standard Specification", electronic file available at https://www.systemc.org/, 2003.

/11/ USB Specification 2.0, electronic file available at http://www.usb.org/developers/ docs/, 2000.

/12/ C. Peacock, "USB in a Nutshell", electronic file available at http://www.beyondlogic.org/ usbnutshell/usb-in-a-nutshell.pdf, 2002.

/13/ P. Puhar, A. Žemva, "Simulation-based functional verification of a video processing IP block", 43th MIDEM Conference Proceedings, pp. 171-176, 2007.

Primož Puhar, B.Sc.
LEA d.o.o.
Finžgarjeva 1A, 4248 Lesce
primoz.puhar@lea.si

Prof. Dr. Andrej Žemva,
University of Ljubljana
Faculty of Electrical Engineering,
Tržaška 25, 1000 Ljubljana, Slovenia

Prispelo (Arrived): 04.02.08 *Sprejeto (Accepted):* 28.5.08

IZVEDBA REKURZIVNIH DIGITALNIH SIT S PLC KRMILNIKOM

Aleksandar Dodič[1], Rudolf Babič[2]

[1]Lučka uprava Rijeka, Hrvatska
[2]Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko, Maribor, Slovenija

Kjučne besede: digitalna obdelava signalov, digitalni filtri, rekurzivni filtri, eliptični filtri, PLC krmilnik

Izvleček: V delu je opisan postopek načrtovanja in izvedbe rekurzivnih digitalnih filtrov prve, druge in tretje stopnje s programirnim logičnim krmilnikom. Osnovni namen je izločanje neželenih signalov nizkih frekvenc pri obdelavi analognih vrednosti v sistemih industrijskega krmiljenja. Za praktično izvedbo smo uporabili PLC krmilnik SIEMENS S7-315-2DP. Predvideli smo nizkoprepustno frekvenčno karakteristiko, ki smo jo zagotovili z eliptičnim filtrom. Pri frekvenci vzorčenja 50 Hz smo mejno frekvenco izbrali pri 3 Hz. S popolnim zapisom vrednosti koeficientov filtra smo dosegli pričakovano slabljenje motilnih signalov v zapornem frekvenčnem področju 49 dB. Primerjavo med simulacijsko izvedbo, dobljeno v MATLAB programskem okolju in praktično izvedbo digitalnega filtra druge in tretje stopnje s PLC krmilnikom smo ocenili s primerjavo impulznih odzivov in njihovega odstopanja. Digitalni filter tretje stopnje smo načrtali v kaskadni obliki s strukturama prve in druge stopnje. Pri digitalnem filtru druge stopnje smo podrobno analizirali odstopanje impulznega odziva izvedenega filtra s PLC krmilnikom pri zapisu koeficientov popolni in pri zapisu v skrajšani obliki. Amplitudni frekvenčni karakteristiki sta podani za drugo in tretjo stopnjo. Amplitudna frekvenčna karakteristika je ustreza postavljenim zahtevam. Slabljenje 40 dB dosežemo že pri dvakratni mejni frekvenci, kar je bistveno bolje kot pri filtru druge stopnje.

IIR Digital Filter Implementation With PLC Controler

Key words: Digital signal processing, Digital filters, IIR filter, Eliptic filter, PLC controller

Abstract: In this paper the design and realization of first, second and third order recursive digital filter (IIR digital filter) with PLC controller is described. The purpose of the filter structure is eliminating unwanted signals at processing of analog signals in industrial proccess automation with PLC controllers. In practical realization the Siemens S7-315-2DP PLC controller is used. In this application the eliptical filter is chosen, cutt-off frequency was 3 Hz and sampling frequency was 50 Hz. With format long values of filter coefficients we obtain the expected stop band attenuation of 40 dB as is designed with MATLAB software. For the comparision of PLC controller realization of digital filter and simulated structure with MATLAB, the obtained values of impulse response are used. The results are sumarised in Table 2 for second order digital filter for format long and shortened mode format of coefficients entry and in Table 5 for third order digital filter only for format long coefficients. In fig. 1 the magnitude response of the empiric digital filter is shown for different values of smoothing parameter **k.** In fig. 2, 3 and 6 the masks as a part of the organization block of Siemens PLC controler for entry of the coefficients and input variables for the first, second an third order digital filter structure are presented. In fig. 4 the magnitude frequency response of 2^{nd} order low pass eliptic IIR filter with cut-off frequency of 3 Hz is shown. The third order recursive filter is designed as caskaded form with cascade structures of the first and the second order. For this operation the MATLAB FDA toolbox is used for second order structure (SOS structure) calculation. The first and second order structures for eliptic recursive digital filter the SOS matric is presented in equation (10) and in fig. 5 the general cascade realization of two structures recursive digital filter for third and also for fourth order digital filters is shown. The impulse response and the magnitude frequency response of 3^{rd} order eliptic IIR digital filter with cut off frequency of 3 Hz are shown in Fig. 7 and 8 respectively. In comparison to the second order digital filter where the attenuation of 40 dB is obtained at 15 Hz with third order digital filter structure the attenuation of 40 dB is obtained at 7 Hz.

1. Uvod

Obdelava analognih signalov je pogosta naloga v skoraj v vsakem avtomatizacijskem procesu katerega obdelujemo s PLC krmilnikom. Analogne vhodne vrednosti, ki jih dobimo na vhodu ali kot rezultate delnih izračunov, so praviloma spremenjene zaradi raznih motenj. Tako je večkrat potrebno z dodatno opremo, programsko ali aparaturno, odpraviti motnje, da bo proces katerega krmilimo, potekal brezhibno.

Ker uporabljeni senzorji in pretvorniki ne ločijo motilnih signalov od koristnih, dobi krmilni sistem poleg koristnih signalov še neželene motilne signale. Motnje so posebej nevarne, če se frekvenca ujema z mehansko resonančno frekvenco katere od naprav v sistemu. V takih primerih postane krmilje popolnoma neuporabno. Motnje lahko odpravimo na več načinov: z analognim filtrom na vhodu krmilnika ali s programsko izvedenim digitalnim filtrom znotraj PLC krmilnika.

Analogni filter predstavlja preprost način izločanja motilnih signalov. Največ se uporablja RC vezje, včasih zadostuje le kondenzator C, saj je upornost R običajno že zajeta kot porazdeljena upornost v dolžini kablov, upornosti kontaktov, ter v notranji upornosti pretvornika in vira signala. Težave nastopajo, če se spremeni upornost vhodne zanke, saj se s tem spremeni tudi mejna frekvenca. Razlike so prisotne tudi pri uporabi napetostnih in tokovnih vhodov v PLC krmilnik. Pri tokovnih vhodih, ko so prisotne manjše upornosti, motnje v vhodni zanki še težje odpravimo. Pri tokovno krmiljenih virih, potrebujemo za glajenje tudi tuljave, ki pa se ne uporabljajo tako pogosto kot kondenzatorji. Zato analogno filtriranje v takih primerih ni običajno. Analogni filter postane tudi povsem neuporaben, ko bi hoteli izločiti motnje v signalih, ki predstavljajo rezultat vmesnih izračunov več različnih spremenljivk.

V članku je opisan način filtriranja z digitalnim filtrom. Digitalni filter je programsko izveden v PLC krmilniku, ki se upo-

rablja za krmiljenje osnovnega industrijskega procesa. Preproste rešitve uporabljajo algoritme za glajenje, ki delujejo kot nerekurzivni filter in se uporabljajo po iskustvenih navodilih. V našem primeru pa smo načrtali in izvedli rekurzivno obliko filtra druge in tretje stopnje z natančno določenimi parametri povezanimi z mejno frekvenco, ojačenjem v prepustnem frekvenčnem področju in slabljenjem v zapornem frekvenčnem področju. Koeficiente filtra smo izračunali s pomočjo programskega paketa MATLAB.

2. Osnovne digitalne filtrske strukture

Izhodišče za oblikovanje filtrske strukture sta sistemska funkcija H(z) in diferenčna enačba. Za vsak linearni, časovno neodvisni diskretni sistem obstaja več enakovrednih struktur. Strukture so lahko kanonične in nekanonične oblike. Za filtrsko strukturo pravimo, da je kanonične oblike, če je število zakasnilnih členov natančno enako stopnji filtra /1/. Nekanonične strukture običajno uporabljamo pri procesorjih z aritmetiko s fiksno vejico, ker imajo le en seštevalnik. Za izvedbo digitalnih filtrov 1. in 2. stopnje s PLC krmilnikom lahko uporabimo direktno kanonično ali direktno transponirano kanonično strukturo. Obe sta enakovredni. Razlika je le v številu seštevalnikov in množilnikov potrebnih za realizacijo. Za strukture višjih stopenj praviloma uporabljamo kaskadno /2/ ali vzporedno povezavo struktur 1. in 2. stopnje ne glede ali gre za rekurzivno ali nerekurzivno izvedbo.

3. PLC krmilnik

Programirni logični krmilnik (PLC) je namensko mikroračunalniško vezje, ki se uporablja na področju industrijske avtomatizacije in daljinskega vodenja procesov. V začetku so bili signali v PLC-ju le digitalni. Novejši PLC krmilniki obdelujejo digitalne in analogne signale. To pomeni, da je z njimi možno programsko izvesti tudi digitalno filtriranje. Tako izveden filter je ekvivalenten dejanskemu, fizičnemu filtru in predstavlja samostojno enoto.

3.1 Takt

Čas izvajanja enega takta je zelo pomemben kriterij za oceno hitrosti PLC krmilnika. Njegov podatek podaja informacijo o času, potrebnem za opravilo vseh funkcij v določenem času. Izvajanje programa PLC krmilnika je ciklično, to pomeni, da se spremembe vrednosti spremenljivk znotraj enega takta programa ne bodo upoštevale in izvedle.

3.2 Ciklična prekinitev

Če želimo zelo hitro obdelavo nekaterih signalov ali funkcij, npr. alarmov, PID regulatorjev itn., lahko uporabimo različne sheme izvajanja programa s pomočjo prekinitvenih funkcij. Te nam omogočajo direkten pristop do vhodnih in izhodnih spremenljivk in tudi sprotno obdelavo, ne glede na osnovni takt PLC programa. Način pristopa do teh funkcij je različen pri različnih proizvajalcih PLC-jev in tudi programske opreme. Običajno obstajajo specialne procedure oz. bloki, ki omogočajo prekinitve krmiljene z dogodki ali pa so prekinitve krmiljene časovno.

Ciklično prekinitev lahko izvedemo na več načinov, pri vezju S7-300/400 je zagotovljena z organizacijskim blokom OB35. Obstajata dva možna načina: v OB35 generiramo impulz katerega v glavnem programu uporabimo kot uro, ki določa periodo vzorčenja ali v OB35 pokličemo funkcijo v kateri je implementiran digitalni filter. V tem primeru je digitalni filter izveden kar v organizacijskem bljoku OB35.

Interval prekinitve OB35 mora biti daljši od časa potrebnega za obdelavo funkcije, ki jo pokličemo iz OB35. Če ta pogoj ni izpolnjen, vezje S7-300/400 kliče prekinitev za obdelavo napake določene z organizacijskim blokom OB80 /3/.

4. Pogosto uporabljeni digitalni filtri v PLC krmilniku

Izvedba digitalnih filtrov v PLC krmilniku je dodatna naloga za inženirje, ki načrtujejo sisteme s PLC krmilniki, saj večine motilnih signalov ni mogoče odpraviti s preprostimi analognimi filtrskimi vezji. Ker pa običajno ti inženirji nimajo dovolj znanja o digitalni obdelavi signalov, je v priročnikih mogoče najti nekaj vrst digitalnih filtrov, ki se uporabljajo za manj zahtevne aplikacije. Ti digitalni filtri so vedno nizkoprepustni in njihovo delovanje je določeno le s t.i. koeficientom glajenja. V nadaljevanju bomo opisali najpogosteje uporabne rešitve.

4.1 Gladilni filter s premikajočim povprečenjem

To je nerekurzivni filter, ki se uporablja za glajenje. Osnovna oblika /4/ je podana z

$$y[n]=\frac{1}{M}\cdot\sum_{k=0}^{M-1}x[n-k], \qquad (1)$$

kjer je M število uporabljenih preteklih vhodnih vrednosti. Izhodni signal predstavlja povprečje M preteklih vhodnih vrednosti in ima zakasnitev (M-1)/2 taktov. Čim večji je M, tem boljše je glajenje. Ker so v praktičnih aplikacijah prisotne omejitve vezane na zakasnitev, se za M izbirajo vrednosti od 50 do 200.

Prenosno funkcijo dobimo po preureditvi (1) v obliki:

$$H(z)=\frac{1}{M}\sum_{n=0}^{M-1}z^{-n}=\frac{1}{M}\cdot\frac{1-z^{-M}}{1-z^{-1}}=\frac{z^{M}-1}{M\left[z^{M-1}(z-1)\right]} \qquad (2)$$

4.2 Empirični filter

To je izkustvena oblika digitalnega filtra v PLC aplikacijah. Uporabniku običajno ni potrebno poznati teoretičnega ozadja teorije digitalne obdelave signalov in načrtovanja filtrov. Gre za rekurzivno nizkopasovno filtersko strukturo prve stopnje, brez natančno določene frekvence vzorčenja /5/.

Uporaba je omejena na glajenje vhodnega signala. Sprotna izhodna vrednost je določena z nastavkom:

$$Y(n) = \frac{(Y(n-1)\cdot k) + X(n)\cdot(100 - k)}{100} \quad (3)$$

Lastnosti so določene s parametrom k: pri k=0 ni filtriranja, pri k=100 pa dosežemo max. filtriranje

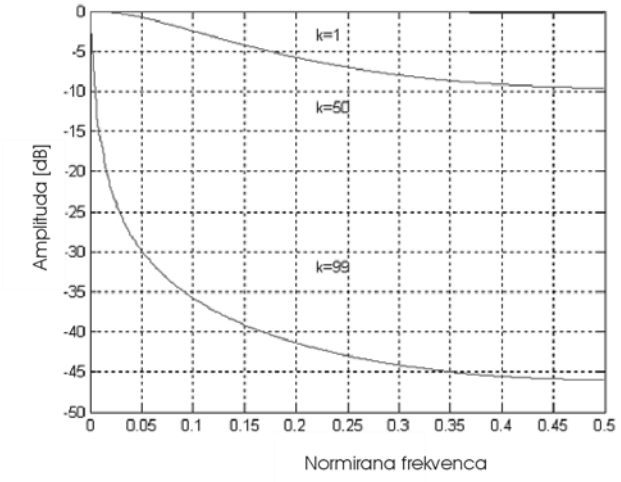
k=0 → *Y(n) = X(n)* - vhod direktno na izhod, ni vpliva prejšnje vrednost izhoda

k=99 → *Y(n) = 0.99Y(n-1)+0.01X(n)* - max filtriranje, vpliv vhodnega signala najmanjši

Prenosna funkcija ima za k=1 naslednjo obliko:

$$H(z) = \frac{99\cdot z}{100\cdot z - 1}, \quad (4)$$

Na sliki 1 so prikazane amplitudne vrednosti frekvenčne karakteristike filtra za različne vrednosti koeficienta **k**.



*Slika 1: Amplitudna karakteristike empiričnega filtra za različne vrednosti koeficienta **k***
*Fig. 1: Magnitude response of the empiric filter for different values of parameter **k***

Frekvenca vzorčenja oz. časovne razmere v izrazu niso zajete. Zato tudi ni povezave med **k** in mejno frekvenco. Ker frekvenčnega spektra motilnega signala običajno niti ne poznamo, se **k** določi s poizkušanjem, tako da se doseže najboljši rezultat izločanja motilnih signalov. Probleme lahko pričakujemo v primerih, ko je prisoten spremenljiv čas izvajanja programa PLC-ja in ko postane večji od časa vzorčenja. V določenih razmerah lahko postane tako načrtovan filter povsem neuporaben.

4.3 Rekurzivni digitalni filter 1. stopnje

Tudi za običajno rekurzivno strukturo 1. stopnje se v preprostih aplikacijah uporabljajo približni postopki načrtovanja, ki so vezani na parameter koeficienta glajenja /5/. Pogosto se uporablja naslednji zapis:

$$y(n) = \frac{1}{L}\cdot\left(S\cdot y(n-1) + (L-S)\cdot x(n)\right) \quad (5)$$

Zapis je podoben nastavku (4), le da je v (5) dodatno prisoten še parameter L, ki določa občutljivost karakteristike filtra na spremembe koeficienta S. S primerjavo (3) in (5) ugotovimo, da je S dejansko koeficient glajenja. Za L se uporabljajo vrednosti med 12 in 20, običajno je v takšnih izrazih vrednost 16. Ko uporabimo pri (5) še *z*- transformacijo, sledi prenosna funkcija:

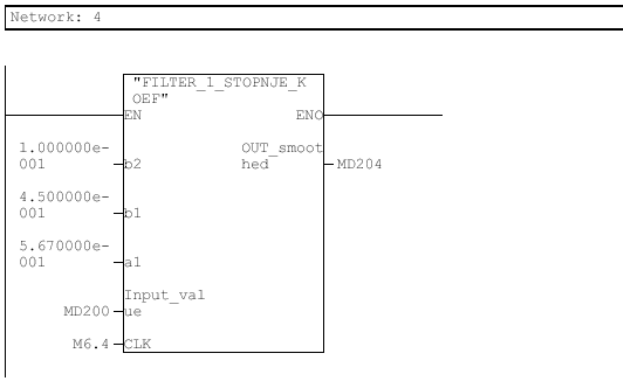
$$H(z) = \frac{L-S}{L - S\cdot z^{-1}} \quad (6)$$

S primerjavo dobljene prenosne funkcije s splošno obliko prenosne funkcije digitalnega filtra 1. stopnje v obliki,

$$H(z) = \frac{b_2 + b_1\cdot z^{-1}}{1 + a_1\cdot z^{-1}}, \quad (7)$$

vidimo, da je $b_2 = 0$ $b_1 = \frac{L-S}{L}$ in $a_1 = \frac{S}{L}$.

Če v editorju S7, ki je programsko orodje v Siemens-ovem programskem paketu in se uporablja za kreiranje in urejevanje struktur avtomatizacije za PLC krmilnike iz družine S7, napišemo takšno funkcijo, bomo dobili za vnos spremenljivk, oz. za branje izhodne vrednosti (v LAD načinu editorja S7) naslednjo "masko", prikazano na sliki 2:



Slika 2: Maska za vnos koeficientov za filter 1. stopnje
Fig. 2 : Mask for entry of filter coefficients for 1st order filter

Vrednosti koeficientov b_2, b_1, a_1, naslovi spremenljivk Input_value, CLK in OUT_smoothed, so podani le kot primer, pri dejanskem klicu funkcije bodo na teh mestih simboli «???». Če vnesemo napačni tip spremenljivke (BIT, BYTE, WORD ali REAL), editor avtomatsko zavrže takšen vnos in ga obarva rdeče.

4.4 Rekurzivni digitalni filter 2. stopnje

Vse lastnosti opisane filtra 1. stopnje, lahko razširimo tudi na filter 2. stopnje. V tem primeru se ne uporabljajo več empirične poenostavitve. Ker gre za splošen primer, ga bomo določili s koeficienti prenosne funkcije. Zanj smo uporabili sledeče zahteve: frekvenca vzorčenja Fs= 50 Hz, mejna frekvenca Fm=3 Hz, slabljenje v prepustnem frekvenčnem področju Ap=1 dB in slabljenje v zapornem frekvenčnem področju As=40 dB.

S pomočjo MATLAB-a smo dobili naslednje vrednosti koeficientov, ki so za popolno in okrajšano obliko prikazani v

razpredelnici 1. Okrajšano obliko dobimo pri izračunu, če izpustimo inštrukcijo *format long*.

Tabela 1: Koeficienti eliptičnega IIR filtra 2. stopnje v popolni in okrajšani obliki
Table 1: Coefficients of 3rd order elliptic filter in format long and in shortened mode

koeficienti	Popolna oblika	Okrajšana oblika
b2	0.03679945320611	0.0367
b1	0.04153456075581	0.041534
b0	0.03679945320611	0.041534
a2	1.0	1.0
a1	1.53728291988207	- 1.53
a0	0.66646479475237	0.66

Prvih 20 vzorcev impulznega odziva simulacijskih rezultatov in praktične izvedbe digitalnega filtra s popolno in okrajšano dolžino koeficientov je skupaj z ustreznima odstopanjima podanih v razpredelnici 2.

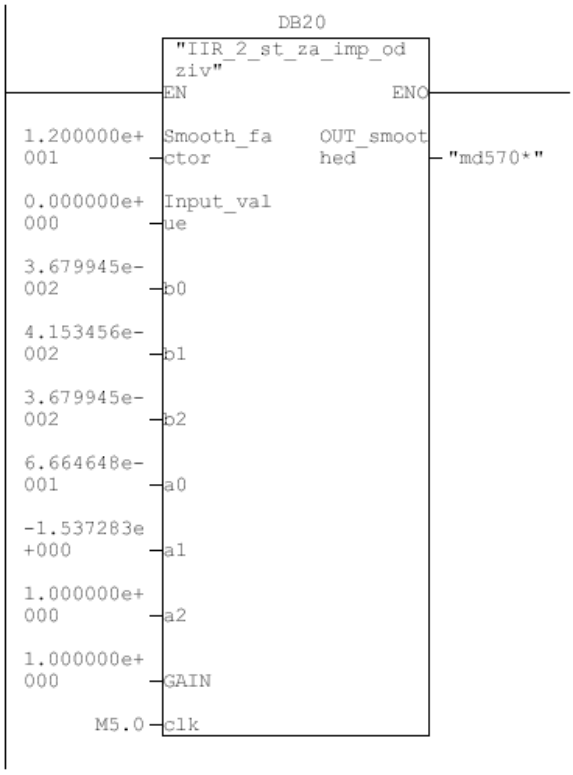
Primerjava impulznih odzivov digitalnega filtra druge stopnje izvedenem s PLC-jem s simulacijskimi rezultati dobljenimi z MATLAB-om, pokaže, da je v primeru natančnih vrednosti koeficientov (format long) odstopanje zanemarljivo. Zlahka pa vidimo, da se pojavlja občutno večje odstopanje pri vnosu koeficientov v okrajšani obliki.

Pri izvedbi digitalnega filtra druge stopnje pride v primerjavi s filtrom prve stopnje do razlike tudi pri sami izvedbi, saj je klic funkcije oz. funkcijskega bloka možen samo iz organizacijskega bloka OB35 z asinhrono prekinitvijo. Drugače bi dobili v primeru uporabe več filtrov nestabilne rešitve zaradi prekrivanja vmesnih rezultatov. Zaradi tega je izračun izhodne vrednosti opravljen tudi s funkcijskim blokom (FB) in ne več s funkcijo (FC). Vsakemu funkcijskem bloku je priključen pripadajoči podatkovni blok (DB) v katerem so shranjene vrednosti vmestnih spremenljivk in prejšnjih izhodnih in vhodnih vrednosti. Ker je spremenljivk več, je procedura pri preslikavi spremenljivk ob koncu izračuna nekoliko bolj zahtevna. Maska za vnos koeficientov, vhodnih in izhodnih spremenljivk je prikazana na sliki 3.

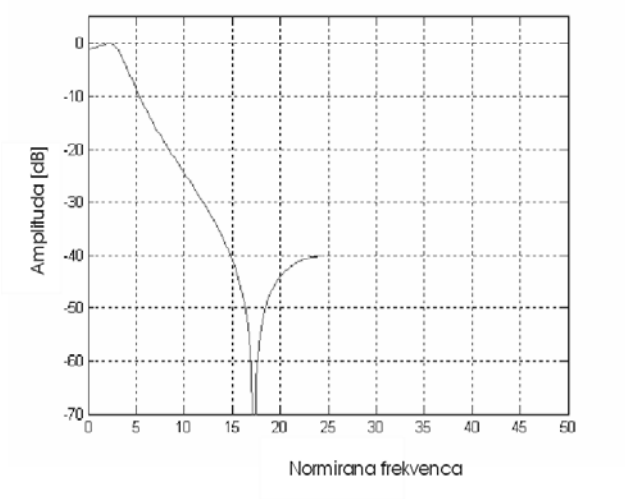
Ker je v našem primeru frekvenca vzorčenja bistveno večja od mejne frekvence filtra, ni težav s prekrivanjem spektrov **/6/**. Frekvenčna karakteristika nizkoprepustnega eliptičnega filtra s popolnimi vrednostmi koeficientov je podana na sliki 4.

Tabela 2: Primerjava 20 vrednosti impulznega odziva digitalnega filtra druge stopnje
Table 2: The comparison of 20 values of impulse response of digital filter of 2nd degree

odtipek št.	Simulacijski rezultati (MATLAB)	Praktična izvedba digitalnega filtra s PLC-jem	Odstopanje v (%)	Praktična izvedba digitalnega filtra s PLC-jem	Odstopanje v (%)
0	36,79940	36,799400	0	36,80000	-0,001630461
1	98,10570	98,105700	0	97,83800	0,272868957
2	163,09010	163,090000	6,13158E-05	162,10400	0,604635107
3	185,33170	185,332000	-0,000161872	183,44600	1,017472996
4	176,21300	176,213000	0	173,68400	1,435194906
5	147,372000	147,373000	-0,000678555	144,66200	1,838883913
6	109,11370	109,114000	-0,000274943	106,70100	2,211179714
7	69,51980	69,519800	0	67,77200	2,514103896
8	34,15117	34,151200	-8,78447E-05	33,27470	2,566442087
9	6,16752	6,167530	-0,00016214	6,17794	-0,168949594
10	13,27933	13,279300	0,000225915	12,50900	5,80097038
11	-24,52450	24,524500	0	23,21630	5,334257579
12	-28,85090	-28,850900	0	-27,26490	5,497228856
13	-28,00700	-28,007300	-0,001071161	-26,39260	5,764273217
14	-23,82700	-23,827000	0	-22,38580	6,048600327
15	-17,96300	17,963000	0	16,83120	6,300729277
16	-11,73400	-11,734300	-0,002556673	-10,97710	6,450485768
17	-6,06728	-6,067280	0	-5,68634	6,27859601
18	-1,50661	-1,506620	-0,000663742	-1,45523	3,410305255
19	1,72754	1,727540	0	1,52648	11,63851488
20	3,65982	3,659820	0	3,29597	9,941745769



Slika 3: Maska za vnos koeficientov in vhodnih spremenljivk rekurzivnega filtra 2. stopnje
Fig. 3 : Mask for entry of coefficients and input variables of 2nd order IIR filter



Slika 4: Amplitudna frekvenčna karakteristika IIR digitalnega filtra druge stopnje z mejno frekvenco 3 Hz
Fig. 4: The magnitude frequency response of 2nd order IIR filter with cut-off frequency 3 Hz

4.5 Rekurzivni digitalni filtri višje stopnje

Filtri višjih stopenj se običajno realizirajo kot kaskadna vezava več struktur prve in druge stopnje in le izjemoma v direktni obliki. Čeprav z digitalnim filtrom druge stopnje dosežemo dokaj dobre rezultate v primerjavi z opisanimi izkustvenimi filtri, lahko s filtrom višje stopnje le dosežemo ali večje dušenje ali bolj strmo upadanje amplitudne frekvenčne karakteristike nad mejno frekvenco. Pri tem pa moramo paziti le na to, da takšen filter ne moti delovanje PLC-ja in da ne povzroča prevelike časovne zakasnitve osnovnega programa.

V našem primeru bomo prikazali izvedbo digitalna filtra tretje stopnje s katerim smo zadostili postavljenim zahtevam. Za to je potrebno uporabiti le dve kaskadni strukturi. V ta namen smo kreirali poseben funkcijski blok (FB), s katerem smo lahko analizirali obnašanje filtra.

Primerjavo med simulacijskimi rezultati, dobljenimi v MATLAB programskem okolju in praktično izvedbo s PLC krmilnikom smo ocenili s primerjavo odstopanja impulznih odzivov. Fazne razmere nam niso bile pomembne in jih nismo upoštevali. Za snemanje impulznega odziva smo modificirali funkcijski blok rekurzivnega filtra druge stopnje na način, da je bil vhod urinega impulza - «clk» - aktiven. Ta vhod je lahko aktiven, če prožimo funkcijski blok z notranjo uro, generirano v programu. Če FB pokličemo iz OB35, ta vhod ni aktiven, oz. ga moramo deaktivirati. To pomeni da smo ob vsakem proženju dobili na izhodu le eno izhodno vrednost. Pri tem je bila na vhodu vrednost 1000.0 le ob prvem ciklu, in nič pri vseh ostalih.

Načrtali, simulirali in izvedli smo eliptični IIR filter 3. stopnje. Zanj smo uporabili iste zahteve kot za filter druge stopnje.

Koeficiente filtra smo izračunali v MATLAB okolju neposredno oziroma s pomočjo FDA toolbox-a, ki je tudi del programskega paketa MATLAB. Dobljeni koeficienti v popolni obliki so za eliptični filter prikazani v razpredelnici 3.

Tabela 3: Koeficienti eliptičnega IIR filtra 3. stopnje
Table 3: The coefficients of 3rd order elliptic filter

b3= 0,0136 426 503 0127	a3= 1,00
b2=- 0,0018 096 880 4567	a2= -2,5336 859 770 7381
b1=- 0,0018 096 880 4567	a1= 2,2493 565 017 9860
b0= 0,0136 426 503 0127	a0= -0,6920 046 002 1359

Za pretvorbo filtra višje stopnje v kaskadno obliko smo uporabili MATLAB programsko orodje (FDA toolbox ali inštrukcijo ***zp2sos*),** ki nam omogočata določitev struktur druge stopnje ali SOS (second order structure) matriko v obliki:

$$SOS = \begin{bmatrix} b_{21} & b_{11} & b_{01} & 1 & a_{11} & a_{01} \\ b_{22} & b_{12} & b_{02} & 1 & a_{12} & a_{02} \\ . & . & . & . & . & . \\ b_{2L} & b_{1L} & b_{0m} & 1 & a_{1m} & a_{0m} \end{bmatrix} \quad (8)$$

V prvi vrstici so podani koeficienti prve strukture (števec in imenovalec), v drugi druge itn. Pri tem je *m* število struktur 2. stopnje.

Če poznamo SOS matriko, zapišemo splošno obliko prenosne funkcije v obliki:

$$H(z) = g \cdot \prod_{k=1}^{L} \frac{b_{2k} + b_{1k} \cdot z^{-1} + b_{0k} \cdot z^{-2}}{1 + a_{1k} \cdot z^{-1} + a_{0k} \cdot z^{-2}}, \quad (9)$$

kjer je z *g* označeno skupno ojačenje. Za digitalni filter tretje stopnje je SOS matrika naslednja:

$$SOS = \begin{bmatrix} 1{,}00 & 1{,}00 & 0{,}00 & 1{,}00 & -0{,}818337350 & 0{,}00 \\ 1{,}00 & -1{,}132649302 & 1{,}00 & 1{,}00 & -1{,}715348626 & 0{,}845622651 \end{bmatrix} \quad (10)$$

in ojačanje g = 0,013642650301

Za prenosno funkcijo digitalnega filtra z dvema kaskadnima strukturama,

$$H(z) = g \cdot \frac{b_{2a} + b_{1a} \cdot z^{-1} + b_{0a} \cdot z^{-2}}{1 + a_{1a} \cdot z^{-1} + a_{0a} \cdot z^{-2}} \cdot \frac{b_{2b} + b_{1b} \cdot z^{-1} + b_{0b} \cdot z^{-2}}{1 + a_{1b} \cdot z^{-1} + a_{0b} \cdot z^{-2}}, \quad (11)$$

nam SOS matrika direktno podaja vrednosti koeficientov posameznih struktur a in b. Njeni koeficienti so podani v razpredelnici 4.

Tabela 4: Koeficienti posameznih kaskadnih struktur eliptičnega IIR filtra 3. stopnje
Table 4: The coefficients of both cascaded structures for 3rd order eliptic filter

b2a	=	1,00
b1a	=	1,00
b0a	=	0,00
a1a	=	-0,818337350
a0a	=	0,00

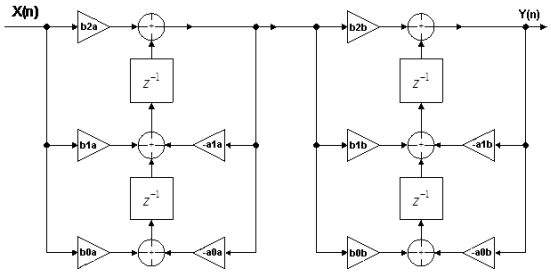
b2b	=	1,00
b1b	=	-1,1326493 02
b0b	=	1,00
a1b	=	-1,715348626
a0b	=	0,8456226513

Vidimo, da sta pri digitalnem filtru tretje stopnje koeficienta b_{0a} =0 in a_{0a}=0. Ojačanje bomo zato upoštevali pri prvi strukturi, ker ima manj koeficientov in je manj množenj.

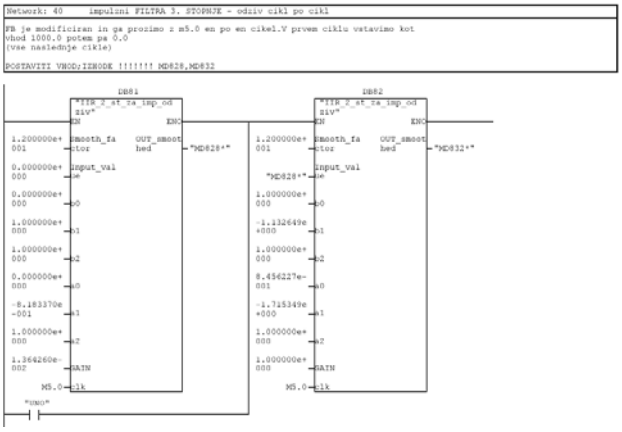
Splošna kaskadna izvedbena oblika za rekurzivne digitalne filtre tretje stopnje je prikazana na sliki 5. Ker gre za splošno obliko z dvema kaskadnima strulturama drugih stopenj, je uporana tudi za filter četrte stopnje.

Na sliki 6 je prikazana maska za vnos koeficientov v posamezni kaskadni strukturi digitalnega filtra. Izhod iz prve strukture MD 828 je vhod v drugo strukturo. Izhod iz filtra je MD 832.

Primerjava impulznih odzivov digitalnega filtra tretje stopnje izvedenem s PLC-jem v kaskadni obliki in s simulacijskimi rezultati dobljenimi z MATLAB-om, pokaže, da je v



Slika 5: Splošna kaskadna izvedbena oblika rekurzivnega digitalnega filtra z dvema strukturama
Fig. 5: General cascade realization for two structures recursive digital filter



Slika 6: Del organizacijskega bloka OB1 digitalnega filtra s prikazom vnosa koeficientov in vhodnih spremenljivk v prvo in drugo strukturo
Fig. 6: A part of the organisation block for entry of the coefficients and input variables of 3rd order IIR filter in the first and second cascaded structure

primeru natančnih vrednosti koeficientov (format long) odstopanje zanemarljivo, kot smo to že ugotovili pri filtru druge stopnje. Prvih 20 vzorcev impulznega odziva je podanih v razpredelnici 5.

Celotni impulzni odziv izvedenega digitalnega sita s PLC-jem je prikazan na sliki 7, ustrezna frekvenčna karakteristika, ki smo jo s pomočjo MATLAB-a izračunali na osnovi impulznega odziva pa na sliki 8:

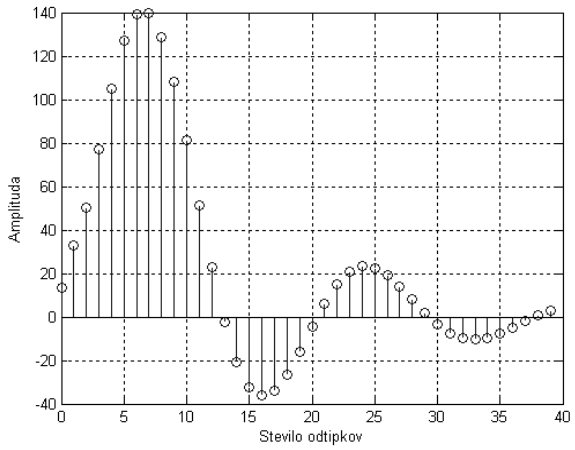
Amplitudna frekvenčna karakteristika je ustreza postavljenim zahtevam. Slabljenje 40 dB dosežemo že pri 7 Hz, skoraj pri dvakratni mejni frekvenci, kar je bistveno bolje kot pri filtru druge stopnje, ko smo zahtevano slabljenje dosegli pri 15Hz.

5. Zaključek

V članku smo prikazali možnosti izvedbe eliptičnega digitalnega filtra prve, druge in tretje stopnje v rekurzivni obliki s

Tabela 5: Primerjava prvih 20 vrednosti impulznega odziva digitalnega filtra tretje stopnje
Table 5: The comparision of first 20 values of impulse response of digital filter of 3rd degree

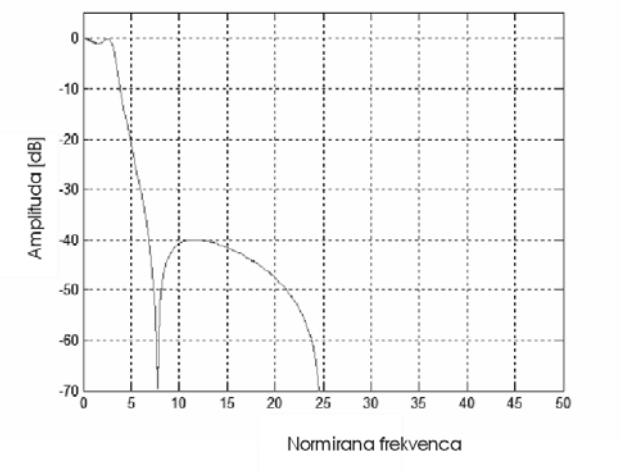
odtipek št.	Simulacijski rezultati (MATLAB)	Praktična izvedba digitalnega filtra s PLC-jem	ODSTOPANJE v %
0	13,642650301273000	13,6426	0,000368706
1	32,756503712783000	32,7564	0,000316617
2	50,497821912436000	50,4977	0,000241421
3	77,347995718632000	77,3478	0,000253036
4	105,0555793178170	105,055	0,000551439
5	127,13935612429400	127,139	0,000280105
6	139,34892221216300	139,349	-5,58223E-05
7	139,83616962670000	139,783	0,0380228
8	128,70340523262500	128,703	0,000314858
9	108,10092058526500	108,101	-7,34635E-05
10	81,12485120469300	81,1248	6,31184E-05
11	51,45073780478000	51,4507	7,34776E-05
12	22,68763570368100	22,6877	-0,000283398
13	-2,10893694285100	-2,10885	0,004122591
14	-20,77181759385600	-20,7717	0,000566122
15	-32,18556365668600	-32,1855	0,00019778
16	-36,28428240826100	-36,2843	-4,84831E-05
17	-33,91036398052200	-33,9104	-0,00010622
18	-26,57448525769500	-26,5746	-0,000431776
19	-16,16379328939600	-16,164	-0,00127885
20	-4,64461306675500	-4,64483	-0,004670642



Slika 7: Impulzni odziv IIR digitalnega filtra 3. stopnje
Fig. 7: Impulse response of 3th order IIR digital filter

PLC krmilnikom. Ker senzorji in pretvorniki, ki jih uporabljamo na področju avtomatizacije procesov s PLC krmilniki, ne ločijo motilnih signalov od koristnih, dobi procesni sistem poleg koristnih signalov še neželene motilne signale. Zato je potrebno z dodatno opremo, aparaturno ali programsko, odpraviti motnje, da bo proces katerega krmilimo, potekal brezhibno. Pri vedno zahtevnejših aplikacijah avtomatizacije industrijskih procesov se pojavljajo zahteve po digitalnem odpravljanju neželenih vhodnih signalov.

Glede na to, da nas pri naših aplikacijah vpliv nelinearnega faznega kota ni problematičen, smo za filtriranje motilnih signalov izbrali rekurzivno filtrsko strukturo. Uporabili smo koeficiente eliptičnega filtra. Prevladalo je dejstvo, da ima za dano strmino upadanja frekvenčne karakteristike v pre-



Slika 8: Amplitudna frekvenčna karakteristika IIR digitalnega filtra tretje stopnje z mejno frekvenco 3 Hz
Fig. 8: The magnitude frequency response of 3th order IIR digital filter with cut off frequency 3 Hz

hodnem frekvenčnem področju najmanjšo stopnjo. Tudi valovitost amplitudne frekvenčne karakteristike ni bila problematična, če le upoštevamo mejne vrednosti sistema.

Pri izvedbi filtrske strukture v PLC sistemu smo prikazali, da je ciklična prekinitev pri nekaterih aplikacijah zelo pomembna, in da ne moremo vstaviti filtra v programsko zanko, če čas urinih taktov ni konstanten, če je odvisen od pogojev razvejanja osnovnega programa. Takšna odločanja znotraj programa lahko podaljšajo čas izvajanja takta tudi za velikostni razred do 10 krat.

Najprej so v delu opisani empirični filtri, ki se uporabljajo na podlagi izkušenj in ne zahtevajo poglobljenega teoretičnega znanja. V nadaljevanju pa smo prikazali postopke načrtovanje in izvedbe rekurzivnih struktur prve, druge in tretje stopnje. Struktura rekurzivnega filtra predstavlja povezavo z empiričnimi oblikami gladilnih filtrov, vendar z natančno določenimi parametri filtriranja, strukturi druge in tretje stopnje pa zahtevata posebne postopke načrtovanja in izvedbe s PLC krmilnikom. Koeficiente filtrov smo izračunali v MATLAB okolju neposredno oziroma s pomočjo FDA toolbox-a. Pri digitalnem filtru druge stopnje smo podrobno analizirali odstopanje impulznega odziva izvedenega filtra s PLC krmilnikom pri zapisu koeficientov v popolni in pri zapisu v skrajšani obliki.

Za pretvorbo filtra tretje stopnje v kaskadno obliko smo uporabili MATLAB programsko orodje, ki omogoča določitev struktur druge stopnje ali SOS (second order structure) matriko. S prikazano SOS matriko je kaskadna izvedba filtra tretje stopnje enostavno določena. S slikami so prikazani deli organizacijskih blokov PLC krmilnika s prikazom mask za vnos koeficientov in vhodnih spremenljivk v posamezne strukture digitalnih filtrov. Primerjava impulznih odzivov digitalnega filtra tretje stopnje izvedenem s PLC-jem v kaskadni obliki s simulacijskimi rezultati dobljenimi z MATLAB-om, tudi pokaže, da je v primeru natančnih

vrednosti koeficientov (format long) odstopanje zanemarljivo. Večja odstopanja se pojavijo le pri vpisovanju koeficientov v okrajšani obliki. Odstopanja so rezultat rekurzivnega izračuna izhodne vrednosti. Zato se je potrebno zapisu koeficientov v okrajšani obliki izogibati, saj v ničemer ne zmanjšuje niti aparaturnih, niti časovnih zahtev. Ne glede na način zapisa koeficientov, ni bilo nevarnosti za nestabilnost izvedenih filtrov.

Tudi amplitudna frekvenčna karakteristika ustreza postavljenim zahtevam. Slabljenje 40 dB dosežemo skoraj že pri dvakratni mejni frekvenci, kar je bistveno bolje kot pri filtru druge stopnje. Pri 7 Hz in ne pri 15 Hz.

6. Literatura

/1./ L. Milić, Z. Dobrosavljević, Uvod u digitalnu obradu signala, Elektrotehnički fakultet Beograd, 1999.

/2. / R. Babič, Dinamika izhodnega signala pri kaskadni obliki izvedbe nerekurzivnih digitalnih sit Informacije MIDEM. - ISSN 0352-9045. - Letn. 31, št. 3 (2001), str. 152-158.

/3. / SIEMENS S7-300 manual, programmable controller, hardware and instalation SIEMENS AG, 1998, EWA 4NEB 710 6084-002 01

/4. / L. Milić, M. Đurić, Rekurzivni digitalni filtri, Naučna knjiga Beograd, Beograd, 1982

/5. / Aleksandar Dodič, Izvedba rekurzivnih digitalnih sit s PLC krmilnikom, magistrska naloga, Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko, Maribor, 2007.

/6. / Sanjit K.Mitra, Digital Signal Processing-A computer Based Approach, Mc.Graw Hill, 2002.

Mag. Aleksandar Dodič, univ. dipl. inž., Lučka uprava Rijeka, Riva 1, 51 000 Rijeka, Hrvatska, e-mail: sandro.dodic@portauthority.hr

Izr. prof. dr. Rudolf Babič, Univerza v Mariboru, Fakulteta za elektrotehniko računalništvo in informatiko, 2000 Maribor, Smetanova 17, Slovenija

Prispelo (Arrived): 08.03.07 *Sprejeto (Accepted):* 28.5.08

MICROPHONE TRANSFER FUNCTION ADAPTATION USING A BI – QUAD FILTER AND DCL

J. Stergar[1], D. Miletić[1], C. Beaugeant[2], B. Tramblay[2]

[1]UNI Maribor, Faculty of EE and Computer Science, Maribor, Slovenia
[2]Siemens AG, formal ICM Mobile Phones, Munich, Germany

Key words: Bi-Quadratic filter, microphone transfer function, dynamic compression, adaptation, noise robustness, audio signal enhancement

Abstract: A suitable adaptation of the microphone in the audio path of a mobile device is a very sensitive tasks. The conformation of the frequency response characteristic to the GSM standards is inevitable. Achieving the highest correlation between the specified GSM frequency response specification with the microphone and speaker characteristic is a delicate matter. In this paper we will present tests performed for microphone transfer function adaptation on a mobile phone using a Bi-Quad filter. Therefore a cascading with a IIR filter of the II. order was applied. The main goal of our test was to evaluate the influence of a recursive filter in the audio input path to the recognition rate of the embedded recognizer. The focus of our tests was to simulate the cascaded components with only one substitute having similar (almost equivalent) frequency response characteristics. By doing so we assumed that the microphone used was a high quality microphone with a linear frequency response. Further enhancement of the audio quality was performed by a more sophisticated dynamic compression and limitation algorithm (DCL). For the evaluation tests the Motiv and Aurora 3 databases were applied to the audio input. The tests have shown that with appropriate adaptation of the cascaded components an improvement of the recognition rate is realizable. The recognition rate of the mobile phone embedded recognizer was enhanced for over 6% indicating that the proposed approach sensibly contributes to speech signal enhancement.

Adaptacija prenosne funkcije mikrofona z Bi – Quad filtrom in DCL

Kjučne besede: Bi-Quad filter, mikrofonska prenosna funkcija, dinamično zgoščanje, adaptacija, šumna robustnost, izboljšanje avdio signala

Izvleček: Ustrezna adaptacija mikrofonske prenosne funkcije avdio poti mobilnega telefona je zelo občutljiva naloga. Skladnost odzivne frekvenčne karakteristike z GSM standardi je obvezna. Doseči kar najvišje sovpadanje med standardiziranim GSM odzivom prenosnega telefona ter dejansko prenosno karakteristiko mikrofona in zvočnika je dokaj težavno. V tem članku bomo predstavili preskuse za adaptacijo mikrofonske prenosne funkcije mobilnega telefona z uporabo rekurzivnega filtra II. stopnje t.i. Bi-Quad filtra. Zato smo v kaskado z vhodno avdio potjo telefona vstavili ustrezen adaptacijski filter. Glavni cilj naših preskusov je bil oceniti vpliv kaskadiranega filtra na uspešnost razpoznavanja vsajenega razpoznavalnika telefona. Pri izvedenkah eksperimentih smo se osredotočili na kaskadiranje zgolj enega samega člena (rekurzivnega filtra) s podobno oz. ekvivalentno frekvenčno odzivno karakteristiko kot veleva GSM standard. Pri tem smo uporabili visokokakovosten mikrofon z linearnim frekvenčnim odzivom. Dodatno izboljšanje kakovosti avdia je bilo izvedeno z naprednejšim algoritmom dinamičnega zgoščanja in omejevanja (DCL). Za evalvacijo razpoznavanja vsajenega razpoznavalnika smo uporabili Motiv in Aurora govorni bazi z omejenim naborom ukaznih besed. Preskusi so pokazali, da lahko z ustrezno adaptacijo prenosne funkcije mikrofona izboljšamo uspešnost razpoznavanja. Z izboljšanjem govornega signala smo povečali uspešnost razpoznavanja vsajenega razpoznavalnika za 6%.

1 Introduction

In this article we will present tests regarding the influence of a II. order recursive filter implementation on speech recognition in a mobile device. This kind of the so called Bi-Quad filter is usually used in mobile devices in cascade with a microphone and a loudspeaker.

The purpose of cascading a filter in the audio path of a recognizer is to adjust the frequency response – the transfer function of microphone – to a specific GSM standard. Recursive filter and microphone or loudspeaker together must form a suitable transfer function which is defined with one of the GSM standards.

A cornerstone of quality in telecommunication is the acoustic quality of terminals. Nevertheless a combination of several factors influences this quality. Essential to the terminals are the physical characteristics of the transducers. This leads to possible distortions of signals like deviation from the ideal frequency response or even nonlinearities. Other factors are coupled with the environment and the use case in which the terminals are operated (e.g., noisy environment or hands-free mode), causing degradations like echo and unintelligible speech for the far-end listener. In principle, most effects can be reduced by elaborate acoustical or mechanical designs.

However, without the help of digital signal processing algorithms, the acoustic quality remains insufficient /3/.

Considerable amount of varying background noise is a problem for all mobile devices such as cell phones or speech controlled car systems. Automatic systems are much more sensitive to the variability of the acoustic signal than humans. Therefore the recognition error rates of speech recognition systems using standard methods usually rise considerably in these conditions /1/.

Besides the quest for robust features two main lines of research aimed at increasing performance of speech recognizers:

- speech signal enhancement and
- model adaptation.

Although some adaptation techniques achieve very good performance, their use in embedded systems is only of limited interest. This is due to the fact that recognizers operating in mobile phones are subject to constantly changing environments and little or no adaptation data.

In contrast, speech enhancement techniques require no training; therefore they are suitable for embedded systems and additionally provide "real-time" improvement of recognition rates. For a resource constrained mobile phone, speech signal enhancement has the added advantage that the same program code can be used to improve not only the recognition rates of the speech recognizer but also the quality of the speech signal for the far-end talker during a voice call. Of course, different tunings of the enhancement algorithm have to be found for both cases in order to optimize for a machine or a human being the listener /2/.

The first issue to deal with regarding speech enhancement in mobile phones is the quality and characteristics of transducers. The frequency characteristics of microphones and loudspeakers do not necessarily comply with the requirements given.

For acoustic shock prevention dynamic compression with signal limitation was additionally applied.

2 Dynamic Compressor and Limiter

Acoustic shock is reduced by limiting an input signal based on tones detected through frequency domain analysis. Further enhancement of the audio quality is performed by a more sophisticated dynamic compression and limitation (DCL) algorithm. The basic principle of the DCL is the following: by amplification of medium signal levels speech intelligibility is improved. For each consecutive frame, the power (PWR) of the signal is computed and weighted with the power of the previous frame. Depending on this energy, a first gain is applied on the signal according to the curve in Fig. 3. An amplification is applied for frames whose power belongs to the interval [Lim E, Lim L]. For low power signal, no amplification is applied and for high power frames, a limitation to Lim L is applied. In addition, a general gain C is applied to the signal according to the shift up head-room, so that a dynamic gain according to the dotted curve is applied to each consecutive frame.

The DCL provides speech signals which sound more 'direct' and 'present' by the reduction of the dynamic range and limitation of the signal to a maximum level. If such property is welcome for speech conversation scenario [1], it is also foreseen that speech recognizer could be positively influenced by an enhancement of such a speech presence: Amplifying speech region where the information of the signal is the most important is a priori a good way to increase the performance of speech recognition.

Moreover, for scenario where the Signal to Noise Ratio (SNR) is not too low, the influence of the noise is reduced. Indeed, in such scenario the noise level is lower the threshold Lim E and that "noise only" frames are less amplified than "speech + noise" frames. As a result, the noise period influence on speech recognizer is reduced compared to speech periods.

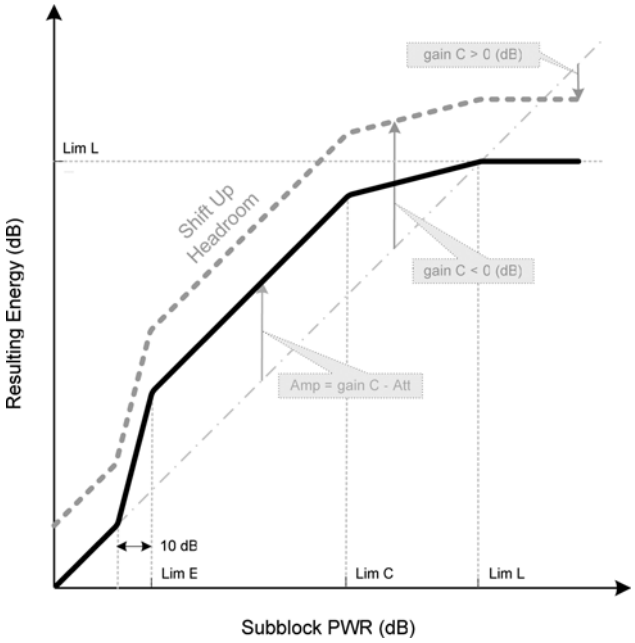


Fig. 1: The DCL function.

Likewise, sudden bursts of noise whose energy is higher then the threshold Lim L are limited by the DCL. It involves that their influence on speech recognizer is reduced as well.

Such remarks show that a priori the DCL should enhance the performance by enhancing the 'presence' of speech and reducing the influence of static and burst noise.

3 The Bi-Quadratic adaptation filter

More and more frequently occurs that the speech input signal exhibits a "flatter" spectrum, for example when a hands-free installation is used, employing a microphone with linear frequency response. Conventional recognizers are designed to be independent of the input with which they operate, and, they are without any knowledge of the characteristics of this input. If microphones with different characteristics are likely to be connected up to the mobile phone, or more generally if the recognizer is likely to receive acoustic signals exhibiting different spectral characteristics, there are cases in which the Very Smart Recognizer (VSR) embedded in our case, operates in a sub-optimal manner. In this context, a main purpose of the microphone transfer function adaptation is to improve the speech signal making it less dependent on the spectral characteristics.

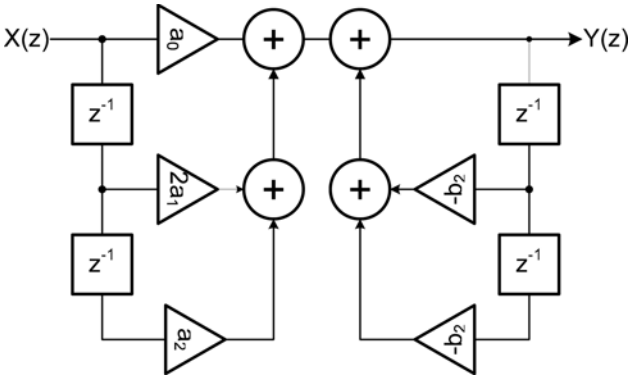


Fig. 2: Block diagram of a typical bi-quad filter.

The purpose of cascading a filter in the audio path of a recognizer is therefore to adjust/adapt the frequency response – the transfer function of microphone – to a specific standard.

The purpose of cascading a filter in the audio path of a recognizer is to adjust the transfer function of the microphone to a specific standard. A convolution function of the different characteristics of the filter and microphone in cascade is performed and together they must form a suitable transfer function which is defined with one of the GSM standards /5/.

$$H(z) = g\frac{a_0 + 2\cdot a_1 z^{-1} + a_2 z^{-2}}{1 + 2\cdot b_1 z^{-1} + b_2 z^{-2}} \qquad (1)$$

A II. order recursive filter – the Bi-Quadratic filter – was implemented to adapt the microphone transfer function. The Bi-Quad digital filter is a common name for a two-pole, two-zero recursive filter which name was derived from the transfer function structure of the filter. This kind of the Bi-Quadratic filter is typically used in mobile devices in cascade with a microphone and loudspeaker.

Practically this is an Infinite Impulse Response (IIR) filter of the second order. The transfer function for this kind of a filter can be defined from its block diagram (Figure 2).

It can be seen that this is a common two-pole, two-zero digital filter with a typical transfer function (Figure 3) /6/.

Using the shift theorem for z transforms, the difference equation for the Bi-Quad can be written by inspection of the transfer function as:

$$y(n) = \sum_{i=0}^{M} b_i x(n-i) - \sum_{j=1}^{N} a_j y(n-j) \qquad (2)$$

where x(n) denotes the input signal sample at time n, and y(n) is the output signal /8/.

In most fixed-point arithmetic schemes (such as two's complement, the most commonly used) there is no possibility of internal filter overflow. That is, since there is fundamentally only one summation point in the filter, and since fixed-point overflow naturally "wraps around" from the largest positive to the largest negative number and vice versa, then as long as the final result y(n) is "in range", overflow is avoided, even when there is overflow of intermediate results in the sum. This is an important, valuable, and unusual property of the used filter structure. There are twice as many delays as are necessary.

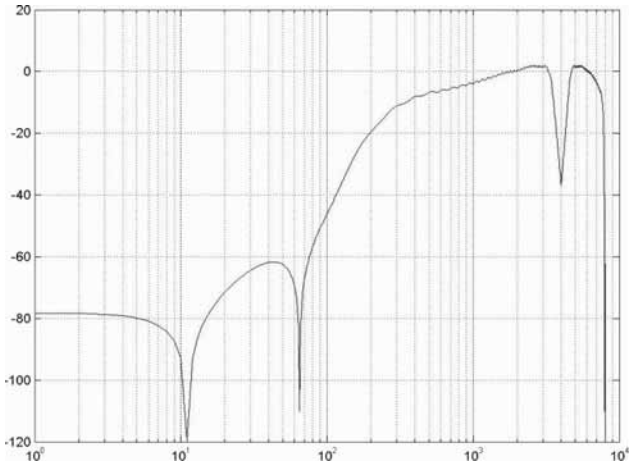


Fig. 3: A characteristic frequency response of the bi-quadratic filter (handset parameterisation).

As a result, the bi-quad structure is not canonical with respect to delay. In general, it is always possible to implement an N^{th}-order filter using only N delay elements. It is a very useful property of the Bi-Quad implementation that it cannot overflow internally in two's complement fixed-point arithmetic: As long as the output signal is in range, the filter will be free of numerical overflow. Most IIR filter implementations do not have this property /9/.

4 Used test material

4.1 The MoTiV database

The database MoTiV was recorded after the initiative between the industrial partners Philips, Siemens, Bosch, and Volkswagen in the subproject Man-Machine Interaction /7/. In total, 35 hours of hands-free multi-channel recorded speech data from about 640 drivers were collected in seven different mid- to upper-class ranged cars. All recordings were simultaneously made at least by two microphones, which had been fixed on the car ceiling at the A-beam and in the middle between driver and passenger.

For our experiments only a subset of recorded material was used. It consisted of 26 words (mostly command words) in German language with 100 diverse samples for each word. As mentioned all of these samples were recorded in different car environments therefore samples with different amount of noise were included. Beside that all samples were recorded by both genders. (Table 1).

The original samples had to be preprocessed (down-sampled) because of the embedded Very Smart Recognizer (VSR) limitations. Little memory and low computational power on the used mobile phone VSR supported only samples with f_s = 8 kHz for processing.

Table 1: The command words used from Aurora database.

Motiv		Aurora
ändern	löschen	
aus	Navigation	ZWEI
Ende	nein	ZWO
halt	Radio	SECHS
Hauptmenü	Start	FUENF
Hilfe	Stop	VIER
Information	stumm	ACHT
ja	suchen	DREI
Karte	Telefon	NEUN
Kassette	wählen	SIEBEN
Korrektur	weiter	EINS
lauter	wiederholen	NULL
leiser	zurück	

4.2 The Aurora 3 Database

The Aurora 3 database is a database of digits. This database is a subset of the SpeechDat-Car database in German language which has been collected as part of the European Union funded SpeechDat-Car project. It contains isolated and connected German digits spoken in the following noise and driving conditions inside a car: High/low speed good/rough road, stopped with motor running, town traffic /10/. Only digits from 0 to 9 are included. The samples are of different kind. A single sample consists of one or more digits. The frequencies of the digit appearance in the samples corpus differ. The used database was recorded on two different channels (ch0 and ch1). The ch0 is the primary channel where much less additive noise is present compared to channel ch1. The whole database has 3118 samples proportionally distributed in each channel. For the VSR recognition testing the samples had to be converted.

5 The Very Smart Recognizer

The Very Smart Recognizer (VSR) is a software-only speech recognition component especially designed for mobile terminals by Siemens AG. Featuring a modular architecture and flexible configuration it is particularly well suited for the support of voice commands /11/.

For the recognition experiments we used a speaker independent HMM (Hidden Markov Model) based VSR (V4.50). Speaker independent HMM-based technology offers command-and-control and digit dialling, i.e., recognition of commands and phone numbers without requiring a training phase. Natural number (e.g. twenty-five, ninety-eight) recognition applies the same technology while improving the usability of voice dialling. HMM based recognizers imply a higher implementation complexity and need appropriate speech databases in many languages. Today's memory and computational resources in 3G phones are facilitating deployment of such technology.

The VSR used for testing was in form of an executable and was activated with the appending parameters:

<melParamFile> <hmmFile> <vocabularyFile> <sampleFile>

In the options field a specific output file format specification for the recognition module was possible. In the melParamFile different kinds of parameters for VSR were included. The most important parameter for testing was the noise reduction parameter (NSR) set for all tests. The hmmFile carried a phonetic description for the used database. The vocabularyFile included a vocabulary of currently used database. The sampleFile represented the input data file with the recognition samples. This sample file was conveyed to the VSR input.

Fig. 4: The reference diagram for global recognition experiments.

6 Experiments

In the preliminary tests we applied the Bi-Quad module as speech enhancement pre-processing unit immediately before the VSR unit. For tests samples with different noise characteristics were deliberately selected (MoTiV/Aurora 3) for the evaluation of the recognition results with different DCL parameterisations.

The origin test framework consisted of the pre-processing module in cascade with the VSR module (Figure 4). This framework was used for separate tests using the Bi-Quad filter or DCL. Lastly a general I. order high-pass filter was also tested (Figure 5).

We started the tests with the objective of global recognition rates for each database, which has been later used as reference to other performed experiments. In addition a variety of experiments had been performed.

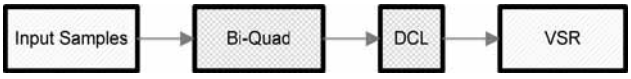


Fig. 5: The extended reference diagram supplemented with the DCL in cascade with the high-pass bi-quad filter.

The first step in all experiments as already mentioned was the evaluation of the global recognition rate for each database (VSR, Figure 10-13).

In the first experiment a Bi-Quadratic filter was inserted into the recognition path adapting the microphone transfer function. In the experiment our intent was to examine the speech recognition effectiveness/improvement of the VSR using different parameterisations of the Bi-Quad filter: high-pass (HP), handset and hands-free (Figure 7).

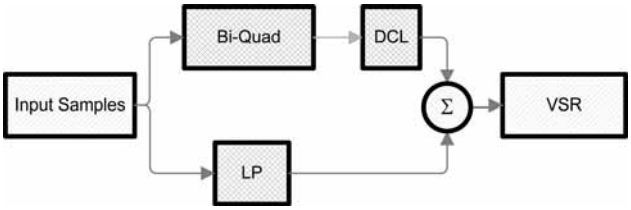


Fig. 6: The reference diagram combining the paths of the high-pass Bi-Quad + DCL with the low-pass branch.

In the second experiment a DCL module was inserted into a cascade before the VSR recognition stage replacing the Bi-quad (Figure 4). The Bi-Quad as well as the DCL executable were applied to every sample in the two databases. Therefore a comparison of the results with the global recognition rate could be performed (Figure 7).

In the second experiment another test was performed. A Bi-Quad filter was inserted before the DCL module and the high-pass, hands free and handset parameterisation of the cascaded Bi-Quad was examined (Figure 5). Our intent was to evaluate the influence of the inserted DCL on the VSR recognition.

In the third experiment another test scenario was applied. The recognition path was split into a high-pass filter branch in cascade with the DCL and a low-pass filter branch (Figure 6). The separated signals were summarised before the final recognition stage. Using the depicted method we separated the signal (information) from noise and applied the DCL only to the noisy part of the signal.

The goal of this experiment was to evaluate the impact of the DCL stage on the final recognition rate after summarisation.

All recognition rates (Word Correct Rates) were estimated with:

$$WCR\,(\%) = \frac{H}{N_r} = \frac{N_r - (D+S)}{N_r} \qquad (3)$$

where N_r represents the total number of words in the reference corpus, S the number of substituted words in the confusion matrix, D the number of words deleted from the confusion matrix, and H the number of correctly recognised words /12/. For all experiments an executable for the confusion matrix generation was used gathering the results for the word and global, recognition rates with the substitutions and deletions for each word tested.

The high/low pass filters as well as the BiQuad were realized in MATLAB environment.

7 Test results

All experiments were mainly performed with different DCL and filter parameterisations on the MoTiV database. The additional tests were performed with the Aurora 3 database and were used as a confirmation/rejection reference for the test results gained with the MoTiV database.

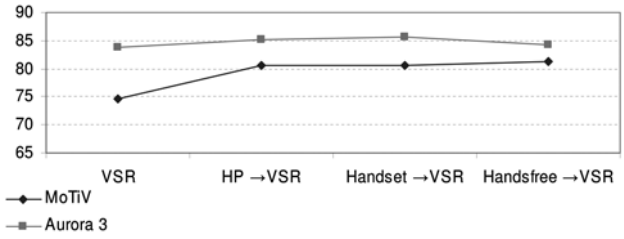


Fig. 7: The WCR for the MoTiV and Aurora 3 database using different Bi-Quad parameterisations.

Firstly our objective was to test the influence of the Bi-Quad adaptation of the transducers transfer function. The experiments show that all of the three parameterisations – high-pass, hands-free and handset contribute to increased the WCR of the implemented VSR in average over 6% (Figure 7). The promising recognition enhancement formed the foundation for the experiments with the DCL cascade. Our goal was to estimate if the cascading of the DCL preserves the gained improvement in recognition of the VSR and furthermore if only a single DCL parameterization would be sufficient for database independent recognition. Therefore in parallel supplementary experiments with the Aurora 3 database were performed.

With the applied DCL we also performed tests using different cut-off frequency ($f_{cut\text{-}off}$) variations. We estimated if any enhancements are possible for the different cut-off window. Slightly enhancements for some test scenarios using $f_{cut\text{-}off}$ = 350Hz were observed but the overall recognition results were best by using the $f_{cut\text{-}off}$ = 300Hz.

Further tests have been performed using the Aurora 3 database separately evaluating the global recognition rates for samples in both channels. The graphs indicate that the DCL parameterisation used for the MoTiV database almost preserves the gained recognition accuracy of the VSR (Figure 8). The recognition rate of the Aurora 3 database just slightly degrades in spite of the DCL parameterisation being optimised on the MoTiV database.

The experiments indicate that the used Bi-Quad considerably improves the overall recognition rate on the embedded VSR. Furthermore we can assume that the applied version of the DCL does not essentially degrade the recognition robustness.

Our deduction was already foreseen after making the DCL parameterization for the MoTiV database, since speech samples can differ in many parameters not just from loudness but most important from the level and characteristic of the added noise.

8 Conclusion

In order to make speech recognition more robust preprocessing influencing the physical characteristics of the mobile device transducers can be applied. The amount of varying background noise is a problem for all mobile devic-

es especially for automatic systems which are much more sensitive to the variability of the acoustic signal than humans. Therefore a pre-processing scenario was introduced improving speech recognition error rates with microphone transfer function adaptation limiting the background noise variations with a DCL preserving speech robustness.

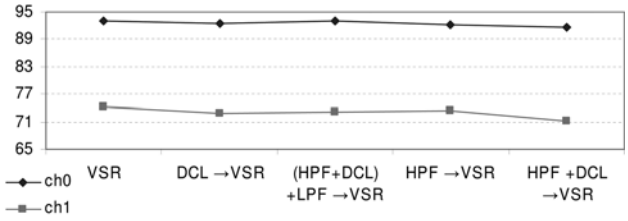


Fig. 8: The WCR for the supplemental experiments on the Aurora 3 database with the applied DCL.

We can observe that the applied Bi-Quad module considerable improves the Word Correct Rate (>6%) of the recognition module. Experiments also show that the DCL module applied for background noise variations elimination does not essentially degrade the gained improvement with the transfer function adaptation. Hence an assumption can be made that the introduced framework is preserving the speech robustness of the embedded VSR. There are strong indices that there can be a single DCL parameterization which would guarantee constant recognition results for arbitrary speech samples.

9 References

/1/ F. Hilger, H. Ney: Quantile Based Histogram Equalization for Noise Robust Large Vocabulary Speech Recognition, IEEE Transactions On Speech And Audio Processing, pp: 845 – 854, Volume 14, Issue 3, 2006.

/2/ S. Aalburg, C. Beaugeant, S. Stan, T. Fingscheidt, R. Balan, J. Rosca; Single- And Two-Channel Noise Reduction For Robust Speech Recognition In Car, Proceedings of ISCA Workshop, Multi-Modal Dialogue in Mobile Environments, Germany, 2002.

/3/ C. Beaugeant, M. Schönle, I. Varga; Challenges of 16 kHz in Acoustic Preand Post-Processing for Terminals, IEEE Communications Magazine, May 2006.

/4/ H. Gustafsson, I. Claesson, U. Lindgren; Low-Complexity Feature-Mapped Speech Bandwidth Extension," IEEE Transactions On Speech And Audio Processing, pp: 577- 588, Vol. 14, Issue 2, 2006.

/5/ 3GPP TS 26.131, "Technical Specification Group Services and System Aspects, Terminal Acoustic Characteristics for Telephony, Requirements (Release 6),", 2004.

/6/ U. Zölzer, Digitale Audiosignalverarbeitung, die 2. durchgesehene Auflage - Stuttgart : Teubner, 1997.

/7/ D. Langmann, H. R. Pfitzinger, T. Schneider, R. Grudszus, A. Fischer, M. Westphal, T. Crull, U. Jekosch, CSDC - The MoTiV Car Speech Data Collection. LREC98, 1998.

/8/ A. V. Oppenheim and R. W. Schafer; Digital Signal Processing, Englewood Cliffs, NJ, Prentice-Hall, 1975.

/9/ J. O. Smith; Introduction to Digital Filters with Audio Applications, W3K Publishing, http://books.w3k.org/, 2007.

/10/ Evaluations and Language resources Distribution Agency (http://www.elda.org/).

/11/ I. Varga et. all; ASR in Mobile Phones - An Industrial Approach, IEEE Transactions On Speech And Audio Processing, Vol. 10, No. 8, 2002.

/12/ I. McCowan, D. Moore, J. Dines, D. G.-Perez, M. Flynn, P. Wellner, H. Bourlard; On the Use of Information Retrieval Measures for Speech Recognition Evaluation, IDIAP Research Report 04-73, 2005.

/13/ 3GPP TS 26.071, AMR speech CODEC, http://www.3gpp.org/ ,2007.

Janez Stergar, Dejan Miletić
University of Maribor
Faculty of Electrical Engineering and Computer Science
Smetanova 17, 2000 Maribor, Slovenia
janez.stergar@uni-mb.si

Christophe Beaugeant, Bruno Tramblay
Siemens AG, ICM Mobile Phones
Grillparzerstrasse 10-18, 81675 Munich, Germany

Prispelo (Arrived): 27.03.08 *Sprejeto (Accepted):* 28.5.08

ANALIZA, MODELIRANJE IN SIMULACIJA VPLIVA PROMETA APLIKACIJ ZA IZMENJAVO DATOTEK P2P NA ZMOGLJIVOST OMREŽIJ

Matjaž Fras, Jože Mohorko, Žarko Čučej

Fakulteta za elektrotehniko, računalništvo in informatiko
Univerza v Mariboru, Maribor, Slovenija

Kjučne besede: P2P promet, samopodobnost, aplikacija, simulacija

Izvleček: V zadnjih letih doživlja delež prometa, ki ga ustvarjajo P2P aplikacije za izmenjavo datotek skokovito rast. S svojo količino in intenzivnostjo znatno vpliva na zmogljivost omrežij. Namen članka je predstavitev metode za analizo in modeliranje takšnega prometa za potrebe simulacij. Na osnovi izmerjenega prometa smo ocenili statistične parametre za modeliranje tovrstnega prometa. V simulacijskem okolju OPNET smo zgradili testno omrežje, kjer se pretakajo poleg ostalih prometov še promet aplikacij P2P, ki smo ga modelirali s predlagano metodo. S simulacijami smo preučili vpliv P2P prometa na zmogljivost omrežja.

Analysis, modeling and simulation of P2P file sharing traffic impact on networks' performances

Key words: P2P traffic, self-similarity, application, simulation

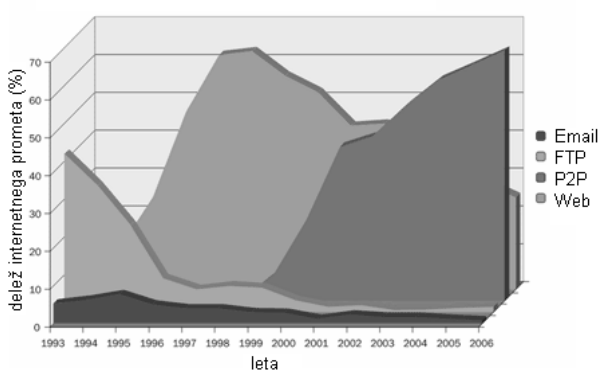
Abstract: Last year's CachneLogic research shows that, in January 2006, P2P traffic accounted for approximately 71% of all internet traffic, and was the main reason for internet traffic increasing. Figure 1 shows the trend of P2P traffic increase. The pioneer of P2P file sharing applications is very well-known Napster, which was created in 1999 by Shawn Fanning and used for sharing music files. Napster was finally destroyed by the music industry after Court proceedings. Napster, as a concept, had a great influence on developing file-sharing applications, such as eMula, Kazza, µTorrent, LimeWire, Skype etc.

In our research, we paid attention to the modeling of network traffic caused by P2P file sharing applications (P2P traffic). We measured P2P traffic using the Wireshark network traffic capture tool as shown in Figure 2. We created a self-similarity and long-range dependence analysis for the measured tests P2P traffics, the results of which are shown in Table 1, and in Figures 3 and 4. We also estimated the parameters for a statistical description of a P2P network separately for the processes of packet-size and inter-arrival time, which represent the main network traffic processes. The histogram method for distribution parameters' estimation is shown in Figure 5. These estimated distribution parameters were used for modeling a customized application of P2P network traffic in OPNET simulation tool. Using simulations with simple test networks, as shown in Figure 6, we represented the possibilities of simulating the impact of P2P network traffic on network performances. We show how the volume of P2P users impacts on other applications' performances (web time response shown in Figure 7 and link utilization shown in Figure 8.)

1. Uvod

Trendi uporabe P2P (peer-to-peer) aplikacij za izmenjavo datotek iz leta v leto naraščajo, kar je pokazalo tudi poročilo CachneLogic research /1/. Tako je bilo januarja leta 2006 kar 71 procentov vsega svetovnega prometa prav prometa P2P, kar je posledica populizacije P2P aplikacij. Slika 1 kaže deleže prometa različnih aplikacij celotnega spletnega prometa zadnjih nekaj let. Začetek razvoja P2P aplikacij sega v leto 1999, ko je Shawn Fanning z Univerze Northeastern iz Bostona razvil prvo večjo in razvpito P2P aplikacijo z imenom Napster, ki je omogočala ilegalno izmenjavo glasbenih datotek. Glasbeni industriji je s podporo glasbenih izvajalcev le uspelo Napsterju s pomočjo sodišča preprečiti velik pohod. Kljub temu pa je Napster povzročil revolucionarni vpliv na načine (legalne in nelegalne) distribucije intelektualne lastnine. Pripomogel je k velikem razcvetu P2P aplikacij v začetku 21. stoletja, kot so eMula, eDonkey, Kazza, LimeWire, µTorrrent, Skype…

V članku bomo predstavili metode merjenja, analize in modeliranju omrežnega prometa, ki ga povzročajo P2P



Slika 1: Poročilo CahneLogic o rasti internetnega prometa med leti 1993 in 2006 /1/.

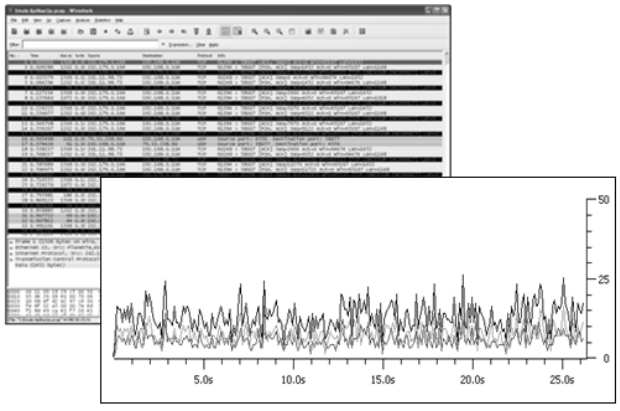
aplikacije (P2P promet). S pomočjo analize prometa običajnega uporabnika P2P aplikacij smo določili parametre statističnega opisa prometa. Te smo uporabili pri modeliranju za namene simuliranja takšnega prometa v simulacijskih orodjih, kot je npr. OPNET. Analizo P2P prometa smo izvedli na podlagi večih testnih vzorcev P2P prometov, ki smo

jih izmerili s pomočjo programa za zajemanje omrežnega prometa (vohljači) v različnih omrežjih z različnimi prenosnimi zmogljivostmi. Izvedli smo analizo omrežnega prometa s stališča samopodobnosti prometa s pomočjo ocene Hurstovega parametra ter dolgega območja odvisnosti na podlagi avtokorelacijske funkcije /2, 3, 4, 5/. V okvirju analize prometa smo določili porazdelitve, ki se najbolje prilegajo histogramom izmerjenega prometa in ocenili pripadajoče parametre statističnega opisa naključnih procesov omrežnega prometa in sicer procesa velikosti paketov in časa med paketi. Ta dva procesa smo nato uporabili pri modeliranju aplikacije po meri v simulacijskem orodju OPNET. V simulacijskem orodju smo nato modelirali še primer testnega omrežja v katerem smo pokazali vpliv P2P aplikacij na omrežne zmogljivost, ter odziv standardnih aplikacij kot so email, web browsing, FTP,…

Članek sestavljajo naslednja poglavja. Drugo poglavje podaja metodo merjenja P2P prometa v omrežjih. Tretjo poglavje podaja kratko matematično ozadje samopodobnosti in dolgega območja odvisnosti, ter rezultate analize izmerjenih testnih prometov. Četrto poglavje podaja modeliranje prometa P2P aplikacije na podlagi opravljene analize. Peto poglavje podaja kratek opis testnega omrežja v simulacijskem okolju OPNET. V šestem poglavju so rezultati simulacij, kjer smo preučevali vpliv P2P aplikacij na lastnosti omrežja glede na število uporabnikov P2P aplikacij. Na koncu sledi še zaključek.

2. Merjenje P2P prometa

Merjenje prometa, ki ga ustvarjajo aplikacije za izmenjavo datotek (P2P promet) smo izvedli s pomočjo programa za zajemanje prometa (popularno imenovan vohljač) Wireshark /12/. Tako ustvarjen promet za posamezne uporabnike tovrstnih aplikacij smo zajemali v različnih omrežjih z različnimi prenosnimi zmogljivostmi. Vsaka P2P aplikacija za izmenjavo datotek v P2P omrežju ima vlogo tako odjemalca kot tudi ponudnika. Zato smo zajet promet razdelili na promet, ki ga uporabnik prejema (vloga odjemalca) in promet, ki ga oddaja (vloga strežnika). Pri meritvi prometa smo s filtri izločili promet drugih aplikacij.



Slika 2: Uporabniški vmesnik Wireshark vohljača, ter časovni potek zajetega prometa aplikacije P2P v paketih na časovno enoto in sicer celotni promet, promet odjemalca in ponudnika.

Tabela 1 prikazuje osnovne parametre izmerjenih testnih P2P prometov, ki smo jih zajeli s pomočjo Wireshark vohljača.

Tabela 1: Primeri izmerjenih P2P testnih prometov naključnega uporabnika P2P aplikacij.

testni promet	1. testni promet (eMula)		2. testni promet (µTorrent)	
	promet strežnika	promet odjemalca	promet strežnika	promet odjemalca
srednja vrednost (p/s)	101,66	88,96	1469,281	1093,795
srednja vrednost (kb/s)	87,513	22,094	1724,955	637,167
testni promet	3. testni promet (eMula)		4. testni promet (µTorrent)	
	promet strežnika	promet odjemalca	promet strežnika	promet odjemalca
srednja vrednost (p/s)	71,940	49,831	39,044	28,394
srednja vrednost (kb/s)	71,940	37,24	37,176	13,344

3. Analiza prometa

3.1 Samopodobnost in dolgo območje odvisnosti

Analizo prometa smo izvedli z vidika samopodobnosti in dolgega območja odvisnosti. Opis omrežnega prometa z modelom samopodobnosti /2, 4, 5, 6, 7/ temelji na teoriji fraktalov. Ta model je nadomestil starejše modele omrežnega prometa, kot sta Poissonov in Markov /3/. Samopodobnost temelji na samopodobnosti časovnih zaporedij, ki jih podaja naslednja definicija /5, 9 10/:

Naj bo $X = (X_t, t = 0, 1, 2, \ldots)$ kovariančno stacionarni stohastični proces s stacionarno srednjo vrednostjo $\mu = E/X_t/$, končno varianco $\sigma^2 = E[(X_t - \mu)^2]$, avtokovariančno funkcijo $\gamma(k) = E[(X_t - \mu)(X_{t+k} - \mu)]$, ki je odvisna le od vrednosti k, in avtokorelacijsko funkcijo $r(k)$;

$$r(k) = \frac{\gamma(k)}{\sigma^2} = \frac{E\left[(X_t - \mu)(X_{t+k} - \mu)\right]}{E\left[(X_t - \mu)^2\right]}, \qquad k = 0,1,2,\ldots \quad (1)$$

Potem se avtokorelacijska funkcija asimptotično približuje

$$r(k) \approx k^{-\beta} L_1(k), \quad k \to \infty, \quad 0 < \beta < 1, \quad (2)$$

kjer je $L_1(k)$ počasi se spreminjajoča funkcija v neskončnosti in za katero velja $\lim_{t\to\infty} = L_1(tx)/L_1(t) = 1$ pri vseh $x > 0$ (npr. $L_1(t)$ = konstanta, $L_1(t) = \log(t)$). Če časovno zaporedje izpolnjuje pogoj v enačbi 2, potem takšno časovno zaporedje opisujemo kot samopodobno časovno zaporedje. Merilo samopodobnosti predstavlja Hurstov parameter, katerega določimo iz parametra β in sicer:

$$H = 1 - \frac{\beta}{2}, \qquad 0 < \beta < 1 \quad (3)$$

Iz (3) sledi, da se v primeru samopodobnega prometa vrednost Hurstovega parametra med 0,5 in 1.

Samopodobnost lahko vsebuje lastnost dolgega območja odvisnosti, pri katerem je vrednost novega stanja zelo močno odvisna od prejšnjih stanj. Značilnost takega procesa je visok nivo spremenljivosti na različnih časovnih skalah. V primeru agregiranega procesa na dolgih časovnih skalah pa proces ne teži h glajenju. Definirajmo stohastični proces, ki izpolnjuje relacijo 2, ter avtokorelacijsko funkcijo samopodobnosti drugega reda s predpisom $r(k) = \gamma(k)/\sigma^2$. Za vrednosti $0 < H < 1$, $H \neq 0{,}5$ velja

$$r(k) \approx H(2H-1)k^{-2H-2}, \qquad r \to \infty \tag{4}$$

Za vrednosti $0{,}5 < H < 1$ se avtokorelacijska funkcija $r(k)$ asimptotično obnaša kot $ck^{-\beta}$ za vrednosti $0 < \beta < 1$, kjer je konstanta $c > 0$, $\beta = 2 - 2H$ in velja

$$\sum_{k=-\infty}^{\infty} r(k) = \infty \tag{5}$$

Avtokorelacijska funkcija pada hiperbolično (počasi), zato ni sumabilna. Procesi s hiperboličnim upadanjem $r(k)$ so stacionarni procesi z dolgim območjem odvisnosti. Kadar ima avtokorelacijska funkcija $r(k)$ procesa $X(t)$ končno vsoto, potem je proces s kratkim območjem odvisnosti. Eksponentno padanje avtokorelacijske funkcije

$$r(k) \approx \rho^{k}, \quad k \to \infty, \quad 0 < \rho < 1 \tag{6}$$

povzroči, da ima avtokorelacijska funkcija končno vsoto

$$0 < \sum_{k=-\infty}^{\infty} r(k) < \infty \tag{7}$$

3.2 Analiza izmerjenih testnih P2P prometov

Za izmerjene testne promete iz tabele 1 smo opravili analizo samopodobnosti z oceno Hurstovega parametra in testom dolgega območja odvisnosti na podlagi avtokorelacijske funkcije. Hurstovega parametra H kot merilo samopodobnosti ne moremo natančno izračunati, ampak ga lahko le ocenimo /4, 5, 6, 7/. Obstaja več različnih metod ocenjevanja, njihovi rezultati medsebojno bolj ali manj odstopajo. Pri tem ni kriterijev, s katerimi bi ugotovili, katera metoda bo dala najboljši rezultat. Najpogosteje uporabljene metode za oceno Hurstovega parametra so variančna, R/S, periodogram, Whittle in Arby-Veitch metoda. Ocenjene rezultate parametra H z omenjenimi metodami za testne promete iz tabele 1 podaja tabela 2. Slika 3 prikazuje primere estimacije Hurstovega parametra z variančno, R/S in periodogram metodo za 4. testni promet (promet µTorrent odjemalca) iz tabele 1.

Na osnovi analize različnih testnih primerov P2P prometa smo ugotovili, da je omrežni promet aplikacij P2P v večini primerov samopodoben saj so ocenjene vrednosti Hurstovega parametra v večini primerov večje od vrednosti 0,5. Z analizo dolgega območja odvisnosti na podlagi avtokorelacijske funkcije smo videli, da ta pada eksponentno ter teži k vrednosti 0. To pomeni, da takšen omrežni promet vsebuje lastnost kratkega območja odvisnosti. Iz izmerjenih testnih prometov aplikacije P2P vidimo, da se ti najbolj razlikujejo v srednjih vrednosti prometa tako v bitih in paketih na časovno enoto. Srednje vrednosti P2P prometa se v večini primerov s časom spreminja. Srednja vrednost (okensko povprečenje) P2P prometa je odvisna odvisno od razpoložljivih soležnikov (peers), sejalcev (seeds), pijavk (leech) in jat (swarm), ki so na voljo v P2P omrežju za datoteke, ki jih prenašamo. V nekaterih primerih se lahko zgodi, da aplikacije µTorrent težijo že kar h konstantni prenosni hitrosti, sploh če imamo na voljo veliko pasovno širino.

Tabela 2: Ocenjene vrednosti Hurstovega parametra izmerjenih testnih prometov različnih P2P aplikacij (tabela 1) z različnimi metodami za oceno parametra H.

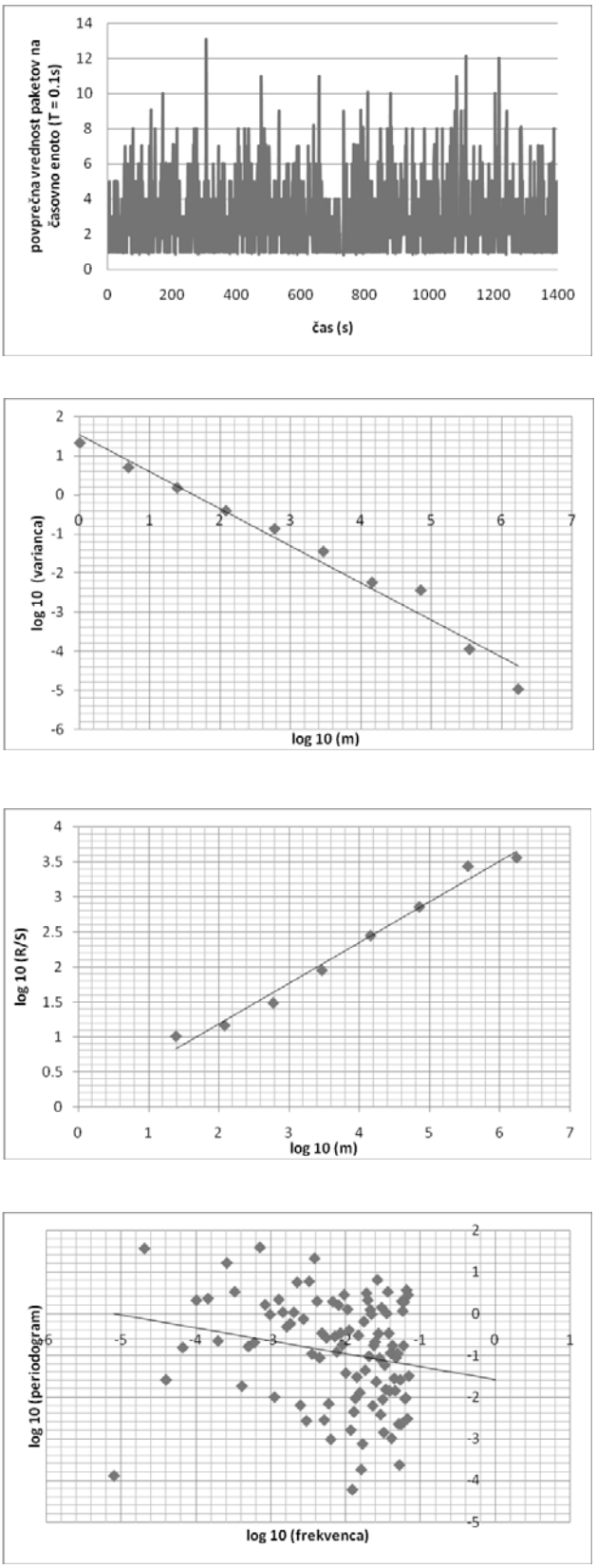
testni promet	1. testni promet (eMula)		2. testni promet (µTorrent)	
	promet strežnika	promet odjemalca	promet strežnika	promet odjemalca
variančna metoda	0,682	0,482	0,7	0,803
R/S metoda	0,509	0,504	0,533	0,457
metoda periodogram	0,765	0,994	0,806	1
Whittle estimator	0,535	0,586	0,58	0,689
Arby-Veich	0,551	0,557	0,719	0,685
testni promet	3. testni promet (eMula)		4. testni promet (µTorrent)	
	promet strežnika	promet odjemalca	promet strežnika	promet odjemalca
variančna metoda	0,382	0,699	0,464	0,525
R/S metoda	0,406	0,406	0,505	0,505
metoda periodogram	0,454	0,343	0,589	0,655
Whittle estimator	0,545	0,538	0,537	0,559
Arby-Veitch	0,544	0,568	0,627	0,613

Slika 4 prikazuje primer avtokorelacijske analize dolgega območja odvisnosti omrežnega 4. testnega prometa (promet µTorrent odjemalca).

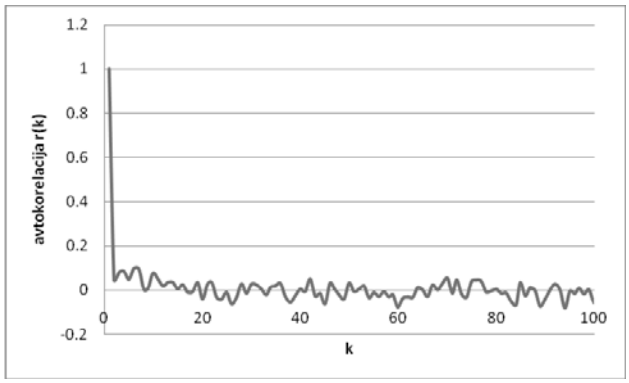
4. Modeliranje P2P aplikacije

Modeliranje omrežnega prometa v simulacijskih okoljih pogosto temelji na porazdelitvah verjetnosti za naključne procesa prometa kot sta procesa velikosti paketov in časa med paketi. Naključne procese omrežnega samopodobnega prometa opisujemo s počasi pojemajočimi porazdelitvami verjetnosti (heavy tailed distributions) /4, 5/. Pri počasi pojemajočih porazdelitvah verjetnost upada hiperbolično, za razliko z hitro pojemajočimi porazdelitvami, kot je to na primer eksponentna porazdelitev, kjer verjetnost upada eksponentno. Eksponentno porazdelitev opisuje naslednja porazdelitvena funkcija/16/:

$$p(x) = \frac{1}{\mu} \cdot e^{-\frac{x}{\mu}} \tag{8}$$



Slika 3: Ocenjevanje Hurstovega parametra H *za 4. testni promet (promet µTorrent odjemalca) P2P aplikacij (zgoraj-levo) z različnimi metodami. Zgoraj-desno ocenjevanje z variančno metodo, spodaj-levo z R/S metodo in spodaj-desno z metodo periodograma.*



Slika 4: Avtokorelacijska funkcija r(k) *procesa za 4. testni promet P2P prometa (promet µTorrent odjemalca).*

Parameter $1/\mu$ večkrat nadomestimo s parametrom λ, ki ga imenujemo parameter intenzivnosti.

Paretova porazdelitev je najpreprostejša počasi pojemajoča porazdelitev verjetnosti in jo opišemo z naslednjo porazdelitveno funkcijo/17/:

$$p(x) = \alpha k^{\alpha} \cdot x^{-\alpha-1}, \quad k \le x, \quad \alpha, k > 0 \tag{9}$$

Parameter α imenujemo tudi parameter oblike (shape). Parameter k imenujemo parameter lokacije (local), ki je najmanjša možna vrednost naključne spremenljivke x.

Poleg Paretove porazdelitve iz družine počasi pojemajočih porazdelitev, je pomembna tudi Weibullova /18/:

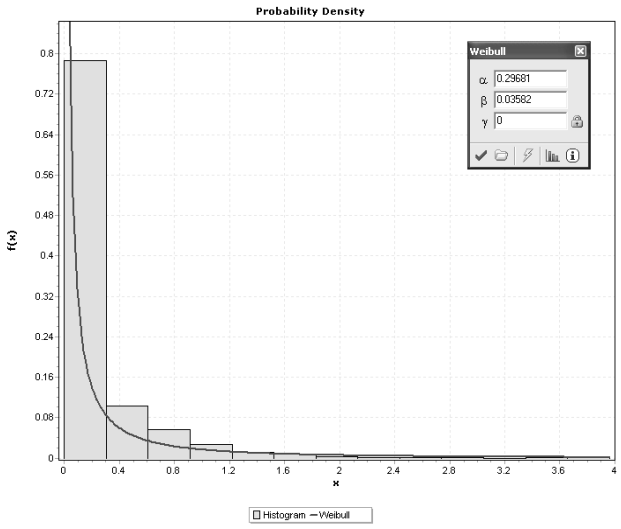
$$p(x) = \frac{\alpha}{k} \cdot \left(\frac{x}{k}\right)^{\alpha-1} \cdot e^{-\left(\frac{x}{k}\right)}, \quad x \ge 0, \quad \alpha, k > 0 \tag{10}$$

Parameter α imenujemo parameter oblike. Parameter k imenujemo parameter lokacije.

Verjetnostne porazdelitve ter pripadajoče parametre porazdelitve stohastičnega procesa navadno ocenimo na osnovi histogramov, ki nam služijo kot referenca pri izbiri najbolj prilegajoče porazdelitve ter izračunu parametrov izbrane porazdelitve. Histogram je diskretni približek zvezne funkcije gostote verjetnosti.

Pri analizi izmerjenega omrežnega prometa, statistično ovrednotimo oba naključna procesa prometa na podlagi histogramov in sicer proces velikosti paketov in procesa časa med paketi. Oceno porazdelitve in parametrov omrežnega P2P prometa smo izvedli s pomočjo orodja EasyFit, ki je prikazano na sliki 4. Natančnost aproksimacije histogramov s porazdelitvami verjetnosti ocenimo s pomočjo testov kot sta Kolmogorov-Smirnov ali Chi-squere /13, 14, 15, 20/.

Na podlagi opravljene analize izmerjenih testnih prometov (tabeli 1 in 2) smo zasnovali model P2P aplikacije za izmenjavo datotek. V simulacijskem okolju OPNET imamo na voljo veliko različnih načinov opisa omrežnega prometa /11/. Odločili smo se za modeliranje prometa s pomočjo aplikacij saj nam simulacijsko okolje ponuja že v naprej



Slika 5: Primer ocene parametrov Weibullove porazdelitve za proces časa med paketi na podlagi izmerjenega histograma.

pripravljene aplikacije, kot so npr. email, FTP ali spletno brskanje ter tudi možnost modeliranja aplikacije po meri, ki smo jo uporabili za modeliranje P2P prometa. P2P aplikacijo za izmenjavo datotek v simulaciji smo modelirali na podlagi aplikacije »database access«, v katero smo vnesli parametre, ki smo jih ocenili s pomočjo opisanih metod. Za opis procesa časa med transakcijami smo izbrali Weibullovo porazdelitev, medtem ko smo za opis procesa velikosti transakcij izbrali eksponentno porazdelitev. Za eksponentno porazdelitev smo se odločili na podlagi analize prometa, kjer smo s pomočjo avtokorelacijske funkcije ugotovili, da samopodoben promet aplikacij P2P v večini primerov vsebuje lastnost kratkega območja odvisnosti. Vrednost Hurstovega parametra je okoli 0,5, kar predstavlja tudi mejo med lastnostjo dolgega in kratkega območja odvisnosti. Izmerjeni prometi prav tako ne kažejo ekstremnih konic, kar je lastnost prometov z dolgim območjem odvisnosti, katere je primerno modelirati s Paretovo porazdelitvijo za proces velikosti paketov. P2P aplikacijo za izmenjavo datotek smo definirali tako, da nam v smeri od strežnika do odjemalca povzroči okoli 100kb/s kar znaša okoli 35paketov/s ter v obratni smeri (odjemalec strežnik) približno polovico prometa strežnik-odjemalec.

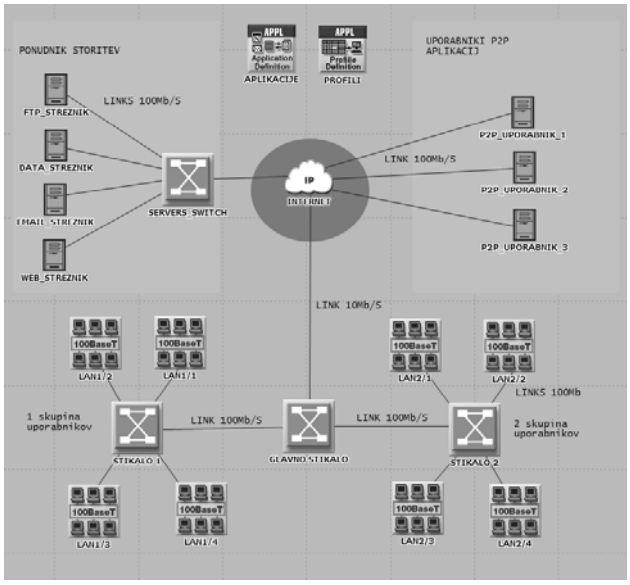
5. Testno omrežje

Za preučevanje vpliva P2P aplikacij za izmenjavo datotek na zmogljivost omrežja smo v simulacijskem okolju OPNET sestavili manjšo testno omrežje, ki predstavlja omrežje podjetja. Testno omrežje sestavljata dve manjši skupini uporabnikov, v katerih je po 40 (4x10) uporabnikov. V celotnem omrežju je torej 80 uporabnikov razdeljenih v skupine po 10 uporabnikov povezanih preko stikal, kot je prikazano na sliki 6. Vsaka skupina je povezana na stikalo s povezavo 100Mb/s. Stikali skupin sta povezana na glavno stikalo s 100Mb/s povezavo. To stikalo je povezano z internetom in dejansko predstavlja pasovno širino, ki jo lahko končni uporabnik zakupi od ponudnika internetnih storitev. Ta povezava predstavlja zakupljen ali najet vod z prenosno zmogljivostjo 10Mb/s. Na internet imamo priklopljene strežnike, ki predstavlja strežniško arhitekturo ponudnika internetnih storitev. S temi strežniki smo podprli standardne aplikacije, ko so brskanje po internetu, elektronska pošta in FTP prenos. Vsi strežniki so povezani na strežniško stikalo, ki je povezano v internet s povezavo 100Mb/s. Prav tako smo na internet povezali strežnike 3 zunanje P2P uporabnike na katere se bodo povezovali uporabniki P2P aplikacij iz podjetja. S simulacijo smo se skušali čim bolj približati realnim razmeram v komunikacijskih omrežjih.

V simulacijskem orodju smo modelirali tri različne scenarije in sicer:

1. Vsi uporabniki (80) internet storitev uporabljajo le standardne aplikacije in sicer elektronsko pošto, FTP prenosi datotek in brskanje po spletu.
2. Ob tem da vsi uporabniki (80) uporabljajo standardne aplikacije prva skupina 40 uporabnikov uporablja še P2P aplikacije.
3. Ob tem, da vsi uporabniki (80) uporabljajo standardne aplikacije uporabljajo tudi vsi aplikacije P2P.

Z različnim številom uporabnikov P2P aplikacij v testnem omrežju smo skušali pokazati, kako takšna uporaba vpliva na zmogljivosti omrežja. V simulacijskem okolju smo spremljali različne parametre omrežja, kot so zakasnitve, odzivni čas spletne strani, izkoriščenost povezav, kot bo prikazano v naslednjem poglavju.

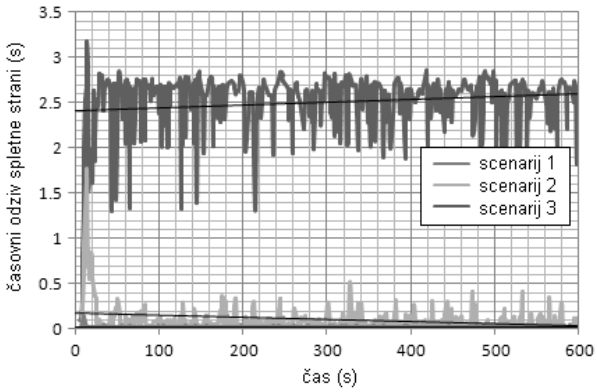


Slika 6: Testno omrežje za simuliranje vpliva prometa aplikacij P2P na zmogljivost omrežja

6. Rezultati simulacij

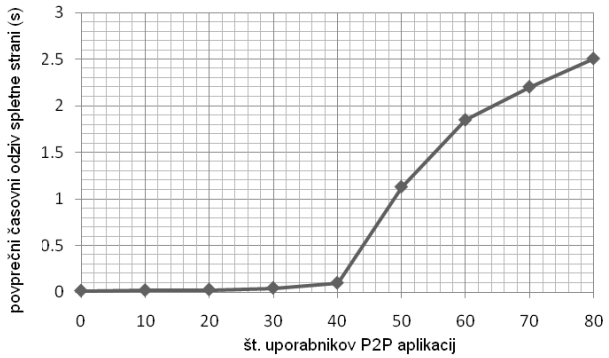
Iz rezultatov simulacij treh scenarijev testnega omrežja iz slike 6, v katerem smo spreminjali število uporabnikov P2P aplikacij za izmenjavo datotek vidimo, kako uporaba le teh

vpliva na razmere v celotnem omrežju. S povečevanjem uporabnikov P2P aplikacij se poveča P2P promet v omrežju, kar ima negativen vpliv na druge uporabljene aplikacije v omrežju. Graf na sliki 7 prikazuje odzivni čas spletne strani za vse tri scenarije. Kot smo tudi pričakovali, se v primerih večje uporabe P2P aplikacij povečuje odzivni čas spletnih strani.



Slika 7: Odzivnost spletne strani v odvisnosti od števila P2P uporabnikov na omejeni pasovni širini 10Mb/s za tri različne scenarije.

Graf na sliki 8 prikazuje vpliv števila uporabnikov P2P aplikacij za deljenje datotek na povprečni časovni odziv spletne strani.

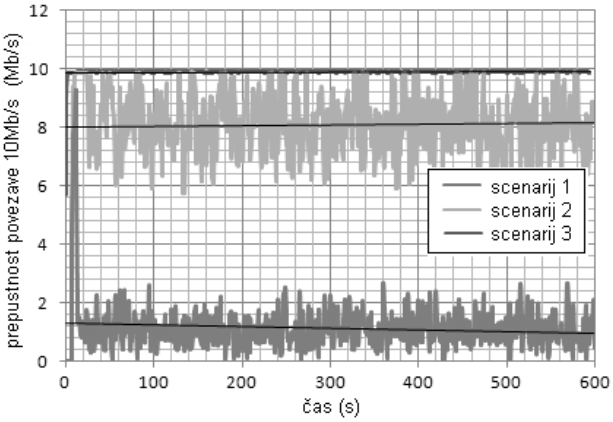


Slika 8: Povprečni časovni odziv spletne strani v odvisnosti od števila uporabnikov P2P aplikacij za deljenje datotek v testnem omrežju.

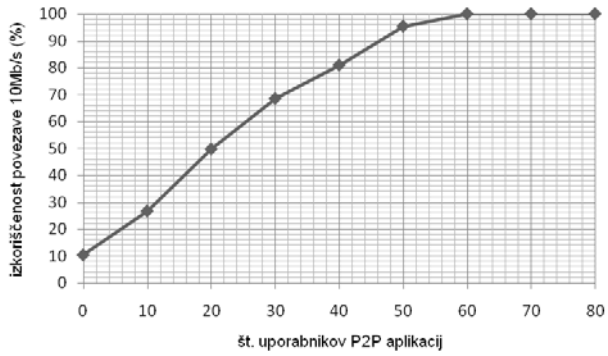
Slika 9 prikazuje izkoriščenost najete povezave, ki znaša 10Mb/s. V primeru uporabe standardnih aplikacij je izkoriščenost okoli 18 odstotna, v primeru naraščanja uporabnikov pa ta strmo narašča in pride v primeru, ko vsi uporabniki uporabljajo P2P aplikacije, v zasičenje.

Slika 10 prikazuje izkoriščenosti najetega voda 10Mb/s podjetja v odvisnosti od števila uporabnikov P2P aplikacij za deljenje datotek v testnem omrežju na sliki 6.

S simulacijo takšnega omrežja omogoča vpogled delovanja takšnega omrežja ter oceno vpliva posameznih parametrov na zmogljivost omrežja.



Slika 9: Prepustnost prenosne povezave men internetom in glavnim stikalom kapacitet 10Mb/s.



Slika 10:Izkoriščenost najetega voda v odvisnosti od števila uporabnikov P2P aplikacij.

6. Zaključek

V članku smo predstavili metode analize in modeliranja omrežnega prometa na osnovi izmerjenega prometa, za namene simulacij. Kot primer uporabe metod smo izbrali preučevanje vpliva uporabe P2P aplikacij za izmenjavo datotek na lastnosti omrežja. Z meritvami in analizo smo ugotovili, da je v večini primerov promet P2P aplikacij samopodoben, vendar pa ne vsebuje lastnosti dolgega območja odvisnosti, kar smo upoštevali v postopku modeliranja za opis statističnih procesov prometa. Iz rezultatov simulacije smo videli, da uporaba takšnih aplikacij zmanjša zmogljivost omrežja in negativno vpliva na ostale aplikacije, ki so nato deležne večjih zakasnitev.

Seveda pri uporabi P2P aplikacij za izmenjavo datotek v javnih ustanovah in podjetjih ni potencialno škodljiva le zaradi zmogljivosti omrežja. Negativno vpliva tudi na delovno storilnost, povečuje možnost vdorov virusov, vzpodbuja nelegalno uporabo programske opreme in multimedijskih vsebin.

Tehnični vidik teh problemov bi lahko rešili z zagotavljanjem kakovosti storitev (QoS), vendar se pa sistemski administratorji po navadi zatečejo k drugim rešitvam, kot so na primer administrativne prepovedi, zaklepanje vtičev (portov), omejevanje pravic v operacijskih sistemih, itd.

Reference

/1/ http://torrentfreak.com/bittorrent-the-one-third-of-all-internet-traffic-myth/.

/2/ W. E. Leland, M. S. Taqqu, W. Willinger in D. V. Wilson, On the self-similar nature of Ethernet traffic (Extended version), IEEE/ACM Transactions on Networking, Vol.2, pp.1–15, 1994.

/3/ V. Paxon in S. Floyd, Wide area traffic: the failure of Poisson modeling, IEEE/ACM Transactions on Networking, 3(3): 226–244, 1995.

/4/ O. Sheluhin, S. Smolskiy in A. Osin (Author), Self-Similar Processes in Telecommunications, John Wiley & Sons, 2007.

/5/ K. Park in W. Willinger, Self-Similar Network Traffic and Performance Evaluation, John Wiley & Sons, 2000.

/6/ T. Karagiannis, M. Molle in M. Faloutos, Understanding the limitations of estimation methods for long-range dependence, University of California.

/7/ M. Fras, J. Mohorko in Z. Cucej, Estimating the parameters of measured self similar traffic for modeling in OPNET, IWSSIP Conference, 2007, Maribor, Slovenia.

/8/ W. Willinger, M. S. Taqqu, R. Sherman in D. V. Wilson, Self-similarity through high-variability: statistical analysis of Ethernet LAN traffic at the source level, IEEE/ACM Transactions on Networking, 5(1): 71–86, 1997.

/9/ H. Yőlmaz, IP over DVB: Management of self similarity, Master of Science, Bođaziçi University, 2002.

/10/ M. Z. Jiang, Analysis of wireless data network traffic, Master of Applied Science, Simon Fraser University, Vancouver, Canada, 2000.

/11/ J. Mohorko, M. Fras in Ž. Čučej, Modeling of IRIS replication mechanism in tactical communication network with OPNET, IWSSIP Conference, 2007, Maribor, Slovenia.

/12/ A. Orebaugh, G. Ramirez in J. Beale, Wireshark & Ethereal Network Protocol Analyzer Toolkit, Syngress Publishing, Inc., 2007.

/13/ Chakravarti, Laha, in Roy, (1967). Handbook of Methods of Applied Statistics, Volume I, John Wiley and Sons, pp. 392-394.

/14/ http://en.wikipedia.org/wiki/Kolmogorov-Smirnov_test

/15/ http://en.wikipedia.org/wiki/Chi-square_goodness-of-fit_test.

/16/ http://en.wikipedia.org/wiki/Exponential_distribution.

/17/ http://en.wikipedia.org/wiki/Pareto_distribution.

/18/ http://en.wikipedia.org/wiki/Weibull_distribution.

/19/ R. Jamnik, Verjetnostni račun in statistika, DZS, Ljubljana, 1987.

/20/ B. Vujičić, N. Cackov, S. Vujičić in L. Trajković, Modeling and Characterization of Traffic in Public Safety Wireless Networks, Simon Fraser University, Vancouver, Canada, SPECTS 2005

Matjaž Fras, Jože Mohorko, Žarko Čučej
Fakulteta za elektrotehniko,
računalništvo in informatiko
Univerza v Mariboru, Smetanova 17,
2000 Maribor, Slovenija

Tel: (+386 2) 220-7120 Email: matjaz.fras1@uni-mb.si

Prispelo (Arrived): 04.04.07 *Sprejeto (Accepted):* 28.5.08

PREDSTAVITEV OMREŽJA UMTS IN NJEGOVA SIMULACIJA S POMOČJO SIMULACIJSKEGA ORODJA OPNET MODELER

Jože Mohorko, Saša Klampfer

Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko, Maribor, Slovenija

Kjučne besede: UMTS omrežje, 3G infrastruktura, Opnet Modeler, gradniki segmentov, kvaliteta sprejema, odzivni čas, aplikacije, domena.

Izvleček: Članek opisuje posamezne gradnike omrežja UMTS (Universal Mobile Telecommunication System), in njegovo simulacijo s programom Opnet Modeler. V fazi simuliranja omrežja se bomo omejili na dejavnike, ki vplivajo na kakovost posamezne storitve (aplikacije), predvsem pa na vpliv oddaljenosti in kvaliteto sprejema. Slednjo smo ovrednotili s pomočjo odzivnih časov pri spletni aplikaciji in telefoniji IP (ang. Internet Protocol).

Presentation of UMTS network and his simulation using OPNET Modeler

Key words: UMTS network, 3G infrastructure, Opnet Modeler, segment build elements, receiving quality, response time, applications, domain.

Abstract: Now days the use of mobile communications, applications and mobile data transfer are rapidly increased. The standardization work of GSM-based systems has begun in the 1980s, when the developing of unique radio communications system for Europe, at 900MHz, has started. Since then many modifications have been made in order to fulfill the increasing demand from the operators and cellular users.

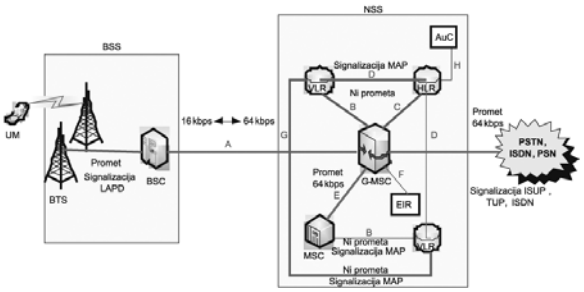
This paper describes a Universal Mobile Telecommunications Service (UMTS) network and an example of data reception, when web application and voice over internet protocol (VoIP) are used. The UMTS represent third generation (3G), broadband packet based transmission of text, video, digitized voice and multimedia at data rates up to 2 megabits per second (Mbps). The UMTS is intention for consistent set of services to mobile computers and phone users, no matter where they are located in the world. Third generation technology is based on the Global System for Mobile (GSM) communication standard. It is also endorsed by major standards, bodies and manufacturers, as the planned standard for mobile users around the world. Because UMTS is today fully available, computers and phone users can be constantly attached to the internet wherever they are, wherever they travel and as they roam, will have exactly the same set of capabilities. Users will have access through a combination wireless and satellite transmissions. Even today, some places are not fully covered with UMTS, so in that case, users can use multimode devices that switch to the currently available technology, such as GSM 900 and 1800 where UMTS is not yet available. UMTS offers many different applications like: light and heavy web browsing, reading web mail, VoIP quality speech, video conferencing, base access, telnet session, file transfer, file copy, GSM and PCM quality speech, SMS, MMS and so on. We pick out only two applications, and that is light web browsing and VoIP, which are today very popular for most population on the world, because many applications allow user to connect with the whole world at any place and any time.

Under the third generation partnership project (3GPP) the third generation of UMTS cellular system was developed. The main parts of this systems are UMTS terrestrial radio access network (UTRAN), based on wide division multiple access (WCDMA) radio technology and GSM/EDGE radio access network (GERAN) based on global system for mobile telecommunications (GSM/enhanced data rates for global evolution (EDGE) radio technology). On the other hand the UMTS can be divided into three major parts: User Equipment (UE) that interfaces with the user and radio interface, UTRAN that handles all radio-related functionality, and the Core Network (CN), which is responsible for switching and routing calls and data connections to external networks. These elements are shown in Fig. 1. In second section the base UMTS elements and their function are presented on Fig. 2 and Fig. 3. Section 3 shortly represents the capabilities of the Opnet Modeler program and describes types of networks, which can be simulated. Section 4 represents the construction of UMTS network in the Opnet Modeler and describes the web application and VoIP. The web browsing is simulated in UMTS network application and object response time and page response time were observed. We located UMTS mobile node with name »Web2« near UMTS transmitter. If we compare positions for both nodes (Fig. 4), is very simple to say, that is node »Web2« closer to transmitter than node »Web1«. Both nodes are placed on fix position and both have identical parameters and identical settings, because of that, we can simply estimate remote distance influence at interdependence with object response time and page response time. Results of the simulations are shown on Fig. 5-6 and 8-9. Fig. 7 shows us UMTS structure for VoIP application. Section 5 concludes the paper.

1 Uvod

V zgodnjih osemdesetih letih se je pričel hiter razvoj analognih celičnih sistemov, ki predstavljajo prvo generacijo mobilnih telekomunikacij (1G). Problematika prve generacije je nastopila že ob razvoju tovrstnih sistemov, saj je vsaka država razvijala svoj sistem, kar je privedlo do omejene uporabnosti. Ta omejitev je bila povod za ustanovitev skupine »Groupe Speciale Mobile« (GSM) v Franciji, ki se je nato leta 1989 preimenovala v »Global System for Mobile telecommunications« in preselila v Anglijo. Skupina je bila zadolžena za postavljanje kriterijev mobilnih brezžičnih omrežij, ki so se nanašali na kvaliteto govora, nizke stroške za terminale in storitve, mednarodno gostovanje, upravljanje terminalov, spektralno učinkovitost, združljivost s PSTN in kasnejšim ISDN itd. Leta 1995 je bil pod okriljem iste skupine zaključen standard druge generacije mobilnih telekomunikacij (2G), katere predstavnik je bilo implementirano GSM mobilno omrežje. Le to v osnovi vsebuje tri osnovne podsisteme in sicer, omrežni podsistem NSS (ang. Network SubSystem), upravljaljski omrežni sistem NMS (ang. Network Management System) in podsistem

bazne postaje BSS (ang. Base Station Subsystem). NSS vsebuje mobilni stikalni center MSS (ang. Mobile Switching Centre) in stikalne prehode GMSC (ang. Gateway Switches), ki predstavljajo točke povezav med GSM mobilnim omrežjem in PSTN, ISDN napravami. NMS vključuje opremo za delovanje in vzdrževanje mobilnega omrežja OMC (ang. Operation and Maintenance Centre), medtem ko podsistem bazne postaje BSS vključuje opremo za upravljanje radijskega vmesnika med mobilno postajo in celičnim radijskim omrežjem.



Slika 1: Zgradba GSM omrežja
Fig. 1: GSM Architecture

Omrežje je sestavljeno iz mobilne postaje (MS), bazne oddajne postaje (BTS), nadzornika bazne postaje (BSC), podsistema bazne postaje (BSS), mobilnega stikalnega centra (MSC), authentifikacijskega registra (AuC), registra domače lokacije (HLR) in registra gostujoče lokacije (HLR), kot prikazuje slika 1. GSM uporablja kombinacijo dveh tehnologij dostopov med katere spadata TDMA (ang. Time Division Multiple Access) in FDMA (ang. Frequency Division Multiple Access) način. Z uporabo kombinacije obeh so-dostopov je bilo v GSM omrežju zagotovljenih več digitalnih prenosnih frekvenc, in sicer 450MHz, 900, 1800 in 1900 MHz. Višji kot je frekvenčni pas, večja je dodatna kapaciteta v prenosnem kanalu. Kljub hitremu razvoju GSM tehnologije in velike uporabnosti le te, je kmalu postalo jasno, da bo za sodobne aplikacije obstoječo tehnologijo potrebno nadgraditi. Ker je GSM omrežje v osnovi tokovno komutirano, se omejitev odraža v okrnjenem naboru razpoložljivih aplikacij med katere spadajo prenos zvoka, SMS sporočil, E-mail sporočil in WAP. Hiter razvoj IP tehnologije je prisilil snovalce k nadgradnji omrežja v smeri paketne komutacije, saj IP omrežja temeljijo zgolj na paketnih tehnologijah. V ta namen je prva izpeljanka do 3G omrežij (R99) že vsebovala oba načina komutacije, s čimer se je nabor aplikacij dodatno razširil. Dodatno so bile vpeljane še MMS, Web in Video aplikacije. Nadaljnji razvoj je potekal v smeri paketno komutiranega omrežja, ki ga definira izdaja 5-6 (ang. Release 5-6), ki predstavlja današnje UMTS omrežje tretje generacije (3G). Prehod iz 2G v 3G omrežje ni bil direkten, temveč so se vmes pojavile alternative kot so HSCSD, GPRS in EDGE, ki so predvsem operaterjem omogočile lažji prehod. Vmesna skupina sodi v družino 2.5G. Generacija 2.5G se razlikuje od predhodne predvsem v višji pasovni širini, ki je bila dosežena z združevanjem časovnih rezin. Dandanes so se komaj do dobra uveljavila 3G omrežja, pa vendarle že na vrata trka četrta generacija HSPDA (4G), ki se bo po navedbah lahko ponašala z znatno višjo pasovno širino (do 14 Mbps).

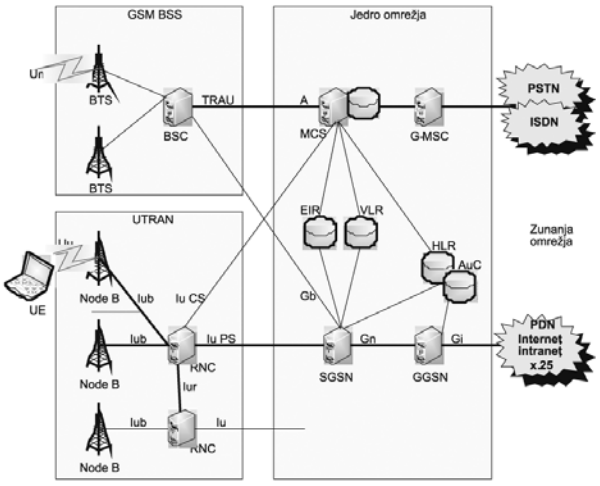
V tem članku bomo v drugem poglavju podrobneje opisali razvoj omrežja UMTS in podali njegove gradnike. V tretjem poglavju bomo predstavili aplikacijo UMTS omrežja in ga simulirali s pomočjo paketa Opnet Modeler. V četrtem poglavju bomo predstavili eksperimentalne rezultate in članek zaključili s petim poglavjem /1/.

Slika 2: Potek razvoja mobilnih generacij
Fig. 2: Development procedure of mobile generations

2 Poti razvoja UMTS omrežja

UMTS je tretja generacija (3G) mobilnih komunikacijskih sistemov, ki zagotavlja niz širokopasovnih storitev v svetu brezžičnih in mobilnih komunikacij. Predstavlja nizek strošek za operaterje, hkrati pa je sposoben zagotavljati pasovne širine do 2Mbps. Pasovna širina seveda ni konstantna v celotnem področju temveč je odvisna od trenutnega položaja. Tako se pasovne širine na podeželju gibljejo do 144 kbps, v primestnem okolju do 384 kbps in v neposredni bližini oddajne postaje (urbano okolje) tudi do 2 Mbps. UMTS ohranja zmožnost globalnega gostovanja v drugi generaciji mobilnega omrežja (GSM/GPRS), hkrati pa vpeljuje številne izboljšave. Tehnologija tretje generacije je namenjena prenosu slik, grafičnih vsebin, video komunikacij ter ostalih multi-medijskih informacij kot sta zvok in podatki (ang. data). Razvoj UMTS-a je temeljil na postopnem približevanju kompletnemu IP omrežju z razširjeno drugo generacijo (2G) mobilnih omrežij, kjer se za prenos podatkov po zračnem vmesniku uporablja prostrani CDMA (ang. Wide-band Code Division Multiple Access). GPRS tako predstavlja konvergenčno točko med 2G tehnologijami in paketno komutirano domeno 3G UMTS omrežja, hkrati pa takšna točka skrbi za brezhibno predajo zveze med UMTS in GSM omrežji. Srce celotne arhitekture predstavlja jedro omrežja, ki ga prikazuje slika 3.

Jedro UMTS-a temelji na topologiji GSM/GPRS in skrbi za komutacijo, usmerjanje in funkcije podatkovne baze uporabniškega prometa. K jedru omrežja pripadajo tokovno komutirani elementi, kot so MSC, VLR in GMSC ter paketno komutirana gradnika SGSN, GGSN. Registri EIR, HLR in AuC so skupni obema načinoma komutacije. Gradniki SGSN, GGSN, Node B in RNC so prikazani na slikah 4 in 7, kjer smo slednje uporabili za izgradnjo UMTS omrežja v simulacijskem okolju Opnet Modeler.



Slika 3: Arhitektura UMTS omrežja
Fig. 3: Architecture of UMTS network

Metoda prenosa podatkov v jedru temelji na asinhronskem načinu ATM (ang. Asynchronus Transfere Mode). ATM prilagoditvena druga plast (AAL2) upravlja in skrbi za tokovno komutirane povezave, medtem ko za paketno komutacijo skrbi paketni povezavni protokol (AAL5), kateri še izvaja pravilno dostavo podatkov. Podrobnejši opis gradnikov sledi v nadaljevanju.

V prehodno/vmesno družino 2.5G tako spadajo HSCSD, GPRS in EDGE. Za HSCSD (ang. High Speed Circuit Switched Data) je značilna povezava do 57.6 kbps. Ker HSCSD temelji na tokovni orientiranosti ima uporabnik ves čas na voljo konstantno pasovno širino, ne glede na to ali jo potrebuje ali ne. Nivo višje glede pasovne širine se nahaja GPRS, ki ponuja povezave do 114 kbps (teoretično tudi do 171 kbps), vendar slednji že bazira na paketno orientirani domeni. Z uvedbo EDGE (ang. Enhanced Data Rates for GSM/Global Evolution) sistema so operaterji lahko zagotavljali kapacitete prenosnih kanalov, ki so že primerljive z UMTS-om. V osnovi gre zgolj za spremenjen način modulacije 8-PSK (ang. Phase Shift Key), katera zagotavlja višje prenosne hitrosti. Z optimalno povezavo 384 kbps predstavlja konkurenco 3G sistemom.

Dandanes UMTS ponuja različne kategorije storitev in aplikacij med katere spadajo internetni dostop (sporočanje, prenos videa in glasbe, zvok in video čez internetni protokol, bančništvo, trgovanje preko spleta), intranet in extranet dostop (spletna pošta, asistenca na poti, mobilna prodaja, tehnične storitve, dostop do baz, video telefonija, konferenčne seje) ter multimedijsko sporočanje kamor sodita SMS in MMS.

Soobstoj obeh, GSM in UMTS omrežij je v prvotni izvedbi zahteval tokovno in paketno komutirano področje (ang. domain) vendar različica UMTS omrežja, kot ga poznamo danes vsebuje zgolj paketno komutirano področje, saj se ves promet prenaša preko IP paketne domene, medtem ko se govorni promet prenaša kot VoIP (ang. Voice over IP). Ključna elementa paketne domene sta SGSN (ang. Serving GPRS Support Node), ki predstavlja podporno vozlišče za strežni GPRS in GGSN (ang. Gateway GPRS Support Node), ki predstavlja podporno vozlišče za GPRS prehod. SGSN shrani naročniški profil uporabnika, ki je registriran v SGSN, hkrati pa upravlja informacije o njegovi lokaciji ter usmerja paketni promet po jedru omrežja do drugega SGSN ali pa do primernega GGSN. Slednji povezuje jedro omrežja z zunanjimi paketnimi omrežji, kot je internet in skrbi za usmerjanje prometa v takšna omrežja oziroma iz takšnih omrežij do mobilnega terminala. Vmesnik, ki se nahaja med enoto SGSN in enoto GGSN imenujemo protokol tuneliranja GTP (ang. GPRS Tunelling Protocol).

Tokovno komutirano domeno sestavljajo komutacijski center za mobilne storitve MSC, komutacijski center za prehodne mobilne storitve GMSC in register gostujočih naročnikov VLR (ang. Visitor Location Register). MSC in GMSC imata enako vlogo kot SGSN ter GGSN v paketni domeni, VLR pa predstavlja register z informacijami o uporabnikih, hkrati pa skrbi za registracijo uporabnika, ko le ta pride v območje nove bazne postaje. Register domačih naročnikov HLR, register za identifikacijo opreme EIR in avtentikacijski center AuC so skupni obema domenama (področjema). Prvi izmed njih vsebuje podatke o naročnikih in sodeluje z VLR registrom, medtem ko ostala dva služita za preverjanje, varnost in identifikacijo strojne opreme (mobilnega terminala). HLR, AuC, VLR, EIR, MSC so elementi omrežnega podsistema.

Centrala mobilnih uslug MSC (ang. Mobile Switching Center) predstavlja telekomunikacijske, prenosne in dodatne usluge. Njena naloga je skrb za komutacijo zvez in iskanje prostih zvez med mobilnimi postajami ter mobilnimi postajami in naročniki javnega omrežja. MSC črpa potrebne informacije tako iz podatkovnih baz NSS, kot tudi iz podatkovnih baz znotraj BSS in OSS. Področje, ki ga pokriva ena centrala mobilnih uslug imenujemo MSC področje.

Register domače lokacije HLR (ang. Home Location Register) predstavlja bazo podatkov, katera vsebuje vse informacije o naročnikih, ki domujejo v danem MSC področju (GSM in ISDN identifikacija, naročene telekomunikacijske, prenosne in dodatne storitve). HLR hrani podatke o trenutnem položaju vseh njenih domačih mobilnih postaj, ne glede na to ali se mobilna postaja trenutno nahaja na njenem področju, ali pa celo gostuje v katerem drugem MSC področju. Te podatke potrebuje centrala v primeru dohodnega klica, da lahko ugotovi v katero celico mora poslati klic oziroma, da klic usmeri v ustrezno MSC področje. V kolikor se v MSC področju pojavi mobilna postaja, katere status ni »domač« za to področje, bo ta MSC poslal zahtevo po podatkih tako imenovanemu »domačemu« MSC-ju. Ta mu pošlje podatke o naročniku iz svoje HLR baze. V trenutku, ko prvi MSC dobi podatke jih vpiše v svoj VLR register.

Register gostujoče lokacije VLR (ang. Visitor Location Register) ima podobno funkcijo kot njegov predhodnik HLR, le

da vsebuje podatke o vseh gostujočih uporabnikih. Ta lastnost omogoča vzpostavljanje tudi odhodnih klicev. VLR v bistvu ni nič kaj drugega, kot uporabnikova dinamična baza podatkov, ki potrebuje intenzivno izmenjavo podatkov z njegovim HLR-jem in se mora ohranjati vse dokler uporabnik ne odide v drugo MSC področje.

Avtentifikacijski center AuC (ang. Autehntication Center) vsebuje shranjene podatke, ki so potrebni za preučevanje prisluškovanja na radijskem prenosu ter podatke za preprečevanje uporabe omrežja neregistriranim uporabnikom. Iz tega vidika vsebuje številne šifrirne ključe za šifriranje in dešifriranje začetnih podatkov radijskega prenosa, postopke za ugotavljanje pristnosti SIM kartice itd. Dostop do baze je skrbno varovan. AuC vključuje tudi EIR (ang. Equipment Identity Register) register.

Vsaka mobilna postaja ima svojo identifikacijsko številko, ki se nahaja pod kratico IMEI (International Mobile Equipment Identity). Register EIR vključuje sezname s številkami mobilnih postaj, katerim je dostop dovoljen in katerim ne. S pomočjo takšnega registra se odkriva telefonske aparate, ki so bili odtujeni (kraja).

Podsistem bazne postaje sestavljata dva ključna dela, in sicer BTS in BSC.

Bazno oddajno/sprejemna postaja BTS (ang. Base Transceiver Station) ima nalogo zagotavljanja potrebnih frekvenčnih kanalov, oziroma nosilcev za vzpostavitev dvosmernih (ang. duplex) radijskih zvez do mobilne postaje, ki se trenutno nahaja v njenem dosegu. BTS postaja lahko pokriva področje ene ali več celic. Znotraj ene BSS je lahko tudi do nekaj sto baznih postaj. BSS opravlja naloge šifriranja koristnih informacij zaradi prisluškovanja, pretvarja radijski prenos v digitalno obliko, katera se uporablja v fiksnem delu, izvaja najrazličnejše meritve signala med prenosom ter ga nato posreduje BSC enoti, tvori časovno izravnavo med oddanimi in sprejetimi signali ipd. BTS vsebuje modulatorje, demodulatorje, kanalske koderje, dekoderje, naprave za digitalni prenos…

Nadzornik bazne postaje BSC (ang. Base Station Controler) opravlja funkcijo nadzorne in upravljaljske postaje v BSS podsistemu, ki je neposredno povezan z MSC, hkrati pa je preko podatkovnega omrežja X.25 povezana z OSS podsistemom. BSC na osnovi teh dveh povezav pridobi dodatne podatke, ki jih potrebuje za svoje delovanje. Ob tem BSC skrbi še za predajo zveze med celicami (ang. handover), dodeljevanje frekvenčnih nosilcev radijskim zvezam, zagotavlja potrebne kvalitete glede na rezultate meritev, ki mu jih posredujejo BTS enote (popravlja oddajne moči BTS-a, sporoča MS-u potrebno oddajno moč, popravlja časovno razliko med oddanimi in sprejetimi signali, preklaplja na boljšo radijsko zvezo znotraj celice itd.)

Operacijski in vzdrževalni podsistem OSS podpira enega ali več OMC-jev, ki izvajajo nadziranje in vzdrževanje delovanja celotnega sistema. OMC (ang. Operation Maintenance Center) vzdržuje in nadzira delovanje vseh elementov omrežja, kot so MSC, BSC, BTS, MS. Iz tega razloga vsebuje podatke o fizični strukturi omrežja (število posameznih elementov), podatke o organizaciji frekvenčnega plana ter opreme, ki poganja sistem.

Omrežje UTRAN je sestavljeno iz več podsistemov, ki so preko Iu vmesnika povezani z nosilnim omrežjem. Vsak podsistem RNS (ang. Radio Network Subsystem) sestavlja ena radijska kontrolna enota (kratica RNC) ter eno ali več vozlišč (baznih postaj BTS). Vsak takšen podsistem upravlja z radijskimi kapacitetami celic, katere pokriva. Za vsako povezavo med mobilnim terminalom in omrežjem je definiran po en podsistem radijskega omrežja (RNS), kot strežni podsistem radijskega omrežja SRNS, ki je odgovoren za radijsko povezavo med mobilnim terminalom in omrežjem dostopa. Kadar nastane potreba, ki presega zmogljivosti SRNS, je lahko ob tem definiran še namenski podsistem DRNS (ang. Drift RNS). Slednji se uporabi pri prehodu med dvema celicama, hkrati pa v takšnem primeru SRNS predstavlja Iu vmesnik do nosilnega omrežja.

UTRAN omrežje vsebuje dva ključna gradnika in sicer bazno postajo in omrežni radio kontroler RNC. Bazna postaja je v osnovi enaka enoti BTS v GSM omrežju, le, da ta za razliko od predhodne zagotavlja podporo za UMTS radijski vmesnik. RNC predstavlja srce novega dostopnega omrežja. Vsi sklepi, odločitve in preverjanja o obratovanju omrežja se določajo v tem segmentu, v samem RNC centru pa se nahaja visoko hitrostno paketno stikalo, s katerim se omogoči zadovoljiva prepustnost prometa. RNC tako vsebuje podporni mehanizem za povezovanje z mobilno postajo, ki ima dovoljenje komuniciranja znotraj njegovega področja. Primarna naloga RNC člena se navezuje na zagotavljanje kvalitetne in učinkovite paketne povezave s paketnimi elementi jedra omrežja, kamor spada SGSN. Prav tako kot BSC tudi RNC skrbi za kontrolo nad radijskim prenosom, kvaliteto prenosa, oddajno močjo, vzpostavlja in prekinja povezave in skrbi za mehko predajo zveze med celicami. RNC lahko premore še dodatno funkcionalnost, kot je upravljanje radijskih resursov RRM (ang. Radio Resurse Managment) ipd.

Hrbtenično omrežje predstavlja most med jedrom omrežja in UTRAN omrežjem, hkrati pa omogoča širokopasovni dostop ter medsebojno povezovanje med uporabniki. Temelj hrbteničnega omrežja predstavljajo paketne tehnologije in sicer IP oziroma ATM. To pomeni, da je hrbtenično omrežje sestavljeno iz skupine IP usmerjevalnikov ali ATM vozlišč medsebojno povezanih s pomočjo povezav točka–točka.

Osnovna postaja uporablja CDMA metodo so-dostopa, katera širi signal v vse smeri, kar omogoča mnogo boljšo izrabo pasovne širine, hkrati pa operaterju omogoči lažji način povečevanja kapacitete prenosnega kanala na določenih predelih. Način so-dostopa iz vseh strani najdemo v mnogih literaturah tudi pod imenom WCDMA (ang. Wide-band CDMA) /1/, /2/, /3/.

3 Aplikacije UMTS omrežja

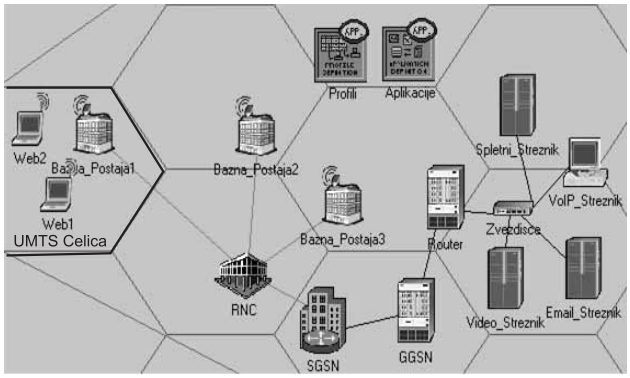
Aplikacije, ki so nam dandanes na voljo lahko simuliramo s simulacijskimi orodji. Opnet Modeler je vodilno razvojno okolje v industriji, ki se uporablja za modeliranje in simuliranje komunikacijskih mrež, hkrati omogoča konstruiranje in študije telekomunikacijskih infrastruktur, posameznih naprav, protokolov ter aplikacij. Orodje stremi k objektno orientiranemu modeliranju, ki vključuje grafične urejevalnike (urejevalnik enot in procesov). Ustvarjeni kontinuirani modeli predstavljajo zrcalo strukture dejanskih omrežij in omrežnih komponent, s čimer se model toliko bolj usklajuje z dejanskim omrežjem ali njegovim segmentom. Prisotna je podpora za vse tipe komunikacijskih mrež z naprednimi tehnologijami. Simulacijski jezik bazira na seriji hierarhičnih urejevalnikov, ki vzporedno ponazorijo strukturo protokolov, opreme, mreže. Zagotovljena je tudi animacija dogajanja v omrežjih, kar še dodatno poenostavi razumevanje delovanja posameznega elementa. Na ta način lahko tvorimo najrazličnejše topologije omrežij kot so »fast ethernet«, »WiFi«, UMTS, GSM, »coax ethernet« itd.

UMTS premore aplikacije spletnega deskanja, prebiranje spletne pošte, FTP prenos, kopiranje datotek, video konference, video telefonijo, VoIP telefonijo, SMS, MMS, dostop do baz, telnet seje in še in še. V takšnem okolju lahko definiramo aplikacije, katerim lahko dodelimo poljubne utežnostne stopnje in sicer na primer manjše obremenjevanje omrežja z določeno aplikacijo (ang. Light) in težje obremenjevanje omrežja z določeno aplikacijo (ang. Heavy). Ker je dandanes priljubljeno mobilno spletno deskanje smo si za preizkus UMTS omrežja izbrali aplikacijo lahkega spletnega deskanja ter vpliv parametra oddaljenosti na kvaliteto storitve, kar bomo predstavili v četrtem poglavju /4/.

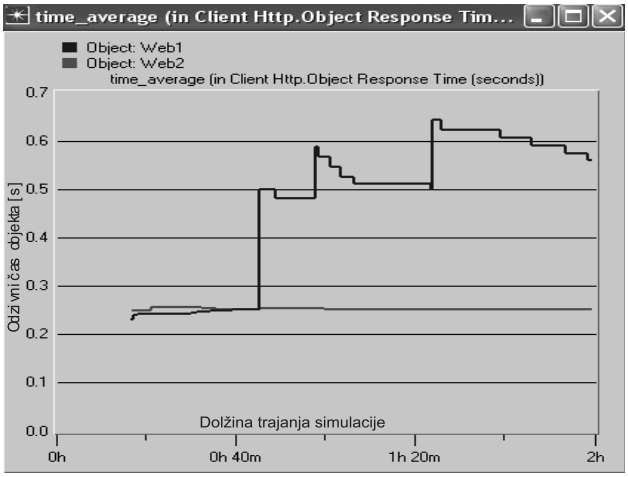
4 Vpliv parametra oddaljenosti na aplikacijo spletnega deskanja in VoIP

Hitrosti prenosa v UMTS omrežju so neposredno v odvisnosti glede na oddaljenost enote od oddajnika. To pomeni, da se lahko enota v neposredni bližini oddajnika povezuje z mnogo višjo hitrostjo (do 2 Mbit/s) v primerjavi s tisto, ki se nahaja na samem robu sprejemne zmogljivosti (zgolj 144 kbit/s in manj). To tezo bomo potrdili z uporabo spletne aplikacije na sledeč način: uporabnika smo postavili v UMTS celico in sicer tako, da je eden izmed njiju v neposredni bližini oddajnika, drugi pa na večji oddaljenosti.

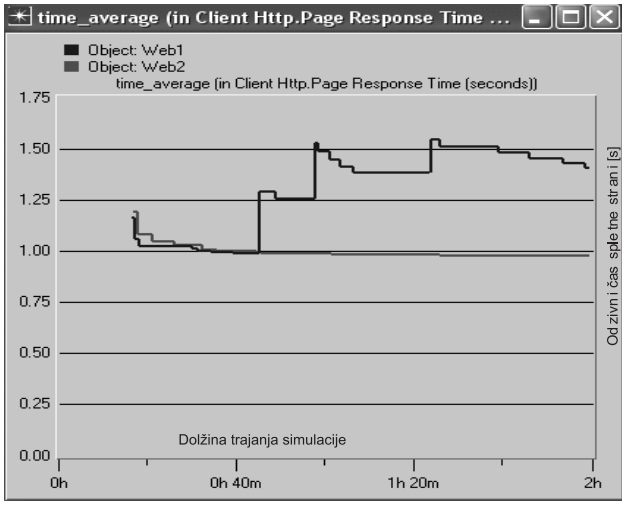
Simulacijsko strukturo sestavljata dva uporabnika spletne aplikacije, ki se nahajata znotraj iste UMTS celice. Do aplikacije, ki se nahaja na spletnem strežniku se povezujeta preko radijskega dostopa in WCDMA vmesnika do bazne postaje, katera je naprej preko ATM-OC3 fizične povezave priključena na radijsko nadzorno postajo RNC, ta pa je naprej preko identične povezave priključena na SGSN enoto. GGSN uporablja za dostop do SGSN in usmerjevalnika PPP-DS3 povezavo. Strežniki so z zvezdiščem medsebojno povezani z 10 Base-T tipom povezave, katera tudi predstavlja most med zvezdiščem in usmerjevalnikom. Na zvezdišče so priključeni strežnik spletne pošte, video, spletne ter VoIP aplikacije, izmed katerih je v uporabi samo spletni strežnik, kamor se UMTS enoti povezujeta do želene storitve. Določitev, katero aplikacijo izmed vseh navedenih naj enoti uporabljata, določata definiciji profila in aplikacije. UMTS enota z imenom »Web2« se nahaja bližje oddajnika v primerjavi z mobilno enoto »Web1«. Iz tega razloga pričakujemo boljše rezultate tiste enote, ki je bližje oddajniku. Naša teza se je izkazala za pravilno, saj je imela enota »Web2« opazno manjši odzivni čas objekta, hkrati pa tudi manjši odzivni čas spletne strani. Sliki 5-6 prikazujeta oba odzivna časa v sekundah, tekom dve urnega simulacijskega obdobja. Maksimalni odzivni čas objekta postaje na večji oddaljenosti znaša 0.63 sekunde, medtem ko je za postajo v neposredni bližini slednji bolj kot ne konstanten. Odzivni čas spletne strani za enoto na večji oddaljenosti ne preseže vrednosti 1.51 sekunde.



Slika 4: Fiksna postavitev UMTS spletnih odjemalcev v neposredno bližino oddajne enote
Fig. 4: UMTS user unit placement in UMTS cell



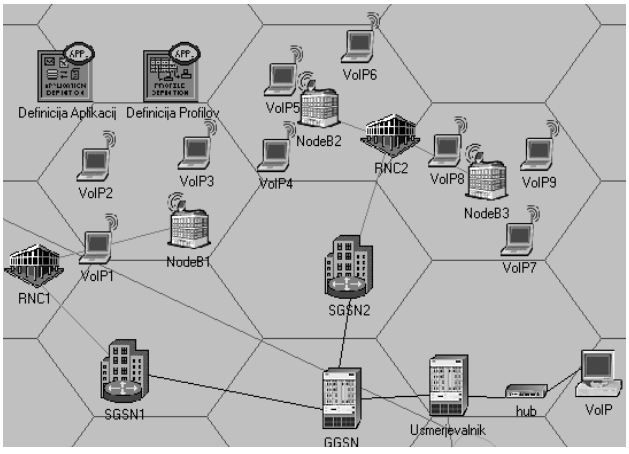
Slika 5: Odzivni čas objekta za uporabnika »Web1« (modra) in uporabnika »Web2« (rdeča), kjer je prvi izmed navedenih oddaljen dlje od oddajnika
Fig. 5: Object response time



Slika 6: Odzivni čas spletne strani za uporabnika »Web1« (modra) in uporabnika »Web2« (rdeča)
Fig. 6: Page response time

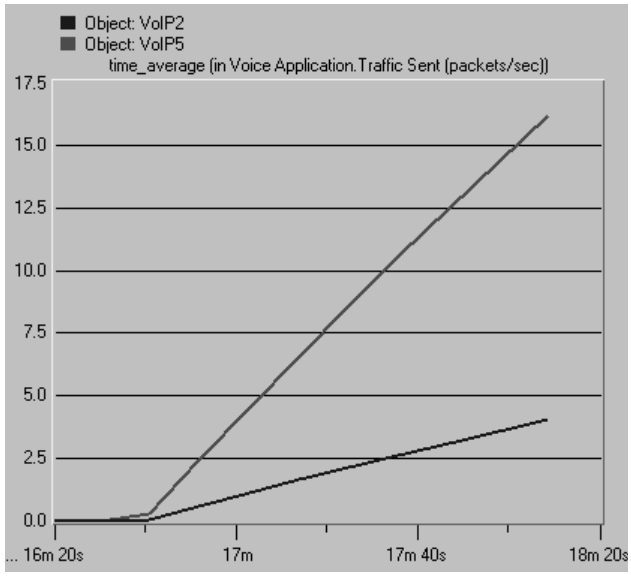
Poudariti je potrebno, da imata oba UMTS odjemalca identične nastavitve, s čimer smo zagotovili enakopravnost obeh enot, s tem pa izluščili vpliv parametra oddaljenosti, ki nas zanima za tovrstno aplikacijo. Vsekakor obe sliki ponazarjata izrazito povečanje odzivnega časa enote na večji oddaljenosti.

Da bo predstava res popolna si še oglejmo, kako vpliva parameter oddaljenosti mobilne enote na VoIP aplikacijo. Tudi v tem primeru so nastavitve vseh postaj enake, opazujemo pa količino ustvarjenega in naknadno prejetega prometa, ki ga ustvarjata enoti »VoIP2« in »VoIP5«, ki se nahajata na različnih oddaljenostih.

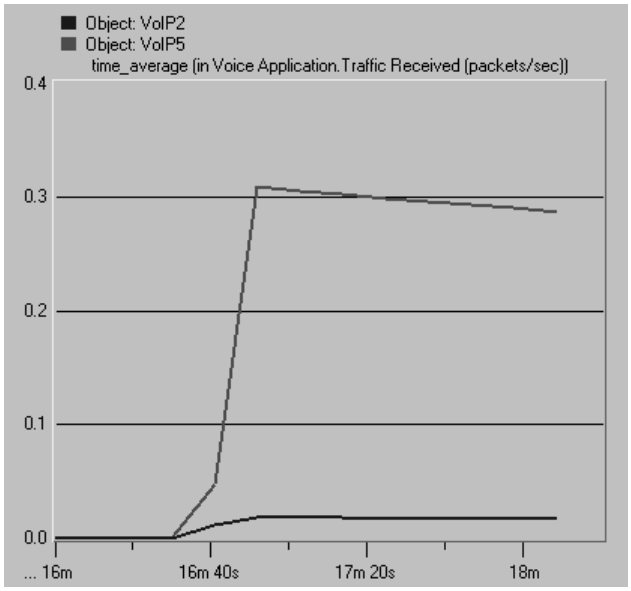


Slika 7: Zgradba UMTS s pomočjo predhodno opisanih gradnikov
Fig. 7: UMTS structure consist with before described base elements

Postaja »VoIP5« lahko zaradi neposredne bližine oddajnika ustvarja mnogo več prometa ter ga pošilja v omrežje po znatno večji pasovni širini, kot enota »VoIP2«, ki je veliko bolj oddaljena. Večja pasovna širina enote »VoIP5« pa ne pomeni samo večjo količino ustvarjenega prometa, temveč tudi večjo količino sprejetega prometa na časovno enoto, kot je prikazano na slikah 8-9.



Slika 8: Količina prometa VoIP – oddajanje
Fig. 8: VoIP traffic capacity – transmitting



Slika 9: Količina prometa VoIP – sprejemanje
Fig. 9: VoIP traffic capacity – receiveing

Rezultati na slikah 8-9 so prikazani v številu poslanih paketov na sekundo v simulacijskem obdobju dveh minut. V najboljšem primeru dosežemo pri oddajanju 16 oddanih paketov na časovno enoto v najslabšem pa v povprečju 4.2 paketa. Ker poznamo velikost paketa (2048 bitov), lahko izračunamo bitno hitrost povezave, ki znaša za optimalni primer 262 kbps, za najslabši primer pa 68 kbps. Povezave, ki jih dobimo s krajšim računskim postopkom vendarle nekoliko odstopajo od teoretičnih navedb, vsekakor pa optimalne 2 Mbps nismo nikakor dosegli.

5 Zaključek

Rezultati simulacij nam nazorno prikazujejo, kako že sama oddaljenost enot od oddajnika vpliva na kvaliteto storitve, zraven tega pa še lahko na kvaliteto vplivajo drugi dejavniki, ki so nam še kako dobro znani iz brezžičnih povezav (oblika reliefa, kvaliteta pokritosti reliefa s signalom, vegetacija, klima, podnebje ipd.) Na osnovi tega lahko podamo zaključni sklep, da se je za izvajanje kvalitetnih storitev potrebno posluževati drugih mehanizmov, kateri enotam na večji oddaljenosti in slabših sprejemnih pogojih lahko zagotavljajo boljšo kvaliteto (prilagajanje oddajne moči za posamezno enoto, ustrezne modulacije, komprimiranje).

6 Literatura

/1/ John Wiley & Sons Ltd, Ham Holma, Antti Toskala, WCDMA for UMTS – Third Edition, The Atrium, Southern Gate, Chichester, West Sussex, Engald, 2004

/2/ John Wiley & Sons Ltd, David Soldani, Man Li, Renaud Cuny, QoS and QoE Management in UMTS Cellular Systems, The Atrium, Southern Gate, Chichester, West Sussex, Engald, 2006

/3/ 3G Tutorial, UMTS overview by UMTS World

/4/ John Wiley & Sons Ltd GSM, GPRS and EDGE Performance, The Atrium, Southern Gate, Chichester, West Sussex, Engald, 2003

Jože Mohorko, Saša Klampfer
Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko,
Smetanova 17, 2000 Maribor
sasa.klampfer@uni-mb.si

***Saša Klampfer** je diplomiral leta 2007 na Fakulteti za elektrotehniko, računalništvo in informatiko, v Mariboru in je podiplomski študent na FERI Maribor. Njegovo raziskovalno področje zajema telekomunikacijske sisteme, modeliranje komunikacijskih sistemov, robotske sisteme in regulacijske sisteme.*

Prispelo (Arrived): 24.04.07 *Sprejeto (Accepted):* 28.5.08

CORE-BASED DESIGN WITH PARASITIC-AWARE APPROACH FOR MEDIUM POWER AMPLIFIER AT 900 MHz, 2.4 GHz, 3.5 GHz AND 5.85 GHz

Arjuna Marzuki[1], Amiza Rasmi[2], Zaliman Sauli[3], Ali Yeon Md Shakaff[3]

[1]School of Electrical and Electronic Engineering,
Universiti Sains Malaysia, Seri Ampangan, Penang, Malaysia.
[2]Telekom Research & Development Sdn Bhd, Malaysia
[3]School of Microelectronics, Universiti Malaysia Perlis, Malaysia

Key words: medium power amplifier,parasitic aware approach, RFIC, MMIC

Abstract: Practical core-based design suitable for medium power amplifier (MPA) is presented. The core circuit is developed and applied at 0.9 GHz, 2.4 GHz, 3.5 GHz and 5.85 GHz. Parasitic-aware design flow is introduced in the whole approach. 5.85 GHz MPA achieves a P1dB of 16.5 dBm, PAE of 15.8% and gain of 4.5 dB at the 12 dBm power input under a low power supply of 2.5V. The maximum current, I_{max} is 77 mA and the power consumption of the device is 192.50 mW. 3.5 GHz MPA achieves a P1dB of 18.2 dBm, PAE of 26.5% and gain of 7.98 dB at the 10.2 dBm power input under a low power supply of 3.0V. The maximum current, I_{max} is 79 mA and the power consumption of the device is 237 mW. 2.4 GHz MPA achieves a P1dB of 17 dBm, PAE of 20.1% and gain of 7.0 dB at the 10 dBm power input under a low power supply of 3.0V. The maximum current, I_{max} is 79 mA and the power consumption of the device is 237 mW. 0.9 GHz MPA achieves a P1dB of 14.2 dBm, PAE of 11% and gain of 4.2 dB at the 10 dBm power input under a low power supply of 3.0 V. The maximum current, I_{max} is 79 mA and the power consumption of the device is 237 mW. Lastly, simulated results almost match the measurement results shows the advantages of applying parasitic information to the core circuit for MPA designs and the effectiveness of core-based design approach in Radio Frequency Integrated Circuit (RFIC) and Monolithic Microwave Integrated Circuit (MMIC).

Načrtovanje ojačevalnikov srednjih moči upoštevajoč parazitne vplive pri frekvencah 900MHz, 2.4GHz, 3.5GHz in 5.85GHz

Kjučne besede: ojačevalniki srednjih moči, upoštevanje parazitnih vplivov, RFIC, MMIC

Izvleček: V prispevku predstavimo praktično izvedbo načrtovanja ojačevalnikov srednjih moči (MPA-Medium Power Amplifier). Osrednje vezje smo razvili in uporabili pri frekvencah 0.9 GHz, 2.4 GHz, 3.5 GHz and 5.85 GHz. Metode načrtovanja so take, da ves čas vodimo računa o parazitnih vplivih. Pri 5.85 GHz MPA smo dosegli P1dB pri 16.5 dBm, PAE 15.8% in ojačanje 4.5 dB pri 12 dBm vhodne moči in pri nizki napajalni napetosti 2.5V. Največji tok, I_{max} je 77mA, poraba moči pa 192.5mW. Pri 3.5 GHz MPA smo dosegli P1dB pri 18.2 dBm, PAE 26.5% in ojačanje 7.98 dB pri 10.2 dBm vhodne moči in pri napajalni napetosti 3.0V. Največji tok, I_{max} je 79mA, poraba moči pa 237mW. Pri 2.4 GHz MPA smo dosegli P1dB pri 17 dBm, PAE 20.1% in ojačanje 7.0 dB pri 10.2 dBm vhodne moči in pri napajalni napetosti 3.0V. Največji tok, I_{max} je 79mA, poraba moči pa 237mW. Pri 0.9 GHz MPA smo dosegli P1dB pri 14.2 dBm, PAE 11% in ojačanje 4.2 dB pri 10 dBm vhodne moči in pri napajalni napetosti 3.0V. Največji tok, I_{max} je 79mA, poraba moči pa 237mW.

Merjeni rezultati se ujemajo s simuliranimi, kar potrjuje pravilen pristop k izvedbi MPA z upoštevanjem parazitnih vplivov pri načrtovanju radiofrekvenčnih (RFIC) in mikrovalovnih (MMIC) integriranih vezij.

1. Introduction

The wireless communication industry has grown rapidly in recent years. The growing Wireless LAN (WLAN) has generated increasing interest in technologies that enable higher data rates and capacity than initially deployed systems. LAN applications have driven the demand for personal wireless communications terminals, and these items need to be low-operating voltage and small size /1/. Power amplifiers among these terminals play a very important role in these systems. So, the application ambit this power amplifier is the key component for researching the advance systems of WLAN and other wireless network systems.

The Gallium Arsenide (GaAs) Pseudomorphic High Electron Mobility Transistor (PHEMT) has good performances on the frequency range, noise figure, output power, and high efficiency with low distortion /2, 3, 4, 5/. Because of its superior performance over the metal oxide semiconductor (MOS) transistors, GaAs transistors have been used extensively to build the Radio Frequency (RF) power amplifiers and play an important role in the wireless communications. PHEMT power amplifiers are making serious inroads into handset cellular (800 MHz to 2.3 GHz) and Wireless LAN (WLAN) (2.4 GHz to 5.85 GHz) applications /6/. GaAs technology has lower R&D cost than CMOS R&D cost is another factor which lures companies to use the technology in power amplifier design /7/.

In system level design, RF platform-based design is normally applied to reduce design cycles /8/. It is often for designing a multi-band, multi-mode IC, reconfigurable reference platform design approach is normally employed /9/. Minimum number of sub-blocks and definition common block is very much useful for first time right IC. Due to uncertainty in layout and parasitic, high frequency integrated circuit design normally needs number of design cycles. Parasitic-aware design flow /10/ is introduced to reduce number of design iterations. The combination of platform-based and parasitic-aware approach could reduce the design cycles and offer flexible block for first time right IC. This approach is very useful for highly integrated multi-standard application integrated RF Front-end silicon-based design /11/. As GaAs technologies become acceptable for RFIC application /7/, the approach can also be applied here.

The design of core circuit and the final design for a 5.85 GHz, 3.5 GHz, 2.4 GHz single-ended medium power amplifier (MPA) for wireless LAN application and 0.9 GHz single-ended medium power amplifier for handset cellular are described in details in this work.

This paper is organized as follows. We first give an introduction to the application of MPA, the technology and the design approach. The following section details out the methodology, design and simulation results. Finally, experimental results and conclusion are discussed in the last two sections.

2. Design

2.1 Design Methodology

Typical MMIC or RFIC design process is to start with topology analysis with respect to specifications. Topology is then simulated at schematic level to verify the performance against the specification. The design is then convert into the layout and post-layout simulation is done to verify the performance against the schematic simulation. If the performance is not similar to the specifications, the design layout has to be modified. The process is repeated until the specifications are met. Multi-band RFIC designs use many approaches; wideband design, parallel design and single design with flexible matching components. This work discusses core-based design approach which can also deliver Multi-band RFIC.

The design flow in Figure 1 is used in this work to give full considerations for the effects that parasitic have on circuit performance. A common block or core circuit which satisfies the specifications at all interested frequencies must be figured out first /8/. The core schematic circuit must be simulated with layout with known parasitic performance. This approach will reduce the design iterations.

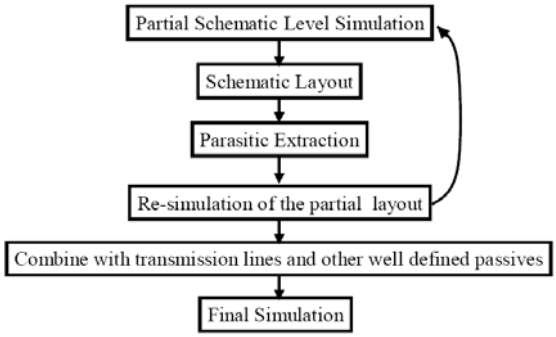


Fig. 1: Parasitic-Aware Design Flow /10/

2.2 Core Circuit Design

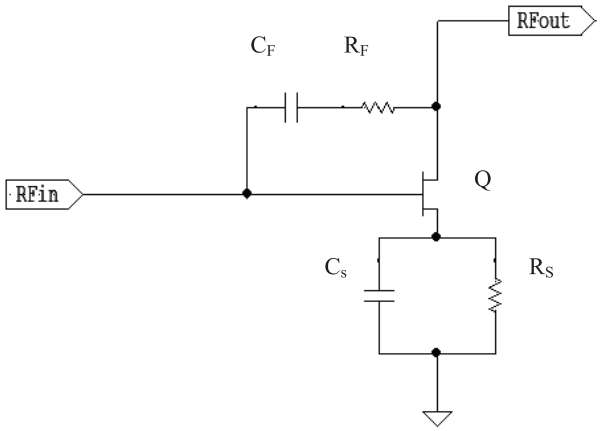


Fig. 2: Core circuit.

From Figure 2, resistor R_F forms the feedback and capacitor C_F is added to allow for independent biasing of the gate and drain of the transistor. C_F can normally be chosen so that it is large enough to be a short circuit over the frequency of interest. In addition, the effect of feedback is to make the input and the output impedances more convenient for matching. Q is depletion-mode transistor which requires negative biasing voltage, R_s is used to set the biasing condition. C_s is used to short the R_s at all interested frequencies. The configuration of the core circuit is very similar to shunt-series amplifier.

From Figure 2, the closed-loop gain, A_V

$$A_V = -\frac{1}{\left(\frac{G_F(1+\frac{1}{A_{OL}})}{G_S}+\frac{1}{A_{OL}}\right)} \quad (1)$$

where G_F, G_S, A_{OL} is conductance of feedback resistance, conductance of source resistance and open-loop gain respectively. The closed-loop gain will be equal to an open-loop gain if G_F approaches zero. In this case, the open loop gain is referring to transistor gain without feedback topology. An amplifier with the resistor feedback can achieve self matching /13/.

Details of high frequency small-signal analysis can be found in Thomas Lee's book /12/. This topology offers wide bandwidth /13/, which is suitable candidate for the MPA at different frequency.

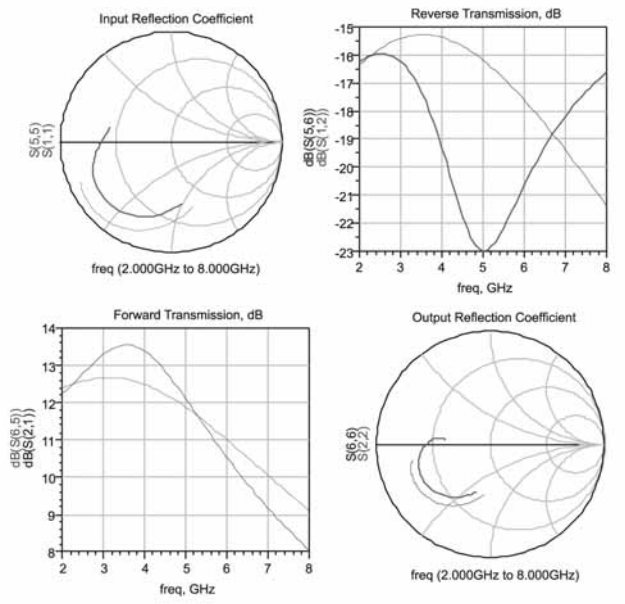


Fig. 3: Core circuit simulation results

From Figure 3, the performance of core circuit is compared. S(1,1), S(1,2), S(2,1) and S(2,2) are partial schematic level simulation results. The layout information is added to circuit by adding transmission line model between transistor source and capacitor. S(5,5) S(5,6), S(6,5) and S(6,6) are schematic simulation results. From Figure 3, it can be concluded that parasitic information does affect the input and output reflection coefficient of the core circuit. The input impedance and output impedance of the core circuit is important information for MPA design. All circuits use active and passive models from the foundry with transistor; number of finger (NOF) = 10, unit gate width (UGW) = 100 µm, C_F = 8 pF, C_s = 2 pF, R_F = 500 Ω and R_s = 10 Ω. The core circuit with parasitic information is modeled as modified-transistor and later used in MPA design.

2.3 Medium Power Amplifier Design

The complete schematic designed MPA are shown in Figures 4 and 5 where L_G, L_S, L_D and L_O are all implemented on-chip. The inductors L_G and L_S are chosen to provide the desired input impedance. The inductor L_D is a current source for the MPA and used for output power matching. The capacitor C_{in} at the input is used for input matching. The capacitor C_C at the output plays a role for both DC block and output matching. The capacitor C_O is used for network matching. The outputs are matched for high compression point, P1dB.

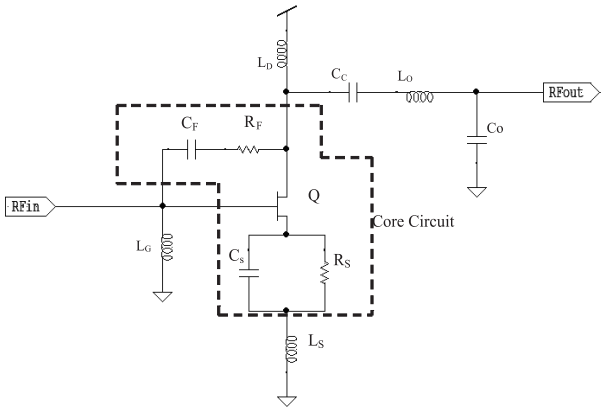


Fig. 4: Circuit schematic of the PHEMT single-ended medium power amplifier for RF frequency of 2.4 GHz, 3.5 GHz, and 5.85 GHz.

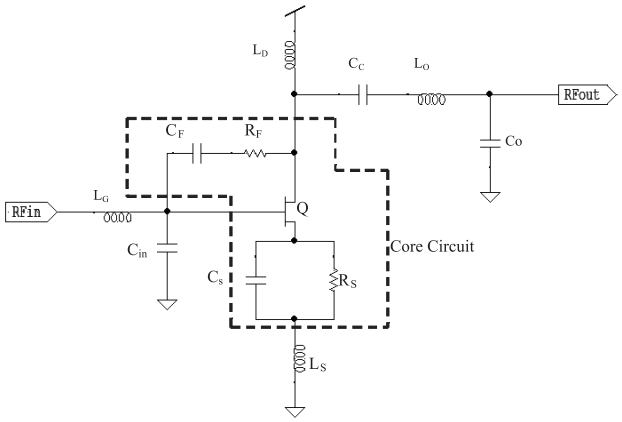


Fig. 5: Circuit schematic of the PHEMT single-ended medium power amplifier for RF frequency of 0.9 GHz.

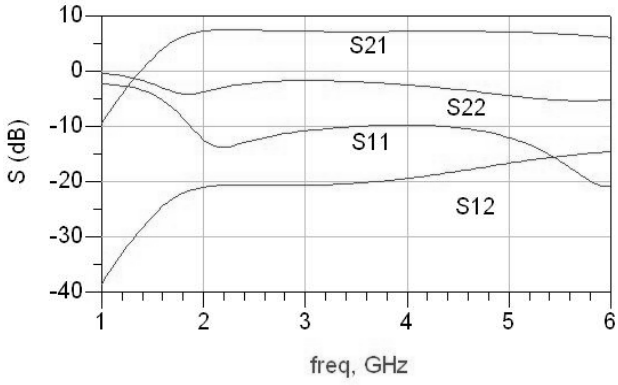
2.3.1 Result & Discussion

The single-ended medium power amplifiers are shown in Figure 4 and 5 are simulated in 0.15µm GaAs PHEMT process technology using ADS simulator /14/. The supply voltage, V_{DD} for this simulation is 2.5 V to 3.0 V.

i) MPA at 5.85 GHz

The small-signal performance of the single-ended MPA is shown in Figure 6 over 1 to 6 GHz. The linear gain (S(21)) obtained is 6.3 dB, S(12) is -14.8 dB, input return loss is 20.6 dB and output return loss is 5.4 dB at a frequency of 5.85 GHz and V_{DD} s 2.5 V.

Figure 7 shows a stability factor, K as a function of frequency for this single-ended MPA. At 5.85 GHz, a stability factor, K for this device is 1.172. The MPA is in unconditionally stable condition due to the stability factor for the MPA is higher than 1 at the whole range of frequency.



Fgi. 6: Gain, input return loss and output return loss as a function of frequency for medium power amplifier at 5.85 GHz.

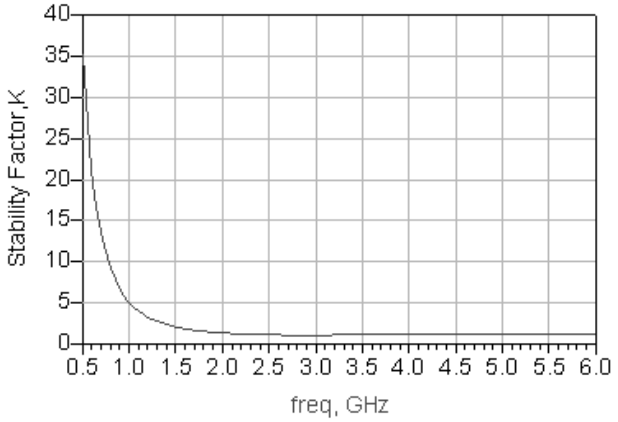


Fig. 7: Stability factor, K of medium power amplifier at 5.85 GHz.

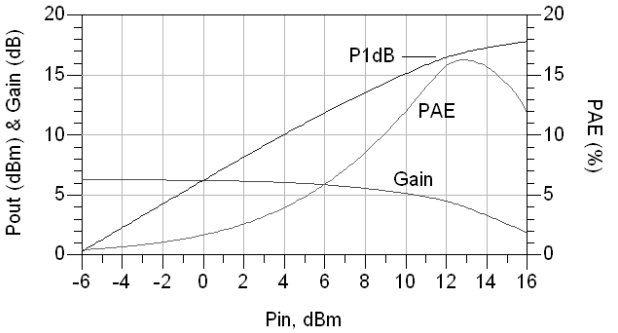


Fig. 8: Output power, power added efficiency and power gain versus input power for medium power amplifier at 5.85 GHz.

Figure 8 shows the output power, power gain and the power added efficiency, PAE as a function of input power, respectively. The MPA has an output power of 16.5 dBm at 1dB gain compression (P1dB), a power gain of 4.5 dB and the power added efficiency (PAE) of 15.8% for an input power, P_{in} of 12 dBm.

Figure 9 shows the maximum available gain, MAG, associated power gain and gain as a function of frequency for the simulated PHEMT medium power amplifier. At 5.85 GHz, the MAG is 8.04 dB and the associated power gain is 7.993

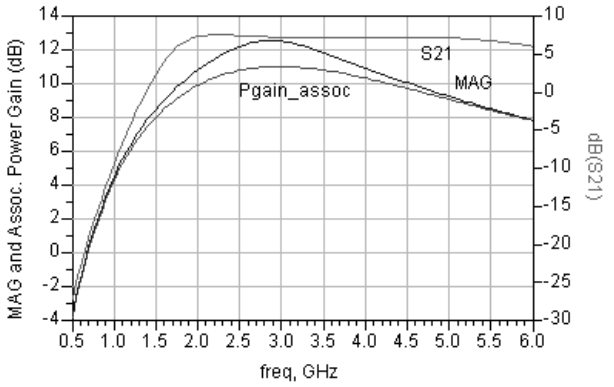


Fig. 9: Maximum available gain, MAG, associated power gain and gain of medium power amplifier at 5.85 GHz.

dB. The MAG is the maximum available gain at all frequencies with the output condition matched to 50 Ohm.

ii) MPA at 3.5 GHz

The small-signal performance of the single-ended MPA is shown in Figure 10 over 1 to 6 GHz. The linear gain (S(21)) obtained is 11.4 dB, S(12) is -18.8 dB, input return loss is 18.1 dB and output return loss is 10.4 dB at a frequency of 3.5 GHz and V_{DD} is 3.0 V.

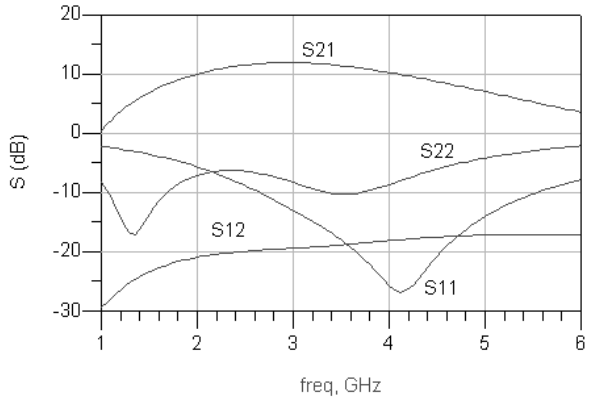


Fig. 10: Gain, input return loss and output return loss as a function of frequency for medium power amplifier at 3.5 GHz.

Figure 11 shows a stability factor, K as a function of frequency for this single-ended MPA. At 3.5 GHz, a stability factor, K for this device is 1.305. The MPA is in unconditionally stable condition due to the stability factor for the MPA is higher than 1 at the whole range of frequency.

Figure 12 shows the output power, power gain and the power added efficiency, PAE as a function of input power, respectively. The MPA has an output power of 18.2 dBm at 1dB gain compression (P1dB), a power gain of 7.98 dB and the power added efficiency (PAE) of 26.5% for an input power, P_{in} of 10.2 dBm.

Figure 13 shows the maximum available gain, MAG, associated power gain and gain as a function of frequency for

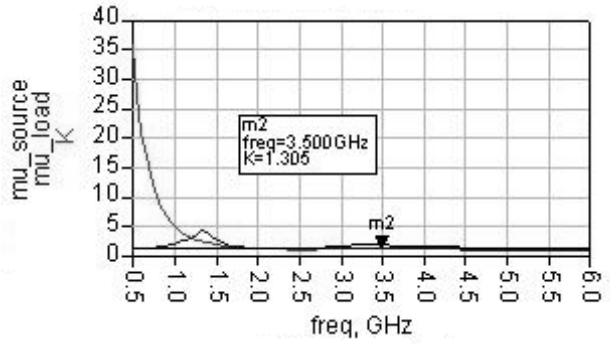


Fig. 11: Stability factor, K of medium power amplifier at 3.5 GHz.

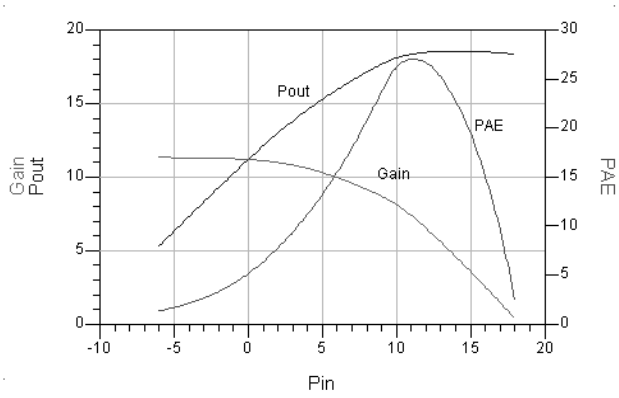


Fig. 12: Output power, power added efficiency and power gain versus input power for medium power amplifier at 3.5 GHz.

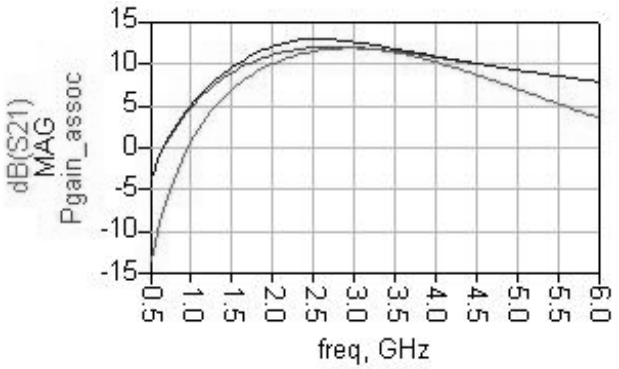


Fig. 13: Maximum available gain, MAG, associated power gain and gain of medium power amplifier at 3.5 GHz.

the simulated PHEMT medium power amplifier. At 3.5 GHz, the MAG is 11.8 dB and the associated power gain is 11.5 dB. The MAG is the maximum available gain at all frequencies with output condition matched to 50 Ohm.

iii) MPA at 2.4 GHz

The small-signal performance of the single-ended MPA is shown in Figure 14 over 1 to 6 GHz. The linear gain (S(21)) obtained is 10.9 dB, S(12) is -20.5 dB, input return loss is 6.1 dB and output return loss is 5.6 dB at a frequency of 2.4 GHz and V_{DD} is 3.0 V.

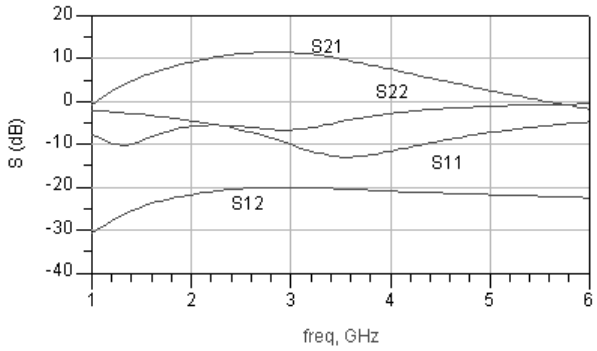


Fig. 14: Gain, input return loss and output return loss as a function of frequency for medium power amplifier at 2.4 GHz.

Figure 15 shows a stability factor, K as a function of frequency for this single-ended MPA. At 2.4 GHz, a stability factor, K for this device is 1.233. The MPA is in unconditionally stable condition due to the stability factor for the MPA is higher than 1 at the whole range of frequency.

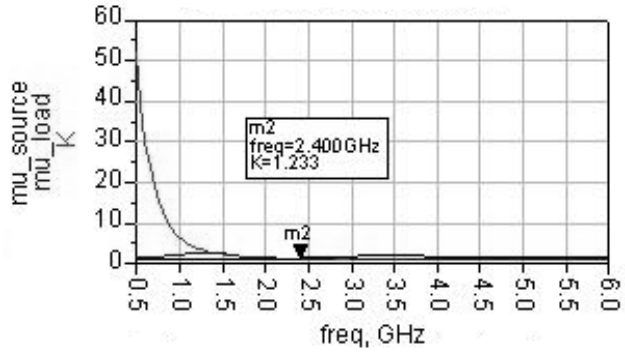


Fig. 15: Stability factor, K of medium power amplifier at 2.4 GHz.

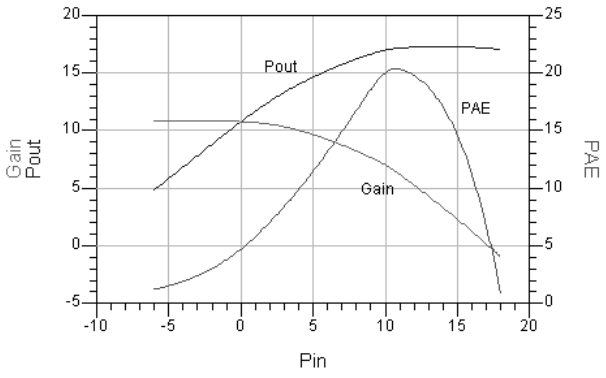


Fig. 16: Output power, power added efficiency and power gain versus input power for medium power amplifier at 2.4 GHz.

Figure 16 shows the output power, power gain and the power added efficiency, PAE as a function of input power, respectively. The MPA has an output power of 17.0 dBm at 1dB gain compression (P1dB), a power gain of 7.0 dB and the power added efficiency (PAE) of 20.1% for an input power, P_{in} of 10 dBm.

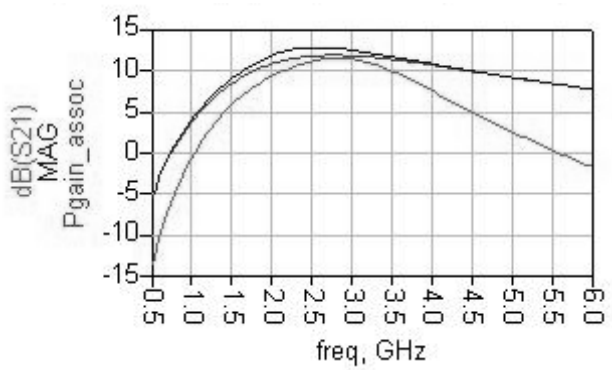


Fig.e 17: Maximum available gain, MAG, associated power gain and gain of medium power amplifier at 2.4 GHz.

Figure 17 shows the maximum available gain, MAG, associated power gain and gain as a function of frequency for the simulated PHEMT medium power amplifier. At 2.4 GHz, the MAG is 12.79 dB and the associated power gain is 11.68 dB. The MAG is the maximum available gain at all frequencies with the output condition matched to 50 Ohm.

iv) MPA at 0.9 GHz

The small-signal performance of the single-ended MPA is shown in Figure 18 over 1 to 6 GHz. The linear gain (S(21)) obtained is 10.1 dB, S(12) is -19.8 dB, input return loss is 13.4 dB and output return loss is 12.1 dB at a frequency of 0.9 GHz and V_{DD} is 3.0 V.

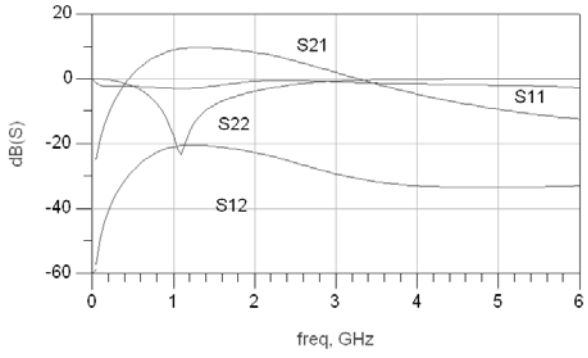


Fig. 18: Gain, input return loss and output return loss as a function of frequency for medium power amplifier at 0.9 GHz.

Figure 19 shows a stability factor, K as a function of frequency for this single-ended MPA. At 0.9 GHz, a stability factor, K for this device is 1.536. The MPA is in unconditionally stable condition due to the stability factor for the MPA is higher than 1 at the whole range of frequency.

Figure 20 shows the output power, power gain and the power added efficiency, PAE as a function of input power, respectively. The MPA has an output power of 14.2 dBm at 1dB gain compression (P1dB), a power gain of 4.2 dB and the power added efficiency (PAE) of 11% for an input power, P_{in} of 10 dBm.

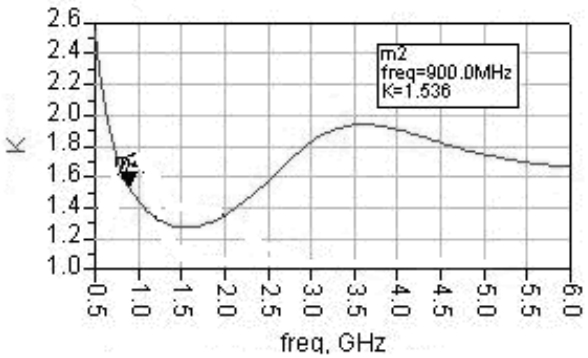


Fig. 19: Stability factor, K of medium power amplifier at 0.9 GHz.

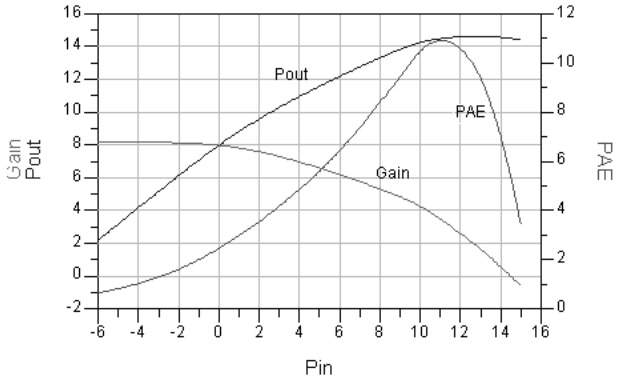


Fig. 20: Output power, power added efficiency and power gain versus input power for medium power amplifier at 0.9 GHz.

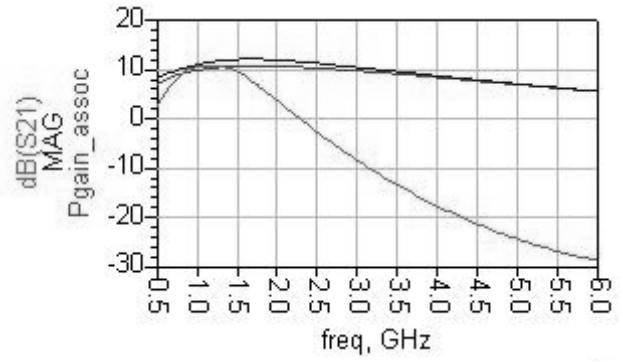


Fig. 21: Maximum available gain, MAG, associated power gain and gain of medium power amplifier at 0.9 GHz.

Figure 21 shows the maximum available gain, MAG, associated power gain and gain as a function of frequency for the simulated PHEMT medium power amplifier. At 0.9 GHz, the MAG is 10.6 dB and the associated power gain is 9.5 dB. The MAG is the maximum available gain at all frequencies with the output condition matched to 50 Ohm.

Table 1 shows the summary of MPA performance for this work at different frequency. It is also shown that poor return losses at some frequencies could be attributed due to insufficient matching at the input and output of the MPA.

Table 1: Summary of single-ended MPA performance.

Parameter	This work			
Frequency (GHz)	5.85	3.5	2.4	0.9
S21 (dB)	4.5	11.4	10.9	10.1
PAE (%)	15.8	26.5	20.1	11
P1dB (dBm)	16.5	18.2	17.0	14.2
S12 (dB)	-14.8	-18.8	-20.5	-19.8
S11 (dB)	-20.6	18.1	6.1	13.4
S22 (dB)	-5.4	10.4	5.6	12.1
Voltage Supply (V)	2.5	3.0	3.0	3.0
L_g (nH)	2.16	2.16	2.16	10.95
C_{in} (pF)	n/a	n/a	n/a	0.8
L_s (nH)	0.372	0.372	0.372	0.372
L_D (nH)	5.288	5.288	4.128	5.288
C_C (pF)	1	4.2	7.7	4.5
L_O(nH)	1.35	1.65	2.436	5.288
C_O (pF)	0.76	1	1.4	2.1
Application	802.11a Wireless LAN	802.16 WiMAX	802.11a/b/g Wireless LAN	Handset cellular

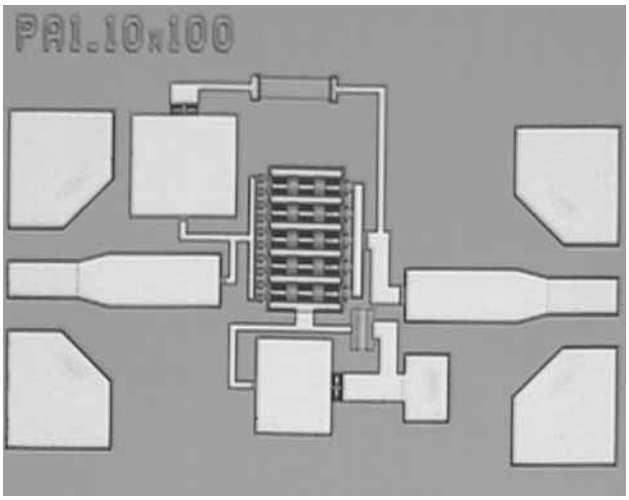


Fig. 22: Core circuit microphotograph

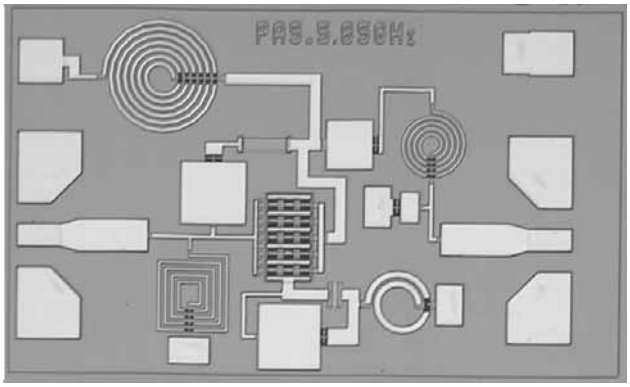


Fig. 23: 5.85 GHz MPA microphotograph

Therefore more design optimization /15/ is required to improve the performance of the MPA in this work.

Some of the passive devices such as capacitors and inductors have been characterized and presented in conference /16, 17/. MPA of 5.85 GHz results have also been presented in a conference /18/.

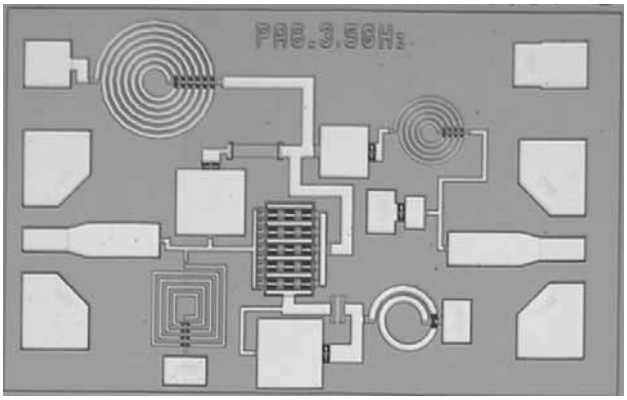


Fig. 24: 3.5 GHz MPA microphotograph

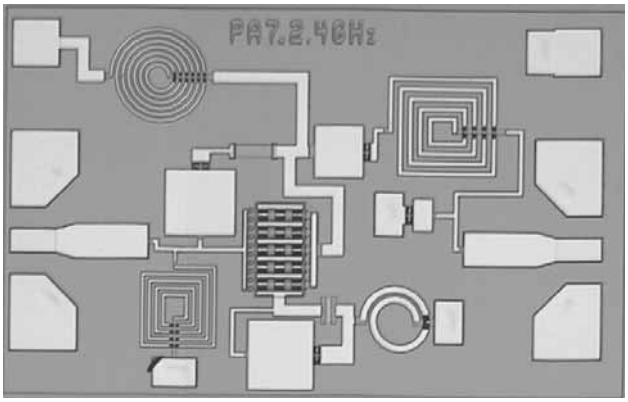


Fig. 25: 2.4 GHz MPA microphotograph

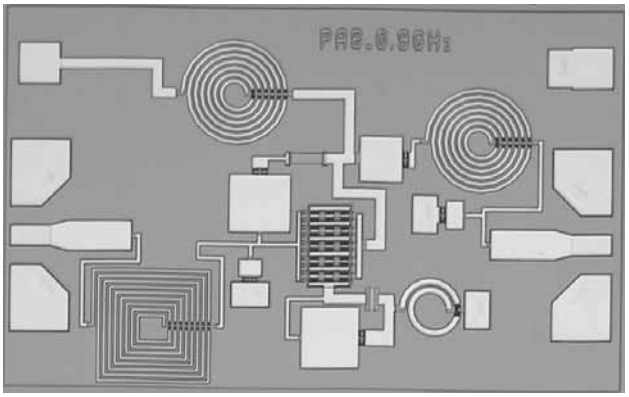


Fig. 26: 0.9 GHz MPA microphotograph

3. Experimental Results

Figure 22 shows photo of fabricated core circuit while Figure 23, 24, 25, 26 shows 5.85 GHz, 3.5 GHz, 2.4 GHz and 0.9 GHz MPA microphotographs respectively. A standard on wafer measurement methodology is used in characterizing the core circuit. In the S-Parameter results (Figure 27), the orders of the charts are S11 and S22 (Unit Smith Chart), S21 Magnitude (dB), and S12 Magnitude (dB). The S-parameter measurement is performed in the

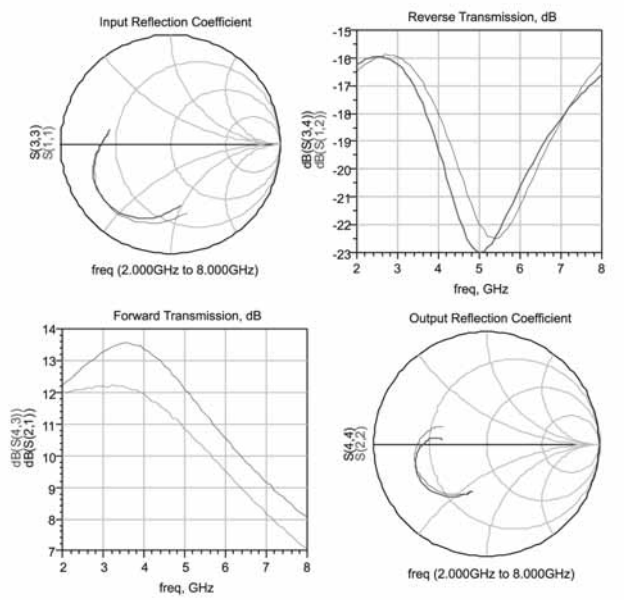


Fig. 27: Simulation and measurement data for core circuit.

frequency range from 2 GHz to 8 GHz and the bias conditions is 2.5 V of drain voltage, V_{DS} and a gate voltage, V_{GS} is 0 V. The drain current, I_D is 83 mA.

It can be seen from Figure 27, the partial schematic layout simulation (S(3,3), S(4,3), S(3,4) and S(4,4)) is almost similar to the measurement results. Unfortunately due to test equipments issue, no MPA measurement has been made.

4. Conclusion

A core-based and parasitic-aware design flow is presented and MMIC designs using the same core circuit for all interested frequencies are also presented. Measurement results of core circuit very much similar to the parasitic-aware schematic simulation result, this has shown the potential of the approach in designing multi-standard and multi-application RFIC and MMIC.

A single-ended medium power amplifier using 0.15 µm GaAs power PHEMT process technology with a gate width of 100 µm and 10 fingers is presented. A 5.85 GHz MPA achieved a P1dB of 16.5 dBm, PAE of 15.8% and gain of 4.5 dB at the 12 dBm power input under a low power supply of 2.5V. The maximum current, I_{max} is 77 mA and the power consumption of the device is 192.50 mW. Other results; a linear gain (S21) of 6.3 dB, input return loss of 20.6 dB, output return loss of 5.4 dB and a stability factor, K of 1.172 at RF frequency of 5.85 GHz. This MPA is suitable for IEEE 802.11a wireless LAN applications.

3.5 GHz MPA achieved a P1dB of 18.2 dBm, PAE of 26.5% and gain of 7.98 dB at the 10.2 dBm power input under a low power supply of 3.0V. The maximum current, I_{max} is 79 mA and the power consumption of the device is 237 mW. Other results; a linear gain (S21) of 11.4 dB, S(12) of -18.8 dB, input return loss of 18.1 dB, output return loss of 10.4 dB and a stability factor, K of 1.305 at RF frequency of 3.5 GHz. This MPA is suitable for IEEE 802.16 WiMAX applications.

2.4 GHz MPA achieved a P1dB of 17 dBm, PAE of 20.1% and gain of 7.0 dB at the 10 dBm power input under a low power supply of 3.0V. The maximum current, I_{max} is 79 mA and the power consumption of the device is 237 mW. Other results; a linear gain (S21) of 10.9 dB, S(12) of -20.5 dB, input return loss of 6.1 dB, output return loss of 5.6 dB and a stability factor, K of 1.233 at RF frequency of 2.4 GHz. This MPA is suitable for IEEE 802.11a/b/g wireless LAN applications.

0.9 GHz MPA achieved a P1dB of 14.2 dBm, PAE of 11% and gain of 4.2 dB at the 10 dBm power input under a low power supply of 3.0 V. The maximum current, I_{max} is 79 mA and the power consumption of the device is 237 mW. Other results; a linear gain (S21) of 10.1 dB, S(12) of -19.8 dB, input return loss of 13.4 dB, output return loss of 12.1 dB and a stability factor, K of 1.536 at RF frequency of 0.9 GHz. This MPA is suitable for handset cellular applications.

Even though there is no measurement on MPA designs, authors believe based on the core circuit measurement results the targeted simulation results of MPA can be achieved.

Acknowledgement

The authors gratefully acknowledge the support of TM Research & Development Sdn. Bhd. for this work under Project number R05-0607-0.

References

/1./ A. Raghavan. H. Deukhyoun, M. Moonkyun, A. Sutono, L. Kyutae, and J. Laskar, "A 2.2-V Operation, 2.4-GHz Single-Chip GaAs MMIC Transceiver for Wireless Applications," *2002 IEEE MTT-S International Microwave Symposium Digest*, Vol. 2. pp. 1019 –1022.

/2./ C. E. Weitzel, "RF Power Amplifiers for Cellphones," 2003, GaAs MANTECH Inc.

/3./ Chien -Chang Huang, Sung-Mao Lee, and Kuan-Yu Chen. 2005, "GaAs PHEMT Characterization for OFDM Power Amplifier Application," *10th International Symposium on Microwave and Optical Technolog,* pp. 767-770.

/4./ A. Platzker, S and Bouthillete, "Variable Output, High-Efficiency Low-Distortion S-band Power Amplifiers*," 1995 IEEE MTT-S Int. Microwave Symp. Dig*, pp. 441-444.

/5./ J.Komiak, S. Wang, and T. Roger. 1997, "High Efficiency 11watt Octave S/C-band PHEMT MMIC Power Amplifier," *1997 IEEE MTT-S Int. Microwave Symp. Dig*, pp. 1421-1424.

/6./ Kohei Fujii, Henrik Morkner, and Edward Brown, "A Novel Low Cost Enhancement Mode Power Amplifier MMIC in SMT Package for 7 to 18 GHz Applications*," 2004.12th GaAs Symposium*, pp. 599-602.

/7./ Chin-Chun Meng and Tzung-Han Wu, " A 5 GHz RFIC Single Chip Solution in GaInP/GaAs HBT Technology," *Microwave Journal*, Vol. 51, No.2, February 2008, pg. 132.

/8./ Peter Baltus, "Platform-Based RF-System Design," *Analog Circuit Design*, Springer 2006, pp. 195–213.

/9./ Daniel L. Kaczman, Manish Shah, Nihal Godambe, Mohammed Alam, Homero Guimaraes, Lu M. Han, Mohammed Rachedine, David L. Cashen, William E. Getka, Charles Dozier, Wayne P. Shepherd, Karl Couglar , "A single-chip tri-band (2100, 1900, 850/800 MHz) WCDMA/HSDPA cellular transceiver ," *IEEE Journal of Solid-State Circuits*, Vol. 41, May 2006 pp. 1122 - 1132.

/10./ Yanxin Wang, "Millimeter Wave Transceiver Frontend Circuits In Advanced SiGe Technology With Considerations for On-Chip Passive Component Design And Simulation," PHD Thesis, Cornell University, 2006.

/11./ Paolo Rossi, Antonio Liscidini, Massimo Brandolini, Francesco Svelto , "A variable gain RF front-end, based on voltage-voltage feedback LNA, for multistandard application," *IEEE Journal of Solid-State Circuits*, Vol. 40, Mar 2005, pp. 690 -697.

/12./ Thomas H. Lee, "The Design of CMOS Radio-Frequency Integrated Circuits," Cambridge Univ Press, 2006 pp. 284-288.

/13./ A.Marzuki, T Zainal, A Zulkifli, N Mohd-Noh, and Z. A. Abdul-Aziz, " A Broadband RF Feedback Amplifier Design with Simple Feedback Network," *2004 RF and Microwave Conf*., pp.1-4.

/14./ Agilent Technologies., Agilent ADS, 2005A.

/15./ Arjuna Marzuki, Zaliman Sauli and Ali Yeon Md Shakaff, "A Practical High Frequency Integrated Circuit Power-constraint Design Methodology Using Simulation-based Optimization," *UKMEC2008*, London, 2008.

/16./ Rasidah Sanusi, Ahmad Ismat Abdul Rahim and A. Marzuki, "Effect of MIM Capacitor on the Performance of Low Noise Amplifier for Wireless LAN Applications," 2007 *ROVISP*, Malaysia.

/17./ Norhapizin.K, Azmi Ismail, Ahmad Ismat A.R. and A.Marzuki, "Characterization of Spiral Inductor based on 0.15ĕm GaAs pHEMT Technology for RF Application," *2007 ROVISP*, Malaysia.

/18./ Amiza Rasmi, Mohd Azmi Ismail, Ahmad Ismat Abd Rahim and A. Marzuki,"0.15ĕm Pseudomorphic HEMT Medium Power Amplifier for Wireless LAN Application," *2007 ROVISP*, Malaysia.

Arjuna Marzuki,
School of Electrical and Electronic Engineering,
Universiti Sains Malaysia, Seri Ampangan,
14300 Nibong Tebal, Penang, Malaysia.
Tel.: +604 599 6021; Fax: +604 594 1023 E-mail: eemarzuki@eng.usm.my

Amiza Rasmi,
Telekom Research & Development Sdn Bhd, Malaysia

Zaliman Sauli,
Ali Yeon Md Shakaff
School of Microelectronics,
Universiti Malaysia Perlis, Malaysia

Prispelo (Arrived): 03.05.07 *Sprejeto (Accepted):* 28.5.08

A CMOS MEMBERSHIP FUNCTION CIRCUIT EMPLOYING SINGLE CURRENT DIFFERENCING BUFFERED AMPLIFIER

Mahmut Tokmakçi, Mustafa Alçi

Erciyes University, Engineering Faculty, Dept. of Electrical&Electronics Engineering, Kayseri, Turkey

Key words: MFC, CDBA, PSPICE, CMOS

Abstract: The author proposes a new fully integrable Membership Function Circuit (MFC) using a Current Differencing Buffered Amplifier (CDBA) which employs two second generation current conveyor (CCII) and a voltage buffer. This MFC achieves basic membership functions such as trapezoidal, triangle, S-shape, and Z-shape. The characteristics (width, height, and position) of the implemented MFC are easily adjusted by varying left and right voltages and bias currents. Since the proposed MFC is implemented with a single CDBA block with simple structure, it can be capable of high-speed operation and integrated as a circuit to cover small area of a chip. The behaviour of the proposed MFC has been verified by PSPICE using the model parameters with 0.5 mm MIETEC CMOS process. The proposed MFC is voltage-input current-output and CMOS based structure with low supply voltages (± 1.5V). Therefore, it is suitable for both current-mode and low-voltage fuzzy and neural hardware.

Izvedba CMOS MFC vezja z uporabo enojnega tokovnega diferenčnega ojačevalnika

Kjučne besede: MFC, CDBA, PSPICE, CMOS

Izvleček: V prispevku predlagamo izvedbo popolnoma itegriranega MFC vezja (Membership Function Circuit) z uporabo CDBA, ki uporablja dva CCII druge generacije in napetostni vmesnik. S pomočjo tako izvedene MFC pridemo do osnovnih funkcij, kot so trapezoidalna, trikotna, S-oblike in Z-oblike. Karakteristike MFC zlahka prilagajamo s spreminjanjem napetosti in napajalnih tokov. Ker je MFC izveden z enim CDBA blokom z enostavno strukturo, je hiter in se da integrirati na majhno površino čipa. Vedenje MFC smo preverili s pomočjo programa PSPICE z uporabo modelnih parametrov procesa 0.5 µm MIETEC CMOS. MFC je izveden z napetostnim vhodom in tokovnim izhodom s tehnologijo CMOS pri nizki napajalni napetosti +-1.5V.

1. Introduction

The Membership Function Circuit (MFC) or fuzzifier is one of the most important units in the fuzzy logic controllers (FLC). The various MFC hardware have implemented in literature /1-10/. A high-speed digital MFC based on BiCMOS technology has been proposed in /1/ but the fabrication cost is high. The other MFC designs with current-mode analogue circuits were proposed in /2-3/. However the speed of these circuits is low. Then, the sub-threshold membership function circuit was proposed in /4/. Although this MFC has low power consumption, the output current linearity and accuracy of the circuit is low. Most of the membership function circuits in literature /5-6/ have been designed to provide two membership functions as triangle and trapezoidal shapes in general. In addition to these functions, for generating Z-shape and S-shape membership functions are required extra circuits in original membership function circuit /7-8/. A voltage-input/current-output programmable Gaussian function network with capacitors for the programmability is introduced in /9/. But, the capacitors in network can be refreshed to maintain an accurate programmed value. Also, the reference current needs to be adjusted to control the amplitudes of the output current in their design. The other MFC with good programmable features is presented in /10/. In this MFC, all using transistors are operated in weak inversion region and narrow input current range. The dynamic range of this circuit is small and speed is low for general applications because of the inherent limitations of transistors in weak inversion /11/.

In this study, a new MFC using CDBA is presented. The proposed MFC has capability of generating four standard membership functions without extra devices. Also, it can be operated high speed and implemented simple structures with easy design automation.

The outline of this paper is as follows. Section II briefly defines a basic Current Differencing Buffered Amplifier (CDBA) and proposed Membership Function Circuit (MFC) is theoretically described in detailed. Section III evaluates a current-mode MFC with PSPICE simulation experiments. In Section IV, the overall conclusions are given.

2. Circuit description

The fuzzification block maps the measured fuzzy input variable(s) of a fuzzy system into a suitable range that corresponds to the universe of discourse, and then converts the crisp input value into a fuzzy set. In many fuzzy and neuro-fuzzy applications, a Gaussian or triangular function is generally used in the fuzzification process of the Fuzzy Logic Controllers (FLCs).

In this study, we used a CDBA into our proposed MFC. The modified circuit structure of the CDBA in /12/ and

circuit symbol is shown in Fig. 1. The characteristic equation of this element can be given as

$$V_p = V_n = 0, \quad I_z = I_p - I_n, \quad V_w = V_z \qquad (1)$$

Here, current through z-terminal follows the difference of the current through p-terminal and n-terminal. Input terminals, p and n, are internally grounded. A possible CMOS realization of CDBA consisting of a differential current controlled current source (DCCCS) followed by a voltage buffer is shown in Fig. 1. The CDBA offers the well-known advantages of the CFA and CCII, such as high slew-rate, wide bandwidth and simple implementation. According to the above equations, this element converts the difference of the input currents I_p and I_n, into the output voltage V_w, through the impedance which will be connected to terminal 'z'. Therefore, the CDBA can be considered as a trans-impedance amplifier and, from this viewpoint, it is similar to the current feedback amplifier (CFA) /13/. Furthermore, since the CDBA can be considered as a collection of current-mode and voltage-mode unity gain cell, this element is free from many parasitic effects and is expected to be suitable for high-frequency operation /14-16/. It can be operated in both current-mode and voltage-mode in a wide frequency range and can also be implemented with CMOS technology.

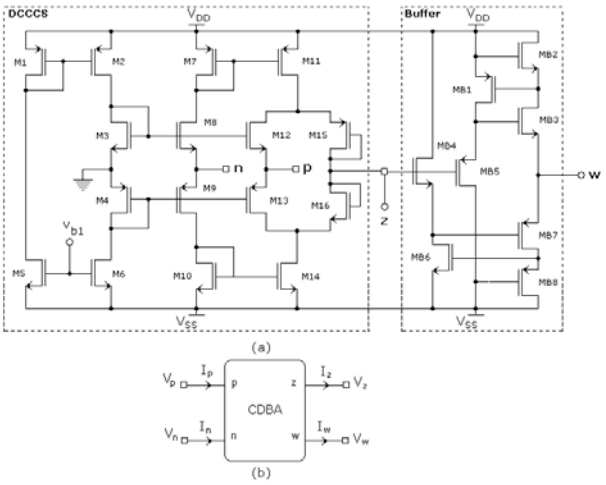


Fig. 1 (a) Simplified circuit of modified CDBA and (b) its symbol

The proposed MFC is shown in Fig. 2 which composed of two P channel MOS based cascode current mirrors, a single CDBA, two bias current sources (I_{BR}, I_{BL}), and four N channel MOS transistors. The PMOS current mirrors are carried out reversing right and left currents from input NMOS transistors. The CDBA is operated to take difference between right and left currents from PMOS cascode mirrors.

The MOSFETs, M1, M2, M3, and M4, are the identical transistors, working in saturation region. In condition that $V_{GS}>V_T$ and $V_{DS}>V_{GS}-V_T$, the expression of drain current for the simple MOS transistor operating in saturation region is

$$I_{ds} = \mu C_{ox} \left(\frac{W}{L}\right)(V_{GS} - V_T)^2 \qquad (2)$$

$$I_{ds} = K(V_{GS} - V_T)^2 \qquad (3)$$

where **K** is trans-conductance parameter. μ, C_{ox}, W, and L stand for carrier effective mobility, gate oxide capacitance per unit area, width, and length of the channel, respectively.

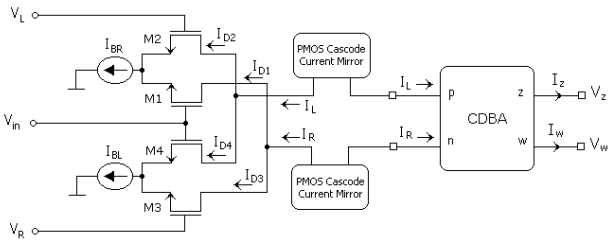


Fig. 2. The block representation of proposed CDBA-based MFC

In Fig.2, V_L and V_R can be adjusted to determine left and right side of membership function such as trapezoidal. The branch currents of proposed MFC are given as follows:

$$I_{D1} = \frac{I_{BR}}{2} + V_{D1}\frac{K_1}{2}\sqrt{\frac{2I_{BR}}{K_1} - V_{D1}^2} \qquad (4)$$

$$I_{D2} = \frac{I_{BR}}{2} - V_{D1}\frac{K_2}{2}\sqrt{\frac{2I_{BR}}{K_2} - V_{D1}^2} \qquad (5)$$

$$I_{D3} = \frac{I_{BL}}{2} + V_{D2}\frac{K_3}{2}\sqrt{\frac{2I_{BL}}{K_3} - V_{D2}^2} \qquad (6)$$

$$I_{D4} = \frac{I_{BL}}{2} - V_{D2}\frac{K_4}{2}\sqrt{\frac{2I_{BL}}{K_4} - V_{D2}^2} \qquad (7)$$

where difference voltages, V_{D1} and V_{D2}, are given by

$$V_{D1} = V_{in} - V_L \quad \text{and} \quad V_{D2} = V_{in} - V_R \qquad (8)$$

Also, right and left currents of MFC, I_R and I_L, can be obtained with KCL as follows:

$$I_R = I_{D1} + I_{D3} \quad \text{and} \quad I_L = I_{D2} + I_{D4} \qquad (9)$$

The output current, I_{out}, is equal to difference of left and right currents of MFC and it is obtained from z-terminal of CDBA.

$$I_{out} = I_L - I_R \qquad (10)$$

In condition that $K_1=K_2=K_3=K_4=K$ and $I_{BR}=I_{BL}=I_B$, the output current of MFC can be given in Eq. (11).

$$I_{out} = K\left[(V_{in} - V_L)\sqrt{\frac{2I_B}{K} - (V_{in} - V_L)^2} + (V_{in} - V_R)\sqrt{\frac{2I_B}{K} - (V_{in} - V_R)^2}\right] \qquad (11)$$

V_L must be greater than V_R in order to generate the Gaussian-type curve. The drain current equations of the MOSFETs are valid when voltage from gate-to-source, V_{GS}, is higher than V_T threshold voltage of related transistor. Hence, the following condition must be as follows:

$$\sqrt{\frac{2I_B}{K_i}} < V_{Di} < 2\sqrt{\frac{2I_B}{K_i}} \qquad i = 1,2 \qquad (12)$$

3. Simulation results

In order to verify the above given theoretical analysis, a new MFC is designed using the proposed configuration of Fig.2. The behaviour of the implemented CDBA-based MFC was confirmed with 0.5 µm MIETEC CMOS process parameters by PSPICE simulation experiments. The device dimensions of the channel width (W) and the channel length (L) for M_1, M_2, M_3, and M_4 are 10 mm and 4 µm, respectively. For all transistors operated in saturation region, bias currents of differential pairs, I_{BR} and I_{BL}, are adjusted to 10 µA. The power supply is V_{DD}= -V_{SS}=1.5 V.

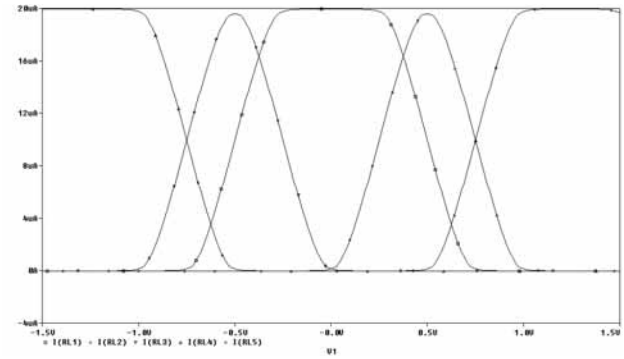


Fig. 3. The basic membership function output graphs of proposed MFC

The simulation result of basic four membership functions such as trapezoidal, S and Z shape obtained by proposed MFC is shown in Fig. 3. Here, right voltage V_R must be bigger than left voltage V_L. For example, in this study, triangular membership shape is obtained on condition that V_R-V_L=0.5 V.

Fig. 4 (a) and (b) show the width adjusting of MFC towards right and left, respectively. In this DC analysis result, the one of the adjusting voltages is fixed a constant value for extension to right or left direction of membership function. Also, the height of membership function (amplitude of output current) can be varied with adjusting bias currents, I_{BR} and I_{BL}, as shown in Fig. 5. Here, bias currents are varied between [10 µA; 50 µA] in 5 µA steps for V_R –V_L= 1V, i.e., trapezoidal membership shape centred 0 V. Also, the CDBA based MFC has capable of high-speed operations because of using CDBA structures in this design has got high frequency range. The frequency response of the MFC is shown in Fig. 6. The bandwidth of the MFC is about 266 MHz (for -20 dB) as shown in Fig. 6. Low power low voltage technology offers the possibility of connecting the bulk and the source of transistors. In this case, the input current swing increases but the bandwidth is reduced substantially, as it is expected

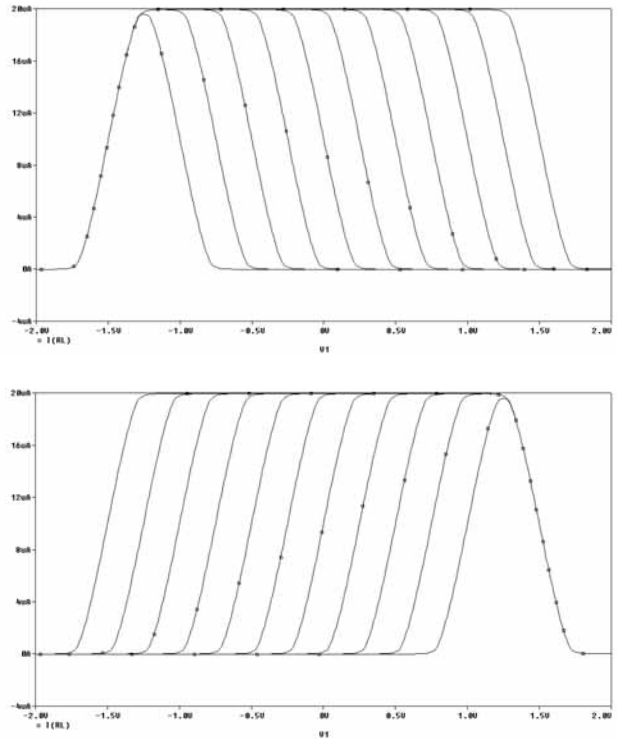


Fig. 4. Width adjusting of proposed MFC (a) Right directions, and (b) Left directions

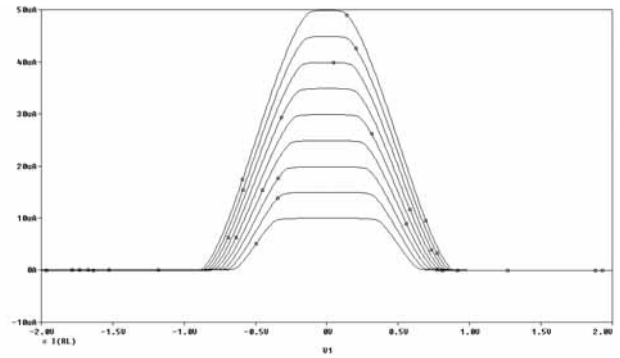


Fig. 5. Width adjusting of membership function with varied bias currents, I_{BR} and I_{BL}, for trapezoidal shape centered 0V. (I_{BR} = I_{BL} and between [10 µA; 50 µA] with 5 µA steps)

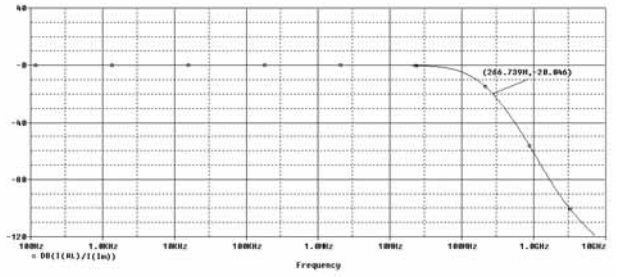


Fig. 6. Frequency response of the proposed CDBA based Membership Function Circuit (MFC)

4. Conclusion

A new and fully integrable membership function circuit (MFC) is proposed by using single CDBA. The performance of proposed MFC is analyzed based upon PSPICE simulation experiments. This MFC operates low supply voltages as ± 1.5 V. Also, the proposed MFC has got both wide input current range and high frequency response. The proposed circuit configuration consists of a single CDBA and four N channel MOS transistors operated in saturation region. This MFC is suitable for both current-mode and voltage-mode fuzzy hardware because of a voltage buffer is available in CDBA structure. Basic four membership functions can be implemented by this MFC. The characteristics such as width, height, and position of implemented membership functions are easily adjustable by varying left and right voltages, and bias currents.

References

/1/ D.G. Zrilic, J. Ramirez-Angulo, B. Yuan, "Hardware Implementations of Fuzzy Membership Functions, Operations, and Inference", *Computers and Electrical Engineering*, vol. 26, pp. 85-105, 2000.

/2/ T. Kettner, C. Heite, and K. Schumacher, "Analog CMOS Realization of Fuzzy Logic Membership Functions", *IEEE Jour. of Solid-State Circuits*, vol. 28, no.7, pp. 857-861, 1993.

/3/ M. Tokmakçý, M. Alçý, and R. Kýlýç, "A simple CMOS-based Membership Function Circuit", *Analog Integrated Circuits and Signal Processing*, vol. 32, pp. 83-88, 2002.

/4/ S. Collins, G.F. Marshall, "Subthreshold Membership Function Circuit", *Electronics Letters*, vol. 34, no. 11, pp. 1113-1114, 1998.

/5/ I. Baturone, S. Sánchez-Solano, Á. Barriga, and J. L. Huertas, 'Implementation of CMOS Fuzzy Controllers as Mixed-Signal Integrated Circuits', *IEEE Trans. on Fuzzy Systems*, Vol.5, No.1, pp.1-19, Feb.1997.

/6/ S. Bouras, M. Kontronakis, K. Suyama, and Y. Tsividis, 'Mixed Analog-Digital Fuzzy Logic Controller with Continuous-Amplitude Fuzzy Inferences and Defuzzification', *IEEE Trans. on Fuzzy Systems*, vol.6, No.2, 205-215, 1998.

/7/ R.G. Carvajal, A. Torralba, F. Colodro, and L.G. Franquela, "Mixed-Signal CMOS Fuzzifier with Emphasis in Power Consumption*", IEEE 42nd Midwest Symp. on Circuits and Systems*, vol.2, pp. 929-933, 2000.

/8/ P. Saavedra, A. López, J. Zrilic, and J. Ramírez-Angulo, "New Analog Current-mode/Voltage-mode Fuzzifier with continuously Adjustable Parameters", *IEEE Proc. 40th Midwest Symposium on Circuits and Systems*, USA, vol.1, pp. 31-34, 1997.

/9/ J. Choi, B. J. Sheu, and J. C.-F. Chang, "A Gaussian Synapse Circuit for Analog VLSI Neural Networks," *IEEE Trans. on Very Large Scale Integration (VLSI) System*s, vol. 2, no. 1, pp. 129-133, March 1994.

/10/ L. Tigaeru, "Programmable Analog Membership Function Circuit for Hybrid-mode Fuzzy Systems", *Electronics Letters*, vol. 39, no.8, pp. 642-644, 2003.

/11/ E. Seevinck, and R.J. Wiegerink, "Generalized Translinear Circuit Principle", *IEEE Jour. of Solid-State Circuits*, vol.26, no.8, pp.1098-1102, 1991.

/12/ C. Acar, and S. Özoğuz, "A Versatile Building Block: Current Differencing Buffered Amplifier Suitable for Analog Signal Processing Filters," *Microelectronics Journals*, vol.30, pp.157-160, 1999.

/13/ K. Manetakis, C. Toumazou, 'Current-feedback opamp suitable for CMOS VLSI technology', *Electronics Letters,* vol. 32, no.12, pp.1090-1092, 1996.

/14/ A. Toker, S. Özcan, O. Çiçekoğlu, and C. Acar, "Current-mode allpass filters using CDBA and a new high Q bandpass filter configuration", *IEEE Trans. on Circuits and Systems-II: Analog and Digital Signal Processing*, vol.47, no.9, pp.949-954, 2000.

/15/ W. Tangsrirat, W. Surakampontorn, and N. Fujii, "Realisation of leaprog filters using current differencing buffered amplifiers," *IEICE Transaction on Fundemantals*, vol. E86-A, no.2, pp.318-326, 2003.

/16/ S. Celma, J. Sabadell, and P. Martinez, "Universal filter using unity-gain cells," Electronics Letters, vol.31, no.21, pp.1817-1818, 1995.

Mahmut Tokmakçi, Mustafa Alçi
Erciyes University, Engineering Faculty, Dept. of Electrical&Electronics Engineering, 38039, Kayseri, Turkey

e-mail: tokmakci@erciyes.edu.tr; malci@erciyes.edu.tr

Prispelo (Arrived): 30.04.07 *Sprejeto (Accepted):* 28.5.08

AN EXAMPLE OF AN INTEGRATED GPS AND DR POSITIONING SYSTEM DESIGNED FOR ARCHEOLOGICAL PROSPECTING

Franc Dimc [1], Branko Mušič [2], and Radko Osredkar [3]

[1] University of Ljubljana, Faculty of Maritime Studies and Transportation, Portorož, Slovenia

[2] University of Ljubljana, Department of Archeology, Ljubljana, Slovenia

[3] University of Ljubljana, Faculty of Computer and Information Science, Ljubljana, Slovenia

Key words: dead reckoning, integrated navigation, geophysical exploration, MEMS accelerometer

Abstract: Geophysical prospecting at archeological sites requires a flexible and relatively inexpensive positioning system, capable of determining positions within a circle with 30 cm error radius. GPS positioning technology is capable to achieve this accuracy only under favorable conditions, but can be supplemented, for more general use, with a dead reckoning (DR) MEMS accelerometer positioning system. In this contribution we describe such a hybrid GPS-DR system, developed for archeological applications, capable of reducing the trajectory orientation errors, which for the GPS alone can amount to ± 1 °, by 60 %. The improvement in the position data to the required precision in archeological fieldwork is demonstrated.

Primer integriranega lokacijskega sistema GPS in seštevne navigacije, namenjenega arheološkemu raziskovanju

Kjučne besede: seštevna navigacija, integrirana navigacija, geofizikalne raziskave, MEMS akcelormeter

Izvleček: Geofizikalne raziskave na arheoloških najdiščih zahtevajo prožen in relativno cenen sistem pozicioniranja, ki je spodoben določati položaj znotraj kroga napake s polmerom 30 cm. GPS navigacijska tehnologija je sposobna zagotoviti takšno natančnost le v ugodnih razmerah, ki jih pri terenskem delu običajno ni mogoče zagotoviti. Vendar pa se jo da z inercialnim navigacijskim sistemom, ki temelji na uporabi MEMS pospeškometrov, podpreti. V tem prispevku opisujemo hibriden sistem pozicioniranja, temelječ na GPS in seštevni navigaciji (dead reckoning – DR – navigacija), razvit za potrebe arheološke prospekcije, ki je sposoben zmanjšati orientacijsko napako GPS sistema (običajno v bližini ± 1 °) za 60 %. Pokažemo, da sistem lahko zagotovi pri arheološkem delu zahtevano natančnost.

1. Introduction

There is considerable commercial and research interest in a variety of applications for accurate location systems. Different systems are characterized by different levels of accuracy, coverage, cost of installation, operation and maintenance, frequency of location updates, etc /1/, /2/. Not surprisingly, in any application a system has its respective strengths and weaknesses. For example, a standard GPS-based system can achieve good levels of accuracy but in practice the coverage is often limited. GSM-based positioning has been advocated in various applications and claimed to achieve meaningful levels of accuracy at a very low cost, however not in the accuracy range of the GPS-based positioning /3/. It seems quite obvious that the choice of any particular location system for a specific application is highly application-specific. In this paper we consider a hybrid system for positioning in a challenging real world application: geophysical surveying of an archeological site. A typical application of the positioning system in an archeological prospecting setting is an amalgamation of position data with a ground penetrating radar, and/or ground conductivity or magnetic anomaly meters, etc /4/, /5/.

Geophysical measurements for extensive surveys at archaeologically unexplored sites, usually performed via common terrestrial geodetic procedures, can be considerably aided by the satellite navigation methods. The idea behind applying inexpensive, single frequency GPS systems to geophysical surveys in archaeological prospecting arises from their simple use, and the need for flexibility and effectiveness in such work. A sufficient degree of flexibility can only be assured by complete autonomy of the surveying team in the planning and execution of the fieldwork. GPS positioning technology is well suited for autonomous geophysical prospecting /5/, /6/, in particular in large-scale evaluation projects, where a comprehensive, detailed survey is not as important as establishing the background levels and acquiring a good understanding of the effects of the geology and pedology. Its efficient application is important for surveys in regions where geodetic fixed points for terrestrial position measurements are not accessible or available. In our previous work we have tested various types of inexpensive GPS receivers to asses the util-

ity of the technique for archeo-geophysical surveying /5/, /7/, and demonstrated that it is possible to compensate for the errors, usually encountered in GPS positioning, to a considerable extend, and refine the results by post-processing of the measurement data. Under favorable conditions a 30 cm relative uncertainty of the kinematical positioning was achieved. This is considered sufficient in the type of surveying under discussion /8/.

It is characteristic of the GPS positioning methods that the accuracy of relative positions fixes is usually considerably better than the accuracy of absolute position fixes. However, with a comprehensive post-processing effort a very accurate static fix by single frequency GPS receiver is also possible /9/. We have attempted to improve the performance of relatively inexpensive GPS receivers by first establishing reference points of known position in the field. A roving receiver then continuously corrects its position outputs by calculating its pseudo range. This can be done relatively easily at a test site under favorable measuring conditions. However, an important practical distinction favors test sites from real life archeological sites in that at the later favorable conditions are seldom encountered. Therefore methods of augmenting a GPS system were sought. As in our previous work, we are concerned solely with horizontal positioning, which is of primary importance, leaving the more difficult problem of vertical positioning to a future project.

2. Hybrid GPS-MEMS accelerometer DR location system

Several methods of augmenting a GPS positioning system suggest themselves. It has been demonstrated /10/, that often a fusion of GPS and dead reckoning (DR), implemented e.g. by a compass and accelerometers, shows the best accuracy, significantly improving the accuracy of GPS alone. Such a fusion is by no means self evident, especially if relatively cheap and readily available micro machined MEMS accelerometers are used for augmenting the GPS system. There is a significant difficulty associated with the use of such MEMS sensors: their electrical parameters are not well defined and each manufactured device must be individually calibrated prior to use and relatively frequently recalibrated /11/. However, in this paper we attempt to demonstrate the viability and possibilities of a hybrid GPS-MEMS accelerometer DR location system for archeological prospecting fieldwork.

Despite a variety of detection schemes employed, every accelerometer can be modeled as a simple mass-spring-damper system, where the mass on the spring changes its position relative to its supporting frame with the input acceleration, forming a second-order system:

$$m\ddot{x} + c\dot{x} + kx = ma_{input}$$

where x is the displacement of the "proof" mass m of the accelerometer with respect to its frame, a_{input} is the external (input) acceleration, k is the suspension coefficient, and c is the damping coefficient of the system /12/. Determination of the displacement of the device along a trajectory with time proceeds by numerical integration of the input acceleration. Due to small errors of various sources in the accelerometer, integrator circuitry etc, long time integration results in an accumulation of errors and double integration compounds the errors ($\sim t^2$). Without frequent resetting of the measured position of the device to the actual position, huge errors can result. On the other hand, integration over short periods of time (in the order of 1 minute or less) can give satisfactory results, making supplementation of temporarily unavailable GPS signals with DR position data possible, or a correction of them, if the source of the GPS error can be detected. One such known source of GPS errors that can be corrected in a kinematical setting with a DR procedure is the error, due to the time delay between the actual position and the position indicated by the GPS system after the measurement and computation. The correction procedure implies a time triggered data acquisition procedure for both parts of the hybrid positioning system. The fusion of the DR and GPS is able to reduce the average GPS error by approximately 50% /10/ in a general setting and has been reported to reduce the maximum error by 70 % in favorable circumstances. It is perhaps remarkable, that integration of more than two positioning systems into one usually does not result in any further improvement of the accuracy /10/.

Due to the nature of the application of the hybrid positioning system being considered in our investigation, i.e. geophysical surveying of an archeological site, the system needs to be as simple as possible. It is basically comprised of 3 sensors: a GPS receiver and two MEMS accelerometers, augmented by a magnetic compass. The later serves principally as an orientation sensor, which is needed as an accelerometer by itself does not differentiate between a dynamic and the gravitational accelerations. In principle the data from the GPS receiver alone could be used to derive the required device orientation, eliminating the need for a compass. In practice, however, this proved to be unreliable. On the other hand, orientation data from 2 GPS receivers proved sufficiently accurate and reliable for the purposes of the proposed application. In this case a second GPS receiver replaces the compass, resulting in a hybrid system that has proved itself to be even easier to manage than the one with a compass. The practical difference between the two is basically due to different and incommensurable sampling rates from the GPS receiver and the compass, often causing timing errors. We need to stress again that system devised is tailored to a particular application and is therefore perhaps no suited or optimal in other applications.

The hybrid location system we have developed consists of a GPS receiver (FlexPak, NovAtel) with choke-ring antenna, two perpendicularly mounted two-dimensional MEMS accelerometers (ADIS16201, Analog Devices) and a magnetic compass (HMR3300, Honeywell). Sensors are driv-

en by a microcontroller A T90S4433 (Atmel), and their data transferred to a personal computer, used for the data processing, via two ISM radio band links (ER 400TRS, Low Power Radio Solutions). During use of the system the GPS position data are combined and corrected with the acceleration data by the standard DSP techniques /13/. Walking speeds attainable by a prospector in a field rarely exceed 1 m/s, and his kinematical position (actually, the position of the device he carries or drags behind him) can thus be deduced from accelerometer data sampled below 10Hz. A synchronization of various positioning device outputs of the hybrid system involves a matching algorithm which takes typical delay times in consideration. An effective coupling of the DR data with the GPS data is then achieved by Kalman filtering. The scheme of the system is shown on Fig. 1 and the system itself in Fig 2. Fig. 3 illustrates the whole positioning system, mounted on a carrier together with a ground penetrating radar, in use.

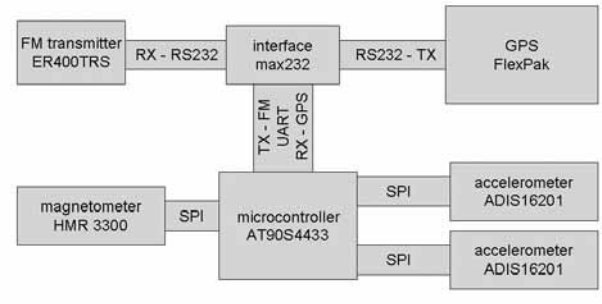


Fig. 1. A block diagram of the GPS/MEMS dead reckoning (DR) hybrid device.

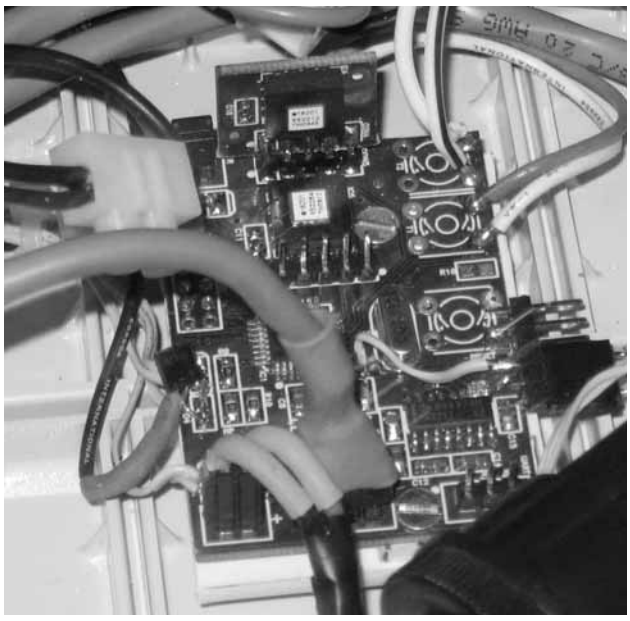


Fig. 2. A detail of the hybrid system: the two perpendicularly mounted MEMS chips with the controller board. Due to the interferences caused by the GPS receiver, the magnetometer unit is placed separately.

Fig. 3. Ground Penetrating Radar (GPR), the position of which is recorded by the proposed hybrid positioning system, comprising a GPS receiver SR 20, georeferenced by GS 20 (Leica), and the dead reckoning (DR) accelerometer subsystem, in use at an archeological site. Both positioning devices are placed at the top of the GPR antenna.

3. Results and discussion

The starting point of the present investigation is the result, established in our previous work /14/, that a GPS positioning system is not perfectly capable of assuring the required accuracy of 30 cm in all circumstances. The absolute distances between two locations in a field, as measured with such a system, are generally compatible with the required precision, but the orientation of the measured trajectory of the surveyor in the field is not. Even in favorable conditions the orientation of a measured straight trajectory deviates from the actual trajectory of a surveyor by ± 1 °, resulting in a positioning error that exceeds the required accuracy after only 17 m displacement along a straight trajectory. To a certain degree this can be compensated for by frequent referral to the established reference points in the field and/or corrected by the calibrated DR acceleration data.

Calibration of the DR sub-system of the integrated positioning device has been preformed along a 15 m straight track, measured with a laser measuring device, with well defined starting and end points. The instrument, positioned on a sledge, was dragged along the chosen path at a typical rate of somewhat under 1 m/s, and 60 runs were performed in several sessions on different days. Fig.4 represents a typical run along the test track. The waviness of the trajectory in an illustration of the fact, that for a human operator, dragging the sledge with the measuring equipment behind him, it is considerably easier to aim for the final position than to accurately follow a predetermined path, even if this is straight. The final position data were analyzed according to standard procedures /15/ and com-

pared to the actual position of the end point. The measurements from a single session comprise a cluster of data form which the average final position for that session is calculated. This position varies between different sessions, with a compound rms final position error for all clusters together being 25 cm, which corresponds to a ± 0.95 ° direction error. This is comparable to the direction error of the GPS measurements alone. However, separately considering only clustered data from a single session, the results show a considerably more favorable picture. The rms error of a cluster of data is 15 cm in the longitudinal direction, and 11 cm in the transverse direction. We attribute this difference to the fact that the disturbing forces during a run act primarily parallel to the path taken, with the transverse disturbances being considerably smaller. The direction accuracy, which is of primary concern to us, depends only upon the scatter of the data in the transverse direction and the stated rms in this direction represents an angular error of only 0.40 °, an improvement of 60 % relative to the GPS positioning. This is entirely within the range of reported performances of similar systems in different applications /10/. By combining the absolute displacement data, as measured by the GPS subsystem, with DR direction data, our hybrid system is thus capable of assuring the accuracy of positioning as required by the archeological field work.

The origins of clustering of the accelerometer position data between sessions are not understood at this time. In the field work it would be of considerable practical importance to eliminate the clustering. On the other hand, to large extend the question can be avoided by a resetting of the DR subsystem at the end point of the test track or, in the field, at the previously established reference points. The overall reliability and robustness of the DR subsystem is demonstrated by the fact that, regardless of the clustering, the average of the absolute distances traveled by the device over all sessions, as measured by the DR method, is within 1 cm of the actual distance between the starting an end points of the 15 m test track. Thus we conclude that the scatter of the final position data is caused by random and not systematic errors.

The determination of the actual trajectory taken by the surveyor in the field proceeds by measuring the staring and end points of segments of the trajectory taken. Starting form a known position, the pseudo position of the end point of each approximately 10 – 15 m long segment is first determined from the GPS data, and then corrected by the DR direction data. This position is then taken as the starting point of the next segment and its endpoint determined in the same manner. From the known positions of the starting and end points of a segment, the actual position of a point on a segment can then be determined by the standard interpolation of its pseudo position. The segmentation of the trajectory thus takes advantage of the relatively accurate absolute distance data offered by the GPS part of the hybrid system, and the relatively accurate direction data of the DR subsystem, whose limited range is not exceed-

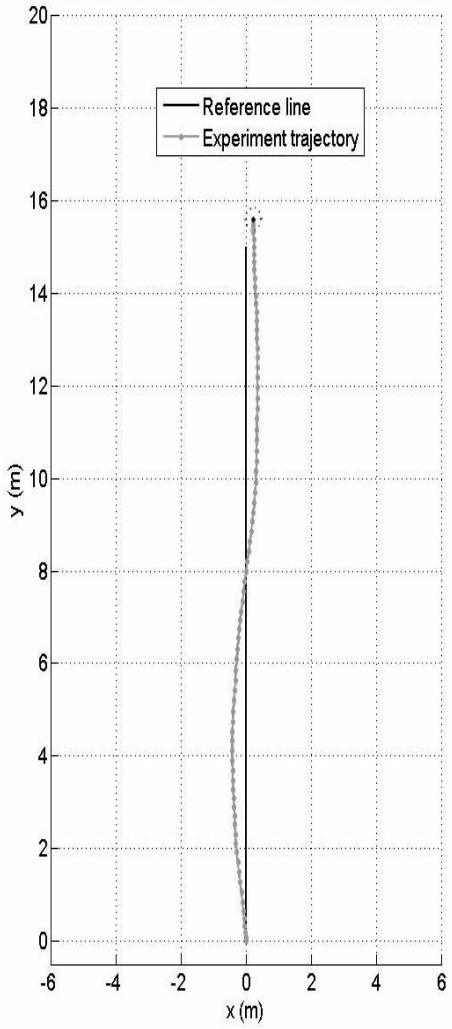


Fig.4. A typical run along the test track. The waviness of the trajectory illustrates the difficulty of dragging the sledge with the measuring equipment along a predetermined path. For a human operator it is considerably easier to aim for the final position, compensating for digressions, which results in a wavy trajectory.

ed by the segment length. A successful implementation of the proposed GPS and MEMS DR sensors fusion in realistic surveying conditions is significantly augmented by an appropriate analytical model of the data acquisition and manipulation process. The model is tuned by fitting its outputs to the positions of the established reference points, and a correspondence between the hybrid system outputs and the subsequent DSP processing is thus established. The model is still being improved, but its original version and the described hybrid positioning system have been extensively tested and used at Tanagra, a large and unexplored archeological site in Greece. The data obtained there are still being evaluated and results will be reported elsewhere.

A completely different approach to accurate positioning with the same hybrid system as described above is suggested by Cho and Park /16/: the change in position of a

surveyor is calculated by the estimated average length of his step, the number of steps he takes, and their directions. Correct step length estimate is crucial for the success of the method. The acceleration signal pattern from the MEMS sensors is used to count the steps. However, the unavoidable misalignments of the device are a source of serious errors, and the magnetic compass azimuth data must be used in the tilt compensation algorithm. With this approach we have been unable to reproduce or even approach the precision of the hybrid system, described above, even on a level, smooth test-field, and at limited displacements. Thus the validity of such an approach in a more general situation is questionable.

4. Conclusion

Geophysical prospecting at archeological sites requires a flexible and relatively inexpensive positioning system, capable of determining the positions of test equipment within a circle with 30 cm error radius. As the GPS positioning technology is capable of achieving this accuracy only under favorable conditions, seldom encountered in fieldwork, we have developed a hybrid positioning system, comprising of a GPS receiver and (DR) MEMS accelerometers, which is capable of augmenting the GPR positioning data. The described system is capable of reducing the GPS trajectory orientation errors, which usually amount to ± 1 °, by 60 %, thus assuring the required accuracy. We report on the calibration of the system in test surroundings, while and extensive survey of an actual archeological site, which has been performed with the system, is still being evaluated and will be reported at a later date.

References

/1/ J. Hightower, G. Borricello, Location Systems for Ubiquitous Computing. Computer, 34 (2001), 57–66.

/2/ M. Hazas, J. Scott, J. Krumm, Location-Aware Computing Comes of Age, Computer, 37 (2004), 95–97.

/3/ A. Varshavsky, et al., Are GSM Phones THE Solution for Localization?, WMCSA 2006 (2006).

/4/ G. Barratt, V. Gaffney, H. Goodchild, S. Wilkes, Survey at Wroxeter Using Carrier Phase, Differential GPS Surveying Techniques, Department of Ancient History and Archaeology, Birmingham University Field Archaeology Unit, Univ. of Birmingham, Archaeological Prospection, 7 (2000), pp. 133–143.

/5/ F. Dimc, B. Mušič, R. Osredkar, Surveying for geophysical exploration, using a single-frequency global navigation satellite system receiver capable of 30 cm horizontal kinematical positioning uncertainty, RMZ - Materials and Geoenvironment, 55 (2008), pp. 85-110.

/6/ C. Gaffney, C. J. Gatter, Revealing The Buried Past. Geophysics for Archaeologists, Tempus Publishing Ltd., 2003.

/7/ F. Dimc, B. Mušič, R. Osredkar, Attaining Required Positioning Accuracy in Archeo-geophysical Surveying by GPS, Proc. of 12th International Power Electronics and Motion Control Conference, Portorož, Slovenia, Aug. 30 – Sept. 1, 2006, pp. 2037-2040.

/8/ J. Van Sickle, GPS for Land Surveyors, Taylor and Francis, New York, 2001, p.74.

/9/ T. Beran, et al., High-Accuracy Point Positioning With Low-Cost Receivers. Navigation, Journal of The Institute of Navigation, 54 (2007), pp. 53-63.

/10/ U. Steinhoff, et al., How Computer Vision Can Help in Outdoor Positioning, Lecture notes comput. sci., (2007), pp. 124-141.

/11/ D. Jurman, et al., Inertial and Magnetic Sensors: the Calibration Aspect, Informacije MIDEM 37(2007)2, pp. 67–72.

/12/ C. Acar, A. M. Shkel, Experimental Evaluation and Comparative Analysis of Commercial Variable-capacitance MEMS Accelerometers, University of California Postprints, Paper 256 (2003)

/13/ O. Mezentsev, Sensor Aiding of HSGPS Pedestrian Navigation, PhD Thesis, Department of Geomatics Engineering, University of Calgary, March 2005.

/14/ F. Dimc, B. Mušič, R. Osredkar, The fixes derived from the C/A code of the single frequency GPS receiver attain the demands for the archeo-geophysical surveying, Elektrotehniški vestnik, 74 (2007), 255-260.

/15/ J. Drnovšek, J. Bojkovski, G. Geršak, I. Pušnik, Metrologija, Univ. of Ljubljana, Fac. of Electrical Engineering, 2004, p. 132

/16/ Cho S. Y., Park C. G., MEMS based Pedestrian Navigation System, The Journal of Navigation, The Royal Institute of Navigation, Vol. 59, No.1, pp. 135 - 153,2006.

Franc Dimc, University of Ljubljana, Faculty of Maritime Studies and Transportation, Nautical Engineering Department, Pot pomorščakov 4, 6320 Portorož, Slovenia.
Tel.: +386 (0) 5 6767 293; Fax.: +386 (0) 5 6767 295
E-mail: franc.dimc@fpp.uni-lj.si

Branko Mušič, University of Ljubljana, Department of Archeology, Zavetiška ulica 5, 1000 Ljubljana, Slovenia

Radko Osredkar, University of Ljubljana, Faculty of Computer and Information Science, Tržaška cesta 25, 1000, Ljubljana, Slovenia

Prispelo (Arrived): 08.04.08 *Sprejeto (Accepted):* 28.5.08

PIEZO 2009

Electroceramics for End users IV

Belvedere Hotel, Zakopane, Poland

March 1-4, 2009

www.piezo2009.com

IMPORTANT DATES

Abstract deadline: Nov 1, 2008

Notification of acceptance: Nov 15, 2008

Early registration: Dec 1, 2008

GENERAL CHAIRS

Prof. Marija Kosec, Jožef Stefan Institute, Slovenia

Prof. Marc Lethiecq, University of Tours, France

PIEZO 2009

Electroceramics for End-users IV is the next scientific event in the series of conferences dedicated to advances in electroactive, particularly piezoceramic, materials and devices.

It was established by the POLECER Thematic network and continued by the MIND Network of Excellence, starting in Interlaken, Switzerland, 2002. Following conferences were:

Capri, Italy, 2003

Bled, Slovenia, 2003

Courmayeur, Italy, 2004

Abeyé les Vaux de Cernay, France, 2005

Hafjell, Norway, 2006

Liberec, Czech Republic, 2007

ABOUT ZAKOPANE

Zakopane is located near the "Tatra Mountains" in southern Poland, about 110 km from Krakow.

CONTACT

Piezo 2009 Secretariat

Konstantin Astafiev

konstantin.astafiev@piezoinstitute.com

Piezo INSTITUTE

The European Institute of Piezoelectric Materials and Devices

CONFERENCE TOPICS

- Piezoelectrics, ferroelectrics, relaxors, tunable materials, multiferroics etc.
- Processing: bulk, thick and thin films, LTCC, integration, novel lithography-free deposition methods etc.
- Lead free and sustainable materials
- MEMS
- Materials and devices for high frequency ultrasonic medical imaging
- Other ultrasonic applications (NDE, underwater, ranging etc.)
- Concepts and materials for very high temperature applications
- High power devices (medical therapy etc.)
- Energy harvesting concepts and devices
- Actuators (for ink jet printers, fuel injection etc.)
- Modelling

INVITED SPEAKERS

Dragan Damjanovic, EPFL, LC, Switzerland

Richard Greaves, MEGGITT Sensors, UK & Switzerland

Andrzej Nowicki, IPPT, Poland

Carsten Schuh, Siemens AG, Corporate Technology

Eberhard Hennig, PI Ceramics, Germany

Susan Trolier-McKinstry, Ceramic Science and Engineering, Penn State, USA

Satoshi Wada, University of Yamanashi, Japan

Tomasz Zawada, InSensor A/S, Denmark

TUTORIALS

Design of ultrasonic transducers for medical imaging and NDE applications

Processing of piezoelectric materials (thin/thick films and ceramics)

CD ROM - dvajset letnikov revije Informacije MIDEM

Revija Informacije MIDEM je krepko vstopila v svoje tretje desetletje.
Ob tem jubileju smo pripravili posebno izdajo zadnjih dvajsetih letnikov v elektronski obliki.

Izkoristite svojevrstno priložnost, da izpopolnite svojo strokovno knjižnico.

Naročite svoj osebni izvod.

Za člane društva MIDEM je cena CD ROMa 50,00EUR, za ostale pa 60,00EUR.

Naročila sprejema uredništvo revije.

Pišite na: iztok.sorli@guest.arnes.si

MIDEM
DRUŠTVO-SOCIETY
IMAPS Slovenia Chapter

M I D E M

Strokovno društvo za mikroelektroniko,
elektronske sestavne dele in materiale
MIDEM pri MIKROIKS
Stegne 11, 1521 Ljubljana
SLOVENIJA

TEL.: +386 (0)1 5133 768
FAX: +386 (0)1 5133 771
Email / WWW
iztok.sorli@guest.arnes.si
http://paris.fe.uni-lj.si/midem/

MIDEM SOCIETY REGISTRATION FORM

1. First Name ..... Last Name .....

   Adress .....

   City .....

   Country ..... Postal Code .....

2. Date of Birth .....

3. Education (please, circle whichever appropriate)

   PhD MSc BSc High School Student

3. Profession (please, circle whichever appropriate)

   Electronics Physics Chemistry Metallurgy Material Sc.

4. Company .....

   Adress .....

   City .....

   Country ..... Postal Code .....

   Tel.: ..... FAX: .....

   Email .....

5. Your Primary Job Function

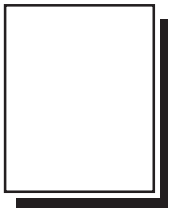
   Fabrication Engineering Facilities QA/QC
   Management Purchasing Consulting Other

6. Please, send mail to a) Company adress b) Home Adress

7. I wil regularly pay MIDEM membership fee, 25,00 EUR/year

MIDEM member recive Journal "Informacije MIDEM" for free !!!

Signature ..... Date .....



MIDEM at MIKROIKS
Stegne 11

1521 Ljubljana
Slovenija

Informacije MIDEM
Strokovna revija za mikroelektroniko, elektronske sestavine dele in materiale

NAVODILA AVTORJEM

Informacije MIDEM je znanstveno-strokovno-društvena publikacija Strokovnega društva za mikroelektroniko, elektronske sestavne dele in materiale - MIDEM. Revija objavlja prispevke s področja mikroelektronike, elektronskih sestavnih delov in materialov. Ob oddaji člankov morajo avtorji predlagati uredništvu razvrstitev dela v skladu s tipologijo za vodenje bibliografij v okviru sistema COBISS.
Znanstveni in strokovni prispevki bodo recenzirani.

Znanstveno-strokovni prispevki morajo biti pripravljeni na naslednji način:

1. Naslov dela, imena in priimki avtorjev brez titul, imena institucij in firm
2. Ključne besede in povzetek (največ 250 besed).
3. Naslov dela v angleščini.
4. Ključne besede v angleščini (Key words) in podaljšani povzetek (Extended Abstract) v angleščini, če je članek napisan v slovenščini
5. Uvod, glavni del, zaključek, zahvale, dodatki in literatura v skladu z IMRAD shemo (Introduction, Methods, Results And Discsussion).
6. Polna imena in priimki avtorjev s titulami, naslovi institucij in firm, v katerih so zaposleni ter tel./Fax/Email podatki.
7. Prispevki naj bodo oblikovani enostransko na A4 straneh v enem stolpcu z dvojnim razmikom, velikost črk namanj 12pt. Priporočena dolžina članka je 12-15 strani brez slik.

Ostali prispevki, kot so poljudni cčlanki, aplikacijski članki, novice iz stroke, vesti iz delovnih organizacij, inštitutov in fakultet, obvestila o akcijah društva MIDEM in njegovih članov ter drugi prispevki so dobrodošli.

Ostala splošna navodila

1. V članku je potrebno uporabljati SI sistem enot oz. v oklepaju navesti alternativne enote.
2. Risbe je potrebno izdelati ali iztiskati na belem papirju. Širina risb naj bo do 7.5 oz.15 cm. Vsaka risba, tabela ali fotografija naj ima številko in podnapis, ki označuje njeno vsebino. Risb, tabel in fotografij ni potrebno lepiti med tekst, ampak jih je potrebno ločeno priložiti članku. V tekstu je treba označiti mesto, kjer jih je potrebno vstaviti.
3. Delo je lahko napisano in bo objavljeno v slovenščini ali v angleščini.
4. Uredniški odbor ne bo sprejel strokovnih prispevkov, ki ne bodo poslani v dveh izvodih skupaj z elektronsko verzijo prispevka na disketi ali zgoščenki v formatih ASCII ali Word for Windows. Grafične datoteke naj bodo priložene ločeno in so lahko v formatu TIFF, EPS, JPEG, VMF ali GIF.
5. Avtorji so v celoti odgovorni za vsebino objavljenega sestavka.

Rokopisov ne vračamo. Rokopise pošljite na spodnji naslov.

Uredništvo Informacije MIDEM
MIDEM pri MIKROIKS
Stegne 11, 1521 Ljubljana, Slovenia
Email: Iztok.Sorli@guest.arnes.si
tel. (01) 5133 768, fax. (01) 5133 771

Informacije MIDEM
Journal of Microelectronics, Electronic Components and Materials

INSTRUCTIONS FOR AUTHORS

Informacije MIDEM is a scientific-professional-social publication of Professional Society for Microelectronics, Electronic Components and Materials – MIDEM. In the Journal, scientific and professional contributions are published covering the field of microelectronics, electronic components and materials.
Authors should suggest to the Editorial board the classification of their contribution such as : original scientific paper, review scientific paper, professional paper…
Scientific and professional papers are subject to review.

Each scientific contribution should include the following:

1. Title of the paper, authors' names, name of the institution/company.
2. Key Words (5-10 words) and Abstract (200-250 words), stating how the work advances state of the art in the field.
3. Introduction, main text, conclusion, acknowledgements, appendix and references following the IMRAD scheme (Introduction, Methods, Results And Discsussion).
4. Full authors' names, titles and complete company/institution address, including Tel./Fax/Email.
5. Manuscripts should be typed double-spaced on one side of A4 page format in font size 12pt. Recommended length of manuscript (figures not included) is 12-15 pages
6. Slovene authors writing in English language must submit title, key words and abstract also in Slovene language.
7. Authors writing in Slovene language must submit title, key words and extended abstract (500-700 words) also in English language.

Other types of contributions such as popular papers, application papers, scientific news, news from companies, institutes and universities, reports on actions of MIDEM Society and its members as well as other relevant contributions, of appropriate length , are also welcome.

General informations

1. Authors should use SI units and provide alternative units in parentheses wherever necessary.
2. Illustrations should be in black on white paper. Their width should be up to 7.5 or 15 cm. Each illustration, table or photograph should be numbered and with legend added. Illustrations, tables and photographs must not be included in the text but added separately. However, their position in the text should be clearly marked.
3. Contributions may be written and will be published in Slovene or English language.
4. Authors must send two hard copies of the complete contributon, together with all files on diskette or CD, in ASCII or Word for Windows format. Graphic files must be added separately and may be in TIFF, EPS, JPEG, VMF or GIF format.
5. Authors are fully responsible for the content of the paper.

Contributions are to be sent to the address below.

Uredništvo Informacije MIDEM
MIDEM pri MIKROIKS
Stegne 11, 1521 Ljubljana, Slovenia
Email: Iztok.Sorli@guest.arnes.si
tel.+386 1 5133 768, fax.+386 1 5133 771

ORGANI DRUŠTVA MIDEM - BOARDS OF MIDEM

Predsednik društva MIDEM
MIDEM President

Prof. dr. Slavko Amon, univ. dipl. inž. fiz., Fakulteta za elektrotehniko, Ljubljana, Slovenija

IZVRŠILNI ODBOR MIDEM - MIDEM EXECUTIVE BOARD

Podpredsednika
Vice-presidents

Prof. dr. Marija Kosec, univ. dipl. inž. kem., Institut "Jožef Stefan", Ljubljana, Slovenija
Dr. Iztok Šorli, univ. dipl. inž. fiz., MIKROIKS, d. o. o., Ljubljana, Slovenija

Tajnik
Secretary

Igor Pompe, univ. dipl. inž. el., Ljubljana, Slovenija

Člani izvršilnega odbora MIDEM
Executive Board Members

Prof. dr. Lojze Trontelj†, univ. dipl. inž. el., Ljubljana, blagajnik
Darko Belavič, univ. dipl. inž. el., HIPOT-RR, d. o. o., Šentjernej
Prof. dr. Bruno Cvikl, univ. dipl. inž. fiz., Fakulteta za gradbeništvo, Maribor
Mag. Leopold Knez, univ. dipl. inž. el., Iskra TELA, d. d. , Ljubljana
Dr. Miloš Komac, univ. dipl. inž. kem., ARRS, Ljubljana
Jožef Perne, univ. dipl. inž. el. , IskraEmeco, d. d., Kranj
Prof. dr. Marko Topič, univ. dipl. inž. el.; Fakulteta za elektrotehniko, Ljubljana
Prof. dr. Anton Zalar, univ. dipl. inž. met., Institut "Jožef Stefan", Ljubljana
Dr. Werner Reczek, Infineon, Avstrija
Prof. dr. Giorgio Pignatel, University of Perugia, Italija
Prof. dr. Leszek J. Golonka, Technical University, Wroclaw, Poljska

Nadzorni odbor
Supervisory Board

Prof. dr. Franc Smole, univ. dipl. inž. el., Fakulteta za elektrotehniko, Ljubljana
Mag. Andrej Pirih, univ. dipl. inž. el., Iskra - Zaščite, d. o. o. , Ljubljana
Dr. Slavko Bernik, univ. dipl. inž. kem., Institut "Jožef Stefan", Ljubljana

Častno razsodišče
Court of Honour

Franc Jan, univ. dipl. inž. fiz., Kranj
Prof. dr. Radko Osredkar, univ. dipl. inž. el., Fakulteta za elektrotehniko, Ljubljana
Mag. Milan Slokan†, univ. dipl. inž. kem., Ljubljana

Sponzorji MIDEM

MIDEM Sponsors

JAVNA AGENCIJA ZA RAZISKOVALNO DEJAVNOST RS, Ljubljana

INSTITUT "JOŽEF STEFAN", Odsek za elektronsko keramiko, Ljubljana

FAKULTETA ZA ELEKTROTEHNIKO - LMFE, Ljubljana

FAKULTETA ZA ELEKTROTEHNIKO - LMSE, Ljubljana

FAKULTETA ZA ELEKTROTEHNIKO - LPVO, Ljubljana

MIKROIKS, d.o.o., Ljubljana

TC SEMTO, Ljubljana

ISKRA ZAŠČITE d.o.o., Ljubljana

RLS, d.o.o., Ljubljana

Informacije MIDEM, 38(2008)1, Ljubljana, MAREC 2008, ISSN 0352-9045, UDK 621.3:(53+54+621+66)(05)(497.1)=00

Strokovna revija za mikroelektroniko, elektronske sestavne dele in materiale

Journal of Microelectronics, Electronic Components and Materials

Publikacija informacije MIDEM izhaja po ustanovitvi Strokovnega društva za mikroelektroniko, elektronske sestavne dele in materiale - MIDEM kot nova oblika publikacije Informacije SSOSD, ki jo je izdajal Zvezni strokovni odbor za elektronske sestavne dele in materiale SSOSD pri Jugoslovanski zvezi za ETAN od avgusta 1969 do 6. oktobra 1977 in publikacije Informacije SSESD, ki jo je izdajala Strokovna sekcija za elektronske sestavne dele, mikroelektroniko in materiale - SSESD pri Jugoslovanski zvezi za ETAN od 6. oktobra 1977 do 29. januarja 1986.

Časopis Informacije MIDEM je vpisan v register časopisov pri Ministrstvu za informiranje pod registrsko številko 809.