

**BI**  **ICT**

**IT'17**  
ŽABLJAK

**XXII**

*međunarodni naučno - stručni skup*

**INFORMACIONE  
TEHNOLOGIJE**

*SADAŠNJOST I BUDUĆNOST*

Urednik  
Božo Krstajić

**IT'17**

**INFORMACIONE  
TEHNOLOGIJE**

**- SADAŠNJOST I BUDUĆNOST -**

**Urednik**  
*Božo Krstajić*

*Zbornik radova sa XXII međunarodnog naučno - stručnog skupa  
INFORMACIONE TEHNOLOGIJE - sadašnjost i budućnost  
održanog na Žabljaku od 27. februara do 04. marta 2017. godine*

Zbornik radova  
INFORMACIONE TEHNOLOGIJE - sadašnjost i budućnost 2017

*Glavni urednik*  
Prof. dr Božo Krstajić  
Elektrotehnički fakultet Univerziteta Crne Gore

*Izdavač*  
Univerzitet Crne Gore  
Elektrotehnički fakultet  
Džordža Vašingtona bb., Podgorica  
[www.etf.ucg.ac.me](http://www.etf.ucg.ac.me)

*Tehnička obrada*  
Aleksandra Radulović  
Centar informacionog sistema Univerziteta Crne Gore

Doc. dr Žarko Zečević  
Elektrotehnički fakultet Univerziteta Crne Gore

Ivan Jokić  
Elektrotehnički fakultet Univerziteta Crne Gore

*Tiraž*  
150

Podgorica 2017.

*Sva prava zadržava izdavač i autori*

## **Organizator**

Elektrotehnički fakultet, Univerzitet Crne Gore

## **Suorganizator**

BIO-ICT Centar izvrsnosti

## **Skup podržao**

Institute of Electrical and Electronics Engineers - IEEE

## Programski odbor

Dr Novak Jauković, Elektrotehnički fakultet, UCG, Podgorica, MNE  
Dr Ljubiša Stanković, Elektrotehnički fakultet, UCG, Podgorica, MNE  
Dr Zdravko Uskoković, Elektrotehnički fakultet, UCG, Podgorica, MNE  
Dr Vujica Lazović, Ekonomski fakultet, UCG, Podgorica, MNE  
Dr Branko Kovačević, Elektrotehnički fakultet, UB, Beograd, SRB  
Dr Milorad Božić, Elektrotehnički fakultet, UBL, Banja Luka, RS BIH  
Dr Miroslav Bojović, Elektrotehnički fakultet, UB, Beograd, SRB  
Dr Zoran Jovanović, Elektrotehnički fakultet, UB, Beograd, SRB  
Dr Božidar Krstajić, Elektrotehnički fakultet, UIS, I. Sarajevo, BIH  
Dr Milica Pejanović-Đurišić, Elektrotehnički fakultet, UCG, Podgorica, MNE  
Dr Despina Anastasiadou, Research & Development Innovation Academy, Solun, GRC  
Dr Dejan Popović, Elektrotehnički fakultet, UB, Beograd, SRB  
Dr Gabriel Neagu, National Institute for Research & Development in Informatics, Bucharest, ROU  
Dr Božo Krstajić, Elektrotehnički fakultet, UCG, Podgorica, MNE  
Dr Tomo Popović, Fakultet za informacione sisteme i tehnologije, UDG, Podgorica, MNE  
Dr Milovan Radulović, Elektrotehnički fakultet, UCG, Podgorica, MNE  
Dr Le Xie, Texas A&M University, College Station, TX, USA  
Dr Sašo Gelev, Elektrotehnički fakultet, UGD, Radoviš, MKD  
Dr Budimir Lutovac, Elektrotehnički fakultet, UCG, Podgorica, MNE  
Dr Zoran Veljović, Elektrotehnički fakultet, UCG, Podgorica, MNE  
Dr Igor Radusinović, Elektrotehnički fakultet, UCG, Podgorica, MNE  
Dr Alex Sprintson, Texas A&M University, College Station, TX, USA  
Dr Nikša Tadić, Elektrotehnički fakultet, UCG, Podgorica, MNE  
Dr Miloš Daković, Elektrotehnički fakultet, UCG, Podgorica, MNE  
Dr Milutin Radonjić, Elektrotehnički fakultet, UCG, Podgorica, MNE  
Dr Ana Jovanović, Elektrotehnički fakultet, UCG, Podgorica, MNE  
Dr Vesna Rubežić, Elektrotehnički fakultet, UCG, Podgorica, MNE  
Dr Biljana Stamatović, Fakultet za informacione sisteme i tehnologije, UDG, Podgorica, MNE  
Dr Ramo Šendelj, Fakultet za informacione sisteme i tehnologije, UDG, Podgorica, MNE  
Dr Stevan Šćepanović, Prirodno-matematički fakultet, UCG, Podgorica, MNE  
Dr Goran Šuković, Prirodno-matematički fakultet, UCG, Podgorica, MNE

## Organizacioni odbor

Dr Božo Krstajić, Elektrotehnički fakultet, UCG, Podgorica, MNE - Predsjedavajući konferencije

Dr Milovan Radulović, Elektrotehnički fakultet, UCG, Podgorica, MNE - Finansijski rukovodilac konferencije

Dr Zoran Veljović, Elektrotehnički fakultet, UCG, Podgorica, MNE

Dr Ana Jovanović, Elektrotehnički fakultet, UCG, Podgorica, MNE

Dr Saša Mujović, Elektrotehnički fakultet, UCG, Podgorica, MNE

Dr Tomo Popović, Fakultet za informacione sisteme i tehnologije, UDG, Podgorica, MNE - Rukovodilac izdavačke djelatnosti

Dr Žarko Zečević, Elektrotehnički fakultet, UCG, Podgorica, MNE - Tehnički rukovodilac konferencije

Vladan Tabaš, dipl.ing., Čikom, Podgorica, MNE

Aleksandra Radulović, CIS UCG, Podgorica, MNE - Kontakt za informacije

Ivan Jokić, Elektrotehnički fakultet, UCG, Podgorica, MNE - Kontakt za informacije

Amar Kapić, Elektrotehnički fakultet, UCG, Podgorica, MNE - Kontakt za informacije

# P R E D G O V O R

Poštovani učesnici i čitaoci,

Iza nas je jedna od uspješnijih konferencija “INFORMACIONE TEHNOLOGIJE – sadašnjost i budućnost” (IT’17) održana od 27. februara do 04. marta 2017. godine na Žabljaku, a ovo je zbornik radova. Programski odbor i recenzentski tim je izvršio selekciju kvalitetnih radova koji su prezentovani na konferenciji i publikovani ovdje, a najbolji će biti prošireni i objavljeni u časopisu Elektrotehničkog fakulteta Univerziteta Crne Gore u Podgorici ("ETF Journal of Electrical Engineering"). Zahvaljujući vama, dragi autori, iz godine u godinu podižemo kvalitet radova koji su sadržani u ovom zborniku, a prvi put smo sponzorisni i od strane IEEE asocijacije.

Pred vama su 72 autorska rada, prezentovana u 11 sekcija, i sažeci 5 predavanja po pozivu koja su realizovana u plenarnom dijelu programa. Radovi, teme i predavači su birani po kvalitetu kao i tematici sadržaja kako bi konferencija razmatrala aktuelne momente u razvoju ICT-a, a šira stručna javnost i svi zainteresovani su bili u prilici da prate dešavanja na Skupu preko videolinka.

Organizator je registrovao oko 170 učesnika širom regiona i Evrope, 100 studenata sa sva 3 crnogorska univerziteta i veliki broj prisutnih i online posjetilaca (po našoj procjeni oko 100). Konferencija je bila dobro medijski ispraćena i podržana od državnih institucija, lokalne samouprave i velikog broja ICT kompanija.

Konačno, i Durmitor je svojim čarima i vremenskim uslovima omogućio svim učesnicima da u potpunosti uživaju u njegovim ljepotama i pružio nezaboravne utiske prirode ovog kraja. Ono što se moglo osjetiti je jedan zajednički duh i sinergija koje su nas vezivali svih 5 dana i nadahnuli novom snagom i voljom za radom i budućim druženjima. Organizator se nada da svi učesnici dijele naše mišljenje da je ovo jedan od najboljih (ako ne najbolji) IT u posljednjih 10 godina. To nas obavezuje da sljedeći IT bude bar kao ovaj, uz dodatni kvalitet, nove teme, nove ljude, nove akcije i isti duh konferencije.

Sve detalje o ovom, prošlim i narednom skupu možete naći na web adresi konferencije [www.it.ac.me](http://www.it.ac.me).

Prof. dr Božo Krstajić

# SADRŽAJ

## PREDAVANJA PO POZIVU

Др Сениша Влајић Развој софтверских компоненти заснован на генератору кода и софтверским патернима The Development of Software Components Based on the Code Generator and Software Patterns.....	1
Др Душан Савић Развој софтвера заснован на моделу случајева коришћења и MDD приступу Model and Use Case Driven Software Development Process .....	2
Др Илија Антовић Аутоматско генерисање корисничког интерфејса засновано на случајевима коришћења – SilabUI приступ Automatic User Interface Generation Based on Use Cases – The SilabUI approach.....	3
Dr Saša D. Lazarević Razvoj C# Cross-Platform aplikacija upotrebom .NET Core 1.0 C# Cross-Platform Applications' Development using .NET Core 1.0 .....	4
Tomislav B. Šekara SAVREMENE METODE PROJEKTOVANJA INDUSTRIJSKIH SISTEMA AUTOMATSKOG UPRAVLJANJA MODERN DESIGN METHODS OF INDUSTRIAL AUTOMATIC CONTROL SYSTEMS .....	5

## AUTORSKI RADOVI

Amar Kapić, Žarko Zečević, Milovan Radulović and Božo Krstajić A VARIABLE STEP SIZE PERTURB AND OBSERVE ALGORITHM FOR MAXIMUM POWER POINT TRACKING.....	6
Sreten Ćorić PRIMJER IMPLEMENTACIJE PRIVATNOG CLOUD RJEŠENJA EXAMPLE OF THE IMPLEMENTATION OF PRIVATE CLOUD SOLUTION .....	10
Milan Zejak, Milovan Radulović HIBRIDNI KONCEPT SMART HOME SISTEMA AN EXAMPLE OF A HYBRID SMART HOME SYSTEM .....	14
Igor Miljanić CA SISTEM KABLOVSKIH TELEVIZIJA – MEHANIZAM KONTROLE I MONITORING CA SYSTEM WITHIN CABLE OPERATED PROVIDERS - A MECHANISM OF CONTROL AND MONITORING .....	18



Vladimir Radulović, Milovan Radulović LED RASVJETA TUNELA PRIMJER TUNEL BUDOŠ NA PUTU PODGORICA NIKŠIĆ TUNNEL LIGHTING WITH LED EXAMPLE OF ROAD TUNNEL BUDOŠ ALONG ROAD PODGORICA-NIKŠIĆ.....	22
Ivan Jokić, Žarko Zečević, Zdravko Uskoković, Milovan Radulović, Božo Krstajić A NEW METHOD FOR SYNCHROPHASOR ESTIMATION.....	26
Ivan R. Bulatović, Vladan M. Radulović DEVELOPMENT OF COMBINATION WAVE GENERATORS FOR SIMULATION OF A LIGHTNING SURGES IN LOW-VOLTAGE AC POWER CIRCUITS .....	30
Milutin Ostojić, Martin Čalasan UTICAJ PLANIRANOG 400KV DALEKOVODA ČEVO - PLJEVLJA, NA DIONICI KROZ OPŠTINU ŽABLJAK, SA ASPEKTA ELEKTRIČNOG I MAGNETNOG POLJA IMPACT OF 400KV TRANSMISSION ČEVO-PLJEVLJA, AT SECTION THROUGH THE MUNICIPALITY OF ZABLJAK, FROM THE ASPECT OF ELECTRIC AND MAGNETIC FIELDS ...	34
Milan Kočović, Zoran Miljanić, Ilija Vujošević USE OF DYNAMIC PROGRAMMING FOR SHORT-TERM HYDRO-THERMAL COORDINATION .....	38
Mladen Mitkovski, Vlatko Cingoski, Sašo Gelev PERFORMANCE MODELLING OF PV ARRAYS UNDER SPECIFIC WORKING CONDITIONS .....	42
Boris Miljanić, Martin P. Čalasan, Vladan Vujičić DETERMINATION OF INDUCTION MACHINE PARAMETERS BY USING PARTICLE SWARM OPTIMIZATION WITHOUT CONSIDERING IRON LOSSES .....	46
Aleksandar Dedić JEDAN PRIMJER KONTROLE INVERTORA NAPONA KORIŠĆENJEM MIKROKONTROLERA ONE EXAMPLE OF VOLTAGE INVERTER CONTROL USING A MICROCONTROLLER .....	50
Jelena Končar, Sonja Leković, Zita Petrović Katai E-GOVERNMENT AS A BUSINESS MODEL OF ECOMMERCE IN THE REPUBLIC OF SERBIA .....	54
Tatjana Stanivuk IMPLEMENTACIJA E-UČENJA U POMORSTVU IMPLEMENTATION OF E-LEARNING IN MARITIME AFFAIRS.....	58
Marina Matijević, Ramo Šendelj UNAPREĐENJE ICT INFRASTRUKTURE OBRAZOVNOG SISTEMA CRNE GORE PRIMJENOM CLOUD COMPUTING-A USING CLOUD COMPUTING IN THE EDUCATION SYSTEM OF MONTENEGRO.....	62
Marina Matijević, Ivana Ognjanović ZAŠTITA PODATAKA U OBRAZOVNOM SISTEMU CRNE GORE UPOTREBOM OPENPGP STANDARDA DATA PROTECTION IN THE EDUCATIONAL SYSTEM OF MONTENEGRO USING OPENPGP STANDARD .....	66

Nataša Popović, Aleksandra Radulović, Isidora Rovčanin Dragović, Oleg Cmiljanić, Sehija Dizdarević IMPACT OF WEB - BASED LEARNING MENAGEMENT SYSTEMS ON EDUCATION AT THE FACULTY OF MEDICINE IN PODGORICA, MONTENEGRO .....	70
Ana Grbović ICS IN THE HPP PERUCICA AS A NATIONAL CRITICAL INFRASTRUCTURE.....	74
Vladimir Bigović, Sandra Đurović, Snežana Šćepanović TESTIRANJE UPOTREBLJIVOSTI GRAFIČKOG KORISNIČKOG INTERFEJSA OPERATIVNIH SISTEMA ZA MOBILNE UREĐAJE USABILITY TESTING OF GRAPHICAL USER INTERFACE FOR MOBILE OPERATING SYSTEMS.....	78
Sanja Bauk, Tatijana Dlabac O NEKIM ICT RJEŠENJIMA ZA POVEĆANJE BEZBJEDNOSTI NA RADU U INVAZIVNIM OKRUŽENJIMA ON SOME ICT SOLUTIONS FOR INCREASING SAFETY AT WORK IN INVASIVE ENVIRONMENTS .....	82
Stanko Janković, Biljana Stamatović AUTOMATED CLASSIFICATION OF WEB APPLICATION PARTS FROM A BLACKBOX SECURITY TESTING PERSPECTIVE .....	86
Marko Č. Bošković, Tomislav B. Šekara, Milan R. Rapaić, Vidan Govedarica TUNING RULES OF PI/PID CONTROLLERS WITH ROBUSTNESS CONSTRAINT FOR INDUSTRIAL PROCESSES WITHOUT DEAD - TIME.....	90
Ivica Vuković SISTEMI ZA NAPAJANJE I HLAĐENJE GREEN DATA CENTRA GREEN DATA CENTER POWER AND COOLING SISTEMS.....	94
Snežana Šćepanović, Ivan Mraković BARRIERS TO ADOPTION AND SUCCESS OF SCRUM IN PRACTICE.....	98
Bojana Tošić, Dragan Đurić, Radovan Sekulić, Jelena Simićević CLOUD COMPUTING SOLUTIONS FOR HEALTHCARE SYSTEM IN ME: CHALLENGES AND ISSUES .....	102
Jurica A. Žarković OSVRT NA KONTINUIRANI MONITORING GLIKEMIJE KOD PACIJENATA SA DIJABETESOM REVIEW OF THE METHODS OF CONTINUOUS MONITORING OF BLOOD GLUCOSE IN PATIENTS WITH DIABETES .....	106
Jelena Badnjar, Nataša Popović, Miroslav Radunović, Tomo Popović RETINAL BLOOD VESSEL SEGMENTATION: DIJKSTRA FOREST AND MULTI – SCALE LINE DETECTION METHODS.....	110
Milija Bajčeta, Petar Sekulić, Slobodan Djukanović RETINAL BLOOD VESSELS SEGMENTATION USING NEURAL NETWORKS AND FEATURES BASED ON LINE OPERATORS .....	114

Salko Kovačić, Emina Đulić, Admir Šehidić IMPROVING THE SECURITY OF ACCESS TO NETWORK RESOURCES USING THE 802.1X STANDARD IN WIRED AND WIRELESS ENVIRONMENTS .....	118
Armin Lucević, Lejla Taraniš THE USE OF INFORMATION TECHNOLOGY IN THE HEALTHCARE SYSTEMS, CASE OF SERBIA .....	122
Arsenije Maliković, Dragan Filipović, Budimir Lutovac PROJEKTOVANJE SPREGNUTIH MIKROTRAKASTIH VODOVA DESIGN OF COUPLED MICROSTRIP LINES .....	126
Dragan Filipović, Vladan Durković POTENCIJAL, ELEKTRIČNO POLJE I PODUŽNA KAPACITIVNOST TANKOG LINIJSKOG PROVODNIKA U UGLU IZMEĐU DVIJE ZEMLJENE PROVODNE POLURAVNI POTENTIAL, ELECTRIC FIELD AND CAPACITANCE OF A THIN LINE CONDUCTOR IN AN ANGLE BETWEEN TWO GROUND HALFPLANES .....	130
Teodora S. Veljović, Boban P. Bondžulić, Ljubiša D. Tomić, Bojan Č. Milanović, Rade M. Pavlović ODREĐIVANJE ŠIRINE I DIVERGENCIJE LASERSKOG SNOPA LASER BEAM WIDTH AND DIVERGENCE MEASURING .....	133
Luka Lazović, Vesna Rubežić, Ana Jovanović CPW NAPAJANA MIKROTRAKASTA MONOPOL ANTENNA ZASNOVANA NA MODIFIKOVANOM SIERPINSKI FRAKTALU CPW FEEDED MICROSTRIP MONOPOL ANTENNA BASED ON MODIFIED SIERPINSKI FRACTAL .....	137
Ana Perošević, Ivana Milašević, Danijela Joksimović, Dragana Drakulović, Dijana Đurović, Slavka Stanković PHYSICAL-CHEMICAL PARAMETERS AND PHYTOPLANKTON IN BOKA KOTORSKA BAY .....	141
Mandić Milica, Gvozdenović Slađana, Joksimović Aleksandar, Pešić Ana CONDITION OF CULTURED SEA BASS AND SEA BREAM AND OCCURRENCE OF THEIR SPAWNING IN THE CAGES IN BOKA KOTORSKA BAY .....	145
Marko Nikolić, Slađana Gvozdenović, Sergey Kholodkevich, Tatiana Kuznetsova, Anton Kurakin VALVE MOVEMENTS IN BIVALVES AS A BEHAVIORAL BIOMARKER .....	149
Nikola Bulatović, Nikola Žarić, Ana Pešić, Mirko Đurović, Trifun Savić, Aleksandar Joksimović REAL-TIME MEASUREMENTS ANALYSIS OF SEAWATER PHYSICAL PARAMETERS WITH SEMAR BUOY SYSTEM.....	153
Elma Hot, Balša Bajagić ANOMALY DETECTION AND PREPROCESSING OF SOIL DATABASE OF MONTENEGRO.....	157
Trifun Savić, Balša Bajagić, Mirko Knežević, Milutin Radonjić METHOD FOR MEASURING RELEASED AMOUNT OF WATER IN SMART IRRIGATION SYSTEM.....	161

Nikola Bulatović, Tanja Vukićević, Rabina Šabotić, Vanja Asanović MEASUREMENTS CALIBRATION ANALYSIS OF NOVEL IOT PLATFORMS .....	165
Gojko Gogić, Slavica Tomović, Igor Radusinović, IMPLEMENTATION AND PERFORMANCE EVALUATION OF OPENFLOW DATA-CENTRE NETWORK TESTBED .....	169
Božidar Škrbić, Slobodan Đukanović A BRIEF OVERVIEW OF IOT ARCHITECTURE AND WEB PLATFORMS.....	173
Žarko Zečević, Tomo Popović, Božo Krstajić CLOUD BASED SOLUTION FOR AUTOMATIC IMAGE MOSAICKING AND GEOREFERENCING .....	177
Nikola Žarić, Boban Melović, Miomir Jovanović ONLINE PROMOCIJA POLJOPRIVREDNIH PROIZVODA: STUDIJA SLUČAJA BIOPORTAL.ME ONLINE PROMOTION OF AGRICULTURAL PRODUCTS: CASE STUDIES BIOPORTAL.ME.....	181
Petar Sekulić, Milija Bajčeta, Slobodan Djukanović RETINAL BLOOD VESSELS SEGMENTATION USING SUPPORT VECTOR MACHINE AND MODIFIED LINE DETECTOR .....	185
Miljan Bigović, Žarko Zečević, Luka Filipović, Božo Krstajić KOMPARACIJA RAČUNARSKOG MODELA MOLEKULA SA EKSPERIMENTALNIM REZULTATIMA COMPARISON OF COMPUTER MODEL OF MOLECULE WITH EXPERMINTAL RESULTS .....	189
Bojan Đaković, Budimir Lutovac KEY PERFORMANCE INDICATORS VISUALIZATION THROUGH “BIMP” APPLICATION.....	193
Ivan Mraković KVALITATIVNA KOMPARATIVNA ANALIZA VODOPADNOG, SPIRALNOG I SCRUM MODELA RAZVOJA SOFTVERA QUALITATIVE COMPARATIVE ANALYSIS OF WATERFALL, SPIRAL AND SCRUM SOFTWARE DEVELOPMENT MODELS .....	197
Branko Džakula VISOKOTEHNOLOŠKI (SAJBER) KRIMINAL I ANATOMIJA BEZBJEDNOSNIH RIZIKA U INFORMACIONOKOMUNIKACIONIM SISTEMIMA CYBER CRIME AND ANATOMY OF SECURITY RISKS IN INFORMATION-COMMUNICATION SYSTEMS .....	201
Katerina Cekova, Jasna Cekova, Sašo Gelev HOME AUTOMATION SYSTEM CONTROL BASED ON SPEECH COMMANDS.....	205
Darko Naumovski, Sašo Gelev, Vlatko Cingoski and Biljana Petrevska MALOLETNICKA RAČUNARSKA PORNOGRAFIJA U REPUBLICICI MAKEDONIJI JUVENILE COMPUTER PORNOGRAPHY IN MACEDONIA .....	209

Vasilij N. Kruglov, Andrey V. Chiryshv, Artem V. Kruglov DEVELOPING OF THE SIMULATION SYSTEM TO MODEL THE FILLING LAYER OF SPHERICAL GRANULATED PARTICLES .....	213
Aleksandar Tudzarov, Goce Stefanov, Sašo Gelev NOVEL TECHNIQUE FOR AUTHENTICATION & ENCRYPTION IN NEXT GENERATION NETWORKS .....	217
Zoran S. Veličković, Zoran N. Milivojević, Marko Z. Veličković INSERTOVANJE VODENOG ŽIGA SKREMBLOVANOG GMSAT ALGORITMOM U DWT-SVD DOMENU INSERTING THE WATERMARK SCRAMBLED BY GMSAT ALGORITHM IN DWT-SVD DOMAIN .....	221
Rade M. Pavlović, Vladimir S. Petrović, Ljubiša D. Tomić UPOREDNA ANALIZA MERA ZA DINAMIČKU PROCENU USPEŠNOSTI SJEDINJAVANJA VIDEO SEKVENCI COMPARATIVE STUDY OF DYNAMIC MEASURES FOR VIDEO FUSION.....	225
Miloš Brajović, Miloš Daković, Ljubiša Stanković A HIGHER ORDER TIME-FREQUENCY REPRESENTATION WITH REDUCED CROSS-TERMS .....	229
Tamara Šević, Milenko Andrić, Boban Bondžulić, Dimitrije Bujaković, Dejan Ivković IZDVAJANJE OBELEŽJA RADARSKIH CILJEVA OBRADOM SPEKTROGRAMA RADAR TARGETS FEATURE EXTRACTION USING SPECTROGRAM .....	233
Zoran Milivojević, Dijana Kostić, Zoran Veličković, Darko Brodić PERFORMANCE OF THE SERBIAN MATRIX SENTENCE TEST IN PRESENCE OF THE BABBLE NOISE .....	237
Dragan S. Janković, Eleonora M. Milić, Aleksandar M. Milenković REALIZACIJA MOBILNOG MEDICINSKOG PODSETNIKA INTEGRISANOG SA MEDICINSKIM INFORMACIONIM SISTEMOM MEDIS.NET IMPLEMENTATION OF MOBILE MEDICAL APPOINTMENTS INTEGRATED WITH THE MEDICAL INFORMATION SYSTEM MEDIS.NET.....	241
Aleksa Srdanov, Selena Vasić, Nada Ratković Kovačević, Dragan Milovanović THE APPLICATION OF SIMULATED INTUITION IN MINIMIZING THE NUMBER OF MOVES IN GUESSING THE SERIES OF IMAGINED OBJECTS .....	245
Stevan Čakić, Aleksandar Popović ALAT UNITY3D U KONCEPTU PROŠIRENE REALNOSTI UNITY3D TOOL IN THE AUGMENTED REALITY CONCEPT .....	249
Lazar Radinović, Aleksandar Popović GEO-LOKACIJSKI SERVIS ZA JAVNI PREVOZ GEOLOCATION SERVICE FOR PUBLIC TRANSPORT .....	253
Kosta Pavlović, Goran Šuković DEEP LEARNING TECHNIQUES FOR CLASSIFICATION OF HANDWRITTEN DIGITS .....	257

Boris Tuponja, Goran Šuković APPLYING NEAT ON THE GAME 2048 .....	261
Aleksa Srdanov, Vladan Đulaković, Dragan Milovanović ČETIRI EKSTREMA VEZANA ZA SUDOKU FOUR EXTREMES RELATED TO SUDOKU .....	265
Aleksandar Tudzarov, Goce Stefanov, Sašo Gelev CLIENT BASED SELECTION ALGORITHM FOR HYBRID WIRELESS NETWORKS .....	269
Uglješa Urošević, Zoran Veljović RJEŠENJA ZA POBOLJŠANJE D2D KOMUNIKACIJA U 5G MREŽAMA SOLUTIONS FOR D2D IMPROVEMENTS IN 5G NETWORKS .....	273
Jelena T. Šuh, Snežana Lj. Stefanovski Pajović ZAŠTITA OD DOS NAPADA PRIMENOM REMOTELY TRIGGERED BLACK HOLE TEHNIKE FILTRIRANJA SAOBRAĆAJA PROTECTION AGAINST DOS ATTACKS USING THE REMOTELY TRIGGERED BLACK HOLE TECHNIQUE FOR TRAFFIC FILTERING .....	277
Jovana Mrkić, Enis Kočan INDEX MODULATION FOR PERFORMANCE IMPROVEMENT OF OFDM BASED SYSTEMS .....	281
Aleksandra Lopušina, Enis Kočan, Milica Pejanović - Đurišić MEETING 5G REQUIREMENTS THROUGH MILLIMETER WAVE COMMUNICATIONS .....	285
Vuk Radman, Milutin Radonjić ARDUINO-BASED SYSTEM FOR SOIL MOISTURE MEASUREMENT .....	289

**Предавач: проф. др Синиша Влајић ([vlajic@fon.bg.ac.rs](mailto:vlajic@fon.bg.ac.rs)),  
Факултет организационих наука, Универзитет у Београду**

Тема предавања

## **Развој софтверских компоненти заснован на генератору кода и софтверским патернима**

Када се развије неки сложени софтверски систем у њему се може учити пуно различитих, међусобно повезаних елемената који обезбеђују жељену функционалност програма. Неки од тих елемената могу да садрже нека општа (генеричка) својства која се могу поново користити у развоју неких других софтверских система. Издвајање таквих елемената у независне, поново употребљиве јединице програма представља процес дефинисања софтверских компоненти. Оно што представља проблем у процесу прављења софтверских компоненти јесте хетерогеност њених елемената који могу да припадају различитим технологијама. Тако се на пример у развоју сложених Web апликација, може у исто време користити Јава технологије (JSP, JSF, Servleti,...), HTML језик за опис структуре Web страница, XML језик за дефинисање формата података, Java Script језик за дефинисање функционалности Web страна, EL језик за опис израза Web страна, JSON за дефинисање формата података који се размеђују између клијента и сервера и SQL упитни језик у раду са релационим системима за управљање базама података (MySQL, SQL server, Oracle, ...). Поставља се питање, како од једног оваквог сложеног софтверског система можемо направити софтверску компоненту која ће моћи поново да се користи у изградњи нових софтверских система. Одговор на ово питање представља тему овог предавања. У њему ћемо изложити оригинални приступ (СИЛАБ-Луча) који се развија у оквиру Катедре за софтверско инжењерство на Факултету организационих наука у Београду, који прати процес развоја софтверске компоненте коришћењем сопственог генератора програмског кода и скупа софтверских патерна који дефинишу макро и микро архитектуру софтверске компоненте.

Topic of the lecture

## **The Development of Software Components Based on the Code Generator and Software Patterns**

When a complex software system is developed, there are a lot of different, interconnected elements that provide the desired functionality of the program. Some of these elements may contain some general (generic) properties that can be reused in developing other software systems. Extraction of these elements in an independent, reusable units of the program is the process of defining software components. The problem in the process of creating software components is related to heterogeneity of its elements that may belong to different technologies. For example, in the development of complex Web applications we can simultaneously use Java technologies (JSP, JSF, Servlets, ...), HTML language for description of the structure of Web pages, XML language for defining data formats, Java Script language for defining functionalities of Web pages, the EL language for description of expressions of Web pages, JSON for defining formats of the data exchanged between client and server and the SQL query language to work with relational database management systems (MySQL, SQL server, Oracle, ...). The question is, how from such a complex software system we can make a software component that will be used again in the construction of new software systems. The answer to this question is the theme of this lecture. We will expose the original approach (SILAB-Luca), which is being developed within the Department of Software Engineering at the Faculty of Organizational Sciences in Belgrade, which follows the process of development of the software components using our own program code generator and a set of software patterns that defines the macro and micro architecture of software components.

Предавач: др Душан Савић ([dules@fon.bg.ac.rs](mailto:dules@fon.bg.ac.rs)),  
Факултет организационих наука, Универзитет у Београду

Тема предавања

## Развој софтвера заснован на моделу случајева коришћења и MDD приступа

Моделом вођени развој (Model Driven Development, MDD) софтвера подразумева коришћења модела као примарног артефакта током свих фаза развоја софтвера. Већина MDD приступа је вођена анализом (analyst-driven), што значи да се иницијални MDD модел креира као резултат предходно одрађене фазе прикупљања захтева и анализе захтева. Најчешће се овај иницијални MDD модел креира ручно. Један од главних питања у вези са MDD јесте како направити аутоматски иницијални MDD модел.

У оквиру Лабораторије за софтверско инжењерство, Факултета организационих наука почетком 2007. године инициран је и покренут Silab пројекат. Основни циљ овог пројекта је био да се стандардизује, униформише и формализује начин спецификације корисничких захтева на основу који би се омогућило аутоматско генерисање различитих компоненти софтверског система. SilabMDD приступ настао је као кључни резултат Silab пројекта. Први пут овај приступ је коришћен за реализацију Kostmod 4.0 пројекта за потребе Министарства одбране Краљевине Норвешке.

Као део SilabMDD приступа предложена је Silab-UCMDDM метода. Silab-UCMDDM метода наглашава значај и неопходност коришћења 3 међусобно конзистентна и компламентарна модела: а) модел случајева коришћења, б) доменски модел и с) модел прелаза стања. Дефинисане је веза између ова три модела која се огледа првенствено у чињеници да акције случајева коришћења треба да се специфицирају у складу са доменским моделом, а да предуслови и постуслови за извршење случајева коришћења треба да буду дефинисани у складу са моделом прелаза стања. Спецификација захтева у оквиру Silab-UCMDDM методе омогућена је преко посебно развијеног доменски специфичног UCDSL језика. Предложени UCDSL језик је имплементиран преко JetBrains MPS алата за мета-програмирање (Meta Programming System, MPS). Овај језик је интегрисан у оквиру SILAB-MDDTOOLSET алата који се може користити како додатак (plug-in) унутар алата као што су: MPS и IntelliJ IDEA.

Topic of the lecture

## Model and Use Case Driven Software Development Process

Model driven development (MDD) refers to the systematic use of models as primary artifacts throughout the software development lifecycle. The most of MDD approach are analyst-driven that means that the initial MDD model is derived replacing existing analyst model. Therefore, analyst model as a result of Requirements Engineering or Enterprise Modeling process is usually manually analyzed to create the initial MDD model. One of the major issues associated with MDD is how to automatically create the initial MDD model.

SilabMDD approach emerged as a key result of Silab Project which was initiated in 2007 in the Software Engineering Laboratory at Faculty of Organizational Sciences, University of Belgrade. The main goal of this project was to enable automated analysis and processing of software requirements in order to achieve automatic generation of different parts of a software system. As a proof of concept, Silab project was used for the Kostmod 4.0 project, which was implemented for the needs of the Royal Norwegian Ministry of Defense. After several years of using this project in developing different intensive software system we are established a SilabMDD approach.

As a part of SilabMDD approach, the Silab-UCMDDM method is proposed. Silab-UCMDDM emphasizes the importance and necessity of using 3 mutually consistent and complementary models: a) the use case model, b) the domain model and c) the state machine model. Direct link among these three models are proposed, which are primarily reflected in the fact that the specification of use case actions should be based on the domain model, while the preconditions and post-conditions for executing use cases should be defined in the state machine model. This model state machine model is used for a clear and precise definition of use cases. Requirements specification within Silab-UCMDDM method is enabled via its own domain specific language (UCDSL). The proposed UCDSL language was performed using the JetBrains MPS tool for meta-programming (Meta-Programming System JetBrains MPS). UCDSL language is integrated within the SILAB-MDDTOOLSET tool that can be used as an add-on (plug-in) for environments such as MPS and IntelliJ IDEA.



Предавач: др Илија Антовић ([ilijaa@fon.bg.ac.rs](mailto:ilijaa@fon.bg.ac.rs)),  
Факултет организационих наука, Универзитет у Београду

Тема предавања

## **Аутоматско генерисање корисничког интерфејса засновано на случајевима коришћења - SilabUI приступ**

Пројектовање и имплементација корисничког интерфејса су активности које изискују највише напора и времена у процесу развоја софтверског система. Предавање ће бити усмјерено на аутоматизацију процеса пројектовања и имплементације корисничког интерфејса. Аутоматизација је заснована на уоченим везама које постоје између софтверских захтева, прије свега специфицираних коришћењем случајева коришћења, и резултујућег корисничког интерфејса, узимајући у обзир карактеристике циљних имплементационих технологија и типова апликација. Биће разматрани принципи и карактеристике које треба да буду саставни дио алата за аутоматизацију развоја корисничког интерфејса, а који су имплементирани у SilabUI приступу.

Topic of the lecture

## **Automatic User Interface Generation Based on Use Cases – The SilabUI approach**

The design and implementation of user interfaces are two of the most effort and time consuming activities of the software development process. The lecture will address the automation of the process of designing and implementing user interfaces. The automation is based on the identified connections between software requirements, particularly use cases, and the resulting user interface of the application, taking into account the characteristics of the targeted implementation technologies and application types. We will also discuss the principles and characteristics which should be an integral part of the tool for automation of user interfaces development, that are all implemented in SilabUI approach.

Predavač: **Dr Saša D. Lazarević** ([slazar@fon.bg.ac.rs](mailto:slazar@fon.bg.ac.rs)),  
Fakultet organizacionih nauka, Univerzitet u Beogradu

Tema predavanja

## **Razvoj C# Cross-Platform aplikacija upotrebom .NET Core 1.0**

Ovaj rad govori o tome kako možemo razviti C# cross-platform aplikacije pomoću .NET Core 1.0, ASP.NET Core 1.0, Entity Framework Core 1.0 i Visual Studio Code. Razvoj cross-platform aplikacija znači biti u stanju kako pisati kôda tako i pokretanje istog na operativnim sistemima drugačijim od Windows-a, odnosno kako napraviti aplikacije na .NET Core koji se može cross-platform hostovati na Windows-u, Linux-u, Docker kontejner-u i Mac OS X-u. U ovom radu pokrićemo sledeće teme:

- Razumijevanje razvoja cross-platform aplikacija
  - Alati i okviri
  - Instaliranje Visual Studio Code i ASP.NET Core
- Razvoj cross-platform web aplikacija pomoću ASP.NET Core
  - Projektovanje aplikacije
  - Izgradnja ASP.NET Core web aplikacije
  - Pokretanje ASP.NET Core web aplikacije
  - Hostovanje ASP.NET Core web aplikacija u Microsoft Azure
- Razumijevanje .NET Core command-line alata
  - Razumijevanje DNX alata
  - Razumijevanje CLI alata

Topic of the lecture

## **C# Cross-Platform Applications' Development using .NET Core 1.0**

This paper is about how we can develop C# cross-platform applications using .NET Core 1.0, ASP.NET Core 1.0, Entity Framework Core 1.0, and Visual Studio Code. Cross-platform development means being able to both write code and run the results on operating systems other than Windows, i.e. how to build an .NET Core application that can be hosted cross-platform on Windows, Linux, a Docker container, and Mac OS X. In this paper, we will cover the following topics:

- Understanding cross-platform apps development
  - Tools and Frameworks
  - Installing Visual Studio Code and ASP.NET Core
- Development of cross-platform web applications using ASP.NET Core
  - Designing the application
  - Building the ASP.NET Core web application
  - Running the ASP.NET Core web application
  - Hosting ASP.NET Core web app in Microsoft Azure
- Understanding .NET Core command-line tools
  - Understanding the DNX tools
  - Understanding the CLI tools

Predavač: **Tomislav B. Šekara** ([tomi@etf.rs](mailto:tomi@etf.rs)),  
Elektrotehnički fakultet Univerziteta u Beogradu, Srbija

Tema predavanja

## **SAVREMENE METODE PROJEKTOVANJA INDUSTRIJSKIH SISTEMA AUTOMATSKOG UPRAVLJANJA**

U okviru ovog predavanja biće predstavljene savremene metode projektovanja industrijskih sistema automatskog upravljanja. Posebno će biti istaknute metode optimizacije (podešavanja) parametara regulatora pod ograničenjima na robusnost i osjetljivost na mjerni šum i odgovarajuće strukture sistema automatskog upravljanja za tako projektovane regulatore. Biće dat akcenat na PI/PID regulatore zbog velike primjene u industriji koja prelazi 93% u odnosu na sve ostale regulatore prema Honeywell's ispitivanjima. Osim optimizacionih metoda biće prikazane i analitičke metode projektovanja regulatora. Biće pokazano da je predložena metodologija primjenljiva na široku klasu industrijskih procesa čije se ponašanje može opisati linearnim matematičkim modelima sa konačno i beskonačno stepeni slobode, uključujući tu konvencionalne sisteme konačne dimenzije, sisteme sa kašnjenjem, sisteme u čijem ponašaju dominiraju talasni i transportni problemi, kao što su problemi prenosa mase i energije, sistemi koji se opisuju frakcionim diferencijalnim jednačinama. Na kraju će biti prikazane metode karakterizacije industrijskih procesa primjenom kritičnog eksperimenta (fazno zaključana petlje i relejne tehnike) u cilju formiranja jednog opšteg modela procesa koji će nam poslužiti kao osnova za adekvatno projektovanje regulatora široke klase industrijskih procesa prethodno navedenom metodologijom.

Topic of the lecture

## **MODERN DESIGN METHODS OF INDUSTRIAL AUTOMATIC CONTROL SYSTEMS**

In this lecture modern design methods of industrial automatic control systems will be presented. The optimization (tuning) methods of controller parameters under constraints on robustness and sensitivity to the measurement noise will be especially highlighted as well as the appropriate control structures of the automatic control systems for such designed controllers. The focus will be on PI/PID controllers due to their large application in industry which exceeds 93% in comparison with all the other controllers according to Honeywell's surveys. In addition to these optimization methods, different controller analytical design methods will be presented. It will be shown that suggested methods are applicable to the wide class of industrial processes whose behavior can be described with linear mathematical models with finite and infinite degrees of freedom, including conventional finite dimension systems, time-delayed systems, systems whose behavior is dominated by a wave and transport problems such as mass and energy transfer problems, systems which are described with fractional differential equations. At the end, different methods for characterization of industrial processes applying critical experiment (phase locked loops - PLL and relay techniques) will be shown. The idea behind these characterization methods is to form a general process model which will be used to design adequate controllers for a wide class of industrial processes using above mentioned control methods.

# A Variable Step Size Perturb and Observe Algorithm for Maximum Power Point Tracking

Amar Kapić, Žarko Zečević, Milovan Radulović and Božo Krstajić

**Abstract** — In recent years many maximum power point tracking (MPPT) algorithms were developed. The algorithms variegate in parameter required, initial parameter requirement, convergence speed, cost, implementation complexity, sensitivity, ability to track true maxima, and in other respects. MPPT algorithm is implemented in energy converter through which photovoltaic system connects to grid or to the consumer. To reduce steady state error, in this paper a MPPT algorithm with variable step size is proposed. Simulations results show that proposed algorithm exhibits better performance compared to the considered ones.

**Keywords** — Forgetting factor, Incremental Conductance (InCond), Maximum Power Point Tracking (MPPT), Perturb and Observe (P&O), Photovoltaic (PV) module.

## I. INTRODUCTION

THE efficiency of solar cell is very low and it amounts 20-25% which is equivalent to the percent of solar energy which turns into electrical energy, [1]. To extract as much as possible power from the solar panel, it is necessary to bring operating point into maximal power position. Consequently, large number of algorithms are implemented in maximum power point tracking. MPPT algorithms can be classified into two categories: conventional and soft computing approach, [2]. Conventional algorithms are most used due to their simple implementation, computational complexity and low cost. As the most significant representatives stand out perturb and observe (P&O) [3], [4] and incremental conductance (InCond) [5]. Also, in the mentioned group fractional open circuit voltage, fraction short circuit current, current sweep and many others are classified. Soft computing methods are well adopted for handling nonlinearity and they are better in exhibiting steady state performance. Fuzzy logic, neural network, differential evolution, particle swarm optimization (PSO), pilot cells, cuckoo search and many others are classified in soft computing methods. Soft computing methods are much slower in practice, they demand additional equipment and they are not economically acceptable. All of that gives advantage to conventional methods. Overview of various MPPT techniques are discussed in [6] and [7].

P&O algorithm decreases or increases duty cycle of converter (to perturb) by observing the change of the

operating point direction from the photovoltaic system (to observe). InCond algorithm is based on the fact that slope of the  $P$ - $V$  characteristic is equal to zero at maximum power point, positive on the left side from maximum power point and negative on its right side. By comparing power change with voltage change, between two consecutive steps, direction of the change in the next step can be determine. In these algorithms the process of increasing or decreasing the duty cycle with fixed step size is repeated periodically until the MPP is reached. A large value of step size will cause large change in duty cycle, leading to large fluctuations in steady state. Therefore, it is necessary to find optimal step of duty cycle change, because the relatively small step causes slower reaching of the MPP which causes huge losses in energy production. Several solutions decreasing oscillations around MPP and improving the convergence speed were proposed [8], [9]. Solutions are based on variable step size usage. However, these approaches do not solve the sudden irradiation change problem. If it comes to irradiation increase, power will become high, while voltage change will stay low. Fraction becomes large so that leads to temporary, short-term confusion of algorithm in attempt of MPPT. The solution presented in [10] outperforms the previously published methods.

Papers [10]-[13], also show adaptive algorithms when sudden changes of irradiation are made, but additional sensors are required for their realization which increase implementation cost.

In this paper, in order to overcome mentioned problem modified P&O algorithm is proposed. The algorithm is using forgetting factor and knowing fraction of power change and voltage change to get variable step size in order to effectively reduce oscillations. In iteration when there is sudden change in irradiance, fixed step size gets activated to achieve faster convergence speed.

## II. MODEL OF PV MODULE

Solar cell is basically semiconducting diode (PN junction) which converts solar radiation into electrical energy by photovoltaic effect. The one diode model PV cell is mostly used in literature because of its optimal balance in complexity and precision. Equivalent scheme consists of photo current source, diode, a series resistor caused by the electrode and the parallel resistance caused by PN junction leakage current, [14].

In order to increase output power the cells are connected in series, parallel or by their combination in PV module and equivalent circuit is shown in Fig. 1.

Authors are with the Faculty of Electrical Engineering, University of Montenegro, Džordža Vašingtona bb, 81000 Podgorica, Montenegro (e-mails: [amark@ac.me](mailto:amark@ac.me), [zarkoz@ac.me](mailto:zarkoz@ac.me), [milovanr@ac.me](mailto:milovanr@ac.me), [bozok@ac.me](mailto:bozok@ac.me)).

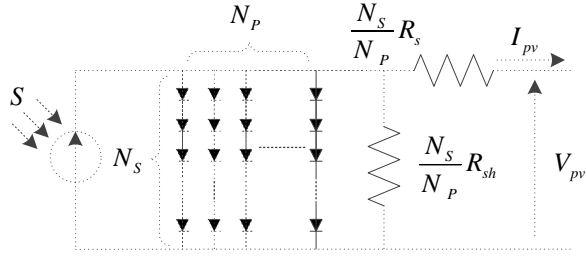


Fig. 1. Equivalent circuit of PV cell [13]

The PV module can be described by nonlinear equation shown in the form, [15]:

$$I_{pv} = N_p I_{ph} - N_p I_{rs} \left( e^{\frac{q(V_{pv}/N_s + I_{pv}R_s/N_p)}{AkT}} - 1 \right) - \frac{\frac{N_p}{N_s} V_{pv} + I_{pv} R_s}{R_{sh}}, \quad (1)$$

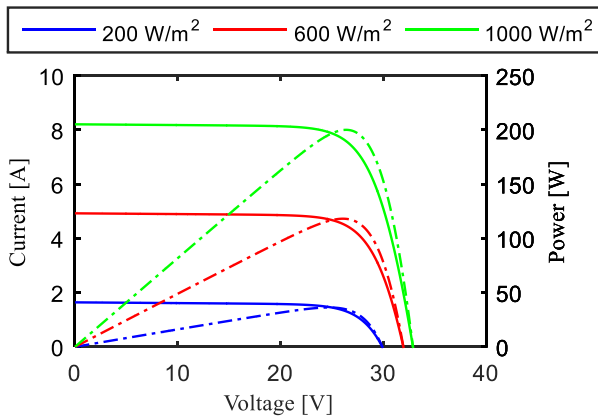
where,  $I_{pv}$ ,  $I_{ph}$ ,  $I_{rs}$ ,  $V_{pv}$ ,  $q$ ,  $A$ ,  $k$ ,  $T$ ,  $N_s$ ,  $N_p$  are output current, light generated current or photo current, diode reverse saturation current, output voltage, electron charge, diode ideality constant, Boltzmann's constant, cell absolute temperature, number of cells connected in series, number of cells connected in parallel, respectively.

The equation (1) is nonlinear and, therefore, iterative methods are used to solve it. One of them is Newton-Rapson method:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad (2)$$

where,  $x_n$  is the value of  $x$  at  $n^{th}$  instant,  $f'$  is the derivative of  $f$  with respect to  $x$ . Variable  $x$  can be either  $I_{pv}$  or  $V_{pv}$ .

Voltage and current of PV module are highly related to irradiance and temperature. The  $V$ - $I$  and  $P$ - $V$  characteristics under different irradiance levels are shown in Fig. 2.

Fig. 2.  $V$ - $I$  and  $P$ - $V$  characteristics at the constant temperature (25 °C)

The specifications for the PV module used in this paper are given in [16].

### III. MPPT ALGORITHMS

#### A. Perturb and Observe (P&O) algorithm

P&O algorithm in discrete steps changes the voltage of one-way connection between photovoltaic system and energy converter. The next step is determined by observing the sign of power change and the sign of voltage change. If the change of power is positive, the voltage is moved in the same direction. In the other words, if voltage is increased in the previous step it will be increased in the next step as well and vice versa. If the sign of power change is negative, respectively, if the power compared to the previous condition is decreased, the sign of voltage change is changed. This rules are summarized in Table 1:

$\Delta V$	$\Delta P$	Next Perturbation
+	+	+
+	-	-
-	+	-
-	-	+

#### B. Incremental Conductance (InCond) algorithm

The InCond is the second widely used MPPT algorithm. According to  $P$ - $V$  curve shown in Fig. 1, one can observe the following rules:

$$\begin{aligned} \frac{I}{V} &= -\frac{dI}{dV}, \text{ in the point of maximum power,} \\ \frac{I}{V} &> -\frac{dI}{dV}, \text{ on the left side,} \\ \frac{I}{V} &< -\frac{dI}{dV}, \text{ on the right side.} \end{aligned} \quad (3)$$

InCond operates on the principle of voltage readjusting, respectively, current. If the algorithm is based on voltage change, by increasing or decreasing the value of reference voltage, the algorithm is trying to achieve MPP. In the MPP reference voltage is equal to  $V_{MPP}$ , and it can come to its change if voltage change is no longer equal to zero and that change is caused by change of atmospheric conditions.

Conventional P&O and InCond algorithms use the fixed step size. The smaller step size, the smaller convergence speed and steady state oscillations. In order to increase convergence speed, a various MPPT algorithm with variable step size are proposed. In [2], the P&O step size is modified taking into account a number of changes in gradient. Also, proposed method is skilled to avoid drift in case of rapid change in irradiance. Variable step-size incremental-resistance MPPT algorithm is shown in [17]. Compared to the conventional InCond, this algorithm show improvement in the MPPT response speed and accuracy at steady state.

### IV. PROPOSED ALGORITHM

In order to achieve faster tracking of the MPP and reduce steady state power oscillations simultaneously, in

this paper the P&O with variable step of voltage change is proposed:

$$\Delta v_{n+1} = \lambda \cdot \Delta v_n + \gamma \cdot \left| \frac{\Delta P_n}{\Delta V_n} \right|, \quad (4)$$

where,  $\Delta v_{n+1}$  is variable step,  $\Delta v_n$  value of variable step in previous iteration,  $\Delta P_n$  power change,  $\Delta V_n$  voltage change,  $\lambda$  forgetting factor and  $\gamma$  constant. Proposed variable step size in  $n+1$  iteration is proportional to step value in the previous iteration and well-known fraction of power change and voltage change [18]. The parameter  $\lambda$  is forgetting factor and it should be chosen in the range [0-1]. The influence of the power change on step size is weighted by positive parameter  $\gamma$ .

When a sudden change in irradiance occurs, the step size is increased, since the power change has also been increased. On the other hand, the forgetting factor provides the small step values in the steady state. In this way power oscillations will be reduced. To ensure the stability, the steps size should be limited by some maximal value. Constant  $\gamma$  is determined experimentally, as in the papers [11] and [12].

Also, in algorithm has been added one more condition that serves for checking whether there has been sudden change in irradiance. If the irradiance increase or decrease more than  $100 \text{ W/m}^2$  it will come to readjusting of step size to some maximal value in order to quickly get desired state. To determine the irradiance another sensor is needed, which increases implementation cost. In order to avoid additional sensor, approximation that shows the current is proportional to irradiance is taken, [2]. Hence, if there is a change of irradiance, proportional there will be the change of current:

$$|I(k) - I(k-1)| > 0.1 \cdot I_{scn}, \quad (5)$$

where  $I(k)$  and  $I(k-1)$  are the values of current in current and previous iteration, respectively, while  $I_{scn}$  is short circuit current by irradiation of  $1000 \text{ W/m}^2$ . The flowchart of the modified P&O algorithm is presented in Fig. 4.

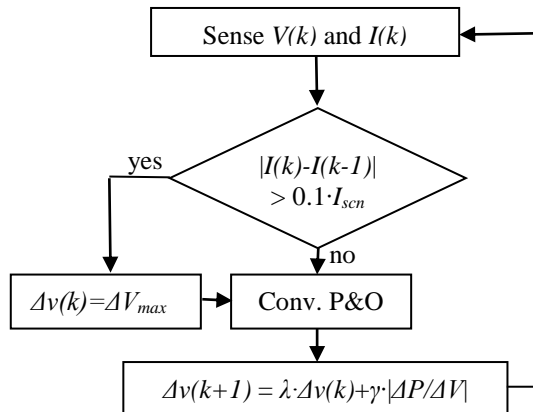


Fig. 4. Flowchart of proposed algorithm

## V. RESULTS AND DISCUSSIONS

The proposed algorithm is compared with modified P&O shown in [2] and improvement of InCond described in [16]. As a performance indicator the theoretical efficiency of algorithms is used:

$$\eta = \frac{\int P(t) dt}{\int P_{MPP}(t) dt},$$

where  $P$  represents the output power of PV panel, while  $P_{MPP}$  represents the theoretical, maximum value of the output of the panel. In this study, the comparison of algorithms is performed for a time varying irradiation, at a constant temperature of  $25 \text{ }^\circ\text{C}$ .

In the first experiment a sudden change in irradiance from  $200 \text{ W/m}^2$  to  $600 \text{ W/m}^2$  and vice versa are considered. Forgetting factor  $\lambda$  and constant  $\gamma$  are determined experimentally, and they are equal to 0.6 and 0.15, respectively. The PV output power of the considered algorithm is shown in Fig. 5. Table 2 shows their efficiency and efficiency of P&O and InCond. We see that the best results are achieved by proposed algorithm. That is as well noticeable in Fig. 6, where fluctuations of proposed algorithms are smallest around MPP. Also, with the change of irradiation algorithm has the fastest adjustment to the new condition. When it comes to the steady state almost full match with theoretical MPP is achieved.

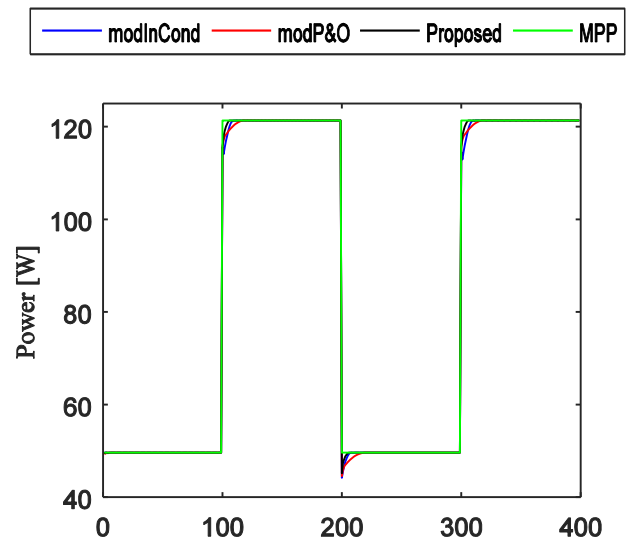
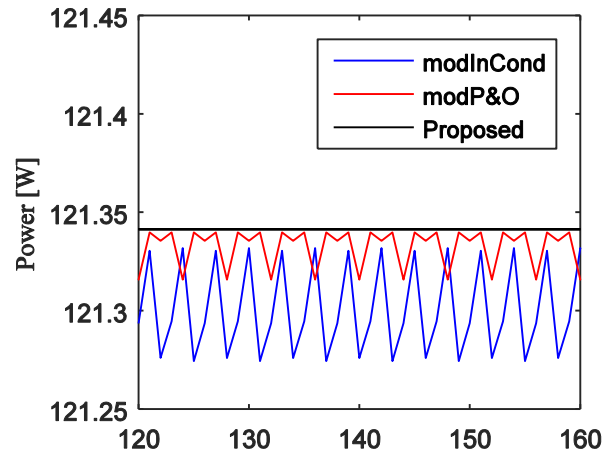


Fig. 5. PV module output power the examined algorithms



(a)  $S = 600 \text{ W/m}^2$ ,  $P_{MPP} = 121.34 \text{ W}$

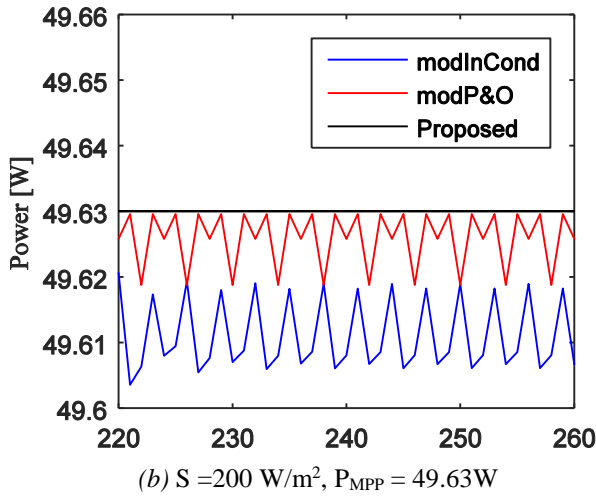


Fig. 6. PV output power in the steady state

TABLE 2: EFFICIENCY OF SIMULATED ALGORITHMS

Algorithm	$\eta$ [%]
P&O	99.28
InCond	99.28
modP&O [5]	99.73
modInCond [19]	99.72
Proposed	99.89

In the second experiment, the algorithms are tested in the case when irradiance is changed at a rate of  $1 \text{ W/m}^2/\text{s}$ . The same value of the forgetting factor as in previous experiment is used here, since the constant  $\gamma$  is set to the value 0.6. Fig. 7 shows PV output power of the considered algorithms. It can be seen that proposed algorithm has the best tracking performance. On the other hand, modP&O needs certain time to adapt to exposed changes. The efficiency of the considered algorithm is given in Table 3. Note that the proposed algorithm is the most efficient,  $\eta = 99.70\%$ .

TABLE 3: EFFICIENCY OF ALGORITHM IN SECOND EXPERIMENT

Algorithm	modInCond	modP&O	Proposed
$\eta$ [%]	99.58	99.22	99.70

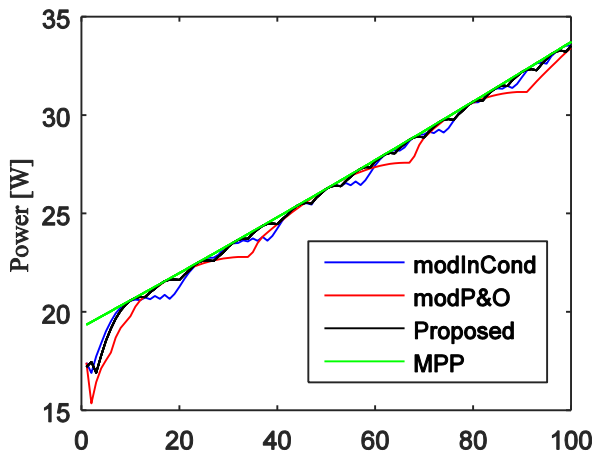


Fig. 7. Constant changes in irradiance

## VI. CONCLUSION

In this paper, a new modified P&O algorithm with variable step size is presented. Simulations results, obtained in Matlab, show that proposed algorithm have noticeable improvements compared to the existing ones. Proposed algorithm gives excellent results in case of sudden changes, as well as in the case of constant growth in irradiance. Further researches will be focused on improvement and adjustment of algorithm in case of partial shading. Also, the proposed algorithm will be implemented on real photovoltaic system.

## REFERENCES

- [1] G. M. Masters, "Renewable and Efficient Electric Power Systems", John Wiley & Sons, New Jersey, 2004
- [2] J. Ahmed, Z. Salam, "A Modified P&O Maximum Power Point Tracking Method with Reduced Steady State Oscillation and Improved Tracking Efficiency", to be published
- [3] N. Femia, G. Petrone, G. Spagnuolo, and M. Vitelli, "Optimization of Perturb and Observe Maximum Power Point Tracking Method", *IEEE Trans. Power Electron.*, vol. 20, no. 4, pp. 963-973, Jul. 2005
- [4] N. Femia, D. Granozio, G. Petrone, G. Spagnuolo, and M. Vitelli, "Predictive and adaptive mppt perturb and observe method", *IEEE Trans. Aerosp. Electron. Syst.*, vol. 43, no. 3, pp. 934-950, Jul. 2007
- [5] F. Liu, S. Duan, F. Liu, B. Liu, Y. Kang, "A Variable Step Size INC MPPT Method for PV Systems", *IEEE Trans. Ind. Electron.*, vol. 55, no. 7, pp. 2622-2628, Jul. 2008
- [6] T. Esmar, P. L. Chapman, "Comparison of Photovoltaic Array Maximum Power Point Tracking Techniques", *IEEE Trans. Energy Conv.*, vol. 22, no. 2, Jun. 2007
- [7] B. Subudhi, R. Pradhan, "A Comparative Study on Maximum Power Point Tracking Techniques for Photovoltaic Power Systems", *IEEE Trans. Sust. Energy*, vol. 4, no. 1, Jan. 2013.
- [8] Y. Yang and F. P. Zhao, "Adaptive Perturb and Observe MPPT Technique for Grid-connected Photovoltaic Inverters," *Procedia Engineering*, vol. 23, pp. 468-473, 2011.
- [9] W. Xiao and W. G. Dunford, "A Modified Adaptive Hill Climbing MPPT Method for Photovoltaic Power Systems", in *Proc. 35th Annu. IEEE Power Electron. Spec. Conf.*, pp. 1957-1963, 2004.
- [10] F. Zhang, K. Thanapalan, A. Procter, S. Carr, and J. Maddy, "Adaptive Hybrid Maximum Power Point Tracking Method for a Photovoltaic System," *Energy Conversion, IEEE Transactions on*, vol. 28, pp. 353-360, 2013.
- [11] S. K. Kollimalla and M. K. Mishra, "Novel Adaptive P&O MPPT Algorithm for Photovoltaic System Considering Sudden Changes in Weather Condition", pp. 653-658, 2013.
- [12] S. K. Kollimalla and M. K. Mishra, "Variable Perturbation Size Adaptive P&O MPPT Algorithm for Sudden Changes in Irradiance", *IEEE Trans. Sust. Energy*, to be published
- [13] S. K. Kollimalla and M. K. Mishra, "A Novel Adaptive P&O MPPT Algorithm Considering Sudden Changes in Irradiance", *IEEE Trans. Energy Conversion*, to be published
- [14] A. Jain, S. Sharma, and A. Kapoor, "Solar cell array parameters using Lambert W-function", *Solar Energy Materials and Solar Cells*, vol. 90, no. 1, pp. 25-31, 2006.
- [15] H. L. Tsai, C. S. Tu and Y. J. Su, "Development of Generalized Photovoltaic Model using MATLAB/SIMULINK," in *Proc. World Congress on Engineering and Computer Science 2008, WCECS '08, October 22-24, 2008, San Francisco, USA*
- [16] KC200GT High Efficiency Multicrystal Photovoltaic Module Data Sheet, Kyocera Corporation, Japan (<http://www.kyocera.com.sg/products/solar/pdf/kc200gt.pdf>)
- [17] Q. Mei, M. Shan, L. Liu, J. Guerrero, "A Novel Improved Variable Step-Size Incremental-Resistance MPPT Method for PV Systems," *IEEE Trans. Ind. Electron.*, vol. 58, pp 2427-2434, Jun. 2011.
- [18] A. Pandey, N. Dasgupta and A. K. Mukerjee, "Design issues in implementing MPPT for improved tracking and dynamic performance," in *Proc. IEEE IECON*, 2006, pp. 1957-1963, Jul. 2007

# Primjer implementacije privatnog cloud rješenja

Sreten Ćorić, Elektrotehnički fakultet, Univerzitet Crne Gore

**Sadržaj** — U radu je prikazana implementacija privatnog IaaS cloud rješenja. Primjena ovakvog rješenja je moguća kod kompanija koje imaju potrebu da za svoje poslovne procese posjeduju moderan informacioni sistem kojeg karakterišu efikasnost, skalabilnost, odgovarajući nivo raspoloživosti i bezbjednosti, u cilju kvalitetnog izvršavanja poslovnih procesa koji se oslanjaju na informaciono-komunikacione tehnologije.

**Glavne reči** — Cloud, klaster, IaaS, PaaS, SaaS, storidž, virtualizacija, VMware, LAN, SAN, FC, vSwitch, VMFS.

## I. UVOD

Cloud računarstvo može biti definisano kao moderan tip računarstva u kojem se dinamički skalabilni i često virtualizovani resursi obezbjeđuju kao servisi preko interneta ili u okviru privatne LAN mreže. Kao takvo, cloud računarstvo je postalo značajan tehnološki trend, i mnogi eksperti očekuju da će ovaj trend promijeniti paradigme računarstva i IT tržište.

Kod cloud tehnologije, korisnici koriste razne uređaje da bi pristupali programima, disk prostorima, i razvojnim aplikativnim platformama preko mreže, a sve kroz servise koje obezbjeđuju provajderi cloud-a. Prednosti cloud tehnologije uključuju i smanjenje investicionih troškova kod proširenja i obnove hardvera informacionih sistema.

Cloud računarstvo se može posmatrati kao kolekcija servisa, koja se prikazuje kao višeslojna računarska arhitektura. Aktuelni servisi u cloud računarstvu su IaaS (Infrastructure as a Service), PaaS (Platform as a Service) i SaaS (Software as a Service) [1].

Modeli razvoja cloud rješenja jasno prikazuju kategoriju cloud okruženja i ona su uglavnom podijeljena po vlasništvu, veličini i načinu pristupa. Aktuelni modeli razvoja cloud-a su privatni, javni, hibridni i community cloud [2].

Privatni cloud, koji je implementiran u radu, je platforma koja egzistira na cloud orijentisanim bezbjednim okruženjima (virtualizaciji). Ova platforma je osigurana sigurnosnom infrastrukturom kojom upravlja IT sektor neke kompanije. Pristup ovakvom cloud-u je dozvoljen samo autorizovanim korisnicima. Za implementaciju i održavanje ovakvog cloud-a u kompaniji uglavnom su zaduženi zaposleni u IT sektoru, i njihova odgovornost je da privatni cloud svojim karakteristikama zadovolji poslovne potrebe kompanije. Pod ovim karakteristikama se podrazumijevaju parametri kao što su: redundantnost, raspoloživost, performanse, modularnost i agilnost privatnog cloud rješenja. Sigurnosni rizici koji postoje u

svim cloud rješenjima se kod privatnog cloud-a na određene načine mogu efikasno umanjiti.

IaaS sloj u cloud računarstvu podrazumijeva da se, uz korišćenje tehnologije virtualizacije, kreira cloud, koji obezbjeđuje IT infrastrukturu koju karakteriše visoka dostupnost, raspoloživost, pouzdanost i performanse. IaaS usluga dostavlja virtualne servere, sa odgovarajućim/željenim hardverskim resursima, performansama i stepenom raspoloživosti, u cilju zadovoljavanja potreba krajnjeg korisnika za kvalitetno izvršavanje određenih poslovnih procesa. Među glavnim razlozima zbog kojih kompanije uvode IaaS rješenje su smanjenje finansijskih izdataka za nabavku nove hardverske infrastrukture i vremena potrebnog za ovaj postupak. IaaS rješenje obezbjeđuje brže i jednostavnije kreiranje serverske infrastrukture koja će podržati izvršavanje i isporuku novih poslovnih proizvoda za internu upotrebu u kompaniji ili za tržište. [7]

Implementacijom IaaS rješenja, postiže se unifikacija hardverske infrastrukture u kompaniji koja je preduslov za optimalno upravljanje troškovima kod održavanja jednog informacionog sistema.

## II. SVRHA I KONCEPT IMPLEMENTIRANOG CLOUD RJEŠENJA

Implementiran privatni IaaS cloud je primjenljiv u praksi i namjenjen je za potrebe kompanija srednje veličine, koje posjeduju do 100 virtualnih servera. Cilj je da se dobije sistem koji je bezbjedan i koji omogućava nastavak isporuke servisa u slučaju otkaza bilo koje komponente na bilo kojem nivou u infrastrukturi kompanijskog informacionog sistema.

Ovakav privatni IaaS cloud zadovoljava aktuelne standarde i posjeduje sve ključne karakteristike cloud-a ovog tipa. Privatni IaaS cloud treba da podržava efikasno kreiranje i izvršavanje virtualnih servera koji su neophodni za funkcionisanje kritičnih IT servisa u kompaniji.

U osnovi cloud rješenja se nalazi virtualizaciona platforma koju čine dva VMware vSphere klastera, koja su instalirana na dvije lokacije povezane redundantnim Ethernet linkovima. Za realizaciju rješenja obezbjeđeno je šest fizičkih hostova, četiri Ethernet LAN (Local Area Network – lokalna Ethernet mreža) sviča L2 nivoa, četiri FC (Fibre Channel – brzi komunikacioni protokol) sviča, dva FC storidž sistema, dva firewall-a i odgovarajuće mrežne konekcije u okviru oba VMware klastera. LAN svičevi se koriste za umrežavanje fizičkih servera na LAN mrežu, a FC svičevi za povezivanje fizičkih servera na storidž sisteme kroz SAN (Storage Area Network - mreža za konekciju servera na storidž disk prostor) mrežu. Svi hostovi imaju identične procesorske performanse 2 x 6Core x 2,30GHz CPU i količinu radne memorije 160GB RAM.

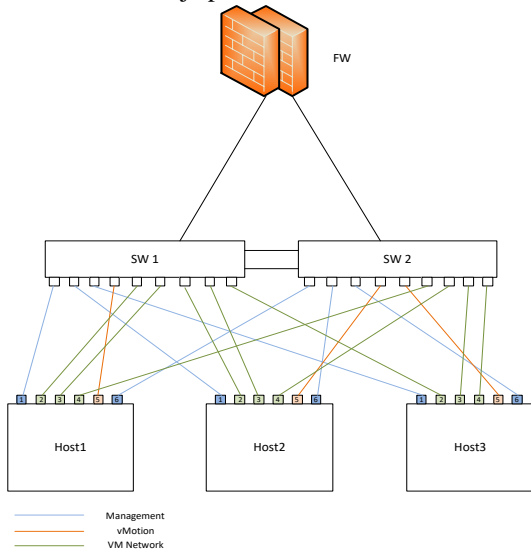


S obzirom da su oba VMware klastera instalirana i konfigurirana na identičan način, u radu je, zbog jednostavnijeg opisa i razumijevanja, prikazana fazna implementacija jednog od dva klastera (Poglavlje III).

### III. IMPLEMENTACIJA CLOUD RJEŠENJA

#### A. Konfiguracija LAN mreže i instalacija VMware hipervizora

LAN konfiguracija je jedna od vitalnih komponenti VMware HA (High Availability – visok stepen dostupnosti) klastera. Šema veza LAN mreže jednog od dva VMware klastera je prikazana na Sl.1.



Sl.1. LAN mreža jednog VMware klastera

S obzirom da je svaki od hostova opremljen sa po šest mrežnih adaptera, za LAN konfiguraciju su ispoštovane najbolje prakse u cilju redundantnosti i visokog stepena raspoloživosti servisa koje će sistem podržavati.[3] Cijeli sistem se nalazi iza firewall-a, i na taj način se vrši kontrola saobraćaja u oba smjera, ka i od sistema. Za povezivanje hostova u klasteru koriste se dva LAN sviča i na taj način je ostvarena redundansa na nivou fizičke mreže. Po dva mrežna adaptera na serverima (1 i 6) će se koristiti za Management mrežu (mreža za pristup i administraciju sistema i VMware management task-ove). Jedan mrežni adapter je konfigurisan za vMotion (VMware funkcionalnost migracije “uživo” virtualnih mašina između fizičkih servera u klasteru). Preostala tri mrežna adaptera su konfigurirana za VM Network (mreže koje koriste sve virtualne mašine koje se izvršavaju na VMware klasteru). Na LAN svičevima je definisan VLAN za Management mrežu i trunk portovi za VM network mrežu.

Instalacija VMware hipervizora se obavlja na isti način na svim ESXi hostovima u klasterima. Instalacioni disk se mount-uje na server i instalacija se obavlja na lokalne SAS diskove na hostovima koji su konfigurirani u RAID1 niz. Nakon instalacije hipervizora vrši se DNS i Management IP konfiguracija hostova i odabir adekvatnih mrežnih adaptera za ovu namjenu (1 i 6), u skladu sa Sl.1. Preporuka je da se, u cilju eventualne integracije VMware okruženja sa servisom aktivnog direktorijuma i olakšane administracije i kontrole pristupa, kao DNS serveri definišu IP adrese domen kontrolera u sistemu. Na samim

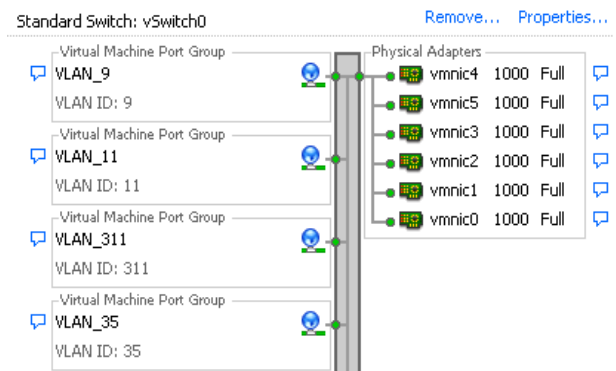
host serverima su definisane statičke IP adrese iz definisanog VLAN-a za Management mrežu.

#### B. Instalacija vCenter servera i kreiranje klastera

Svaki od dva klastera u cloud rješenju imaju po jedan instaliran VMware vCenter server koji služi za konfiguraciju i upravljanje VMware klasterom od strane administratora sistema. vCenter serveri su realizovani kao virtualni u okviru samog okruženja a njihova visoka raspoloživost je obezbeđena kroz samo učešće u HA konfiguraciji. Nakon instalacije vCenter servera, okruženju je moguće pristupiti i kroz VMware vSphere Web Client (alat za pristup, konfiguraciju i administraciju vSphere okruženja). Sljedeći koraci su kreiranje vSphere datacentra, klastera i učlanjivanje po tri hosta u svaki od klastera.

#### C. Konfiguracija virtualne LAN mreže

Virtualna LAN mreža je implementirana kroz funkcionalnost standardnog virtualnog sviča na svakom VMware hostu ponaosob (VMware Standard vSwitch) u okviru kojeg su definisani VLANovi za Management mrežu i virtualne mašine. Na Sl.2. je prikazan virtualni svič na jednom od VMware hostova.



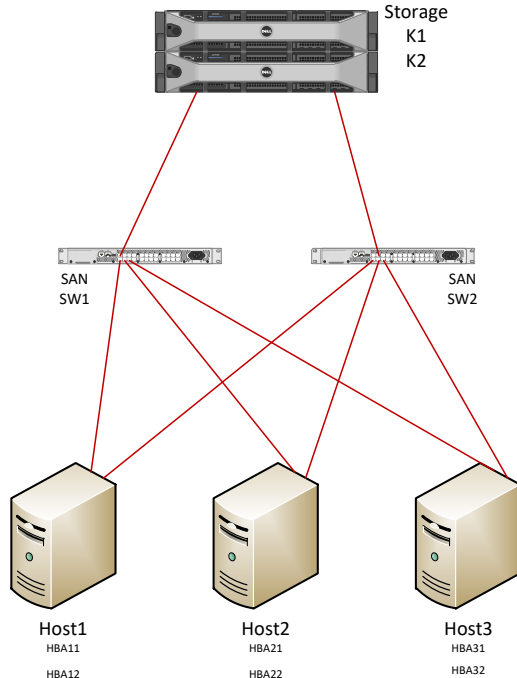
Sl.2. Virtualni svič na VMware hostu

Važno je istaći da je za pravilan rad HA funkcionalnosti neophodna identična konfiguracija virtualnog sviča na svim hostovima u klasteru, kako bi u slučaju failover-a virtualne mašine mogle vršiti komunikaciju nakon prelaska na drugi host. Hostovi u klasteru vrše međusobnu hartbeat razmijenu, u svrhu validacije egzistiranja u istom. Ukoliko host prestane da prima hartbeat od svih drugih hostova iz klastera, on ICMP protokolom provjerava konekciju do svoje mrežne izolacione adrese (gateway adrese) da bi provjerio je li izolovan sa mreže. Ukoliko može da ostvari ping sa gateway-em, onda host nije izolovan sa mreže i pokreće se failover procedura prenosa virtualnih mašina sa hostova koji su izolovani. S obzirom da je implementirana redundansa na nivou Management mreže, ostvaruje se pouzdanija detekcija problema na ovom nivou. Hartbeat se šalje preko više putanja i smanjuje se broj slučajeva izolacije hosta sa mreže. [3]

Redundansa na nivou mrežnih adaptera u okviru jednog servera člana klastera je implementirana koristeći NIC Teaming (korišćenje više mrežnih adaptera u timu) funkcionalnost za Management i VM Network, a za vMotion je korišćena Stand By (postojanje rezervnog mrežnog adaptera koji se aktivira u slučaju “pada” primarnog adaptera) funkcionalnost.

#### D. Storiđž umrežavanje i konfiguracija

Analogno LAN mreži koja je prikazana na SL.1., SAN mreža VMware klastera se realizuje na sličan način, povezivanjem HBA (Host Bus Adapter – kartica za povezivanje servera na storiđž sistem) kartica sa hostova na FC (SAN) svičeve, na koje je konektovan FC storiđž. Na SL.3. je prikazana SAN mreža jednog VMware klastera.



Sl. 3. Šema SAN konekcija jednog od klastera u sistemu

Da bi se postigla redundantnost sistema u dijelu SAN konfiguracije (slično LAN konfiguraciji), svaki od servera je opremljen sa dva HBA, koji su konektovani na različite SAN svičeve. Storiđž sistem u okviru kojeg postoje dva kontrolera je takođe konektovan na oba SAN sviča.

Storiđž umrežavanje, tj. povezivanje servera na storiđž sistem, u cilju mapiranja disk prostora (dodjele i prikaza disk prostora) sa storiđž-a na servere, realizuje se odgovarajućim zoniranjem na SAN svičevima. Elemente jedne zone čine WWN (World Wide Name) jedinstveni identifikatori sa uređaja koji se povezuju na storiđž sistem posredstvom SAN sviča. WWN identifikator je definisan u okviru HBA kartice na serveru, i može se napraviti njegova paralela sa MAC (Media Access Control) jedinstvenim identifikatorom na Ethernet kartici. Jedan WWN se istovremeno može nalaziti u više zona, tako da kreiranje i upravljanje zoning konfiguracijom zahtijeva adekvatno planiranje i dizajn strukture zona za određeni servis, naročito u slučajevima kada zoning konfiguracija sadrži više nezavisnih servisa u sebi. Postoje dva tipa zoning konfiguracije: tzv. soft i hard zoniranje [6].

Za realizaciju storiđž umrežavanja u sistemu koji je implementiran u ovom radu se koristi soft zoniranje koje karakterišu zone koje sadrže WWN identifikatore servera i storiđž-a, bez obzira na koje se fizičke portove SAN sviča ovi WWN-ovi povezuju. U Tabeli 1. su prikazane dvije zone koje su konfigurisane na SAN svičevima u cilju optimalnog funkcionisanja VMware klastera i storiđž sistema.

TABELA 1: KONFIGURISANE ZONE NA SAN SVIČEVIMA

SW	Zona	Članovi zone
SAN SW1	1	K1, Host1_HBA11, Host2_HBA21, Host3_HBA31
SAN SW2	2	K2, Host1_HBA12, Host2_HBA22, Host3_HBA32

Ovakvom zoning konfiguracijom, omogućeno je da se serveri u klasteru “vide” sa storiđž sistemom, i stvoreni su uslovi da je sa strane storiđž sistema moguće kreirati FC hostove (hostove VMware klastera) i izvršiti dodjelu disk prostora istima. Na SL.4. su prikazani kreirani FC hostovi (tri hosta koji čine VMware klaster) na storiđž sistemu.

Name	Status	Host Type	# of P...	Host Mappings
ESXi1	✓ Online	Generic	2	Yes
ESXi2	✓ Online	Generic	2	Yes
ESXi3	✓ Online	Generic	2	Yes

Sl.4. Kreirani FC hostovi na storiđž sistemu

Kod kreiranja hostova važno je imati tačnu evidenciju WWN-ova za određeni host, kako se ne bi desilo da se kod kreiranog određenog hosta nađe WWN od drugog hosta. Na SL.5. je prikazan kreirani storiđž disk prostor kojeg će koristiti VMware hostovi (storiđž volume).

Name	State	Cap...	Pool	Host Mappings
VMware_15K	✓ Online	4.07 TiB	VMware_15K	Yes

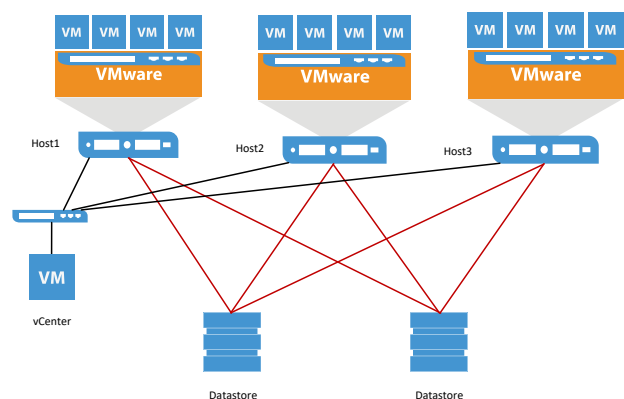
Sl.5. Storiđž volume za VMware klaster

Nakon mapiranja volume-a (prikaza volume-a FC hostovima) potrebno je izvršiti formatiranje particije VMFS fajl sistemom. VMFS je VMware-ov fajl sistem koji se koristi za smještanje i izvršavanje virtualnih mašina. Formatiranje se vrši na jednom od ESXi hostova i nakon toga se vrši Rescan volume-a na ostalim ESXi hostovima. Na SL.6. je prikazan disk prostor formatiran VMFS fajl sistemom.

Volume Properties	
General	Format
Datastore Name: VMware_15K	File System: VMFS 5.60
Total Capacity: 4.07 TB	Maximum File Size: 2.00 TB
	Block Size: 1 MB

Sl.6. Disk prostor formatiran VMFS fajl sistemom

U ovom momentu virtualizaciona platforma je spremna za instalaciju i izvršavanje virtualnih mašina. Na SL.7. se nalazi logička šema realizovanog VMware klastera.



Sl.7. VMware klaster na jednoj lokaciji

U osnovi šeme se nalaze LAN i SAN mreže, disk prostor sa storidž sistema (Datastore) sa instaliranim VMware hipervizorom na hostovima i virtualnim serverima na njima.

Ethernet LAN mreža je prikazana crnom bojom, a SAN mreža crvenom bojom. Fajlovi virtualnih mašina su smješteni na Datastore prostor u okviru storidž-a, tako da se svaka od mašina, po potrebi, može izvršavati na svakom od fizičkih hostova u klasteru.

Kroz navedene faze implementacije u poglavlju III, na identičan način, vrši se implementacija VMware HA klastera na drugoj lokaciji, i nakon toga međusobno redundantno Ethernet povezivanje oba VMware klastera u jednu cjelinu.

#### E. Replikacija između lokacija

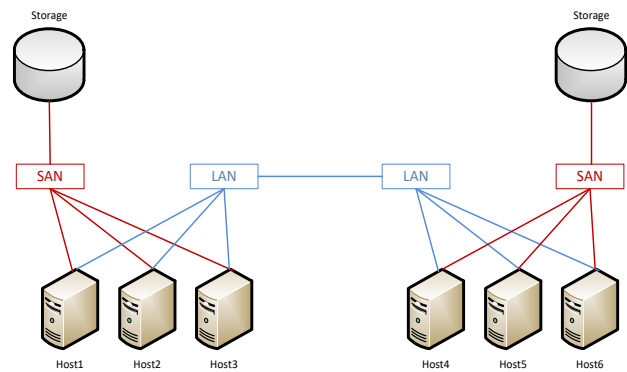
S obzirom da će za određeni broj virtualnih mašina koje se koriste za obavljanje kritičnih poslovnih procesa, biti neophodno obezbjediti dostupnost na obje lokacije, implementirano je rješenje za replikaciju virtualnih mašina sa jedne na drugu lokaciju, na nivou VMware hipervizora. Za ove potrebe izvršena je instalacija vSphere Replication virtualnih appliance-a na oba klastera u sistemu, i njihova konfiguracija. Ovi appliance-i se izvršavaju kao virtualne mašine na klasterima i imaju ulogu da upravljanju konfigurisanim asinhronim replikacijama virtualnih mašina sa jednog VMware klastera na drugi ili obrnuto. Ovom funkcionalnošću se postiže mogućnost da se u slučaju potrebe izvrši brz recovery virtualne mašine na drugoj lokaciji sa ažurnim stanjem i adekvatnim RPO (Recovery Point Objective – tačka oporavka) i RTO (Recovery Time Objective – vrijeme oporavka) parametrima.

VMware vSphere Replication je funkcionalnost koja obezbjeđuje zaštitu podataka i disaster recovery rješenje na nivou virtualne mašine. Ovo rješenje je u potpunosti integrisano sa vCenter serverom.

#### IV. PRIVATNO IAAS CLOUD RJEŠENJE

Implementirana virtualizaciona platforma, u formi dva povezana VMware klastera na dvije lokacije, sa adekvatnim kontrolama pristupa i nivoom bezbjednosti, čini visokoraspoloživi privatni IaaS cloud. Za upravljanje ovakvim rješenjem neophodno je stručno IT osoblje koje će u svakom momentu biti spremno da odgovori na zahtjeve poslovnih procesa (efikasno isporučiti nove virtualne servere odgovarajućih karakteristika) i koje će se brinuti o svim aspektima funkcionisanja jednog ovakvog sistema. Ovako realizovan sistem stvara mogućnost da nema potrebe za ad-hoc investicijama u hardver, te se ulaganja mogu blagovremeno planirati.

Ukupne performanse sistema, procesorska moć, radna memorija i disk prostor se dobijaju zbirom pojedinačnih performansi svakog od fizičkih hostova u sistemu i raspoloživog storidž prostora. Dakle na raspolaganju je ukupno 6x2x6x2.30GHz=165.6GHz CPU i 6x160GB=960GB RAM memorije i 2x4,07TB=8,14TB korisnog storidž prostora. Na Sl.8. se nalazi šema implementiranog cloud rješenja.



Sl. 8. Implementirano privatno IaaS cloud rješenje

Ovaj model cloud-a je odgovarajući za kompanije koje imaju servise i podatke isuviše osjetljive (npr. povjerljivi poslovni podaci ili kritični servisi) za smještaj u javni ili hibridni cloud, i to ga čini konkurentnim u odnosu na cloud provajder rješenja.

#### V. ZAKLJUČAK

Prije nego što kompanija donese odluku o korišćenju cloud tehnologije za potrebe svojih poslovnih procesa, potrebno je izvršiti procjenu opravdanosti i analizu rizika (sa funkcionalnog i bezbjednosnog aspekta) u odnosu na cloud rješenje koje se čini prihvatljivim.

Pored toga što rješenje prikazano u radu ima svoju praktičnu primjenu i karakteristike modernog privatnog IaaS cloud-a, moguće ga je dalje unapređivati u infrastrukturnom i funkcionalnom pogledu. Osim nadogradnje u cilju podizanja performansi i bezbjednosti, moguća je i eventualna transformacija privatnog IaaS cloud-a u privatni PaaS cloud u budućem periodu.

#### LITERATURA

- [1] Borko Furht, Armando Escalante, "Handbook of Cloud Computing", 2010
- [2] V.Victories, "4 Types of Cloud Computing Deployment Model you need to Know" 2015
- [3] "VMware ESXi and vCenter Server Documentation", VMware Available: <https://pubs.vmware.com/>
- [4] Gustavo Santana, "Data Center Virtualization Fundamentals", Cisco Press 2014
- [5] "VMware Infrastructure Architecture Overview", Available: <https://www.vmware.com/>
- [6] "Understanding Fibre Channel Zones", Enterprise Networking Planet, Available: <http://www.enterprisenetworkingplanet.com/>
- [7] Rajat Kandpal, Vishal Kumar, "IaaS Implementation of a Private Cloud using Open Source Technology", International Journal of Computer Applications, 2013

#### ABSTRACT

This paper presents the implementation of the private IaaS cloud solution. The application of such solution is possible for companies, that need to have a modern information system for their business processes, which is characterized by efficiency, scalability, availability and the appropriate level of security, in order to quality execution of business processes that are leaning on ICT.

#### EXAMPLE OF THE IMPLEMENTATION OF PRIVATE CLOUD SOLUTION

Sreten Ćorić, Faculty of Electrical Engineering,  
University of Montenegro

# Hibridni koncept Smart Home sistema

Milan Zejak, Milovan Radulović, Member IEEE

**Sadržaj** - U radu su opisani različiti načini izvedbe Smart Home sistema: žičani, bežični i hibridni koncept. Pokazano je koji razlozi opredjeljuju veliki broj proizvođača i korisnika ka hibridnom konceptu izvedbe. Takođe je razmatrana problematika daljinskog pristupa Smart Home sistemu čime se postiže potpuna efektivnost i opravdanost ulaganja sa aspekta poboljšanja energetske efikasnosti samog objekta.

U radu je prikazan primjer hibridnog Smart Home sistema kog je moguće primijeniti na postojeću, standardnu, električnu instalaciju. Opisan je jedan postupak izrade *custom made* aplikacijama za upravljanje ovakvim sistemima.

**Gljučne reči** — Smart Home, energetska efikasnost, daljinski pristup, *cutom made* aplikacije.

## I. UVOD

Šezdesetih godina prošlog vijeka, svima je postalo jasno da progres na polju elektronskih uređaja kako u industriji tako i u domaćinstvima može vremenom početi da predstavlja veliki problem po pitanje snabdijevanja električnom energijom. Sa tim ciljem je formirano više radnih grupa, istraživačkih centara i sl. sa ciljem da se podigne svijest javnosti o načinima smanjenja potrošnje električne energije a isto tako, sa istim ciljem, da se izvrši pritisak na proizvođače aparata i opreme. Tako je nastao pojam ‘Energetska efikasnost’.

U proteklih pola vijeka dosta toga se uradilo po pitanju energetske efikasnosti. Proizvođači razne opreme/aparata su uočili da im niska cijena a isto tako i niska potrošnja električne energije njihovih proizvoda, obezbjeđuje bolju konkurentnost na tržištu pa samim tim i profit. Nesumnjivo je da na energetska efikasnost u jednom domaćinstvu utiče dobar izbor kućnih aparata u smislu izbora uređaja koji troši manje električne energije za normalan rad, ali ipak, na energetska efikasnost isključivo utiču krajnji korisnici.

*Smart Home* predstavlja koncept koji definiše boravište koje uključuje računarsku ili drugu komunikacijsku mrežu u svrhu povezivanja ključnih električnih i elektronskih uređaja i usluga, te omogućava da se njima daljinski upravlja, prati i pristupa [1]. Ovaj koncept omogućava korisniku da kreiranjem scenarija (više operacija koje će se izvršiti u istom trenutku) omogućiti sistemu da u zavisnosti od zadatih parametara i ulaznih informacija automatski izvršava određene operacije. Na taj način korisnik može učiniti svoj dom energetska efikasnijim a samim tim i bezbjednijim što je i osnovni cilj ovog koncepta.

U ovom radu biće riječi o Smart Home sistemima,

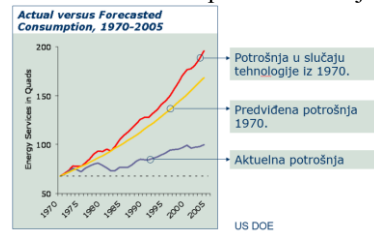
Milan Zejak, (autor za kontakte), ION Solution doo Podgorica, Kruševo 95, 84000 Bijelo Polje, Crna Gora (telefon: 382-67-004020; e-mail: milanzejak@rocketmail.com).

Milovan Radulović Elektrotehnički fakultet u Podgorici, Univerzitet Crne Gore Džordža Vašingtona bb, 81000 Podgorica, Crna Gora (telefon 382-67-313309, e-mail: milovanr@ac.me).

njihovim načinima izvedbe, daljinskom pristupu i uticaju ovakvih sistema na energetska efikasnost sa posebnim osvrtom na prednosti hibridnog načina realizacije sistema.

## II. ENERGETSKA EFIKASNOST SA STANOVIŠTA POTROŠNJE ELEKTRIČNE ENERGIJE

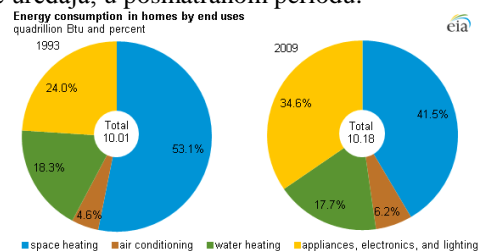
Poslije drugog svjetskog rata, počela je brza elektrifikacija naseljenih mjesta u velikom dijelu svijeta. Brzim razvojem tržišta i sve lakšim načinima da srednja klasa društva dodje u posjed raznih električnih i elektronskih uređaja nametnulo se pitanje, da li će proizvođači električne energije (hidroelektrane, termoelektrane i sl.) moći da podmire potrebe klijenata. Rezultati studije iz 1970 godine[2], prikazani na Sl. 1., daju predviđanja po pitanju potrošnje električne energije u tom trenutku. Bilo je jasno da se nešto mora preduzeti kako bi se tolika potreba za električnom energijom umanjila jer tadašnje elektrane, a ni one čija se izgradnja planirala u budućnosti, ne bi mogle da podmire očekivane potrebe. Tako je nastao koncept ‘Energetska efikasnost’ koji je imao za cilj da, što je više moguće, smanji upotrebu svih energenata a da ne naruši funkcionalnosti sistema, održi komfor na istom nivou i ispuni očekivanja korisnika.



Slika 1. Predviđanja 1970. godine o potrošnji energenata u budućnosti [2]

U SAD-u je 2013 procijenjeno da 37% ukupne potrošnje električne energije čine rezidencijalni objekti (stambeni objekti – kuće, stambene zgrade), komercijalni objekti (javni objekti - restorani, bolnice, škole, itd) 35%, a industrija 27% ukupne potrošnje [3].

Na Sl. 2 prikazana je potrošnja električne energije na nivou jednog rezidencijalnog objekta [4]. u SAD-u, kao i koliko se, procentualno, promijenio udio potrošnje električne energije sa aspekta različitih tipova potrošača i namjene uređaja, u posmatranom periodu.



Slika 2. Promjene potrošnje energije na nivou jednog domaćinstva 1993 - 2009. godine [4]

### III. SMART HOME KONCEPT I RJEŠENJA

Cilj *Smart Home* koncepta je povećanje energetske efikasnosti, sigurnosti i komfora u jednom domaćinstvu. Koncept nastoji da omogući pristupanje, upravljanje i praćenje što većeg broja sistema i podsistema u domaćinstvu.

U proteklih 10 godina, razvojem računara i interneta, sve je više proizvođača opreme za *Smart Home* sisteme. Oni nude širok asortiman proizvoda koji komuniciraju međusobno standardizovanim ili nestandardizovanim protokolima. Neki od standardizovanih protokola su: KNX, X10, C-Bus, Zigbee, Z-Wave, itd, [5], [6].

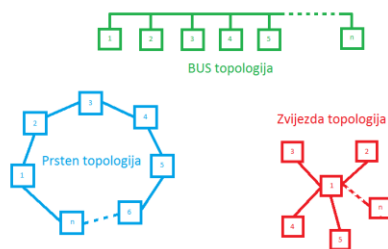
Savremeni *Smart Home* sistem omogućava lokalni i daljinski pristup, upravljanje i nadzor sledećih uređaja i sistema: podno grijanje, razni sistemi klimatizacije, rasvjeta, alarmni sistem, sistem za dojavu i zaštitu od požara, ozvučenje i multimedije, roletne i zavjese, priprema sanitarne tople vode, mjerenje potrošnje električne energije, parno i električno grijanje prostora, itd.

Postoje tri načina izvedbe ovakvih sistema: žičani, bežični i hibridni.

#### A. Žičani način izvedbe *Smart Home* sistema

Kod ovog načina implementiranja *Smart Home* sistema, svi kontroleri bilo da su u razvodnom ormaru ili u prostoru moraju biti međusobno povezani posebnim prenosnim medijumom. Obično se kao prenosni medijum koristi kabl Jh(St)H 2x2x0.8mm ili Jh(St)H 1x2x0.8mm, a sve to zavisno od proizvođača opreme ili protokola koji se koristi za komunikaciju između uređaja. Najčešći protokoli koji se koriste u ovakvoj vrsti komunikacije su: KNX, Buspro, X10, Insteon, itd, [6].

Kod ovog načina implementiranja postoje tri vrlo popularne topologije: prsten, zvijezda, i *bus*, prikazane na Sl. 3.



Slika 3. Topologije u žičanoj izvedbi *Smart Home* sistema

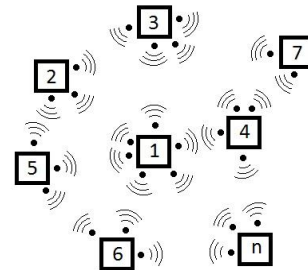
Vrsta topologije koja će se koristiti zavisi od proizvođača opreme. Ukoliko proizvođač dozvoljava kombinovanu topologiju tim bolje, jer se po nekada dešava da dođe do prekida komunikacionog medijuma, pa se takav problem može prevazići primjenom neke druge topologije ako su za to ispunjeni odgovarajući uslovi.

Ovakav način izvedbe je preporučljiv ukoliko se *Smart Home* koncept implementira na objektima u fazi izgradnje. Na već izgrađenim objektima, ovakav način izvedbe povlači i dodatne, ne tako zanemarive, troškove.

#### B. Bežični način izvedbe *Smart Home* sistema

Kod bežičnog načina implementiranja *Smart Home* sistema, svi kontroleri međusobno komuniciraju bežičnim

putem. Najčešći protokoli koji se koriste u bežičnoj komunikaciji među uređajima su: Z-Wave, Zigbee, Bluetooth, WiFi, itd [5]. Ovaj način izvedbe ima određenih ograničenja. Maksimalan broj uređaja koji mogu sačinjavati jednu ovakvu mrežu je, od prilike, dva puta manji nego u slučaju žičane izvedbe. Takođe, pokazalo se da je domet radio signala kojim elementi sistema međusobno komuniciraju najčešće 10 metara zbog fizičkih prepreka u objektu, iako proizvođači najčešće navode da je maksimalan domet 100m. Međutim, ovaj način izvedbe je izuzetno zahvalan kada su u pitanju već izgrađeni objekti. Proizvođači proizvode elemente kojima mogu primijeniti *Smart Home* koncept a sve to bez ikakvih izmjena u oblasti električnih instalacija. Bez obzira na nedostatke, ovaj način izvedbe je svojom funkcionalnošću i sve nižom cijenom, u posljednje vrijeme postao izuzetno popularan. Topologija ovakvog načina izvedbe je prikazana na Sl. 4.



Slika 4. Topologija u bežičnoj izvedbi *Smart Home* sistema

Postavlja se pitanje na koji onda način bežični uređaji funkcionišu, s obzirom na to da se većina njih (konkretno prekidači i senzori pokreta) postavljaju na mjestima gdje postoji samo dovodna faza, tačnije nije prisutan neutralni provodnik. Aktualna su dva pristupa prevazilaženja ovog problema. Prvi pristup je prisustvo baterije unutar prekidača i senzora. Ta baterija snabdijeva ove uređaje energijom potrebnom za rad. Međutim, baterije imaju svoj vijek trajanja. Bez obzira koliko taj vijek bio dug, svakako se nekada baterija mora promijeniti. To iziskuje dodatne troškove održavanja i mogućnost prekida u funkcionisanju sistema. Drugi pristup je upotreba specijalnih kondenzatora izuzetno visoke kapacitivnosti a malih dimenzija. Kondenzatori, ovog tipa, su izuzetno skupi pa se samim tim povećava cijena uređaja. Kondenzatori se pune svaki put kada prekidač propusti električnu energiju prema potrošaču. U ovom dijelu još puno toga treba poboljšati i danas je izuzetno atraktivna istraživačka tema.

#### C. Hibridni način izvedbe *Smart Home* sistema

Hibridni način implementacije *Smart Home* sistema podrazumijeva kombinaciju dvije prethodno opisane izvedbe. Našao je široku primjenu u praksi jer, pored elemenata u prostoru koji međusobno komuniciraju radio signalom, navedeni problemi se, primjenom ovog koncepta, mogu riješiti i u samom razvodnom ormaru gdje se elementi povezuju žičano.

Veza između protokola za bežičnu i protokola za žičanu komunikaciju se postiže preko uređaja koji se popularno naziva bridge ili most. Kod ovog načina izvedbe neophodan je i veoma važan element u sistemu. Zavisno od proizvođača, u ovom uređaju može biti implementiran i ethernet port pa samim tim, predstavlja vezu sa trećim

protokolom, TCP/IP. Veza sistema sa TCP/IP protokolom je od izuzetne važnosti ne samo za ovaj način izvedbe već i za ostala dva. Veza sa TCP/IP protokolom omogućava povezivanje na računarsku mrežu, a samim tim se ostvaruje i veza sa Internetom. Ovakav pristup je od ključnog značaja kada je u pitanju daljinski pristup sistemu posredstvom Interneta. Inače svako konvertovanje signala iz jednog protokola u drugi je izuzetno kompleksno i podložno greškama i smetnjama, zato bridge-ovi mogu biti izuzetno skupi. Proizvođači opreme za Smart Home sisteme pretežno preferiraju svoje, nestandardizovane protokole za komunikaciju među uređajima, pa su veoma česti i bridge-ovi koji predstavljaju vezu nestandardizovanih sa standardizovanim protokolima u automatičnom načinu.

#### IV. DALJINSKI PRISTUP

Daljinski pristup omogućava korisniku da sa udaljene lokacije pristupi *Smart Home* sistemu i upravlja određenim potrošačima, sistemima i podsistemima, zavisno od želje i potrebe. Nekada se pristup udaljenim sistemima obavljao najčešće preko GSM-a. U proteklih petnaestak godina, razvojem internet tehnologija, postao je izuzetno popularan jedan novi koncept, *Internet of things*. Ovaj koncept se odnosi na mrežu fizičkih objekata sa ugrađenom elektronikom, *software*-om i senzorima, međusobno povezanim posredstvom računarske mreže, koji mogu međusobno komunicirati, razmjenjivati informacije i uticati posredno ili neposredno jedni na druge [7]. Konkretno, na nivou jednog domaćinstva to se odnosi na sve elektronske uređaje koji se tu koriste. Zahvaljujući ovom konceptu i sam *Smart Home* koncept je dobio na značaju i smislu a daljinski pristup se obavlja posredstvom računarske mreže. Daljinski pristup posredstvom računarske mreže je brži, sigurniji i efikasniji od svakog prethodnog vida takvog pristupa.

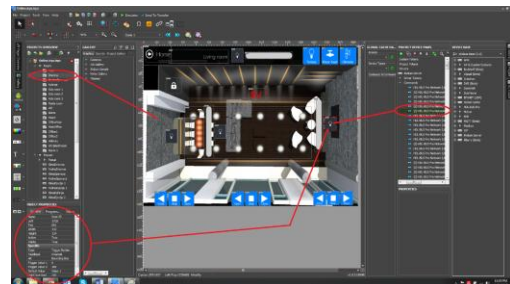
Kako je koncept *Smart Home*-a u posljednje vrijeme postao izuzetno interesantan a i daljinsko upravljanje jednim takvim sistemom izuzetno važno, *smartphone* je dobio priliku da pruži još jedan servis više. Mogućnosti *smartphone*-a iskoristili su brojni proizvođači opreme za *Smart Home* sisteme uvidjevši da na taj način mogu pojačati konkurentnost na tržištu. Napravili su svoje aplikacije za upravljanje *Smart Home* sistemom koje je moguće instalirati na *smartphone*-u i besplatne su. Na taj način, korisnicima su omogućili pouzdan daljinski pristup sistemu bez dodatnih troškova. Dovoljna je samo konekcija na Internet i nije bitno da li se pristupa sa *smartphone*-a, računara ili tableta, *interface* je isti.

Većina aplikacija, koje se koriste za upravljanje *Smart Home* sistemima nijesu *custom made*, tj. urađene su kao šablon i korisnik ne može mijenjati njihov izgled i dizajn shodno želji i potrebi. Smisao svake aplikacije, bilo da je za *smartphone*, računar ili tablet, je da bude što više funkcionalna i jednostavna za korišćenje. Takođe, vodi se računa i o tome da se ne potroši mnogo vremena na prilagođavanje aplikacije jednom *Smart Home* sistemu. Naime, ima jako jednostavnih ali i jako komplikovanih *Smart Home* sistema, zavisno od korisnikove želje i potrebe. Samim tim, sve potrošače, sisteme i podsisteme kojima upravlja *Smart Home* sistem treba povezati i sa

aplikativnim dijelom kako bi se korisniku pružila mogućnost da im daljinski upravlja, prati i pristupa. Međutim, pojedini korisnici spremni su platiti visoku cijenu kako bi dobili ono šta žele. U tu svrhu se koriste *custom made* aplikacije.

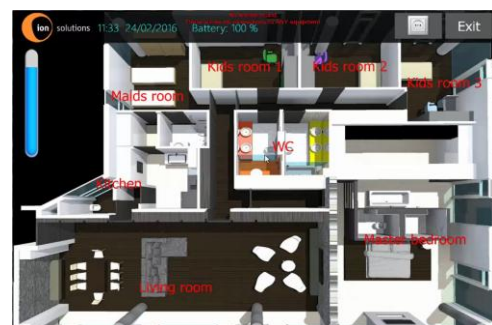
*Custom made* aplikacije se kreiraju po želji i potrebi korisnika, počev od dizajna, izgleda pa i funkcionalnosti. Jedan vrlo popularan alat za kreiranje *custom made* aplikacija je "iRidium Studio", ruske kompanije iRidium.

Kreiranje *custom made* aplikacije upotrebom ovog alata, bazira se na unošenju fotografija i ikonica koje će činiti *interface* aplikacije i dizajn izgleda dugmadi čijim pritiskom će se kontrolisati pojedini potrošači, sistemi i podsistemi. Grafički objekti se zatim povezuju sa adresama kontrolera i njihovim digitalnim i analognim izlazima, Sl. 5.



Slika 5. Kreiranje *interface*-a aplikacije u "iRidium Studio"

Nakon kreiranja aplikacije, željenog izgleda i funkcionalnosti, sve se to sačuva u formatu fajla čijom se aktivacijom na nekoj radnoj stanici instalira server aplikacija. Izgled jedne gotove i funkcionalne *custom made* aplikacije, kojom se upravlja jednim stanom u Podgorici, napravljene upotrebom prethodno opisanog alata, prikazan je na Sl. 6 [8].



Slika 6. Izgled jedne *custom made* aplikacije [8]

Kao završni dio implementacije korisniku se instalira klijent aplikacija na neki od uređaja preko kojih želi pristupiti sistemu, npr. *smartphone*, i poveže ista sa server aplikacijom. Aplikacija se obično preuzme sa iTunes-a ili Google Play-a.

Sve aplikacije moraju imati neki server na kome će se izvršavati. To može biti neki računar ili ga proizvođač pravi u sklopu *bridge* modula. Ukoliko je izabran računar, onda to može biti minimalne konfiguracije, na primjer *Raspberry Pi*, jer ovakve aplikacije nijesu zahtjevne.

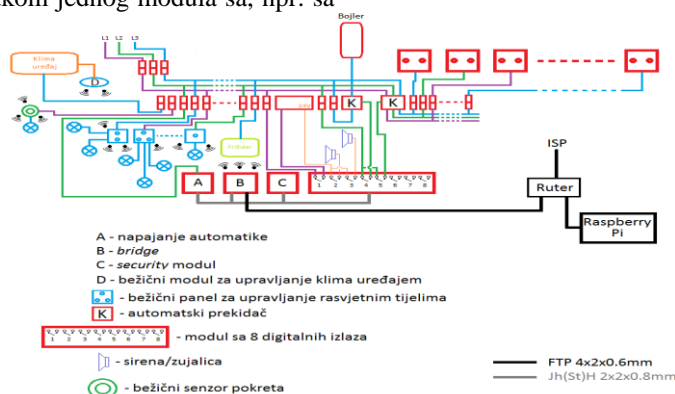
Pristup serveru sa udaljene lokacije se vrši posredstvom računarske mreže. Tu postoje dva načina za pristup. Jedan

je zakup statičke javne IP adrese, inače skuplji, a drugi, upotrebom dinamičkog DNS servera, inače kompleksniji ali jeftiniji i češće se primjenjuje. Takođe, ne treba zaboraviti ni *port forwarding* (*port mapping*).

*Smart Home* sistem se kreira shodno korisnikovim mogućnostima, želji i potrebi. Osnovni elementi sistema, opet zavisno od proizvođača, najviše učestvuju u cijeni (licenca za aplikaciju, *bridge* i napajanje automatike). U pitanju je vrijednost manja od 1000€, za sve. Mađutim, kada se uzme u obzir da dodatkom jednog *security* modula čija cijena ne prelazi par stotina evra, na postojeći *Smart Home* sistem, koji zamjenjuje alarmnu i požarnu central, a za povredu prostora u alarmnom stanju koristi već postojeće senzore za detekciju pokreta koji su implementirani takođe u okviru *Smart Home* koncepta, dobija se mnogo jeftinije rješenje nego da je korisnik kupovao oba sistema zaštite po na osob. Takođe, na već prethodno pomenuto, dodatkom jednog modula sa, npr. sa

8 digitalnih izlaza čija, cijena, ne prelazi stotinu evra, dodatkom jednog bežičnog modula za upravljanje klima uređajem sa cijenom, takođe, ne većom od stotinu evra, dodatkom jednog bežičnog senzora pokreta, sa cijenom od približno stotinu evra i dodatkom nekoliko bežičnih panela za upravljanje rasvjetom, čija cijena ne prelazi pedesetak evra pojedinačno, dobijamo jedan skroman i vrlo efikasan *Smart Home* sistem čija ukupna cijena sa implementacijom ne prelazi par hiljada evra.

Modul sa 8 relejnih izlaza je sasvim dovoljan da se na njega priljucke alarmna sirena, zujalica i njihovo napajanje (3 kanala), pobuda za automatski prekidač kojim se kontroliše bojler (1 kanal), pobuda za automatski prekidač kojim se kontrolišu sve utičnice opšte potrošnje (1 kanal) a ostatak (3 kanala) se može ostaviti kao rezerva za eventualno proširenje sistema. Blok šema jednog hibridnog koncepta *Smart Home* sistema je data na slici 7.



Slika 7. Blok šema jednog jednostavnog *Smart Home* sistema sa hibridnim načinom izvedbe

## V. ZAKLJUČAK

Opisan je *Smart Home* koncept, njegove karakteristike, svrha i pristup. Zaključujemo da je upotrebom jednog ovakvog sistema poželjno upravljati bar rasvjetom, sistemima grijanja i hlađenja i pripreme tople sanitarne vode kao i utičnicama opšte potrošnje. Na taj način se korisniku pruža mogućnost da svoj dom učini energetske efikasnijim, komfornim i bezbjednim.

Hibridni način izvedbe ima najveći šansu za konkretnu primjenu na već izgrađenim objektima. Takođe, u današnje vrijeme zahvaljujući sve većoj konkurenciji na tržištu, cijena jednog jednostavnog *Smart Home* sistema nije više toliko visoka kao nekada, pa je jedan ovakav sistem postao dostupan široj populaciji.

Daljinski pristup je izuzetno važan u ovom konceptu. Sa njim, ovaj koncept ima pravi smisao. Sama mogućnost daljinskog uključenja i isključenja nekog sistema po želji i potrebi korisnika doprinosi energetskej efikasnosti i komforu.

## LITERATURA

- [1] Filip Rudan, "SIMULACIJSKI MODEL PAMETNE KUĆE", strana 6 - FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA - SVEUČILIŠTE U ZAGREBU, jun 2015.

- [2] dr Velimir Čongradac, Energetska efikasnost u poslovno-stambenim objektima, slajd 9. Fakultet tehničkih nauka, Univerzitet u Novom Sadu.
- [3] U.S. Energy Information Administration, Annual Energy Outlook, December 2014. Dostupno na: <https://www.epa.gov/energy/about-us-electricity-system-and-its-impact-environment>
- [4] U.S. Energy Information Administration, Residential Energy Consumption Survey. Dostupno na: <https://www.eia.gov/consumption/residential/>
- [5] Home Automation Device Protocol (HADP): A Protocol Standard for Unified Device Interactions, Thomas Gonnot, Won-Jae Yi, Ehsan Monsef, Jafar Saniie.
- [6] Buses, Protocols and Systems for Home and Building Automation, Ondřej Nývlt.
- [7] Интернет\_ствари, dostupno na: [https://sr.wikipedia.org/wiki/Интернет\\_ствари](https://sr.wikipedia.org/wiki/Интернет_ствари)
- [8] Custom made aplikacija za upravljanje *Smart Home* sistemom, princip rada se može pogledati na link-u: <https://www.youtube.com/watch?v=sPuMG1Mepgl&feature=youtu.be>

**Abstract** - In this paper I describe different ways of implementing the Smart Home system: wired, wireless and hybrid concept. It has been shown a reasons why large number of manufacturers and users decides for hybrid concept. I also discussed about the problem of remote access for Smart Home systems, which achieves full effectiveness and justify the investment in terms of improving the energy efficiency of the building.

This paper presents an example of a hybrid Smart Home system which can be applied to existing, standard, electrical installation. Also, there was discussion about the custom made applications for managing these systems.

# CA sistem kablovskih televizija - mehanizam kontrole i monitoring

Igor Miljanić, Student, ETF

**Sadržaj** - U radu je analiziran način funkcionisanja CA (eng. *Conditional Access*) sistema, sistema uslovnog pristupa kablovskih televizija, zajedno sa različitim mehanizmima kontrole i praćenja njegovog rada. Kontrolne opcije zajedno sa CA sistemom čine jednu funkcionalnu cjelinu. Takođe objašnjena je opravdanost njegove primjene, kao i posledice koje on proizvodi krajnjem korisniku.

**Ključne riječi** – algoritam, enkripcija, skrembling, zaštita..

## I. UVOD

DA bi shvatili način rada digitalnog televizijskog sistema, dat je pregled njegovih funkcionalnih cjelina. U tu svrhu oslonili smo se na radove relevantnih tijela i autora iz date oblasti.

Današnji operateri digitalnih televizija nude mogućnost prikazivanja programa po želji. Jedna od ključnih komponenti digitalne televizije je sistem uslovnog pristupa, odnosno CA sistem. CA sistemi omogućavaju operateru digitalne televizije enkriptovanje signala[1], čime se signal određenog TV sadržaja ili signal kompletnog kanala blokira za neautorizovane potencijalne korisnike. Naravno, ovakvi sistemi zahtijevaju posebne algoritme zaštite koji dodatno usložnjavaju sistem.

Implementacija enkripcije u CA sistemima ključna je stavka koja određuje da li će se digitalni sadržaj prikazivati ili ne. Sam sistem funkcioniše putem servisa i namijenjen je i za prenos ostalih digitalnih tipova signala, digitalnih podataka, digitalnog radija ili interaktivnih servisa. Nivoi zaštite digitalne televizije zavise od kvaliteta enkripcije koja se obavlja nad skremblovanim signalom. Česte pojave poput poruka o nedostatku signala prave problem samom provajderu usluga od koga se očekuje odgovarajući nivo usluge koji može zadovoljiti krajnjeg korisnika. Prilikom pojave ovih problema operater ne može znati da li je u pitanju problem sa samim uređajem, odnosno risiverom, samom konekcijom ili možda CA sistemom. Problem je što su CA sistemi u tolikoj mjeri zatvoreni da otklanjanje ovakvih problema pravi poteškoće samim operaterima, što na kraju utiče na nezadovoljstvo krajnjih korisnika. Možda nezadovoljstvo krajnjih korisnika nije u toj mjeri izraženo u manjim sredinama poput Crne Gore, ali u sredinama u kojima provajder radi i sa par stotina hiljada korisnika, ovakve stvari operateru donose određene neugodnosti koje mogu

dovesti i do otkazivanja pretplate. Međutim, postoje neki relativno jednostavni koraci kojima je moguće osmotriti i detektovati najčešće probleme koji se javljaju prilikom prenosa digitalnog TV signala. Mnogi digitalni operateri u okviru različitih tehnologija za prenos signala koriste CA monitoring.

## II. PRINCIP RADA CSA ALGORITMA

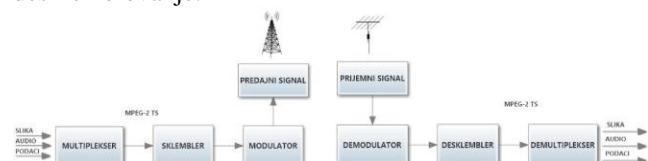
CSA (eng. *Common Scrambling Algorithm*) je algoritam enkripcije namijenjen prenosu digitalne televizije razvijen od strane ETSI-a (eng. European Telecommunications Standards Institute) tokom devedesetih godina. Do 2002. godine CSA je bio strogo čuvana tajna sa ciljem da se CSA implementira u okviru hardvera, čime bi se onemogućili napadi na sam sistem.

U Evropi CA sistemi za satelitski, zemaljski i kablovski broadcast upravo koriste CSA algoritam za enkripciju i dekripciju TV signala. Enkripcija i dekripcija signala[2] naziva se skremblovanje (eng. *scrambling*) i deskremblovanje (eng. *descrambling*) signala. Iako CSA algoritam predstavlja standard za digitalni broadcast, provajderi mogu iskoristiti različite varijante algoritama za enkripciju digitalnog signala.

Proces skremblovanja je proces pri kome zvuk, slika i podaci postaju nečitljivi, u cilju ograničavanja pristupa, odnosno, posmatrajući digitalnu televiziju, onemogućavanje korisnicima pristup sadržaju za koji nemaju pretplatu. Enkripcija je proces zaštite sigurnosnog ključa (kontrolne riječi, jednog od sistema zaštite) kojeg šaljemo sa skremblovanim signalom, pri čemu je zadatak ključa da omogući deskremblovanje signala. Glavna uloga CA sistema pri broadcast-u jeste odlučivanje koji risiveri odnosno STB uređaji mogu primiti određene programske servise ili individualne programe.

CA sistem uglavnom se sastoji od dva podsistema:

- Podsistema za skremblovanje, čija je funkcija skremblovanja signala i onemogućavanje neovlašćenog pristupa TV kanalima, kao i deskremblovanje signala u okviru korisničkog risivera.
- Podsistema za kontrolu pristupa koji obrađuje kontrolne pristupne poruke i određuje da li je potrebno obavljati deskremblovanje. Na sl. 1 je prikazana šema prenosa digitalnog TV signala sa kolima za skremblovanje i deskremblovanje.



Sl. 1. - Šema prenosa digitalnog TV signala



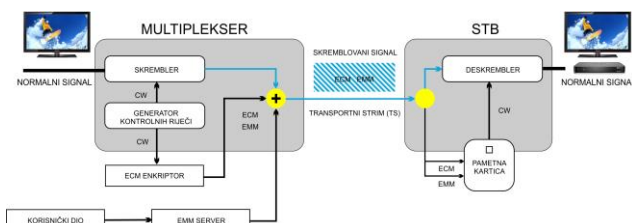
Signal prolazi kroz skrembler, koder koji enkriptuje signal koristeći kontrolnu riječ CW (eng. *Control Word*), dok se enkripcioni ključ generiše nasumično. Signal na izlazu iz kodera je nečitljiv dok se ne unese odgovarajući ključ. Dekoder dekriptuje kodirane signale takođe koristeći CW ključ da bi reprodukovao početni signal. U normalnim DVB sistemima kodiranje se obavlja preko multipleksera, dok se dekodiranje obavlja u samom STB (eng. *Set Top Box*) uređaju. Zbog sigurnosnih razloga kontrolna riječ se mijenja veoma često, uglavnom svakih 10 sekundi, ali i ovaj vremenski interval zavisi od izbora provajdera. Zaglavlje transportnog strima-TS (eng. *Transport Stream*) sadrži dva kontrolna bita koji definišu CW kontrolnu riječ koje treba dekodirati.

Iako signali poput video, audio i ostalih dodataka namijenjeni digitalnom prenosu imaju različita svojstva, moguće ih je sastaviti u jedan digitalni signal olakšavajući proces prenosa podataka. Najčešći sistem koji bi obavljao sastavljanje svih signala baziran je MPEG standardu. U svakom slučaju metode skremblovanja nisu standardizovane, pa samim tim tip skremblovanja zavisi od provajdera, a u nekim slučajevima i od države.

#### A. ECM (Entitlement Control Message)

Najvažniji zadatak CA sistema je slanje ključa dekoderu na što sigurniji način. Ključ, odnosno kontrolna riječ smješta se u kodiranoj ECM[1]poruci implementiranoj u transporni strim. Većina CA sistema koristi sistem pametnih kartica radi sigurnosti, dok neki CA sistemi daju prioritet softverskim komponentama u odnosu na pametne kartice. Pametna kartica prima ECM u okviru signala, izvršava dekriptovanje a zatim šalje kontrolnu riječ STB risiveru radi dekodiranja signala.

Vremenski rok trajanja kontrolne riječi je često ograničen od strane provajdera što dovodi do situacije da se kontrolna riječ mijenja i do nekoliko puta u minuti. Da bi risiver mogao da deskrembluje podatke on u svakom trenutku mora biti informisan o trenutnoj vrijednosti kontrolne riječi.



Sl. 2. - Princip rada ECM-a i EMM-a

Za CA sistem slanje kontrolnih riječi ka STB risiveru veoma je zahtjevan zadatak. Na sl. 2 vidimo da signal prvo ulazi u multiplekser, nakon čega dolazi do skremblera, dok se u generatoru kontrolnih riječi generiše kontrolna riječ. Kontrolna riječ istovremeno šalje se skrembleru i ECM enkriptoru radi procesa enkripcije. Trenutna i sledeća kontrolna riječ šalju se ECM enkriptoru koji nakon enkripcije ECM šalje u multiplekser radi daljeg dodavanja u transporni strim. Sada TS ulazi na ulaz STB-a. Normalni signal bi išao direktno u deskrembler, dok ECM ide direktno na pametnu karticu. Pametna kartica čita ECM, dekriptuje ga i šalje kontrolnu riječ STB uređaju. STB

uređaj prima kontrolnu riječ i vrši deskremblovanje signala što dovodi do izlaznog signala koji nije više enkriptovan, i koji krajnjem korisniku daje regularnu sliku.

Sadržaj ECM poruke nije unikatan već može da varira od provajdera do provajdera, ali uglavnom sadrži sledeće informacije:

- Dvije kontrolne riječi – na ovaj način omogućava se slanje trenutne ali i sledeće ključne riječi radi lakšeg dekodiranja.
- Informacije o datumu i vremenu – pametne kartice na ovaj način mogu da odluče da li je korisniku omogućeno gledanje kanala ili ne.
- Identifikacija kanala – može biti jedinstvena po kanalu ili dijeljena između svih kanala u okviru paketa pretplate.

Identifikator kanala omogućava pametnoj kartici da provjerava da li je kanal dostupan u okviru svoje tabele pristupa.

#### B. EMM (Entitled Management Message)

U CA sistemu EMM enkripcija se koristi radi slanja podataka pametnoj kartici, poput informacija da li je korisniku dozvoljeno gledanje određenog servisa, da li je korisniku dozvoljeno gledanje kanala idućeg mjeseca i slično. STB prima EMM i prosleđuje ih pametnoj kartici radi dalje obrade. Na taj način pametna kartica koristi informacije od EMM-a za ažuriranje internih pristupnih kontrola bazi podataka. Baza podataka sadrži listu kanala i video na zahtjev - VOD (eng. *Video on demand*) koji su dostupni korisniku. EMM može omogućiti pregledanje sadržaja VOD-a u vremenskim intervalima od nekoliko sati do nekoliko mjeseci, kao i omogućavanje gledanja sadržaja na jednom ili više kanala. Kada pametna kartica primi ECM ona taj podatak upoređuje sa svojom bazom podataka čime odlučuje da li korisnik može pristupiti kanalu ili ne. Ukoliko korisnik ima pravo pristupa, dekodirana kontrolna riječ šalje se STB uređaju koja zatim deskrembluje signal.

Već smo vidjeli (sl. 2) da signal, na strani predaje, prvo ulazi u multiplekser a zatim se šalje skrembleru. Skrembler prima CW kontrolnu riječ iz generatora kontrolnih riječi i skrembluje ga tako da su ECM i EMM importovane u transportni strim. Kad god se dozvoli pristup krajnjem korisniku za novi servis EMM se generiše od strane CAS-a i taj novi generisani EMM šalje se multipleksersu radi daljeg importa u transportni strim. Na drugom kraju pametna kartica prima EMM i ECM signal, dekriptuje ga i šalje deskrembleru. Tek nakon što deskrembler primi kontrolnu riječ počinje deskremblovanje signala i na izlazu korisnik dobija traženi servis.

Za razliku od ECM-a, EMM poruke se ne mijenjaju toliko često, već se mijenjanju u mjesečnim intervalima. Naravno, ovaj način mijenjanja EMM-a može biti problematičan za provajdere, jer predstavlja vid zaštite koji može biti meta napada. Svaki provajder može izabrati svoj vremenski interval mijenjanja EMM-a. Sadržaji ECM-a i EMM-a nisu standardizovani i kao takvi zavise isključivo od CA sistema u upotrebi. Kontrolne riječi mogu biti prenošene i kroz različite ECM-ove odjednom, omogućavajući korišćenje različitih CA sistema.

EMM poruke enkriptovane su metodama poznatim isključivo CA provajderu. Ipak, poznato je da sadrže

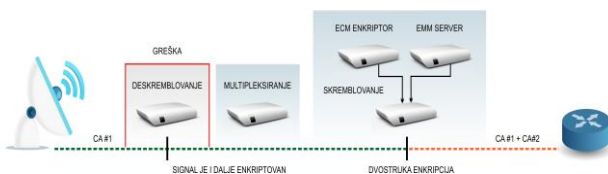
informacije poput dužine polja EMM-a, polje za adresu, kao i enkriptovani dio.

### III. KONTROLE GREŠAKA ULAZNOG SIGNALA

Većina kablovskih i IPTV operatera prima značajni dio ulaznih signala u enkriptovanoj formi, najčešće preko satelita. To znači da će kanali dolaziti od različitih provajdera, pri čemu će svaki provajder imati različite CA sisteme.

Prvi korak pri radu digitalne televizije je dekodiranje ulaznih signala koje se uglavnom obavlja direktno u multiplekserima.

Ukoliko ulazni signal iz nekog razloga ne može biti deskremblovan, biće blokiran za korisnika. Tada imamo situaciju da će signal ponovo, drugi put, biti enkriptovan operatorovim CA sistemom (sl. 3). Korisnicima ovakav signal prikazuje se kao "crna slika".



Sl. 3. - Deskremblovanje ulaznog signala nije uspjelo i enkriptovani signal se dovodi do skremlera.

Problemi sa skremblovanjem najčešće se javljaju zbog hardverskih grešaka ili problema sa ECM i EMM porukama. Da bi se ovi problemi izbjegli poželjno je izvršiti monitoring signala nakon deskremlinga, u momentima prije nego što se signali enkriptuju. Najjednostavniji pristup bi bio provjera skremblovanih kontrolnih bita transportnih strim paketa. Da bi se ipak monitoring izvršio na što pravilniji način potrebno bi bilo detaljno analizirati dolazni signal, naročito video PID (eng. *Proportional Integral Derivatv*). Najsigurniji način provjere bi bio sama provjera da li je omogućeno audio i video dekodiranje, i u slučajevima kada je to moguće da se generiše kratak prikaz audio i video signala radi pravilne procjene. Pored ovog načina kontrole grešaka poželjna je i detekcija signala pri zamrznutoj slici ili detekcija problema poput "crnog ekrana" radi što preciznijeg monitoringa.

#### A. Kontrola i monitoring EMM-a i ECM-a

EMM poruke neprekidno se prenose od CA sistema prema multiplekseru koji EMM poruke unose u transportni strim. Ovo se radi putem mrežnih protokola, preko IP adrese i UDP/TCP protokola. U realnim situacijama monitoring EMM-a sa ciljem detekcije svih problema nije moguće isprojektovati. Dodavanje opreme koja generiše alarme u slučajevima kada EMM bitrate pada ispod određenog nivoa omogućuje značajniju zaštitu, samim tim detekciju i rješavanje grešaka prije nego što te greške dođu do krajnjeg korisnika.

Prilikom prebacivanja na kanal, STB uređajima su potrebne kontrolne riječi radi deskremlovanja sadržaja određenog kanala ali i radi dekodiranja istog. Ovo znači da je transmisioni interval ECM-a veoma važan bez obzira na vrijeme prebacivanja između kanala. Netačan interval

ponavljanja kod ECM-a uglavnom je uzrokovan netačnom konfiguracijom multipleksera ili internih problema u okviru rada multipleksera. Ukoliko ECM ne postoji STB uređaj će biti u nemogućnosti da deskremluje signal čime se korisniku na prijemu prikazuje poruka tipa 'Nema signala'. Do greške nedostajućih ECM-ova uglavnom dolazi zbog greške CA sistema i ECM enkriptora, pogrešne konfiguracije multipleksera, mrežne greške između multipleksera i CA sistema.

Da bi se provjerilo ECM ponavljanje potrebno je pronaći sve ECM-ove za sve servise u signalu. ECM-ovi se uglavnom signaliziraju tako što ostavljaju CA deskriptor u program info sekciji PMT tabele (eng. *Program Map Table*). Ovo omogućava da se isti ECM prenese na PID komponente servisa. Kada se pronađe lista ECM-ova ona treba da primi sve podatke, različite vrste ECM PID-ova i da provjeri interval između različitih ECM paketa.

#### B. Monitoring kontrolne riječi

Kontrolna riječ koja se koristi za enkripciju transportnih stream paketa konstantno se mijenja u vremenskim intervalima. Ukoliko kontrolne riječi ne budu promijenjene ECM može iskoristiti postojeće kontrolne riječi radi dekripcije signala dokle god se ovaj problem ne riješi. U ovakvim situacijama krajnji korisnici mogu imati pristup kanalu čak iako im je operater zabranio pristup.

Gubitak kontrolnih riječi najčešće je prouzrokovan:

- Problemom između multipleksera, skremlera i CA sistema. Ovi problemi se najčešće dešavaju kad je CA sistem fizički udaljen od multipleksera, npr. prilikom VPN (eng. *Virtual Private Network*) konekcije.
- Problemom sa ECM enkriptorom CA sistema. Ovo može biti prouzrukovano hardverskim i softverskim greškama, samim operativnim sistemom ili greškama CA aplikacija koje se nalaze na serveru.
- Problemom sa multiplekserom skremlera. Ovi problemi najčešće se javljaju progrednom konfiguracijom uređaja ili softverskim problemima na samom multiplekseru.

Monitoring kontrolnih riječi, odnosno njegovog mijenjanja omogućava operateru detekciju ovih situacija čak i prije nego što se pojavi problem. U sistemima bez monitoringa ovi problemi mogu potrajati nedjeljama ili mjesecima.

Nakon što signal prođe multiplekser koji obavlja skremblovanje postavlja se set PID-ova[3] koji moraju biti skremblovani u svakom trenutku kao i set PID-ova koji se ne skremluje (čisti PID-ovi). Da bi se obavile ove promjene dovoljno je provjeriti kontrolne bite skremlovanja. Oprema za monitoring trebala bi da ima mogućnost za pregled PID-ova odnosno za deskremlovanje.

#### C. Monitoring EMM povratnog vremena

Parametar koji je od velikog značaja za provajdere je povratno vrijeme EMM-a. Provajder uglavnom želi da iskoristi što manje raspoloživog opsega za EMM saobraćaj. Veliki provajderi imaju oko 100 hiljada

iskoristi što manje raspoloživog opsega za EMM saobraćaj. Veliki provajderi imaju oko 100 hiljada pametnih kartica, što povlači situaciju gdje je potrebno izvršiti autorizaciju EMM-ova za kanale za koji su korisnici prijavljeni. EMM-ovi se uglavnom šalju mnogo ranije tako da za STB uređaje koji su uvijek priključeni na mrežu povratno vrijeme EMM-a nije od velikog značaja. Međutim, povratno vrijeme EMM-a je bitno u situacijama kada je STB ugašen na duži vremenski period, jer tada pametna kartica mora biti ažurirana prije deskremlovanja signala. Ovaj vremenski period traje dosta dugo (ova procedura bi trebala najduže da traje oko 15 min), što dovodi do žalbi korisnika, što bi svaki provajder nastojao da izbjegne. EMM playout ima različite prioritete za EMM-ove, tako da u situacijama gdje korisnik kupi pristup novom setu kanala EMM-ovi se uglavnom stavljaju na veći prioritet i samim tim se ponavljaju češće. CA sistem provajderi ne žele da dijele detaljnije informacije o EMM-ovima, tako da je za monitoring povratnog vremena EMM-a potrebno znati njegov format bez enkripcije. To znači da ukoliko bi trebalo analizirati povratno vrijeme EMM-a potrebno je postići dogovor sa provajderom kao i doći do specifikacije STB uređaja koje provajder koristi.

#### *D. Provjera stanja ECM-a u odnosu na CW promjene*

U svakom PID zaglavljju postoje dva bita koja dekeru govore da li treba da koriste parne ili neparne kontrolne riječi. ECM uglavnom sadrži dvije kontrolne riječi, što mu omogućava da istovremeno nosi obje kontrolne riječi koje se trenutno koriste, kao i kontrolne riječi koje će se koristiti za skremlovanje prilikom iduće promjene kontrolne riječi[4]. Ovo omogućava STB uređaju da uvijek ima na raspolaganju kontrolne riječi koje su potrebne za deskremlovanje sadržaja. Ukoliko promjene ECM-a nisu sinhronizovane sa promjenom kontrolnog ključa koji se koristi za enkripciju signala, može se desiti da STB uređaj neće biti u stanju da deskremluje signal, čime dolazi do pojave „crne slike”.

ECM ne bi smio da se mijenja prerano, jer STB uređaj treba imati dovoljno vremena da deskremluje sve TS pakete enkriptovane sa predhodnim skremling ključem, prije nego što se sam ključ promijeni. Takođe, ECM ne bi smio ni da se mijenja prekasno, jer tada STB uređaj mora imati dovoljno vremena da dobije kontrolne riječi od pametne kartice i dovoljno vremena da ove informacije sačuva, prije nego što počne skremlovanje signala.

## IV. ZAKLJUČAK

Pregledom postojećih tehnologija u radu CA sistema, zajedno sa pratećim komponentama za njegovu kontrolu dolazi se do zaključka da su digitalizacijom omogućene brojne prednosti u odnosu na još uvijek postojeće, analogno stanje. Da bi sve ovo funkcionisalo kako treba i pružilo sve bitne funkcije koje digitalna televizija donosi potrebno je ispoštovati određene standarde i procedure zaštite signala, a sve u cilju što sigurnije razmjene podataka između provajdera i krajnjeg korisnika. Kontrolom i monitoringom CA sistema provajderi mogu omogućiti što kvalitetnije usluge digitalne televizije svojim korisnicima. Iako sistemi kontrole i monitoringa zahtijevaju određene resurse, prije svega finansijske, njihova implementacija, dugoročno gledano, može doprinjeti razvijanju kvalitetnijih provajderskih servisa koje mogu prevazići trenutno dostupne okvire digitalne televizije.

## LITERATURA

- [1] ETSI TS 102 367 (2005, januar). DAB Conditional access. Dostupno:[http://www.etsi.org/deliver/etsi\\_ts/102300\\_102399/102367/01\\_01\\_01\\_60/ts\\_102367v010101p.pdf](http://www.etsi.org/deliver/etsi_ts/102300_102399/102367/01_01_01_60/ts_102367v010101p.pdf)
- [2] CS265 Section 2 (2002, januar). Digital Video Broadcasting Conditional Access Architecture. Dostupno: <http://www.cs.sjsu.edu/~stamp/CS265/projects/papers/dvbca.pdf>
- [3] ETSI TS 102 978 (Jul, 2008). Digital Audio Broadcasting DAB Dostupno:[http://www.etsi.org/deliver/etsi\\_ts/102900\\_102999/102978/01\\_01\\_01\\_60/ts\\_102978v010101p.pdf](http://www.etsi.org/deliver/etsi_ts/102900_102999/102978/01_01_01_60/ts_102978v010101p.pdf)
- [4] N. Helberger, “Controlling access to content : regulating conditional access in digital broadcasting.” Universiteit van Amsterdam, 2005, pp. 50–70.

## ABSTRACT

In this work a brief analysis of CA system within cable operated providers is presented with key parts of CA system described, as well key elements designed for monitoring and control of all important functions that make CA one functional unit. This paper also analyze impact of CA system and its components on end users.

### **CA SYSTEM WITHIN CABLE OPERATED PROVIDERS - A MECHANISM OF CONTROL AND MONITORING**

Igor Miljanić

# LED rasvjeta tunela

## primjer tunel Budoš na putu Podgorica Nikšić

Vladimir Radulović, Milovan Radulović, *Member, IEEE*

**Sadržaj** — U radu je razmatrana problematika osvjjetljenja putnih tunela. Sa aspekta energetske uštede neophodno je primijeniti izvore svjetla koji imaju najbolju svjetlosnu iskoristivost. Potrebno je ispuniti i zahtjeve standarda koji su normirali svjetlotehničke parametre unutar tunela i u prilaznoj i odlaznoj zoni. Savremena tehnologija proizvodnje LED (light emitting diode) izvora svjetlosti dostigla je nivo svjetlosne iskoristivosti koji je znatno veći od konvencionalnih izvora. Samim time nameće se osnovni razlog primjene Led izvora kao rješenja za što manji utrošak električne energije. U radu je prikazana metodologija proračuna i zahtjevi standarda koji se moraju ispuniti pri realizaciji tunelske rasvjete. Kao primjer prikazana je implementacija Led rasvjete u Tunelu Budoš na magistralnom putu Podgorica Nikšić. Tunel Budoš je prvi tunel u Crnoj Gori koji je osvijetljen primjenom LED tehnologije.

**Ključne reči** — energetska efikasnost, Led rasvjeta, tunelsko osvjjetljenje.

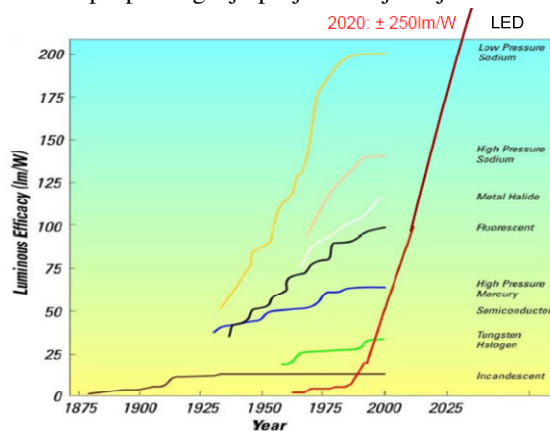
### I. UVOD

DRUMSKI prevoz igra veoma važnu ulogu u svakodnevnom životu. U cilju obezbjeđenja sigurnosti saobraćaja neophodno je da saobraćajnice, zavisno od njihove kategorije, ispunjavaju uslove definisane odgovarajućim standardima. Kao jedan od važnih segmanata kompleksnih saobraćajnica ističe se pitanje veoma zahtjevnog putnog objekta kao što je tunel. U ovom radu analiziraćemo samo jedan segment tunelske infrastrukture koji ima za cilj da obezbijedi što komforniju i sigurniju upotrebu sa aspekta vizuelnog komfora vozača. Pravilnim projektovanjem tunelske rasvjete neophodno je ostvariti takav kvalitet osvjjetljenja koji će omogućiti vozačima adaptaciju između vanjske i unutrašnje sjajnosti površine kolovoza i osvijetljenosti ne narušavajući njihovu koncentraciju i sigurnost [1], [2], [3].

Realizacija različitih nivoa osvijetljenosti tokom dana i noći predstavlja poseban problem koji iziskuje upotrebu savremenih sistema automatske regulacije. Za razliku od osvjjetljenja tunela tokom noći, što predstavlja relativno prost zadatak, dnevno osvjjetljenje tunela se mora analizirati mnogo kompleksnije. Tokom noći je neophodno održati nivo rasvjete koji omogućava kako u prilaznom

dijelu, tako i u unutrašnjosti tunela, adekvatno praćenje rute, uvažavajući uslove definisane standardima za različite kategorije saobraćajnica, kao i drugih svjetlotehničkih veličina osim sjajnosti kolovoza. Promjena uslova prouzrokovanih geografskom orijentacijom, meteo situacijom i topologijom zemljišta na ulaznim ili izlaznim portalima tunela zahtijeva posebnu pažnju pri analizi i projektovanju rasvjete prilaznim saobraćajnicama i pojedinim zonama unutar tunela [3].

Pored energetske efikasnosti i ovako zahtjevnih uslova i promjene režima rada daju nesagledivu prednost primjeni LED izvora svjetlosti. pri rješavanju navedenih pitanja. U današnjem trenutku cijena LED izvora svjetlosti opada, energetska iskoristivost (Sl.1.) raste, a upravljivost izvora svjetla ovog tipa je mnogo jednostavnija nego kod konvencionalnih izvora svjetla. Navedeni razlozi opredjeljuju proizvođače da nastave sa poboljšanjima karakteristika, a projektante da se opredjeljuju za izbor LED izvora pri predlaganju projektnih rješenja.



Sl. 1. Svjetlosna iskoristivost izvora svjetlosti [4]

### II. TUNELSKA RASVJETA

Sigurnost drumskog saobraćaja usko je povezana sa kvalitetnom rasvetom najkritičnijih objekata na putevima, a to su tuneli. Ograničavajuće karakteristike kontinualnog održavanja tunela kao i drugačiji vremenski uslovi unutar samih tunela, zahtijevaju ispravno funkcioniranje rasvjete.

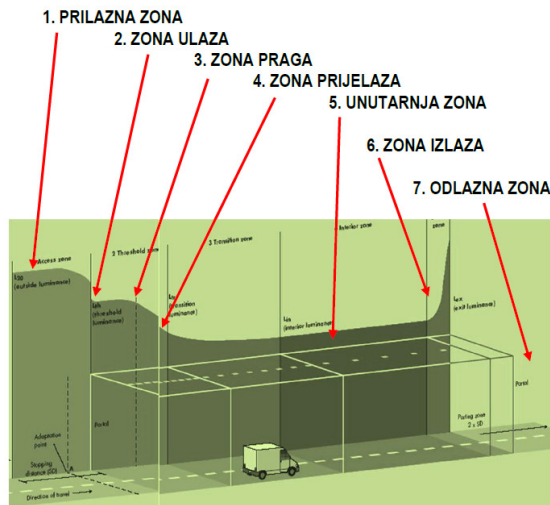
Primaran cilj tunelske rasvjete je da omogući siguran i pouzdan ulaz, prolaz i izlaz iz tunela. Da bi ispunili ove uslove neophodno je izborom odgovarajuće tunelske rasvjete omogućiti brzo prilagođavanje vozača na svjetlo unutar tunela, identifikaciju potencijalno opasnih prepreka, kao i neometan prolaz motornih vozila kroz tunel, bez smanjenja brzine kretanja vozila. Ovi se zahtjevi ogledaju na jedan način tokom dana kada je kontrast jačine svjetlosti između unutrašnjosti tunela i vanjskog prostora

Vladimir Radulović, (autor za kontakte), D O O Taving, Milovana Pekovića 1, 84300 Nikšić, Crna Gora (telefon: 382-67-212002; e-mail: radulovicv@mail.ru).

Milovan Radulović, Elektrotehnički fakultet u Podgorici, Bulevar Džordža Vašingtona bb, 81000 Podgorica, Crna Gora, (telefon 382-67-313309, e-mail: milovanr@.ac.me).

veliki, kao i na drugi način tokom noći, kada vrijedi inverzni režim od dnevnog.

Zbog izloženog, pri planiranju tunelske rasvjete treba obratiti pažnju na potrebu vizualnog prilagođavanja i na 7 ključnih zona rasvjete tunela kako prikazuje Sl. 2.



Sl. 2. Zone tunelske rasvjete

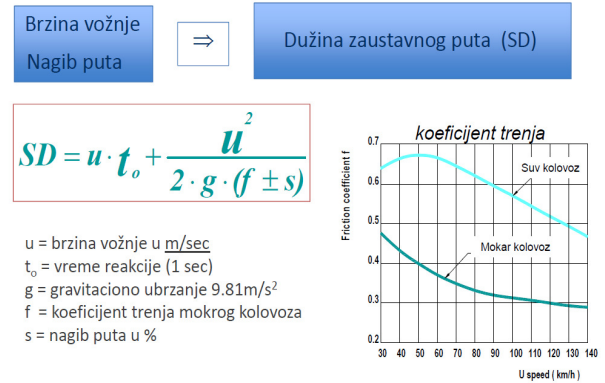
Jedna od metoda za proračun vizuelnog prilagođavanja je tzv. „L20 metoda“. Ona upoređuje prosječnu sjajnost okoline, neba i kolovoza u vizuelnom konusu pod uglom od 20° koji je centriran na vizuelnoj liniji vozača na samom početku pristupne zone [4], [5].

Vizualno prilagođavanje između visoke i niske sjajnosti kolovoza prilikom vožnje nije trenutna pojava. Uzrok tome je prostorno i vremensko prilagođavanje. U slučaju prostornog prilagođavanja velika razlika sjajnosti između unutrašnjosti i vanjskog prostora tunela uslovljava adaptaciju čula vida. Pojava efekta „crne tačke“ prouzrokuje kod vozača osjećaj neprijatnosti i nesigurnosti. Kod vremenskog prilagođavanja važno je to što je ljudskom oku potrebno više vremena da se navikne prilikom promjene iz svjetlije u tamno nego li u obrnutom slučaju. Vrijeme potrebno da se sama adaptacija vida izvrši tretirano je kao kritičan činilac [3]. Obje adaptacije vezane su za promjenjivi sadržaj definisanog vidnog polja vozača na koji utiču slijedeći parametri: svjetla površina neba, površina asfalta dionice puta ispred vozača, okolina oko ulaza u tunel, okolina puta ispred vozača, otvor i portal ulaza tunela.

Na vrijednost sjajnosti prilazne zone, utiče: direktno sunčevo isijavanje, konstrukcija ulaznog portala, okolina izgrađenog portala (zasađenost rastinja), bočne strane tunela u usjeku i procenat prisutnog neba, konfiguracija terena (ravničarski ili planinski), površina i trasa pristupnog puta u dužini najmanje 200 m prije ulaska, površina puta u dužini zone praga u tunelu, klimatski uslovi, vegetacija i okolina (kiša, snijeg, magla, zelenilo,...), kao i svijetli sloj atmosfere između položaja vozača i ulaza u tunel.

Kako vizuelna adaptacija ne bi uticala na sigurnost upravljanja vozilom potrebno je „amortizovati“ istu adekvatnom rasvjetom. Posebno je potrebno posvetiti pažnju rasvjeti prilazne i ulazne zone tunela. Ulaskom u tunel, ulazna zona može se uporediti sa zaustavnim putem.

Na Sl. 3 prikazane su relacije i parametri neophodni za proračun zaustavnog puta. Zaustavni put predstavlja hipotetičku distancu potrebnu da bi vozač u slučaju nekog problema mogao adekvatno reagovati uzimajući u obzir njegovu moć percepcije, vrijeme reakcije i površinu kolovoza [4].



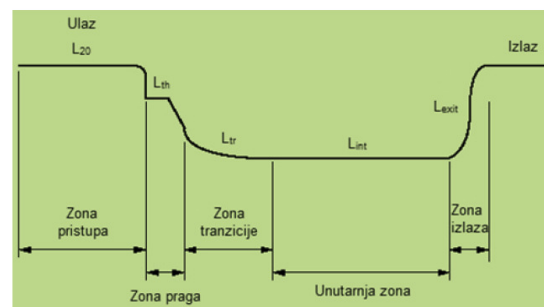
Sl. 3. Postupak proračuna zaustavnog puta [4]

Osvjetljenje ulazne zone tunela mora biti tako projektovano da vozač, u svakom trenutku, može da uoči bilo kakvu prepreku, sa rastojanja od tunelskog portala koje odgovara njegovoj mogućnosti da reaguje i zaustavi vozilo.

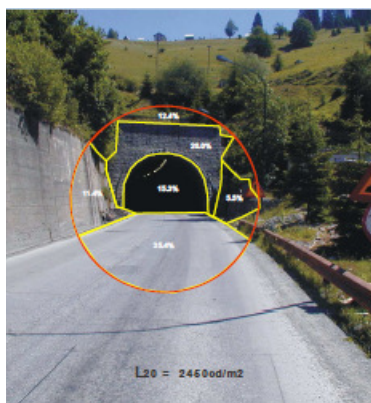
Postupak određivanja nivoa sjajnosti u pojedinim zonama tunela, prikazanim na Sl. 4., predstavlja zadatak formiranja ulaznih podataka na osnovu kojih projektant određuje tip svetiljki i njihov optimalan raspored. Kratak opis formiranja ulaznih parametara dat je u nastavku:

$$L_{20} = \gamma L_C + \rho L_R + \varepsilon L_E + \tau L_{th} \quad (1)$$

gdje je: L<sub>20</sub> – sjajnost prilazne zone, L<sub>C</sub> - sjajnost neba, γ - procenat neba u vidnom polju, L<sub>R</sub> - sjajnost puta, ρ - procenat puta u vidnom polju, L<sub>E</sub> - sjajnost okruženja, ε = procenat okruženja u vidnom polju, L<sub>th</sub> - sjajnost na ulazu u tunel, τ = procenat ulaza u vidnom polju, pr čemu važi γ + ρ + ε + τ = 100 %. Pod vidnim poljem se podrazumijeva baza konusa formirana prema dužini zaustavnog puta. Sl. 6 prikazuje postupak određivanja procentualnih koeficijenata na osnovu fotografije ulaznog portala tunela. Pored ovog eksperimentalnog pristupa, koji se najčešće koristi pri rekonstrukciji postojećeg sistema rasvjete, postoji i teorijski pristup koji se primjenjuje pri projektovanju rasvjete tunela u izgradnji. Određivanje sjajnosti iz relacije (1). vrši se na osnovu parametara datih u Tabeli I.



Sl. 4. Nivoi sjajnosti u pojedinim zonama tunela [1]



Slika 5. Eksperimentalno određivanje sjajnosti prilazne zone na osnovu fotografije (primjer tunel Lokve ulaz od Berana V=60km/h, SD=60m) [4]

TABELA I ODREĐIVANJE SJAJNOSTI PRILAZNE ZONE L 20 [5]

Smjer voznje	L <sub>C</sub> (nebo) kcd/m <sup>2</sup>	L <sub>R</sub> (put) kcd/m <sup>2</sup>	L <sub>E</sub> (okruženje) kcd/m <sup>2</sup>			
			Stijene	Zgrade	Snijeg	Vegetacija
N-sjev.	8	3	3	8	15 (V,H)	2
E - W	12	4	2	6	10 (V)	2
S - jug	16	5	1	4	15 (H) 5 (V) 15 (H)	2

Napomena: V se odnosi na vertikalne, H na horizontalne površine

TABELA II / CIE 88: 2004 I CEN/TR 14380: 2003 – ANEKS A1 [5]

Brzina (km/h)	$k = L_{th} / L_{20}$
< 60 km/h	0.05
80 km/h	0.06
100 km/h	0.08
120 km/h	0.10

TABELA III / CEN/TR 14380: 2003 – ANEKS A2 [5]

zaustavni put (m)	60	100	160
klasa tunela			
4	0.05	0.06	0.10
3	0.04	0.05	0.07
2	0.03	0.04	0.05
1	ne zahtijeva se proračun		

Na osnovu određene sjajnosti prilazne zone (zone pristupa) određuje se sjajnost zone praga L<sub>TH</sub> primjenom Tabele II Korekcija na osnovu dozvoljene brzine i gustine saobraćaja vrši se pomoću Tabele III

Određivanje potrebnog nivoa sjajnosti kolovoza unutrašnje zone L<sub>in</sub> vrši se pomoću Tabele IV.

TABELA VI VRIJEDNOST SJAJNOSTI U UNUTRASNJOJ ZONI CEN/TR 14380: 2003 ANEX A1, CIE 88:2004 [5]

DUGI TUNELI		
Zaustavni put (m)	Protok saobraćaja	
	Nizak	Visok
160 m	6 cd/ m <sup>2</sup>	10 cd/ m <sup>2</sup>
60 m	3 cd/ m <sup>2</sup>	6 cd/ m <sup>2</sup>

Ovim postupkom određuje se dio ulaznih parametara koji pružaju mogućnost projektantu da primjenom odgovarajućeg softvera pristupi proračunu osvjjetljenja za izabrani tip svjetiljke. Raspored svjetiljke i njihov broj uslovljen je ne samo ispunjavanjem prikazanih uslova nego i drugim svjetlotehničkim parametrima koji utiču na kvalitet rasvjete u tunelu.

### III. OSVJETLJENJE TUNELA BUDOŠ

Projekat rekonstrukcije instalacije osvjjetljenja tunela „Budoš” realizovan je u cilju izmjene dotrajale instalacije osvjjetljenja koja je vremenom izgubila svoju pravu funkciju pa njeno investiciono održavanje u daljem periodu nije imalo smisla [6].

#### A. Rješenje prije rekonstrukcije

U ranijoj instalaciji osvjjetljenja tunela bilo je postavljeno 300 svjetiljki. Kao izvore svjetlosti primijenjene su natrijumove sijalice niskog pritiska različitih snaga i broja (2x180W, 2x90W, 2x55W). Ukupna instalisana snaga iznosila je P<sub>intunel</sub> = 76480 W. Uključujući rasvjetu prilaznih zona, instalisana snaga se uvećava, pa ukupna instalisana snaga rasvjete objekta iznosi P<sub>inobjekta</sub> = 78800 W [6].

Postojeći način upravljanja rasvjetom predviđao je rad u tri dnevna režima, koji zavise od vanjske (dnevne) osvjjetljenosti, i jedan noćni režim.

Dnevni režimi rada su dati kao:

- 100 %-ni režim rada,
- 50 %-ni režim rada,
- 17 %-ni režim rada.

Komandom pripadajućih fotoreleja (po tri sa obje strane tunela) upravljalo se dnevnim režimom rada :

- 17 %-ni - pri spoljnoj osvjjetljenosti iznad 2 000 lx,
- 50 %-ni - pri spoljnoj osvjjetljenosti iznad 7 000 lx,
- 100 %-ni - pri spoljnoj osvjjetljenosti iznad 24 000 lx.

Posebni fotoreleji uključuju osvjjetljenje prilaznih dionica kada je vanjski osvjjetljaj manji od 10 lx.

#### B. Novo projektovano resenje tunela „Budoš“ [6]

Sa aspekta napajanja električnom energijom zadržano je osnovno napajanje instalacija osvjjetljenja. Prije provođenja proračuna bilo je potrebno, u skladu sa standardima odrediti potrebne ulazne veličine. Kako je dozvoljena brzina kretanja vozila 60 km/h, a nagib puta 0% određena je dužina zaustavnog puta, koja je u ovom slučaju oko 60m [6].

Proračun nivoa sjajnosti prilazne zone je izvršen u projektnom birou firme „Minel – Schreder“, korišćenjem fotografija i tabela sadržanih u preporukama CIE. Dobijene su vrijednosti za: sjeverni prilaz od L<sub>20</sub> = 3.657 cd/m<sup>2</sup> i južni prilaz od L<sub>20</sub> = 3.396 cd/m<sup>2</sup>

Nivo sjajnosti zone praga se izračunava preko formule:

$L_{th} = k \times L_{20}$ , gdje je k – koeficijent zavisen od dužine zaustavnog puta i primjenjenog sistema osvjjetljaja (dat po preporukama CIE). Sjajnost zone praga može biti konstantna duž čitave zone, ili opada u drugoj polovini zone (do 40 % sjajnosti u prvoj polovini zone) što je i primjenjeno u ovom slučaju radi racionalnijeg rešenja osvjjetljenja. U slučaju tunela „Budoš“ k = 0,05, pa dobijamo L<sub>th</sub> = 182,85 cd/m<sup>2</sup> za sjeverni ulaz, odnosno L<sub>th</sub> = 169,80 cd/m<sup>2</sup> za južni ulaz.

Usvojeni raspored baznih svjetiljki i adaptacionih svjetiljki obezbjeđuje nivo zone praga od 189,8 cd/m<sup>2</sup>, a ona se na kraju zone praga smanjuje na 40 % vrijednosti.

Nivo sjajnosti prelazne zone kreće se od nivoa sjajnosti na kraju zone praga, do nivoa sjajnosti unutrašnje zone. Dat je relacijom:

$$L_{tr} = L_{th}(1,9 + t) - 1,4 \quad (2)$$

gdje je  $L_{tr}$  minimalna dozvoljena sjajnost tranzicione zone u koju vozač dopijeva poslije vremena  $t$  (sec). Smanjenje sjajnosti u ovoj zoni se, u praksi, postize uz pomoć nekoliko skokovitih promjena, pri čemu tako dobijena stepenasta kriva mora biti iznad teorijske krive („CIE kriva“) definisane relacijom (2).

Minimalna sjajnost unutrašnje zone (konstantna sjajnost) uslovljena je dozvoljenom brzinom vožnje i gustine saobraćaja. Prema preporukama predviđena srednja vrijednosti sjajnosti unutrašnje zone je 3 cd/m<sup>2</sup>.

Minimalna osvjetljenost tunela, cijelom njegovom dužinom, pri radu sigurnosnog osvjetljenja iznosi 10 lx. Sigurnosno osvjetljenje se obezbjeđuje korišćenjem svjetiljki baznog osvjetljenja (noćno osvjetljenje), sa kojima se postiže sjajnost od 3 cd/m<sup>2</sup>, s tim što će fluks tih svjetiljki biti smanjen na 50 %.

Za osvjetljenje u tunelu namjenjene su svjetiljke u funkciji baznog i adaptacionog osvjetljenja. Kao bazne usvojene su svjetiljke: GL2 C/5102/80LEDs@500 mA/Symetric (51 kom) instalisane snage 127W, sa relejem za smanjenje snage (za 40%) i svjetlosnog fluksa (za 50%), da bi se koristila i za sigurnosno osvjetljenje tunela (pri ispadu distributivnog napajanja). Za adaptaciono osvjetljenje usvojene su svjetiljke: GL2C/5102/96LEDs@700 mA/Symetric (206 kom) instalisane snage 226W i GL2 C/5102/80LEDs@700 mA/Symetric (12 kom) instalisane snage 190W. Simulacija je izvršena na softveru za dizajn osvjetljenja Ulysse firme Schreder

Kompletna instalacija osvjetljenja (uključenje isključenje pojedinih svjetiljki) ima više režima rada:

- režim 100 % rada - rade sve bazne svjetiljke ( puna snaga i fluks) i sve adaptacione svjetiljke;
- režim 75 % rada -sve bazne svjetiljke ( puna snaga i fluks) i 162 adaptacione svjetiljke
- režim 50 % rada - rade sve bazne svjetiljke ( puna snaga i fluks) i 104 adaptacione svjetiljke
- režim 25 % rada - rade sve bazne svjetiljke ( puna snaga i fluks ) i 50 adaptacionih svjetiljki
- režim nocnog rada - rade sve bazne svjetiljke ( puna snaga i fluks) i sve adaptacione svjetiljke su isključene.

Svjetlosnim intezitetom LED svjetiljke se može upravljati na dva načina. Prvi način se zasniva na promjeni jačine struje kroz LED diodu, dok drugi način koristi impulsno napajanje LED diode u kratkim vremenskim intervalima (Pulse Width Modulation –frekvencije 200Hz). Prednost LED rasvjete je gotovo trenutni odziv na promjenu napona, te korisnici ne primjećuju treperenje LED rasvjete.

Realizovanim rješenjem u odnosu na godišnju cijenu električne energije prije rekonstrukcije od 8.454,60 €,

dolazimo do nove cijene od 4.897,96 €. Iako je instalisana snaga smanjena oko 16%, zahvaljujući realizovanom upravljanju ostvarene su energetske uštede od oko 42%

#### IV. ZAKLJUČAK

Osnovna prednost LED izvora svjetla je u načinu njihove ugradnje u svjetiljke, gdje zbog karakteristike zračenja izvora svjetla postižu dvostruko veću iskoristivost, pa se ostvaruju niži troškovi utrošene električne energije. Druga prednost je bolja regulacija nivoa osvjetljenosti čime se postiže veći stepen sigurnosti i komfora. Zbog mogućnosti regulacije ova tehnologija može na jednostavan način odgovoriti uslovima standarda. Treća prednost je duži životni vijek čime se takođe postiže ušteda.

Date prednosti trasiraju put LED tehnologiji kao prvom izboru prilikom projektovanja tunnelske rasvjete. Navedene karakteristike će izazvati smanjenje pogonskih troškova korisnika javne rasvjete, gubitaka u distributivnoj mreži, ali je neophodno razvijati metode prigušenja i izbjegavanja pojave viših harmonika u distributivnoj mreži.

U radu je prikazana prva realizacija tunnelske rasvjete primjenom LED tehnologije u Crnoj Gori koja pokazuje prednosti i opravdanost primjene navedene tehnologije osvetljenja objekata putne infrastrukture.

#### LITERATURA

- [1] Canan Perdahci, Tunnel Lighting with LED, 7<sup>th</sup> International Ege Energy Symposium, June, Turkey, 2014.
- [2] Li Shuguang, An optimal model for tunnel lighting control systems, Journal Tunneling and Underground Space Tehnology, Elsevier, vol. 49, pp 328-335, 2015.
- [3] M. Kostić, "Vodic kroz svet tehnike osvetljenja", Minel Schreder Beograd, 2000
- [4] Ana Drndarević, Minel -Schröder, Beograd, Osnovne smernice za projektovanje tunnelskog osvetljenja, predavanje u ICG, 2015.
- [5] CEN, "Road Lighting", EN 13201, 2003
- [6] Vladimir Đurišić, Glavni projekat rekonstrukcije instalacije osvetljenja tunela Budoš, "S.P.A.K" d.o.o. – Podgorica, 2015

#### ABSTRACT

In this paper we've discussed problems of road tunnels lighting. In terms of energy savings, it is necessary to apply light sources which have the best light efficiency. It is needed to realize the standardized requirements for the light-technical parameters inside the tunnel and in the ingoing and outgoing zone. Modern technology of production of LED light sources has reached the level of luminous efficiency for these sources, which is much greater compared to conventional sources. Therefore, the primary reason for the application of LED sources as a solution for the smaller consumption of electricity is forced on. The budget methodology and standards requirements that must be met during the realization of tunnel lighting have also been presented in this paper. As an example, the implementation of LED lighting in the road tunnel Budoš along main road Nikšić-Podgorica has been presented. Budoš tunnel is the first tunnel in Montenegro that is illuminated by LED technology.

#### TUNNEL LIGHTING WITH LED

Example of road tunnel Budoš along road Podgorica-Nikšić

Vladimir Radulović, Milovan Radulović

# A New Method For Synchrophasor Estimation

Ivan Jokić, Žarko Zečević, Zdravko Uskoković, Milovan Radulović and Božo Krstajić

**Abstract** — As a highly nonlinear system, power system changes its operational conditions continuously. With regard to constant growth of demand for electrical energy, power systems have become more and more complex. That makes the stability of the grid as a high challenge, and priority in energetics. Synchronized phasor measurement and Phasor Measurement Units (PMUs) have become a subject of interest in modern power systems. Their primarily function is wide area monitoring. The algorithm proposed in this paper uses a phase voltage/current measured signal for estimating the parameters of the fundamental component. Simulation results conducted in the paper show that proposed algorithm has better performances, compared to the considered ones.

**Keywords** — complex circular signal, synchrophasor, synchronized phasor measurement

## I. INTRODUCTION

RECENT development of the power systems, caused by constant increasing demand for electrical energy, and consequently, by growth of production capacity of the renewable sources of energy, requires improvements in the control and maintenance of the system.

Synchronized monitoring of the relevant parameters, both at local, and regional (global) level, has been recognized as promising step in the process of improving the stability of power system [1].

Application of wide area measurement systems (WAMS) into power systems has been established with aim to suppress the occurrence of major blackouts worldwide [1]. Recently, WAMS systems are mainly based on phasor measurement units (PMU). Data from PMUs are time synchronized, using GPS technology, and thus, give the possibility for observing the system as a whole. PMUs estimate the parameters of the voltage/current signal, namely amplitude, frequency and phase [2], [3]. The aim of using PMUs is to monitor the phasors, both in the steady state and under dynamic conditions. Recently, those devices are considered for implementing in protection applications [3]. To secure synchronized operation, and interoperability of the PMUs, the Standard C37.118.1 – 2011 is created by IEEE [4].

The fastest algorithm since now, for phasor estimation is one cycle Fourier filter, known as discrete Fourier transform (DFT) in literature [5].

The DFT based algorithms adopt steady-state concept of phasor [6]. They consider parameters constant within computational window. It makes DFT algorithms very accurate under steady state conditions, but during dynamics, estimated values may become inaccurate. Beside mentioned characteristics, DFT based algorithms are able to reject harmonics presented in signal, but also show unsatisfactorily results at off nominal frequencies and in the presence of decaying DC [7]. In order to overcome off nominal frequency issue, there are with many papers elaborating this topic [8] - [10]. Further, there are some papers that elaborate accuracy analysis of DFT based estimators [11]-[13].

In order to improve DFT under dynamic conditions, a Taylor polynomial method [14], and also Taylor<sup>K</sup> Fourier transform has been proposed. As they include the whole harmonic set into the signal model, these methods have the ability of better rejection around the full set of harmonic frequencies [5].

In time domain, also Kalman filters have been proposed as a method, which estimates depend on its state space signal model. In case of corresponding the input signal and its method, Kalman filters show extraordinary results [15]. Recent development of this method enabled estimation of the dynamic phasor and its first derivative [16].

Newton's iterative method has achieved high accuracy over a wide range of frequency changes [17]. In order to estimate parameters under dynamic conditions, also Prony method can be used [5]. Recently, phase based estimators of phasor and frequency has been developed. They use different methods for calculating parameters, for example, polynomial phase model [18].

This paper provides new method for estimating phasors, and frequency of a sinusoidal signal. In proposed method a complex signal is created from both present and delayed samples of the measured signal. Created signal enables estimation of the frequency. After estimating the frequency, another transformation is applied, which transforms previous complex signal into new, complex, circular signal. Amplitude value, and also phase angle, are computed from the equivalent parameters of the circular complex signal.

The paper is organized as follows: In Section II the signal model is presented. Also phasor form is introduced. Within this section, two representative algorithms are explained. The proposed algorithm is described in Sec. III. Simulation results under tests from standard [4] are given in Sec. 0. Finally, conclusion is derived in Sec. V.

Ivan Jokić, Žarko Zečević, Zdravko Uskoković, Milovan Radulović and Božo Krstajić are with Faculty of Electrical Engineering in Podgorica, University of Montenegro, Džordža Vašingtona bb, 81000 Podgorica, Montenegro (e-mails: [ivan.j@ac.me](mailto:ivan.j@ac.me), [zarkoz@ac.me](mailto:zarkoz@ac.me), [zdravko@ac.me](mailto:zdravko@ac.me), [milovanr@ac.me](mailto:milovanr@ac.me), [bozok@ac.me](mailto:bozok@ac.me)).



## II. ESTIMATION OF SYNCHROPHASORS

Fundamental component of the power system phase voltage signal, in discrete form, can be represented as follows:

$$V_a(n) = A_a \cos(\omega_0 nT + \phi_0) = A_a \cos(2\pi f_0 nT + \phi_0), \quad (1)$$

Where  $A_a$  is amplitude of a signal,  $f_0$  is a fundamental frequency,  $T$  is the sampling interval,  $\omega_0$  is angular frequency and  $\theta_0$  is the initial phase angle. Signal is commonly presented in phasor form:

$$V_a = \left(\frac{V_m}{\sqrt{2}}\right) e^{j\phi} = V_r + jV_i. \quad (2)$$

Using the GPS technology, measurements from the PMU devices are time synchronized.

Numerous algorithms for synchrophasor estimation are proposed in literature. The two widely used representative algorithms are described below, and lately compared with the proposed method.

### A. LES algorithm

LES method is iterative, and uses least-error-squares-based technique to estimate synchrophasors [2]. Method can be described in matrix form, as follows:

$$\mathbf{Y} = \mathbf{A}\mathbf{x} + \boldsymbol{\varepsilon}, \quad (3)$$

Where  $\mathbf{Y}$  is a column vector of measured samples,  $\mathbf{A}$  contains predefined model of the signal,  $\mathbf{x}$  is a column vector which contains estimated values of phasors, and  $\boldsymbol{\varepsilon}$  is error between measured and estimated value of the signal.

Synchrophasors are estimated by solving (3):

$$\mathbf{x} = [\mathbf{A}^T \mathbf{A}]^{-1} \mathbf{A}^T \mathbf{Y}. \quad (4)$$

After computing phasors in each iteration, frequency can also be estimated, using two last iteration results [2]. Values of matrix  $\mathbf{A}$  are then updated, considering new values of the frequency, and the procedure repeats in every reporting frame [4].

### B. DFT based algorithm

DFT based dynamic FIR filter is proposed as a reference algorithm for estimation of synchrophasors in standard [4]. Parameters of fundamental component are estimated as:

$$\hat{X}(i) = \frac{2}{G} \sum_{k=-M/2}^{M/2} x(i+k) \times W(k) \times e^{-j(i+k)T\omega_0}. \quad (5)$$

In previous relation,  $\mathbf{W}$  presents FIR filter coefficients [4],  $k=-M/2, \dots, M/2$ ,  $M$  is window length,  $F_{tr}$  is intersection frequency of the low pass filter,  $T$  is a sampling period,  $h(k)$  is a Hamming window, and  $G$  represents sum of the filter coefficients [4].

## III. STRUCTURE OF THE PROPOSED ALGORITHM

The proposed algorithm estimates parameters of the fundamental component of a measured signal.

In the first step, a complex signal  $V_k$  is created from the measured signal, both present and delayed samples:

$$V_k(n) = V_a(n) + jV_a(n-N) = V_n(n) + jV_b(n), \quad (6)$$

Where  $n$  is iteration index,  $T$  is a sampling period,  $N$  is delay of the signal, and  $j$  is an imaginary number.

Next value of the signal  $V_k$  is estimated using its preceding value and two coefficients,  $h$  and  $g$ , as it is done in [19]:

$$\tilde{V}_k(n+1) = V_k(n)h + jV_k^*(n)g, \quad (7)$$

where coefficients  $g$  and  $h$  are equal:

$$h = \cos(\omega T) + j \frac{\sin(\omega T)}{\sin(\omega NT)}, g = \frac{\sin(\omega T)}{tg(\omega NT)}. \quad (8)$$

Applying equation (7) to the computational window, next relation is derived:

$$\begin{bmatrix} V_k(n - \frac{N}{2}) \\ \vdots \\ V_k(n) \\ V_k(n+1) \\ \vdots \\ V_k(n + \frac{N}{2}) \end{bmatrix} = \begin{bmatrix} V_k(n - \frac{N}{2} - 1) & V_k^*(n - \frac{N}{2} - 1) \\ \vdots & \vdots \\ V_k(n-1) & V_k^*(n-1) \\ V_k(n) & V_k^*(n) \\ \vdots & \vdots \\ V_k(n + \frac{N}{2} - 1) & V_k^*(n + \frac{N}{2} - 1) \end{bmatrix} \begin{bmatrix} h \\ g \end{bmatrix}, \quad (9)$$

where  $N$  represents length of the computational window.

By solving the equation (9), coefficients  $g$  and  $h$  are obtained. If we denote the matrix on the left side of the equation (9) by  $\mathbf{Y}$ , and matrix on the right side by  $\mathbf{M}$ , coefficients  $g$  and  $h$  are then calculated as follows:

$$\begin{bmatrix} h \\ g \end{bmatrix} = [\mathbf{M}^T \mathbf{M}]^{-1} \mathbf{M}^T \mathbf{Y}. \quad (10)$$

Those values can be used for frequency estimation [19]:

$$\tilde{f}(k) = \frac{1}{2\pi T} \sin^{-1}(\Im(h(n) + \alpha(n)g(n))), \quad (11)$$

where coefficient  $\alpha$  depends on  $g$  and  $h$  coefficients [19]. When frequency is estimated, another transformation is applied to signal  $V_k$ :

$$r(n) = W_1^* V_k(n) + jW_2^* V_k(n) = V_a e^{j\omega n T}. \quad (12)$$

The defined transformation is achieved using two coefficients,  $W_1$  and  $W_2$  [20]. By resolving the conditions for obtaining the proposed transformation, next theoretical values are computed for them:

$$W_1 = 1 - j \frac{\cos(\omega NT)}{\sin(\omega NT)}, W_2 = \frac{1}{\sin(\omega NT)}. \quad (13)$$

This transformation results with complex circular signal,  $r(n)$  rotating at power system frequency  $f$ , and with amplitude equal to phasor amplitude  $A_a$ . Finally, amplitude and phase angle of a fundamental component of measured signal are equal to those of signal  $r(n)$ :

$$A_a(n) = |r(n)|, \theta(n) = \text{angle}(r(n)). \quad (14)$$

Computational complexity of the proposed algorithm is significantly higher compared to DFT based method, given

that pseudo inverse of  $N \times 2$  matrix is calculated. On the other side, the LES based algorithm requires pseudo inversion of  $N \times 50$  matrix that contains signal model, which makes LES based algorithm considerably more complex than the proposed one.

#### IV. RESULTS OF SIMULATIONS

Simulations are made in accordance with standard [4]. Sampling frequency of 1/1250 Hz, and reporting speed of 25 frames per second, are used for all three algorithms. For the proposed method and LES based algorithm, 50 samples window length is used. The second signal delay in the proposed method,  $N = 5$ , is used. Window length of 280 samples is used for DFT method.

In the case of various frequency values, within the range  $50\text{Hz} \pm 5\text{Hz}$ , the steady state total vector error (TVE) of all the used methods is far under maximum allowed value, as showed in Table 1. DFT based method, proposed in standard, has the biggest TVE.

TABLE 1: STEADY STATE TVE TEST, FREQUENCY CHANGES

Frequency (Hz)	Range		Maximum TVE	
	+- 5 Hz		1%	
	Prop. M.	DFT	LES	
45	5.778e-12	0.0306	4.932e-12	
50	7.766e-12	0.0349	7.757e-12	
55	8.276e-12	0.0021	6.879e-12	

In Table 2, TVE is showed, for the proposed amplitude values of the input signal. Proposed algorithm has similar results as LES, but significantly better than DFT.

TABLE 2: STEADY STATE TVE TEST, AMPLITUDE CHANGES

Amplitude (p.u.)	Range		Maximum TVE	
	10% - 200%		1%	
	Prop. M.	DFT	LES	
0.1	7.771e-12	0.0349	7.764e-12	
0.8	7.771e-12	0.0349	7.764e-12	
1.2	7.785e-12	0.0349	7.757e-12	
2.0	7.766e-12	0.0349	7.757e-12	

In Table 3, frequency error (FE) is presented for the steady state, under different frequency values of the signal. The FE of the proposed method is considerably lower than the maximum FE [4]. LES method has also low FE value, while DFT based method approaches the maximum FE, as frequency arises.

TABLE 3: STEADY STATE FE TEST, FREQUENCY CHANGES

Frequency (Hz)	Range		Maximum FE	
	+- 5 Hz		0.005 Hz	
	Prop. M.	DFT	LES	
45	2.487e-13	0	4.263e-14	
50	3.482e-13	4.594e-14	4.263e-14	
55	5.045e-13	0.0016	4.974e-14	

A test with ramp change of system frequency is performed, and its TVE and FE results are presented in Table 4 and Table 5, respectively. From Table 4, it can be

concluded that the proposed method evidently better operates under dynamic change of frequency, with regard to used representative methods. The results from Table 5, indicate better capabilities of the proposed method in estimating the frequency, under ramp change of frequency, than other two used methods.

TABLE 4: RAMP FREQUENCY TVE TEST

Ramp Rate [Hz/s]	Ramp rate		Maximum TVE	
	$R_f = \pm 1$ [Hz/s]		1 [%]	
	Prop. M.	DFT	LES	
-1.0	0.0201	0.2301	0.064	
1.0	0.0195	0.2051	0.061	

TABLE 5: RAMP FREQUENCY FE TEST

Ramp Rate [Hz/s]	Ramp rate		Maximum FE	
	$R_f = \pm 1$ [Hz/s]		0.005 [Hz]	
	Prop. M.	DFT	LES	
-1.0	0.0037	0.0214	0.0215	
1.0	0.0036	0.0212	0.0213	

For simulating the dynamic conditions, next form of the input signal is adopted in standard [4]:

$$X_a = X_m [1 + k_x \cos(\omega t)] \times \cos[\omega_0 t + k_a \cos(\omega t - \pi)], \quad (15)$$

Where  $X_m$  is the amplitude of a signal,  $k_x$  and  $k_a$  are amplitude and phase modulation coefficients, respectively, while  $\omega$  is the modulation angular frequency. TVE under amplitude modulation is shown on Figure 1. Proposed method has noticeably better results regarding other two methods. Frequency error values under the same test, shown in Figure 2, indicate that proposed method is more sensitive in estimating the frequency, under amplitude modulation, comparing to other two algorithms. Another type of proposed modulation in standard is phase modulation. Figure 3 and Figure 4 present TVE and FE diagrams, respectively, for considered methods. Both the TVE and FE error for the proposed algorithm are noticeably smaller than for other two used methods.

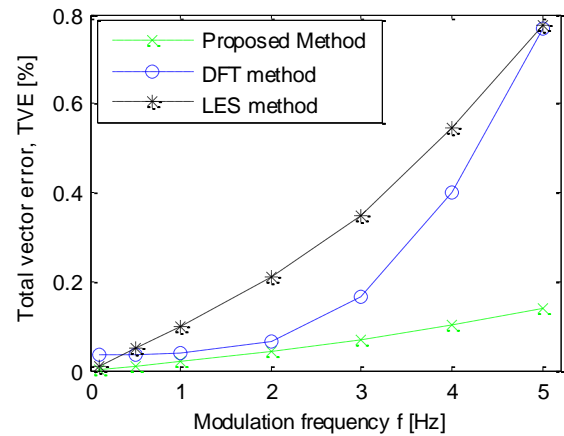
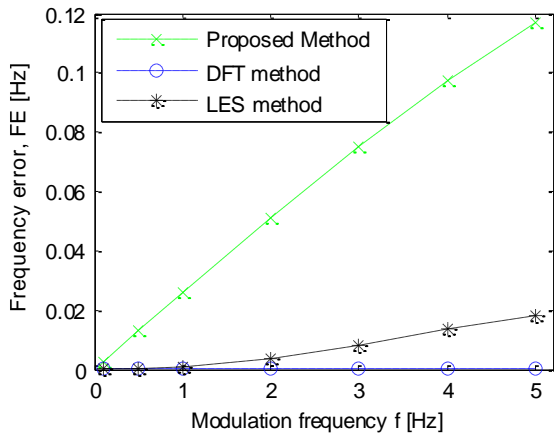
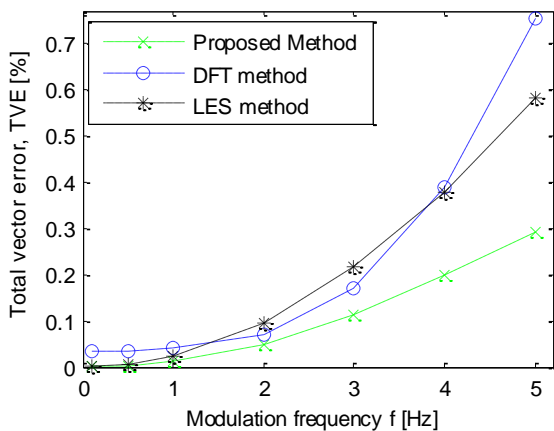
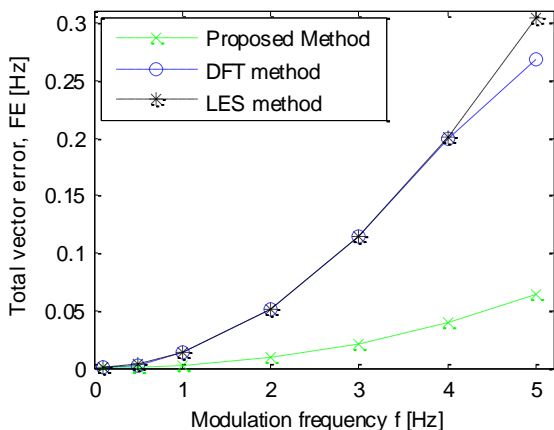


Figure 1. TVE for  $x_a=0.1$ ,  $x_k=0$

Figure 2. FE for  $x_a=0.1$ ,  $x_k=0$ Figure 3. TVE for  $x_a=0$ ,  $x_k=0.1$ Figure 4. FE for  $x_a=0$ ,  $x_k=0.1$ 

## V. CONCLUSION

Proposed method represents a new method for estimating the synchrophasors. Method consists of two transformations. From first, frequency is estimated. This is used for calculating coefficients, which define second transformation. Amplitude and phase angle of the resulting signal corresponds with amplitude and phase angle of the input signal.

Realized experiments are in compliance with the standard [4]. Results from experiments are compared with the results of two widely used algorithms, and make

proposed algorithm competitive and perspective.

## REFERENCES

- [1] Arulampalam Atputharajah, Tapan Kumar Saha, „Power System Blackouts – Literature review“, *Fourth International Conference on Industrial and Information Systems, ICIS 2009*, 28-31 Dec. 2009
- [2] Sarasij Das, Tarlochan Sighu, „A Simple Synchrophasor Estimation Algorithm Considering IEEE Standard C37.118.1-2011 and Protection Requirements“, *IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT*, vol. 62, no. 10, Oct. 2013
- [3] Aranya Chakraborty, Pramod P. Khargonekar, „Introduction to Wide –Area Control of Power Systems“, American Control Conference (ACC), Jun. 17-19, 2013
- [4] *IEEE Standard for Synchrophasor Measurements for Power Systems*, IEEE Standard C37.118.1-2011 (Revision of IEEE Standard C37.118-2005)
- [5] Jose Antonio de la O Serna, Senior Member, *IEEE*, „Synchrophasor Estimation Using Prony’s Method“, *IEEE Trans. On Instrumentation and Measurement*, Vol. 62, No. 8, Aug 2013
- [6] J. A. de la O Serna and K. Martin, „Improving phasor measurements under power system oscillations“, *IEEE Trans. Power Syst.*, vol. 18, no. 1, pp. 160–166, Feb. 2003.
- [7] Sarasij Das, Tarlochan Sighu, „A Simple Synchrophasor Estimation Algorithm Considering IEEE Standard C37.118.1-2011 and Protection Requirements“, *IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT*, vol. 62, no. 10, Oct. 2013
- [8] G. Benmouyal, „System and algorithm for exact compensation of fundamental phasors“, Aug. 23., 2005, Patent 6 934 654
- [9] G. C. Zweigle, L. S. Anderson, A. Guzman-casillas „Apparatus and method for estimating synchronized phasors at predetermined times referenced to an absolute time standard in an electrical system“, Jan. 20, 2009, Patent 7 480 580 B2.
- [10] D. Hart, D. Novosel, H. Yi, B. Smith, and M. Egolf, „A new frequency tracking and phasor estimation algorithm for generator,“ *IEEE Trans. Power Del.*, vol. 12, no. 3, pp. 1064–1073, Jul. 1997.
- [11] D. Macii, D. Petri, and A. Zorat, „Accuracy analysis and enhancement of DFT-based synchrophasor estimators in off-nominal conditions,“ *IEEE Trans. Instrum. Meas.*, vol. 61, no. 10, pp. 2653–2664, Oct. 2012.
- [12] G. Barchi, D. Macii, and D. Petri, „Accuracy of one-cycle DFT-based synchrophasor estimators in steady-state and dynamic conditions,“ in *Proc. IEEE Int. Instrum. Meas. Technol. Conf.*, May 2012, pp. 1529–1534.
- [13] D. Macii, D. Petri, and A. Zorat, „Accuracy of DFT-based synchrophasor estimators at off-nominal frequencies,“ in *Proc. IEEE Int. Workshop AMPS*, Sep. 2011, pp. 19–24.
- [14] J. A. de la O Serna, „Dynamic phasor estimates for power system oscillations,“ *IEEE Trans. Instrum. Meas.*, vol. 56, no. 5, pp. 1648–1657, Oct. 2007.
- [15] P. Zarchan, H. Musoff, and F. K. Lu, *Fundamentals of Kalman Filtering: A Practical Approach*, 3rd ed. AIAA (American Institute of Aeronautics & Ast, Sep. 2009.
- [16] J. A. de la O Serna and J. Rodriguez, „Instantaneous Dynamic Phasor Estimates with Kalman Filter,“ *IEEE* 2010.
- [17] V. V. Terzija, M. B. Djuric, B. D. Kovacevic, „Voltage phasor and local system frequency estimation using Newton type algorithm,“ *IEEE Trans. Power Del.*, vol. 9, no. 3, pp. 1368–1374, Jul. 1994.
- [18] Pasquale Cuccaro, Daniele Gallo, Carmine Landi, Mario Luiso, Gianmarco Romano, „Phase-based Estimation of Synchrophasors“, *IEEE Instrumentation and Measurement Society*
- [19] X. Xia; Z. Blazic, D. Mandić, „Complex-Valued Least Squares Frequency Estimation for Unbalanced Three-Phase Power Systems“, *IEEE Trans Instrumentation and Measurement*, vol.64, no. 3, pp. 638-648, 2015.
- [20] Ž. Zečević, B. Krstajić, T. Popović, „Improved frequency estimation in unbalanced three-phase power system using coupled orthogonal constant modulus algorithm“, *IEEE Transactions on Power Delivery*, Volume: PP, Issue: 99, June 2016.

# Development of Combination Wave generators for simulation of a lightning surges in low-voltage AC power circuits

Ivan R. Bulatović and Vladan M. Radulović, *IEEE*

**Abstract** — A wide variety of surges caused by lightning that are impinging at the service entrance of the buildings with low voltage AC power circuits define so called “surge environment” of these electrical circuits. According to IEEE standards, this surge environment can be described and defined through location categories and standard surge test waveforms. For purposes of simulations and analysis it is necessary to develop models of surge generators that can deliver standard surge test waveforms. In this paper electrical circuits of surge generators of Combination Wave surge, as standard surge test waveform, are proposed. These circuits are modeled and simulated in MATLAB Simulink and obtained results show that values of all waveforms characteristics very well satisfy tolerances given by standards, confirming validity of the proposed surge generators models.

**Key words** — electromagnetic transient programs, low voltage AC power circuits, representative surges, surge generators.

## I. INTRODUCTION

**D**ISTURBANCES on low-voltage AC power circuits can be classified in many ways, such as a voltage increase or reduction from the nominal rms value, a voltage or current wave shape variation, and surge waveform characteristics [1]. The latter include amplitude, duration, rise time, frequency of ringing, polarity, energy-delivery capability, amplitude spectral density, position with respect to the phase of the mains waveform, and frequency of occurrence.

A surges on the AC mains are generally described as a time-domain phenomenon. These surges do not exceed one-half period of the normal mains waveform (fundamental frequency) in duration and they can be periodic or random events and can appear in any combination of line, neutral, or grounding conductors.

Surge voltages and surge currents occurring in low-voltage AC power circuits are originated from two major sources, lightning and switching. Lightning surges can be

described under two distinct scenarios [2]:

- Scenario I: surges impinging at the service entrance of the building;
- Scenario II: direct strike into the structure of lightning protection system.

In the paper, only scenario I will be considered. The database of published surveys (recorded events as well as experiments and computations), anecdotes, and observed failure rates summarized in IEEE C62.41.1 [2] illustrate the wide variety of surges that can be expected to occur in low-voltage AC power circuits.

In order to analyze transient response of equipment and/or performance of surge protective devices in case of surges caused by lightning in low-voltage AC power circuits it is necessary to describe and to model so called “surge environment” in these circuits. According to IEEE C62.41.1 [2] this surge environment can be described through location categories and definition of standard surge test waveforms.

Very scare information regarding electrical circuits of the generators that deliver standard surge test waveforms are presented in the existing literature. Also, manufacturers of commercial surge generators that can be founded on market never give electrical circuits of the generators, but only theirs characteristics and specifications, most often without values of discrepancies in waveforms according to standards.

In the paper, models of surge generators that can deliver Combination Wave surge as one of the standard surge test waveforms for different location categories according to standards are developed and simulated. The Combination Wave surge involves two waveforms, an open-circuit voltage and a short-circuit current. The Combination Wave surge is delivered by a generator that applies a  $1.2/50\mu\text{s}$  voltage wave across an open circuit and an  $8/20\mu\text{s}$  current wave into a short circuit.

Obtained results show that proposed realization of Combination Wave Generators (CWG) very well satisfy tolerances regarding characteristics of surge waveforms given both by IEC and IEEE standards.

## II. CONCEPT OF LOCATION CATEGORIES

As a first step toward a reduction of the complex database on surge occurrences, the concept of location

Ivan R. Bulatović, Crnogorski elektroenergetski sistem, Bulevar Svetog Petra Cetinjskog 18, 81000 Podgorica, Montenegro (telephone: 382-67-862996; e-mail: [ivan.bulatovic@cgcs.me](mailto:ivan.bulatovic@cgcs.me)).

Vladan M. Radulović (corresponding author), Elektrotehnički fakultet u Podgorici, Univerzitet Crne Gore, Džordža Vašingtona bb, 81000 Podgorica, Montenegro (telephone: 382-69-537605; e-mail: [vladanra@ac.me](mailto:vladanra@ac.me)).

categories is proposed in IEEE C.62.41.2 [3]. The propagation of voltage surges and current surges is a phenomenon that does not recognize arbitrary boundaries, but will be influenced by the characteristics of the physical components of the AC power installation. The concept of location category rests on the considerations on dispersion and propagation of surge currents and surge voltages [2].

For surge currents presented at the service entrance of a building, the increasing impedance opposing the flow of surge currents further into the building (with or without the crowbar effect of a flashover that can occur at the meter or in the service-entrance equipment) reduces the surge current that can be delivered along the branch circuits.

In contrast, a voltage surge, with an amplitude below the point of flashover of clearances and presented at the service entrance of a building, can propagate, practically unattenuated, to the end of a branch circuit when no low-impedance load (equipment or local surge protective device - SPD) is present along the branch circuit. Therefore, the location category concept can be described as a staircase of surge voltages and a downhill slope of currents.

IEEE C62.41.1 [2] recognizes three Location Categories: A, B and C. Location Category A applies to the parts of the installation at some distance from the service entrance. Location Category C applies to the external part of the structure, extending some distance into the building. Location Category B extends between Location Categories C and A. Here is worthwhile to be mentioned that similarly to IEEE standards, IEC standards define Lightning Protection Zones.

### III. REPRESENTATIVE SURGE TEST WAVEFORMS

Evaluation of the ability of equipment to withstand variety of surges that can be expected to occur in low-voltage AC power circuits, or of the performance of SPDs in dealing with these surges, can be facilitated by a reduction of the database to a few representative stresses [3].

IEEE C62.41.2 [3] and IEC 61000-4-5 [4] define Combination Wave surge as standard surge-testing waveform. The Combination Wave surge is delivered by a generator that applies a 1.2/50 $\mu$ s voltage wave across an open circuit and an 8/20 $\mu$ s current wave into a short circuit. Equations for the 1.2/50 $\mu$ s voltage and 8/20 $\mu$ s current as separate waveforms are given in IEEE C62.45 [5]. The exact waveforms of voltage and current delivered by the Combination Wave Generator (CWG) is determined by the generator and the impedance of the equipment under test (EUT) to which the surge is applied.

Traditionally, the 1.2/50  $\mu$ s voltage waveform was used for testing the basic impulse level of insulation (BIL), which is approximately an open circuit until the insulation fails. The 8/20  $\mu$ s current waveform was used to inject large currents into surge-protective devices. Because both the open-circuit voltage and short circuit current are different aspects of the same phenomenon, such as an overstress caused by lightning, it is necessary to combine

them into a single waveform when the load is not known in advance. These two surges are combined in one Combination Wave surge.

#### A. Combination Wave surge

The Combination Wave surge is delivered by a generator that can apply a 1.2/50  $\mu$ s voltage wave across an open circuit and an 8/20  $\mu$ s current wave into a short circuit. The exact waveform that is delivered is determined by the generator and the impedance to which the surge is applied. A plot of the nominal open-circuit voltage is shown in Fig. 1. and a plot of the nominal short-circuit current is shown in Fig. 2.

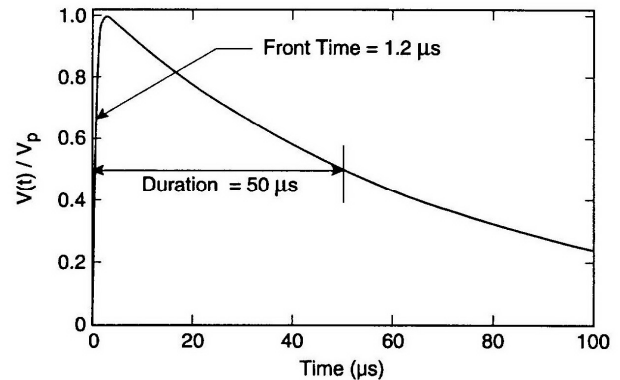


Fig. 1. Combination Wave surge open-circuit voltage

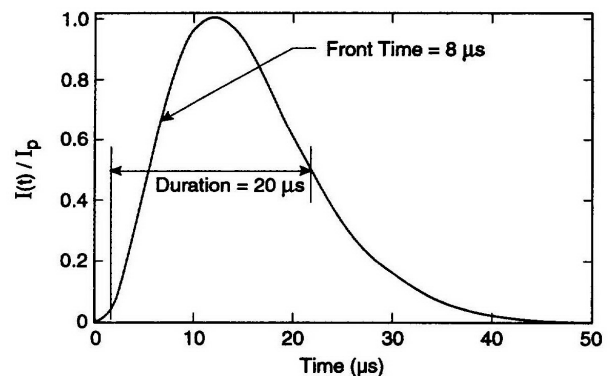


Fig. 2. Combination Wave surge short-circuit current

Open-circuit voltage waveform is defined through following characteristics:

- Front time: 1.2  $\mu$ s  $\pm$  0.36  $\mu$ s
- Duration: 50  $\mu$ s  $\pm$  10  $\mu$ s

The front time for voltage waveforms is  $1.67 \cdot (t_{90} - t_{30})$ , where  $t_{90}$  and  $t_{30}$  are the times of the 90% and 30% amplitude points on the leading edge of the waveform, as defined in IEC and IEEE standards.

The duration is defined as the time between virtual origin and the time of the 50% point on the tail. The virtual origin is the point where a straight line between the 30% and 90% points on the leading edge of the waveform intersects the  $V = 0$  line.

Short-circuit current waveform is defined through following characteristics:

- Front time:  $8 \mu\text{s}$  (+1.0, -2.5)  $\mu\text{s}$
- Duration:  $20 \mu\text{s}$  (+8, -4)  $\mu\text{s}$

The front time for current waveforms is  $1.25 \cdot (t_{90} - t_{10})$ , where  $t_{90}$  and  $t_{10}$  are the times of the 90% and 10% amplitude points on the leading edge of the waveform, as defined in IEC and IEEE standards.

Duration is defined as the time between virtual origin and the time of the 50% amplitude point on the tail. The virtual origin is the time that a straight line between the 10% and 90% amplitude points on the leading edge of the waveform intersects the line.

The value of either the peak open-circuit voltage  $V_p$  or the peak short-circuit current  $I_p$  according to location category are defined and summarized in Table 1.

TABLE 1: EXPECTED MAXIMUM VOLTAGE AND CURRENT SURGES FOR COMBINATION WAVE SURGE

Location category	Amplitude		Effective impedance ( $\Omega$ )
	Voltage (kV)	Current (kA)	
A	6	0.5	12
B	6	3	2

It is necessary to be mentioned that IEC 61000-4-5 [4] also defined Combination Wave surge test waveforms with identical specifications except that open-circuit voltage and short-circuit current waveforms can have undershoot up to 30% of its maximum value as consequence of surge generator design. This is more often case for short-circuit current waveform than for open-circuit voltage waveform. Plot of short-circuit current waveform of Combination Wave generator with 30% undershoot is given in Fig. 3.

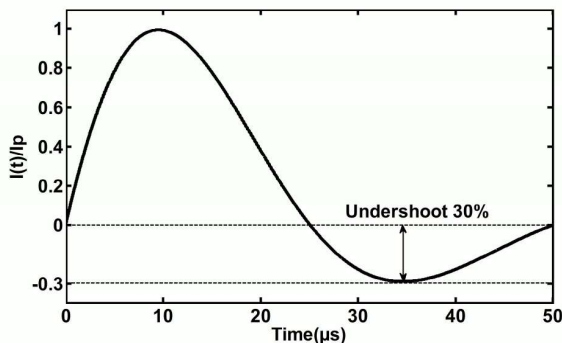


Fig. 3. Short-circuit current waveform with 30% undershoot

These differences regarding undershoot of Combination Wave short-circuit current can have significant influence on deposited energy, as well as on residual voltage on equipment under test [6].

#### IV. DEVELOPMENT OF SURGE GENERATORS DELIVERING STANDARD SURGE TEST WAVEFORMS

First step in analysis of possible influence of surges, caused by lightning, on low-voltage AC power circuits is development of surge generators models that can deliver surge test waveforms according to values and tolerances defined by IEEE C62.45 [5].

In following subchapters electrical circuits of surge

generators delivering Combination Wave surge are presented and analyzed.

Realization of the proposed electrical circuits of the CWG depends on location categories as well as on short-circuit current waveform undershoot.

##### A. CWG without short-circuit current undershoot

Electrical circuit of CWG without short-circuit current undershoot (according to IEEE C62.45 [5]) is given in Fig. 4.

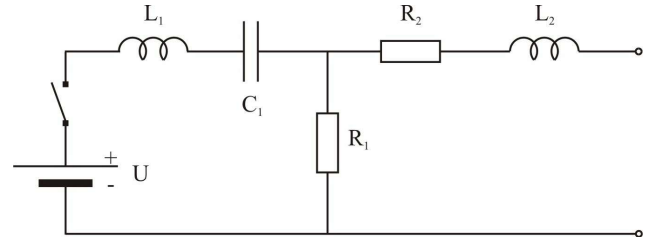


Fig. 4. Electrical circuit of CWG without undershoot

Parameters of elements for proposed realization of CWG without short-circuit current undershoot are:

- For Location Category A:  $U=10.24\text{kV}$ ,  $C_1=1.95\mu\text{F}$ ,  $L_1=15\mu\text{H}$ ,  $L_2=25\mu\text{H}$ ,  $R_1=37\Omega$ ,  $R_2=8.25 \Omega$ .
- For Location Category B:  $U=6.247\text{kV}$ ,  $C_1=12.5\mu\text{F}$ ,  $L_1=2.45\mu\text{H}$ ,  $L_2=4\mu\text{H}$ ,  $R_1=5.83\Omega$ ,  $R_2=1.41 \Omega$ .

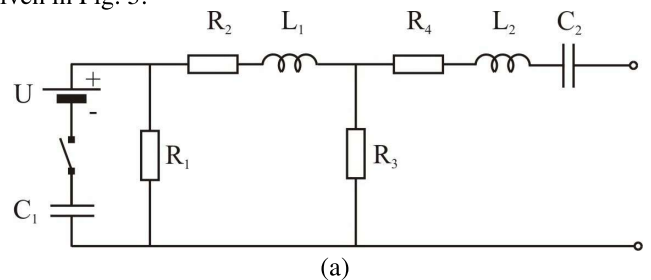
Characteristics of the surges delivered by CWG from Fig. 4., obtained through simulation in MATLAB Simulink are given in Table 2.:

TABLE 2: CHARACTERISTICS OF SURGES DELIVERED BY CWG WITHOUT SHORT-CIRCUIT CURRENT UNDERSHOOT

Location Category	A	B
Front time open-circuit voltage 1.2/50 $\mu\text{s}$	1.21 $\mu\text{s}$	1.26 $\mu\text{s}$
Duration open-circuit voltage 1.2/50 $\mu\text{s}$	52.04 $\mu\text{s}$	52.60 $\mu\text{s}$
Front time short-circuit current 8/20 $\mu\text{s}$	6.79 $\mu\text{s}$	6.70 $\mu\text{s}$
Duration short-circuit current 8/20 $\mu\text{s}$	21.20 $\mu\text{s}$	21.66 $\mu\text{s}$
Effective impedance $U_{max}/I_{max}$	11.98 $\Omega$	1.99 $\Omega$

##### B. CWG with short-circuit current undershoot

Electrical circuit of CWG with exactly 30% short-circuit current undershoot (according to IEC 61000-4-5 [4]) is given in Fig. 5.



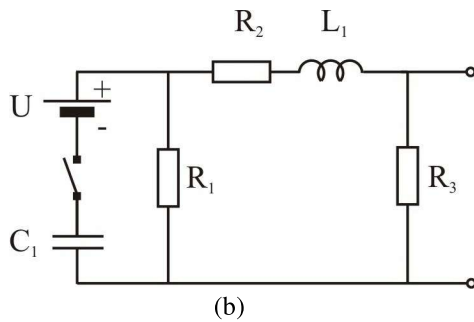


Fig. 5. Electrical circuit of CWG with 30% short – circuit current undershoot (a) for Location Category A, (b) for Location Category B

Parameters of elements for proposed realization of CWG with 30% short-circuit current undershoot are:

- For Location Category A:  $U=6.425\text{kV}$ ,  $C_1=5.75\mu\text{F}$ ,  $C_2=2\mu\text{F}$ ,  $L_1=10\mu\text{H}$ ,  $L_2=42\mu\text{H}$ ,  $R_1=21\Omega$ ,  $R_2=0.9\Omega$ ,  $R_3=23\Omega$ ,  $R_4=5.45\Omega$ .
- For Location Category B:  $U=6.425\text{kV}$ ,  $C_1=5.75\mu\text{F}$ ,  $L_1=10\mu\text{H}$ ,  $R_1=21\Omega$ ,  $R_2=0.9\Omega$ ,  $R_3=23\Omega$ .

From Fig. 5. it can be noticed that electrical circuit of CWG with 30% short-circuit current undershoot for Location Category A differs from the same one for Location Category B only in additional elements  $C_2$ ,  $L_2$  and  $R_4$ .

Characteristics of the surges delivered by CWG from Fig. 5., obtained through simulation in MATLAB Simulink are given in Table 3.:

TABLE 3: CHARACTERISTICS OF SURGES DELIVERED BY CWG WITH 30% SHORT-CIRCUIT CURRENT UNDERSHOOT

Location Category	A	B
Front time open-circuit voltage 1.2/50 $\mu\text{s}$	1.24 $\mu\text{s}$	1.24 $\mu\text{s}$
Duration open-circuit voltage 1.2/50 $\mu\text{s}$	46.95 $\mu\text{s}$	46.74 $\mu\text{s}$
Front time short-circuit current 8/20 $\mu\text{s}$	7.27 $\mu\text{s}$	7.36 $\mu\text{s}$
Duration short-circuit current 8/20 $\mu\text{s}$	20.12 $\mu\text{s}$	18.33 $\mu\text{s}$
Effective impedance $U_{max}/I_{max}$	11.85 $\Omega$	2.01 $\Omega$

It can be noticed that values of all characteristics of proposed realization of CWG given in Tables 2. and 3. very well satisfy tolerances given by standards, which confirms validity of the proposed realizations of CWGs.

As it can be seen from Figs. 3., 4. and 5., all electrical

circuits of surge test waveforms generators have breaker intended to allow that energy from DC source be delivered in different RLC circuits in desired time moment.

## V. CONCLUSION

First step in analysis of possible influence of surges caused by lightning on low-voltage AC power circuits is development of surge generators that can deliver surge test waveforms according to values and tolerances defined by standards.

Combination Wave surge is one of the standard surge test waveforms defined both by IEEE and IEC standards. The Combination Wave surge is delivered by a generator that applies a 1.2/50 $\mu\text{s}$  voltage wave across an open circuit and an 8/20 $\mu\text{s}$  current wave into a short circuit.

Models of surge generators that can deliver Combination Wave surge for different location categories according to standards are developed and presented in the paper.

From obtained results it can be noticed that values of all specific characteristics very well satisfy tolerances given by standards, confirming validity of the proposed realization of surge generators and approve their use in different simulation models. Application of the surge generators models in electromagnetic transient programs can provide further researches of surges influence in low-voltage AC power circuits offering several advantages: parameter analysis, changes in circuit topology, application of different types of equipment under test etc.

## REFERENCES

- [1] *IEEE Recommended Practice for Monitoring Electrical Power Quality*, IEEE Standard 1159, 1995.
- [2] *IEEE Guide on the Surge Environment in Low-Voltage (1000 V and Less) AC Power Circuits*, IEEE Standard C62.41.1, 2002.
- [3] *IEEE Recommended Practice on Characterization of Surges in Low-Voltage (1000 V and Less) AC Power Circuits*, IEEE Standard C62.41.2, 2002.
- [4] *Electromagnetic compatibility (EMC) - Part 4-5: Testing and measurement techniques - Surge immunity test*, IEC Standard 61000-4-5, 2005.
- [5] *IEEE Recommended Practice on Surge Testing for Equipment Connected to Low-Voltage (1000 V and Less) AC Power Circuits*, IEEE Standard C62.45, 2002.
- [6] V. Radulovic and S. Skuletic, "Influence of Combination Wave Generator's Current Undershoot on Overvoltage Protective Characteristics," *IEEE Trans. on Power Delivery*, vol. 26, no. 1, pp. 152-160, January 2011..
- [7] S. Skuletic and V. Radulovic, "Effective Protection Distance from Cascade Coordinated Surge Protective Devices to Equipment in Low-Voltage AC Power Circuits," presented at the 43<sup>rd</sup> International Universities Power Engineering Conference UPEC, Padova, Italy, 1-4 September 2008.
- [8] J. He, Y. Yuan, J. Xu, S. Chen, J. Zou and R. Zeng, "Evaluation of the Effective Protection Distance of Low-Voltage SPD to Equipment", *IEEE Trans. Power Delivery*, vol. 20, no. 1, pp. 123-130, 2005.

# Uticaj planiranog 400kV dalekovoda Čevo-Pljevlja, na dionici kroz opštinu Žabljak, sa aspekta električnog i magnetnog polja

Milutin Ostojić i Martin Čalasan, *Student Member, IEEE*

**Sadržaj** — U ovom radu je prikazana procjena vrijednosti električnog i magnetnog polja planiranog 400kV dalekovoda Lastva-Pljevlja, na dionici kroz opštinu Žabljak. U cilju analize procijenjenih vrijednosti u radu su navedeni mogući uticaji električnog i magnetnog polja, kao i granične norme koje važe u Crnoj Gori, kao i one koje je u svojim smjernicama definisao ICNIRP (*International Commission for Non-Ionizing Radiation Protection*). Analizom procijenjenih vrijednosti električnog i magnetnog polja posmatrane dionice dalekovoda pokazano je će njihove vrijednosti biti daleko manje od važećih graničnih vrijednosti, kao i od vrijednosti električnog i magnetnog polja brojnih uređaja u domaćinstvu.

**Glavne reči** — električno polje, magnetno polje, procjena, granične norme.

## I. UVOD

ELEKTROMAGNETNO polje (EM) se sastoji od dvije komponente - električnog i magnetnog polja koje interaktivno djeluju jedna na drugu. Električno polje je izazvano od strane naelektrisanih objekata preko potencijala (napona), dok je magnetno polje posljedica toka struje između naelektrisanih objekata različitih električnih potencijala [1].

Uticaj EM polja na okolinu zavisi od intenziteta zračenja, vremenske izloženosti zračenju i frekvencije EM polja. Polazeći od frekvencije EM polja, a sa stanovišta zračenja, razlikuju se statička polja, polja niskih i polja visokih učestanosti. Za razliku od statičkog električnog i magnetnog polja, naizmjenična polja u ljudskom tijelu indukuju električne struje, pa su istraživanja uticaja električnih i magnetskih polja niskih učestanosti (ispod 100 kHz) na čovjeka usmjerena na proučavanje efekata tih indukovanih naizmjeničnih struja [2]. Međutim, kod ovih polja (frekvencija polja manja od 100kHz) energija apsorpcije je zanemarljivo mala i ne može uticati na porast tjelesne temperature. Sa druge strane, izlaganje EM poljima učestanosti iznad 100 kHz može dovesti do značajnog apsorbovanja energije, pa se na tim

učestanostima proučava i uticaj EM polja na povećanje tjelesne temperature ili temperature u pojedinim djelovima tijela. Upravo iz tih razloga, za ograničavanje izlaganja stanovništva štetnom dejstvu električnih i magnetnih polja postoje brojni međunarodni i nacionalni propisi, smjernice i preporuke [2-5]. U Crnoj Gori je takođe usvojen zakon, sa odgovarajućim pravilnicima, koji definiše ovu oblast [6].

U ovom radu prikazana je procjena uticaja električnog i magnetnog polja planiranog 400kV dalekovoda Čevo-Pljevlja, na dionici opštine Žabljak. U tom cilju, u Poglavlju II, navedene su najvažnije međunarodne organizacije koje se bave ovom problematikom, kao i granične norme izlaganja stanovništva elektro-magnetnim poljima. Nakon toga, u Poglavlju III, dat je pregled metoda za proračun električnog i magnetnog polja elektro-energetskih objekata. U Poglavlju IV prikazani su rezultati proračuna električnog i magnetnog polja planiranog 400kV dalekovoda na konkretnom mjestu na trasi dalekovoda koje je najkritičnije sa aspekta intenziteta zračenja. U istom poglavlju izvršena je analiza procijenjenih rezultata rada i izvršeno poređenje dobijenih vrijednosti sa graničnim vrijednostima, kao i sa vrijednostima električnog i magnetnog polja brojnih uređaja u domaćinstvu. Na kraju, u zaključku je dat osvrt na cjelokupan rad.

## II. GRANIČNE NORME IZLAGANJA EM POLJA

Za ograničavanje izlaganja stanovništva štetnom dejstvu električnih i magnetskih polja postoje međunarodni i nacionalni propisi, smjernice i preporuke [3-5]. Najpoznatiji međunarodni dokumenti su smjernice:

- Međunarodne komisije za zaštitu od nejonizujućeg zračenja (*International Commission on Non-Ionizing Radiation Protection – ICNIRP*) (Tabele 1 i 2; „f“ - frekvencija),
- Svjetske zdravstvene organizacije (World Health Organization – WHO)
- Međunarodne agencije za istraživanje raka (IARC – WHO International Agency for Research on Cancer).

Crna Gora je donijela zakon o nejonizujućem zračenju (“Sl. list Crne Gore”, Br 35/13 od 23.07.2013. godine), koji se počeo primjenjivati od 01.07.2015. godine [6]. Takođe, donešeno je i desetak drugih podzakonskih akata koji su objavljeni u Službenom listu Crne Gore, a koji tretiraju problematiku elektromagnetnog zračenja. U “Pravilniku o granicama izlaganja elektromagnetnim poljima”, definisane su granične vrijednosti izlaganja

*Milutin Ostojić*, Elektrotehnički fakultet u Podgorici, Univerzitet Crne Gore, Džordža Vašingtona bb, 81000 Podgorica, Crna Gora (telefon: 382-67-624-347; e-mail: [milutino@t-com.me](mailto:milutino@t-com.me)).

*Martin Čalasan*, Elektrotehnički fakultet u Podgorici, Univerzitet Crne Gore, Džordža Vašingtona bb, 81000 Podgorica, Crna Gora (telefon: 382-67-615-237; e-mail: [martinc@t-com.me](mailto:martinc@t-com.me)).



elektromagnetnih polja. Granične vrijednosti izlaganja EM polju za opštu populaciju prema ovom pravilniku imaju iste vrijednosti kao preporuke INCRIP-a. Međutim, referentni nivoi relevantnih fizičkih veličina za izloženost stanovišta elektromagnetnim poljima u području povećane osjetljivosti za pojedinačnu frekvenciju su strožijih vrijednosti u odnosu na preporuke INCRIP-a (Tabela 3).

TABELA 1: GRANIČNE VRIJEDNOSTI ELEKTRIČNOG I MAGNETSKOG POLJA KOJA JE PROPISAO ICNIRP 2010 - OPŠTA POPULACIJA.

Frekvencija	E [kV/m]	B [T]
1Hz – 8Hz	5	$4 \cdot 10^{-2} \cdot f^2$
8Hz – 25Hz	5	$5 \cdot 10^{-3} \cdot f$
25Hz – 50Hz	5	$2 \cdot 10^{-4}$
50Hz – 400Hz	$2.5 \cdot 10^2 \cdot f$	$2 \cdot 10^{-4}$
400Hz – 3kHz	$2.5 \cdot 10^2 \cdot f$	$8 \cdot 10^{-2} \cdot f$
3kHz – 10MHz	$8.3 \cdot 10^{-2}$	$2.7 \cdot 10^{-5}$

TABELA 2: GRANIČNE VRIJEDNOSTI ELEKTRIČNOG I MAGNETSKOG POLJA KOJA JE PROPISAO ICNIRP 2010 - PROFESIONALNO OSOBLJE.

Frekvencija	E [kV/m]	B [T]
1Hz – 8Hz	20	$0.2 \cdot f^2$
8Hz – 25Hz	20	$2.5 \cdot 10^{-2} \cdot f$
25Hz – 300Hz	$5 \cdot 10^2 \cdot f$	$1 \cdot 10^{-3}$
300Hz – 3kHz	$5 \cdot 10^2 \cdot f$	$0.3 \cdot f$
3kHz – 10MHz	$1.7 \cdot 10^{-1}$	$1 \cdot 10^{-4}$

TABELA 3: REFERENTNI NIVOI RELEVANTNIH FIZIČKIH VELIČINA ZA IZLOŽENOST STANOVIŠTA ELEKTROMAGNETNIM POLJIMA U PODRUČJU POVEĆANE OSJETLJIVOSTI ZA POJEDINAČNU FREKVENCIJU.

Frekvencija	E [V/m]	B [T]
1Hz – 8Hz	1250	$1 \cdot 10^4 \cdot f^2$
8Hz – 25Hz	1250	$1.25 \cdot 10^3 \cdot f$
25Hz – 50Hz	1250	50
50Hz – 400Hz	$62.5 \cdot f$	50
400Hz – 3kHz	$62.5 \cdot f$	$20 \cdot f$
3kHz – 10MHz	21	7

### III. NAČINI PRORAČUN ELEKTRIČNOG I MAGNETNOG POLJA DALEKOVODA

Za proračun električnog i magnetskog polja oko energetske objekata mogu se koristiti različite metode. Neke od njih zasnovane su na rješavanju Maxwell-ovih jednačina uz definisane početne uslove [7], neke polaze od talasnih jednačina električnog i magnetskog polja (Helmholtz-ove jednačine) i koriste metod konačnih elemenata uz upotrebu softverskog paketa FEMM [8], PDFtool softverskog paketa MATLAB [9] ili upotrebom HIFREQ programa softverskog paketa CEDEGS [10].

U ovom radu, u cilju procjene efektivnih vrijednosti električnog i magnetnog polja planiranog 400kV dalekovoda Čevo-Pljevlja iskorišćen je metod opisan u [11]. Prema opisanom metodu, pri proračunu električnog i magnetskog polja provodnika niske učestanosti zemlja se može posmatrati kao paramagnetna poluprovodna sredina, što znači da u njoj preovladavaju kondukcione struje u odnosu na struje pomjeraja, pa se zemlja ponaša kao dobar

provodnik. Zbog toga se, pri proračunu električnog polja može primijeniti metod ogledanja, a pri proračunu magnetskog polja može se smatrati da se provodnici nalaze u homogenoj sredini (vazduh). Kad se odredi polje koje stvara jedan provodnik, onda se primjenom metoda superpozicije mogu odrediti polja koja stvaraju svi provodnici dalekovoda.

Kod korišćenog metoda podrazumijeva se i sljedeće:

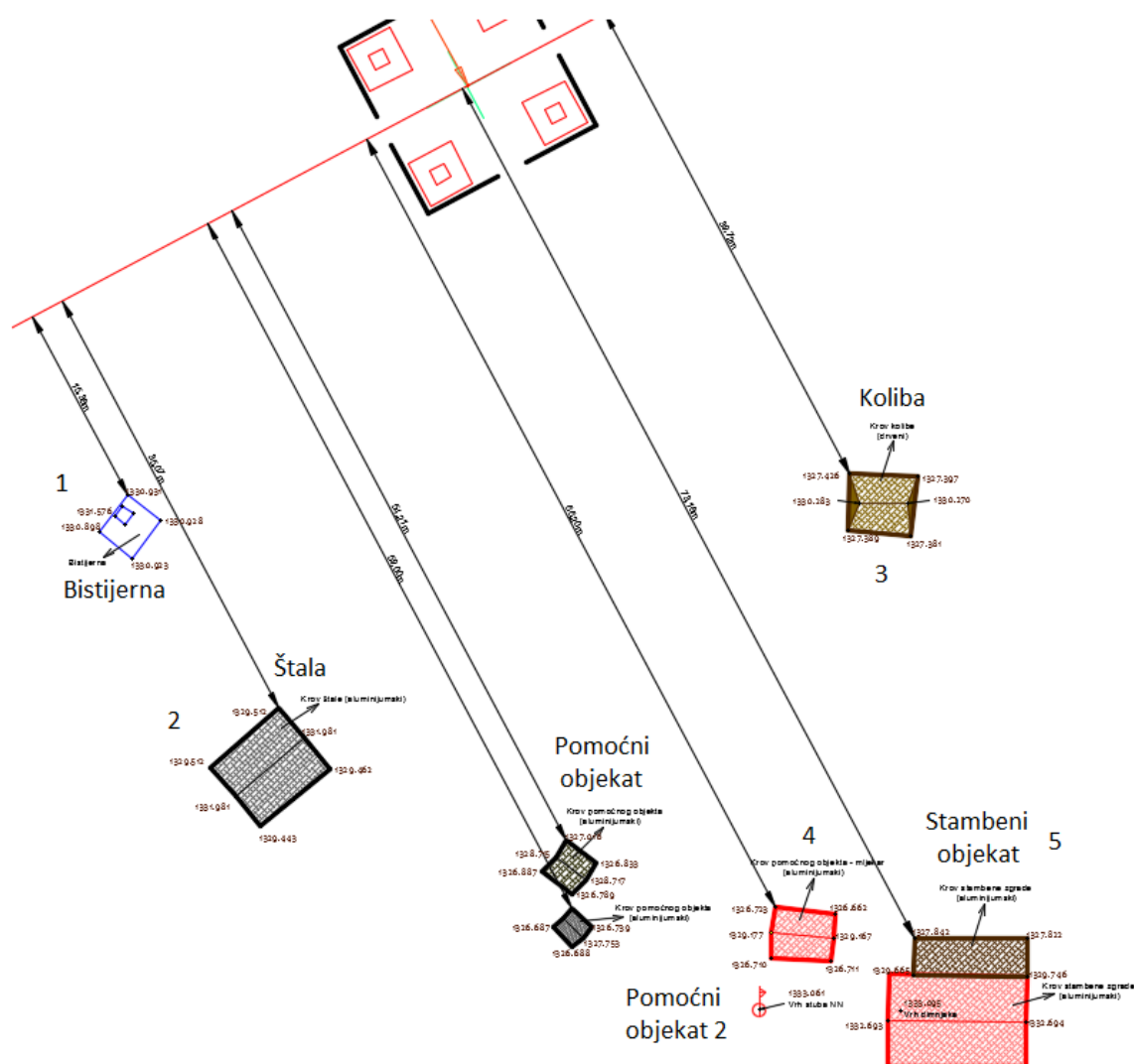
- ✓ provodnici su pravi i beskonačno dugački, a nalaze se na visini tačke poprečnog preseka dalekovoda u kojem se proračunavaju lančanice i polja;
- ✓ proračun se sprovodi po pravcu koji je normalan na osu dalekovoda (tzv. lateralni profil);
- ✓ proračun se sprovodi za najnepovoljniji slučaj, tj. uz pretpostavku da je teren ravan, da na odabranom mjestu nema objekata (kuća, ograda, drveća i drugih objekata) koji mogu da utiču na snižavanje nivoa nejonizujućeg zračenja;
- ✓ pri proračunu magnetnog polja zanemaruje se povratni put struje kroz zemlju;
- ✓ teren je ravan i sve tačke terena su na visini centralne linije;
- ✓ zemlja je savršeni provodnik i
- ✓ dielektrična konstanta je nezavisna od vremenskih uslova i jednaka dielektričnoj konstanti vakuuma.

### IV. PRORAČUN ELEKTRIČNOG I MAGNETNOG POLJA 400kV DALEKOVODA ČEVO-PLJEVLJA NA TERITORIJI OPŠTINE ŽABLJAK

Proračuni elektromagnetskog polja planiranog 400kV dalekovoda su urađeni su na mjestima na trasi dalekovoda koja se smatraju kritičnim sa aspekta inteziteta zračenja. To su mjesta na kojima je visina provodnika najmanja, na kojima dalekovod prelazi preko ili blizu stambenih i drugih objekata, regionalnih i lokalnih puteva i šumskih staza. GOOGLE Earth prikaz najkritičnije tačke (blizina stambenog objekta) prikazan je na Slici 1. Dispozicija postojećih objekata u odnosu na planirani dalekovod prikazana je na slici 2.



Sl. 1. GOOGLE Earth položaj postojećeg najkritičnijeg objekata u blizini planiranog 400kV dalekovoda.



Sl. 2. Dispozicija postojećih objekata u odnosu na planirani 400kV dalekovod

Tabela I - Jačine električnih i magnetnih polja određenih aparata u domaćinstvu [3]

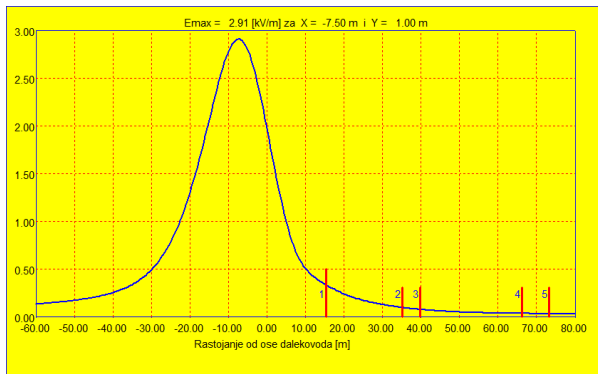
<i>ELEKTRIČNI</i> <i>APARATI</i>	<i>Jačina električnog polja</i> <i>(V/m) na rastojanju od 30cm</i>	<i>Jačina magenetnog polja (μT) na</i> <i>raznim rastojanjima</i>		
		<i>3cm</i>	<i>30cm</i>	<i>100cm</i>
Pegla	120	8-30	0.12-0.3	0.01-0.03
Frižider	120	0.5-1.7	0.01-0.25	<0.01
Mikser	100	-	0.6-10	-
Toster	80	-	-	-
Fen za kosu	80	6-2000	0.01-7	0.01-0.03
TV prijemnik u boji	60	2.5-50	0.04-2	0.01-0.25
Aparat za kafu	60	-	-	-
Usisivač	50	200-800	2-20	0.13-2
Električna rerna	8	1-50	0.15-0.5	0.01-0.04
Sijalica sa užarenim vlaknima	5	-	-	-
Mašina za veš	-	0.8-50	0.15-3	0.01-0.15
Računar	-	0.5-30	<0.01	-
Električni pokrivač	250	-	1.3-3.3	-

Proračuni električnog i magnetnog polja su izvršeni:

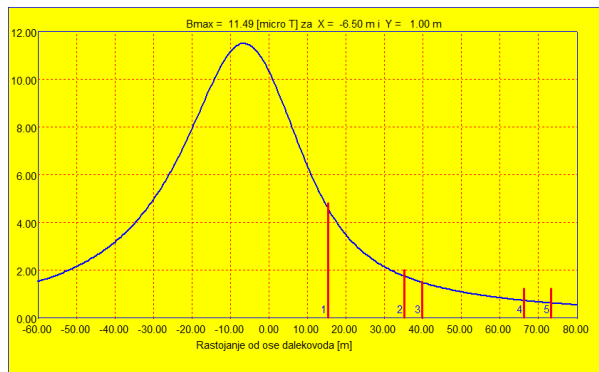
- ✓ za konfiguracije glave čelično rešetkastog stuba tipa "jela" koji su projektom odabrani za ovaj dalekovod,
- ✓ za najvišu dozvoljenu vrijednost napona od 420kV za ovaj naponski nivo.
- ✓ za dozvoljenu struju uzeta od 2000A.

✓ za visinu od 1m iznad površine tla.

Raspodjele jačine električnog polja  $E$ [kV/m] i magnetske indukcije  $B$ [ $\mu$ T], u odnosu na osu dalekovoda prikazane su na Sl. 3 i 4. Crvenim vertikalnim linijama sa oznakama 1 do 5 (vidi sliku 2) prikazane su ivice objekata koje su bliže dalekovodu.



Sl. 3. Raspodjela električnog polja (objekti 1...5 su označeni na slici 2).



Sl. 4. Raspodjela magnetne indukcije (objekti 1...5 su označeni na slici 2).

Na osnovu prikazanih rezultata vidi se da su nivoi polja kod svih objekata (vidi Sliku 2) niži od graničnih vrijednosti koje su propisane za zone povećane osjetljivosti, tj. niži su od 1.25 kV/m i 50 $\mu$ T (Tabela 3). Štaviše, jačine električnog polja kod objekata 2, 3, 4 i 5 će biti srazmjernih vrijednosti kao jačine električnog polja brojnih uređaja u domaćinstvu (Tabela 4). Takođe, vrijednosti magnetne indukcije predmetnog dalekovoda u blizini posmatranih objekata će biti mnogo manje od vrijednosti magnetne indukcije koju formiraju brojni uređaji u domaćinstvu (vidi Tabelu 4).

#### I. ZAKLJUČAK

U ovom radu je prikazan uticaj planiranog 400kV dalekovoda Čevo-Pljevlja, na dionici kroz opštinu Žabljak, sa aspekta električnog i magnetnog polja. U tu svrhu analizirana je vrijednost električnog i magnetnog polja na najkritičnijoj lokaciji sa stanovišta boravka ljudi.

Pokazano je da je vrijednost i električnog i magnetnog polja daleko manja od graničnih vrijednosti električnog i magnetnog polja koje definiše zakon o nejonizujućem zračenju, a koji važi u Crnoj Gori. Takođe, pokazano je da će dalekovod formirati električno i magnetno polje čije će vrijednosti biti manje u poređenju sa električnim i

magnetnim poljem brojnih uređaja koji se koriste u domaćinstvu.

#### LITERATURA

- [1] J. V. Surutka, "Elektromagnetika", Akademska misao, Beograd, 2005.
- [2] J. Jovanović, et. al. „Štetni efekti elektromagnetnih polja ekstremno niskih frekvencija“, *Acta Medica Medianae*, UDK: 614.875:537.8, pp. 54-58, 2010.
- [3] M.S. Čalović, M.M. Mesarović, "Elektromagnetna polja industrijske frekvencije: priroda, neželjeni efekti i zaštita od njihovih štetnih uticaja", *Elektroprivreda*, LXIV, Broj 4, pp. 341-352, 2011.
- [4] M. Ostojić, M. Čalasan, B. Novosel, "Elektromagnetno zračenje dalekovoda i trafostanica", *Časopis Pogled - Inženjerska komora Crne Gore*, ISSN 2336-9175, jul 2013, pp. 50-56.
- [5] "ICNIRP Guidelines for limiting exposure to time-varying electric and magnetic fields (1 Hz – 100 kHz)", *Health Physics*, Vol. 99, No. 6, pp. 818-836, 2010
- [6] PRAVILNIK O GRANICAMA IZLAGANJA ELEKTROMAGNETNIM POLJIMA, Službeni list Crne Gore, Broj: 09-101/49-2014, Podgorica, 23. januar 2015. Godine
- [7] Veličković D.M., Radulović J.J., Božić M.D., „Elektromagnetsko Polje Dalekovoda“, *JUKO CIGRE*, Herceg Novi, pp.1-8, 1997.
- [8] Petković D.M., Krstić D.D., Stanković V.B., „Effect of Electric Fields on Humans in The Immediate Vicinity of 110 kV power Lines“, *Facta Universitates*, Vol. 3, No 1, pp. 63-72, 2006.
- [9] Tupisie S., Isaramongkolrak A. Pao-la-or P. „Analysis Electromagnetic Field Effect Using FEM For Transmission Lines Transposition“, *Proc. of World Academy of Science*, Vol. 41, May 2009., pp. 870-874
- [10] Nikolovski S., Marić P., Baus Z., „Elektromagnetic Field Calculation of Transformer Substation 400/110 kV Ernestovo Using the CEDEGS Software“, *Journal of Electrical Engineering*, Vol. 58, No. 4, pp. 207-213, 2007.
- [11] D. Filipović, M. Ostojić, "Proračun električnog i magnetskog polja u blizini dalekovoda", *I Savjetovanje CG KO CIGRE*, R-B2-03, pp.1-8., oktobar 2009.

#### ABSTRACT

This paper presents the estimation of the electric and magnetic field of planned 400kV transmission power line Čevo-Pljevlja at section through the municipality of Žabljak. In order to analyze the estimated value, in this paper the possible effects of electric and magnetic fields, as well as limit norms applicable in Montenegro, and norms approved by ICNIRP (International Commission for Non-Ionizing Radiation Protection) are noted. The analysis of the estimated values of the electric and magnetic fields of the observed power transmission lines is shown that their values will be less than the applicable threshold values. Also, their values will be less than values of the electric and magnetic field of many household devices.

#### IMPACT OF 400KV TRANSMISSION ČEVO-PLJEVLJA, AT SECTION THROUGH THE MUNICIPALITY OF ZABLJAK, FROM THE ASPECT OF ELECTRIC AND MAGNETIC FIELDS

Milutin Ostojić, Martin Čalasan

# Use of dynamic programming for short-term hydro-thermal coordination

Milan Kocovic Spec. Sci, Doc. dr Zoran Miljanic, Prof. emeritus Ilija Vujosevic

**Abstract** — This paper presents the application of dynamic programming method in the hydro-thermal coordination, which is based on finding the optimal trajectory of the use of water in reservoir storage with the limitation of initial and final storage volumes. Initial and final volumes of the reservoir present the input data in the optimization and which were obtained in short-term planning of water consumption. Also, this paper presents the modeling and the possible approximations of the input-output characteristic of steam and hydroelectric plants. For the purpose of hydro-thermal coordination, the program was implemented using MATLAB software and its graphical user GUI interface. The mentioned program includes hydro-thermal coordination that was made on the principle of dynamic programming using Dijkstra algorithm.

**Keywords** — Dynamic programming, Dijkstra algorithm, Hydro-thermal coordination

## I. INTRODUCTION

THE COORDINATION of the system for production of electrical power with hydroelectric plants is usually more complex than scheduling of an all-thermal generation systems. The hydroelectric plants may be coupled both electrically and hydraulically (downstream multiple plants). There are no two hydroelectric systems in the world that are alike. There are many reasons for this fact/ the natural differences in the watersheds, reservoirs storage as well as so many different restrictions on the operation of the system, such as: controlling floods, regular discharge of water for irrigation, limited water release because of navigability of river at all time or availability of recreation center downstream of reservoir storage. Sudden changes, with high-volume releases of water, may induce high waves with potentially damaging effect for downstream object [1].

The goal of hydro-thermal coordination is to use the amount of water that is given through short-term scheduling of water consumption in the best possible way, which means to obtain the minimum production costs of thermal plants [1].

The coordination of the operations of hydroelectric plants involves, naturally, the scheduling of water releases. The long-range hydro-scheduling problem involves the

Milan Kocovic Spec. Sci. (corresponding author), Faculty of electrical engineering in Podgorica, University of Montenegro, Studentska street nn, 81000 Podgorica, Montenegro (e-mail: kocovicmilan@gmail.com).

Doc. dr Zoran Miljanic, Faculty of electrical engineering in Podgorica, University of Montenegro, Dzordza Vasingtona Street no number, 81000 Podgorica, Montenegro (e-mail: zormi@ac.me).

Prof. emeritus Ilija Vujosevic, Faculty of electrical engineering in Podgorica, University of Montenegro, Dzordza Vasingtona Street no number, 81000 Podgorica, Montenegro (e-mail: ilijav@ac.me).

long-range forecasting of water availability that is based on meteorological and statistical analyse. Short-range hydro-scheduling involves scheduling of amount of water that can be used for generation in short-term (1 day to 1 week). The coordination of hydro - scheduling must be in coordination with water inflow in order to keep the amount of water within an specified range [1].

## II. HYDROELECTRIC PLANT MODEL

The amount of energy that is available in a unit of stored water is equal to the product of the weight of water and the height of the water fall. From that standpoint, it can be concluded that a small amount of water and a big drop produce energy which can be produced with a big amount of water and small drop. For example, 1kWh of energy can be obtained in the following way:

$$1\text{kWh}=24.45\text{m}^3\cdot 15\text{m} \text{ and } 1\text{kWh}=15\text{m}^3\cdot 24.45\text{m}.$$

The amount of energy that is available in the unit quantity of water is not equivalent to electrical energy obtained via that quantity and that fall and it must be reduced by a certain value, caused by energy losses in the energy conversion. Power provided by the water is proportional to the flow of water in  $\text{m}^3/\text{s}$  and the coefficient of energy conversion, which takes into account the net head (gross head - flow losses into the penstock) and the coefficient of efficiency of the turbine-generator. Flow losses typically is the amount of about 5% of the gross head, while in the case of hydro power plants with low pressure and a case of long penstocks, these losses can be much higher. Also, a factor of the efficiency of the turbine is load, and efficiency for smaller loads may fall below 70%. Of course, it should take into account the existence of loss in the conversion of mechanical energy of the turbine into electrical energy of the generator[1].

Typical performance characteristic of hydropower plants is characteristic of incremental water rate  $\delta[\text{m}^3/\text{MWh}]$ , that is shown on Fig.1. for three hydroplants.

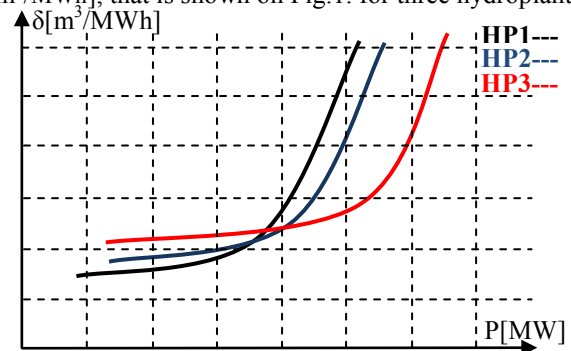


Fig.1. Incremental water rate versus output power

The rise in incremental water rate with increasing unit output results primarily from the increased hydraulic

losses with the increased flow. With the increasing number of hydro units, the hydraulic losses are being increased, primarily due to rise in afterbay which produces decreasing in net and gross head [1].

Fig.2. is showing the basic hydroelectric plant characteristics that represent the dependence of water flow  $q[\text{m}^3/\text{h}]$  and output power  $P[\text{MW}]$ .

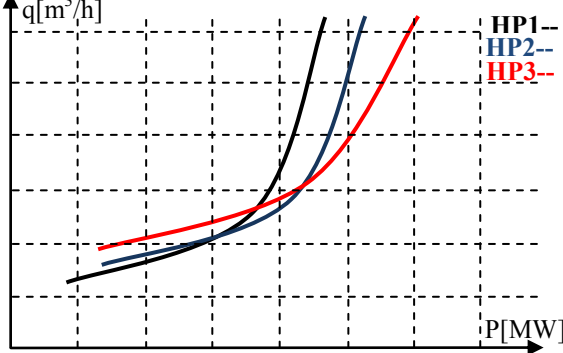


Fig. 2. Basic hydroelectric plant characteristics

This basic energy characteristic of hydroplant has such a form that it can be approximated by a quadratic equation, namely. It can be written in the following form:

$$q = A + B \cdot P + C \cdot P^2 \quad (1)$$

Where  $A$ ,  $B$  and  $C$  are coefficients of quadratic equation which are calculated on the principle of achieving the minimum squared deviations between the actual value of the flow  $q_i$  in the point  $P_i$  and the approximated value of the flow  $q_i$  in the same power  $P_i$  [2].

In case it is known  $n$  points  $(P_i, q_i)$ , then the principle of calculation of the coefficients  $A$ ,  $B$  and  $C$  consists of the following steps:

$$J = \sum_{i=1}^n (A + B \cdot P_i + C \cdot P_i^2 - q(P_i))^2 \quad (2)$$

Where is  $J$  sum of quadratic deviations between the actual and the approximated value of flow.

The goal is to find the minimum of the function  $J$  depending on  $A$ ,  $B$  and  $C$  coefficients, as shown below [2].

$$\frac{\partial J}{\partial A} = \sum_{i=1}^n 2 \cdot (A + B \cdot P_i + C \cdot P_i^2 - q(P_i)) = 0 \quad (3)$$

$$\frac{\partial J}{\partial B} = \sum_{i=1}^n 2 \cdot P_i \cdot (A + B \cdot P_i + C \cdot P_i^2 - q(P_i)) = 0 \quad (4)$$

$$\frac{\partial J}{\partial C} = \sum_{i=1}^n 2 \cdot P_i^2 \cdot (A + B \cdot P_i + C \cdot P_i^2 - q(P_i)) = 0 \quad (5)$$

$$n \cdot A + \left( \sum_{i=1}^n P_i \right) \cdot B + \left( \sum_{i=1}^n P_i^2 \right) \cdot C = \sum_{i=1}^n q(P_i) \quad (6)$$

$$\left( \sum_{i=1}^n P_i \right) \cdot A + \left( \sum_{i=1}^n P_i^2 \right) \cdot B + \left( \sum_{i=1}^n P_i^3 \right) \cdot C = \sum_{i=1}^n P_i \cdot q(P_i) \quad (7)$$

$$\left( \sum_{i=1}^n P_i^2 \right) \cdot A + \left( \sum_{i=1}^n P_i^3 \right) \cdot B + \left( \sum_{i=1}^n P_i^4 \right) \cdot C = \sum_{i=1}^n P_i^2 \cdot q(P_i) \quad (8)$$

Solving this system, namely finding the values of coefficients  $A$ ,  $B$ , and  $C$ , it is obtained the approximate curve for the planning of generation of hydroplant. That

curve is sufficiently precise for economical scheduling [2].

When planning production, in addition to the energy characteristic of hydroplant, it is necessary to have information of the minimum  $q_{min}$  and maximum  $q_{max}$  allowed water flow that also provides data of the minimum and maximum power of hydroelectric plant.

### III. MODEL OF STEAM PLANT

#### A. Input-output characteristic of steam plant

Electrical energy of steam plant is produced by conversion of heat energy to mechanical than mechanical to electrical energy. The curve of heat consumption shows the consumption of heat energy per hour versus the output power of steam plant. More often used characteristic of steam plant is curve of consumption fuel, i.e. fuel cost curve. The total cost of operation includes the fuel cost and the cost of labor, supplies, and maintenance that are expressed as fixed percentage of the incoming fuel costs. Fig.3. shows typical example of input-output characteristic of steam plant. [3].

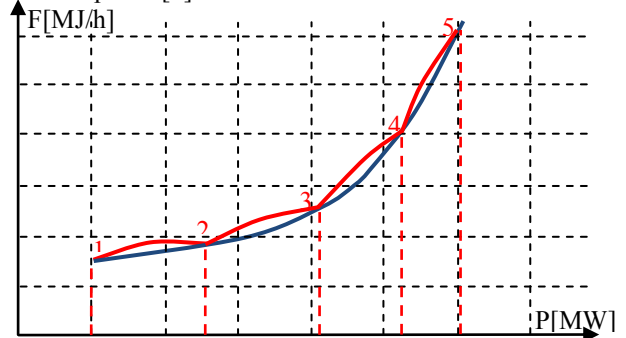


Fig. 3. Fuel cost characteristic of steam plant

On Fig. 3. by points 1 to 5 are indicated the processes of opening the valve. In the vicinity of these points there are discontinuities as the product of increased losses in valves in that range [3]. The curve of fuel cost can be defined by polynomial:

$$F = \alpha + \beta \cdot P + \gamma \cdot P^2 \quad (9)$$

The coefficients  $\alpha$ ,  $\beta$  i  $\gamma$  are calculated on principle of minimum squared deviation between real and approximated value of fuel cost, similar as finding coefficients  $A$ ,  $B$  and  $C$  in previous chapter.

#### B. Optimal operation of an all-thermal system

Typical operation of this group of system is operation of thermal generation units on same bus. Assume that  $\alpha_i$ ,  $\beta_i$  and  $\gamma_i$  are calculated for all generation units. The total fuel cost  $F$  for  $m$  thermal units that supply the demand  $P_D$  can be expressed as:

$$F = \sum_{i=1}^m (\alpha_i + \beta_i P_i + \gamma_i P_i^2) \quad (10)$$

If transmission losses are negligible, it can be written next equation:

$$P_D = \sum_{i=1}^m P_i, \quad (11)$$

$P_D$  - power demand of the system, and  $P_i$  - power that is produced by  $i$  steam plant.

The most used method for solving constrained minimization problems is technique of Lagrange multipliers. With constraint:

$$P_D - \sum_{i=1}^m P_i = 0, \quad (12)$$

it is obtained:

$$\hat{F} = F_T + \lambda \cdot \left[ P_D - \sum_{i=1}^m P_i \right], \quad (13)$$

$$F_T = \sum_{i=1}^m F_i(P_i), \quad (14)$$

and where  $\lambda$  is multiplier which is unknown at the beginning.

Lagrange method is based on the introduction of penalization for any deviation from the set limits, and the minimum is achieved by equating the first derivatives of the functions with 0.

$$\frac{\partial F_i}{\partial P_i} - \lambda = 0, \quad (15)$$

$$\lambda = \frac{\partial F_1}{\partial P_1} = \frac{\partial F_2}{\partial P_2} = \frac{\partial F_3}{\partial P_3} = \dots = \frac{\partial F_m}{\partial P_m}. \quad (16)$$

It is concluded that each optimal operation unit should share the load such that their incremental costs  $\lambda$  are equal. For equation of fuel cost  $F_i(P_i)$  for  $i$  units with Lagrange method it is obtained:

$$\beta_i + 2\gamma_i P_i - \lambda = 0. \quad (17)$$

From equation (17) it follows that:

$$P_i = \frac{\lambda - \beta_i}{2\gamma_i}, \quad (18)$$

$$P_D - \sum_{i=1}^m P_i = 0, \quad (19)$$

$$P_D - \sum_{i=1}^m \left( \frac{\lambda - \beta_i}{2\gamma_i} \right) = 0, \quad (20)$$

$$\lambda = \frac{2P_D + \sum_{i=1}^m \frac{\beta_i}{\gamma_i}}{\sum_{i=1}^m \frac{1}{\gamma_i}}. \quad (21)$$

When value of incremental cost  $\lambda$  was calculated, distribution of power is calculated from equation (18)[2].

#### IV. HYDRO-THERMAL SCHEDULING WITH DYNAMIC PROGRAMMING

If we consider that the input-output characteristic of the steam plant is given by the relation:

$$F = \alpha + \beta \cdot P_T + \gamma \cdot P_T^2, \quad (22)$$

while the characteristic of hydroplant is given by:

$$q = A + B \cdot P_H + C \cdot P_H^2, \quad (23)$$

for  $P_H > P_{min}$  and  $q=0$  for  $P_H=0$ .

For such interval of planning  $j$ , it can be written that volume of storage in the end of that interval is:

$$V_j = V_{j-1} + n_j(r_j - q_j - s_j), \quad (24)$$

where are:  $V_j$ - volume storage in the end of interval  $j$ ,  $V_{j-1}$ - volume storage in the end of interval  $j-1$ ,  $n_j$  - duration of interval  $j$ ,  $r_j$  - inflow of interval  $j$ ,  $q_j$  - flow of interval  $j$ ,  $s_j$  - spillage of interval  $j$ .

If the volume storage on the start of interval  $j$  is  $V_{j-1}=V_i$  and on the end of interval  $j$   $V_j=V_k$ , and spillage is  $s_j=0$ , than the flow  $q_j$  is equal:

$$q_j = \frac{V_i - V_k}{n_j} + r_j. \quad (25)$$

For the obtained flow  $q_j$  it is calculated output power of hydroplant  $P_j$ , and when its power  $P_j$  was reduced for the power of demand  $P_D$ , it is obtained the remaining power that is needed to produce by steam plant. That remaining

power is optimized on the basis of the principle that is that is known as equal incremental cost-loading principle. That principle is given in chapter III.B. If it is obtained that the optimized power of an steam plant is less than the allowed minimal power, than its obtained power is equated with minimal power. After that, process of steam optimization is run again. An analogous procedure is carried out for the case when constraint of maximal power was reached.

The problem of hydro-thermal coordination with dynamic programming is based on finding optimal trajectory of water release whereby the total fuel cost of steam plant must be minimized.

If we consider  $i$  as denoted the volume states on the start of interval  $j$ , and with  $k$  as denoted volume states in the end of interval  $j$ . Further, let with  $TC_k(j)$  is denoted the total cost from the start of scheduling period to the end of period  $j$  for the reservoir storage state  $V_k$ , and with  $PC(i, j-1; k, j)$  is denoted product cost of the thermal system in period  $j$  from an initial volume  $V_i$  to volume  $V_k$  of end of period [1]. Further, algorithm of dynamic programming is given with next equations:

$$TC_k(0) = 0, \quad (26)$$

$$TC_k(j) = \min[TC_k(j-1) + PC(i, j-1; k, j)]. \quad (27)$$

Therewith, the trajectory must be chosen so that the constraints of initial and final volume state are satisfied, such as it is shown on Fig. 4. [1].

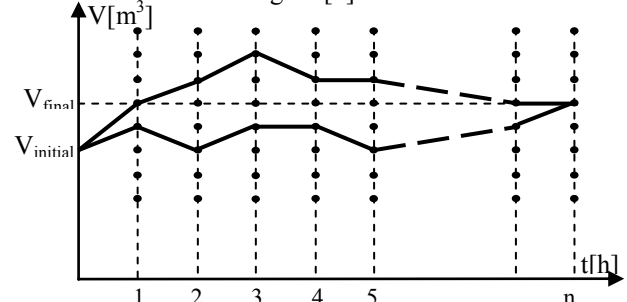


Fig. 4. Possible trajectories of water release

The value of volume  $V_i$  and  $V_k$  can be limited, because of the possible limits of volume states or because of some other non-technical constraints [1].

#### V. THE SOFTWARE IMPLEMENTATION OF DYNAMIC PROGRAMMING FOR HYDRO-THERMAL COORDINATION

##### A. Review of program of hydro-thermal coordination

The program of hydro-thermal coordination, that is realized using MATLAB and GUI interface, is shown on Fig. 5. The program was implemented for coordination of production from a hydroelectric plant and more steam plants.

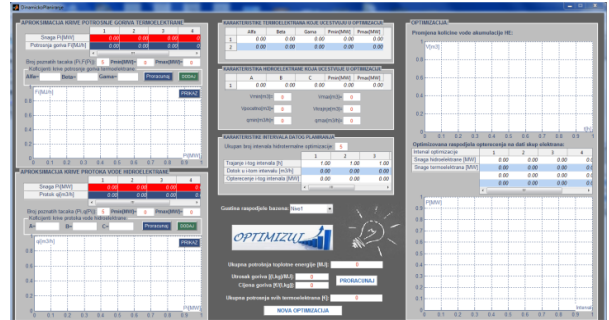


Fig.5. Display of program of hydro-thermal coordination  
Entering basic data is performed via the graphical user interface of interactive objects. Entering parameters that

describing the fuel cost curve of steam plant is carried out by directly entering the coefficients  $\alpha$ ,  $\beta$  and  $\gamma$  or by entering the table of known fuel consumption, on the basis of which it is calculated the aforementioned coefficients. Analogously it is performed input/calculation of  $A$ ,  $B$  and  $C$  coefficients of hydropower characteristic. Besides the power consumption curve coefficients, program requires data of the maximum and minimum power that can be provided by plant. Hydropower plants, for description, need more inputs, unlike steam power plants, which are: minimum and maximum allowed flow rates, minimum and maximum level of the volume of water in the reservoir storage, the initial amount of water in the reservoir and the ultimate level of storage capacity that is needed to accomplish by proper use of water for a given period of planning.

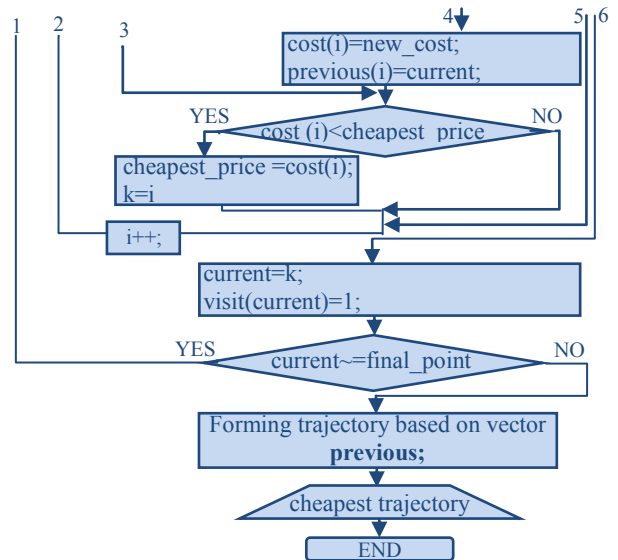
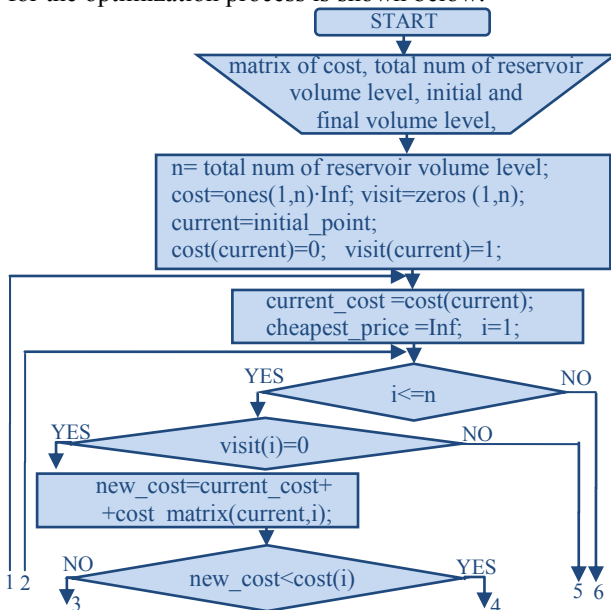
Output data of optimization are shown in tabular and graphical. The table shows the optimized load distribution on a given set of plants and also that load distribution is shown graphically. Also, the graphic is displayed trajectory of volume change of reservoir storage.

The level of precision optimization is adjusted by changing the level of density distribution of the reservoir storage. Of course, the higher the density distribution of reservoir results the greater precision of optimization.

The optimization process begins by forming a matrix of path costs between the individual points of the reservoir volume and by forming vector of redistribution of load between the plants for given path of the changes in the volume reservoir, i.e. it is being formed three-dimensional vector.

*B. Use of Dijkstra's algorithm to determine the cheapest trajectory of storage volume changes*

After matrix of costs were formed, the proces of finding the optimal trajectory is being started. That proces of optimization is realized by using Dijkstra's algorithm that is modified and adapted for this problem of optimization. Dijkstra's algorithm is executed over a matrix of distance between nodes, but for this problem that matrix was replaced with matrix of cost. The algorithm that was used for the optimization process is shown below.



*C. Use of the realized program of concrete example of hydro-thermal coordination*

In addition, it is given an example for hydro-thermal coordination that is taken of reference [1], which was realized by gradient principle for short-term optimization:

HP:  $q=330+4.97 \cdot P_H$ ,  $0MW \leq P_H \leq 1000MW$ .

SP:  $F=500+8 \cdot P_T+0.0016 \cdot P_T^2$  [MBtu/h],

$150MW \leq P_T \leq 1500MW$ , Fuel cost: 1.15 [R/MBtu].

Demand:

1th day: 24:00-12:00 1200 MW, 12:00-24:00 1500 MW

2th day: 24:00-12:00 1100 MW, 12:00-24:00 1800 MW

3th day: 24:00-12:00 950 MW, 12:00-24:00 1300 MW

Hydro-accumulation:

100000 acre-ft-at the beginning, 60000 acre-ft-at the end,  $60000 \text{ acre-ft} \leq V \leq 120000 \text{ acre-ft}$ , 2000 acre-ft/h - inflow.

The optimization of previous example with the realized program produces the total fuel cost of 693428R for the whole period of planing, while the total fuel cost, which is obtained by gradient principle, is 709877R.

VI. CONCLUSION

Efficient and economical use of existing power facilities are some of the basic requirements of the power system. Today there are more methods for optimization, namely coordination of electricity production. The advantage of dynamic programming is that there are no iterations and tests of convergence in its execution, so in case of real request it can be always expected an acceptable solution into practice.

REFERENCES

- [1] Allen J. Wood, Bruce F. Wollenberg, Gerald B. Sheblé, "Power Generation, Operation, and Control", december 2013.
- [2] Mohamed E. El-Hawary, "Electrical Power System-Design and Analysis", 1994. god.
- [3] Prof. dr Ilija Vujosevic, "Eksploatacija i planiranje elektroenergetskih sistema", Lectures, Faculty of electrical engineering in Podgorica, unpublished.
- [4] Prof. dr Slobodan Djukanovic, "Programiranje I", Lectures, Faculty of electrical engineering in Podgorica, unpublished.
- [5] "Dijkstra's Algorithm", Melissa Yan, Massachusetts Institute of Technology-MIT, Available: <http://math.mit.edu/~rothvoss/18.304.3PM/Presentations/1-Melissa.pdf>, Date of access: january 2017.
- [6] "Dijkstra's algoritam za odredjivanje najkracega puta u mrezi", Polytechnic faculty in Rijeka, Available: [http://www.veleri.hr/files/datoteke/nastavni\\_materijali/k\\_promet\\_sl/Kvantitativne\\_4\\_Promet.pdf](http://www.veleri.hr/files/datoteke/nastavni_materijali/k_promet_sl/Kvantitativne_4_Promet.pdf), Date of access: december 2016.

# Performance Modelling of PV Arrays Under Specific Working Conditions

Mladen Mitkovski, Vlatko Cingoski and Saso Gelev

**Abstract** — In this paper modeling and simulation of PV arrays under specific working conditions using Matlab/SIMULINK simulation package is presented. A method for simulation of a PV arrays under specific working condition, especially for analyzing the the so-called hot spot heating effect which usually occurs for non-evenly highlighted solar cell is developed and presented. Mitigation of such problems using bypass diodes is also investigated and the inclusion of these diodes is separately investigated for serially and parallelly connected set of solar modules. The obtained simulated results are also presented and briefly discussed.

**Keywords** — PV arrays, mathematical modeling, hot spot effect, bypass diode.

## I. INTRODUCTION

The need for electricity every year constantly increases. Although the majority of electricity today is still produced by means of burning of fossil fuels, due to negative effects such as pollutant emissions, the global warming and climate changes, renewable energy sources such as solar energy gradually play a significant role in the electricity generation field.

Advances in photovoltaic (PV) solar energy technology contribute significantly towards tremendous increase of solar cells utilization as electricity generators worldwide. Power production using solar cells, tend to become the most popular and reliable source of energy among renewable energy sources in the near future, mostly because PV systems exhibit significantly low carbon dioxide production proving themselves as one of the most environmentally friendly power generation sources.

The solar cells are essentially semiconductors and their operating principle is based on the photovoltaic effect which occurs when the so-called  $p-n$  junction, the connection area between one  $p$ -type and one  $n$ -type semiconductor are joint together, is exposed to light. The low electric potential produced by a single solar cell that depending on the technology used ranges between 0.5 and 0.8 V, is not sufficient for commercial use. However, by

connecting multiple solar cells in series and/or in parallel in PV arrays or so-called photovoltaic modules or panels, one could generate sufficient voltage and currency for practical commercial use.

As the number and size of used modules increases, the possibility that each separate module works under different working conditions during the operation of the whole PV array becomes very common. Cases such as partial shadowing of one or several modules by the surrounding objects, or simply their partial clouding are very common and very unfavorable for the PV generators. PV generators working under these conditions might generate significantly less electricity, their efficiency drops strongly and very often some damages to the structure of the PV array occurred. The main reason is the process called “hot spot heating.” During this process, the generated power in one solar cell working under normal conditions are partially or even entirely consumed by other, shadowed solar cell, converted into additional heat inside the cell producing local overheating and cell damages [1]. A simple yet efficient method for mitigation of these negative effects can be achieved by adding bridge i.e. bypass diodes in parallel with each solar cell.

In this paper, a new practical approach for performance modeling of solar cell working under specific working conditions with and without bypass diodes is investigated. The model was developed in the Matlab/SIMULINK environment and was successfully tested on several parallel and serial connected solar cells. In the proposed method, each solar cell and its bypass diode were modeled separately, providing opportunity for testing various connecting patterns between modules and investigate the efficiency of the whole PV array. Developed model was tested under various specific working conditions such as parallel or serial shadowing of a group of solar cells, various values for intrinsic impedances of the solar cell, temperature and solar radiation intensity in order to grasp the influence of the shadowing effect on the hot spot heating, efficiency of the PV array and the benefits of the bypass diodes.

## II. MODEL OF THE SINGLE SOLAR CELL

The physical properties of a single solar cell were analyzed using an equivalent electric circuit as shown in Fig.1. This electric circuit consists of the current generator  $I_{pv}$ , the serial resistance  $R_s$  and the parallel resistance  $R_p$ , which amounts for several technological parameters of the real solar cells such as contact resistance, impurities or micro defects during the production, and a diode [2].

Mladen Mitkovski, Faculty of Electrical Engineering, University „Goce Delcev“, str. „22 October“ bb, Radovish 2420, Macedonia (tel: 389-32-550033; e-mail: mladen.20234@student.ugd.edu.mk).

Vlatko Cingoski, (corresponding author) Faculty of Electrical Engineering, University „Goce Delcev“, str. „22 October“ bb, Radovish 2420, Macedonia (tel: 389-32-550-033; e-mail: vlatko.cingoski@ugd.edu.mk).

Saso Gelev, Faculty of Electrical Engineering, University „Goce Delcev“, str. „22 October“ bb, Radovish 2420, Macedonia (tel: 389-32-550033; e-mail: saso.gelev@ugd.edu.mk).



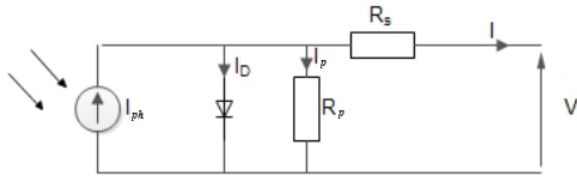


Fig. 1. Real solar cell equivalent electric circuit.

The intensity of the generated solar cell current  $I$  is given as:

$$I = I_{ph} - I_0 \left[ e^{\frac{V - I \cdot R_s}{n \cdot V_T}} - 1 \right] - \frac{V + I \cdot R_s}{R_p} \quad (1)$$

where,  $I_{ph}$  is the current generated by the current generator, the  $I_0$  is the diode's reverse saturation current,  $V_T = kT/q$  is the thermal voltage,  $q$  is the electron charge,  $k$  is the Boltzmann constant,  $T$  is the temperature of the  $p$ - $n$  junction and  $n$  is the diode's ideality factor.

Taking into consideration that a single solar cell voltage at full load ranges between 0.5 to 0.8 VDC and the power generated by a single solar cell is too low, an array of solar cell is usually constructed by interconnecting single solar cells in  $n_p$  parallel and/or  $n_s$  serial branches. Therefore, the total generated power  $P$  by such constructed PV array becomes:

$$V_s = n_s \cdot V_{cell}; \quad I_p = n_p \cdot I_{cell} \quad (2)$$

$$P = V_s \cdot I_p = n_s \cdot n_p \cdot V \cdot I = n_s \cdot n_p \cdot P_{cell} \quad (3)$$

where,  $P_{cell}$ ,  $V_{cell}$  and  $I_{cell}$  are values of generated power, voltage and current of a single solar cell.

Very often different solar modules which belong to a single PV array or panel have to operate under different working conditions. Cases of partial clouding or shadowing of some models are very common and lead to conditions where different solar cells exhibit abnormal working effects which could affect not only those cells but also other interconnected cells in the whole system. Among these effects the most unwanted is the effect of "hot spot heating" effect shown in Fig. 2.

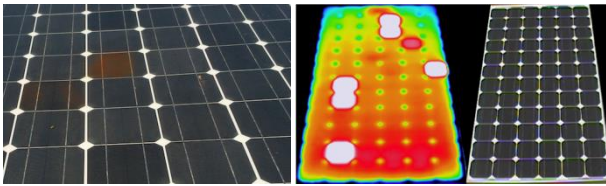


Fig. 2. Hot spot heating of PV panel, a) visual "hot spot" damage, b) infra-red testing of a visibly undamaged panel.

### III. HOT SPOT EFFECT

The shading even of just single cell linked in a long PV array made of strings of serially or parallel connected solar cells could halve the power output of the whole array. To closely investigate this effect, we show an example of a solar module with  $n$  serially connected cells. One shadowed cell is presented separately with its equivalent scheme as shown in Fig. 3. When all other cells are highlighted, the same current  $I$  flow through all cells and the voltage at the ends of the whole module is  $V$ . In case  $n^{\text{th}}$  cell is clouded or shadowed, due to the inverse polarization of its diode, the current through the diode should be zero, thus the entire current generated in the remaining  $(n-1)$  cells of the whole module  $I$ , must pass through resistances  $R_p$  and  $R_s$  of the  $n^{\text{th}}$  cell, causing

voltage drop and reducing the output voltage of the whole module to value  $V_1$ . If we assume that  $(n-1)^{\text{th}}$  cell module still generate the same amount of current  $I$  as in the previous case, then the total voltage  $V_1$  becomes:

$$V_1 = V_{(n-1)} - (R_s + R_p) \cdot I \quad (4)$$

If we further assume that the voltage inside the module is equally distributed through all cells, then:

$$V_{(n-1)} = \left( \frac{n-1}{n} \right) \cdot V \quad (5)$$

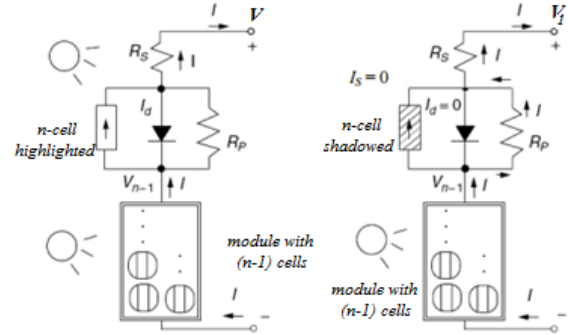
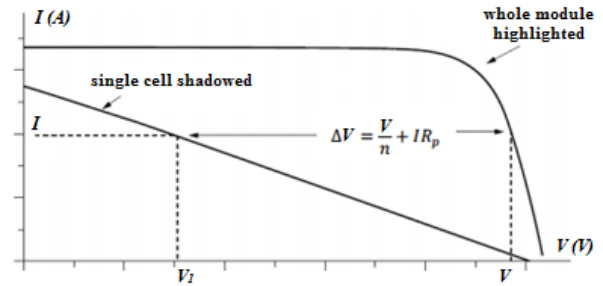


Fig. 3. Comparison between two modules without and with shadowed single PV cell.

By including (5) into (4) and rearranging, one could get the value of the voltage drop at the end of the whole module as a result of shadowing of a single PV cell:

$$\begin{aligned} \Delta V = V - V_1 &= \frac{V}{n} + I \cdot (R_p + R_s) \\ &\cong \frac{V}{n} + I \cdot R_p \quad (R_p \gg R_s) \end{aligned} \quad (6)$$

The effect of the single cell shadowing on the  $V$ - $I$  characteristic of the PV module is presented graphically in Fig. 5. Obviously, the power consumed by shadowed cell is converted into heat, which can cause local overheating and damage of the cell and the whole module.


 Fig. 4. The influence of the shadowed cell on the  $V$ - $I$  characteristic of a single PV module.

Mitigation of these negative effects can be achieved by connecting a so-called *bridging* or *bypass diode* in parallel with each cell, as shown in Fig. 5. When the entire module is evenly highlighted, this diode does not conduct electricity, but if the cell is in shadow, the voltage drop of the cell allows current to flow through the diode instead through the resistance. In comparison with evenly highlighted module, in case of a single shadowed PV cell, the voltage drop in the bypass diode is only about 0.6 V which is far less than the voltage drop of overshadowed cell without bypass diode.

In practice, bridging each solar cell with bypass diode is unfeasible. Instead, bridging is usually done for the whole module, or using of a few diodes for bridging a block of cells within a module. Although these diodes do not have a

major impact on the effect of shade in a single module, they play an important role when several modules are connected in series.

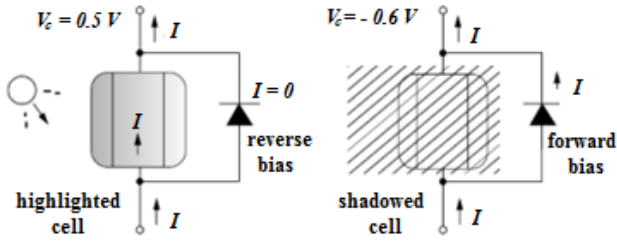


Fig. 5. Mitigation of hot spot effect using bypass diode.

#### IV. MODELING AND SIMULATION OF SOLAR PV CELLS

##### A. Matlab/SIMULINK modeling of solar PV arrays

Matlab<sup>®</sup> is a highly-organized programming language for application in engineering and technology. It incorporates computation, visualization and programming within a simple to use environment where problems and solutions are presented in a familiar and user-friendly mathematical notation and collection of procedures called toolbox. One of those frequently used toolboxes is SIMULINK<sup>®</sup> toolbox, software module for modeling, simulation and analysis of dynamic systems, that supports linear and nonlinear systems, time-continuously and discrete physical models, or any hybrid of these.

There are three options for modeling of solar PV arrays in SIMULINK<sup>®</sup>. The first option includes modeling tools that can implement any differential or algebraic equations into highly complex mathematical model. The second option is given by Simscape<sup>™</sup>, which enables direct modeling of the solar cell with the physical components of its equivalent electric circuit (e.g. resistors, diodes) and by implementation of the same mathematical equation. Finally, the third option, comprise modeling of more complex systems using the library called SimElectronics<sup>®</sup>, which contains a special computation block called "Solar Cell". This block is a source of solar PV power that includes into the model the solar-induced currents and temperature dependences [3].

The "Solar Cell" block is formed by a solar cell equivalent electric circuit consisting of resistor  $R_s$  connected in series with a parallel combination of current source, two exponential diodes and parallel resistor  $R_p$ . Its output current is defined by the following equation:

$$I = I_{ph} - I_{s1} \cdot \frac{\left( e^{\frac{V+I \cdot R_s}{n_1 \cdot V_T}} - 1 \right) - I_{s2} \cdot \left( e^{\frac{V+I \cdot R_s}{n_2 \cdot V_T}} - 1 \right)}{R_p} \quad (7)$$

It is visible that (7) is just a better representation of (1) including two diodes instead of one. Thus, all the parameters in (7) are the same as in (1). The new ones are  $I_{s1}$  and  $I_{s2}$ , and  $n_1$  and  $n_2$  which represent the inverse saturation currents and the ideality factor for the first and the second diode, respectively.

The model „Solar Cell“ used for analysis is given in Fig. 6. The „PV module“ on the left side of Fig. 6, consists of a set of 60 sub-modules named „I-60“ interconnected each other as show in Fig. 7a. Each „I-60“ sub-model is separately modeled using one constant value called „PS Constant“ which represents the value of the sunlight

intensity for each solar cell separately. This new approach, enable us to separately define the working conditions for each solar cell in the module and in the whole PV array independently, setting some of them fully highlighted and some of the partially or fully shadowed.

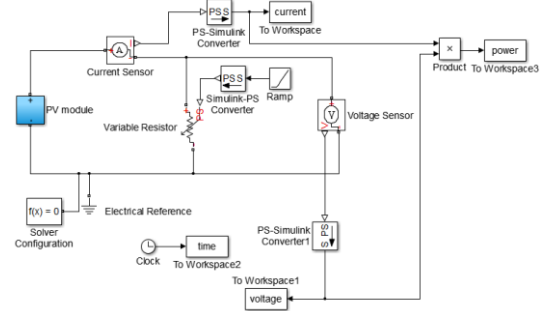


Fig. 6. Matlab/SIMULINK/Simscape analyzed model.

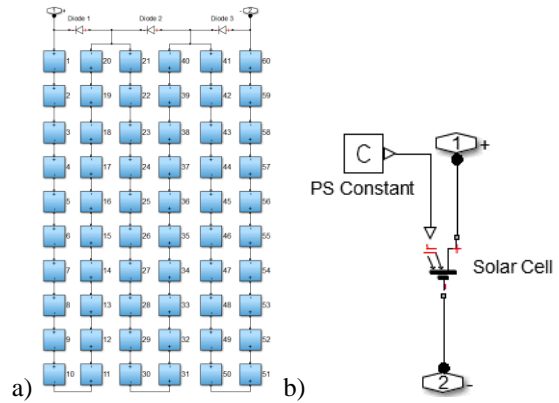


Fig. 7. Solar cell model, a) the whole block, b) single cell.

#### V. OBRAINED RESULTS

##### A. Normal working conditions

For the simulation, we used the data provided from the Canadian Solar polycrystalline PV module CS6P-255P [4]:  $I_s = 2.7098 \cdot e^{-10}$  [A],  $I_{ph} = 9.0012$  [A],  $n = 1.0$ ,  $R_s = 0.0056$  [ $\Omega$ ], and  $R_p = 39$  [ $M\Omega$ ]. The  $I-V$  and  $P-V$  characteristics obtained in the simulation are given in Fig. 8, and Table I.

TABLE I. RESULTS UNDER NORMAL WORKING CONDITIONS.

G [ $W/m^2$ ]	T [ $^{\circ}C$ ]	$P_m$ [W]	$V_m$ [V]	$I_m$ [A]
1000	25	255,33	29,98	8,52

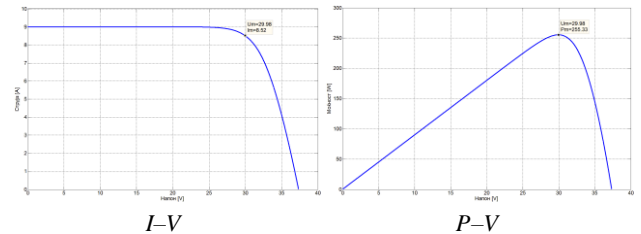


Fig. 8. Characteristic of solar cell under normal conditions.

##### B. Results obtained under hot spot conditions

To investigate the effect of the hot spot effect on the output characteristics of the solar PV array, two specific cases showed in Fig. 9 were investigated:

- Case #1: horizontal shadowing (cells No. 1, 20, 21, 40, 41 and 60), and
- Case #2: vertical shadowing (cells No. 51 – 60).

The obtained results are given in Fig. 10 and Table II.

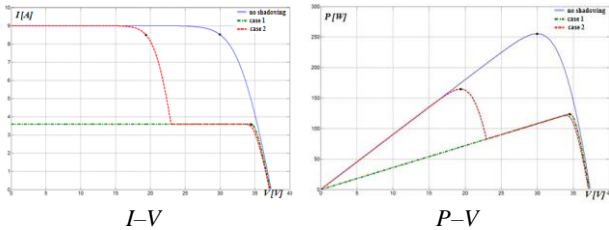
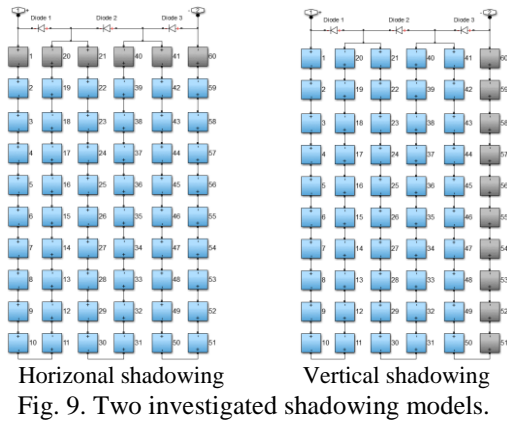


TABLE II. RESULTS UNDER HOT SPOT CONDITIONS.

	$P_m$ [W]	$V_m$ [V]	$I_m$ [A]
No shadowing	255,33	29,98	8,52
Case 1	123,53	34,47	3,58
Case 2	164,70	19,38	8,50

C. The influence of the bypass diode

To analyze the influence of the bypass diode on the power output of PV cells, two separate cases were investigated: a serial connection of one highlighted and one shadowed cell (Fig. 11), and a parallel connection of one highlighted and one shadowed cell (Fig. 12).

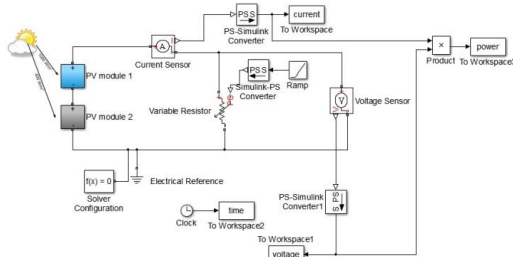


Fig. 11. A model with serially connected solar cells.

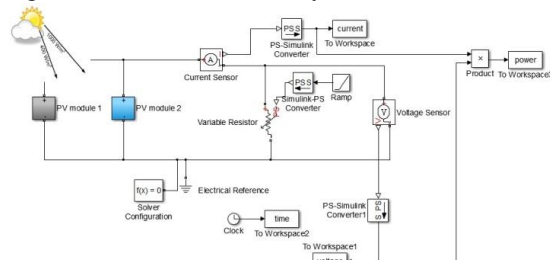


Fig. 12. A model with parallelly connected solar cells.

The obtained results for serial connection are given in Table III, and in Fig. 13, while for parallel connection the results are given in Table IV and Fig. 14.

One can easily conclude that the bypass diode has larger influence in cases of serially connected solar cells than in cases of parallel connected. This could be expected since

in case of serially connected cells the same current should pass through all cells, while in case of parallel connection, each solar cell contributes separately with its generated current. Additionally, the results show that inclusion of bypass diode for mitigation the problem of hot spot effect is advantageous and should be always considered.

TABLE III. RESULTS FOR SERIAL CONNECTION.

	$P_m$ [W]	$V_m$ [V]	$I_m$ [A]
No shadowing	510,67	59,96	8,52
Without bypass diode	226,33	64,45	3,51
With bypass diode	238,77	28,16	8,48

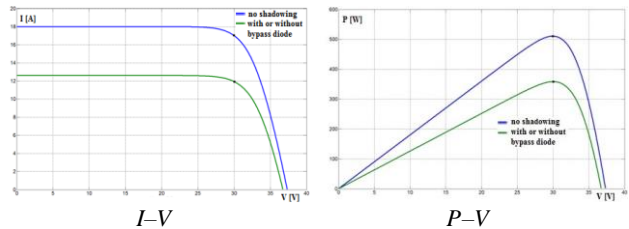
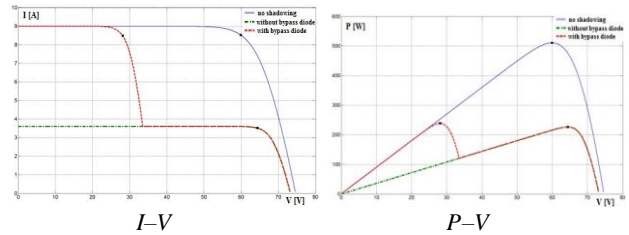


Table IV. RESULTS FOR PARALEL CONNECTION.

	$P_m$ [W]	$U_m$ [V]	$I_m$ [A]
No shadowing	510,67	29,98	17,03
With / Without bypass diode	358,48	30,04	11,93

VI. CONCLUSIONS

We presented a method for modeling of solar PV arrays under specific working conditions. The large portion of the research was focus on investigation of the influence that hot spot effect has on the power performance of PV arrays and the potential mitigation methods for solving of such problem. The obtained results show that using adequately selected and positioned bypass diodes might offer cheap and efficient solution.

REFERENCES

- [1] E. Molenbroek, D.W. Waddington, and K.A. Emery, "Hot Spot Susceptibility and Testing of PV Modules," in Conference: Photovoltaic Specialists Conference, 1991, pp. 547-552.
- [2] B. Dimitrovska, M. Kukuseva, G. Stefanov, V. Cingoski, and V. Sarac, "Analysis and performance of two PV cells under various weather conditions," International Journal for Science, Technics and Innovations for the Industry MTM (Machines, Technologies, Materials), 5/2016, pp. 14-17.
- [3] D. Verma, S. Nema, A. M. Shandilya, S. K. Dash, "Matlab(Simacape) Simulation and Experimental Validation of Solar Photovoltaic System for Performance Analysis Under Varying Environment and Mismatch Conditions," ELELIJ, Vol 4, No 3, August 2015, pp. 39-54.
- [4] Canadian Solar, <http://www.canadiansolar.com/>

# Determination of induction machine parameters by using Particle Swarm Optimization without considering iron losses

Boris Miljanic, Martin P. Calasan, *Student Member, IEEE*, Vladan Vujcic, *Member, IEEE*

**Abstract** - In this paper the determination of induction machine parameters by using Particle Swarm Optimization (PSO) technique is presented. An equivalent circuit without considering the iron losses is analyzed. The circuit parameters are found as the result of the error minimization function between the estimated and maker data. In order to determine the parameters, the values of the phase current and power factor for three slip values of the machine have been used. The accuracy of the usage PSO is analyzed by determining the relative error between the optimized and real value of the phase current and power factor for all three slip value. The obtained results are also compared with results found in literature which are obtained by using Genetic Algorithm (GA). It is concluded that PSO is more efficient in solving the parameter estimation problems.

**Keywords** - Induction motor, PSO, GA, parameters.

## I. INTRODUCTION

THE asynchronous machines, or induction machines, have been widely used due to their robustness, easy application, flexibility, ability to work in harsh environments as well as low cost. The performance characteristics of an induction motor are usually determined from its equivalent circuit. However, the manufacturers do not provide the equivalent circuit parameters, and, therefore, these parameters need to be determined for the purpose of the detailed analysis of steady-state or dynamic operation.

The induction machine parameters are generally determined via the classical no-load and locked-rotor tests [1, 2]. However, these tests cannot be implemented easily. Namely, the no-load and blocked rotor tests are time-consuming tasks, especially if the motor is already coupled to driving equipment. The second classical method for

induction machine parameters determination is based on the acceleration and deceleration tests [3]. However, this method requires special experimental setup for measurements. Beside classical methods, in the literature, different optimization methods for induction machine parameters determination can be found. Concretely, the induction machine parameters can be obtained by using the genetic algorithm - GA [4-8], Particle Swarm Optimization - PSO [9-10], the Bacterial Foraging Technique - BFT [11], the Artificial Neural Network - ANN [12-13], the Artificial Bee Colony Algorithm - ABCA [14], the Shuffled frog-leaping algorithm - SFLA [15], the Dynamic Encoding Algorithm - DEA [16], the Differential Evolution algorithm - DE [17]. Methods based on numerical iterative technique can also be found in literature, such as Newton-Raphson algorithm [18] or Levenberg-Marquardt algorithm [19]. Regardless of which method is used, the machine parameters are computed by using data from machine nameplate and catalog data, as it is presented in [20-21], or by using some results obtained by measurements [19].

In this article, a study has been carried out to determine the model parameters for the steady-state operation of a single rotor circuit induction motor by using PSO algorithm. This algorithm has already been used to this end in certain number of papers [9, 10]. However, the induction machine in “*d*” and “*q*” reference frame is concerned in [9]. On the other hand, in [10], for approximate circuit model the problem formulation uses the starting torque, the maximum torque and the full load torque manufacturer data to estimate the stator resistance, the rotor resistance and the stator leakage reactance parameters. The magnetizing reactance parameter is not considered in this model. For deep bar circuit model, the problem formulation uses the starting torque, the maximum torque, the full load torque, the full load current and the full load power factor manufacturer data to estimate the parameters. In this article, RMS phase current and power factor data that correspond to three different values of the machine slip will be used for determining parameters of the induction machine.

The paper is organized as follows. In Section II the equivalent circuit of the induction machine is presented. Short description of PSO algorithm is given in Section III. The estimation of equivalent circuit parameters by using PSO, and its comparison with results obtained by using GA is presented in Section IV. Concluding remarks are

First author, Boris Miljanic is a student of the Faculty of Electrical Engineering, University of Montenegro, Džordža Vašingtona bb, 81000 Podgorica, Montenegro (phone: 382-69-897-177; e-mail: boris.miljanic@live.ac.me).

Corresponding author, Martin Calasan, is with Faculty of Electrical Engineering, University of Montenegro, Džordža Vašingtona bb, 81000 Podgorica, Montenegro (phone: 382-67-615-237; e-mail: martinc@ac.me).

Third author, Vladan Vujcic, is with Faculty of Electrical Engineering, University of Montenegro, Džordža Vašingtona bb, 81000 Podgorica, Montenegro (phone: +382-20-245-839; e-mail: vladanv@ac.me).

given in Conclusion section.

## II. EQUIVALENT CIRCUIT OF THE INDUCTION MACHINE

Fig. 1 shows an equivalent circuit of the induction machine where  $R_1$ ,  $X_1$ ,  $R_2$ ,  $X_2$ , and  $X_m$  represent the stator resistance, stator leakage reactance, rotor resistance, rotor leakage reactance, and magnetizing reactance, respectively [5].

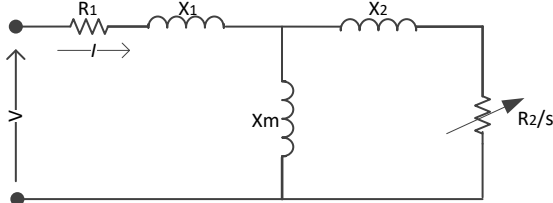


Fig. 1. Equivalent circuit of induction motor without considering the iron losses.

Based on the circuit in Fig. 1, the stator current can be computed as follows:

$$\underline{I} = \frac{\underline{V}}{R_{eq} + jX_{eq}}, \quad (1)$$

where  $R_{eq} + jX_{eq}$  is the equivalent circuit impedance. The equivalent circuit resistance  $R_{eq}$  and reactance  $X_{eq}$  in (1) are equal to:

$$R_{eq} = R_1 + \frac{\frac{X_m^2 R_2}{s}}{\left(\frac{R_2}{s}\right)^2 + (X_2 + X_m)^2}, \quad (2)$$

$$X_{eq} = X_1 + \frac{\left(\frac{R_2}{s}\right)^2 X_m + X_2 X_m (X_2 + X_m)}{\left(\frac{R_2}{s}\right)^2 + (X_2 + X_m)^2}$$

The power factor can be computed as follows:

$$\cos(\varphi) = \cos\left(\tan^{-1} \frac{X_{eq}}{R_{eq}}\right). \quad (3)$$

## III. PARTICLE SWARM OPTIMIZATION ALGORITHM

The PSO is a stochastic optimization method, which uses swarming behaviors observed in flock of birds [9, 10].

The PSO concept consists of changing the velocity of each particle toward its personal best ( $pbest$ ) and global best ( $gbest$ ) locations in each iteration. Acceleration is weighted by a random term, with separate random number generating for acceleration toward  $pbest$  and  $gbest$  locations.

Let  $X$  and  $V$  denote the particle's position and its corresponding velocity in search space, respectively. At

iteration  $K$ , each particle  $i$  has its position defined by:

$$X_i^K = [X_{i,1}, X_{i,2}, \dots, X_{i,N}], \quad (4)$$

and the velocity is defined as

$$V_i^K = [V_{i,1}, V_{i,2}, \dots, V_{i,N}] \quad (5)$$

in search space  $N$ . Velocity and position of each particle in the next iteration can be calculated as

$$V_{i,n}^{k+1} = W \times V_{i,n}^k + C_1 \times rand_1 \times (pbest_{i,n} - X_{i,n}^k) + C_2 \times rand_2 \times (gbest_{i,n} - X_{i,n}^k), \quad (6)$$

where  $i = \{1, 2, \dots, m\}$ ,  $n = \{1, 2, \dots, N\}$ , and

$$X_{i,n}^{k+1} = \begin{cases} X_{i,n}^k + V_{i,n}^{k+1} & X_{\min,i,n} \leq X_{i,n}^{k+1} \leq X_{\max,i,n} \\ X_{\min,i,n} & X_{i,n}^{k+1} < X_{\min,i,n} \\ X_{\max,i,n} & X_{i,n}^{k+1} > X_{\max,i,n} \end{cases} \quad (7)$$

where  $m$  is the number of particle in the swarm,  $N$  is the number of dimensions in a particle,  $K$  is the pointer of iterations,  $V_{i,n}^k$  is the velocity of particle  $i$  at iteration  $k$ ,  $W$  is the weighting factor,  $C_j$  is the acceleration factor,  $rand_j$  is the random number between 0 and 1,  $X_{i,n}^k$  is the current position of particle  $i$  at iteration  $k$ ;  $pbest_i$  is the personal best of particle  $i$  and  $gbest_i$  is the global best of the group.

It should be noted that the first term of formula (7) is the initial velocity of particle, which reflects the memory behavior of particle; the second term "cognition part", represents the private thinking of the particle itself and the third part is the "social" part, which shows the particles behavior stem from the experience of other particles in the population.

The following weighting function is usually used in (7)

$$W = W_{\max} - (W_{\max} - W_{\min}) \times \frac{Iter}{Iter_{\max}} \quad (8)$$

where  $W_{\max}$  and  $W_{\min}$  are the initial and the final weight, respectively,  $Iter$  is the current iteration number and  $Iter_{\max}$  is maximum iteration number [10].

In the above procedures, the convergence speed of each particle could be influenced by the parameters  $C_1$  and  $C_2$  (acceleration factors). The optimization process will modify the position slowly, if the value of  $C_j$  is chosen to be very low. On the other hand, the optimization process can become unstable, if the value of  $C_j$  is chosen to be very high [10].

The constraints of the used optimization technique in the present research work are  $\{R_1, X_1, R_2, X_2, X_m\}$  which must be bounded within some pre-specified limits. These limits

may be mounted as follows

$$\begin{aligned}
 R_1^{\min} &\leq R_1 \leq R_1^{\max} \\
 X_1^{\min} &\leq X_1 \leq X_1^{\max} \\
 R_2^{\min} &\leq R_2 \leq R_2^{\max} \\
 X_2^{\min} &\leq X_2 \leq X_2^{\max} \\
 X_m^{\min} &\leq X_m \leq X_m^{\max}
 \end{aligned} \quad (9)$$

where the superscripts *min* and *max* speak for the minimum and the maximum values of the respective variables. The lower bound ranges enable obtaining accurate results, while speed of convergence depends on constants  $C_1$  and  $C_2$ .

#### IV. ESTIMATION OF EQUIVALENT CIRCUIT PARAMETERS

The approach described in the previous section is applied to a three-phase induction machine (0.75 kW, 380V, 50Hz, 2 poles). Phase current and power factor data for three values of the slip are given in Table 1 [5]. The PSO parameters used in this paper are presented in Table 2.

The criterion for selecting the best individuals in the PSO algorithm is the objective function [5]. The objective function  $F_{\text{objfunct}}$ , used in the present study, is the average error in the power factor and input current for various load points (slip value):

$$F_{\text{objfunct}} = \sum_{i=1}^n \left( \frac{I_{ci}}{I_{mi}} - 1 \right)^2 + \sum_{i=1}^n \left( \frac{\cos(\varphi_{ci})}{\cos(\varphi_{mi})} - 1 \right)^2. \quad (10)$$

Here  $I_{ci}$  and  $\cos(\varphi_{ci})$  are the values computed by (1) and (3).  $I_{mi}$  and  $\cos(\varphi_{mi})$  are measured or analytical values [5]. The variable varies from 1 to 3.

Table 3 shows the equivalent circuit parameters estimated by using PSO as well as with genetic algorithm [5]. The comparison of the value of the phase current and power factor, for parameters determined by using GA and PSO, are presented in Table 4. The phase current – slip and the power factor – slip characteristics determined by using results presented in Table 3 are shown in Figs. 2 and 3, respectively.

As it can be seen from Figs. 2-3, and from Table 3, PSO enable obtaining much more accurate result in comparison with usage of GA.

TABLE 1: SETS OF USED DATA

Stator current	Slip	Power factor
1.86	0.06	0.62
2.39	0.10	0.74
3.07	0.15	0.78

TABLE 2: PSO PARAMETERS.

$c_1$	$c_2$	$W_{\min}$	$W_{\max}$	$Iter_{\max}$
0.5	0.1	0.2	0.5	200

TABLE 3: ESTIMATED CIRCUIT PARAMETERS.

Parameter	GA [ $\Omega$ ] [5]	PSO [ $\Omega$ ]
$R_1$	10.28	10.1145
$X_1$	8.19	7.4935
$R_2$	10.48	10.5247
$X_2$	19.21	19.9740
$X_m$	143.17	144.1155

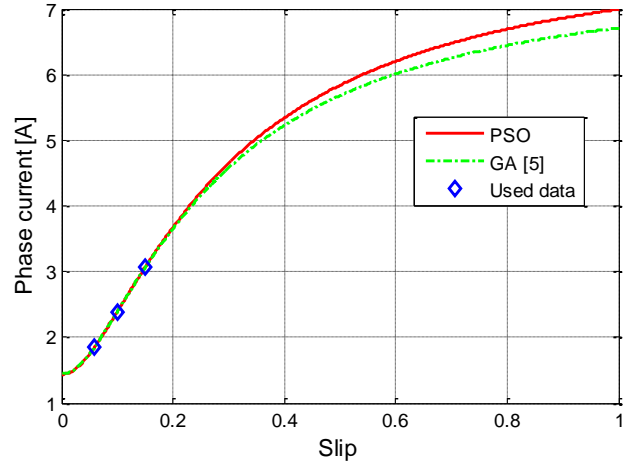


Fig. 2. Phase current – slip characteristics.

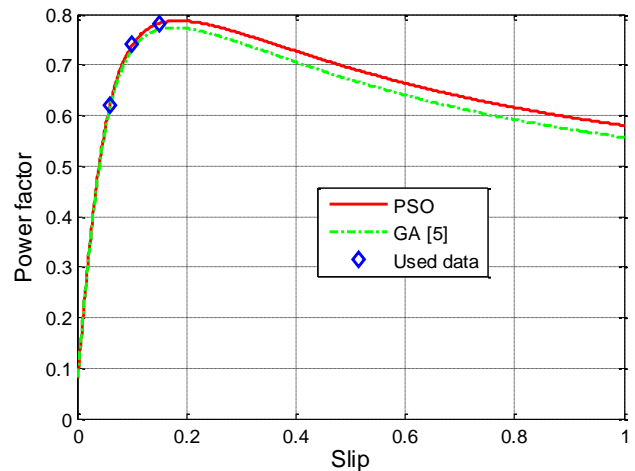


Fig 3. Power factor – slip characteristics.

TABLE 4: Comparison of results

Stator current	Stator current - GA	Error GA	Stator current - PSO	Error PSO
1.86	1.863	-0.0030	1.8588	0.0012
2.39	2.3945	-0.0045	2.3925	-0.0025
3.07	3.0617	0.0083	3.0688	0.0012
Power factor	Power factor - GA	Error GA	Power factor - PSO	Error PSO
0.62	0.6157	0.0043	0.6207	-0.0007
0.74	0.7292	0.0108	0.7375	0.0025
0.78	0.7699	0.0101	0.7815	-0.0015

## V. CONCLUSION

An efficient approach for determining the equivalent circuit parameters of squirrel cage induction motors, based on usage of PSO, is presented.

Although the technique of using PSO for determining parameters of induction motor is not new, novel approach for its implementation have been proposed in this paper. Namely, values of the phase current and power factor that correspond of some values of slip are used as input data in calculations. The obtained results are compared with results found in literature (obtained by using GA). It is shown that PSO enable obtaining much more accurate results in comparison with usage of GA. Also, the usage of PSO is simpler, faster, less intrusive and cheaper than the conventional experimental methods for estimating the equivalent circuit parameters of induction motors.

Furthermore we have obtained more accurate results compared to the results presented in [5], it cannot be concluded that parameters obtained by using PSO are representative for the entire slip range since in this procedure three values of phase current (and power factor) for three different small values of the slip have been used. In the future work this algorithm should be applied for a widespread slip range. Also, different objective function will be considered.

## REFERENCES

- [1] Test Procedure for Polyphase Induction Motors and Generators. IEEE Standard 112, 2004.
- [2] Rotating Electrical Machines. Part 2-1: Standard Methods for Determining Losses and Efficiency from Tests, IEC Standard 60034-2-1, 2007.
- [3] M. Despalatovic, Martin Jadric, B. Terzic, "Identification of Induction Motor Parameters from Free Acceleration and Deceleration Tests", *Automatica*, Vol. 46, pp.123-128, 2005.
- [4] P. Simonik, P. Hudecek, P. Palacky, and S.David, "Estimation of Induction Machine Electrical Parameters Based on the Genetic Algorithms", *Progress In Electromagnetics Research Symposium Proceedings*, KL, Malaysia, March 2012, pp. 999-1002.
- [5] I. Kostov, V. Spasov, V. Rangelova, "Application of genetic algorithms for determining the parameters of induction motors", *Technical Gazette*, vol. 16, no. 2, pp. 49-53, 2009.
- [6] B. Abdelhadi, A. Benoudjit, N. N. Said, "Identification of Induction Machine Parameters Using a New Adaptive Genetic Algorithm" *Electric Power Components and Systems*, vol. 32, pp. 767-784, 2004.
- [7] K. S.Huang, W. Kent, Q. H. Wu, and D. R. Turner, "Parameter identification for FOC induction motors using genetic algorithms with improved mathematical model." *Electric Power Components and Systems*, vol. 29, no. 3, pp. 247-258, 2001.
- [8] H. R. Mohammadi, A. Akhavan, "Parameter Estimation of Three-Phase Induction Motor Using Hybrid of Genetic Algorithm and Particle Swarm Optimization", *Journal of Engineering*, vol. 2014, pp.1-6, 2014.
- [9] C. Guangyi, G. Wei, H. Kaisheng, "On Line Parameter Identification of an Induction Motor Using Improved Particle Swarm Optimization", *Proceedings of the 26th Chinese Control Conference July*, Zhangjiajie, Hunan, China, 2007, pp.745-749.
- [10] V. P. Sakthivel, R. Bhuvaneswari, and S. Subramanian, "Multiobjective parameter estimation of induction motor using particle swarm optimization," *Engineering Applications of Artificial Intelligence*, vol. 23, no. 3, pp. 302-312, 2010.
- [11] V. P. Sakthivel, R. Bhuvaneswari, S. Subramanian, "Bacterial Foraging Technique Based Parameter Estimation of Induction Motor from Manufacturer Data", *Electric Power Components and Systems*, vol. 38, pp. 657-674, 2010.
- [12] M.A. Jirdehi, A. Rezaei, "Parameters estimation of squirrel-cage induction motors using ANN and ANFIS", *Alexandria Engineering Journal*, vol. 55, 2016., pp. 357-368.
- [13] M. Wlas, Z. Krzeminski, and H. A. Toliyat, "Neural-network-based parameter estimations of induction motors," *IEEE Trans. Ind Electr*, vol. 55, no. 4, pp. 1783-1794, 2008.
- [14] A.G. Abro, J.M. Saleh, "Multiple-global-best guided artificial bee colony algorithm for induction motor parameter estimation", *Turkish Journal of Elect Eng & Comp Sci*, Vol. 22, pp. 620 - 636, 2014.
- [15] M. Gomez-Gonzalez, F. J. Perez, "Shuffled frog-leaping algorithm for parameter estimation of a double-cage asynchronous machine", *IET Electric Power Applications*, vol. 6, Iss. 8, pp. 484-490, 2012.
- [16] J.W. Kim, S.W. Kim, "Parameter identification of induction motors using dynamic encoding algorithm for searches (DEAS)", *IEEE Trans Energy Conv*, vol. 20, pp. 16-24, 2005.
- [17] M. Ali, M. Pant, A. Abraham, V. Snasel, "Modified differential evolution algorithm for parameter estimation in mathematical models", *IEEE International Conference on Systems Man and Cybernetics*, pp 2767-2772, 2010.
- [18] J. Pedra, F. Córcoles, "Estimation of Induction Motor Double-Cage Model Parameters from Manufacturer Data", *IEEE Trans Energy Conv*, vol. 19, no. 2, pp. 310-317, sept. 2004.
- [19] M. Akbaba, M. Taleb, A. Rumeli, "Improved estimation of induction machine parameters", *Electric Power Systems Research*, vol. 34, pp.65-73, 1995.
- [20] M. H. Haque, "Determination of NEMA Design Induction Motor Parameters From Manufacturer Data", *IEEE Trans Energy Conv*, vol. 23, No. 4, pp. 997-1003, Dec. 2008
- [21] J.M.C. Guimaraes, J.V. Bernardes, A. E. Hermeto, and E.C. Bortoni, "Parameter Determination of Asynchronous Machines From Manufacturer Data Sheet", *IEEE Trans Energy Conv*, vol. 29, No. 3, pp. 689-697, sept. 2014

# Jedan primjer kontrole invertora napona korišćenjem mikrokontrolera

Aleksandar Dedić

**Sadržaj** — U ovom radu objašnjen je rad invertora napona i jedan metod generisanja impulsa za upravljanje tranzistorima. Sklop za upravljanje tranzistorima originalno je dizajniran a program za mikrokontroler prilagođen je tom sklopu.

**KLjučne reči** — Invertor, Mikrokontroler

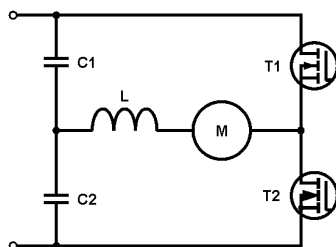
## I. UVOD

Invertor napona, čiji opis je dat u ovom radu, namijenjen je za generisanje monofaznog naizmjeničnog napona vrijednosti 115 V i frekvencije 60 Hz. Osnovna namjena invertora je pogon motora, ali on se može koristiti i za druge namjene. Karakteristike invertora su mogućnost promjene frekvencije u granicama 1 – 70 Hz i mogućnost postepenog povećavanja frekvencije. Invertor generiše napon stepenastog oblika koji je blizak sinusoidalnom u mjeri koja je dovoljno dobra za praktičnu primjenu. Ovdje nijesu navedene karakteristike konkretnog invertora jer one zavise od upotrijebljenih tranzistora.

Za upravljanje tranzistorima u invertoru koristi se originalno dizajnirani upravljački sklop a za formiranje upravljačkih impulsa, mikrokontroler. S obzirom na specifičnost rada upravljačkog sklopa, za mikrokontroler je urađen originalni program. Karakteristika programa je što je on jednostavan i kratak.

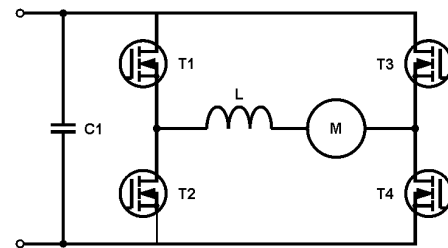
## II. INVERTOR

Invertor služi za dobijanje naizmjeničnog napona sinusoidalnog oblika. U konkretnom slučaju, napon na izlazu nije sinusoidalan nego ima stepenasti oblik ali je i takav oblik kod praktične primjene prihvatljiv u većini slučajeva.



Sl. 1 Polumostni invertor napona

Na slikama 1. i 2. prikazani su najčešće korišćeni oblici invertora: polumostni (slika 1) i mostni (slika 2). Uobičajeno je da se polumostni invertor koristi za generisanje nižih napona, naprimjer 115 V, a mostni za generisanje viših napona, na primjer 230 V. Detaljnija objašnjenja rada ovih invertora mogu se naći u [1], [2] i [3].



Sl. 2 Mostni invertor napona

Upravljački sklop kod mostnog invertora je isti kao kod polumostnog jer se kod mostnog invertora tranzistor T4 otvara istovremeno kad i T1 a T3 istovremeno kad i T2. Induktivnost L služi za filtriranje viših harmonika a stavlja se samo ako to zahtijeva potrošač. U konkretnom slučaju se ta induktivnost nije koristila.

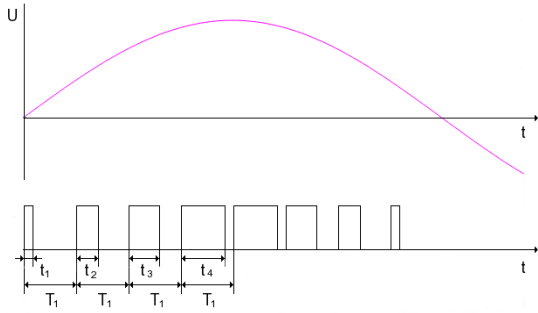
## III. NAČIN RADA INVERTORA

Za dobijanje približno sinusoidalnog napona na izlazu invertora koristi se širinsko – impulsna modulacija. Detaljnija objašnjenja o ovom obliku modulacije data su u [1 - 3]. Kod invertora opisanog ovdje, sinusoida izlaznog napona dijeli se na određeni broj djelova N jednakog trajanja T. Broj djelova N je uvijek konstantan a frekvencija napona se mijenja promjenom trajanja T svih djelova sinusoide. Na slici 3 i 4 prikazane su prve polovine sinusoide i njihovi djelovi. Zbog jednostavnijeg objašnjenja obje sinusoide su podijeljene na 16 djelova odnosno četvrtine periode su podijeljene na 4 dijela. Sinusoida sa slike 3 odgovara frekvenciji 60 Hz a sa slike 4, frekvenciji 50 Hz. Sa prikazanih slika se može vidjeti da je trajanje T<sub>1</sub> manje od trajanja T<sub>2</sub>.

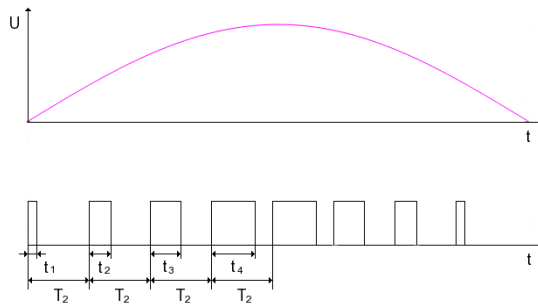
Širinsko-impulsna modulacija dobija se promjenom vremena trajanja impulsa t pri konstantnoj vrijednosti T. Kod ovog načina rada invertora pojedinačne vrijednosti t<sub>x</sub> ostaju konstantne pri svim frekvencijama. Sa slika 3 i 4 vidi se da se vrijednosti T<sub>1</sub> i T<sub>2</sub> razlikuju ali su t<sub>1</sub>, t<sub>2</sub>, t<sub>3</sub>, i t<sub>4</sub> iste u oba slučaja. Kao posljedica ovoga, pri promjeni frekvencije dolazi do promjene odnosa impuls/pauza. Ovim se dobija da se napon na izlazu invertora mijenja proporcionalno sa frekvencijom. Promjena napona sa



frekvencijom je neophodna za pravilan rad motora. Detaljnije o tome se može vidjeti u [3].

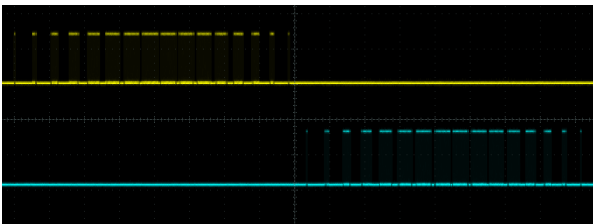


Sl. 3 Širinsko-impulsna modulacija za frekvenciju 60 Hz

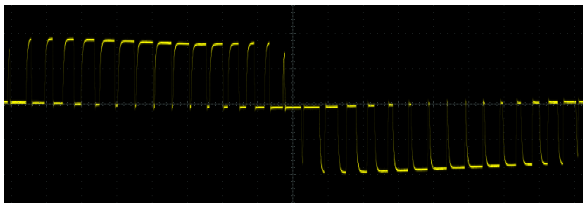


Sl. 4 Širinsko-impulsna modulacija za frekvenciju 50 Hz

Na slici 5 prikazan je oblik upravljačkih impulsa koji se dovode na tranzistore invertora. Sa lijeve strane prikazani su impulsi koji se dovode na gejt tranzistora T1 a sa desne na gejt tranzistora T2.



Sl. 5 Upravljački impulsi



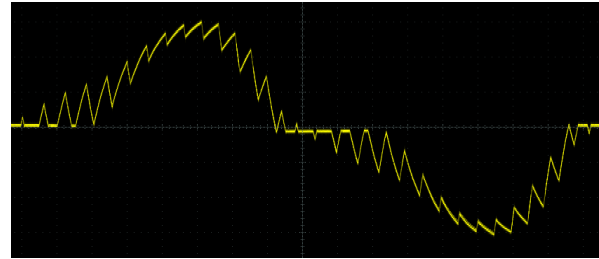
Sl. 6 Napon na izlazu invertora

Na slikama 6, 7 i 8 prikazan je rad invertora za slučaj kada je sinusoida podijeljena na 32 dijela. Ovako mali broj odbiraka uzet je samo zbog jednostavnijeg objašnjenja

njegovog rada. Za praktičnu primjenu poželjno je da broj odbiraka bude najmanje 256.

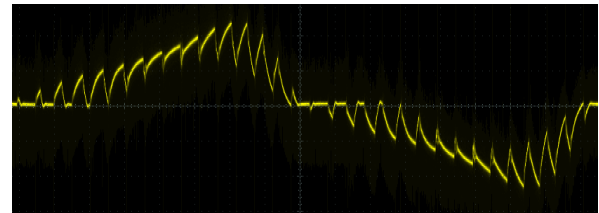
Sva mjerenja, čiji su rezultati prikazani ovdje, vršena su pri frekvenciji 60 Hz.

Na slici 6 prikazan je napon na izlazu invertora a na slici 7 napon na motoru doveden preko induktivnosti 10 mH.



Sl. 7 Napon na motoru

Na slici 8 prikazan je oblik struje kroz motor pri čemu je induktivnost bila isključena. Mjerenje je vršeno sa motorom snage 30 W. Sa slike se vidi da struja ima impulsni oblik.



Sl. 8 Struja kroz motor

Ovaj oblik struje nije dobar jer sa ovakvom strujom motor gubi na snazi. Vezivanje induktivnosti na red sa motorom takođe nije dobro rješenje jer ona vrši korekciju oblika struje ali izaziva smanjenje snage. Preporučljivi način ispravljanja ovog nedostatka je povećanje broja odbiraka sinusoide. Što je veći taj broj, oblik struje se sve više približava sinusoidalnom.

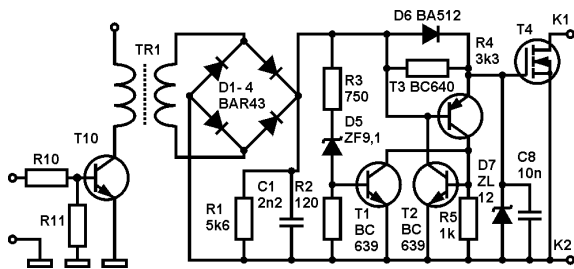
#### IV. UPRAVLJANJE TRANZISTORIMA

Za upravljanje tranzistorima invertora koristi se sklop prikazan na slici 10. Sklop je originalno dizajniran a njegova osnovna namjena je da obezbijedi galvansko odvajanje mikrokontrolera od svih tranzistora invertora.

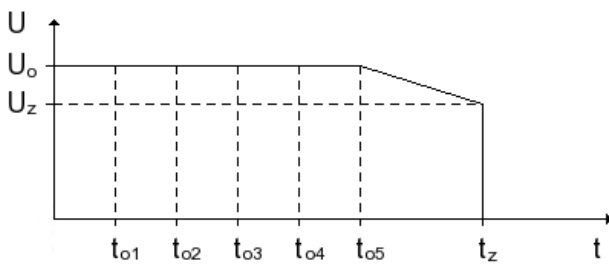
Način rada sklopa sastoji se u tome da, kad treba otvoriti tranzistor T4 invertora, mikrokontroler generiše povorku impulsa. Tranzistor je otvoren sve dok postoje impulsi. Kad prestane generisanje impulsa, tranzistor se zatvara. Karakteristika ovog sklopa je da obezbijuje stabilan napon na gejtu tranzistora invertora u toku vremena u kojem se dovode impulsi iz mikrokontrolera. Ovim se postiže da napon na gejtu može imati proizvoljno trajanje.

Rad ovog sklopa sastoji se u sljedećem: mikrokontroler generiše impulse koji se dovode na bazu tranzistora T10. Ovi impulsi otvaraju i zatvaraju pomenuti tranzistor čime vrše uspostavljanje i prekidanje struje kroz primar transformatora TR. Na sekundaru transformatora

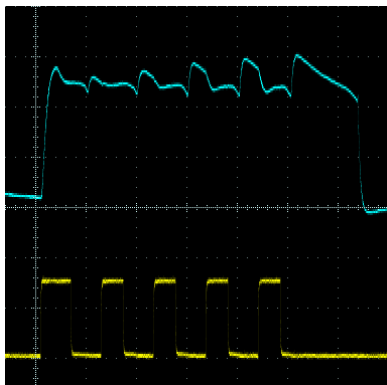
pojavljuje se napon koji se ispravlja diodama D1 – D4 i filtrira kondenzatorom C1. Ovaj napon se dovodi na gejt tranzistora invertora čime se vrši otvaranje tog tranzistora. Da ne bi došlo do oštećenja tranzistora, napon se ograničava na 12 V zener diodom D7. Proces zatvaranja tranzistora počinje kad mikrokontroler prestane da generiše impulse. Tada napon na C1 počinje da opada do vrijednosti pri kojoj se aktivira triger sastavljen od tranzistora T2 i T3 koji spušta napon na gejt tranzistora T4 ispod nivoa napona praga i time ga zatvara. Tranzistor T1 služi da spriječi reagovanje trigerera zbog smetnji.



Sl. 10 Sklop za upravljanje tranzistorima invertora



Sl. 11 Način rada sklopa za upravljanje tranzistorima



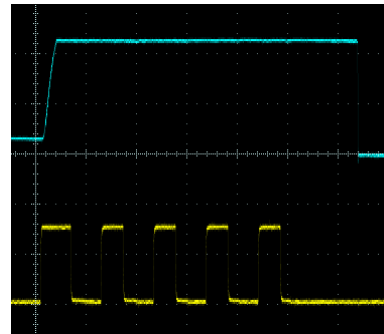
Sl. 12 Napon na kondenzatoru C1

Na slici 11 prikazan je uprošćeni dijagram rada ovog sklopa. U trenucima  $t_{o1} - t_{o5}$  na ulaz se dovode impulsi iz mikrokontrolera pa je napon na gejt tranzistoru T4 konstantan i iznosi  $U_o$ . U trenutku  $t_{o5}$  prestaju da se dovode impulsi pa napon opada do vrijednosti  $U_z$ . U trenutku  $t_z$

reaguje triger koji kratko spaja gejt tranzistora T4 i napon pada na nulu.

Od trenutka  $t_{o5}$  prestanka impulsa do trenutka  $t_z$  zatvaranja tranzistora pojavljuje se određeno kašnjenje koje zavisi od vrijednosti  $R1/C1$  na slici 11. Za date vrijednosti kašnjenje iznosi  $5\mu s$ . Zbog ovog kašnjenja dolazi do proširenja impulsa kod širinsko-impulsne modulacije odnosno do povećanja napona na izlazu invertora u odnosu na traženu vrijednost. Kod praktične primjene ovo povećanje se može zanemariti ali u slučaju potrebe vrši se skraćivanje impulsa.

Na slici 12 prikazan je napon na kondenzatoru C1 (gornji dijagram) u odnosu na impulse na izlazu mikrokontrolera (donji dijagram).



Sl. 13 Napon na gejt tranzistora T4

Na slici 13 prikazan je napon na gejt tranzistora T4 invertora (gornji dijagram) u odnosu na impulse na izlazu mikrokontrolera (donji dijagram).

Frekvencija ovih impulsa iznosi 200 kHz a ona je odabrana prema konstrukciji transformatora.

Eksperimentalno se pokazalo da je rad ovog sklopa izuzetno stabilan i neosjetljiv na smetnje.

Prednosti ovog sklopa u odnosu na postojeća rješenja je dobro galvansko odvajanje, oblik napona koji se dovodi na gejt tranzistora ne zavisi od karakteristika tranzistora i opterećenja jer je obezbijedena velika struja punjenja i pražnjenja kapacitivnosti gejta, male dimenzije transformatora jer radi na frekvencijama nekoliko stotina kiloherca i mogućnost da napon na gejt ima proizvoljno vrijeme trajanja.

Ovaj sklop je predviđen za upravljanje samo jednim tranzistorom invertora. Kod polumostnog invertora moraju postojati 2 a kod mostnog, 4 ovakva sklopa.

## V. MIKROKONTROLER

Za upravljanje tranzistorima koristi se mikrokontroler dsPIC33FJ64GP802 proizvodnje Microchip [4]. Razlog za korišćenje ovog mikrokontrolera je taj što on ima određene hardverske dodatke koji uprošćavaju program za generisanje širinsko-impulsne modulacije.

Za generisanje impulsa za upravljanje tranzistorima invertora koriste se 2 tajmera. Prvi tajmer se koristi za generisanje povorke impulsa frekvencije 200 kHz koji se dovode na tranzistor T10 na slici 10. Kod ove povorke trajanje impulsa iznosi 40% a pauze 60%.

Za širinsko – impulsnu modulaciju koristi se drugi tajmer. U sklopu ovog tajmera koriste se 2 komparatora. Jedan komparator se koristi za određivanje frekvencije na izlazu invertora. U njega se unosi vrijednost trajanja odbirka napona, na primjer T1 sa slike 3 ili T2 sa slike 4. Ova vrijednost se mijenja samo ako se mijenja frekvencija. Drugi komparator se koristi za start i zaustavljanje prvog tajmera. U njega se unose trenutne vrijednosti trajanja impulsa  $t_x$  (slike 3 i 4)

Da bi napon na izlazu invertora imao približno sinusoidalni oblik, potrebno je vrijednost trajanja impulsa  $t_x$  izračunavati korišćenjem trigonometrijskih funkcija. Kako je program za takva izračunavanja izuzetno složen za primjenu u mikrokontroleru, koristi se tabelarni oblik pojedinih vrijednosti  $t_x$ . Tabela se sastavlja samo za prvu četvrtinu sinusoide jer su ostale četvrtine simetrične.

## VI. ZAKLJUČAK

Invertor, čiji je opis dat ovdje, pokazao je zadovoljavajuće rezultate u konkretnoj primjeni. U opštoj primjeni može da dođe do pojave određenih problema pri radu motora kod niskih frekvencija. Ovi problemi nastaju zbog suviše kratkog trajanja naponskog impulsa u odnosu na pauzu usljed čega struja kroz motor dobija impulsni oblik. Ova pojava se može ublažiti ako se puna perioda sinusoide podijeli na veći broj djelova od gore preporučenog. U ovom slučaju treba voditi računa da se trajanje impulsa napona mora proporcionalno smanjiti jer napon mora biti smanjen proporcionalno frekvenciji. Za

praktičnu primjenu u većini slučajeva nije potrebno praviti novu tabelu trajanja impulsa, već samo smanjivati postojeće vrijednosti. Na primjer, ako se sinusoida podijeli na dva puta veći broj djelova od preporučenog, trajanje svakog impulsa treba da bude dvostruko kraće.

## LITERATURA

- [1] N. Mohan, T. M. Undeland, W. R. Robbins, *Power Electronics Converters, Applications and Design*, John Wiley & Sons Inc. 2006
- [2] M. H. Rashid, *Power Electronics Circuits, Devices and Applications*, Pearson Education Inc. 2004
- [3] B. Bose, *Power Electronics and Motor Drives*, Elsevier Inc. 2006
- [4] Microchip Technology Inc, *dsPIC33FJxxxGPx02x04 Data Sheet*, <http://support.microchip.com>

## ABSTRACT

This paper describes the operation of a voltage inverter and a method for pulse generation for transistors control. Transistors control system is originally designed and microcontroller code is adapted to the proposed system.

## ONE EXAMPLE OF VOLTAGE INVERTER CONTROL USING A MICROCONTROLLER

Aleksandar Dedić

# E-government as a business model of e-commerce in the Republic of Serbia

J. Končar, S. Leković, Z. Petrović Katai, *University of Novi Sad, Faculty of Economics in Subotica*

**Content** —There are a lot of basic business models of e-commerce in the global market. Despite the multitude of such business models that are fundamentally similar and often appear in more than one sector, it is important to point out the model e-government. This model is a model of electronic commerce where entities of government purchase and provides products, services or information to businesses or individual citizens. Basic categories and leading business models of e-government are: government-to-citizens G2C, government-to-business G2B, government-to-administration G2G and government-to-employees G2E. E-government in the Republic of Serbia is a form of e-business of government and refers to the delivery of electronic services to different target groups and the public.

**Keywords**— e-commerce business model, e-government, G2C, G2B, G2G, G2E.

## I. INTRODUCTION

ELECTRONIC commerce as one of the most important areas of electronic business has been seen as a new factor to strengthen the competitiveness in the global market. New business models of e-commerce are created thanks to the new forms of business. Tools and applications of e-commerce are being increasingly used to improve the performance of public institutions as well as state and city governments. More convenient access to information and services of the public sector is provided to citizens and businesses with electronic government, through the application of information technology. This creates an efficient and effective management of business transactions of government for citizens and businesses. Defining performance of e-government requires answers to the following questions:

- What are the basic categories of e-government involving all the interactivity between government and its citizens?
- What are the characteristics of transformation and implementation of e-government in the Republic of Serbia?

## II. CATEGORIES OF BUSINESS MODELS OF E-GOVERNMENT

E-government is a model of e-commerce, where e-government exchanges products, services and information to individuals and business entities [1, p. 398]. With that citizens and businesses provides convenient access to government information and services. By conducting transactions with citizens and businesses government leads

to efficient business transactions within the administration itself. In broader terms "e-government offers an opportunity to improve the efficiency and effectiveness of the functions of government and to make government more transparent to citizens and business by providing access to more of the information generated by government, as well as facilitating transactions with and within government" [2, p. 259]. The following business models correspond to this wider category of e-government: government-to-citizens G2C, government-to-business G2B, government-to-government G2G, internal efficiency and effectiveness IEE, and government-to-employees G2E [3, p. 232]. In addition to these business models m-government is increasingly in use as a wireless implementation of e-government, mainly for citizens but is also used for legal entities. E-government thus creates the process of creating single electronic counters in order to improve public electronic services of competent authorities at state level, district and city.

Government to citizens G2C category of business model of e-government covers interactivity between administration and its citizens electronically from anywhere and at any time. Administration services are provided to citizens through civil portal, and are tied to a specific state level of government: state, county and city. Web Sites of administration include all necessary contact information, links to other sites, publications and databases. G2C today in many countries works with the use of wireless mobile devices. The main objectives of G2C model of e-government are related to: creating easy finding points of access to services for individuals, reducing the average time, finding for citizens, meeting the high demands for information to citizens thus improving the overall value of administration for citizens.

Government to business G2B is a category of business model of e-government, and includes interaction between government and enterprises. Business model of e-commerce G2B consists of activities of administration that sells products or provides services to business entities as well as the activities of companies that sell products and services of administration. Objectives of G2B model is related to finding and reviewing the rules and regulations, in order to make transactions with the administration easier, cheaper, fast and comprehensive.

Government to government G2G e-government category includes activities within the administration unit and one between administrations. Thanks to G2G business model leads to faster and more efficient routing of information between federal, state and local governments, at lower costs, while reducing response times of verifying the correct information.

Government to employees G2E includes activities and

services between government units and citizens. Employees in the public administration operate in different geographic locations; where application for G2Eis necessary for effective communication.

Mobile government (m-government) consists of wireless form of implementation of e-government with citizens and businesses. M-government helps availability of public administration information and services to users. Reduction of costs, increased efficiency of services, transformations and modernizations of the public sector organizations all came as a result of m-government.

Business models for e-government allow the organization of the public sector in order to increase efficiency, transparency, ease of access and responsiveness to the demands of the citizens with the use of information technology. The index of the development of e-government in the ten leading countries can be seen in Table 1.

TABLE 1: E-Government Development Index - Top 10 Countries [4]

No.	Country	Index
1.	United Kingdom	0.9193
2.	Australia	0.9143
3.	Republic of Korea	0.8915
4.	Singapore	0.8828
5.	Finland	0.8817
6.	Sweden	0.8704
7.	Netherlands	0.8659
8.	New Zealand	0.8653
9.	Denmark	0.8510
10.	France	0.8456

Effective exercise of the rights and interests of citizens improves the quality of life of all citizens. Prerequisites for providing complete e-services to citizens are related to access to the Internet, adequate communication and information infrastructure, database, and application processing, access and exchange of data. The successful development and implementation of the transformation of e-government requires an effective governance structure that will plan, coordinate, realize and evaluate the e-government strategy.

### III. CHARACTERISTICS OF TRANSFORMATION AND IMPLEMENTATION OF E-GOVERNMENT IN THE REPUBLIC OF SERBIA

The transformation of traditional government services to the implementation of full on-line service presents a very long process some times. The main problem is related to the implementation of maintaining control over data usage. Leading software companies in the global market provide tools and solutions for managing e-government. Agencies of government at global level, thanks to the tools of social networking and its own presence on the site of public social networks provide dissemination of information. The use of new information and communication technologies has enabled citizens to do their jobs from home without waiting in lines, 24 hours a day, 7 days a week, 365 days a year.

All this indicates that the implementation of the concept of e-government is a complex task that requires solving a

series of complex issues and problems. Area of implementation of e-governance should include planning and organizing e-government, creating the right environment for the development of e-government, as well as the development of public services and facilities of e-government. The implementation of e-government is based on the key principles of good governance, namely: transparency, interactivity and civic participation.

The successful implementation of e-government in the Republic of Serbia includes the following phases: development of legislation, building architecture, creation of infrastructure, organization of personnel resources, administration connectivity, as well as the promotion of e-government services.

Effects of implementation of e-government in the Republic of Serbia depend on the individual stakeholders: government, citizens and businesses. Direct effects of implementation are linked to the effectiveness and efficiency, development of infrastructure, broader citizen participation and the creation of new services. E-government is becoming an essential requirement and makes the standard of development of the country.

Benefits of e-government referred to the simplification of services, better and much cheaper; reduced the time in obtaining information or administrative acts; there is greater transparency; increase of revenue; promote economic development; attracting investments, etc.

In contrast to the benefits, e-government problems are related to the possible misuse due to a new way of performing the tasks of the administrative bodies, and also due to the fear of citizens and employees due to uncertainty and resistance to change. The elimination of all these difficulties seek measures of protection of information systems of e-government in order to minimize the risk of cyber crime.

Central national portal [www.euprava.gov.rs](http://www.euprava.gov.rs) has been formed in accordance to the basic principles and the principles of e-government development in the world and in the Republic of Serbia.

Figure 1: The national portal of e-government in the Republic of Serbia



This national portal represents a single point of access to all public services offered by e-government. The basic purpose and role of the national portal is that all citizens

and enterprises have quick and easy access to information and public administration services, regardless of which authority is responsible for providing the services in one place. In this way, citizens are referred to the place where they can obtain the appropriate service, and at one point getting all the information and necessary documents. Portal has the Cyrillic and Latin version, then the English version, but also the languages of national minorities (Hungarian, Ruthenia, Slovak, Croatian and Romanian). In addition, the national portal gives the possibility of registration, use of electronic certificates and personalization.

Strategy for Development of Information Society in the Republic of Serbia, plans for citizens to be able to make all contacts with administrative bodies exercising public powers, courts and health care systems, except those which by their nature require the physical presence of conduct electronically by the year 2020 [5]. The basic prerequisite for the use of e-government is the existence of infrastructure, both on service providers (state and public authorities), and on the side of the users of these services (citizens and businesses).

The development of e-government in the Republic of Serbia has made progress in recent years. This is shown by the data on the development of e-participation of all sides, showing the first fifty countries in 2016.

TABLE 2: Top 50 performers in e-participation in 2016 [6, p. 56]

Rank	Country	Rank	Country
1	United Kingdom	27	Germany
2	Japan	27	Norway
2	Australia	27	India
4	Republic of Korea	27	Sweden
5	Netherlands	32	Chile
5	New Zealand	32	United Arab Emirates
7	Spain	32	Bahrain
8	Italy	32	Ukraine
8	Singapore	34	Russian Federation
8	Canada	37	Brazil
8	Finland	37	Slovenia
12	France	39	Uruguay
12	United States of America	39	Mongolia
14	Austria	39	Ireland
14	Poland	39	Saudi Arabia
14	Mexico	43	Tunisia
17	Israel	43	Luxemburg
17	Morocco	43	Vietnam
17	Lithuania	43	Bulgaria
17	Montenegro	47	Malaysia
17	Serbia	47	Uzbekistan
22	Estonia	47	Azerbaijan
22	China	50	Portugal
22	Denmark	50	Sri Lanka
25	Malta	50	Republic of Moldova
25	Croatia	50	Mauritius
27	Colombia	50	Iceland

Research in 2016 showed that UK is a global leader in e-participation index, while Japan, Australia share second place. Singapore, the US, Morocco and Estonia are still held in high positions according to surveys from 2014 and 2016. China, Mexico, Montenegro and Serbia have competed in the top twenty countries in the first fifty, where they were in the last two years.

The Council for Public Administration Reform is established in the Republic of Serbia and it proposes measures for strategic development of e-government and proposes to the Government to take measures and activities for public administration reform. Improvement of state administration with decentralization and de-concentration requires the establishment of special bodies for the coordination of e-government in local governments whose management is the responsibility of the Working Group on Electronic Government. Professional, administrative and technical tasks for the Working group are doing the Directorate for Electronic Government. It establishes a network of contact persons responsible for the development of electronic administration and all competent authorities.

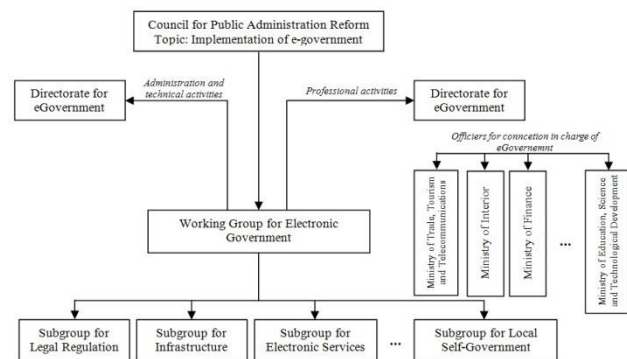


Fig. 1. Special bodies for coordination of e-government in local self-governments [7]

The main objective of e-government in the Republic of Serbia is to make lighter, cheaper and more transparent interaction between government and citizens, governments and companies as well as government agencies themselves. E-government is based on a non-paper procedure which saves time, and the whole process is done much more efficiently and requires investment in training of citizenship in order to become acquainted with the services provided by e-government, and therefore its benefits.

The European Commission has defined the assessment of the state of e-government on the basis of four indicators [8, p. 8]: an online sophistication indicator, an indicator on the number of public services fully available online, user centricity indicators, an assessment of national portals.

E-government performs many functions for individual citizens, companies and other organizations that serve. The use of Internet technologies by governments and government agencies to perform these functions is often called e-government [9, p. 224.]. The sophistication of e-government in the Republic of Serbia defines the quality of services available to citizens and legal entities through the Internet.

On-line availability represents the extent to which Internet technology replaces other forms of communication when using public administration of e-government. Navigation of service to users gives a picture of the state of e-government from perspective of: security of personal data of users; comfort of the user, possibilities of multi-channel access and accessibility of services to all

users. The national e-government portal offers all customers a unique approach to public services.

The development of e-government monitors is followed by the development of new technologies. M-government constitutes to be an important trend that represents the e-government with an emphasis on the use of mobile (portable) devices such as smart phones, tablets, etc.

In the following years for further development of e-government in the Republic of Serbia a crucial influence has development of the Internet, as well as the generally accepted trend of providing electronic services and information in local governments. By implementing e-services with the needs of the user contributes to the development and use of electronic services to different target groups in public.

Most commonly provided services to citizens are as follow: copies of birth certificates, marriage certificates, death certificates; requests for change of residence; identity documents; seeking employment; application of annual personal income tax; social protection benefits; Vehicle Registration; building permits; report to the police; health services; public libraries; enrollment in higher education. The services provided to legal entities are: registration of new companies; submission of statistical data; corporate income tax; social benefits for employees; VAT; licenses relating to the environment; customs declarations and public procurement.

Creating electronic services at the national e-government portal takes care of the needs of citizens and legal entities, as well as on establishing services that are necessary for data exchange between different administrative bodies. Establishing open administration provides transparency of administration, as well as new possibilities for the concept of open data (Open Data Model) in response to the needs of modern society and the increasing importance of information and communication technologies.

The concept of open data predicts a completely free access to data and information of electronic government. It contributes to the economic development on a large extent, transparency of state administration bodies, as well as improving many other services thanks to the large volume of data and information. Transparency, efficiency and better delivery of public services contributes to the growth of economic and social value in the Republic of Serbia.

The basic preconditions for development of e-government in the Republic of Serbia are related to the legal framework, developing economic and technological conditions, as well as education of citizens. The use of e-government services requires the existence of adequate infrastructure on the side of service providers - state bodies and on the side of consumers - individuals and legal entities. Therefore, in the coming period in the Republic of Serbia continuous implementation of the action is necessary in order to permit changes of transition to electronic mode of service and information delivery. E-government in the Republic of Serbia is an area that has a rising expansion trend, thanks to the legal regulations that affect the safety and security of users' rights.

#### IV. CONCLUSION

Electronic government provides opportunities for citizens and businesses to communicate through electronic

media by using new information and communication technologies. Communication through national communications network service establishes a connection between different information systems of public administration bodies, bodies of autonomous provinces and local governments and their web services in the Republic of Serbia. Promotion and development of e-government services, then the reform of the education system and its customization according to requirements of the modern business environment creates conditions for raising the competitiveness of our economy. Strengthening and development of e-government contributes to the achievement of significant savings and increased rationalization of operations in the Republic of Serbia as a whole, but the end result is to raise national competitiveness in global terms. The development of e-government services will increase internal efficiency and competitiveness of the domestic economy.

#### REFERENCES

- [1] S. Lovreta, J. Končar, G. Petković, *Kanali marketinga, trgovina i ostali kanali*, Beograd: Univerzitet u Beogradu Ekonomski fakultet, Univerzitet u Novom Sadu Ekonomski fakultet Subotica, 2013.
- [2] E. Turban, D. King, J. Lang, L. Lai, C. Pollard, D. C. Turban, L. Volonino, I C. Seballos II, *Introduction to Electronic Commerce 3/e*, New Jersey: Prentice Hal, 2011.
- [3] E. Turban, D. King, J. Lee, T-P. Laing, C. D. C. Turban, J. Lang, L. Lia, C. Pollard, S. Murugesan, K. Wuagh, S. Myles, C. Seballos, *Electronic Commerce 2012, A Managerial and Social Networks Perspectives*, 7<sup>th</sup> edition, New Jersey: Prentice Hall, 2012.
- [4] UN E-Government Survey 2016, available at: <https://publicadministration.un.org/egovkb/en-us/Reports/UN-E-Government-Survey-2016>, accessed December 2016.
- [5] *Strategija razvoja elektronske uprave u Republici Srbiji za period od 2015-2018. godine i akcioni plan za sprovođenje strategije za period 2015-2016. godine*, Sl. glasnik RS, br 107/2015.
- [6] *United Nations E-Government Survey 2016, E-Government in Support of Sustainable Development*, Department of Economic and Social Affairs, United Nations, New York, 2016.
- [7] *VRS (2010) Strategija razvoja informacionog društva u Republici Srbiji do 2020. godine*, Vlada Republike Srbije, Službeni glasnik RS, br. 521/2010.
- [8] P. Wauters, M. Nijskens J. Tiebout, *The User Challenge, Benchmarking The Supply Of Online Public Services*, European Commission Directorate General For Information Society And Media, Capgemini, September 2007.
- [9] G. P. Schneider: *Electronic Commerce*, Twelfth Edition, Cengage Learning, 2017.

# Implementacija E-učenja u pomorstvu

Tatjana Stanivuk, Pomorski fakultet u Splitu, Hrvatska

**Sadržaj** — E-učenje je vrlo raširen oblik edukacije, jer omogućuje korisniku relativno lagan pristup traženim informacijama. Pomorski promet i brodarstvo su jedne od najstrože reguliranih gospodarskih grana sa mnoštvom kompleksnih pravila i propisa koja uključuju navigacijska pravila, stručnost i osposobljenost. Danas se E-učenje primjenjuje u pomorstvu, ali ne u tolikoj mjeri koliko u drugim granama gospodarstva. Zašto kad isto ima ogromnu ekonomsku opravdanost.

**Gljučne reči** — analiza ekonomske opravdanosti, obrazovanje na daljinu, pomorstvo, predviđanja.

## I. UVOD

Danas je E-učenje vrlo raširen oblik edukacije. Takva vrsta učenja sastoji se od informacijske i komunikacijske tehnologije - ICT (*engl. Information and communications technology*) koje omogućuju korisniku relativno lagan pristup traženim informacijama. Danas je E-učenje druga najvažnija metoda u organizacijama i tvrtkama [1]. U Republici Hrvatskoj E-učenje je tek u razvojnoj fazi. U okolnim regijama, ono je također na niskoj tehnološkoj razini. Osnovna definicija E-učenja kaže da je to "korištenje multimedije i Interneta u svrhu poboljšanja kvalitete učenja - omogućavanjem pristupa udaljenim izvorima i uslugama te omogućavanjem suradnje i komunikacije na daljinu." (*engl. E-learning Strategy Task Force - ETF*) [2]. Kvalitetan materijal na Internetu i metoda proučavanja materijala uz primjerene instrumente za samo-provjeru znanja glavne su značajke E-učenja naspram klasične nastave (nastavnik – student/učenik). No da li se isto odnosi na sve grane gospodarstva?! Konkretno, u pomorstvu, prije uvođenja takvog sustava, potrebno je dobro odvagati njegove prednosti i nedostatke. Osnovni sustav E-učenja sastoji se od Internet aplikacije putem koje korisnici pristupaju Internet pregledniku (*engl. Internet browser*) preko svojih računala. Treba napomenuti da se takvim Internet aplikacijama koje sadrže program E-učenja može pristupiti samo ako je korisnik spojen na Internet. Takvi programi dostupni su na Internet poslužiteljima (*engl. server*) koji se obično nalaze u nekoj od institucija koje pružaju mogućnost kreiranja tečajeva za E-učenje te koje na taj način postaju mjesta za postavljanje i dijeljenje materijala za samostalno učenje, postavljanje pitanja, rješavanje online zadataka itd. Danas je "Moodle" najpopularniji sustav za E-učenje koji se koristi diljem svijeta. Obrazovna

tehnologija nije ograničena samo na visokoj tehnologiji. Međutim, elektronička obrazovna tehnologija koja se danas naziva E-učenje, postala je dio društva i obuhvaća široku lepezu informacija. Nastavak na E-učenje je M-učenje na kojem se naglašava mobilnost, ali se ne razlikuje previše od elektroničke obrazovne tehnologije.

## II. OBRAZOVANJE NA DALJINU

Odvojenost nastavnika od studenta/učenika temelj je obrazovanja na daljinu (*engl. Computer Mediated Communication - CMC*). Da bi obrazovanje na daljinu bilo učinkovito, komunikacija između nastavnika i studenta/učenika je od velikog značaja. Za učinkovitiju komunikaciju ova metodologija mora utjeloviti one metode koje promiču najbolje komunikacijske mogućnosti. CMC pruža komunikacijske mogućnosti u vidu elektroničke pošte (*e-mail*), razgovora putem Interneta (*engl. Internet Relay Chat - IRC*), interaktivne igre na Internetu (*engl. Multiple user dungeons - MUD*), tematske grupe (*engl. news group*) i računalne konferencije. Ovakav način komunikacije ima važnu ulogu u obrazovanju na daljinu jer isporuku informacija čini učinkovitijom. CMC i njegove mogućnosti pružaju nekoliko komunikacijskih mogućnosti koje čine prijenos komunikacije srodan tradicionalnom učenju [3]. Unutar mogućnosti komunikacije koje CMC model pruža, prava veza može biti uspostavljena između vještine predavanja i raznih interakcija online učenja. Rezultat prije svega ovisi o pripremljenosti nastavnika u poticanju i motiviranju slušatelja.

### A. E-učenje

Prednosti E-učenja su u omogućavanju korisnicima kvalitetnog sudjelovanja u nastavi i onda kada se pitanje udaljenosti, rasporeda i sličnih okolnosti praktički čine nemogućima. Široka dostupnost omogućuje i istovremeno sudjelovanje velikog broja korisnika. Kvalitetan materijal na Internetu i metoda proučavanja materijala uz primjerene instrumente za samo-provjeru znanja glavne su značajke E-učenja naspram klasične nastave. Modernizirana E-učionica otvorena je 24 sata dnevno, što omogućuje efikasno iskorištavanje vremena. Korisnici sami biraju kada će i na koji način pristupiti E-učenju budući da imaju stalan pristup materijalima i nastavi koju polaze. E-učenje omogućuje posebno dinamičnu interakciju između nastavnika – studenta/učenika, kao i studenata/učenika međusobno. Svaki pojedinac doprinosi nastavi pokretanjem, odnosno sudjelovanjem u raspravama koje su vezane za određenu temu. U sklopu sustava omogućena je jednostavna integracija i pristup drugim izvorima relevantnim za gradivo koje se podučava. Glavni nedostatak E-učenja je taj što zahtijeva od korisnika

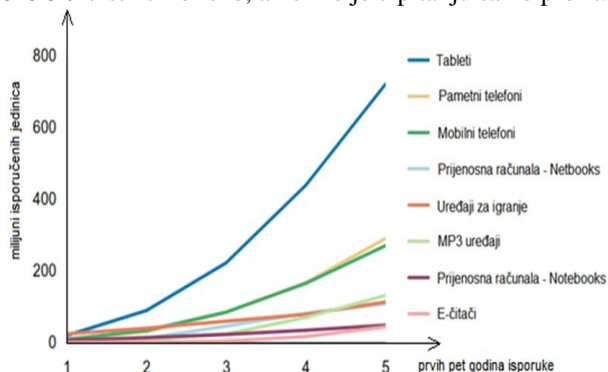
Tatjana Stanivuk, Sveučilište u Splitu - Pomorski fakultet, Rudera Boškovića 37, 21000 Split, Hrvatska (telefon: 385-21-619476; e-mail: tstanivu@pfst.hr).



određena znanja i vještine kako bi se mogli njime koristiti. Bez određene računalne pismenosti, gradivo integrirano u sklopu elektroničkog sustava učenja postaje potpuno beskorisno. Osim tih znanja, za provođenje E-nastave bitno je i da svaki od korisnika ima za to određenu opremu. Ni najkvalitetnija oprema na kojoj se izvodi E-nastava nije u potpunosti pouzdana. Čak ni kada mogući tehnički problemi ne dovedu do prekida u izvođenju E-nastave, svakako će doprinijeti padu koncentracije korisnika, a samim time i padu kvalitete E-učenja. Omogućavanjem samostalnijeg određivanja načina i vremena učenja, E-učenje svojim korisnicima donosi i veću odgovornost. U određenim oblicima E-učenja oni se tako sami moraju motivirati, individualno procjenjivati potrebu za učenjem, što može dovesti do upitnih rezultata i objektivno slabog napretka u procesu učenja. U E-učenju postoji pet ključnih čimbenika, a to su: konzalting ili savjetovanje, sadržaj, tehnologija, usluge i podrška [4].

### B. M-učenje

M-učenje ili mobilno učenje predstavlja vrstu učenja koje se odvija dok korisnik nije na fiksnoj, predodređenoj lokaciji, ili učenje koje se odvija kada korisnik koristi prednosti učenja koje pružaju mobilne tehnologije. M-učenje se fokusira na mobilnost korisnik, na interakciju s prijenosnim tehnologijama, te kao takav postaje važan dio neformalnog učenja. Ključni elementi mobilnih usluga koje su važni za M-učenje: mogućnosti učenja na različitim lokacijama, učenje u osobnom okruženju, pristup bazama podataka i interaktivna povezanost između nastavnika i korisnika, te sposobnost analiziranja informacija [5]. Prednost M-učenja sastoji se u tome da je virtualno dostupan skoro svugdje. Dijeljenje informacija je gotovo trenutno među korisnicima. M-učenje donosi snažnu prenosivost, zamjenjivanjem knjige ili bilježnice s malim uređajem, ispunjenim sa prilagođenim sadržajem učenja. Međutim, da bi M-učenje postalo konkurentno, ono mora ispuniti tri ključna uvjeta: prihvaćanje od strane korisnika, postizanje dobrih rezultata učenja i prihvatljive troškove [5]. Za sada je još uvijek upitno koliko M-učenje može biti uistinu korisno, a koliko je u pitanju samo profit.

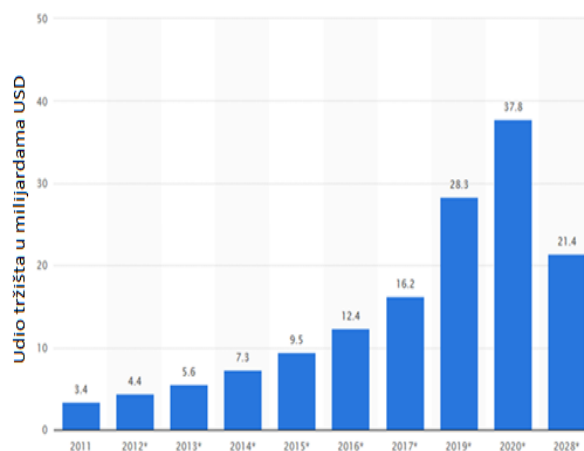


Sl. 1. Ukupna isporuka mobilnih uređaja od 2011. do 2020. godine [6].

Prema istraživanjima, (Sl. 1.) pločasto računalo (*engl. tablet computer*) je uređaj koji se najčešće koristi u kategoriji mobilnih uređaja. Pločasto računalo ili tablet posjeduje većinu funkcionalnosti i mobilnosti posuđenu od pametnih telefona te ima veći zaslon i procesorsku snagu

da zamijeni prijenosno računalo (*engl. notebook*) u mnogim situacijama.

Mogućnost dobivanja novih generacija putem M-učenja igra veliku ulogu u spremnosti tvrtki i institucija da pruži M-učenju priliku u bliskoj budućnosti (Sl. 2).



Sl. 2. Statistička analiza svjetskog tržišta predviđanja potrošnje na M-obrazovanje od 2011. do 2028. godine [7].

### III. EKONOMSKA OPRAVDANOST I PREDVIĐANJA IMPLEMENTACIJE SUSTAVA E-UČENJA NA GLOBALNOM SVJETSKOM TRŽIŠTU

Industrija E-učenja je u svom velikom poletu. Organizacije usvajaju ili nastavljaju usvajati E-učenje, trenutno po stopi od 13% godišnje. Sve veći broj organizacija i tvrtki okreće se E-učenju, jer omogućuje dobitke u raznim sektorima. Dovoljni je pokazatelj 2011. godina kad je 77% američkih korporacija koristilo online učenje, dok je 1995. istu koristilo samo 4% [1]. To znači da se broj poduzeća koja koriste online metode treninga dramatično povećao, ali uglavnom iz razloga što su se tehnološke barijere smanjile i što su se prioriteta kupaca promijenili prema samostalnom učenju.

Analize ukazuju da je zarada od E-učenja do 2017. dosegla čak 12,2 milijardi USD [8]. Globalno tržište je od samostalnog učenja doseglo 32,1 milijardu USD prihoda [8]. Stopa rasta pokazuje kako svaka država prihvaća E-učenje, jer omogućuje mogućnost zarade. Europa i SAD čine 70% globalne industrije E-učenja, međutim najbrže rastuće tržište je Azija [1]. Prema posljednjim istraživanjima najviša stopa rasta E-učenja je u Aziji 17,3%, a zatim u Istočnoj Europi 16,9%, Africi 15,2% i Latinskoj Americi 14,6% [9], [10]. Stopa rasta samostalnog E-učenja po državi: Indija 55%, Kina 52%, Malezija 41%, Rumunjska 38%, Poljska 28%, Češka 27%, Brazil 26%, Indonezija 25%, Kolumbija 20% i Ukrajina 20% [8]. Na Azijskom tržištu, zarada od E-učenje je već dosegla brojku od 11,5 milijardi USD. [8]. Ujedno se predviđa da će se do 2019. godine 50% studentske nastave izvoditi preko Interneta [1]. Glavni korisnici servisa i usluga E-učenja su uglavnom velike tvrtke i čine 30% kupaca usluga E-učenja [8]. Danas, za korištenje E-učenja mnoga sveučilišta traže načine korištenja kroz neprofitne modele koristeći standardni način podučavanja sa primjenom okruženja E-učenja. Također MOOC (*engl.*

*Massive Open Online Course*) tržište je u punom zamahu i konstantno se proširuju usluge i alati. Tržište Sustava za upravljanje učenjem LMS (*engl. Learning Management System Market*) je vrijedilo 2,55 milijarde USD u 2013. sa procijenjenom složenom godišnjom stopom rasta od oko 25,2% [11]. Drugim riječima, očekuje se da će LMS tržište vrijediti više od 7 milijardi USD u 2018., a najveći prihodi očekuju se u Sjevernoj Americi [8].

Svjetsko tržište za proizvode i usluge mobilnog učenja doseglo je 5,3 milijardi USD u 2012. uz godišnju stopu rasta od 18,2% za sljedećih pet godina [8]. Procjenjuje se da će u svijetu, tržište mobilnog učenja, dosegnuti čak 12,2 milijardi USD s početkom 2017. Dok su u 2012. najveći kupci mobilnih proizvoda i usluga za učenje bili: SAD, Japan, Južna Koreja, Kina i Indija, očekuje se da će zaključno s 2017. godinom najveći kupci mobilnih proizvoda i usluga za učenje biti: Kina, SAD, Indonezija, Indija i Brazil [8].

Pozitivan učinak za E-učenje je velik trend prodaje pametnih telefona i sve veći pristup Internetu na globalnoj razini. Prijenosni ili mobilni uređaji poboljšavaju produktivnost rada i proširuju koncept mobilnog učenja u smislu trenda "Ponesi Svoj Vlastiti Uređaj - PSVU" (*engl. Bring Your Own Device - BYOD*) kojeg je u današnje vrijeme jednostavno nemoguće izbjeći. Industrija korporativnog učenja vrijedi gotovo 200 milijardi USD, a E-učenje predstavlja 56,2 milijardi USD od tog dijela [1].

Danas se E-učenje temelji na tri čimbenika: sadržaj, distribucija sadržaja i platforme za E-učenje [12]. Potražnja za programima i uslugama izgleda ovako [8]: 44% tvrtki je namjeravalo kupiti on-line alate i sustave obrazovanja, 41% tvrtki je namjeravalo kupiti sustave za upravljanje učenjem (LMS), 37% tvrtki je namjeravalo kupiti autorske alate / sustave, 33% tvrtki je namjeravalo kupiti alate za učionice i sustave, 29% tvrtki je namjeravalo kupiti sadržaje za razvoj proizvoda i usluga, 27% tvrtki je namjeravalo kupiti Courseware i prezentacijske alate i pripadajući softver, 18% tvrtki je namjeravalo kupiti audio i web konferencijske proizvode i sustave. Trenutno 8% tvrtki koristi MOOC dok drugih 7% misli eksperimentirati sa MOOC. Predviđa se da će se ta brojka povećati do 28% do kraja 2017. godine [8]. Više od 350 tvrtki surađuju s Coursera i Udacity<sup>1</sup> s ciljem izdvajanja najboljih studenata koji bi time postali najbolji kandidati za odgovarajuće poslove [13], [8].

Očekuje se da će online korporativno tržište rasti po stopi od 13% godišnje. Danas, 77% američkih tvrtki nudi online korporativni trening kako bi poboljšale profesionalni razvoj svojih zaposlenika [8]. Može se reći da je korporativno E-učenje jedan od najbrže rastućih tržišta, a vjerojatno i jedno od najperspektivnijih tržišta u obrazovnoj industriji. Navedene statistike jasno pokazuju da E-učenje očekuje daljnji porast te ima i ogroman potencijal za dobit.

<sup>1</sup> Coursera i Udacity - tvrtke koje nude Masivne otvorene online tečajevе (MOOCs). Rade s vrhunskim sveučilištima i organizacijama kako bi neki od njihovih tečajeva bili dostupni na mreži, i nude tečajevе u mnogim granama znanosti.

#### IV. DOSADAŠNJI UTJECAJ E-UČENJA U POMORSTVU I POTREBE DALJNJEG ISTRAŽIVANJA

Danas u pomorskom gospodarstvu postoje multimedijalni tečajevi razvijeni od pojedinih tvrtki (Seagull Maritime AS, Videotel, itd.) s različitim primjenom u pomorstvu. Takvi tečajevi omogućuju učenje na daljinu na brodu ili na kopnu i dostupni su kao fizički medij ili im se može pristupiti putem Interneta. Pružatelji takvih tečajeva nude odgovarajuće module, izrađene za jasne ciljeve učenja, koji bi trebali omogućiti [13]: 1) isplativost i uštedu vremena i novca, 2) poticanje procesa učenja putem videa, teksta, grafike i animacije, 3) angažiranje u sudjelovanju igranja uloga, rješavanje problema, studije slučaja, 4) podršku u pomoći, savjetima potrebnim u korištenju tečajeva, ...

U većini slučajeva se obrazovanje pomoraca odvija kroz dva odvojena sustava. U jednom sustavu obrazovanje se vrši unutar srednjoškolskog i visokog obrazovanja, a u drugom sustavu preko specifičnih programa izobrazbe koji se provode u skladu s odredbama STCW (*engl. The International Convention on Standards of Training, Certification and Watchkeeping for Seafarers*) konvencije. Kako bi se omogućila veća fleksibilnost u obrazovanju radne snage u pomorstvu, Međunarodna pomorska organizacija - IMO (*engl. International Maritime Organization*) je pomoću STCW 78 konvencije uvela standarde obrazovanja za pomorce. Tim korakom su uvedeni minimalni sadržaji kompetentnosti koje pomorac mora ispunjavati, ali su i riješeni problemi priznavanja kvalifikacija i usklađivanja sadržaja različitih nacionalnih obrazovnih programa. IMO je također preko cijelog niza preporučenih programa ponudio obrazovnim institucijama standardizirane vodiče u cilju pojednostavljenja implementacije STCW odredbi. Većina tih sadržaja odnosi se na specifične programe izobrazbe (tečajevе) kojima se moraju steći odgovarajuća znanja i vještine. STCW konvencija je značajno izmijenjena 1995. godine, a druga izmjena dogodila se u Manilli 2010. u kojoj je učenje na daljinu i E-učenje (*engl. Guidance for use of distance learning and e-learning*) prepoznato kao jedan od načina obrazovanja pomoraca, ali isto nije u potpunosti definirano. Pravilna primjena modela izobrazbe je opisana putem definiranih izvedbenih programa i nastavnih sati, međutim nije definirano koliki se dio sadržaja i sati može obraditi na daljinu, a koliki klasičnim oblicima nastave. Na žalost, Hrvatsko Ministarstvo Pomorstva Prometa i Infrastrukture E-učenje nije prepoznalo te postoji problem koji će se u budućnosti morati riješiti [14].

U pomorskom obrazovanju, učenje na daljinu veoma je složen proces. Učenje na brodu zahtjeva veliku koncentraciju i rezervaciju određenog vremena za učenje tj. izbitak iz svakodnevnog redovnog posla na moru. Valja naglasiti da takvo učenje zahtjeva precizno planiranje usklađenosti slobodnog vremena između nastavnika i studenta, ne samo zbog zahtjevnog posla na brodu i nemogućnosti izbivanja iz istoga, već i zbog različitih vremenskih zona, ne mogućnosti dobre Internet veze (ista nije moguća na svim brodovima, neki imaju pristup samo u određenim uvjetima ako su u luci ili vrlo blizu luke, neki mogu komunicirati samo putem elektroničke pošte preko

satelitskog Interneta,...). Ujedno, provedene studije pokazuju da je čovjeku potrebno minimalno 7 do 9 sati sna da bi normalno funkcionirao i uspješno obavljao svakodnevni posao [15]. Kod redovnog obavljanja svoga posla na brodu, pomorci su izloženi velikom stresu i umoru te im je san i odmor od iznimne važnosti. Gdje je tu onda vrijeme potrebno za obrazovanje? Kako isto uskladiti s obzirom na već navedenu problematiku i potrebu konzultacija između nastavnika i studenta na brodu? Opet, konvencionalni sustav učenja i stjecanje određenog znanja zahtjeva vrijeme provedeno u organiziranim nastavnim aktivnostima, tj. na predavanjima, vježbama i drugim oblicima nastave koja se odvijaju u obrazovnim ustanovama na kopnu, a ne na brodu. Međutim, takav standardni način obrazovanja ima nedostataka u optimizaciji rasporeda učenja, jer sama priroda posla pomorcima ne dopušta redovito prisustvovanje nastavi, a oduzima im i vrijeme koje je predviđeno za obitelj i sl.

Pomorski promet i brodarstvo jedne su od najstrože reguliranih gospodarskih grana, sa mnoštvom kompleksnih pravila, smjernica i propisa na regionalnoj i globalnoj razini. Ispitivanja su pokazala da E-učenje može doprinijeti sustavu obrazovanja, međutim granica ili potvrda tomu još uvijek nije dokazana u praksi. [12] Naime, javlja se cijeli niz pitanja na koje treba pronaći odgovor: 1) U kolikom obujmu ili postotku E-učenje može pomoći u samo-učenju u pomorstvu?, 2) Gdje je, u pomorstvu, granica između materijala za samo-učenje dostupnih s Interneta i klasičnog obrazovanja?, 3) Da li pomorac u svom radnom vremenu na brodu može akceptirati ponuđeni materijal kroz E-učenje?, 4) Koliko efektivno pomorac može odvojiti vremena za samo-učenje dok je na brodu, odnosno da li mu je ukupno vrijeme dovoljno za savladavanje odgovarajućih zahtjevnih sadržaja?, 5) Postoji li način da se kontaktira predmetni nastavnik putem informatičkih tehnologija ukoliko pomorcima nisu jasne određene informacije u svezi nastavnog predmeta?...

## V. ZAKLJUČAK

Nedvojbeno je da danas svijet ulazi u doba mobilnosti, a nove tehnologije čine računalo uistinu sveprisutnim, pa time postaju i sastavnim dijelom naših života. Iz takvih tehnologija razvilo se E-učenje kao jedan od oblika obrazovanja. Međutim, takva vrsta učenja ima i svojih negativnih strana jer korisnik mora snositi odgovornost samostalne organizacije učenja, stvaranje samo-motivacije te individualno procjenjivati potrebu za učenjem što može dovesti do upitnih rezultata i objektivno slabog napretka u procesu učenja. Obrazovanje časničkog kadra u pomorstvu propisano je IMO brojem nastavnih sati, te je samim time upitno koliko je određeni dio satnice moguće obraditi na daljinu, a koliko u učionici. Današnja pomorska industrija ima mnogo pravilnika, konvencija i smjernica koje postavljaju granice sigurnosti i učinkovitosti u brodarstvu. Razvojem novih tehnologija u pomorskoj industriji, sustav obrazovanja u pomorstvu znatno se promijenio u zadnjih desetak godina. Naime, nakon drugog svjetskog rata sustav

obrazovanja pomorskog kadra razvijao se proporcionalno zahtjevima tehnološkog napretka industrije. Međutim, brod je ostao relativno opasno mjesto za rad, a život i rad pomorca na brodu i dalje je vrlo stresan i težak. Sa današnjom tehnologijom otvara se mogućnost izobrazbe primjenom E-učenja unutar radnog okruženja na brodovima, a STCW konvencija je prema zadnjim izmjenama i dopunama prepoznala prednosti istog, ali ga nije u potpunosti definirala što ostavlja prostor za buduća istraživanja.

## LITERATURA

- [1] <http://info.shiftelearning.com/blog/bid/247473/18-Mind-Blowing-eLearning-Statistics-You-Need-To-Know>. [pristup: prosinac, 2016]
- [2] <http://marjan.fesb.hr/~ljljana/radovi/E-learning.ppt>. [pristup: prosinac, 2016.]
- [3] E. C. O'Dwyer, "Approaches to distance learning - An evaluation of current methodologies, technologies and operational costs as an alternative means of course delivery for developing country academies," Ph.D. dissertation, World Maritime University, 2001.
- [4] A. Nagy, The Impact of E-Learning, in: Bruck, P.A.; Buchholz, A.; Karssen, Z.; Zeffass, A. (Eds). *E-Content: Technologies and Perspectives for the European Market*. Berlin: Springer-Verlag, 2005, pp.79–96.
- [5] B. Marc, P. Rasmus, "Acceptance, Chances, and Problems of Mobile Learning in Vocational Education in Enterprises," *CEUR Workshop Proceedings*, vol. 955, pp.190-196, 2012.
- [6] <http://blog.commlabindia.com/elearning-development/elearning-on-tablets>. [pristup: prosinac, 2016.]
- [7] <http://www.statista.com/statistics/273960/global-mobile-education-market-volume/>. [pristup: prosinac, 2016.]
- [8] <https://www.linkedin.com/pulse/top-elearning-statistics-facts-2015-you-need-know-colin-thompson>. [pristup: prosinac, 2016.]
- [9] <http://elearningindustry.com/new-report-on-e-learning-market-trends-and-forecast-2014-2016-just-released>. [pristup: prosinac, 2016.]
- [10] H. Crompton, "A historical overview of mobile learning - Toward learner - centered education," Z. L. Berge & L. Y. Muilenburg (Eds.), *Handbook of mobile learning*, pp. 3–14, 2013.
- [11] Brandon Hall Group Research Team, *LMS Trends 2015, Is It Time for Something Different*, Brandon Hall Group LMS Study, 2015.
- [12] <http://www.maritimeprofessional.com/blogs/post/does-elearning-work-in-the-maritime-industry-13878>. [pristup: siječanj, 2017.]
- [13] [http://videotel.com/Maritime\\_Training\\_Solutions/training\\_delivery\\_solutions/Maritime\\_Training\\_Courses/](http://videotel.com/Maritime_Training_Solutions/training_delivery_solutions/Maritime_Training_Courses/). [pristup: siječanj, 2017.]
- [14] *Narodne novine*. Pravilniku o zvanjima i svjedodžbama o osposobljenosti pomoraca. No. 130/13, 45/14 i 124/15. Available: <http://www.propisi.hr/print.php?id=6131>. [pristup: siječanj, 2017.]
- [15] <https://sleepfoundation.org/how-sleep-works/how-much-sleep-do-we-really-need>. [pristup: siječanj, 2017.]

## ABSTRACT

E-learning is widely spread form of education, since it enables the user relatively easy access to the requested information. Maritime transport and shipping trade are among the most regulated industries with great deal of complex rules and regulations, including rules of navigation, navigational expertise and qualifications. Nowadays E-learning is applied in maritime affairs, although not as much as in other industries. Why is that so? Since E-learning provides enormous economic justification.

## IMPLEMENTATION OF E-LEARNING IN MARITIME AFFAIRS

Tatjana Stanivuk

# Unapređenje ICT infrastrukture obrazovnog sistema Crne Gore primjenom Cloud Computing-a

Marina Matijević, *Ministarstvo prosvjete,*  
Prof. dr Ramo Šendelj, *Univerzitet Donja Gorica*

**Sadržaj** — U ovom radu predstavljene su osnovne prednosti Cloud Computing-a kao i njegova primjena u obrazovnom sistemu. Posebna pažnja je posvećena analizi Informacionog sistema obrazovanja Crne Gore, a naročito modulima koji se odnose na MEIS aplikaciju i primjenu Office365 u obrazovno-vaspitnim ustanovama.

**Gljučne riječi** — Bezbjednost, Cloud, Computing, informacioni sistem, infrastruktura, MEIS, obrazovni sistem, Office365, platforma, softver.

## I. UVOD

Cloud Computing se u današnje vrijeme sve više primjenjuje u mnogim sferama a naročito u poslovnim okruženjima, jer se uz pomoć cloud-a mogu postići velike uštede na IT troškovima. Njegova naročito velika primjena je kod malih i srednjih preduzeća koja su zahvaljujući, prednostima cloud-a, mogli priuštiti upotrebu značajnih resursa za svoja poslovanja bez kupovine velikog dijela informatičke opreme i ne zapošljavajući osobe koje bi održavale svu opremu i sistem.

Pored primjene u poslovnom okruženju, Cloud Computing je od posebnog značaja i za velike sisteme, posebno obrazovni pa je, samim tim, našao primjenu i u obrazovnom sistemu Crne Gore.

## II. ANALIZA POSTOJEĆE ICT INFRASTRUKTURE OBRAZOVNOG SISTEMA CRNE GORE

Obrazovni sistem Crne Gore na svim nivoima do nivoa fakulteta je jedan heterogeni sistem čija se heterogenost ogleda, kako u organizacionoj, tako i u geografskoj i funkcionalnoj složenosti. Pored Ministarstva prosvjete u sistemu obrazovanja postoji još veliki broj ustanova: Zavod za školstvo, Centar za stručno obrazovanje, Ispitni centar, Zavod za udžbenike i nastavna sredstva i mreža predškolskih i školskih ustanova, od kojih je 21 državnih i 20 privatnih predškolskih ustanova, 163 državnih i 4 privatne osnovne škole, 50 državnih i 3 privatne srednje, 15 osnovnih muzičkih škola i 3 resursna centra (Sl.1.). U

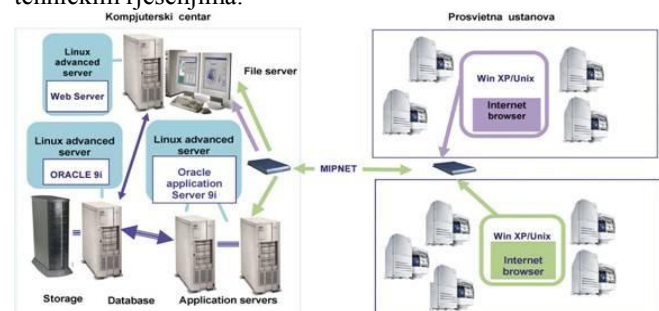
navedenim ustanovama je oko 120.000 djece i oko 13.000 zaposlenih što nam govori da je svaka od navedenih cjelina veoma složen sistem. To često vodi ka traženju parcijalnih rješenja i neusaglašenosti u radu pojedinih podsistema. Drugim riječima, obrazovni sistem Crne Gore spada u grupu heterogenih distributivnih poslovnih sistema.



Sl. 1. Šema obrazovnog sistema Crne Gore, međuinstitucionalna komunikacija i razmjena podataka

Pored velikog broja ustanova, zaposlenih i djece, karakteristika obrazovnog sistema je da se mnoge aktivnosti istovremeno izvode, da je velika geografska udaljenost, kako između ustanova, tako i između mjesta stanovanja učenika i ustanove u kojoj učenik boravi, različiti profili za koje se učenici školuju itd. Zbog svega navedenog neophodno je bilo uspostavljanje Informacionog sistema obrazovanja Crne Gore (MEIS - Montenegrin Education Information System).

Informacioni sistem obrazovanja Crne Gore implementiran je u svim obrazovno-vaspitnim ustanovama, privatnim i državnim, do nivoa visokog obrazovanja (Sl.2.) [9]. S obzirom da je sistem heterogen, njegova je izgradnja morala sadržati razne module koji su se bazirali na raznim tehničkim rješenjima.



Sl. 2. Arhitektura MEIS-a

Marina Matijević, ul. Kralja Nikole 321, 81000 Podgorica, Crna Gora (telefon: 382-67-977-997; email: [matijevic.marina@gmail.com](mailto:matijevic.marina@gmail.com))

Ramo Šendelj, ul. Dalmatinska 78, 81000 Podgorica, Crna Gora (telefon: 382-67-226-333; email: [ramo.sendelj@gmail.com](mailto:ramo.sendelj@gmail.com))

Jedan od djelova Informacionog sistema obrazovanja Crne Gore je MEIS aplikacija. Putem MEIS aplikacije

vodi se mnoštvo podataka koji idu za svakog pojedinačnog zaposlenog i učenika, prate se učenici kroz cijelo obrazovanje do fakulteta. Svaki zaposleni u školi obučava se za korišćenje MEIS aplikacije. U zavisnosti od uloge koju ima, iste obaveze usmjerava i prema MEIS aplikaciji. Dakle, uprava škole kreira odjeljenja, dijeli nastavu i unosi zaposlene u sistem. Razredne starješine vode računa o svom odjeljenju, upisuju učenike u dnevnik, evidentiraju izostanke, disciplinu itd. Svaki nastavnik unosi ocjene za svoj predmet učenicima kojima predaje. Istovremeno, razredne starješine i uprava škole imaju veliki broj kreiranih izvještaja direktno iz sistema. Sistem je nadograđen portalom za roditelje putem kojeg svaki roditelj sa korisničkim imenom i lozinkom ima pristup uspjehu, izostancima, vladanju i disciplini svog djeteta, kao i još jedan dodatni način elektronske komunikacije sa odjeljenjskim starješinom.

U pitanju je web aplikacija koja je dostupna korisnicima sa bilo kojeg uređaja iz škole ili sa neke druge lokacije (Sl.3.). Obrazovno-vaspitne ustanove ne koriste svoje resurse za vođenje evidencije i praćenje obrazovnih procesa a, samim tim, nemaju potrebu da obezbijede odgovarajuće održavanje svojih sistema. Drugim riječima, za sve navedene obrazovno-vaspitne ustanove do nivoa fakulteta, evidencija se vodi u data centru Ministarstva prosvjete koje je zaduženo za razvoj i unaprijeđenje cijelog sistema kao i za njegovu implementaciju i održavanje. Ovakvo rješenje je, sa druge strane, bilo neophodno jer se radi o registrima koji posjeduju lične podatke pa, samim tim, serveri na kojima se nalaze podaci moraju biti fizički smješteni u Crnoj Gori.



Sl.3. Početna stranica MEIS aplikacije

Međutim, pored evidencije podataka i praćena procesa o obrazovnom sistemu, postoje i druge potrebe kao, na primjer, razmjena iskustava nastavnika. Za te potrebe Ministarstvo prosvjete je kreiralo portal za nastavnike, [www.skolskiportal.edu.me](http://www.skolskiportal.edu.me). Portal je osmišljen kako bi unaprijedio korišćenje ICT-a i približio informatički svijet i njegove mogućnosti nastavnom kadru i drugim subjektima u obrazovanju. To svakako podiže i njihov nivo informatičkog znanja.

Portal omogućava nastavnicima objavljivanje radova i

svojih predavanja. Takođe, njegova svrha je omogućavanje razmjene ideja, znanja i iskustava među nastavnicima, podsticanje nastavnog kadra na rad s Internetom, upoznavanje sa novim tehnologijama i sticanje novih znanja i iskustava, kao i podsticanje novih oblika stvaralaštva nastavnika u raznim područjima obrazovanja. Na školskom portalu je omogućena laka pretraga i preuzimanje (download) didaktičkog softvera, kao i lista korisnih linkova.

Pored razmjene iskustava, jedna od neophodnih potreba je elektronska komunikacija. Kako je u obrazovnom sistemu Crne Gore oko 13.000 zaposlenih, jasno je da za obezbjeđivanje mail naloga za sve zaposlene potrebno mnogo resursa. Pored toga, potreba je da svaka ustanova posjeduje zvaničnu mail adresu kao i da određena tijela ili grupe unutar ustanove posjeduju svoje naloge (npr. uredništvo školskog časopisa).

Ograničenje u resursima se javlja i kod razvoja novih servisa koji se mogu ponuditi kao nadogradnja postojećeg sistema, kao na primjer davanje prikaza lokacija svih objekata obrazovno-vaspitnih ustanova na geografskoj mapi sa davanjem i osnovnih informacija o samoj ustanovi (trenutni broj učenika, zaposlenih, kontakt podaci itd). U oblasti obrazovanja postoji značajan broj servisa koji se mogu razviti ali pored hardverskih ograničenja glavni problem predstavlja kadar koji je neophodno zaposliti da bi se cijeli sistem održao.

Zbog velike količine resursa neophodnih za obezbjeđivanje mail naloga ili novih servisa, te finansijskih i kadrovskih ograničenja, primijenilo se drugačije rješenje od onih koji su se do tada primjenjivana na uspostavljanu MEIS aplikacije.

### III. CLOUD COMPUTING

Kada govorimo o Cloud Computing-u, podrazumijeva se da se podaci skladište na Internetu, a ne lokalno na računaru. Međutim, nije riječ samo o skladištenju podataka već o činjenici da se kompletni sistemi nalaze u Cloud-u.

Postoji nekoliko vrsta Cloud-a [1-6] i to:

- privatni
- javni
- hibridni

Privatni cloud je isključivo namijenjen jednom klijentu i može da bude smješten unutar neke organizacije ili kod određenog cloud provajdera.

Javni cloud je namijenjen, kako organizacijama, tako i pojedincima i u vlasništvu je cloud provajdera.

Hibridni cloud povezuje privatni i javni, odnosno predstavlja proširenje privatnog sa resursima javnog clouda.

Ako posmatramo drugačiju podjelu, primijetićemo da postoje tri vrste servisa odnosno usluga zasnovanim na Cloud-u [7] i to:

- infrastrukture
- platforme
- softveri

Infrastruktura kao usluga (Infrastructure as a Service IaaS) podrazumijeva da korisnik može koristiti

informatičku infrastrukturu (servere, mrežnu opremu, skladištenje podataka, operativni sistem itd). Uglavnom se radi o virtuelnim platformama.

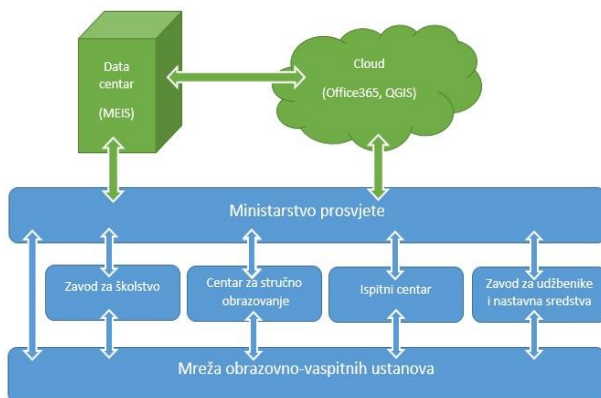
Platforma kao usluga (Platform as a Service – PaaS) odnosi se na razvojnu okolinu. Korisnik razvija sopstvene aplikacije koje pokreće na infrastrukturi cloud provajdera.

Softver kao usluga (Software as a Service – SaaS) pruža korisniku mogućnost upotrebe aplikacija koje su dostupne na infrastrukturi cloud-a. Korisnik pristupa aplikacijama putem interneta sa različitih uređaja.

Sva tri modela različito utiču na bezbjednost u cloud-u i svaka kombinacija cloud servisa nosi različite rizike i zahtjeva različite mjere zaštite.

U cloud-u se bezbjednost može podijeliti u dvije kategorije i to: bezbjednost cloud provajdera i bezbjednost korisnika. Kada govorimo o bezbjednosti cloud provajdera, tu se prvenstveno misli da provajder mora obezbijediti da njegov cloud bude bezbjedan kao i podaci i aplikacije korisnika. Isto tako, korisnik mora biti siguran da cloud provajder vodi računa da na adekvatan način zaštiti njegove podatke [8]. Pored navedenog potrebno je posebnu pažnju obratiti i na pravni mehanizam kojim bi se vodili ukoliko dođe do kompromitovanja podataka koji se nalaze u cloud-u.

#### IV. UNAPREĐENJE ICT INFRASTRUKTURE OBRAZOVNOG SISTEMA CRNE GORE PRIMJENOM CLOUD REŠENJA



Sl.4. Infrastruktura informacionog sistema obrazovanja Crne Gore

U Ministarstvu prosvjete nalazi se DNS (Domain Name System) server edu.me. Uz pomoć DNS servera i Microsoftovih resursa formirana je posebna zona u DNS-u za svaku obrazovno-vaspitu ustanovu u Crnoj Gori do nivoa fakulteta (Sl.4.). Na taj način mail adrese su prepoznatljive za sve obrazovno-vaspitne ustanove i završavaju se sa edu.me. Npr. za osnovnu školu “Kekec”, mail adresa škole [škola@os-kekec.edu.me](mailto:škola@os-kekec.edu.me). Isti oblik mail adrese ima svaki zaposleni škole ([ime.prezime@os-kekec.edu.me](mailto:ime.prezime@os-kekec.edu.me)). Kroz ove mail naloge svakom zaposlenom je omogućeno korišćenje Office365 alata kao što su: Mail, Kalendar, Kontakti, Skype, Online Word, Online Excel, Online PowerPoint, Online OneNote, Yammer, One Drive, Delve, Video, Sway, Sites itd (Sl.5.) [10].



Sl. 5. Aplikacije koje su dostupne zaposlenima i učenicima kroz Office365 sa bilo kojeg uređaja konektovanog na internetu

Pored zaposlenih, svi učenici mogu koristiti Office365, odnosno isti oblik naloga kao i nastavnici.

Drugim riječima, uvođenjem Office365 u obrazovni sistem stvorili su se neophodni uslovi da nastavnici mogu komunicirati putem elektronske pošte ili da koriste prednosti video i audio komunikacije putem Skype-a. Pored toga, mogu da razmjenjuju dokumenta ili rade na istim dokumentima korišćenjem OneDrive-a i Online Office-a (Word, Excel, PowerPoint, OneNote). Isto tako, u situaciji su da dijele kalendare, koriste planere za realizaciju zajedničkih aktivnosti itd.

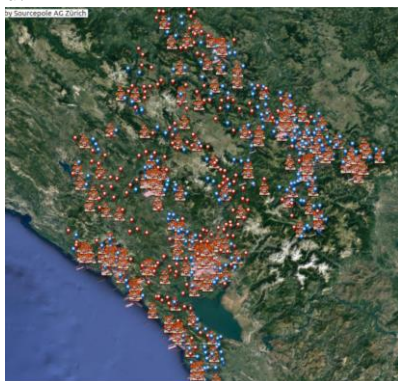
Na isti način mogu da komuniciraju sa učenicima ili dijele nastavne materijale sa njima. Modernizacija nastavnog procesa je neophodna kako bi se podigao kvalitet obrazovanja. Uvođenje informaciono-komunikacionih tehnologija u nastavni proces i proces učenja koje se vrši u skladu s prethodno određenim ciljevima i standardima, u značajnoj mjeri može doprinijeti kvalitetu nastavnog procesa, efikasnosti učenja i boljoj didaktičkoj organizaciji pojedinih nastavnih predmeta.

Svakoj obrazovnoj ustanovi obezbijeđene su iste mogućnosti. Ipak, primjena se od škole do škole u značajnoj mjeri razlikuje. Sve škole koriste naloge za ustanovu kao zvaničnu mail adresu. U značajnom broju škola veliki procenat zaposlenih koristi mail naloge pod domenom škole i dosta koriste sve prednosti Office365. Trenutno, u manjem broju škola učenici koriste ove naloge. Međutim, cilj Ministarstva prosvjete je da se Office365 što više koristi u obrazovne svrhe jer je činjenica da značajno može doprinijeti intenzivnijoj saradnji učenika i nastavnika, kao i nastavnika sa ostalim kolegama iz svoje ili drugih ustanova.

U nekim školama kreirani su školski kalendari u kojem su označene sve planirane aktivnosti u školi (dan škole, sportski dan, eksterno testiranje učenika, izleti itd) i kalendar je podijeljen sa svim zaposlenima u ustanovi.

Ako se razmotra na koji način je najefikasnije unaprijediti postojeći sistem, na primjer, iskoristiti postojeće podatke o objektima obrazovno-vaspitnih ustanova i drugih podataka vezanih za te ustanove (broj zaposlenih, broj učenika itd) mogao bi se iskoristiti QGIS cloud rješenje (za vektorsku obradu i publikovanje prostornih podataka). Na taj način bi se dio podataka iz MEIS baze podataka (koja je na infrastrukturi Ministarstva prosvjete) iskoristio za servis koji bi bio u cloud-u i svaka izmjena podatka u bazi automatski bi se izmjenila i u cloud-u (Sl.6.). Ovo je samo jedan od primjera na koji način se upotrebom cloud sistema može

nadograditi postojeći informacijski sistem obrazovanja, a da to ne zahtjeva ulaganje u dodatnu hardversku infrastrukturu.



Sl. 6. QGIS

Ako posmatramo strukturu Informacionog sistema obrazovanja Crne Gore iz ugla bezbjednosti, možemo reći da je granica oko toga u kojem dijelu se koriste resursi Ministarstva prosvjete, a u kojem dijelu Microsoft cloud dobro postavljena. Registri sa ličnim podacima i sva evidencija o školskim resursima nalazi se u Data centru Ministarstva prosvjete i o njima brinu samo ovlašćena lica, dok se Microsoft cloud koristi za komunikaciju i online kolaboraciju.

#### V. ZAKLJUČAK

Informacioni sistem obrazovanja Crne Gore je homogeni sistem. Pored resursa koji se koriste za određene djelove informacionog sistema (MEIS aplikacija, portal za roditelje, portal za nastavnike itd) koristi se i Microsoft Office365 cloud rješenje koje je u značajnoj mjeri smanjilo troškove za nabavku velike količine neophodne opreme, kao i troškove za održavanje sistema. Drugim riječima, informacijski sistem obrazovanja Crne Gore predstavlja hibridni cloud.

U velikim sistemima poput obrazovnog, upotreba cloud rješenja nije doprinijela samo obezbjeđivanju alata za slanje i prijem pošte, nego je istovremeno omogućeno i korišćenje mnogih drugih alata koji podstiču online kolaboraciju među zaposlenim u obrazovno-vaspitnim ustanovama, kao i između nastavnika i učenika.

Na ovaj način povećava se primjena informaciono-komunikacionih tehnologija u nastavnom procesu za razmjenu iskustava, bržu komunikaciju i bolju obavještenost nastavnika, a sve u cilju modernizacije svih procesa u obrazovno-vaspitnoj ustanovi. To je glavni preduslov za postizanje optimalnih rezultata i povećanje nivoa ICT standarda u našem obrazovnom sistemu.

Pored modernizacije nastavnih procesa i procesa unutar obrazovnih ustanova, a imajući u vidu veličinu i kompleksnost obrazovnog sistema, cloud sistemi mogu doprinijeti da sa značajno unaprijedi i razvije niz novih servisa koji se oslanjaju na informacijski sistem obrazovanja.

Pored značajnih ušteda u nadogradnji i proširenju data centra serverskom infrastrukturu, koja bi zahtjevala nakon određenog vremena obnavljanje, od podjednako

značaja je i manja potreba za dodatnim kadrom iz IT oblasti koji je s jedne strane u deficitu, dok sa druge strane postoje ograničenja u broju zaposlenih, jer sektor obrazovanja mora da ima i značajan broj zaposlenih iz drugih oblasti (nastavnika, pedagoga, psihologa, defektologa, logopeda itd).

Neophodno je unapređenje informacionog sistema obrazovanja upotrebom cloud computing-a, jer za tim postoje velike potrebe i dobra osnova. Pravci unapređenja mogu se ogledati u kreiranju profila svake obrazovno-vaspitne ustanove na kojem može da se prate svi statistički podaci ustanove kao i trendovi upisa učenika. Zatim, na mapi Crne Gore uz QGIS mogu se obraditi i publikovati prostorni podaci o ustanovama i slično.

Što se tiče bezbjednosti cloud sistema, prednost je što o cloud-u brinu stručna lica i što postoje rezervne lokacije. S druge strane, pravna regulativa u Crnoj Gori predviđa da se svi registri u kojima čuvaju lični podaci osoba moraju fizički nalaziti u Crnoj Gori. S tim u vezi, registar učenika i zaposlenih nalazi se u data centru Ministarstva prosvjete, dok se razni servisi mogu kreirati u cloud-u.

#### LITERATURA

- [1] Kevin T. McDonald, "Above the Clouds: Managing Risk in the World of Cloud Computing", IT Governance Publishing, Cambridgeshire, United Kingdom, 2010
- [2] Michael Hugos, Derek Hultitzky, "Business in the Cloud: What Every Business Need to Know About Cloud Computing", Wiley, John Wiley & Sons, Inc, 2010
- [3] Michael Kavis, "Architecting the Cloud: Design Decisions for Cloud Computing Service Models (SaaS, PaaS, and IaaS)", Wiley, John Wiley & Sons, Inc., Hoboken, New Jersey, 2014
- [4] A. Lenk, M. Klems, J. Nimis, S. Tai, T. Sandholm, "What's Inside the Cloud? An Architectural Map of the Cloud Landscape", Software Engineering Challenges of Cloud Computing, 2009. CLOUD '09. ICSE Workshop on, Vancouver, 2009
- [5] Ben Halpert, "Auditing Cloud Computing: A Security and Privacy Guide", Wiley, John Wiley & Sons, Inc, Hoboken, New Jersey, 2011
- [6] Lizhe Wang, Rajiv Ranjan, Junjin Chen, Boualem Benatallah, "Cloud Computing: Methodology, Systems and Applications", CRC Press, Taylor & Francis Group, 2012
- [7] Nedžad Mehić, Miljan Marković, Feđa Lekić, "Model bezbednog upravljanja heterogenim OS okruženjem u klauđu", BISEC, Beograd, 2015
- [8] Igor Ognjanović, R.Šendelj, I.Ognjanović, "Impact analyses of Cyber attacks on Cloud systems", BISEC, Beograd, 2016
- [9] Ministarstvo prosvjete i nauke, "Glavni projekat informacionog Sistema obrazovanja RCG", Podgorica 2004
- [10] Technology in Education, <https://teched-su2011.wikispaces.com/Cloud+Computing>

#### ABSTRACT

This paper presents the main advantages of Cloud Computing as well as its application in the educational system. Special attention is devoted to the analysis of Montenegrin educational information system, particularly to its modules that relate to MEIS application and implementation of Office365 in educational institutions.

#### USING CLOUD COMPUTING IN THE EDUCATION SYSTEM OF MONTENEGRO

Marina Matijević, *Ministry of Education*,  
Prof. dr Ramo Sendelj, *University Donja Gorica*

# Zaštita podataka u obrazovnom sistemu Crne Gore upotrebom OpenPGP standarda

Marina Matijević, *Ministarstvo prosvjete,*  
Doc. dr Ivana Ognjanović, *Univerzitet Donja Gorica*

**Sadržaj** — U ovom radu predstavljene su osnove zaštite podataka, odnosno osnove enkripcije. Predstavljen je OpenPGP standard, odnosno njegove dvije najpoznatije implementacije PGP i GnuPGP, kao i softver koji podržava taj standard i njegov princip rada. Poseban doprinos rada predstavlja kritička analiza upotrebe uz identifikaciju potencijalnih faktora na slabosti sistema bezbjednosti.

**Ključne riječi** — asimetrični algoritam, dekripcija, digitalni potpis, digitalni sertifikat, enkripcija, GPG, hash algoritam, javni ključ, PGP, PKI simetrični algoritam, šifrovanje, tajni ključ.

## I. UVOD

U savremenim društvenim dešavanjima podaci i informacije imaju veliku vrijednost u poslovanju. Njihova zaštita je, za svaku firmu, na primarnom mjestu.

Podatke je sve teže zaštititi jer postoje razne prijetnje koje dolaze spolja i iznutra, pa se sve češće može čuti da je gubljenje podataka česta pojava i problem. Uzroci zbog kojih dolazi do gubitka podataka su različiti. Nekada dođe do gubitka nekog prenosivog uređaja, često zbog presrijetanja elektronske pošte, nedozvoljenog pristupa podacima ili neznanja.

Za svaki navedeni uzrok postoje i određene mjere koje se mogu preduzeti da ne dođe do gubitka ili zloupotrebe određenih podataka ili informacija. U ovom radu posebna pažnja će biti posvećena zaštiti podataka koji se prenose elektronskim putem odnosno elektronskom poštom, jer je razmjena podataka putem elektronske pošte jedan od najzastupljenijih načina u poslovnoj i privatnoj komunikaciji.

Rad je strukturiran na sljedeći način: Poglavlje II daje uvod u enkripciju, dok Poglavlje III predstavlja OpenPGP standarda kao jedan od najpoznatijih i najrasprostranjenijih standarda za enkripciju, a konkretan softver baziran na OpenPGP standardu je opisan u poglavlju IV. Poglavlje V je centralni dio rada koji identifikuje osnovne karakteristike informacionog sistema implementiranog u sistemu obrazovanja u CG, koje potencijalno predstavljaju nedostatke za bezbjednost sistema, koje je neophodno unaprijediti. Poglavlje VI završava rada s glavnim

Marina Matijević, *Ministarstvo prosvjete, ul. Vaka Đurovića bb, 81000 Podgorica, Crna Gora (telefon: 382-20-405-310; email: marina.matijevic@mpps.gov.me)*

Ivana Ognjanović, *Univerzitet Donja Gorica, Donja Gorica, 81000 Podgorica, Crna Gora (telefon: 382-20-410-777; email: ivana.ognjanovic.edu@gmail.com)*

zaključcima i smjernicama za dalje istraživanje.

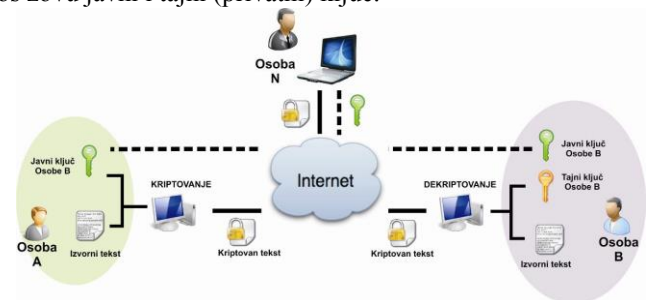
## II. ENKRIPCJA

SMTP je protokol za prenos poruka putem elektronske pošte u tekstualnom formatu. Podatak o pošiljaocu je njegova mail adresa koja je takođe u tekstualnom formatu. S tim u vezi, onaj ko presretne poruku može ili izmijeniti sadržaj poruke ili podatke o pošiljaocu.

Kako danas postoje mnogi pojedinci i organizacije koje imaju razne motive za presrijetanje komunikacije, mora se dodatno voditi računa o načinu na koji se podaci razmjenjuju. U cilju sigurne mrežne komunikacije nastao je protokol za enkripciju i potpisivanje poruka.

Enkripcija (šifrovanje) predstavlja proces u kojem dolazi do izmjene podataka tako da podaci budu nerazumljivi onim osobama kojima nijesu namijenjeni i ne posjeduju određeni ključ. Da bi podaci postali razumljivi i upotrebljivi neophodno ih je dešifrovati, odnosno dekriptovati.

Dvije osnovne vrste enkripcije su simetrična i asimetrična enkripcija. Simetrična enkripcija za enkripciju i dekripciju koristi isti ključ, dok asimetrična za enkripciju koristi jedan, a za dekripciju drugi ključ. Ta dva ključa se još zovu javni i tajni (privatni) ključ.



Sl.1. Postupak asimetričnog kriptovanja

Danas postoje mnogi standardi za enkripciju, a najrasprostranjeniji je OpenPGP standard.

## III. OPENPGP STANDARD

Dvije najpoznatije implementacije OpenPGP standarda su PGP (Pretty Good Privacy) i GnuPG (GNU Privacy Guard) programi.

Phil Zimmermann je tvorac prve verzije PGP programa koji se zasnivao na simetričnom i asimetričnom algoritmu za enkripciju, kao i onom za stvaranje hash koda poruke. Kroz sve kasnije verzije PGP programa, princip rada je ostao zadržan a mijenjao se samo korišćeni algoritam.



Kako je algoritam za asimetričnu enkripciju bio zaštićen patentom, u prvoj verziji RSA (Rivest Shamir Adleman) algoritmom a u drugoj IDEA (International Data Encryption Algorithm), to je dovelo do komercijalizacije softvera i spriječilo njegovo veliko širenje. Prva legalna verzija koja se koristila u SAD je bila 2.5 koja je umjesto RSA algoritma koristila besplatnu RSAREF biblioteku funkcija.

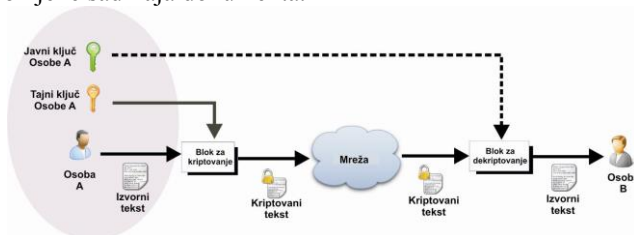
Nastavkom razvoja PGP programa došlo se do verzije 5.0 koja je sadržala mnoge enkripcijske algoritme i to za simetrični dio IDEA, Triple-DES i CAST, a za asimetrični dio RSA, Hellman i DSS/DSA, dok se kao hash algoritam koristilo MD5 ili SHA-1. Tek pojavom verzije 6.0 pojavljuje se mogućnost integracije PGP-a sa MS Outlook-om ili MS Outlook Express-om.

Open Source zajednica je razvila verziju GnuPG (GNU Privacy Guard) programa kojim su eliminisani problemi koji su se javljali kod distribucije PGP programa, kako zbog patenata, tako i zbog zakonskih regulativa.

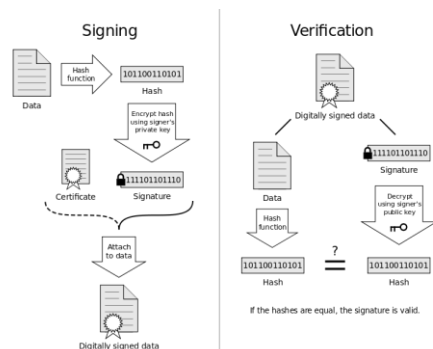
PGP i GnuPG programi omogućavaju enkripciju i digitalni potpis dokumenata i oba programa svoj rad zasnivaju na kombinaciji raznih enkripcijskih algoritama čime se postiže veći nivo sigurnosti.

Dakle, oba programa koriste hash algoritam, simetričnu i asimetričnu enkripciju. Hash algoritam predstavlja skup matematičkih operacija pomoću kojih se iz dokumenata kreira jedinstven sadržaj koji se zove hash kod. Važno je napomenuti da je kod hash algoritma nemoguće dobiti dva hash koda za dva različita dokumenta. Čak i izmjena samo jednog karaktera u dokumentu utiče na izmjenju njegovog hash koda.

Da bi se digitalno potpisao neki dokument neophodno je dobiti hash kod dokumenta primjenom neke od hash algoritama. Dobijeni hash kod se enkriptuje tajnim (privatnim) ključem pošiljaoca i dodaje na početak poruke. Primalac može dekriptovati hash kod ukoliko posjeduje javni ključ pošiljaoca i uporediti hash kod sa samim hash kodom koji je dobio iz dokumenta koji je primio. Ukoliko su kodovi jednaki može biti siguran da nije došlo do promjene sadržaja dokumenta.

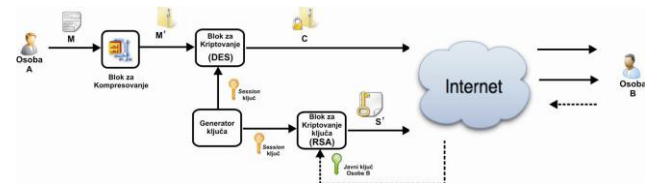


Sl.2. Digitalno potpisivanje cijele poruke



Sl.3. Dijagram primjene digitalnog potpisa i verifikacija

Prije enkripcije dokument se može kompresovati ZIP algoritmom a zatim se enkriptuje simetričnom enkripcijom. Ukoliko dokument treba da bude digitalno potpisan, onda se prvo vrši digitalno potpisivanje dokumenta a nakon toga kompresovanje i enkripcija. Kompresijom dokumenta postiže se bolja zaštita jer se uklanja redundantna informacija iz dokumenta. Neophodno je da se hash kod dobije prije kompresije jer je ZIP nedeterministički algoritam.



Sl.4. Postupak kriptovanja pomoću PGP algoritma

Za simetričnu enkripciju koristi se slučajno generisani ključ i svaki put je različit. Zatim se taj ključ enkriptuje asimetričnim algoritmom uz pomoć javnog ključa primaoca poruke i dodaje se na početku poruke.

Da bi mogla da se primijeni enkripcija asimetričnim algoritmom pomoću javnog ključa, potrebno je, prije toga, generisati barem jedan par ključeva za asimetričnu enkripciju. Generisani ključevi su predugački i teško pamtljivi za ljude pa se često primjenjuje šifrovanje privatnog ključa kratkom šifrom i on se čuva na lokalnom disku. Ta šifra se koristi svaki put kada se enkriptuje neki dokument. Kao što je poznato, da bi primalac mogao pročitati dokument koji mu se pošalje mora dobiti javni ključ iz para generisanih ključeva čiji privatni ključ se koristio za enkripciju.

Iz ugla kriptanalize može se reći da PGP i GnuPG su na visokom nivou kada je sigurnost u pitanju ali i dalje je potrebno obratiti pažnju na određene elemente kako bi eliminisali svaki mogući rizik.

Naime, kako se tajni (privatni) ključ čuva na lokalnom disku i on je šifrovavan nekom kraćom šifrom, potrebno je povesti računa da se šifra pažljivo odabere kako ne bi predstavljala samo pojam iz svakodnevnog života jer, ukoliko neko sazna tu šifru, teoretski, on može dešifrovati tajni ključ.

Nedovoljno jaka šifra nije jedini potencijalni rizik u ovom procesu. Da bi komunikacija mogla pouzdano da se obavlja, primalac mora posjedovati javni ključ pošiljaoca. Javni ključ može da se šalje putem maila, na USB-u, CD-u ili da se javni ključ objavi na nekom serveru kako bi svima bio dostupan. To otvara mogućnost da bilo ko postavi javni ključ a navede da on pripada nekoj drugoj osobi. Na taj način, sve što je usmjereno za osobu, čiji je ključ navodno objavljen, od osoba koja je postavljala ključ na serveru može da se zloupotrijebi.

Ovaj problem je prevaziđen uvođen PKI (Public Key Infrastructure) sistema za izdavanje digitalnog sertifikata. Digitalnim sertifikatom se potvrđuje autentičnost nečijeg javnog ključa. Tačnije, javni ključevi koji treba da se razmjenjuju potpisuju se privatnim ključem organizacije koja pruža usluge izdavanja digitalnog sertifikata.

## IV. SOFTVER BAZIRAN NA OPENPGP STANDARDU

Razvojem PGP standarda razvijao su se i razne verzije softvera a najopštija podjela je: PGP verzija 2.x, PGP verzija 5.x i više i GnuPG.

Od PGP verzije 2.x najnovija verzija programa je PGP 2.6.2. Ovaj program ne posjeduje grafičko okruženje već se može koristiti samo iz komandne linije.

Već od verzija PGP programa 5.x i većih postoji razvijeno grafičko okruženje. Ove verzije programa su komercijalne a najnovija, dostupna na tržištu, je PGP 8.1. Program je dostupan za operativne sisteme Windows i Mac. Pored toga što nudi mogućnost zaštite podataka skladištenih na lokalnom disku, on se prilikom instalacije integriše s aplikacijama za elektronsku poštu (MS Outlook i MS Outlook Express) pa je i razmjena dokumenata putem elektronske pošte sigurna ukoliko se koristi digitalno potpisivanje i enkripcija poslatih poruka.

Open Source zajednica je razvila program GnuPG (GPG) koji je besplatan za korišćenje i nezavistan, za razliku od PGP programa i raznih patenata. Samo osnovna verzija programa nije imala grafičko okruženje. Program je dostupan za Windows, Linux, Unix i Mac operativne sisteme.

Za Windows operativni sistem koristi se program GPG4WIN, koji ima grafičko okruženje Kleopatra. Prvo je neophodno kreirati digitalni sertifikat. Pokrene se aplikacija Kleopatra, zatim se odaberu opcije File-New Certificate-Create a personal OpenPGP key pair, nakon čega se unese ime, prezime i email adresa. Na odabir opcije Create Key, otvara se prozor za unos lozinke. Važno je odabrati dobru lozinku, jer se pomoću nje štiti tajni (privatni) ključ na lokalnom disku. U slučaju da lozinka nije dovoljno jaka, program upozorava i daje mogućnost da se zadrži ista lozinka ili da se unese druga.

Nakon kreiranja sertifikata potrebno je eksportovati sertifikat. U listi My Certificates nalazi se kreirani sertifikat koji se može putem opcije Export eksportovati u fajl koji ima ekstenziju .asc. Takav fajl se može otvoriti samo u Kleopatri i u njemu se nalazi javni ključ sertifikata koji je kreiran. Da bi se postigla još veća sigurnost, javni ključ se često šalje u tekstualnim obliku.

Kada se dobije javni ključ ili preuzme sa interneta veoma je važno da on bude u .asc formatu a ne u .txt prije njegove upotrebe, kako bi se mogao importovati u program Kleopatra (Import Certificates).

Svaka vrsta dokumenta može digitalno da se potpiše, bilo da se radi o office dokumentu ili pdf, fotografiji, audio ili video zapisu itd. U proram Kleopatra se to radi na način što odaberemo opcije File-Sing/Encrypt files, zatim se odabere dokument koji se enkriptuje, bira se, da li želimo dokument da se enkriptuje i potpiše, samo enkriptuje i samo potpiše. Ukoliko želimo dokument enkriptovati i potpisati, odabere se opcija Sign and Encrypt. U narednom koraku potrebno je odabrati za koga se poruka enkriptuje. Ona se može enkriptovati samo za osobe čiji javni ključ posjedujemo i on se nalazi u listi. Sljedeći korak je izbor sertifikata kojim se potpisuje ova poruka. U tom slučaju se bira i unosi šifra sertifikata koji je na početku kreiran. Ukoliko je sve urađeno kako treba, dobija se poruka Singnin and encryption succeeded i formira se fajl sa ekstenzijom .pgp.

Pomoću Kleopatre može se dekriptovati određeni dokument korišćenjem opcije Decrypt/Verify Files nakon čega se odabere dokument. Potrebno je unijeti šifru sertifikata kako bi se dekriptovao fajl.

Kao što je već navedeno, Gnu PG se može koristiti i u Linux operativnom sistemu. Ukoliko se koristi za rad iz komandne linije tada postoji set naredbi koje prate isti proces korišćenja programa za Windows platformu:

- gpg --gen --key (generisanje para ključeva za asimetričnu enkripciju)
- gpg --armor --export (eksportovanje ključa, armor određuje da li će ključ biti zapisan u tekstualnom ili binarnom obliku)
- gpg --import ImeDokumenta (ključ se može čuvati u dokumentu)
- gpg --recipient primalac --encrypt ImeDokumenta (enkriptovanje dokumenta, kao rezultat naredbe dobija se dokument sa ekstenzijom .pgp)
- gpg --recipient primalac --armor --encrypt ImeDokumenta (enkriptovanje dokumenta, kao rezultat se dobija dokument sa ekstenzijom .asc)
- gpg --decrypt ImeDokumenta (dekriptovanje dokumenta)
- gpg --sing ImeDokumenta (potpisivanje dokumenta, ukoliko se koristi opcija armor moguće je kreirati tekstualnu datoteku)
- gpg --sing --detach ImeDokumenta (potpisivanje dokumenta, kao rezultat dobije se dokument sa ekstenzijom .sig)
- gpg --verify ImeDokumenta (provjera autentičnosti digitalnog potpisa)
- gpg --encrypt --sign ImeDokumenta (enkripcija i digitalno potpisivanje dokumenta jednom istovremeno).

## V. ANALIZA UPOTREBE I UNAPREĐENJE ZAŠTITE PODATAKA U OBRAZOVNOM SISTEMU CRNE GORE

U obrazovnom sistemu Crne Gore teži se ka što većoj primjeni informaciono-komunikacionih tehnologija kako u nastavnom procesu tako i u komunikaciji među zaposlenima i online kolaboraciji. Samim tim, potreba za zaštitom podataka postaje sve veća. Sagledavanjem osnovnih karakteristika korisnika sistema, samog sistema, kao i trendova u razvoju IKT, a samim tim i sve više opasnosti po sigurnost podataka u sistemu, mogu se identifikovati sljedeći potencijalno kritični faktori:

**Nivo edukacije zaposlenih.** Zaposleni u obrazovno-vaspitnim ustanovama o zaštiti podataka znaju više samo na teorijskom nivou. Što se tiče primjene, ona je poprilično osnovna, npr. vode računa da obezbijede podatke na svojim računarima od gubljenja, međutim, sama zaštita podataka (enkripcija podataka, upotreba digitalnog potpisa) nije toliko primjenjena.

**Dokumenta sa povjerljivim podacima.** Činjenica je da se u pedagoškoj dokumentaciji, koju vode obrazovno-vaspitne ustanove nalazi i niz podataka koje bi trebali zaštititi (informacije o učenicima), odnosno za koje bi morali obezbijediti adekvatnu zaštitu kada se prosljeđuju ili razmjenjuju među kolegama. U pedagoškoj dokumentaciji se nalaze podaci o maloljetnim osobama

među kojima su takođe i lični podaci, pa je samim tim potreba za zaštitom podataka neophodna. Pored razmjene podataka među zaposlenima često se obavlja i razmjena podataka između obrazovnih ustanova i centara za socijalni rad, kao i sa centrima bezbjednosti.

**Svijest o sajber bezbjednosti.** Za unapređenje i obezbjeđenje bolje zaštite podataka kao i samu komunikaciju i razmjenu podataka za koju je zaštita neophodna, prvo bi trebalo sistemski raditi na podizanju svijesti zaposlenih u obrazovno-vaspitnim ustanovama o značaju i potrebi da se vodi računa o potpunoj zaštiti podataka. Tek kada se razumiju pravi razlozi za primjenu raznih modela za zaštitu podataka tek tada će se i primjena u značajnoj mjeri unaprijediti.

**Specijalistička znanja o sajber bezbjednosti (preventivne i proaktivne mjere).** Nakon podizanja svijesti kod zaposlenih o značaju i potrebi za zaštitom podataka, neophodno je u svim obrazovno-vaspitnim ustanovama obezbjeđivati i odgovarajuće obuke za zaposlene kako bi znali primjenjivati razne modele za zaštitu podataka.

**Tehnički zahtjevi.** Kako je na svim računarima instaliran Windows operativni sistem tada bi najefikasnije bilo instalirati GPG4WIN program i uvesti primjenu GnuPG standarda za zaštitu podataka, jer je i samo korišćenje ovog softvera jednostavno. Pored toga, obrazovni sistem je veliki sistem pa je poželjno da se koristi unificirano rješenje za sve ustanove, da bi se na lakši način pružila potrebna tehnička podrška ustanovama sa centralnog nivoa (Ministarstva prosvjete). GPG4WIN je besplatan open source program što je takođe prednost, jer bi korišćenje nekog komercijalnog programa zahtjevalo izdvajanje značajnih finansijskih sredstava za licence, zbog velikog broja obrazovnih ustanova, odnosno velikog broja zaposlenih i računara.

S druge strane, na osnovu Strategije korišćenja Open Source tehnologija u Crnoj Gori, cilj je da open source rješenja budu što više zastupljena u svim sferama našeg društva pa tako i u obrazovnom sistemu.

Unaprijeđenje primjene zaštite podataka kod zaposlenih u obrazovnom sistemu ima višestruki značaj. Naime, zaposleni u obrazovnim ustanovama su u svakodnevnoj komunikaciji sa djecom, pa samim tim oni najviše utiču na njihovo obrazovanje i vaspitanje. Ukoliko je svijest i znanje nastavnika o zaštiti podataka na visokom nivou, tada će oni svoje znanje iz ove oblasti prenositi i na djecu. Na taj način, obezbjeđen je kontinuitet u podizanju svijesti o značaju zaštite podataka i kod velikog broja osoba, koji su sada djeca, ali koji će za nekoliko godina raditi određene poslove na kojima će zaštita podataka sigurno biti jedna od neophodnih aktivnosti.

## VI. ZAKLJUČAK

Zaštita podataka u današnjem vremenu ima poseban značaj jer je sve više zainteresovanih lica koja na razne načine žele doći do tuđih podataka i iskoristiti ih u razne svrhe (najčešće za ostvarivanje određene finansijske dobiti). Pored odgovornog čuvanja podataka na računaru ključnu ulogu ima i sigurna razmjena podataka jer je činjenica da podatke moramo razmjenjivati. Najčešći model za razmjenu podataka je putem elektronske pošte i

upravo korišćenje digitalnog sertifikata i digitalnog potisa su našli veliku primjenu u tom dijelu. OpenPGP standard se pokazao kao dosta pouzdan kada je riječ o ovom problemu i nudi elementarnu zaštitu na jednostavan način za svakog korisnika i podržan je za sve verzije operativnih sistema.

U obrazovnom sistemu zaštita podataka je od velikog značaja zbog velikog broja informacija koje škole posjeduju o maloljetnim licima, pa to zahtjeva pored fizičkog obezbjeđivanja i zaštite podataka na lokalnim diskovima, i zaštitu podataka koja se prenosi putem elektronske pošte ili dijeljenjem među kolegama. Imajući u vidu veličinu obrazovnog sistema, primjena open source rješenja je pogodna zbog uštede finansijskih sredstava jer ne zahtjeva kupovinu velikog broja licenci.

## LITERATURA

- [1] Xiaoyun Wang, Guangwu Xu, Mingqiang Wang, Xianmeng Meng, "Mathematical Foundations of Public Key Cryptography", CRC Press, Taylor & Francis Group, 2016
- [2] Chuck Easttom, "Modern Cryptography: Applied Mathematics for Encryption and Information Security", McGraw Hill Education, 2016
- [3] Vlastimil Klima, Rosa Toma, "Attack on private signature keys of the OpenPGP format." PGPM programs and other applications compatible with OpenPGP, 2001, <http://eprint.iacr.org/2002/076.pdf>, 2001
- [4] Henk Van Tilborg, "Encyclopedia of Cryptography and Security", Springer Science & Business Media inc, New York, 2005
- [5] Antonie Joux, "Algorithmic Cryptanalysis", CPC Press, Taylor & Francis Group, 2009
- [6] William Stallings, "Cryptography and Network Security", Prentice Hall, 2005
- [7] Ramo Šendelj, Ivana Ognjanović, Matjaž Debevc, Marko Holbl, Marko Kääramees, Massimo Capelli, Florin Ioras, Stefano Guarino, "EU Practice for Cyber Security Education", Univerzitet Donja Gorica, Podgorica, 2015
- [8] Ramo Šendelj, Ivana Ognjanović, Matjaž Debevc, Marko Holbl, Marko Kääramees, Massimo Capelli, Indrachapa Bandara, Flavio Lombardi, "Existing EU Practices for Cyber Security", Univerzitet Donja Gorica, 2015
- [9] Ramo Šendelj, Ivana Ognjanović, "Cyber Security Education in Montenegro: current trends, challenges and open perspectives", 7th annual International Conference on Education and New Learning Technologies (EDULEARN15), Barcelona, Spain, 2015
- [10] Stefan A. Brands, "Rethinking public key infrastructures and digital certificates", The MIT Press, London, England, 2000

## ABSTRACT

This paper presents the basics of data protection and the basics of encryption. The OpenPGP standard is presented and its two best known implementations GnuPG and PGP, as well as software that supports this standard and its working principle. The critical analysis of the use with identification of the potential factors of the weaknesses of the security system represents particular contribution of the work.

## DATA PROTECTION IN THE EDUCATIONAL SYSTEM OF MONTENEGRO USING OPENPGP STANDARD

Marina Matijević, *Ministry of Education*,  
Doc. dr Ivana Ognjanović, *University Donja Gorica*

# Impact of web-based learning management systems on education at the Faculty of Medicine in Podgorica, Montenegro

Nataša Popović, Aleksandra Radulović, Isidora Rovčanin Dragović, Oleg Cmiljanić and Sehija Dizdarević

**Abstract** — In this paper we evaluate the impact of blended web-based learning via Moodle on the education of students at the Faculty of Medicine, University of Montenegro. The two main goals of our study were to determine impact of e-learning technologies on student progress in mastering the course, and to assess learner satisfaction after the introduction of e-learning. To accomplish these goals we compared a group of students who attended the physiology course before Moodle was implemented as an educational tool, with a group of students who attended the physiology course after the introduction of Moodle. We compared formative assessment scores between these two groups. In addition, we analyzed the impact of high vs. low Moodle use on the formative assessment scores, and conducted the survey to assess the satisfaction among Moodle users.

**Keywords** — blended learning, Moodle in physiology, web-based learning.

## I. INTRODUCTION

The development of information technology at the end of the 20<sup>th</sup> century had a significant impact on many aspects of our lives. The Internet is perceived as one of the most important sources of new information by many, which emphasizes the significance of information technologies [1]. Therefore, web-based learning is becoming a new standard in the process of education. In accordance with this notion, MIT attempted to offer all of its courses online [2]. There are numerous course management systems designed to aid web-based learning. Some of them are commercial packages (for example Blackboard), while others are open source and therefore free to download and use (Moodle, Open eClass). Moodle (modular object-oriented dynamic learning environment) is a web-based platform adopted worldwide and translated in more than 30 languages [3], [4]. In 2013, Moodle was adopted by 23% of the higher education institutions in the United States, which makes it the second most common web-based course management system next to Blackboard [5].

Web-based course management platforms can facilitate blended learning and distance learning [6]. Blended learning combines traditional face-to-face education with

the use of informational technologies and computer-mediated delivery of educational material, as opposed to distance learning which is mostly computer-mediated. Research showed that web-based learning has numerous advantages over the traditional style of learning. It gives students the flexibility to choose time and place to study, improves communication among students, and between students and the course instructor. It provides an efficient way to introduce the use of multimedia presentations and additional links to web sites, which results in increased ease of adopting the new information and development of critical thinking. It diversifies the means of student assessment and allows the course instructor to provide feedback to the students in a more efficient and timely manner [7]. On the other hand, one of the main concerns among some course instructors is that attendance of face-to-face lectures will decrease if the lecture notes are made readily available through web-based course management platforms [8]. Despite these concerns, the web-based course management systems are being introduced at the medical education graduate and postgraduate levels, as well as continuing medical education. Medical schools using Moodle as a platform in teaching physiology reported generally positive experience [8], [9]. At the University of Montenegro, an e-learning concept using Moodle as a platform for course management was established in 2007 [10]. Most of the courses at the University of Montenegro use Moodle as a simple repository for lecture notes, some additionally use it for administration of exams and tests, and only a few use it for streaming lectures and delivering the complete course in an online form [10]. Physiology is the first course offered at the Faculty of Medicine to use this type of web-based learning technology.

The two main goals of our study were to determine the impact of e-learning technologies on student progress in mastering the physiology course, and to assess learner satisfaction after the introduction of e-learning.

## II. MATERIALS AND METHODS

### A. Organization of physiology course

Students studying at the Faculty of Medicine, University of Montenegro, can enroll into one of the 3 professional tracks: School of Medicine (to earn the degree of Medical Doctor), School of Dentistry (to earn the degree of Doctor of Stomatology), and School of Pharmacy (to earn the degree of Doctor of Pharmacy). Similar criteria are used for acceptance of the students into their respective programs every year. Physiology is a mandatory course for all students at the Faculty of Medicine during their 2<sup>nd</sup> year of studies, regardless which professional track they belong to.

The physiology course consists of face-to-face theory lectures, practical classes, and seminars during which students give oral presentation on a given topic in front of

Nataša Popović, Faculty of Medicine, University of Montenegro, Kruševac bb, 81000 Podgorica, Montenegro (tel: +382/20/246-651, e-mail: npopovic@ac.me)

Aleksandra Radulović, Center of Information Technology, University of Montenegro, Cetinjska br. 2, 81000 Podgorica, Montenegro (tel: +382/20/414-282, e-mail: saska@ac.me)

Isidora Rovčanin Dragović, Faculty of Medicine, University of Montenegro, Kruševac bb, 81000 Podgorica, Montenegro (tel: +382/20/246-651, e-mail: isidorarovcanin@yahoo.com)

Oleg Cmiljanić, Faculty of Medicine, University of Montenegro, Kruševac bb, 81000 Podgorica, Montenegro (tel: +382/20/246-651, e-mail: olegcmiljanic@yahoo.com)

Sehija Dizdarević, Faculty of Medicine, University of Montenegro, Kruševac bb, 81000 Podgorica, Montenegro (tel: +382/20/246-651, e-mail: sehijabp@hotmail.com)

the entire student audience. The difficulty level of the course and number of credit hours differs depending on the professional track (based on the European Credit Transfer System (ECTS): 11 credits in the Fall, and 11 in the Spring semester for the School of Medicine; 5.5 credits in the Fall and 5.5 credits in the Spring semester for the School of Dentistry; 5 credits in the Fall and 7 in the Spring semester for the School of Pharmacy) At the school of medicine and school of dentistry physiology is a two semester course, with each semester lasting 15 weeks. At the school of pharmacy physiology is divided into two separate courses, each lasting 1 semester.

### B. Participant groups and implementation of information technologies

The study included 113 students enrolled in physiology course during the Fall semester of the 2015-2016 academic year (Group 1: 49 medical, 24 dental, and 40 pharmacy students), and 97 students enrolled in physiology course during the Fall semester of the 2016-2017 academic year (Group 2: 42 medical, 17 dental, and 38 pharmacy students). Distribution of students in 3 professional tracks was similar in both groups (Fig.1). Moodle was introduced as an educational tool in the Spring semester of the 2016-2017 academic year. Therefore, during the Fall semester Group 1 relied very little, while group 2 relied heavily on information technologies. There were no major changes in the teaching staff, curriculum, schedule of classes between the two groups compared.

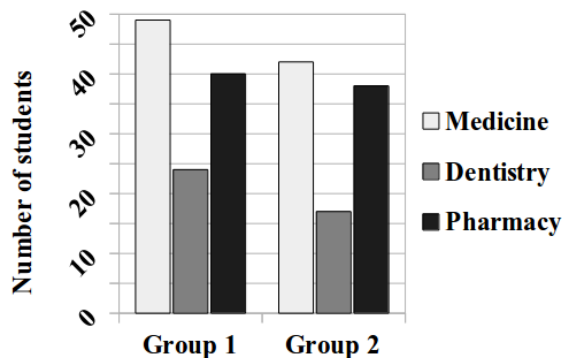


Fig. 1. Group 1- low use of informational technologies. Group 2- high use of information technologies in the form of blended learning. Distribution of students in 3 professional tracks was similar in both groups.

For group 1 the theory lectures were face-to-face, and the lecture notes were available for students to download from a password protected web site repository. During practical classes students observed lab demonstrations in groups and were obligated to personally perform some of the lab exercises. PhysioEx lab simulations and video recordings of lab demonstrations were used only occasionally (expressed as a percent of the total number of hours of practical classes: 8.33% for medical students, 27.3% for dental students and 33.3% for pharmacy students).

For group 2 delivery of theory lectures was still face-to-face. However, the complete curriculum was chronologically organized by weeks and published online through Moodle. In addition, the lecture notes were supplemented with video clips, test questions, and links to

informative web sites. Students had the chance to test their knowledge by taking online practice tests. A student forum feature was available, but was not used. During practical classes, lab demonstrations were replaced completely by video clips and PhysioEx lab simulations. This way recorded demonstrations could be watched individually as many times as necessary during the class, which made the learning process more flexible and adjustable to each students' learning speed. Students were still obligated to personally perform those lab exercises that were focused on developing essential skills for work with patients (for example: performing exam of deep tendon reflexes).

### C. Formative assessment

In the study presented here, to evaluate student learning progress for the group 1, two term tests in paper form were conducted during the course of the Fall semester. Also, paper-based practical quizzes were conducted during practical classes only for the students enrolled to become Medical Doctors.

To evaluate student learning progress for the group 2, the two term tests were conducted in the computerized form through Moodle in the Faculty of Medicine classrooms. In addition, computer-based practical quizzes were conducted through Moodle at the end of the every practical class, which made them more streamlined so they became a standard way of student performance evaluation for all professional program tracks.

Participation in all practical quizzes and term tests was mandatory. In the study presented here, we measured student progress in mastering the course through the results of the paper-based term tests for the group 1, and computer-delivered term tests for the group 2. Since paper-based practical quizzes were not conducted for all 3 professional tracks in group 1, results of the quizzes were not included in the study.

### D. Feedback after formative evaluation

For the group 1 paper-based term tests were graded within 2-3 days and feedback on incorrectly answered questions was provided at the students' request during the student consultation time. Students in the group 2 would get the feedback and results from the term test students immediately after the test was completed.

### E. Grading

Each term test consisted of 60 multiple choice questions. For the purpose of this study the test grades for each student were expressed as % of correctly answered questions in both tests together. The students were then grouped in 5 groups according to their formative testing scores as following: 5 (0-50%), 6 (51-60%), 7(61-70), 8 (81-90%), 10 (91-100%).

### F. Moodle usage

In order to define high-, mid- and low-volume Moodle users, we counted total number of times each student in the group 2 accessed items offered through Moodle during the 15-week period. Moreover, we assumed the number of clicks could be the measure of students' active participation in tests and quizzes, so additionally we counted total number of clicks each student performed while using Moodle. The distribution of Moodle usage data points was verified by finding quartiles. Data points

that were below lower boundary (1<sup>st</sup> quartile–1.5 x interquartile range) and above upper boundary (3<sup>rd</sup> quartile + 1.5 x interquartile range) were considered to be outliers and were excluded from further analysis (9 out of 93, or 9.3% of all points). Data points belonging to the the 1<sup>st</sup> quartile represented low-volume, those in the 2<sup>nd</sup> and 3<sup>rd</sup> quartile mid-volume, and in the 4<sup>th</sup> quartile represented high-volume Moodle user group.

### G. Statistical analysis

Statistical analysis was done by using a t-test and ANOVA in the LibreOffice Linux software package.

### H. Survey

To assess learner satisfaction after the introduction of e-learning, we conducted the survey among the students at the end of the course. The participation in the survey was voluntary and anonymous.

Since Moodle was introduced in the Spring semester of the 2015-2016 academic year the students in group 1 did use Moodle during the Spring semester, so the survey reflects opinions of students who all used Moodle towards the end of the course. We surveyed 113 students: students from group 1 and students of the pharmacy professional track from group 2 (who completed Physiology 1 course). The survey consisted of 3 sections: demographic information, a section with answers rated on a 5 point Likert scale (5 denoting “I strongly agree”, 1 denoting “I strongly disagree”), and a section with open-ended questions. We used a survey model to address all dimensions that have been shown to have significant impact on learner satisfaction [7].

## III. EXPERIMENTAL RESULTS

In order to determine the impact of e-learning technologies on student learning progress, we compared formative test scores between group 1 and group 2. Group 1 scored slightly higher on formative tests compared to the group 2 ( $p=0.005$ , group 1 mean grade of 8.5 vs. group 2 mean grade of 7.9, data not shown).

To address the concern expressed by some instructors that attendance of face-to-face lectures would decrease if the lecture notes are made readily available through web-based course management platforms, we compared attendance of face-to-face lectures between group 1 and group 2, and found no statistically significant difference ( $p=0.12$ , group 1 average attendance 86.5% vs group 2 average attendance of 82.5%, data not shown). Moreover, students who attended face-to-face lectures scored higher grades in both group 1 and group 2 (Fig. 2 (a) and (b)). However, students who belonged to the high-volume Moodle users in group 2 scored the higher grades (Fig. 3). Moodle use shown in the Fig. 3 is measured by number of times each student accessed items offered through Moodle program during the 15-week Fall semester. Moodle use assessed by number of clicks each student made in the same period did not significantly affect the grade (data not shown).

The summary of notable survey results is shown in Table 1. The results are presented as a percent of the total number of surveyed students that had that particular answer. The two most common answers to each question are in bold. Additionally, the survey results showed that 57% students were accessing Moodle on their cell phone,

confirming that a large number of students accepted the innovation readily, and that the change fit very well with their daily habits. 78.8% of students stated they have average computer skills, 46.9% of students stated they have excellent, and 30.1% they have good Internet connection at the place of their residence.

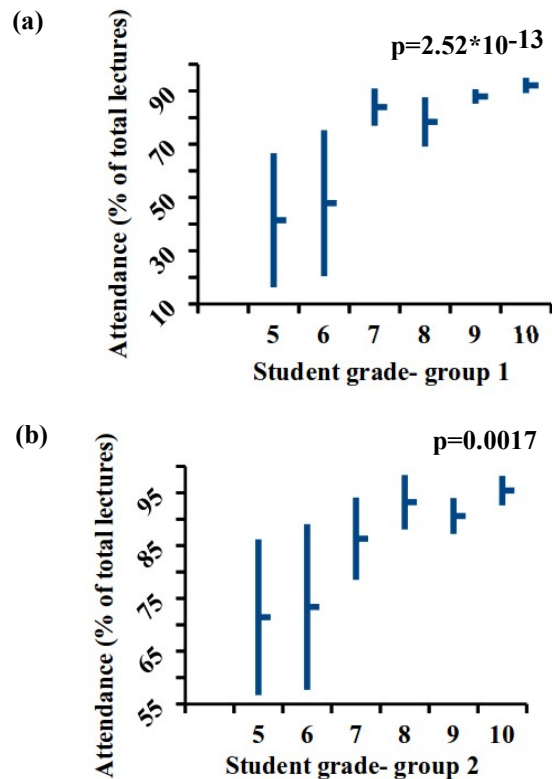


Fig. 2. (a) Students in the group 1 who more regularly attended face-to-face lectures had better grades on formative testing ( $p=2.52 \times 10^{-13}$ , mean attendance for students who scored 5- 41.4%, 6- 47.9%, 7- 83.9%, 8- 78.4%, 9- 87.9%, 10- 92.1%); (b) Students in group 2 who with more regular attendance also had better grades on formative testing ( $p=0.00017$ , mean attendance for individuals grades 5- 71.4%, 6- 73.4%, 7- 86.3%, 8- 93.2%, 9- 90.6%, 10- 95.4%).

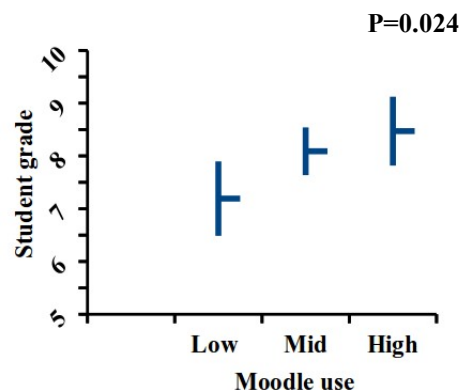


Fig. 3: Students in group 2 who used Moodle more frequently achieved better grades on formative testing ( $p=0.024$ , the mean grade for low volume users was 7.2, mid volume users 8.1, and high volume users 8.4). Moodle use was assessed by the number of times each student accessed items offered through Moodle in a 15-week period.

TABLE 1: SUMMARY OF THE RESULTS

<b>Graded answers</b>	<b>5</b>	<b>4</b>	<b>3</b>	<b>2</b>	<b>1</b>
Moodle was easy to use	<b>68</b>	<b>24</b>	4	2	0
Allows prompt feedback	<b>68</b>	<b>23</b>	4	2	1
Improves time management	<b>35</b>	17	<b>27</b>	15	7
Improves communication among students	19	14	<b>30</b>	<b>25</b>	12
Improves communication between teacher and students	<b>32</b>	<b>26</b>	24	12	6
It complements traditional methods of learning	<b>43</b>	<b>22</b>	16	12	5
It can completely substitute traditional style of learning	<b>28</b>	18	19	<b>23</b>	12
I would recommend its further use in physiology course	<b>45</b>	<b>29</b>	15	4	7

When asked what they did not like about the use of informational technologies, 15% of students felt that hands-on experience would be better than PhysioEx lab simulations and 9.7% of students thought that the school computers are slow and needed to be updated.

#### IV. DISCUSSION

Formative test scores got slightly lower (by 0.5 points) after Moodle was introduced. Other researchers show no difference in knowledge gain after the introduction web-based educational tools [8], [11], [12]. It is important to note that our results are based on formative testing and show only student progress up to the end of the Fall semester, so the practical significance of our finding is yet to be seen after all the students complete class, and their knowledge is evaluated during the the final oral exam. If we look only at the group of students who had computer-based term tests, formative test scores were higher in the high-volume users compared to low-volume Moodle users, which suggests that use of informational technologies might be useful after all.

Introduction of a web based educational tool did not change the attendance of the face-to-face lectures. The attendance was high for both groups (between 82.5% and 86.5%). Moreover, the students who attended face-to-face lectures more regularly had higher scores on the formative tests. This may be due to the fact that students with higher grades tend to have better work ethics, and therefore attend class more regularly regardless the mode of class delivery. However, this finding also underscores the importance of attending face- to- face lectures regularly. In accordance with this, the survey showed that while 65% students thought the informational technologies complemented traditional methods of learning, only 46% of students felt that informational technologies can completely substitute traditional style of learning.

However, most students agreed on several benefits of this innovation in learning: Moodle was easy to use (92%), it allowed prompt feedback (91%), improved their time management (52%), and improved communication between students and the teacher (58%). Seventy four percent of students recommended further use of Moodle in

the Physiology course. Collectively, these results suggest that students perceive the implementation of information technologies in the Physiology course as useful and that they were satisfied with these innovative changes.

Based on the experience of the teaching staff, Moodle is easy to use, has great flexibility with regards to editing and changing the web page, and it does not require any previous experience in web page programming. The use of web based learning platform decreases expenses related to supplying and maintaining laboratory material and equipment and it saves time in test preparation and test grading.

One of the limitations of our study is that the forum feature in Moodle was not used, which is reflected in the results of the student survey that showed only 33% of students felt Moodle improved communication among students. Also, we did not compare how information technologies affected summative assessment scores, because final exams and the survey of additional students will be done at the end of Spring semester of 2017, after all students in the study complete the course.

#### V. CONCLUSION

The study presented here shows promising results, but more experience is needed for further improvement of the quality of education and fine tune the teaching style with the goal to achieve the best possible learning outcomes.

#### REFERENCES

- [1] S. Guarino, E. Leopardi, S. Sorrenti, E. De Antoni, A. Catania, and S. Alagartnam, "Internet-based versus traditional teaching and learning methods," *Clin Teach*. Vol.11, no. 6, pp. 449-53, 2014.
- [2] J. Wu, R. J. Tsai, C. C. Chen, and Y. Wu, "An integrative model to predict the continuance use of electronic learning systems: hints for teaching," *International Journal of E-Learning*, vol. 5, no. 2, pp. 1537-2456, 2006.
- [3] M. Khabbaz, and R. Najjar, "Moodle-based distance language learning strategies: an evaluation of technology in language classroom," *International Journal of applied Linguistics and English Literature*, vol. 4, no. 4, pp. 205-210, 2015.
- [4] "Welcome to the Moodle community forums," Electronic material (Online), Available: <https://moodle.org/course/index.php>
- [5] K. C. Green, "A Profile of the LMS market" *The Campus Computing Project. The National Survey of Computing and Information Technology*, pp. 23, 2013. Available: <http://www.campuscomputing.net/item/2013-campus-computing-survey-0>
- [6] E. Costello, "Opening up to open source: looking at how Moodle was adopted in higher education," *Open Learning: The Journal of Open, Distance e-Learning*, vol. 28, no. 3, pp.187-200, 2013.
- [7] P.-C. Sun, R. J. Tsai, G. Finger, Y.-Y. Chen, and D. Yeh, "What drives successful e-Learning? An empirical investigation of the critical factors influencing learner satisfaction", *Computers and Education*, vol. 50, no. 4, pp.1183-1202, 2006.
- [8] K. Seluakumaran, F. F. Jusof, R. Ismail, R. and Husain, "Integrating and open-source course management system (Moodle) into the teaching of a first-year medical physiology course: a case study," *Adv Physiol Educ.*, vol. 35, pp. 369-377, 2011.
- [9] T. Gazibara, V. Marusic, G. Maric, M. Zaric, I. Vujcic, D. Kisic-Tepavcevic, J. Maksimovic, N. Maksimovic, Lj. Markovic Denic, Lj., S. Sipetic Grujicic, T. Pekmezovic, and A. Grgurevic, "Introducing e-learning in epidemiology course for undergraduate medical students at the Faculty of medicine, University of Belgrade: a pilot study." *J. Med Syst*. Vol. 39, pp. 121, 2015.
- [10] A. Radulovic, S. Loskovska, V. Devedzic, and B. Krstajic, "Practices of distance learning in Western Balkans: the issue of quality," *The Second international Conference on e-Learning*. vol. Elearning-2011, pp. 97-102, 2011.
- [11] E. Felder, M. Fauler, and S. Geiler, "Introducing e-learning/teaching in a physiology course for medical students: acceptance by students and subjective effect on learning," *Adv. Physiol. Educ.*, vol. 37, pp. 337-342, 2013.
- [12] D. A. Cook, D. M. Dupras, W. G. Thompson, and V. S. Pankratz, "Web-based learning in residents' continuity clinics: a randomized, controlled trial," *Acad. Med.*, vol. 80, no. 1, pp.90-97, 2005.

# ICS in the HPP Perucica as a national critical infrastructure

Ana Grbović, HPP Perucica, EPCG

**Abstract:** Because of the frequent examples of sophisticated cyber attacks, such as Stuxnet, Night Dragon and Duqu, it is necessary to implement new security measures in the ICS systems. The paper discusses about the importance of defining ICS as a critical national infrastructure. Security concepts take into account the specific requirements of process control which in general differ significantly from the requirements of Corporate IT. Also, paper describes the concept of deep hierarchy security known as “defense in depth” on the existing system, as well as the possibility of extending and improving the security of ICS (Industrial Control System) and SCADA (Supervisory Control and Data Acquisition) system at HPP Perucica.

**Key words** — critical infrastructure, ICS, SCADA, cyber security, defense-in-depth

## I. INTRODUCTION

Industrial process control systems (ICS) are control and monitoring networks and systems specially designed to support industrial processes. These systems are responsible for monitoring and managing the various processes and operations, such as the distribution of electricity and gas, water treatment, oil refining or rail transport. The largest subgroup of ICS is a SCADA (Eng. Supervisory Control And Data Acquisition), platform that is used in the HPP "Perucica" to control and monitor industrial processes - from simple commands on production units up to the incredibly complex operations associated with protective functions.

In recent years, ICS has gone through a significant transformation from proprietary, isolated systems to open architecture and standard technologies highly interconnected with other corporate networks and the Internet. All this has led to a reduction in costs, ease of use and the possibility of remote control and monitoring from different locations. However, an important drawback derived from the connection to the Internet and open communication networks has increased the sensitivity of computer networks to attacks [1].

SCADA systems' traditional role is changing as the Industrial Internet of Things (IIoT) continues to take a larger role. SCADA systems were not originally designed for cyber security and plants need to adjust to this new

reality. With the development of information systems there is a need for exchange of information between the Process and Corporate information systems, as well as the distribution of data from Process systems to the global internet network, which provides access to a wide range of applications.

ICS systems have roots in proprietary technology that was traditionally isolated from the enterprise information technology (IT) infrastructure. These platforms were not originally designed for cyber security [2].

Industrial process control systems represent a strategic asset against the growing number of catastrophic terrorist attacks that affect critical infrastructure. In the last decade, these systems are faced with a significant number of incidents, including Stuxnet, which has created a lot of concern and discussion among all actors involved in this field.

There are many different causes of incidents at ICS, such as random error of an administrative nature or errors that occur during the system updates, insider attacks and weaknesses. Risks of industrial sabotage or attack under the sponsorship of people whose aim is disruption of critical infrastructure are the most dangerous. These types of threats can have potentially devastating outcomes if they execute successfully. To protect the SCADA system, it is essential to evaluate the risk of its exposure to attacks and implement all necessary security measures.

## II. ICS SAFETY IN NUMBERS

### A. Frequency of the ICS systems

The ICS components available from the Internet (shown on Figure 1) are concentrated in the European region (41.41%) despite the modest numbers in individual countries. South America lags behind the Old World (37.46%), Asia holds the third place (12.39%).

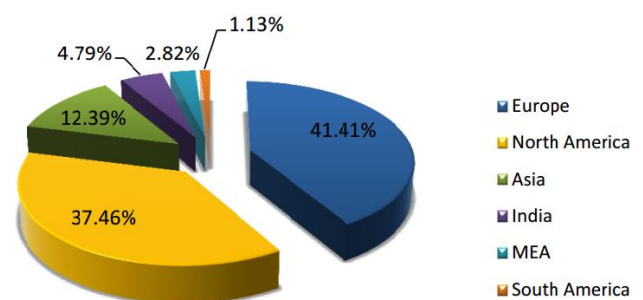


Figure 1. ICS Allocation by Regions

Almost the third part of the ICS systems, which elements can be accessed from the Internet, are located in



the USA (31.3%). Italy follows far behind (6.8%), South Korea rounds out the top three (6.2%). Russia holds the 12th place with 2.3%, and only 1.1% of ICS systems available from the global network are located in China [3].

### B. Types of ICS Systems

Most often the global network contains various SCADA components including HMI. They account for 70% of all detected objects. Another 27% of the ICS components are programmable logic controllers. Various network devices used in ICS networks (referred to as the Hardware in the following chart) were detected in 3% of cases [3].

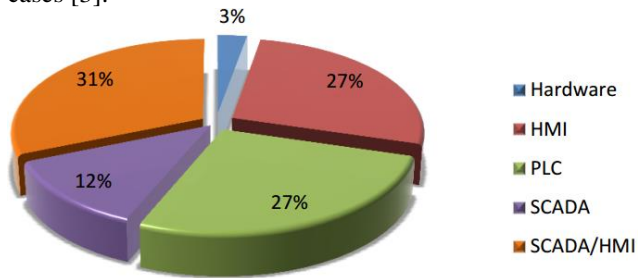


Figure 2. ICS Components

### C. Vulnerabilities of ICS systems

The report chart below indicates that 90% of the examined ICS vulnerabilities occurred from 2011-2015. Since Stuxnet was publicly disclosed in mid-2010, it could have triggered increased interest in discovering control system vulnerabilities and exploits. Also, between 2014 and 2015 there is a 49% increase [4].

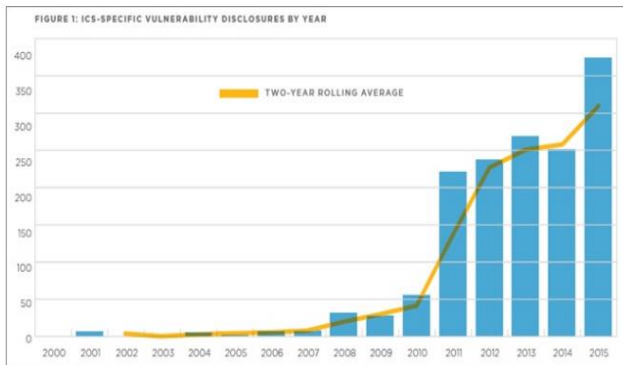


Figure 3. ICS specific vulnerability disclosures by year

Stuxnet is the first virus that was created to attack critical infrastructure of power plants and electrical networks. It is an incredibly sophisticated malware that has a clearly defined target. If, for example, it appeared on a computer, no damage will be made, unless the PC is not connected to the SCADA system specific [5].

## III. CRITICAL REVIEW OF THE CURRENT ICS SECURITY POLICY AND PRACTICE AS A NATIONAL CRITICAL INFRASTRUCTURE

The key documents on cyber security in Montenegro are: Study with defined competences of state bodies in combating cyber crime (adopted in 2012), Cyber security

strategy of Montenegro (adopted in 2013) for the period 2013-2017, Methodology for critical infrastructure selection (adopted in 2014) and annual reports on incident situations on Internet.

The key part of the legislation is the Law on Cyber Security (Official Gazette of Montenegro, nb. 14/10).

The analysis of the current situation in Montenegro shows that critical infrastructure is not officially defined. For the purpose of forming the National CIRT Team, Ministry for Information Society and Telecommunications, in cooperation with the Company IMPACT in Malaysia, made a review of the critical sectors in Montenegro based on the recommendations of international criteria. This recommendations recognize the energy sector as a national critical infrastructure.

The positive side of Cyber Security Strategy from 2013 is that it supports promotion and implementation of cyber security measures. However, emphasis is given to IT cyber security development, mainly realized through different government sectors. Insufficient attention is paid to the potential security of ICS systems with its proprietary protocols.

On the other hand, the Cyber Security Strategy represents a certain step forward when it comes to the application of the principles of sustainable development, in comparison to earlier situation.

Key priorities defined in the Cyber Security Strategy are to set the vision, scope, objectives and priorities, monitor the risk assessment at the national level, take into account existing policies, regulations and capacities, develop the clear management structure, identify and engage stakeholders, establish the mechanisms for the exchange of confidential information, develop of cyber-security contingency plans, organize the cyber security exercises, establish public-private partnerships establish the basic safety requirements, establish mechanisms for incidents reporting, increase public awareness about this issue, engage in international cooperation to harmonize national strategy on cyber security [6].

One of the problems is the lack of the necessary skills for a successful defense of incidents, as well as need for changes and amendments or enactment of new legislation, on the basis of which it would be possible to successfully detect and prosecute persons involved in all forms computer crime.

Also, it is noticed that any kind of examination of incidents on ICS has never be done since annual reports on incident situation does not treat it at all.

After all, development and planning of the national cyber crisis plan is an important factor in the overall planning of the state cyber security. It should be realistic and accurate but also it should take into account all possible participants. This involves the interaction of public and private sector.

An important part of this process is the identification and definition of critical national infrastructure, threats and risks connected to it.

#### IV. IMPLEMENTED SECURITY MEASURES AT HPP PERUĆICA

To limit the exposure to cyber attacks, networks and devices with SCADA protocols, are isolated from other networks (Figure 4). These components are not linked to the Internet. Access to the other networks is strictly regulated by limiting the input / output traffic on strictly necessary protocols.

network SCADA showed how well these links are protected. Following types of connections were identified and estimated:

- Internal LAN and WAN, including business networks
- Internet
- Wireless network equipment, including satellite uplinks
- Relationship with business partners, suppliers and regulatory agencies

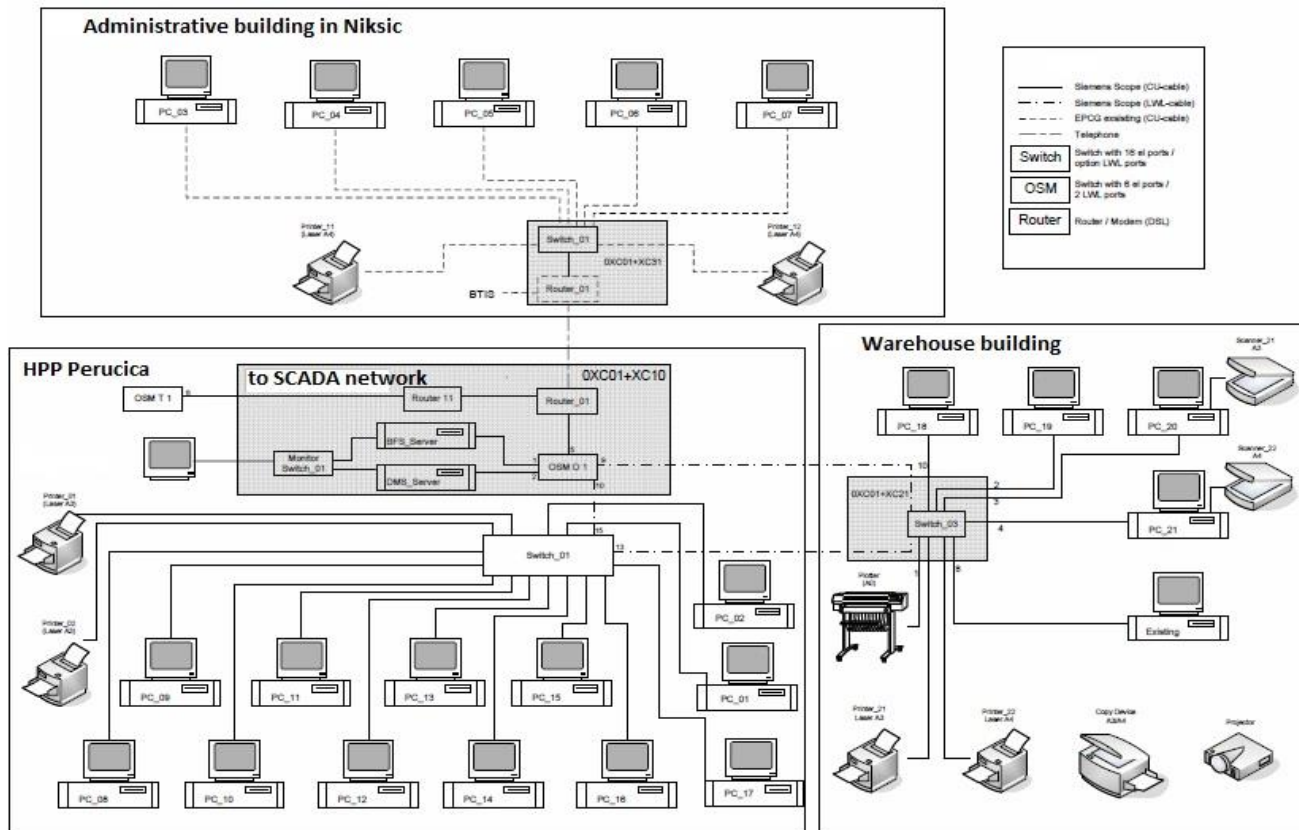


Figure 4. Connection of SCADA network from HPP Perucica with other networks in the company

The following steps represent the activities carried out in the implementation phase of the project, in order to establish an effective cyber security program:

##### A. Network protection strategy based on the principle of defense-in-depth

The basic principle, which is a part of a network protection strategy, is the principle of defence-in-depth. Defense-in-depth takes into account design stage of the development process, and it is an integral part of all technical considerations and decisions relating to the network. Single points of failure are eliminated, while the cyber defense is layered in order to limit and reduce the impact of any security incidents. In addition, the system is protected by the second system in the same layer [7].

##### B. Identification of all connections with the SCADA network

A thorough analysis of the risk assessment indicated the necessity of any connection to the SCADA network. A comprehensive understanding of all connections to the

##### C. Turning off all unnecessary connections to SCADA network

To ensure the highest level of SCADA system security, it is necessary to isolate the SCADA network from other networks to the greatest possible extent. Any connection to another network introduces security risks, especially if the relationship creates a route to the Internet. Although a direct link with other networks can provide an efficient exchange of information, insecure connections are simply not worth the risk.

##### D. Evaluation and strengthening of the security of all remaining connections to the network SCADA

In order to evaluate the level of protection, testing and vulnerability analysis was conducted on all remaining connections to the SCADA network. This information, together with the risks from the process, was used in developing a strong strategy for the protection of all connections to the SCADA network. As the SCADA network is secure as the its weakest connection point,

firewalls and other appropriate security measures have been implemented at every entry point.

*E. Strengthening SCADA network by removing or disabling unnecessary services*

SCADA servers that use commercial or open source operating systems can be exposed to various attacks through standard network services. Removing or turning off unused services and networks, like a demon, reduce the risk of direct attack which is especially important when the SCADA network is connected with other networks.

*F. Place strict controls over any medium that is used in SCADA network*

Modem, wireless, and wire networks used for communication and maintenance represent a significant vulnerability of SCADA networks. A successful attack could allow an attacker to bypass all other controls and to have direct access to SCADA resources. To reduce the risk of such attacks, such incoming access IS disabled and has replaced by the VPN system [8].

V. MEASURES THAT SHOULD BE IMPLEMENTED TO INCREASE SECURITY OF SCADA NETWORK

*A. Do not rely on proprietary protocols and factory settings security*

Some SCADA systems use a unique, proprietary protocols for communication between the device and the server from the field. Often the security of SCADA system based solely on the confidentiality of these protocols. Unfortunately, unclear protocols provide very little "real" security. For this reason, you should avoid proprietary protocols and factory settings, it is necessary need to check whether manufacturers present interfaces allow the safe functionality of the system.

*B. Implement security features provided by the suppliers of equipment and systems*

We should insist on implementation on the security features in the form of product upgrades. Factory security settings (such as the computer network firewalls) are often adjusted to ensure maximum usability and minimum security. All security functions should be to ensure maximum levels of safety.

*C. Conduct internal and external systems of protection against intrusion and establish a 24-hour monitoring system*

In order of effectively respond to cyber attacks, it is necessary to establish a strategy to protect against intrusion by notifying the system administrator about a malicious network activity originating from internal or external sources. Monitoring system for detection would be active 24 hours a day, and this feature, which can be easily set up, would send a notice about the attack by sms message or e-mail. In addition, in order to ensure an effective response to any attack, procedure for responding to incidents should be prepared. As a complement to network monitoring, in order to reveal the existence of suspicious activity, it should be possible to record all the daily logs on the

systems [8].

*D. Establish SCADA "red team" to identify and assess possible scenarios for an attack*

Establish a "red team" for identifying potential scenarios of attacks and potential vulnerabilities of the system. Use different profiles of people who can recognize the weaknesses of the whole network, SCADA systems, physical systems and security controls. People who regularly work on the system have a great insight into the vulnerability of SCADA network, and they should be consulted for identification of potential critical scenarios and the possible consequences.

*E. Clearly define cyber security roles, responsibilities and authority for managers, system administrators and users*

Staff needs to understand specific and define clear and logical roles and responsibilities. In addition, key personnel should be given sufficient authority in the execution of their duties. Establish cyber security organizational structure that defines roles and responsibilities and clearly identify how the cyber security issues escalated, and when notified in an emergency.

*F. Document the network architecture to identify systems that contain important functions or contain sensitive information that requires additional levels of protection*

As part of the process of establishing an effective protection strategy is necessary to develop and document a robust informational security architecture.

VI. CONCLUSION

The paper explains the importance of defining the ISC systems as a national critical infrastructure. Also, it shows the frequency, types and vulnerabilities of SCADA systems. Parer explains the existing security measures implemented in HPP Perucica, as well as measures that could be implemented for security improvement.

LITERATURA

- [1] *Critical Infrastructures and Services*, The European Union Agency for Network and Information Security, 2016 <https://www.enisa.europa.eu/topics/critical-information-infrastructures-and-services/scada?tab=details>
- [2] Keith S. ,Victoria P., Suzanne L., Marshall A. & Adam H.(May 2015). *Guide to Industrial Control Systems (ICS) Security*, U.S. Department of Commerce : National Institute of Standards and Technology Special Publication, (<http://dx.doi.org/10.6028/NIST.SP.800-82r2>)
- [3] Gleb Gritsai Alexander Timorin Yury Goltsev Roman Ilin Sergey Gordeychik Anton Karpin (2012). *Scada safety in numbers*, Positive technologies
- [4] Katherine Brocklehurst (2016). *ICS Vulnerabilities Trend Report: Missed Warnings, exposed industrial environments*. FireEye
- [5] *STUXNET Malware Targets SCADA Systems*. Trend Micro. Jan 2012.
- [6] *Strategy for Cyber Security in Montenegro*, Government of Montenegro, 2013, pp. 8-9.
- [7] *Defense in Depth – Multi-Layer Protection Approach*, Siemens. 2014.
- [8] *21 Steps to Improve Cyber Security of SCADA Networks*, U.S. Department of energy, 2012

# Testiranje upotrebljivosti grafičkog korisničkog interfejsa operativnih sistema za mobilne uređaje

Vladimir Bigović, Sandra Đurović, Snežana Šćepanović, *Fakultet za informacione tehnologije, Univerzitet "Mediteran" Podgorica*

**Sadržaj** — Upotrebljivost mobilnih uređaja i aplikacija predstavlja ključni faktor za uspjeh mobilnog računarstva. Prilikom razvoja i implementacije softvera za mobilne uređaje posebnu pažnju treba posvetiti dizajniranju grafičkog korisničkog interfejsa zbog velike raznolikosti mobilnih tehnologija i značajne različitosti krajnjih korisnika. U ovom radu prikazano je istraživanje upotrebljivosti grafičkog korisničkog interfejsa dva najpopularnija operativna sistema za mobilne uređaje – Android i iOS.

**Ključne riječi** — grafički korisnički interfejs, interakcija čovjek-računar, mobilni uređaji, mobilni operativni sistemi, upotrebljivost.

## I. UVOD

**P**ILIKOM dizajniranja upotrebljivog korisničkog interfejsa jedan od glavnih imperativa je izgradnja efikasnih modela interakcije čovjek-računar [1]. Upotrebljivost predstavlja ocjenu kvaliteta, koja se odnosi na lakoću korišćenja nekog korisničkog interfejsa. Postoji pet komponenti koje definišu upotrebljivost [2]:

- *Lakoća učenja* — Da li je običnom korisniku lako da uspješno obavlja osnovne zadatke, kada se prvi put susreće sa grafičkim korisničkim interfejsom uređaja?
- *Efikasnost* — Kada se korisnici upoznaju sa grafičkim korisničkim interfejsom uređaja, koliko brzo mogu obavljati zadatke?
- *Memorabilnost* — Kada se korisnici nakon perioda nekorišćenja opet susretu sa grafičkim korisničkim interfejsom uređaja, koliko lako ponovo mogu uspostaviti stručnost?
- *Greške* — Koliko grešaka prave korisnici, koliko su kritične te greške, i da li ih korisnici lako ispravljaju?
- *Zadovoljstvo* — Da li je grafički korisnički interfejs uređaja prijatan za korišćenje?

Vrijednovanje upotrebljivosti predstavlja skup metoda "mjerenja" čiji rezultati odgovaraju na pitanje koliko dobro ljudi mogu koristiti neki objekat u svrhu za koju je taj objekat namijenjen. Postoje različite metode vrijednovanja upotrebljivosti [3], među kojima se kao najkorisnije pokazalo testiranje iskustva korisnika, koje se sastoji od tri komponente:

- Pronaći ciljnu grupu odnosno reprezentativne korisnike.
- Korisnicima postaviti reprezentativne zadatke u odnosu na dizajn.
- Posmatrati akcije korisnika, da li uspješno izvršavaju zadatke ili imaju poteškoća.

Istraživanje obuhvata dva zasebna testa: test upotrebljivosti sa test subjektima i test kognitivnog opterećenja. Navedeni testovi su izabrani jer njihovim sprovođenjem rezultate dobijamo analizirajući iskustvo korisnika na licu mjesta, odnosno koristeći "think-aloud" protokol. Nakon testiranja, kombinovanjem metode "think-aloud" i mjerenjem vremena potrebnog za izvršenje zadatka, došli smo do empirijskih mjerljivih rezultata. Uređaji korišćeni u testiranju su Samsung Galaxy S4 (Android "4.4 KitKat" verzija OS-a) i Apple iPhone 5 ("iOS 7" verzija OS-a). U trenutku sprovođenja istraživanja, Android 4.4 KitKat imao je najveći udio među aktivnim uređajima zasnovanim na Android platformi (35.5%), zbog čega je istraživanje sprovedeno koristeći tu, i ako stariju, vrlo popularnu verziju ovog operativnog sistema [4]. Da bi mjerenje bilo korektno, izabran je Apple-ov uređaj sa operativnim sistemom iz istog vremenskog razdoblja-iOS 7. Android je (kao i iOS) od verzije korišćene u testiranju imao više iteracija-verzija, od kojih je najskorija "7.0 Nougat", ali je rasprostranjenost iste toliko mala (~1.0%) da bi njeno korišćenje u istraživanju upotrebljivosti bilo irelevantno zbog nedostupnosti prosječnom korisniku [4].

## II. TEST UPOTREBLJIVOSTI SA TEST SUBJEKTIMA

Prije početka testiranja, test subjekti ispunjavaju upitnik na osnovu kojeg se kategorizuju rezultati. U testu učestvuje 10 test subjekata različitih godina, pola, nivoa obrazovanja i nivoa tehnološkog iskustva. Testove kontrolirše moderator, evidentirajući vrijeme i osiguravajući praćenje instrukcija uz snimanje video kamerom u cilju naknadne evaluacije istraživanja. Sadržaj

Vladimir Bigović, Fakultet za Informacione Tehnologije u Podgorici, Univerzitet Mediteran, Josipa Broza bb, 81000 Podgorica, Crna Gora (telefon: 382-67-747579; e-mail:vladimir.bigovic@gmail.com).

Sandra Đurović, Fakultet za Informacione Tehnologije u Podgorici, Univerzitet Mediteran, Josipa Broza bb, 81000 Podgorica, Crna Gora (telefon: 382-67-747569; e-mail:djurovic.sandra1@gmail.com).

Snežana Šćepanović, Fakultet za Informacione Tehnologije u Podgorici, Univerzitet Mediteran, Josipa Broza bb, 81000 Podgorica, Crna Gora (telefon: 382-67-235915; e-mail:snezana.scepanovic@unimediteran.net)

zadataka na testu kao i vremensko ograničenje za izvršavanje istih su predloženi od strane autora. Korisnicima je postavljeno 10 istovjetnih zadataka na svakoj platformi posebno, počevši sa predodređenog startnog ekrana mjeri se nezavisno i evidentira vrijeme koje je potrebno korisniku da ispuni zadatak. Mjerenje se vrši do 8 minuta, poslije čega se konstatuje neuspješnost testa. Korisnik ne smije koristiti fizičke tastere na uređaju tj. interakcija se svodi samo na ekran (korisnički interfejs operativnog sistema). Jezik prikaza na uređajima je postavljen na maternji jezik test subjekta. Zadaci koriste samo fabričke mogućnosti uređaja bez dodatnih aplikacija i izmjena. Nakon završetka individualnog zadatka moderator uređaj vraća u početno stanje za slijedeći zadatak i opet započinje proces.

U testu upotrebljivosti sa test subjektima se pomoću različitih zadataka mjeri lakoća korišćenja nekog grafičkog korisničkog interfejsa. Na testu su postavljeni slijedeći zadaci.

Počevši sa otključanog glavnog ekrana:

1. Otvoriti web stranicu <http://www.google.com> i uraditi pretragu na riječ “ipad air”;
2. Otkucati SMS poruku sa tekstem “you’re still weak” i poslati je na probni broj;
3. Napraviti fotografiju prostora ispred sebe;
4. Promijeniti ton obavještenja o pozivu na prvo u listi;
5. Promijeniti pozadinsko osvjetljenje ekrana uređaja na minimalno ponuđeno;
6. Počevši sa zaključanog ekrana uređaja, otključati telefon i vratiti na glavni ekran;
7. Počevši sa ekrana standardnog web browser-a uređaja, vratiti se na glavni ekran;
8. Počevši sa otključanog glavnog ekrana, promijeniti pozadinsku sliku ekrana (wallpaper);
9. Počevši sa zaključanog ekrana uređaja, podesiti ugrađeni budilnik uređaja na 5:45h;
10. Počevši sa otključanog glavnog ekrana, podesiti jezik prikaza uređaja na engleski jezik.

U zadacima 6. i 9. test subjekti ne koriste fizičko dugme za otključavanje već opciju “swipe” (fizičko dugme je već bilo pritisnuto). Zadatak 7. nije moguće uraditi bez korišćenja fizičkog dugmeta.

#### A. Test Android OS-a

Testiranje je izvršeno na Samsung-ovom Galaxy S4 uređaju koji sadrži značajno izmijenjenu verziju izvornog Androida, stoga rezultati nijesu konzistentni za ostale uređaje istog OS-a i može biti značajnih odstupanja. U tabeli 1. za svakog test subjekta (S) predstavljeno je vrijeme (mm:ss) potrebno za izradu zadatka (Z).

TABELA 1: REZULTATI TESTA ZA ANDROID OS

	Z.1	Z.2	Z.3	Z.4	Z.5	Z.6	Z.7	Z.8	Z.9	Z.10
S.1	00:16	00:32	00:05	00:26	00:05	00:01	NE	00:50	00:35	01:01
S.2	03:30	01:30	01:00	NE	00:30	00:10	NE	03:00	02:30	NE
S.3	02:00	00:20	00:10	01:05	00:12	00:05	NE	01:00	01:11	04:00
S.4	00:50	00:40	00:11	06:50	00:24	00:05	NE	00:50	00:43	02:03
S.5	03:00	01:01	00:32	NE	00:39	00:10	NE	03:07	02:47	NE
S.6	00:53	00:41	00:10	03:03	00:30	00:14	NE	01:04	01:02	05:00
S.7	01:00	00:45	00:21	06:02	00:27	00:10	NE	03:07	02:00	NE
S.8	01:33	00:50	00:32	04:04	00:55	00:31	NE	04:00	02:39	NE
S.9	00:55	00:43	00:15	NE	00:20	00:10	NE	01:40	00:59	04:01
S.10	00:37	00:20	00:25	02:55	00:19	00:15	NE	02:02	04:00	06:08

Nakon testiranja, analizirajući komentare test subjekata dolazimo do slijedećih opažanja:

- *Nemogućnost razlikovanja aplikacija i vidžeta* — Android kao OS sadrži i aplikacije i vidžete, podjelu naslijeđenu još od desktop operativnih sistema bez jasnog mjesta u mobilnom okruženju. Jednostavnije aplikacije izgledaju kao vidžeti pa čak i iskusnim korisnicima što doprinosi konfuziji i značajno utiče na upotrebljivost.
- *Nedostatak vođenja/uputa korisnika* — kod nekih akcija neće doći do obavještenja da je akcija uspjela/neuspjela. Korisnik mora da nagađa načine da obavi određene akcije. Oslanjanje na pokušaj/grešku kao iskustvo učenja u OS-u je generalno loš dizajn interfejsa.
- *Loša raspoređenost sadržaja* — korisnici nijesu u mogućnosti da pronađu opcije i određena podešavanja jer su raspoređena u tab-ovima koje subjekti ne umiju da aktiviraju. Opcije nijesu postavljene po logičnim kategorijama već često čini se beskonačno jedna unutar druge.
- *Natrpnanost i konfuzija glavnog ekrana* — mnoštvo ikonica i vidžeta neuniformnih veličina, kao i različitih “fabričkih” ekrana zbunjuju korisnika i stvaraju osjećaj “natrpnanosti” ekrana gdje se mnoštvo elemenata bori za pažnju korisnika.
- *Duplikati ikonica i aplikacija* — više ikonica iste aplikacije, kao i više različitih aplikacija koje obavljaju isti zadatak (primjer-postojanje “mail” i “gmail” aplikacije) zbunjuju korisnika, smanjuju efikasnost multipliciranjem notifikacija i usporavaju operativni sistem.
- *Brisanje ikonica ili aplikacija je neintuitivno* — da bi obrisali ikonicu, potrebno je držati je pritisnutom, pa je prenijetu u “kantu” za otpatke koja je smještena u gornjem desnom uglu. S obzirom na veličinu ekrana ovog uređaja samo prevlačenje ikonice predstavlja problem, kao i sama činjenica da neiskusni korisnici nijesu sigurni da li brišu prečicu na glavnom ekranu ili cijelu aplikaciju.
- *Animacije su nelogične i neefikasne* — animacije se u dobrom dizajnu koriste kao pomoć pri objašnjavanju korisniku šta se dešava sa aplikacijom/uređajem oponašajući neke realne pokrete (primjer-vizuelni “ulazak” u novi ekran predstavlja otvaranje aplikacije). Android koristi animacije samo kao “ukrase” sa dodatkom potencijalnog usporavanja uređaja.
- *Nekonzistentan dizajn* — korišćenjem istih principa dizajna kroz cijeli OS postiže se bolje upoznavanje korisnika sa sistemom i svojevrsno vođenje korisnika, na osnovu simbola, boja i ostalih elemenata koji su istovjetni širom sistema. Ovaj operativni sistem sastoji se od mnoštva različitih stilova, boja, sijenki što može da “plaši” novog korisnika.
- *Povremena usporavanja i neoptimizovanost* — usporavanja i “kočenje” sistema na uređaju sa prilično moćnim hardverom značajno smanjuje efikasnost i upotrebljivost korisnika nezavisno od njegovog

tehnološkog iskustva.

- *Prviše podešavanja i mogućnosti* — obim različitih mogućnosti i podešavanja može da bude “zastrašujući” za neiskusnog korisnika.
- *Neintuitivni simboli* — simboli na tasterima koji obavljaju neke funkcije kao i izgled samog tastera bi trebalo da daju korisniku do znanja da je to “prekidač” za neku akciju a ne samo nepotrebni ukras.
- *Više mogućih putanja do istih rezultata* — iste akcije je moguće uraditi na veći broj načina što zbunjuje subjekte i ometa uspostavljanje niza logičkih pravila za upravljanje i snalaženje u OS-u kod korisnika.

### B. Test iOS-a

Testiranje je izvršeno na Apple iPhone 5 uređaju. Obzirom da se operativni sistem nalazi u skoro nepromijenjnom obliku samo na Apple-ovim uređajima, rezultati se mogu smatrati reprezentativnim za sve ostale uređaje ove platforme, sa mogućim manjim odstupanjima. U tabeli 2. za svakog test subjekta (S) predstavljeno je vrijeme (mm:ss) potrebno za izradu zadatka (Z).

TABELA 2: REZULTATI TESTA ZA IOS

	Z.1	Z.2	Z.3	Z.4	Z.5	Z.6	Z.7	Z.8	Z.9	Z.10
S.1	00:11	00:31	00:03	00:22	00:04	00:01	NE	00:56	00:31	00:45
S.2	03:10	01:11	01:05	NE	00:31	00:09	NE	02:20	02:20	NE
S.3	02:05	00:20	00:08	01:15	00:16	00:04	NE	01:00	01:17	03:25
S.4	00:50	00:42	00:13	00:27	00:21	00:05	NE	00:49	00:26	02:34
S.5	02:30	00:59	00:40	NE	00:47	00:04	NE	03:24	02:30	07:07
S.6	00:56	00:48	00:15	03:00	00:31	00:16	NE	01:14	01:02	03:40
S.7	01:01	00:37	00:33	NE	00:25	00:12	NE	04:07	02:11	06:59
S.8	01:20	00:50	00:40	03:34	00:58	00:29	NE	04:05	02:40	07:40
S.9	00:50	00:54	00:19	04:02	00:21	00:10	NE	01:40	00:45	04:11
S.10	00:30	00:21	00:20	02:30	00:19	00:12	NE	02:17	04:10	NE

Nakon testiranja, analizirajući komentare test subjekata dolazimo do slijedećih opažanja:

- *Nenamjerni pristup funkcijama* — određeni elementi mogu biti slučajno aktivirani tokom interakcije sa drugim elementima korisničkog interfejsa (primjer-centar za notifikacije kojem se pristupa prevlačenjem prsta sa vrha ekrana ka dolje). Uzrok tome je nedovoljna ispoliranost određenih funkcija koje su zbog pritiska tržišta morale biti usvojene za kratko vrijeme.

- *Nepoznat dizajn ikonica* — u cilju modernizovanja grafičkog korisničkog interfejsa i odstupanja od oponašanja realnih predmeta u simbolici ikonica i drugih elemenata (praksa veoma prisutna u prethodnim verzijama istog OS-a) mnogi elementi su teško razumljivi tehnološki iskusnijim korisnicima, a potpuno nepoznati novim korisnicima. Pojednostavljanje elemenata je dovelo do smanjenja intuitivnosti.

- *Nejasnost grafičkog korisničkog interfejsa* — elementi koji odrađuju neke akcije i dekorativni elementi se u većini slučajeva ne razlikuju. Novi korisnici treba da nagađaju i na “slijepo” pogađaju interaktivne elemente.

- *Preveliko pojednostavljanje sistema* — i ako svjestan izbor pri dizajnu sistema, ova činjenica može veoma tehnološki obrazovane korisnike da odbije od platforme.

- *Nedovoljna mogućnost prilagođavanja* — takođe Apple-ov svjestan i namjeran izbor u dizajnu sistema u cilju održavanja homogenosti i jedinstvenog izgleda na svim uređajima. Bilo kakav proces dodatnog

prilagođavanja korisničkim potrebama zahtjeva napredna znanja i donekle rizične tehnike “razbijanja” virtuelnih ograničenja koja je postavio Apple (“jailbreaking”).

- *Nedostatak mogućnosti* — i ako nedostatak funkcija nekad može da ima pozitivan efekat na upotrebljivost mobilnog uređaja, u vremenu kada svaki proizvođač pokušava da poveća “vrijednost” svog uređaja dodavajući različite korisne, i ne toliko korisne mogućnosti, nedostatak funkcija se može smatrati nedostatkom u upotrebljivosti uređaja. Oslanjajući se na ogromnu količinu aplikacija dostupnih u svom katalogu, iOS 7 odstranjuje funkcije dostupne “iz kutije” na uređajima konkurentnih mobilnih operativnih sistema.

- *Povremene greške u fabričkim aplikacijama* — u cilju razbijanja ekonomske “simbioze” sa Google-om kao sada već ozbiljnim konkurentom na tržištu mobilnih uređaja (Android), Apple je Google-ove zrele i moćne usluge/aplikacije koje su se nekada nalazile na fabričkim verzijama ovog OS-a zamijenio svojim nedovršenim i značajno sporijim verzijama (primjer-Apple-ov servis “Maps” je zamijenio “Google maps”). Korisnici uvijek žele najkompletnije, najbolje i najbrže usluge na tržištu što ih može odvratiti od iOS platforme.

### III. TEST KOGNITIVNOG OPTEREĆENJA

Prema definiciji kognitivno opterećenje predstavlja zbir mentalnih resursa koji su potrebni za upravljanje sistemom [5]. Indeks kognitivnog opterećenja je kvantifikovani način “mjerenja” kognitivnog opterećenja i predstavlja broj elemenata sa kojim korisnik prethodno treba da bude upoznat da bi uređaj koristio spontano i intuitivno [6].

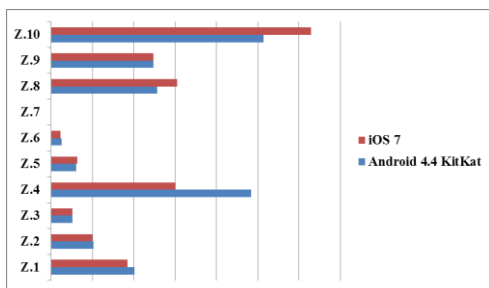
Testiranje smo izvršili samostalno, brojeći aplikacije/vidžete kao i druge ikonice i elemente interfejsa koje sadrži fabrička (početna) verzija datog operativnog sistema. Obzirom na to da Samsung Galaxy S4 uređaj sadrži značajno izmijenjenu verziju izvornog Androida, rezultati nijesu konzistentni za ostale uređaje istog OS-a, i može biti značajnih odstupanja. Ovaj uređaj u fabričkoj verziji imao je kognitivni indeks od 162, tačnije: 104 aplikacije i vidžeta, uz 58 ikona i ostalih elemenata interfejsa.

Obzirom da se iOS 7 nalazi u skoro nepromijenjnom obliku samo na Apple-ovim uređajima, rezultati se mogu smatrati reprezentativnim za sve ostale uređaje ove platforme, sa mogućim manjim odstupanjima. Test je izvršen na Apple-ovom iPhone 5 uređaju. Ovaj uređaj u fabričkoj verziji imao je kognitivni indeks od 40.

### IV. POREĐENJE I ANALIZA REZULTATA

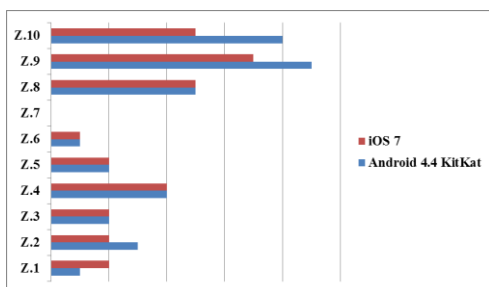
Analiza dobijenih rezultata testa upotrebljivosti sa test subjektima pokazuje da su iOS 7 i Android veoma ujednačeni, uz malu prednost Apple-ovog iOS 7 u broju dodira potrebnih za postizanje željenih akcija.

Na osnovu poređenja mjerenja prosječnog vremena potrebnog za izvršavanje zadatka (Sl. 1.) i minimalnog broja dodira potrebnih za izvršavanje zadatka (Sl. 2.) možemo zaključiti da su subjekti imali manje poteškoća u izvršavanju zadataka na Apple-ovom uređaju.



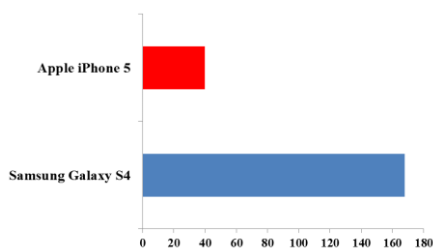
Sl. 1. Poređenje vremena potrebnog za izvršavanje zadatka.

Minimalan broj dodira potrebnih za izvršavanje svih zadataka na Samsung-ovom uređaju je 35, dok je na Apple-ovom uređaju minimalan broj dodira 31.



Sl. 2. Poređenje minimalnog broja dodira potrebnog za izvršavanje zadatka.

Rezultati testiranja kognitivnog opterećenja (Sl. 3.) pokazuju da iOS 7 ima četiri puta bolje rezultate.



Sl. 3. Poređenje indeksa kognitivnog opterećenja.

Prilikom analize neophodno je uzeti u obzir različite načine pristupa Samsung-a i Apple-a u izradi grafičkog korisničkog interfejsa mobilnog operativnog sistema. Apple bira zatvoreniji pristup sa fiksno određenim interfejsom koji garantuje jednostavnost ali se ne može prilagođavati (nudi se samo promjena redoslijeda ikonica aplikacija i pozadine ekrana). U grafičkom korisničkom interfejsu iOS-a koriste se jednostavni oblici i tzv. "ravni" dizajn lišen suvišnih odsjaja, sjenki i efekata, uz minimalistički interfejs u cilju omogućavanja razumljivosti i upotrebljivosti za svakog korisnika nezavisno od nivoa znanja i vještina. Samsung ima drugačiji pristup, i omogućava veliki broj opcija za prilagođavanje grafičkog korisničkog interfejsa mobilnog operativnog sistema svakom korisniku. Ovakav pristup odgovara iskusnijim korisnicima ali može zbunjivati nove korisnike.

Analizom rezultata testiranja možemo zaključiti da je Apple i dalje lider na tržištu kad se govori o upotrebljivosti uređaja. iOS 7 obezbeđuje lakoću korišćenja, prefinjen i moderan dizajn korisničkog interfejsa za širok krug korisnika i veliki broj aplikacija za biznis korisnike.

Samsung grafički interfejs je težak za snalaženje neiskusnim korisnicima. Nedostaci u upotrebljivosti nadomještaju se mnoštvom softverskih mogućnosti i bezbjednosnim karakteristikama koje obezbeđuje KNOX softver [7], fabrički prisutan na svim novim Samsung uređajima. Na taj način Samsung pokušava privući prvenstveno biznis korisnike. Mnogo mogućnosti za prilagođavanje grafičkog korisničkog interfejsa (za svakoga po nešto) i dalje omogućavaju da Samsung bude jedan od lidera na tržištu mobilnih uređaja [8].

## V. ZAKLJUČAK

Područje upotrebljivosti mobilnih uređaja se intenzivno razvija i ima za cilj da osigura kvalitetne mobilne usluge po mjeri korisnika. Rezultati istraživanja i komercijalna praksa pokazuju da je razvoj mobilnih operativnih sistema ujednačeniji nego ikad. Android i iOS su izrasli u zrele operativne sisteme koji mogu da pariraju čak i operativnim sistemima desktop računara. Generalno upotrebljivost mobilnih uređaja još nije postigla zadovoljavajući nivo i slijede dalja istraživanja, identifikacija problema i kreiranje smjernica za unaprijeđenje. U tom kontekstu razvojni inženjeri, eksperti upotrebljivosti i krajnji korisnici moraju saradivati i učestvovati u iterativnim razvojnim procesima softvera za mobilne uređaje.

## LITERATURA

- [1] J. Nielsen, "Designing web usability," New York, NY, United States: Turtleback Books, 1999.
- [2] S. Krug and R. Black, "Don't make me think!: A common sense approach to web usability," 4th ed. Indianapolis, IN: New Riders Publishing, 2005.
- [3] J. Nielsen and R. Budiu, "Mobile usability," Berkeley, CA: New Riders Publishing, 2012.
- [4] Android developers. *Dashboards*. [Online]. Available: <https://developer.android.com/about/dashboards/index.html>. Accessed: Feb. 9, 2017.
- [5] K. Whifton, (1998). *Minimize cognitive load to maximize usability* / Nielsen Norman Group. [Online]. Available: <https://www.nngroup.com/articles/minimize-cognitive-load>. Accessed: Feb. 9, 2017.
- [6] D. A. Menasce, "Load testing of web sites," IEEE Internet Computing, vol. 6, no. 4, pp. 70-74, Jul. 2002.
- [7] SAMSUNG, (2016). *Knox*. [Online]. Available: <https://www.samsungknox.com/en>. Accessed: Feb. 9, 2017.
- [8] I. G. Clifton, "Android user interface design: Turning ideas and sketches into beautifully designed apps," Harlow: Addison-Wesley Educational Publishers, 2013.

## ABSTRACT

Usability of mobile devices and applications is a key factor for the success of mobile computing. During the development and implementation of software for mobile devices, special attention should be paid to the design of graphical user interfaces due to the large diversity of mobile technology and the significant diversity of end-users. This paper presents a study of usability testing of graphical user interface of two most popular mobile operating systems-Android and iOS.

## USABILITY TESTING OF GRAPHICAL USER INTERFACE FOR MOBILE OPERATING SYSTEMS

Vladimir Bigović, Sandra Đurović, Snežana Šćepanović

# O nekim ICT rješenjima za povećanje bezbjednosti na radu u invazivnim okruženjima

Prof. dr Sanja Bauk, Doc. dr Tatijana Dlačić, Pomorski fakultet Kotor, Univerzitet Crne Gore

**Sadržaj** — U radu je dat prikaz nekih komercijalnih ICT rješenja za povećanje bezbjednosti radnika u invazivnim okruženjima (industriji čelika, rudarstvu, građevinarstvu, naftnim platformama, na brodovima i sl.). Ukazano je na neophodnost povećanja bezbjedonosnih mjera, prije svega zbog potrebe zaštite ljudskih života i zdravlja, a potom i zbog povećanja produktivnosti, smanjenja troškova kompanija i osiguravajućih društava. Takođe, ukratko su opisana neka alternativna rješenja, ispitivana na nivou simulacija, u skladu sa realnim potrebama i mogućnostima Luke Bar, koja decenijama posluje u tranzicionim uslovima i suočava se sa problemom obezbjeđivanja sofisticiranih i obično skupih ICT rješenja.

**Ključne reči** — bezbjednost radnika, ICT, invazivna okruženja, Luka Bar.

## I. UVOD

PREMA podacima Međunarodne organizacije rada (ILO-International Labour Organization, eng.) svakih 15 sekundi, 151 radnik u svijetu doživi nezgodu na radu. Na godišnjem nivou, u globalnim razmjerama, broj povreda koje nemaju fatalan ishod je 317 miliona. Ono što zabrinjava još više je činjenica da 321 000 ljudi umre u svijetu svake godine usljed povreda na radu [1]. Radi se o velikom problemu u mnogim industrijama, bez obzira na bezbjedonosne propise i procedure, tako da je nužno smanjenje broja i intenziteta ovih povreda. U tom pravcu, u radu je prvo dat kratak pregled nekih savremenih, sofisticiranih ICT rješenja koja se koriste za ove svrhe. Potom je dat kratak opis nekih alternativnih mogućnosti koje su ispitivane u kontekstu individualnih potreba i preferencija Luke Bar, kao radno invazivnog okruženja koje funkcioniše u tranzicionim uslovima. Korišćeni su sekundarni i primarni literaturni izvori, respektivno.

## II. EW&SS: PRIKAZ 1

U IBM Haifa (Izrael) istraživačkom timu radi se na novoj vrsti platforme za upravljanje bezbjednošću radnika. IBM "Employee Wellness and Safety Solution" (EW&SS) je novo rješenje za praćenje vitalnih parametara radnika i njihove bezbjednosti. Ovaj sistem sakuplja podatke sa gotovo svih vrsta nosivih senzora, kao i senzora za snimanje ključnih parametara okruženja. Na osnovu

prikupljenih i obrađenih podataka, sistem upozorava radnike i osoblje zaduženo za bezbjednost u slučaju neposredne opasnosti. Fokus je na svakom pojedinačnom radniku, s ciljem smanjenja broja nezgoda, odnosno, akcidentnih situacija. Rješenje koje nudi IBM, je neka vrsta tehnološkog *andela čuvara* radnika koji rade u veoma opasnim uslovima. Nova nosiva platforma radi kao svojevrsan sistem upozorenja u realnom vremenu. Ona analizira velike količine podataka koje se prikupljaju sa senzora, a koji su ugrađeni u zaštitnu opremu radnika, tj. *pametne* zaštitne šlemove, zaštitne prsluke, kao i u *pametne* telefone radnika. Senzori u kontinuitetu snimaju puls radnika, kretnje, tjelesnu temperaturu, nivo hidratacije, kao i faktore okruženja: temperaturu, nivo buke, prisustvo toksičnih gasova i dr. [2].

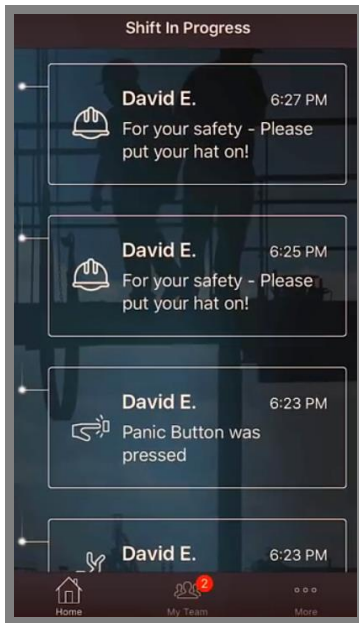
EW&SS otkriva i obavještava o svakoj nepravilnosti i potencijalnom riziku radnika ili nadzornika (Slika 1) u realnom vremenu. Sistem može da upozori na neku određenu opasnost na radnom mjestu, ili da pošalje alarm u slučaju da se desila povreda nekog od zaposlenih. Ovakve mjere bezbjednosti na radu mogu aktivno da spriječe nastanak akcidentne situacije. Npr., sistem daje upozorenje kada je radnik na radnom mjestu bez obaveznog bezbjedonosnog šlema, zaštitnih naočara ili slušalica. Sistem takođe može da prikaže na displeju upozorenje ako je radnik isuviše blizu nekoj mašini ili je na terenu koji je klizav. Takođe, upozorenje se uključuje kada je radnik duže od predviđenog vremena izložen štetnim-toksičnim materijama ili buci; kada je njegov organizam dostigao kritičnu tačku dehidracije, pa čak i u situacijama kada nije dovoljno koncentrisan ili kada je premoren. Platforma takođe identifikuje slučajeve hitnosti, tj. slučajeve pada radnika ili onesvješćenja, kada se izvještaj o tome šalje direktno zaduženima za pružanje pomoći i/ili najbližem radniku koji može takođe da pruži pomoć. Informacije sa senzora i *pametne* zaštitne opreme se direktno pohranjuju u *pametne* telefone zaposlenih, koji mogu da procesuiraju i analiziraju podatke. Neke od prikupljenih informacija se čuvaju u kladu, za potrebe daljih analiza, za studije i poboljšanje bezbjedonosnih propisa i procedura, kao i za njihovo prilagođavanje potrebama određenog zaposlenog.

Nema sumnje, čak i one povrede na radu koje namaju fatalan ishod, imaju izrazito negativan uticaj na radnike, odnosno, njihovo opšte zdravstveno stanje. Ne samo da radnici stradaju, već povrede uzrokuju i velike štete po pitanju produktivnosti, te zbog toga kompanije imaju velike finansijske gubitke. Samo na nivou SAD ovi gubici

Sanja Bauk, Tatijana Dlačić, Pomorski fakultet Kotor, Univerzitet Crne Gore, Dobrota 36, 85330 Kotor, Crna Gora (telefon: 382-32-303184; e-mail: [bsanjaster@gmail.com](mailto:bsanjaster@gmail.com); [tanjav@ac.me](mailto:tanjav@ac.me)).



su oko 220 biliona dolara godišnje, što odgovara gubitku od oko 27 miliona radnih sati godišnje [1]. Ovo naravno pogađa i osiguravajuća društva.



Sl. 1. EW&SS na strani nadzornika: Praćenje statusa senzora na zaštitnoj opremi radnika (izvor: [3]).

EW&SS je istraživački prototip koji se trenutno testira u vodećim SAD kompanijama.

### III. EW&SS: PRIKAZ 2

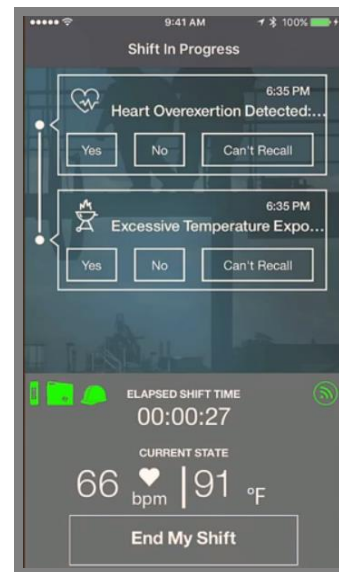
Napredak IoT (Internet of Things, eng. – Internet stvari) je omogućio konektovanje radnika. Radnik koji je svjesniji situacije, zahvaljujući sensorima koji kontrolišu njegove vitalne parametre, kao i parametre okruženja, utiče na određeni način na to da je i njegovo radno okruženje bezbjednije. Nosivi i ugrađeni senzori (u zaštitnu opremu) omogućuju praćenje radnika na radnom mjestu s ciljem sprječavanja pada, radnog preopterećenja, obezbjeđivanja propisnog odstojanja od teške mehanizacije i sl. Nosivi senzori se koriste u IoT-u za prikupljanje, integrisanje i analiziranje senzorskih podataka. Udruženi sa inovativnim kognitivnim mogućnostima i eksternim izvorima, kao što su oni vezani za stanje u okruženju, vremenske prilike i sl., stvaraju veliki potencijal za bolje upravljanje zdravljem i bezbjednošću radnika u pravcu suštinske transformacije savremenih načina života i rada.

Većina povreda na radu se može spriječiti nadgledanjem opšteg fizičkog stanja radnika. Povrede se mogu spriječiti, bilo da se kontrolišu da li se zaštitna oprema propisno koristi, bilo da se kontrolišu vremenska i lokacijska ograničenja u opasnim situacijama-zonama. Zahvaljujući integraciji i prezentaciji kontekstualnih informacija radnicima, oni postaju obaviješteniji o situaciji u neposrednom okruženju. Pri tome se koriste neintruzivni, uvijek aktivni, *svjesni* stanja okoline, senzorski uređaji koji prosljeđuju ključne (kritične) informacije na pravo mjesto i u pravo vrijeme. Iako na tržištu već neko vrijeme postoje fitness narukvice, koje nude individualna rješenja, tj. obavještavaju korisnika o tome koliki mu je puls, koliko je

utrošio kalorija i sl., ovaj IBM istraživački projekat nudi platformu koja koristi prednosti kongitivnog računarstva za analizu informacija prikupljenih sa brojnih senzora. Sakupljanjem i analiziranjem informacija na ovaj način, moguće je otkriti što se zaista dešava u radnom okruženju, tako da se može identifikovati indikativna kombinacija opasnosti, koje bi se u slučaju pojedinačnog posmatranja mogle previdjeti. Na primjer, kombinacija povećane tjelesne temperature radnika, povećanog pulsa i odsustvo njegovih uobičajenih kretnji tokom nekoliko minuta, ukazuju na to da je radnik najvjerovatnije doživio toplotni stres. Bilo koji od ovih pokazatelja, posmatran zasebno, neće aktivirati alarm, ali ako se uzmu u obzir integralno, sistem će ukazati na ozbiljnu situaciju koja zahtijeva intervenciju.

EW&SS platforma je razvijena za potrebe *Northstar Bluescope Steel* korporacije za proizvodnju čelika, sa firmama u Australiji, na Novom Zelandu i u Sjevernoj Americi. Riječ je o kongitivnoj platformi koja koristi *IBM Watson IoT* tehnologiju za nosive senzore, koji omogućuju radnicima da ostanu bezbjedni u opasnim uslovima.

U borbi protiv toplotnog stresa, koji je najčešća negativna pojava na radu u neposrednoj proizvodnji čelika, EW&SS sakuplja podatke sa različitih senzora koji stalno snimaju temperaturu kože radnika, rad srca, galvanski odziv kože, nivo aktivnosti, upoređujući prikupljene informacije sa onima o ambijentalnoj temperaturi i vlažnosti vazduha (Slika 2).

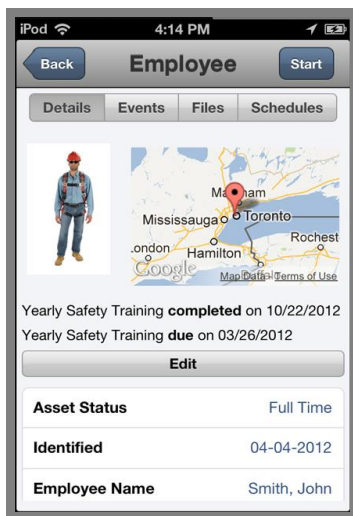


Sl. 2. EW&SS na strani radnika: Praćenje pulsa radnika i temperature u čeličani (izvor: [3]).

EW&SS nudi individualna rješenja za svakog od radnika, dajući im pojedinačno savjete, npr., da naprave pauzu od 10-15 minuta u hladu, uzmu tečnosti i sl., ukoliko njihovi fiziološki parametri prekorače uobičajene vrijednosti. Takođe, mogu se koristiti i senzori za određivanje nivoa radijacije, buke, toksičnih gasova i dr. Prisustvo štetnih gasova se može otkriti korišćenjem senzora ugrađenih u zaštitnu opremu radnika, a koji komuniciraju sa baznim stanicama putem Wi-Fi ili Bluetooth tehnologije [2].

## IV. NEKA DRUGA KOMERCIJALNA RJEŠENJA

*MasterLoch Field ID* softver, između ostalog, omogućuje praćenje zaštine opreme radnika (PPE-Personal Protective Equipment, eng.) u smislu njegovog korišćenja i funkcionalnosti [4]. Field ID app za Google Android, Apple iOS ili Windows Mobile može da snimi poziciju PPE u realnom vremenu i obavijesti nadzornika o stanju PPE. Sistem je usklađen u realnom vremenu sa RFID i GPS uređajima, na bilo kom mjestu i u svakom trenutku u klauđu. Kontrolom bezbjedonosnih pojaseva radnika, sistem obezbjeđuje proaktivnu zaštitu od pada na radnom mjestu (Slika 3). Ukoliko bi sistem imao WISP (Wireless Identification and Sensing Platform, eng.) mogućnosti, ID-u i funkcionalnosti zaštinog pojasa, mogao bi se pridružiti i senzor za detekciju i alarmiranje u slučaju pada radnika (ovi senzori rade na principu uparenih akcelerometara).



Sl. 3. Field ID app za potrebe radnika (izvor: [4])

*3M Science Applied to Life™* koristi RFID tehnologiju za označavanje svakog pojedinačnog dijela PPE, praćenje njegovog korišćenja, funkcionalnosti, istorije i lokacije. Sistem radi na desktop i mobilnim uređajima. EHS (Environmental Health and Safety, eng.) profesionalci ga koriste za inventarisanje, pri inspekcijским pregledima, za snimanje efikasnosti korišćenja PPE i sl. [5].

*Honeywell* rješenja koriste UHF RFID čipove kao svojevrsne sertifikate (birth certificate, eng.) koji uključuju informacije o svakom serijskom broju PPE komponente, datumu proizvodnje, proizvođaču i dr. Zahvaljujući ovom sistemu omogućeno je aktivno prikupljanje i praćenje važnih bezbjedonosnih informacija o svakom radniku i njegovoj PPE, na svakoj lokaciji koja je pokrivena sistemom i u bilo kom trenutku. Putem on-line portala, korisnici mogu da se uloguju, skeniraju PPE, vrše inspekcijски nadzor ili poslove vezane za održavanje, kreiraju i šalju izvještaje, aktiviraju alarm i sl. [6].

*RFIDentity* obezbjeđuje napredne analitičke mogućnosti korisnicima prilikom naručivanja, zamjene i/ili opravke čipovanih dijelova PPE. *Intelligentni* PPE modul omogućuje korisnicima da definišu svoja pravila i prate vrijeme korišćenja, navike radnika i potrebe za opravkom ili nabavkom nove PPE. Sistem nudi optimalna rješenja kompanijama kojima je potrebna skalabilna RFID

infrastruktura, smanjenje operativnih troškova, povećanje produktivnosti i dr. [7]

*RFIDNordic* je sistem za podršku bezbjednosti na radu, za više stotina radnika na naftnim platformama i brodovima. Sistem koristi UHF Gen 2 tagove instalirane na zaštitnim šlemovima radnika. Instaliranjem RFID čitača na ulazu u brod, sistem obezbjeđuje nadzorniku informaciju o tome koliko je radnika na brodu i može da optimizira plan napuštanja broda u slučaju opasnosti. Na naftnim platformama, ovaj sistem se koristi u kombinaciji sa senzorima za mjerenje koncentracije štetnih i zapaljivih gasova, kao i za planiranje napuštanja platforme u slučajevima opasnosti. Takođe se koristi za upozorenje radnika ukoliko su u blizini zone sa restriktivnim pristupom. *RFIDNordic* povezuje korisnike, dobavljače, upravne organe, istraživačke centre, pa tako omogućuje dizajniranje i implementiranje bezbjedonosnih rješenja koja odgovaraju individualnim potrebama korisnika [8].

## V. NEKA ALTERNATIVNA RJEŠENJA

Prethodno opisana, komercijalna rješenja za povećanje bezbjednosti radnika u invazivnim radnim okruženjima, u realnom vremenu, bazirana su na periodičnom lociranju ili kontinuiranom praćenju radnika, provjeravanju da li radnici nose propisanu PPE, da li je ona funkcionalna i sl. [9,10]. Ova rješenja su veoma složena i obično prilično skupa. Prema nekim podacima (npr., OpenRTLS, URL: <https://openrtls.com>), instaliranje ovakvih sistema košta od 3-11 USD po m<sup>2</sup> radne površine. To praktično znači, da ako bi se neki od ovih sistema primijenio, primjera radi, na čitavoj površini Luke Bar, koja je oko 130 ha, to bi značilo da treba izdvojiti više milionski iznos novca. Ako se uzme u obzir činjenica da su radnici koji rade na dokovima i u skladištima daleko više izloženi riziku nego oni koji rade u kancelarijama, iznos za implementaciju nekog od prethodno opisanih rješenja bio bi i u tom slučaju velik sa aspekta Luke. U pravcu premošćavanja ovog problema i upoznavanja menadžera i stejholdera sa nekim alternativnim mogućnostima povećanja bezbjednosti lučkih radnika, vršena su neka simulaciona ispitivanja. Naime, simulacije su realizovane nad lejautom kontejnerskog terminala u Luci, budući da se polazi od pretpostavke da su lučki radnici i pješaci na ovom terminalu najizloženiji rizicima (rad na otvorenom, blizina teške manipulativne i transportne mehanizacije, mogućnost greške pri rukovanju i transportu teških tereta i sl.). Istraživanja su prvo išla u pravcu korišćenja PPE sa aktivnim UHF Gen 2 RFID čipovima [11], čitljivim na udaljenostima od nekoliko stotina metara od čitača povezanih na lokalnu mrežu. Pozadinski sistem obezbjeđivao bi pri tome periodično lociranje radnika (kada su u dometu čitača), kao i provjeru da li nose PPE i da li je ona funkcionalna. Na osnovu simulacija u Opnet i Omnet++ okruženju, pokazalo se da sa deset čitača na terminalu i dvadesetak radnika, sistem funkcioniše na udaljenostima većim od 500 m po osnovu nivoa snage prijemnog signala, SNR i BER, za karakteristične WLAN, Wi-Fi i White-Fi frekvencije. Naravno, glavni nedostatak ovakvog sistema je nemogućnost praćenja radnika, korišćenja i funkcionalnosti njihove zaštitne opreme na radu u realnom vremenu. Takođe, napravljena je serija

eksperimenata za MENET i ZigBee tehnologijama, kao komunikacionim kanalima za povećanje bezbjednosti lučkih radnika. Kod MANET scenarija korišćena je ideja da radnici i nadzornici komuniciraju putem PDA (Personal Digital Assistant, eng.) uređaja, a kod ZigBee-a su BCU (Body Central Unit, eng.) radnika tretirani kao krajnji čvorovi mreže, koji komuniciraju sa fiksnim i pokretnim ruterima (instaliranim na viljuškarima), dok čitavom mrežom i saobraćajem vezanim za bezbjednost upravlja koordinator postavljen na strategijskoj lokaciji lučkog terminala. Simulacije su rađene za različite algoritme rutiranja kod MANET-a (OLSR, TORA, AODV i DSR) i za različite topologije ZigBee-a (zvijezda, stablo i mreža). Detaljni rezultati simulacionih analiza performansi ovih mreža, nad terminalom za generalni i kontejnerski teret u Luci Bar su dati u referenci [12], dok su sumarni rezultati prikazani u Tabeli 1.

TABELA 1: UPOREDNA ANALIZA PERFORMANSI MANET I ZIGBEE

MREŽA ZA POTREBE BEZBJEDNOSTI LUČKIH RADNIKA.	
Scenario	Optimalna tehnologija
5 čvorova	ZigBee (zvijezda; 868 MHz i 2.4 GHz)
10 čvorova	ZigBee (mreža; 2.4 GHz)
15 čvorova	MANET (OLSR rutin protokol)
20 čvorova	MANET (AODV rutin protokol)

Dodatno, rađena su istraživanja u pravcu implementacije klauz usluga u rješavanju pitanja bezbjednosti lučkih radnika u Luci Bar. Na osnovu analize sekundarnih literaturnih izvora i do sada razvijene teorije u oblasti metodološkog okvira za adaptaciju klauz, utvrđeno je da bi kao potencijano rješenje mogao da se razmotri model „outsourcing as a service“, eng. Ovaj model podrazumijeva da su isporučioi ICT rješenja za povećanje bezbjednosti ujedno i njegovi vlasnici. Pri tome je odgovornost za implementaciju, održavanje i razvoj sistema, takođe u nadležnosti isporučioaca. Naravno, prije opredjeljenja za ovakav model autsorsinga u klauz trebalo bi detaljno (pre)ispitati tržište i strategijsko opredjeljenje Luke Bar po pitanju uloge i mjesta ICT-a u njenom poslovanju.

## VI. ZAKLJUČAK

U radu su opisana neka *gotova* ICT rješenja za povećanje bezbjednosti radnika u invazivnim oružanjima (EW&SS, MasterLoch Field ID, 3M Science Applied to Life™, Honeywell, RFIDentity i RFIDNordic). Ukazano je na njihove osnovne sličnosti i specifičnosti, pri čemu se može zaključiti da se radi o sofisticiranim rješenjima sa takoreći globalnim pokrivanjem (GPS ili 3G), velikom preciznošću na nivou pojedinačnih PPE dijelova (RFID), u realnom vremenu. Ova rješenja su skalabilna, bazirana na konceptu Interneta stvari ili Interneta svega i orjentisana u pravcu kontekstualnog računarstva. Radom su obuhvaćena i neka alternativna rješenja, koja su nedavno ispitivana na logičkom i simulacionom nivou u kontekstu potreba i realnih mogućnosti Luke Bar, u cilju omogućavanja direktnog povećanja bezbjednosti radnika i indirektnog povećanja efikasnosti upravljanja zaštitom životne sredine

u Luci. Ideje i rezultati ovih istraživanja detaljno su opisani u referencama [11,12]. Dalja istraživanja, pored onih tehničko-tehnološke prirode, trebalo bi da budu usmjerena: (i) na ispitivanje menadžera i stejkholdera, u smislu koliko su zainteresovani za korišćenje ovakvih ICT rješenja u smanjenju rizika od povreda na radu; kao i (ii) na ispitivanje lučkih radnika i nadzornika, u smislu koliko su spremni da postanu konstitutivni dijelovi ovakvih mreža nosivih senzora.

## LITERATURA

- [1] Asaf Adi, Putting wearables to work for improved safety, URL: <https://www.ibm.com/blogs/research/2016/01/putting-wearables-to-work-for-improved-safety/> (pristup: 02.01.2017)
- [2] Chris O'Connor, Improving worker safety with wearables, Internet of Things blog, URL: <https://www.ibm.com/blogs/internet-of-things/worker-safety-and-wearables/> (pristup: 03.01.2017)
- [3] IBM, Using the Internet of Things to boost employee safety, URL: <https://www.youtube.com/watch?v=wJbAKeLGzrs> (pristup: 03.01.2017)
- [4] MasterLoch, Field ID, URL: <http://fieldid.com/> (pristup: 05.01.2017)
- [5] 3M Science Applied to Life™, Workers Health & Safety, URL: [http://www.3m.com/3M/en\\_US/worker-health-safety-us/?WT.mc\\_id=www.3m.com/ppesafety/](http://www.3m.com/3M/en_US/worker-health-safety-us/?WT.mc_id=www.3m.com/ppesafety/) (pristup: 05.01.2017)
- [6] Honeywell, Enabled Safety Protection, URL: <http://www.enabledproducts.com/> (pristup: 05.01.2017)
- [7] RFIDentity, RFID Industrial Solutions, URL: <http://www.rfidentity.com/> (pristup: 05.01.2017)
- [8] RFID Nordic, URL: <http://www.rfidnordic.se/> (pristup: 05.01.2017)
- [9] Kelm A., Laußat L., Meins-Becker A., Platz D., Khazaei M.J., Costin A.M., Helmus M., Teizer J. 2013. Mobile passive Radio Frequency Identification (RFID) portal for automated and rapid control of Personal Protective Equipment (PPE) on construction sites, *Automation in Construction*, 36, pp. 38-52.
- [10] Barro-Torres S., Fernández-Caramés T.M., Pérez-Iglesias H.J., Escudero C.J. 2012. Real-time personal protective equipment monitoring system, *Computer Communications*, 36, pp. 42-50.
- [11] Bauk S., Schmeink A., Colomer J., An RFID Model for Improving Workers' Safety at the Seaport in Transitional Environment, *Transport*, 31(1), 2016, pp. 1-11.
- [12] Bauk S., Gonzalez D.G., Schmeink A., Avramovic Z.Z., MANET vs. ZigBee: Some simulation experiments at the seaport environment, *Journal of Information Technology and Applications*, 6(2), 2016, pp. 63-72.

## ABSTRACT

The paper presents some commercial ICT solutions for increasing workers' safety in invasive environments (steel industry, mining, construction, oil and gas rigs, onboard ships, etc.). It is pointed to the need to increase safety measures, primarily because of the necessity to protect human lives and health, but also due to the need to increase productivity and reduce the costs, which companies and insurance agencies have. The article also briefly describes some alternative solutions that were examined at the level of simulations, in accordance to the real needs and potentials of the Port of Bar, which has been operating during the decades in the transitional conditions and it is faced with the problem of providing sophisticated and usually expensive ICT solutions.

## ON SOME ICT SOLUTIONS FOR INCREASING SAFETY AT WORK IN INVASIVE ENVIRONMENTS

Sanja Bauk, Tatijana Dlabač

# Automated classification of web application parts from a blackbox security testing perspective

Stanko Jankovic; Biljana Stamatovic

**Abstract** — This paper describes possible methods of increasing the efficiency of automation of blackbox security testing of web applications. It does not focus on the aspect of testing certain parts of the web application and the identification of vulnerabilities, but on the selection of parameters/parts of the web application that should be tested. Of course, the goal is always to test the complete application, but from a blackbox perspective, identifying all the parts of the web application, and at the same time avoiding to test same parts of the web application multiple times, is not always an easy task.

**Keywords** — application, blackbox, security, testing, web

## I. INTRODUCTION

TECHNOLOGY is changing the world with an ever increasing pace, and the security aspect of it was never as important as today. Web sites/applications are no exception of that and are often the favorite target of attackers. But securing web applications is not an easy task. Developers have often to consider functionality, performance, scalability, sustainability, and as if that would not be enough, security and many other factors must not be forgotten. In addition to that, web applications can almost never be considered as a finished product. They are constantly being updated/changed, often by a large number of developers, which makes it hard to keep an overview of all the cogwheels.

Many companies involved in web application development have recognized that security is not something they can manage alone. They often try to improve security by starting bug bounty programs and by that relying on thousands of "whitehat hackers" to identify and report potential security vulnerabilities in their products. Most often, these companies do not give any privileged access to their product to the hackers, but only blackbox access which is basically the level of access any anonymous user would have. That approach allows companies like Facebook [1], Google [2], Microsoft [3], and many other, to hire a large number of testers, without

the need to provide the source code of their products to them.

From a blackbox perspective, an assessment of the security of a web application can be done by just focusing on the inputs and outputs. In this paper, problems of automation of blackbox security testing are presented through three high level steps:

1. Discovering web application content/parts
2. Selection of web application parts to be tested
3. Testing of the selected parts

The first step is usually done manually by browsing through the web application or automated by using a spidering tool. The second step is almost always done manually. The third step can be done manually or automated by using various tools. Many tools exist to automate the first and third step [4][5], but all of them are quite limited with automating the second step. The only way to indirectly, and very roughly automate the second step using the mentioned tools would be to limit the spidering tool to spider only to a certain depth of the web application [6].

The second step is certainly the step with least options for automation, but it is a very important step because selecting too many parts of the application could lead to unnecessarily testing multiple parts of the application which are essentially the same. For example, imagine a web site which contains thousands of blog posts. From a security perspective, it would most probably be enough to test the parameters related to just one blog post, and not all parameters for all blog posts. Similarly, selecting too few parts of the web application could lead to some vulnerabilities to be missed.

But why is the second step almost always done manually? Are there really no ways to automate this step? In fact, the spidering tool from step 1 could technically be limited to spider only content to a certain link depth which would in some way limit the number of parts that need to be selected/tested, but that is a very rough criteria. Why could we not develop an algorithm that would recognize if two parts of a web application are the same from a security perspective or not (for example if two parts represent just two blog posts, or two completely different parts of the web application) and classify them accordingly. Ideally, we would classify the parts of a web application based on the code coverage they trigger on the server side, but that

Stanko Janković, Faculty of information systems and technology, University of Donja Gorica, Donja Gorica, 81000 Podgorica, Montenegro (telephone: 382-67-758679; e-mail: stanko.jankovic@live.com).

Prof. Dr. Biljana Stamatović, Faculty of information systems and technology, University of Donja Gorica, Donja Gorica, 81000 Podgorica, Montenegro (telephone: 382-20-410719; e-mail: biljana.stamatovic@udg.edu.me).

is not something we have access to from a blackbox perspective. The only criteria/indicators for classification would basically be the contents of the HTTP requests[7] and the HTTP responses[8] related to the particular web application parts as that is the only way to interact with an application from a blackbox testing perspective. Could we use these in order to automate the classification of web application parts from a security perspective?

## II. INDICATORS AND ALGORITHM FOR CLASSIFICATION

Before we get into the details of what indicators could be used for the classification of the parts of a web application from a security perspective, we should define what a "part" of a web application actually represents. We will consider each HTTP request that a web application can generate as one of its possible parts, and therefore we are basically speaking about the classification of HTTP requests of a web application.

As mentioned earlier, indicators for such a classification could be elements from the HTTP requests or the corresponding HTTP responses. For example, we could try to use the following:

- the structure of the URL from the HTTP request
- the names of the GET/POST parameters from the HTTP request
- the HTML code from the HTTP response
- the CSS code that the HTML code from the HTTP response uses

We have analyzed the usability of the mentioned indicators on a test site. The test site was a Wordpress site (version 4.5.2) in which we inserted 1000 blog posts and 500 comments using the FakerPress plugin. We also changed the link format to "Post name" in order to make it harder to do the classification based on URL structure only. For the purpose of simplifying the test cases, the tests were done as anonymous user and did not consider the effects of JavaScript, nor did they take into account HTTP request types other than GET and POST, nor those that did not get HTML code as a response.

### A. Structure of the URL from the HTTP request

Because we want our classification algorithm to be usable across most web applications, without the need to modify it, it is clear that the raw URL structure could not be used as an indicator. But if we would normalize the URL structure in the way so that we replace each web folder name in the URL with "string" if it is a string, with "int" if it is an integer or with "file" if it is a filename, we would get a URL structure which we could use to tell if the chances that two HTTP requests should be located in separate classes is high or not. For example, the following links:

- <http://localhost/blog/2/>
- <http://localhost/contacts/>

would become:

- <http://string/string/int/>
- <http://string/string/>

and it is clear that these two HTTP requests should belong in separate classes. Therefore we will use this indicator in our algorithm.

### B. Names of the GET/POST parameters from the HTTP request

If two HTTP requests use parameters with different names, it is very probable that they execute different code on the server side and that they should be put into different classes. But in our tests we concluded that very often one HTTP request can use a lot of different combinations of parameter names (like sort, filter, group...) which led to probably unnecessary creation of additional classes of HTTP requests. Therefore we excluded the use of GET parameter names for classification purposes in our algorithm, but noted that it is still important to keep track of all possible parameter names for certain classes of HTTP requests in order to be able to test all parameters later on.

There was one small exception, which we will mention later on, where we in certain cases still used the POST parameter names for the classification of HTTP POST requests.

### C. HTML code from the HTTP response

Although this was maybe the most promising indicator, it did not perform that well in our tests. It was of course clear that comparing two sets of HTML codes (as strings) from two pages would always be different if we do not perform some normalization of the codes (for example, if we compare the HTML codes of two blog post pages, they will be different, even if from a security perspective those two pages are the same). Therefore we tried to perform the following normalizations in order to make this indicator more usable:

- Removing all text (keeping only HTML tags)
- Removing all text, attributes and double elements (for example if a page has multiple comments)
- Removing all text, attributes, double elements, and elements that are usually only used in combination with text (like paragraph tags, heading tags, annotation tags...)
- Removing all text, attributes, double elements, and elements that are usually only used in combination with text and tags/DOM elements that do not contain any links, forms or buttons
- Removing all text, attributes, double elements, and all tags except div tags

In the tests that we performed, it was evident that the percent similarity between pages that should be placed in the same class are higher than between pages that should be placed in different classes, and almost all normalization steps increased the percent similarity between pages that should be placed in same class. But still, it seemed almost impossible to get the HTML codes into such a form that they are equal if they need to be placed into same class,

and different only if they need to be placed into different classes. Because it would be quite hard to identify a universal threshold for the percent similarity that would be used to split the HTTP requests into classes, we concluded that we would not use this indicator in our algorithm.

#### D. CSS code that the HTML code from the HTTP response uses

Although this indicator did not appear as the most promising one, it showed very good results in our test. Even without any normalization steps, it showed higher percent similarities between pages that should be placed in the same class than the previous indicator. After we performed several steps with different normalization techniques, we achieved the best result by removing all elements that are usually only used in combination with text and tags/DOM elements that do not contain any links (excluding link tags for CSS stylesheets), forms or buttons before we calculated the CSS code that the HTML code from the HTTP response uses. We concluded that this indicator could be used in our algorithm.

### III. ALGORITHM

In our tests, we have focused on HTTP GET and POST requests. It must be noted that we will not be able to use the same indicators for HTTP POST requests as we use for HTTP GET requests. The reason for that is that POST requests most often perform much more sensitive actions on the server side than GET requests. That means that for GET requests it is not a problem to automatically submit them and use the elements from the response for their classification. But for POST requests, it is not really recommended to automatically submit them because we could run into the danger of performing some sensitive actions like sending e-mails or posting comments. Therefore, we cannot use the elements from the HTTP response for classifying HTTP POST requests and our algorithm will behave differently depending on the type of request it is classifying.

#### A. Classifying HTTP GET requests

Our algorithm will put HTTP GET requests into the same class if they have the same URL structure and use the same CSS code (considering the normalization described in II.D).

#### B. Classifying HTTP POST requests

Our algorithm will put HTTP POST requests into the same class if they are generated by forms that have the same action attribute and same parameter names, or if they are generated by forms which have different action attributes, but same parameter names and the pages on which the forms are located belong to the same GET class.

We have tested such an algorithm, combined with an spidering algorithm that spidered 5 links deep, on our test site. The Python script that implemented all that discovered 501 HTTP GET request and 3 HTTP POST requests on our test site which it then classified into 23 classes of HTTP GET requests and 3 classes of HTTP

POST requests. One minor problem could be noticed: out of the 23 HTTP GET classes, 11 classes were "double". That means, for example not all pages that contained blog posts were put into a single class, but were distributed across several classes because of minor differences. But this amount of "double" classes is actually not really a problem. It just means that we would have to test 11 HTTP requests that we do not really have to test. But remembering that before the classification, that amount was not 11, but much higher (489), it becomes clear that the achieved results are very good.

### IV. USING THE CLASSIFICATION FOR SPIDERING EFFICIENCY IMPROVEMENT

Although our algorithm was quite effective in classifying HTTP requests, it was not completely autonomous because we still had to provide the link depth parameter to the spider. In our tests, we spidered 5 links deep. But would we discover more useful content from a security perspective with a higher value? Or would an even smaller value get us the same results? This depends of course on the site we are spidering, but in our case it turned out that this number was too high. As it is shown in table 1, we would get the same number of classes if spidered only 2 links deep, and even with 1 link depth, we would get just one class less, which is anyway a double one.

TABLE 1: RESULTS BASED ON SPIDERING DEPTH

Spidering depth	Number of discovered HTTP GET requests	Number of classes of HTTP GET requests	Number of classes of HTTP POST requests
5	501	23	3
4	393	23	3
3	295	23	3
2	182	23	3
1	114	22	3
0	29	13	2

If we would represent the number of discovered HTTP GET classes based on spidering depth, we would get the following graph:

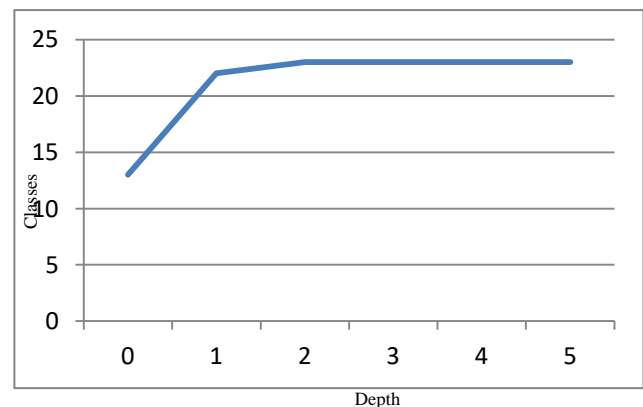


Fig 1. Number of discovered HTTP GET classes based on spidering depth

But how could we automate the selection of the spidering depth value? One way would be to start with a low value and increment it till we notice that the curve from Fig 1 flattens. But is there any better way? Let's say that we are analyzing a web page of a blog post. It makes sense to follow all the links on that blog post in order to see if any useful content lies behind them. But is it necessary to do the same on all pages with blog posts? That means, if we spidered all the links on a page from one class, do we really need to spider all the links on all other pages from the same class? We modified the spidering part of our classification tool/algorithm in a way that it did not require any link depth parameter, but that it just simply did not spider links on pages if it already spidered links on a page from the same class. The results were quite surprising: our tool spidered 68 pages which it classified into 23 classes of HTTP GET requests and 3 classes of HTTP POST requests, and all that without specifying any link depth parameter. That means that it was even more efficient than if we knew in advance which link depth is perfect for our site (look at Table 1).

#### V. CONCLUSION

The algorithm described in this paper was able to classify parts of a web application in a way that would help improve the efficiency of automating the process of blackbox security testing. The classification helps not only with the process of selecting HTTP requests that have to be tested, but also to make the spidering algorithm more efficient.

This paper focused primarily on the testing environment on which the algorithm was tested, but the algorithm showed positive results also on several real world web sites/applications, although it is quite hard to measure/quantify such results because even with manual selection, it is often impossible to tell what the correct classification of some application parts would be from a security perspective.

Although the described algorithm is most probably far from perfect, and although automated tools will not fully replace manual testing that soon, especially not for more complex vulnerability testing or logic flaw testing [9], it is at least clear what some of the obstacles in the automation of blackbox security testing are, and that overcoming those obstacles would certainly help in improving the efficiency of automated blackbox scanning, but could probably also aid manual testing by showing the target from new viewpoints. When it comes to possible future researches on this topic, it is very likely that machine learning/artificial intelligence (AI) algorithms will play a big role in solving the problems mentioned in this paper. Most probably such algorithms will be giving even better results than the algorithm described in this paper and some projects are already showing that AI will play a big role in many fields, including IT security [10].

#### REFERENCES

- [1] Facebook Bug Bounty Program: <https://www.facebook.com/whitehat>
- [2] Google Bug Bounty Program: <https://www.google.com/about/appsecurity/reward-program/>
- [3] Microsoft Bug Bounty Program: <https://technet.microsoft.com/en-us/library/dn425036.aspx>
- [4] Bau, Jason; Bursztein, Elie; Gupta, Divij; Mitchell, John , *State of the Art: Automated Black-Box Web Application Vulnerability Testing*, Security and Privacy (SP), 2010 IEEE Symposium on <http://ieeexplore.ieee.org/document/5504795/>
- [5] Michael, C.C; Ken van Wyk; Radosevich Will; US-CERT; Blackbox security testing tools (downloaded on 02.02.2017): <https://www.us-cert.gov/bsi/articles/tools/black-box-testing/black-box-security-testing-tools>
- [6] Acunetix site crawler/spider setup, (downloaded on 01.12.2016.): <http://www.acunetix.com/support/docs/wvs/acunetix-site-crawler/>
- [7] Balzarotti, Davide; Pellegrino, Giancarlo, *Toward Black-Box Detection of Logic Flaws in Web Application*, San Diego, Conference paper (2014) <http://www.eurecom.fr/en/publication/4207/detail/toward-black-box-detection-of-logic-flaws-in-web-applications>
- [8] HTTP request specifications (downloaded on 08.08.2016): <https://www.w3.org/Protocols/rfc2616/rfc2616-sec5.html>
- [9] HTTP response specifications (downloaded on 08.09.2016.): <https://www.w3.org/Protocols/rfc2616/rfc2616-sec6.html>
- [10] Kumar, Mohit; The Hacker News; *Cloud-AI – Artificially Intelligent System Found 10 Security Bugs in LinkedIn*, (downloaded on 02.02.2017): <http://thehackernews.com/2017/01/artificial-Intelligence-cybersecurity.html>

# Tuning rules of PI/PID controllers with robustness constraint for industrial processes without dead-time

Marko Č. Bošković, Tomislav B. Šekara, Milan R. Rapaić, Vidan Govedarica

**Abstract** — This paper provides simple and an effective algorithm to design controllers with robustness constraint for typical industrial processes without dead time. The design procedure is performed in complex domain and presented method is applicable to a variety of systems. Good set point response without overshoot and good load disturbance rejection are obtained with a single tuning parameter considering the maximum sensitivity and sensitivity to measurement noise. The presented design method is applied to typical low-order representatives of industry processes including stable, integrating and oscillating plants without dead time. The solution to this problem leads to PI and PID controller for the first order and the second order processes, respectively. The obtained results of numerical simulations validate analytical analysis.

**Keywords** — PI/PID control; Tuning rules; Load disturbance rejection; Set point response; Robustness;

## I. INTRODUCTION

THE most of controller design methods start with models of the process and specified requirements on performance and robustness. The main controller design goals are to attenuate load disturbance as much as possible (process control) and to achieve setpoint response (motion control) [1]. This is often performed by using two degrees of freedom control structure where the controller design problem is separated from the set-point tracking.

The dominant control algorithm in process control industry is still PI/PID controller [2,3]. Variety of methods for tuning parameters of PI/PID controllers are reported in [4]. Numerous simple and efficient tuning procedures for PI/PID controllers are developed such as [5-8], as well as the optimization procedures under constraints on robustness are elaborated [9-24]. Besides, different IMC based (Internal Model Control) PID controller tuning methods are developed to control stable processes with

time delay such as [25-27], as well as modified IMC based methods used to control time delayed and unstable processes [28-31]. In order to obtain desired performance it is necessary to tune well controller and achieve robust stability to the model uncertainties. However, good disturbance rejection corresponds to smaller values of IAE (Integrated Absolute Error), which leads to higher sensitivity to the measurement noise and lower robustness. Hence, the control system design goal is to find the best trade-off between these conflicting criteria [16].

This paper is focused solving control problems involving low-order models of plants typically encountered in industry. The controller  $C(s)$  is designed for the first and the second order processes under constraints on robustness and sensitivity to the measurement noise. The controller designed for the first order plants is a PI controller, while for the second order plants PID is obtained. Further considerations are beyond the scope of the present paper, the same approach can be used to control more complex stable and unstable processes, with and without dead time. In this way a high order controller would be obtained, as demonstrated in [31-32], where PID is retrieved by using 2/2 Pade approximation on the complex controller transfer function.

## II. ANALYTICAL COMPLEX-DOMAIN CONTROLLER DESIGN FOR PROCESSES WITHOUT DEAD TIME

The control system architecture is presented in Fig. 1, where  $C(s)$  is controller transfer function. It is assumed that process transfer function  $G_p(s)$  is known.  $G_f(s)$  describes the feed forward from reference signal  $r$  to the control signal  $u$ . The load disturbance, the measurement noise and the output signal, are denoted with  $d$ ,  $n$ ,  $y$ , respectively, as it is shown in Fig. 1.

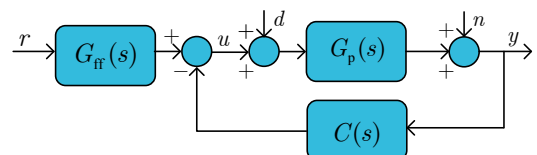


Fig. 1. Block diagram of the control structure with controller  $C(s)$  and process  $G_p(s)$

Wide class of typical industrial processes may be modelled as  $G_p(s) = H(s)e^{-\tau s} / Q(s)$ , where  $Q(s)$  and  $H(s)$  are polynomials  $\deg Q(s) = n \geq \deg H(s)$ ,  $n \leq 2$ ,  $H(0) \neq 0$ . In this paper processes without dead time are considered, i.e.  $\tau = 0$ . Complementary sensitivity function of the

Marko Č. Bošković, Faculty of Technical Sciences, University of Novi Sad, Trg Dositeja Obradovića 6, 21000 Novi Sad, Serbia; (e-mail: marko.boskovic@uns.ac.rs).

Tomislav B. Šekara, School of Electrical Engineering, University of Belgrade, Bulevar kralja Aleksandra 73, 11120 Belgrade, Serbia (e-mail: tomi@etf.rs).

Milan R. Rapaić, Faculty of Technical Sciences, University of Novi Sad, Trg Dositeja Obradovića 6, 21000 Novi Sad, Serbia; (e-mail: rapaja@uns.ac.rs).

Vidan Govedarica, Faculty of Electrical Engineering, University of East Sarajevo, Vuka Karadžića 30, 71123 East Sarajevo, Bosnia and Herzegovina (e-mail: vidangov@yahoo.com).



controlled process  $G_p(s)$  in Fig. 1. is defined by  $T_p(s) = L(s)/(1+L(s))$ , wherein  $L(s) = C(s)G_p(s)$  is the loop transfer function. The design procedure starts with specifying the desired complementary function  $T(s)$  as follows

$$T(s) = \frac{N(s)}{P(s)}, \quad (1)$$

where  $N(s) = 1 + \sum_{j=1}^n \eta_j s^j$ ,  $P(s) = (\lambda s + 1)^{2n}$

$\eta_j \in \mathbb{R}$ ,  $j = \overline{1, n}$  and  $\lambda > 0$  adjustable parameter. The controller  $C(s)$  is defined using (1) to maximally attenuate disturbance  $d$

$$C(s) = \frac{1}{G_p(s)} \frac{T(s)}{1-T(s)} = \frac{1}{H(s)} \frac{N(s)Q(s)}{F(s)} \quad (2)$$

where  $F(s) = P(s) - N(s)$ . Parameters  $\overline{\eta_1, \eta_n}$  are calculated to achieve cancellation of poles of the process  $G_p(s)$  with zeros of the function  $F(s)$ .

Free parameter of the controller  $C(s)$  is time constant  $\lambda > 0$ . Time constant  $\lambda$  is selected to meet desired performance and adequate robustness of the closed loop system. In this paper, the performance index is characterized by the integrated absolute error IAE [33]

$$\text{IAE} = \int_0^{\infty} |e_d(t)| dt \quad (3)$$

where  $e_d$  is control error due to unit step load disturbance applied at the process input as shown in Fig. 1. It is obvious this performance index is reduced to  $\text{IE} = 1/k_i$  (Integrated Error) when  $e$  is positive, which will always be the case when the closed loop system is well-damped. As the measure of robustness, a maximum of the sensitivity function  $M_s = \max_{\omega} |1/(1+L(i\omega))|$  is used in the present research. The sensitivity to the measurement noise is expressed in the form of  $M_n$  defined by

$M_n = \max_{\omega} |G_{un}(i\omega)| = \max_{\omega} |-C(i\omega)/(1+L(i\omega))|$ , while  $M_p$  is the maximum of complementary sensitivity function, ie.  $M_p = \max_{\omega} |L(i\omega)/(1+L(i\omega))|$ . Since the considered process transfer functions  $G_{pi}(s)$ ,  $i=1, \dots, 7$ , are of the first or the second order, the controller transfer function in (2) is reduced to PI or PID controller

$$C_{PI}(s) = \frac{k_p s + k_i}{s}, \quad C_{PID}(s) = \frac{k_d s^2 + k_p s + k_i}{s(T_f s + 1)} \quad (4)$$

where  $k_p$ ,  $k_i$ ,  $k_d$  are proportional, integral and derivative gain and  $T_f$  is filter time constant of PID controller. Analytical tuning formulae of PI/PID controller parameters for typical low-order process models  $G_{p1}(s) = K/s$ ,  $G_{p2}(s) = K/(Ts+1)$ ,  $G_{p3}(s) = K/(s(Ts+1))$ ,  $G_{p4}(s) = K/((T_1 s + 1)(T_2 s + 1))$ ,  $G_{p5} = K\omega_n^2 / (s^2 + 2\zeta\omega_n s + \omega_n^2)$  are given in the Table 1.

### III. SIMULATION ANALYSIS

The proposed design method is applied to typical representatives of industrial processes without dead time. The test batch includes stable, integrating and oscillating plants without dead time.

$$G_{p1}(s) = \frac{3}{s}, \quad G_{p2}(s) = \frac{25}{2s+1}, \quad G_{p3}(s) = \frac{1}{s(s+1)},$$

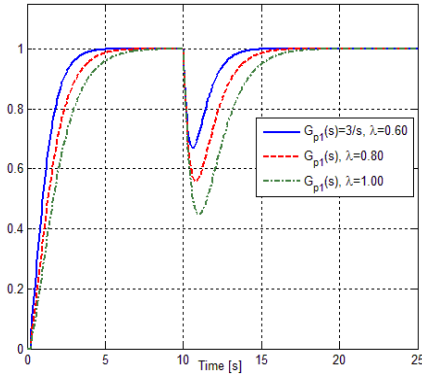
$$G_{p4}(s) = \frac{2}{(10s+1)(5s+1)}, \quad G_{p5}(s) = \frac{4}{s^2+0.16s+1},$$

$$G_{p6}(s) = \frac{5(1+s)}{(2s+1)(5s+1)}, \quad G_{p7}(s) = \frac{5(1-s)}{(2s+1)(5s+1)},$$

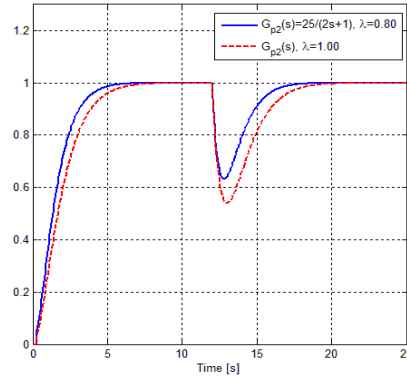
$$G_{p8}(s) = \frac{1}{(s+1)^4}, \quad G_{p9}(s) = \frac{1}{\prod_{k=0}^3 (0.7^k s + 1)}.$$

TABLE 1: THE TUNING FORMULAE OF PI/PID CONTROLLER PARAMETERS FOR TYPICAL INDUSTRIAL PROCESSES

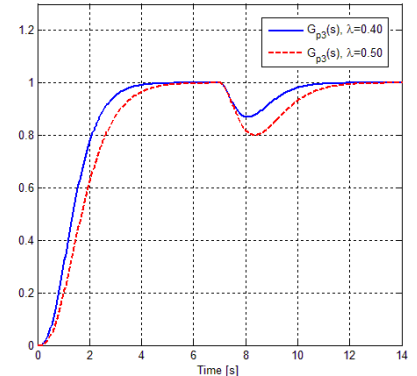
Process model	Controller type	Parameters
$G_{p1}(s) = \frac{K}{s}$	PI	$k_p = \frac{2}{K\lambda}, k_i = \frac{1}{K\lambda^2}$
$G_{p2}(s) = \frac{K}{Ts+1}$	PI	$k_p = \frac{2T-\lambda}{K\lambda}, k_i = \frac{T}{K\lambda^2}$
$G_{p3}(s) = \frac{K}{s(Ts+1)}$	PID	$k_p = \frac{4T^2}{K\lambda^2(4T-\lambda)}, k_i = \frac{T^2}{K\lambda^3(4T-\lambda)}, k_d = \frac{6T^2-\lambda(4T-\lambda)}{K\lambda(4T-\lambda)}, T_f = \frac{T\lambda}{4T-\lambda}$
$G_{p4}(s) = \frac{K}{(T_1 s + 1)(T_2 s + 1)}$	PID	$k_p = \frac{(T_1 + T_2)\lambda^3 - 4T_1 T_2 \lambda^2 + 4T_1^2 T_2^2}{K\lambda^2(4T_1 T_2 - \lambda(T_1 + T_2))}, k_i = \frac{T_1^2 T_2^2}{K\lambda^3(4T_1 T_2 - \lambda(T_1 + T_2))},$ $k_d = \frac{(T_1^2 + T_1 T_2 + T_2^2)\lambda^2 - 4T_1 T_2 (T_1 + T_2)\lambda + 6T_1^2 T_2^2}{K\lambda(4T_1 T_2 - \lambda(T_1 + T_2))}, T_f = \frac{T_1 T_2 \lambda}{4T_1 T_2 - \lambda(T_1 + T_2)}$
$G_{p5}(s) = \frac{K\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}$	PID	$k_p = \frac{\zeta\omega_n^3 \lambda^3 - 2\omega_n^2 \lambda^2 + 2}{K\omega_n^2 \lambda^2 (2 - \zeta\omega_n \lambda)}, k_i = \frac{1}{2K\omega_n^2 \lambda^3 (2 - \zeta\omega_n \lambda)},$ $k_d = \frac{(4\zeta - 1)\omega_n^2 \lambda^2 - 8\zeta\omega_n \lambda + 6}{2K\omega_n^2 \lambda^3 (2 - \zeta\omega_n \lambda)}, T_f = \frac{\lambda}{2(2 - \zeta\omega_n \lambda)}$



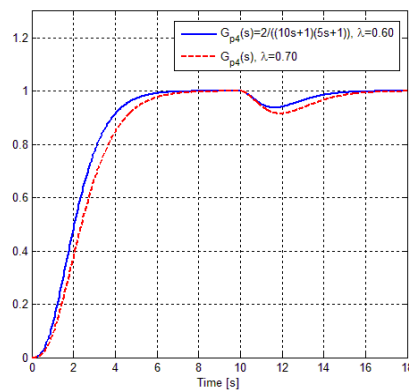
**Fig. 2.** The reference unit step response of process  $G_{p1}(s)$  with load disturbance  $D(s) = -1/s$  starting from  $t=10$  s



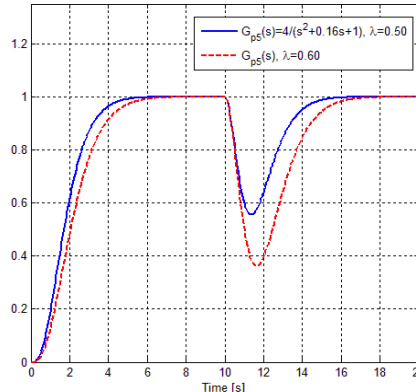
**Fig. 3.** The reference unit step response of process  $G_{p2}(s)$  with load disturbance  $D(s) = -0.1/s$  starting from  $t=12$  s



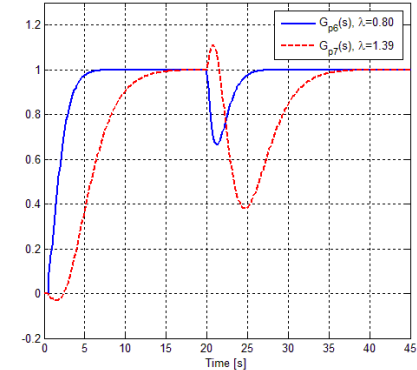
**Fig. 4.** The reference unit step response of process  $G_{p3}(s)$  with load disturbance  $D(s) = -1/s$  starting from  $t=7$  s



**Fig. 5.** The reference unit step response of process  $G_{p4}(s)$  with load disturbance  $D(s) = -5/s$  starting from  $t=10$  s



**Fig. 6.** The reference unit step response of process  $G_{p5}(s)$  with load disturbance  $D(s) = -0.5/s$  starting from  $t=10$  s



**Fig. 7.** The reference unit step response of process  $G_{p6}(s)$  and  $G_{p7}(s)$  with load disturbance  $D(s) = -1/s$  and  $D(s) = -0.5/s$  respectively starting from  $t=20$  s

TABLE 2: THE OBTAINED RESULTS USING PROPOSED CONTROLLER DESIGN METHOD

Process	Control	$\lambda$	$k_p$	$k_i$	$k_d$	$T_r$	IAE	$M_n$	$M_s$	$M_p$
$G_{p1}(s)$	PI	0.60	1.1111	0.9259	-	-	1.0800	1.11	1.00	1.15
$G_{p1}(s)$		0.80	0.8333	0.5208	-	-	1.9201	0.83	1.00	1.15
$G_{p1}(s)$		1.00	0.6667	0.3333	-	-	3.0003	0.67	1.00	1.15
$G_{p2}(s)$	PI	0.80	0.1600	0.1250	-	-	8.0000	0.16	1.00	1.02
$G_{p2}(s)$		1.00	0.1200	0.0800	-	-	12.5000	0.12	1.00	1.00
$G_{p3}(s)$	PID	0.40	6.9444	4.3403	3.1667	0.1111	0.2304	28.50	1.34	1.38
$G_{p3}(s)$		0.50	4.5714	2.2857	2.4286	0.1429	0.4375	17.00	1.33	1.35
$G_{p4}(s)$	PID	0.60	72.2167	30.2986	57.8665	0.1571	0.0329	368.42	1.38	1.49
$G_{p4}(s)$		0.70	53.3474	19.2312	48.9474	0.1847	0.0519	265.01	1.38	1.47
$G_{p5}(s)$	PID	0.50	0.7704	0.5102	0.6934	0.1275	1.9600	5.44	1.33	1.34
$G_{p5}(s)$		0.60	0.4615	0.2964	0.5619	0.1537	3.3738	3.67	1.30	1.25
$G_{p6}(s)$	PID	0.80	7.1974	1.1355	9.3918	4.5470	0.8807	2.06	1.06	1.27
$G_{p7}(s)$	PID	1.39	0.7115	0.1389	0.8704	0.2593	7.6114	3.36	2.00	1.35
$G_{p8}(s)$	PID	0.85	1.0623	0.3567	0.8890	0.0606	2.8043	14.66	1.40	1.00
$G_{p9}(s)$	PID	0.65	0.8619	0.3993	0.5078	0.5988	2.5048	0.85	1.40	1.00

In all examples, the control signal  $u$  in Fig. 1 is defined as

$$u(t) = k_p (br(t) - y_f(t)) + k_i \int (r(t) - y_f(t)) dt - k_d \frac{dy_f(t)}{dt}$$

wherein is  $0 \leq b \leq 1$ , and  $b$  is used for set point weighting as in [12], and  $y_f$  is filtered output signal generated by the first order noise filter. The results of numerical simulations for processes  $G_{p1}$ - $G_{p7}$  are presented in Fig. 2-7. The obtained parameters of PI/PID controllers and quantitative indices of performance and robustness are given in Table 2. From Fig. 2-7 and Table 2 it is obvious that smaller value of  $\lambda$  leads to better load disturbance

response ie. smaller IAE and greater values of sensitivity to the measurement noise  $M_n$ . Adjustable parameter  $\lambda$  is selected to have satisfactory load disturbance rejection and the set-point response. The plant operator can decrease  $\lambda$  to smaller values and get smaller IAE as long as the actuator can provide control signals of appropriate intensities, that is as long as the saturation of the control signal is avoided. The proposed tuning method is applicable to the processes including the left and right half-plane zeros as it is shown in Fig. 7. However, if processes contain left half-plane zeros which are dominant then the tuning rules should be modified as it was recently

reported in [34]. In similar manner, the tuning procedure can be performed to control high-order stable ( $G_{p8}$  and  $G_{p9}$ ) and unstable processes with and without dead time. The idea is to find first the high-order controller transfer function in (2) and then apply in a certain way some of rational approximations such as Pade approximation as in [32] and reduce it to coincident low-order controller. In all cases, it is important to consider compromise between performance and robustness.

#### IV. CONCLUSION

This paper presents an analytical complex domain based procedure to design controllers for typical industrial processes without dead time. It is assumed that plant transfer function is a priori known. The tuning procedure has one free parameter which is selected to obtain satisfactory load disturbance response as well as set-point response and to achieve the robustness of the closed-loop system. The solution leads to PI controller for the first order plants and PID controller for the second order plants, while for higher order processes PID controller may be obtained using Pade approximation. This is done by choosing adequate complementary sensitivity transfer function and taking into account compromise between performance and robustness.

#### ACKNOWLEDGEMENTS

Authors gratefully acknowledge the financial support of Ministry of Education and Technological Development of the Republic Serbia, Grants TR 33020 (T.B.Š.), TR32018 and TR33013 (M.R.R). The author (M.R.R) would like to express gratitude for the project „Kratkoročni projekat od posebnog značaja za razvoj APV, contract number: 114-451-620/2015-03“.

#### REFERENCES

- [1] F.G. Shinskey, "Process-Control Systems. Application Design and Tuning", McGraw-Hill, New York, USA, 4th edition, 1996.
- [2] S. Yamamoto and Hashimoto, "Present status and future needs: the view from Japanese industry", In Arkun and Ray; Eds., Chemical Process Control-CPCIV. Proc. 4th Inter. Conf. on Chemical Process Control, Tex, USA, 1991.
- [3] L. Desborough, R. Miller, "Increasing customer value of industrial control performance monitoring—Honeywell's experience", in: Sixth International Conference on Chemical Process Control, AIChE Symposium Series Number 326, Vol. 98, 2002.
- [4] A. O'Dwyer, „Handbook of PI and PID controller tuning rules“, Imperial College Press 3rd edition, 2009.
- [5] K.J. Åström, T. Hägglund, "Revisiting the Ziegler-Nichols step response method for PID control", *Journal of Process Control*, Vol. 14, pp. 635–650, 2004.
- [6] S. Skogestad, "Simple analytic rules for model reduction and PID controller tuning", *Journal of Process Control*, Vol. 13, pp. 291–309, 2003.
- [7] M. Shamsuzzoha, M. Lee, "IMC-PID controller design for improved disturbance rejection of time-delayed processes", *Industrial & Engineering Chemistry Research*, Vol. 46, no.7, pp. 2077-2091, 2007.
- [8] K.J. Åström, T. Hägglund, "Revisiting the Ziegler-Nichols step response method for PID control", *Journal of Process Control*, Vol. 14, pp. 635-650, September 2004
- [9] M.R. Mataušek, T.B. Šekara, "PID controller frequency-domain tuning for stable, integrating and unstable processes, including deadtime", *J. Process Control*, Vol. 21, pp. 17-27, 2011.
- [10] T.B. Šekara, M.R. Mataušek, "Classification of dynamic processes and PID controller tuning in a parameter plane", *J. Process Control* Vol. 21, Issue 4, pp. 620-626, 2011.
- [11] K.J. Åström, H. Panagopoulos, and T. Hägglund, "Design of PI controllers based on non-convex optimization," *Automatica*, Vol. 34, pp. 585–601, May 1998
- [12] H. Panagopoulos, K.J. Åström and T. Hägglund, "Design of PID controllers based on constrained optimization", *IEE Proceedings-Control Theory and Applications*, ol. 149, pp. 32-40 January 2002.
- [13] B. Kristiansson and B. Lennartson, "Evaluation and simple tuning of PID controllers with high-frequency robustness", *Journal of Process Control*, vol.16, pp. 91-102, February 2006.
- [14] B. Kristiansson and B. Lennartson, "Robust tuning of PI and PID controllers: using derivative action despite sensor noise" *IEEE Control Systems Magazine*, pp. 55-69, February 2006.
- [15] A.J. Isaksson and S.F. Graebe, "Derivative filter is an integral part of PID design", *IEE Proceedings-Control Theory and Applications*, vol.149, pp. 41-45, January 2002.
- [16] O. Garpinger, T. Hägglund, K. J. Åström, „Criteria and Trade-offs in PID Design“, IFAC Proc. Volumes, Vol. 45, Issue 3, pp. 47-52, 2012.
- [17] T. Hägglund, „A unified discussion on signal filtering in PID control“, *Control Engineering Practice*, Vol. 21, pp. 994-1006, Aug. 2013.
- [18] V.R. Segovia, T. Hägglund, K.J. Åström, "Measurement noise filtering for PID controllers", *Journal of Process Control*, Vol. 24, pp. 299-313, 2014.
- [19] K. Soltesz, C. Grimholt, S. Skogestad, „Simultaneous design of proportional-integral-derivative controller and measurement filter by optimisation“, *IET Control Theory&Applications*, pp.1-16, Nov. 2016.
- [20] T.B. Šekara and M.R. Mataušek, "Optimization of PID controller based on maximization of the proportional gain under constraints on robustness and sensitivity to measurement noise ", *IEEE Trans. Automatic Control*, vol. 54, no.1, pp.184-189, Jan. 2009.
- [21] P.O. Larsson, T. Hägglund, "Control signal constraints and filter order selection for PI and PID controllers", in: American Control Conference, San Francisco, CA, USA, 2011, pp. 4994–4999.
- [22] T.B. Šekara and M.R. Mataušek, "Revisiting the Ziegler-Nichols process dynamics characterization", *J. Process Control*, Vol. 20, pp. 360-363, 2010.
- [23] T.B. Šekara, M.R. Mataušek, „Nove formule podešavanja PI regulatora za industrijske procese sa integralnim dejstvom“, *Vojnotehnički glasnik*, Vol. 61, No. 1, pp. 7-20, 2013.
- [24] O. Garpinger, T. Hägglund, „Software-based optimal PID design with robustness and noise sensitivity constraints“, *J. Process Control*, Vol 33, pp. 90-101, 2015.
- [25] L. I. Chien, "IMC-PID controller design-an extension" *IFAC Proceedings Series*, vol. 6 (pp. 147-152), 1988.
- [26] P. S. Fruehauf, L. I. Chien, M. D. Lauritsen, "Simplified IMC-PID tuning rules." *ISA Transactions*, 33(1), 43-59, 1994.
- [27] Y. Lee, M. Lee, S. Park, C. Brosilow, "PID controller tuning for desired closed-loop responses for SI/SO system", *A.I.Ch.E. Journal*, 44(1), 106-115, 1998.
- [28] T. Liu, W. Zhang, D. Gu, "Analytical design of two-degree-of-freedom control scheme for open-loop unstable process with time delay." *J Process Control* 2005; 15:559-72.
- [29] M. Shamsuzzoha, M. Lee, "Analytical design of PID controller cascaded with a lead-lag filter for time-delay processes", *Korean J. Chem. Eng.*, vol. 26, pp. 622-630, 2009.
- [30] M. Shamsuzzoha, M. Lee, "Enhanced disturbance rejection for open-loop unstable process with time delay" *ISA Transactions*, Vol. 48, pp. 237-244, 2009.
- [31] T.B. Šekara, M. B. Trifunović, V. Govedarica, "Frequency Domain Design of a Complex Controller under Constraints on Robustness and Sensitivity to Measurement Noise", *Electronics*, Vol. 15, pp. 40-44, 2011.
- [32] M. Bošković, T.B. Šekara, Milan R. Rapaić, Boris Jakovljević, „Rational approximations to design controllers for unstable processes, including dead time“, pp. 200-204, INDEL, Nov. 2014.
- [33] F.G. Shinskey, "How good are our controllers in absolute performance and robustness?", *Measurement and Control*, vol. 23, pp. 114-121, May 1990.
- [34] M.Č. Bošković, T.B. Šekara, M. R. Rapaić, Vidan Govedarica, „Novel Tuning Rules for Stable Dead-Time Processes with Dominant Left Half-Plane Zero“, INDEL, 3-5 Nov. 2016.

# Sistemi za napajanje i hlađenje *green* data centra

Ivica Vuković, *Elektrotehnički fakultet Podgorica*

**Sadržaj** —Decenijama zanemarivana energetska efikasnost data centara, u posljednje vrijeme sve više privlači pažnju stručne javnosti. Visoki operativni troškovi i velika emisija CO<sub>2</sub> posledica su masivne potrošnje energije potrebne za napajanje i hlađenje data centara. Stoga se unapređivanjem sistema za napajanje i hlađenje data centara može značajno uticati na njihovu energetska efikasnost. U radu je predstavljena uporedna analiza postojećih rješenja u cilju prikaza optimizacije postojećih resursa, poboljšanja energetske efikasnosti i uspješnije redukcije gasova štetnih po okolinu.

**Cljučne riječi** — Hlađenje, UPS, Data centar, energetska efikasnost, 380 VDC, PUE, PSU.

## I. UVOD

**GREEN ICT** tehnologije sve više dobijaju na popularnosti u cilju smanjenja emisije CO<sub>2</sub>, očuvanja životne sredine i ublažavanja sve evidentnijih klimatskih promjena. ICT tehnologije pri svom radu koriste ogromnu količinu energije i odgovorne su za približno 2% globalne emisije CO<sub>2</sub>, što odgovara emisiji CO<sub>2</sub> u vazduhoplovnoj industriji [1]. Najveću odgovornost za ovakav vid potrošnje snose data centri čija kompleksnost onemogućava pronalazak adekvatnog *green* rješenja, odnosno kreiranje data centra koji je utemeljen na primjeni jedne jedinstvene tehnologije. Stoga je od izuzetne važnosti da se temeljno analiziraju i uporede najnovija tehnološka rješenja u cilju poboljšanja energetske efikasnosti.

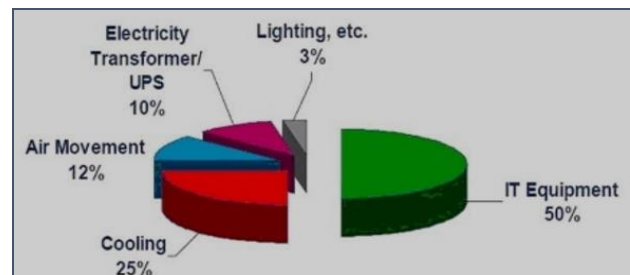
Decenijama unazad, fokus kod data centara usmjeren je na poboljšanje mogućnosti procesiranja, povećanje brzine, kvaliteta usluga, te razvoju cloud sistema što je za rezultat imalo masivnu potrošnju energije od strane IT opreme. Potrošnja nije bila uslovljena samo povećanjem računarske snage već i potrebom za snažnijom opremom za hlađenje IT komponenti. Energetska efikasnost data centara skoro da i nije postojala.

Veliki dio energije u okviru data centra koristi se u druge svrhe, tako da rashladna oprema troši skoro ekvivalentnu količinu energije kao i sami IT sistemi, dok prosječno oko 10% otpada na gubitke pri distribuciji i konverziji električne energije [2]. Otežavajući faktor implementacije *green* data centra predstavlja činjenica da veća potrošnja energije od strane IT opreme ide u prilog sistemima za napajanje koji tada funkcionišu sa većim stepenom iskorišćenja i manjim gubicima. Međutim, veća potrošnja energije rezultira većom količinom toplote koja mora biti efikasno odvedena, što dodatno opterećuje sisteme za hlađenje. Naime, jedan MW energije potreban

za rad današnjeg data centra obično zahtijeva dodatnih 0,7 MW za hlađenje generisane toplote [3].

U cilju postizanja maksimalne efikasnosti pri potrošnji energije i emisiji CO<sub>2</sub>, svaka od komponenti data centara mora biti dizajnirana i upotrijebljena efikasno.

Tri glavne cjeline (sl. 1) iz ugla potrošnje energije, kod data centara su: IT oprema, sistem za napajanje i sistem za hlađenje.



Sl.1. Potrošnja energije u data centru [2]

Mogućnosti uštede energije kod IT opreme ograničene su visokim kvalitetom servisnih usluga. Tehnologije koje se mogu ovdje primijeniti su: virtualizacija, deaktivacija neupotrebljenih servera, konsolidacija manje opterećenih servera i sl.

Ipak, značajna poboljšanja energetske efikasnosti treba tražiti u drugačijoj realizaciji sistema za napajanje i hlađenje u data centru [4]. U saglasnosti sa datom tvrdnjom u sledećem poglavlju biće razmatrane odgovarajuće tehnologije u cilju prikaza efikasnijih energetske rješenja.

## II. SISTEM ZA NAPAJANJE DATA CENTRA

Sistem napajanja kod data centara obično se sastoji od sistema prekidača, backup generatora, UPS-eva, opreme za napajanje uređaja i jedinice za napajanje servera (PSU). Sve ove cjeline tokom konverzije energije ne samo da stvaraju gubitke, već i oslobađaju toplotu, što predstavlja dodatno opterećenje za sistem hlađenja. Pažljivim odabirom i adekvatnim opterećenjem pomenutih komponenti može se značajno povećati efikasnost rada data centara.

### A. Nепrekidni sistem napajanja (UPS)

UPS tipično kombinuje tri funkcije u jedan sistem: vrši odabir aktivnog napajanja (distributivna mreža ili generator), sadrži neki oblik uskladištene energije (električna ili mehanička) i uklanja promjene napona ili harmonijske distorzije kod AC napajanja.

Trenutno najviše dominiraju UPS-evi dvostruke konverzije (AC-DC-AC), koji u cilju obezbeđenja prethodne tri funkcije su veoma energetske neefikasni. Izbjegavanjem jednog centralnog UPS sistema i primjenom više manjih može se postići veća efikasnost koja je zatim dodatno poboljšana pri konsolidaciji servera.

Kao efikasno rješenje nameće se rotacioni Flywheel sistem. U uporednoj analizi Flywheel sistema i statičkog sistema dvostruke konverzije dobijeni su sledeći podaci: pri djelimičnom opterećenju od 25% što je i uobičajen slučaj pri radu data centara, izmjerene vrijednosti su 81 i 76% na stranu Flywheel-a, dok pri nominalnom opterećenju od 100% efikasnost kod Flywheel-a je preko 95%, a kod sistema dvostruke konverzije 90% [5]. Velika prednost Flywheel sistema je što ne zahtijeva baterije za skladištenje energije i samim tim smanjuje količinu potencijalnog hemijskog otpada, te ukupan prostor potreban za smještaj sistema. Moguće uštede energije i uštede troškova održavanja na duži period su ogromne. Ukoliko uzmemo u obzir 10 MW sistem, za period od 20 godina tj. koliko iznosi životni vijek Flywheel sistema, uštede bi iznosile preko 20 miliona dolara [6].

Zahvaljujući naponu u EU od 400/230V, za razliku od nekih drugih djelova svijeta (SAD, Japan itd.), izbjegnuta je potreba za jedinicom za distribuciju napajanja kao i gubici, jer napon ispunjava standarde napajanja serverskog PSU (90–264 VAC).

### B. Serverski PSU i DC sistem napajanja

Zadatak PSU se ogleda u napajanju serverskih komponenti (CPU, ventilatori, memorije itd.) energijom pretvaranjem visokog AC napona u niski DC napon (3.3V, 5V ili 12V). Energetska efikasnost PSU uglavnom zavisi od njegovog opterećenja, temperature, broja električnih kola itd., te energija mora proći mnoge faze konverzije koje dovode do neefikasnosti.

PSU spada među najneefikasnije komponente u današnjim data centrima; mnogi data centri se i dalje isporučuju sa niskom kvalitetnim napajanjima sa efikasnošću od samo 60–70% [14].

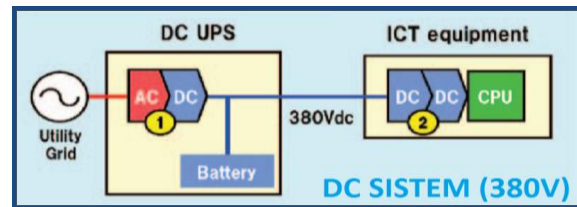
U cilju poboljšanja efikasnosti i obezbeđenja većeg i efikasnijeg faktora opterećenja, pri odabiru serverskog PSU veoma je bitno da mu nominalna snaga ne prelazi previše stvarne potrebe servera.

Nabavkom samo jednog rek PSU umjesto nekoliko serverskih, efikasnost bi mogla biti dodatno poboljšana. Kao dodatno unapređenje mogao bi se razviti algoritam koji bi konsolidacijom servera pri virtualizaciji vodio računa da se opterete oni rek ormari tj. njihova napajanja (i/ili UPS sistemi) koja su u tom trenutku najbliža nominalnom opterećenju.

Zamjena svih PSU sa ekvivalentnim Energy Star napajanjima koja garantuju minimum 80% efikasnosti pri bilo kom opterećenju, takođe se nameće kao jedno od mogućih rješenja [7].

Sledeće moguće rješenje predstavlja upotreba 380 VDC sistema koji pokazuje poboljšanje od 5-8% u odnosu na najbolji u klasi AC sistem [8]. Naime, 380 VDC sistem napajanja je energetski efikasniji jer postoje samo dva koraka pri pretvaranju energije na putu do ICT opreme tako da u poređenju sa UPS-om dvostruke konverzije, DC sistemi eliminišu završni korak inverzije UPS-a i korak ponovnog pretvaranja AC u DC na ulazu u PSU (sl. 2) [9].

DC sistemi osim poboljšanja efikasnosti, povećanja pouzdanosti sa smanjenim brojem komponenti omogućavaju i lakšu integraciju obnovljivih izvora energije u sisteme DC napajanja jer oni izvorno daju jednosmjerni napon (solarni paneli, gorivne ćelije i vjetroturbine) [10].



Sl. 2. Efikasnost kod DC sistema napajanja

DC oprema dostupna je na tržištu [11], ali troškovi su trenutno veći od slične AC opreme.

Naposletku, gubici energije od 1-3% mogu nastati u niskonaponskim rek kablovima jer veliki data centri mogu imati podignute podove dužine preko 100 m, tako da kablovi za napajanje mogu biti prilično dugi [12].

### III. SISTEMI HLAĐENJA KOD DATA CENTARA

Oblast gdje postoji mogućnost za značajna unapređenja, u cilju poboljšanja energetske efikasnosti data centara, predstavljaju sistemi za hlađenje. Pronalaskom načina za oslobađanje manje toplote od strane IT komponenti, rezultiralo bi i manjom količinom toplote sa kojom se mora nositi sistem za hlađenje.

Takođe, neka istraživanja su pokazala da neujednačena temperatura u data centrima takođe može dovesti do značajnog pada pouzdanosti IT sistema [7], što je čest slučaj posebno karakterističan kod sistema za vazdušno hlađenje. U potrazi za efikasnijim sistemom za hlađenje, imperativ je da data centar funkcioniše na odgovarajućoj temperaturi.

Sistemi za hlađenje se grubo dijele na sisteme vazdušnog hlađenja i sisteme hlađene tečnošću uključujući različite podsisteme i medijume za odvođenje toplote.

#### A. Vazdušno hlađenje

Dok je hlađenje tečnošću efikasnije za velike data centre, vazdušno hlađenje je dovoljno za one manjih dimenzija, ili one s značajno manjom količinom toplote koju je potrebno ukloniti.

Vazdušno hlađenje se uveliko upotrebljava u data centrima. Vazduh veoma lako cirkuliše direktno kroz prostor koji sadrži opremu, te na taj način može biti medijum za odvođenje toplote sa servera [13].

Poboljšanje ovog sistema može se ostvariti korišćenjem spoljašnjeg vazduha, primjenom varijabilnih ventilatora, postavljanjem jedinice za hlađenje što bliže serverskoj opremi i sl. [14].

Ipak, najjednostavniji način poboljšanja energetske efikasnosti ovog sistema jeste onemogućavanje miješanja toplog i hladnog vazduha. Veoma je važno da vazduh hladi vruću opremu a ne cijelu prostoriju, što je danas nažalost čest primjer. Moguće rješenje u cilju stroge izolacije hladnih od vrućih djelova prostorije je upotreba PVC barijera koje su veoma praktične, lako se montiraju i ostvaruju uštede energije do 15% jer smanjuju efekat recirkulisanja toplog vazduha [15].

Međutim, nedostaci vazdušnog hlađenja koji sve više usmjeravaju pažnju na alternativne sisteme su [16]:

- Vazduh je nevjerojatno loš provodnik toplote.
- Potrebni su široki kanali za usmjeravanje vazduha ka rek ormarima.
- Ventilatori troše veliku količinu energije, stvaraju buku i uvlače prašinu i druge nečistoće unutar serverskih kućišta.

- Efikasno hlade samo područja preko kojih se vazduh može nesmetano kretati.
- Potrebni su veliki hladnjaci koji zauzimaju ogroman prostor i ujedno su najveći potrošači u sistemu itd.

Jedna od najnovijih ideja predložena u literaturi je povećanje temperature data centara. Neki od vodećih proizvođača servera udruženi sa ekspertima za energetske efikasnosti data centara dijele mišljenje da data centri mogu funkcionisati pri daleko većim temperaturama nego što je to slučaj danas, bez ugrožavanja radne efikasnosti, i uz velike uštede vezane za troškove hlađenja i emisiju CO<sub>2</sub>. Naravno ulazna temperatura vazduha trebala bi biti povećana, tek nakon detaljne analize mogućih implikacija za svaki komad opreme.

Za svako povećanje ulazne temperature za jedan stepen moguće su uštede energije kod data centara od 2-4% [17]. Prvenstveno se misli na energiju potrebnu za rad hladnjaka, zatim potencijalno povećanje broja sati ekonomičnih režima rada itd. [18].

Povećanje temperature predloženo u [12] preporučuje ulazne temperature od 25-30°C, umjesto 18-20°C. Veće temperature bi uveliko olakšale efikasno hlađenje data centara. Praktično nema servera ili mrežne opreme kojima je potrebna ulazna temperatura od 20°C, i nema dokaza da veće temperature uzrokuju češće otkaze komponenti.

Prema Moore-ovom zakonu, sistemi hlađenja zasnovani na vazduhu neće moći još dugo efikasno da se nose sa toplotom generisanom od IT opreme, tako da čak ni *free cooling* spoljašnjim vazduhom neće predstavljati rešenje [19]. Upravo zato velike kompanije poput Microsofta, Google i Amazona, u poslednje vrijeme, naročito vrednuju hlađenje tečnošću velikih data centara.

#### B. Hlađenje tečnošću

Kada serverski ormari oslobađaju mnogo veću toplotu nego što to vazdušno hlađenje može podržati, na scenu stupa hlađenje tečnošću. Tečnost sama po sebi zbog svoje veće gustine i specifične toplote, efikasnije odvodi toplotu od vazduha [20].

Tehnologija koja sve više dobija na popularnosti je direktno hlađenje koristeći vodu ili druge rashladne tečnosti. Ideja kod direktnog hlađenja se ogleda u tome da se hlađenje obavlja što je moguće bliže izvoru toplote kako bi se povećala efikasnost hlađenja. Povećanjem efikasnosti, može se ukloniti više toplote primjenom iste ili manje količine energije. Direktno vodeno hlađenje procesora i memorije može ukloniti od 78% do 85% toplote generisane od strane servera u toku rada [21]. Efikasnije odvođenje toplote omogućava proširenje broja servera u postojećim prostornim kapacitetima.

Odličan primjer hlađenja toplom vodom predstavlja superkompjuter SuperMUC u Minhenu sa više od 241,000 jezgara, brzine preko 6.8 Petaflop/s i sa energetsom uštedom do 40% u odnosu na sisteme hlađene vazduhom [22].

Obično voda koja se koristi u data centrima ima ulaznu temperaturu oko 16°C, a nakon izlaska iz sistema oko 20°C. Da bi se voda ohladila na 16°C potrebna je složena i energetska zahtjevna oprema za hlađenje. U isto vrijeme jedva da postoji mogućnost bilo kakve upotrebe vode od 20°C, jer kao takva previše je hladna da bi se mogla upotrijebiti u bilo kom tehničkom procesu [23]. SuperMUC omogućava povećanje ulazne temperature

vode na 40°C dok voda na izlazu postiže temperaturu do 70°C i kao takva može se ponovo upotrijebiti u drugim tehničkim procesima (npr. za zagrijavanje zgrada) [23]. Ovaj sistem hlađenja ostvaruje uštede na godišnjem nivou od milion eura [24].

Dakle, tehnologija direktnog vodenog hlađenja može se koristiti za drastično smanjenje potrošnje energije.

Nova metoda, kao varijacija na tradicionalni sistem odvođenja toplote putem vode, koristi etapno hlađenje. Ova metoda rješava probleme koji mogu nastati kod prethodnih sistema za hlađenje na bazi vode u slučaju kontakta električne opreme sa vodom. Rashladna tečnost, koja inače zahtijeva manje energije za cirkulaciju unutar sistema, pri kontaktu sa sobnom temperaturom pretvara se u gasovito stanje tako da su potencijalna curenja svedena na minimum. Loša strana je da etapno hlađenje zahtijeva nove vodovodne instalacije [25].

Hlađenje potapanjem podrazumijeva redukciju toplote u hardveru kroz uranjanje u dielektričnu tečnost koja je toplotno provodna. Ova metoda se koristi skoro vijek u industriji električnih transformatora, međutim do nedavno sama dielektrična tečnost je bila veoma agresivna za elektronske komponente i veoma štetna po životnu sredinu.

Green Revolution Cooling (GRC) [26] je riješio ove probleme koristeći posebno mineralno ulje kao sredstvo za odvođenje toplote koje nije škodljivo za elektronske sklopove i koje je potpuno biorazgradivo. Iako mineralno ulje u poređenju sa vodom posjeduje samo 40% specifične toplote, ono je za razliku od nje električni izolator [27]. Osim GRC, prodavci poput 3M, Intel i SGI takođe nude ovaj tip sistema za data centre.

Prema [28] i [29] prednosti hlađenja potapanjem su mnogostruke:

- Poboljšanje pouzdanosti eliminacijom vrućih tačaka uzrokovanih neujednačenim protokom vazduha.
- Uštede do 95% u energiji za hlađenje.
- Jednako efikasno i u toplim i vlažnim predjelima.
- Izmjerena PUE vrijednost od 1.03.
- Potpuna eliminacija buke.
- Preko 15% manja potrošnja energije zbog eliminacije serverskih ventilatora i boljeg termalnog rukovanja.
- 25% manji troškovi ugradnje u poređenju sa tradicionalnim sistemima hlađenih vazduhom.
- Manje potrebe za prostorom, eventualna podrška većih servera i veće gustine data centara.
- Bezbjednija zamjena djelova "na živo" u slučaju otkaza itd.

Zbog svih ovih karakteristika, hlađenje potapanjem bi moglo zamijeniti konvencionalne metode hlađenja vazduhom ili vodom. Jedini ograničavajući faktor predstavlja drugačiji razmještaj opreme od postojećih sistema jer su rek ormari postavljani horizontalno u rezervoare. To uveliko ograničava njegovo usvajanje i dovodi do visokih inicijalnih troškova jer je potrebno odbacivanje postojeće infrastrukture i izgradnja nove.

Da li budućnost data centara leži pod morem? Microsoft je izvršio tromječno testiranje data centra uronjenog na dubini od 9m koji je hlađen okolnom morskom vodom. Održivi aspekt projekta na duže staze predstavlja ideja o povezivanju ovakvih data centara sa kopnom samo sa mrežnim kablom. Napajali bi se pomoću hidrokinetičke

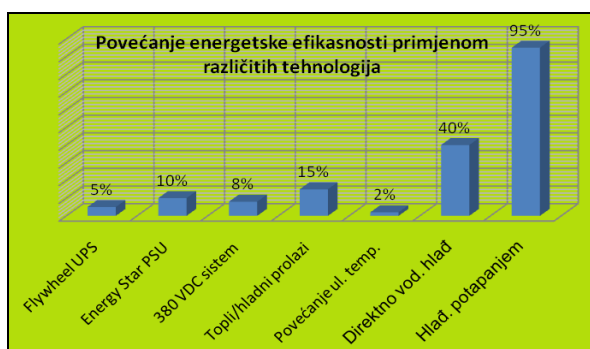
energije talasa ili plime, a bili bi hlađeni pomoću morske vode [19].

Ukoliko uporedimo prethodno navedene sisteme za hlađenje, hlađenje potapanjem izdvaja se kao najefikasnije i najodrživije rješenje sa najmanjom PUE vrijednošću (Tabela 1).

TABELA 1: KOMPARACIJA SISTEMA ZA HLAĐENJE

Sistem za hlađenje	Prosječan PUE
Vazdušni	1.8 [30]
Tečnošću (voda)	1.35 [30]
Tečnošću (potapanjem)	1.03 [28]

Komparacijom predstavljenih tehnologija sistema za napajanje i hlađenje dolazi se do rezultata koji direktno utiču na poboljšanje energetske efikasnosti i smanjenja emisije CO<sub>2</sub> (sl. 3).



Sl. 3. Procentualni prikaz poboljšanja energetske efikasnosti primjenom različitih tehnologija

#### IV. ZAKLJUČAK

Većina današnjih data centara je prilično neefikasna, jer je energetska efikasnost decenijama zanemarevana u odnosu na performanse, kvalitet servisa i kapitalne izdatke. Najčešći slučaj je da današnji data centri koriste samo 50% od ukupne energije za napajanje IT opreme.

Primjenom navedenih tehnologija uveliko se može unaprijediti energetska efikasnost data centara, smanjiti operativni troškovi i značajno smanjiti emisija CO<sub>2</sub>.

#### LITERATURA

- [1] S. Mingay, "IT's Role in a low carbon economy," presented at the *Gartner Symposium/IT Expo*, 2008.
- [2] B. Wehl, E. Teetzel, J. Clidaras, C. Malone, J. Kava, M. Ryan, "Sustainable data centers," in *Magazine XRDS: Crossroads, From Pollution to Pixels*, 2011, pp. 8-12.
- [3] A. Miletić, A. Radulović, B. Krstajić, "Zelena računarska učionica," in *XXI Međunarodni naučno-stručni skup INFORMACIONE TEHNOLOGIJE – sadašnjost i budućnost (IT'16)*, 2016, pp. 121-124.
- [4] Z. Wang, X. Wang, M. Li, "Research and Application of Green Power System for New Data Centers," in *International Telecommunications Energy Conference (INTELEC)*, 2015.
- [5] M. Milad, M. Darwish, "UPS System: How Can Future Technology and Topology Improve the Energy Efficiency in Data Centers?," in *IEEE Power Engineering Conference (UPEC)*, 2014.
- [6] K. Fehrenbacher, Flywheels, Available: <https://gigaom.com/2011/10/31/what-you-need-to-know-about-flywheels/>
- [7] S. Kumar, R. Buyya, "Green Cloud Computing and Environmental Sustainability," in *Harnessing Green It: Principles and Practices*, IEEE Computer society, 1st edition, 2012, pp. 315-339.

- [8] D. Pratt, A. Kumar, "Evaluation of 400V DC Distribution in Telco and Data Centers to Improve Energy Efficiency," in *International Telecommunications Energy Conference (INTELEC)*, 2007.
- [9] H. Yajima, T. Babasaki, "Energy-saving and Efficient use of Renewable Energy by introducing an 380 VDC Power-supply System in Data Centers," in *International Telecommunications Energy Conference (INTELEC)*, 2015.
- [10] D. Demetriou, "Energy Efficiency and Reliability Transformation at the IBM India Software Lab Data Center," in *15th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm)*, 2016.
- [11] IBM zEnterprise EC12 Technical Guide, Available: <http://www.redbooks.ibm.com/abstracts/sg248049.html>
- [12] L. A. Barroso, J. Clidaras, U. Hözl, *The Datacenter as a Computer*, in the Morgan & Claypool Publishers series, 2nd edition, 2013, pp. 70–71.
- [13] A. Woods, "Cooling the data center," in *Communications of the ACM*, vol. 53, no. 4, 2010, pp. 36–42.
- [14] I. Metzger, O. V. Geet, "Data Center Energy Efficiency and Renewable Energy Site Assessment: Anderson Readiness Center," National Renewable Energy Laboratory (NREL) Golden CO, 2014.
- [15] Ekološki zeleni data centar, Available: <http://www.altus-it.hr/o-nama/ekoloski-zeleni-data-centar>
- [16] J. Wallace, Immersion liquid cooling small form factor and server-class systems, Available: <http://smallformfactors.mil-embedded.com/articles/immersion-form-factor-server-class-systems/>
- [17] R. Vokoun, Energy Efficiency in Today's Data Center, Available: <http://www.datacenterknowledge.com/archives/2012/06/28/energy-efficiency-todays-data-center/>
- [18] D. Moss, J. Bean, "Energy Impact of Increased Server Inlet Temperature," *APC by Schneider Electric - White Paper 138*, 2009.
- [19] Microsoft testing undersea immersion cooling, Available: <https://submer.com/microsoft-testing-undersea-immersion-cooling/>
- [20] M. K. Patterson and D. Fenwick, "The state of datacenter cooling," in *Intel Corporation White Paper*, 2008.
- [21] L.A. Parnell, "Combining Cooling Technology and Facility Design to Improve HPC Data Center Energy Efficiency," in *15th IEEE ITherm*, 2016.
- [22] J. Buchholz, Y. Sandoval, E. Volk, "Increase energy efficiency of data center cooling facilities with airflow optimization," in *IEEE International Conference on Systems, Man, and Cybernetics*, 2013.
- [23] <https://www.lrz.de/services/compute/supermuc/systemdescription/>
- [24] IBM's SuperMUC - World's 1st Hot-Water Cooled SuperComputer, Available: <https://www.youtube.com/watch?v=i8VoJgUPxvY/>
- [25] G. Schulz, *The Green and Virtual Data Center*, by Auerbach Publications CRC Press, 2009, p. 143.
- [26] "Green revolution cooling", Available: <http://www.grcooling.com>
- [27] Liquid immersion cooling, Available: <http://whatis.techtarget.com/definition/liquid-immersion-cooling/>
- [28] D. Varma, "Submerged servers in the data center: not as new or nearly as scary as we think," by Green revolution cooling, 2014.
- [29] Immersion Cooling for Data Centers, Available: <https://submer.com/#solution/>
- [30] B. Plankers, "Liquid cooling drives down costs and improves efficiency," Available: <http://searchdatacenter.techtarget.com/opinion/Liquid-cooling-drives-down-costs-and-improves-efficiency/>

#### ABSTRACT

Decades held in back seat, data center's energy efficiency, recently increasingly attracting the attention of experts. High operating costs and high CO<sub>2</sub> footprint are the result of massive consumption of energy in order to power and cool data centers. Therefore, improving of the data centers power and cooling systems can significantly affect their energy efficiency. This paper presents a comparative analysis of existing solutions in order to display the optimization of existing resources, improvement of energy efficiency and the successful reduction of gases harmful to the environment.

#### GREEN DATA CENTER POWER AND COOLING SYSTEMS

Ivica Vukovic

# Barriers to adoption and success of SCRUM in practice

Snežana Šćepanović, Ivan Mraković

**Abstract** — Agile software development methods were created as a reaction against documentation driven, heavyweight traditional ways of developing software. The goal of Agile methodologies is to overcome shortfalls and deficiencies of traditional methods and to allow often and quick changes in software development process. While Agile techniques vary in practices and emphasis, they share common characteristics, including iterative development and a focus on interaction, communication and the reduction of resource-intensive intermediate artifacts. The aim of this paper is to present challenges and obstacles in implementation of Agile software development methodologies in practice focusing on Scrum as one of the widely used Agile methodologies.

**Key words** — Agile software development, Scrum, Software development life cycle

## I. INTRODUCTION

**M**ULTI-MODALITY is perfect term to describe state of IT world looking from software application development perspective. When it comes to software development life cycle (SDLC) models, companies tend to use different ones depending on their scope of business and type of project being developed. There is no universal model which can suit everyone's needs. SDLC models are made in such a way to break the problem of project development in smaller parts which are easier to solve. Each SDLC method - traditional or Agile consists of a few common phases: requirement gathering and analysis, design, implementation or coding, testing, deployment and maintenance.

Agile SDLC methods are developed in order to complement shortfalls and deficiencies of traditional ones. According to VersionOne report [1], 53% of respondents state that less than a half of the teams in their organization use Agile, while only 9% state that everyone uses Agile. Hence, 34% are organizations where majority of their teams are Agile and just 4% are non-Agile or traditionally orientated. Obviously, there is multi-modal environment where various SDLC methods are still in practice.

Companies from different industries tend to adopt Agile methods simply because of clear benefits. Agile places a focus on speeding up the product delivery process by allowing teams to better manage changing priorities (87%),

Snežana Šćepanović, Faculty for Information Technologies, University "Mediterranean", Josipa Broza bb, 81000 Podgorica, Montenegro (tel.: 382-67-235915, e-mail: snezana.scepanovic@unimediteran.net)

Ivan Mraković, Faculty for Information Technologies, University "Mediterranean", Josipa Broza bb, 81000 Podgorica, Montenegro (tel.: 382-69-451428, e-mail: ivanmr@ac.me)

while there is increase team productivity (85%) and improved project visibility (84%) [1]. Even though Agile methods usually create benefits, "Agile adoption always comes with special challenges and accordingly, fundamental organizational changes are necessary for successful outcome" [2, p.1]. Literature research presented in this paper shows that there are various issues and challenges which can obstruct implementation of Scrum as leading Agile methodology. The paper is organized as follows: chapter II represents comparison between Agile and traditional SDLC models; chapter III describes Scrum while chapter IV presents analyses of common obstacles for Scrum implementation. Finally, chapter V presents conclusion.

## II. AGILE VS TRADITIONAL SDLC MODELS

In Agile, software development is carried out through short sequences – iterations. Fig. 1 shows framework of Agile development. It is clear that each iteration covers all SDLC phases. Each phase ends with increment or, simply, deliverable product. At that stage, customers have a chance to change their requirements, to give a feedback and to collaborate in order to get a better product at the end of the following iteration.

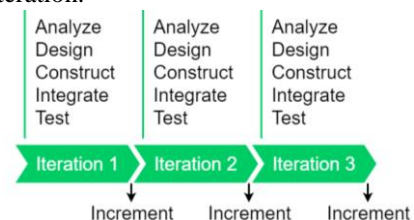


Figure 1. Agile framework (adopted from [3, p.9]).

Main difference between traditional SDLC models (e.g. Waterfall) and Agile approach of software development is in the path of developing. Actually, in Waterfall, there is a strict path which allows moving to the next SDLC phase not before previous one is finished. At the end of phase there are documents or codes but there is no deliverable product. Deliverable product is visible only at the end of the project. Creating predictive SDLC phases and defining clear borders between them is exactly what traditional approach do. Despite long list of deficiencies and lack of efficiency, traditional models are still in use nowadays. Therefore, it is more rarely because customers usually do not know exactly their requirements at the beginning of software development. Due to uncertainty at the start of the project, customers tend to change their requirements later on, causing chaos in developing process. Agile SDLC



models are more sensitive and quicker when it comes to software requirements changes. According to Agile Manifesto [4], advantages of Agile over traditional methods are as follows:

- Individuals and interactions over processes and tools;
- Working software over comprehensive documentation;
- Customer collaboration over contract negotiation;
- Responding to change over following a plan.

Constant chance for remediation is main feature of Agile developing. Discovering defects early is primary benefit of iterative over sequential methods [5]. It is emphasized with comprehensive testing which leads to early discovery of risks and issues of the project.

Katic in [6, p.3] states that the main advantages of Agile over traditional methods are as follows:

- Shorter software developing time up to 80% - shorter delivery of product;
- Greater stability and quality of the system thanks to customers' feedback at early stages;
- Better human resource management;
- More flexible to requirements changes.

In traditional approach, team rules are clearly defined and easily distributed because of strong documentation and clear requirements while in Agile and Scrum, team members have to keep strong collaboration and to complement each other during iterations in order to solve problems en route. It can be embarrassing if team is too big [6]. The commitment of Agile team members must be greater than traditional ones what is considered as one of disadvantages of Agile SDLC models.

Disadvantages of Agile methods are as follows [7]:

- Less emphasis on documentation;
- Needs excellent and experienced developers;
- Very limited planning is done to start the project.

Various Agile SDLC models are used in practice. According to research [1] mostly followed model is Scrum which is practiced by 58% of respondents. Scrum/XP Hybrid as Agile method is practiced by only 10% of teams, followed by Custom Hybrid of multiple methodologies (8%), Scrumban (7%) and Kanban (5%).

### III. SCRUM FRAMEWORK

Scrum is the most used Agile SDLC framework according to researches published by VersionOne and ScrumAlliance [1], [8]. Both reports are carried out globally through various industries such as IT, financial, healthcare, government, insurance and transportation. Based on research there is obviously a dominant usage of Scrum as SDLC model. As ScrumAlliance reports in [8], 42% of respondents use only Scrum in their software developing process while 54% use Scrum in combination with another method. In [1] those values are higher, ranging from 58% of solo-Scrum usage to 68% of Scrum combined with other methods. "Jeff Sutherland developed

the Scrum process in 1993 while working at the Easel Corporation, and Ken Schwaber formalized the process in the first published paper on Scrum at OOPSLA 1995" [9, p. 220].

In Scrum, software developing is carried out through Sprints where clearly defined functionalities are added. During Sprints, which usually last from 2-4 weeks, change of requirements is not allowed. Once the Sprint is completed, customers and stakeholders are welcomed to change their requirements or set the new ones. Scrum as a framework functions thanks to strong collaboration of team members and heavy impact of customers. Therefore, it is dynamic, quick and Agile process.

Scrum consists of development team, events, artifacts and rules set among them. Concept of this method is shown in Fig. 2.

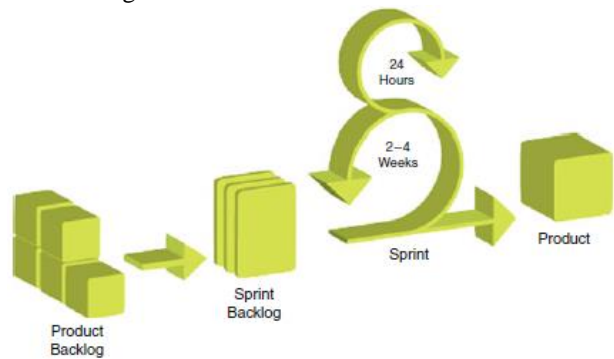


Figure 2. Scrum concept [9, p. 221]

"Scrum relies on self-organizing, empowered teams to deliver the product increments, but also on a customer, or Product Owner, that must provide the development team with a list of desired features, using business value as the priority mechanism" [9, p. 220]. Scrum team consists of all parties involved in project. Apart from customers and stakeholders, it is formed by Product owner, Scrum master and development team.

Development team has freedom to organize itself and collocate activities among its members in order to achieve greatest values from backlogs. Activities of developers are usually overlapped and in case of failure no single person should be blamed for it but the team as a whole. Optimum size of development team is 3-9 members [3]. Because of that, Scrum master has mainly managerial rule reflected in his responsibility to lead other team members in Agile direction towards common goal. Key difference between Scrum master and Product owner is in the fact that Product owner is directly involved in formulation and prioritizing of requirements as well as into shaping of increments.

Scrum as a framework is ideal for delivering product versions – releases at the end of each Sprint. Every feature which has to be accomplished by increment is defined and prioritized in Product backlog. Product owner is in charge of Product backlog and is responsible for gathering requirements usually in form of user stories, e.g. "as a <role>, I want<a feature>, so that I can <accomplish something>" [10, p. 8]. „All user stories intended to be developed during whole project are set into Project

Backlog, while all user stories intended to be developed during next Sprint are set into Sprint Backlog” [11, p. 470].

Developing progress is checked on daily basis during Daily scrums. It is short and time limited meeting organized in order to summarize what has been done and set what has to be done until next Daily scrum. In this mode, Scrum master can evaluate progress and suggest adjustments in order to ensure satisfactory completion of current Sprint. At the end of each Sprint, main difference between Agile and traditional approach is visible in form of releasable product increment. Based on release state, backlogs are updated for next Sprint.

#### IV. SCRUM IMPLEMENTATION BARRIERS

Implementing and maintaining Scrum throughout company is endless cycle of exploration, implementation and releases. Scrum methodology is not easy to adopt. In organization where traditional model is employed, transferring to Scrum can be a real challenge. Without clear understanding of what Scrum entails, there is a reasonable risk of project development failure.

Company culture continues to dominate the top causes of failed Agile projects with “company philosophy or culture at odds with core Agile values” at 46%, and “lack of management support for cultural transition” at 38% together with “lack of support for cultural transition” and “inconsistent Agile practices and process” at the same rate [1]. “Lack of experience with Agile methods” was chosen by 46% of respondents as the second most important reason for Agile failure. Based on research, six most common impediments are extracted and shown in Tab. 1.

TABLE 1: MOST COMMON SCRUM IMPLEMENTATION BARRIERS.

<i>Barrier</i>	<i>Source</i>
Ability to change organizational culture	[1], [8], [12-14]
General organizational resistance to change	[1-2], [12-14]
Pre-existing rigid/waterfall framework	[1-2], [8], [14]
Not enough personnel with the necessary Agile experience	[1-2], [12], [14]
Management support	[1], [8], [14-16]
Business/user/customer availability	[1], [12-17]

Issues shown in Tab. 1 can be further classified. Lopez-Martinez et. al in [14] distinguish four groups of issues, among which two are organization and people referred.

Group of organizational issues consists of:

- Ability to change organizational culture;
- General organizational resistance to change;
- Pre-existing rigid/waterfall framework;
- Management support.

Ability to change organizational culture is listed as top priority barrier in [1] with the score of 55%, and the 2<sup>nd</sup> most important obstacle in [8] with 46%.

As reported in [12], organizational culture is of great concern, as it can significantly impact people who work on the project. In that case, domino effect is created and people start looking their own benefits rather than team ones. Cite in [12, p. 261] perfectly describes difference between anti-Agile culture orientated company and culture how it should look like in Scrum: “The culture...is more on Control and Competences, Scrum is more on Cooperation and Cultivation”. Scrum should be used in organization which culture is Agile orientated, capable to change the work style and capable to induce continuous collaboration and cooperation among all members. Normally, improvements always should start from top level. Leaders have great influence on culture and they are those who can make the things moving towards Scrum. If they do that, others in the team will follow the model. Otherwise, Scrum failure is obvious at the very beginning.

Organizational resistance to change is tagged as organizational issue but it is tightly connected with people because they tend to behave routinely and once they adopt to everyday routine, they resist to change. People are afraid of losing their current place in development process [2], [14]. Dave [13] states that Scrum teams need to change the way of work because Scrum means continuous change. If the company is struggling with Scrum implementation due to resistance to change, than team members should become cross-functional. They have to be located together and close to each other. Otherwise, agility will be impeded. “Customers cannot make their requirements known, because those requirements are constantly changing” [13, p. 19]. Scrum teams have to embrace change. In that manner, even if the requirements are changed later in the project, they can be satisfied. So, if organization needs change, it should start from people because sometimes even top members behave antagonistically towards changes.

Scrum as other Agile SDLC methods are implemented in order to redeem disadvantages of traditional methods. If the company uses Waterfall framework, transition to Scrum can be difficult. It is 3<sup>rd</sup> most common problem in Scrum implementation as reported in [1] with 40% of prevalence among respondents. As explained in ch. II, organization and rules in Waterfall is completely different from Scrum. In [2] is described transition process from Waterfall to Scrum with all challenges encountered. Actually, people face great challenges in their new rules. Transition takes time. It is very hard for people to bear in mind that there should be Scrum Master as team leader without ability to directly act into the process. There is no Project Manager but Product Owner and documentation is minimized. Even if organization decide to reorganize and perform Scrum, it will often find human related obstacles. Team members can find themselves in completely new environment experiencing obligations to finish small tasks during time limited Sprints and documenting user stories in different way. Changing to Scrum means introducing Daily Scrum. Skipping it may cause transition failure due to mistakes in implementation. In 52% of cases, non-

identification of metrics for measurement of Scrum success is why transition process fails [8].

Management support, both technically and financially, is recognized as one of the most important impendent for Scrum implementation [1], [8]. Scrum idea is conceived on self-managed teams but it still needs strong leadership and support from top level management. A lack of understanding of the Scrum reduces the autonomy of the Scrum team. “When Management does not fully understand the Scrum Processes, they are likely to impose large numbers of requirements onto the team which than feels highly pressurized, yielding sprint interruptions” [16, p.8]. Management should be tightly connected with all parties involved in the project. At first, management should support Scrum Master in making good cooperation between Scrum team and Product Owner. In case that customers are continuously setting new requirements, Product Owner should set the team free of that pressure, solving things out on his own. But, it is only possible if “short-term planning and frequent information exchange within the team” [15] are supported from top level management.

Lack of experience and knowledge is another reason of Scrum failure. In [1] this issue is listed on 4th place with frequency of 39%. In [12] Scrum coaches and consultants criticize team members who impede normal Scrum process flow due to their lack of knowledge. “This resulted in making the Agile team spending most of the time arguing about what the books say about the process rather than implementing what the process say” [2, p.2]. Therefore, as in any other process, not experienced team is not able to make a project of the same quality as experienced one. The same goes to team members who are not educated in comparison to educated members. A great number of software projects fail due to poor communication with users in requirement analysis and testing. For Scrum the communication style with users is more significant than the frequency of communication [14].

## V. CONCLUSION

This paper presents an analytical overview of most common barriers in implementation of Agile software development methods in practice focusing on Scrum as one of the widely used Agile methodologies. Presented barriers should be considered during implementation of Scrum especially if the company is new to Agile. Agile methodologies provide an organization or a team with the flexibility to adopt a selected subset of principles and practices based on their culture, their values, and the types of systems that they develop. Organizations that want to be Agile need to provide enough personnel with the necessary Agile experience and create environment that facilitates business/user/customer availability and rapid communication with software development project team members. The research shows that lack of ability to change of organizational culture is one of the main obstacles in implementation of Scrum.

While conducting our research, we found that there is a

lack of literature on techniques to compare the Agile method's effectiveness. It is difficult to distinguish between methods in respect to adequacy and capability of the organization to support the adopted principles and practices specified by the method. Future research will be focused on developing adequate techniques and indicators for assessment of effectiveness of Agile methodologies in practice.

## LITERATURE

- [1] VersionOne. (2016). 10th Annual State of Agile Report. [Online]. Available: <https://versionone.com/pdf/VersionOne-10th-Annual-State-of-Agile-Report.pdf>
- [2] H. Hajjidiab and A.S. Taleb, “Adopting Agile Software Development: Issues and Challenges”, *International Journal of Managing Value and Supply Chains (IJMVSC)*, ISSN: 0976 - 979X, vol. 2 (3), p. 1, 2011.
- [3] N.K. Rad and F. Turley. (2014). *EXIN Agile Scrum Foundation Workbook*. [Online]. Available: <https://www.scribd.com/doc/263185194/EXIN-Agile-Scrum-Training-Manual>
- [4] K. Beck, M. Beedle, A.V. Bennekum, A. Cockburn, W. Cunningham, M. Fowler, J. Grenning, J. Highsmith, A. Hunt, R. Jeffries, J. Kern, B. Marick, R.C. Martin, S. Mellor, K. Schwaber, J. Sutherland and D. Thomas. (2001). Agile Manifesto. [Online]. Available: <http://agilemanifesto.org/iso/en/manifesto.html>
- [5] B. Hipps and M. Sarbiewski. (2010). *The Applications Handbook: A Guide to Mastering the Modern Application Lifecycle*. Hewlett-Packard Development Company, L.P. USA. [Online]. Available: <http://static.ziftsolutions.com/files/8adc69543be1b3f7013c6d9558431dd2.pdf>
- [6] T. Katić. (2009). “Trendovi razvoja aplikacija”, Available: <http://161.53.18.5/zpr/Portals/0/Predmeti/MIT/Trendovi%20razvoj%20aplikacija.pdf>
- [7] R.C.G. Dhanajayan and S.A. Pillai, “SLMBC: spiral life cycle model-based Bayesian classification technique for efficient software fault prediction and classification”, *Soft Computing*, ISSN: 1432-7643, vol. 20 (8), pp. 1-13, 2016.
- [8] ScrumAlliance. (2015). The 2015 State of Scrum Report. [Online]. Available: <https://www.scrumalliance.org/landing-pages/2015-state-of-scrum-report-download>
- [9] M. Cossentino, V. Hilaire, A. Molesini and V. Seidita, *Handbook on Agent-Oriented Design Processes*. ISBN 978-3-642-39974-9, Heidelberg: Springer-Verlag, 2014.
- [10] C. Sims and H.L. Johnson, *Scrum: a Breathtakingly Brief and Agile Introduction*. Dymaxicon, 2012.
- [11] I. Padavić, M. Velić and D. Ljubobratović, “Agilni razvoj programskih proizvoda”, in *Proc. INFOTEH-JAHORINA*, Jahorina, 2011, pp. 466-470.
- [12] T.J. Gandomani and M.Z. Nafchi, “Agile transition and adoption human-related challenges and issues: A Grounded Theory approach”, *Computers in Human Behavior*, ISSN: 0747-5632, vol. 62, pp. 257-266, 2016.
- [13] D. McKenna, *The Art of Scrum: How Scrum Masters Bind Dev Teams and Unleash Agility*, ISBN: 978-1-4842-2277-5, Aliquippa: CA Technologies, 2016.
- [14] J. López-Martínez, R. Juárez-Ramírez, C. Huertas, S. Jiménez and C. Guerra-García, “Problems in the Adoption of Agile-Scrum Methodologies: A Systematic Literature Review”, DOI: 10.1109/CONISOFT.2016.30, in *Proc. 4th CONISOFT*, Puebla, 2016, pp. 141-148.
- [15] T. Dingsøy, T.E. Fægri, T. Dybå, B. Haugset and Y. Lindsjörn, “Team Performance in Software Development: Research Results versus Agile Principles”, *IEEE Software*, ISSN: 0740-7459, vol. 33 (4), pp. 106-110, 2016.
- [16] M. Tanner and A. Mackinnon, “Sources of Interruptions Experienced During a Scrum Sprint”, *The Electronic Journal of Information Systems Evaluation (EJISE)*, ISSN: 1566-6379, vol. 18 (1), pp. 3-18, 2015.
- [17] I. Inayat, S.S. Salim, S. Marczak, M. Daneva and S. Shamshirband, “A systematic literature review on agile requirements engineering practices and challenges”, *Computers in Human Behavior*, ISSN: 0747-5632, vol. 51, pp. 915-929, 2015.

# Cloud computing solutions for healthcare system in ME: challenges and issues

Bojana Tošić, *Institute of Modern technology, Montenegro*; Dragan Đurić, *Institute of Modern technology, Montenegro*; Radovan Sekulić, *MG Soft d.o.o. Montenegro*; Jelena Simićević, *MG Soft d.o.o. Montenegro*

**Abstract:** This paper describes main challenges and critical factors related to introduction of Cloud computing into Health Information System in Montenegro (ME), presented as intermediate findings and results of the ongoing IPA project “Modern technology development for Health Information System in ME” (HISTech).

**Keywords:** Cloud, eHealth, Healthcare, HIS.

## I. INTRODUCTION

Cloud computing is on its way to becoming increasingly powerful tool for extending IT resources and accessibility of IT services in all areas, including Healthcare. Providing more accessible healthcare services to general public, using innovative ICT technologies with reduction of costs associated to traditional means of provision of Healthcare services, has been recognized in national and European eHealth policies and strategies [1][2]. Availability and ease of access is especially important for wide range of primary healthcare services that are essential to general public. The reason for recent expansion of cloud based software solutions in healthcare industry could be found in raised complexity of traditional Health ICT, evolving to become more interactive and patient oriented with more focus on the quality of care.

Introduction of cloud computing introduces many opportunities to healthcare service delivery. Cloud technology enables more efficient usage of application software and data storage, facilitating daily procedures and operations of health protection in a flexible and scalable way through effective Service-Level-Agreements (SLAs) (as in “pay as you use” contracts). [3]

Although, cloud computing demonstrates clear benefits in availability, service delivery and resource allocation

there are some reservations that need to be presented and addressed, mostly having in mind technical requirements (security, privacy, integration, interoperability, etc.), national legislation requirements, financial and other requirements.[4]

Mostly all healthcare services are based on and operate with confidential patient medical and personal information. Security and privacy in storing data on cloud is the primary concern related to the confidential data. National legislation on personal citizen data (including medical records) sets limiting factors on using and especially on storing of confidential personal data on the Cloud. Also, data in the Cloud can be stored on various international locations, where different laws on personal data apply.

Large number of healthcare services require significant exchange of health related information in real time. Databases of diseases, medicines and large number of codebooks are used in various healthcare services, and are already supported by existing Integral Healthcare Information System of ME. This requires high level of integration of Cloud services with already existing data sources in healthcare.

Interoperability with other software solutions (Enterprise resource planning systems - ERP, Laboratory Information Systems – LIS, Radiology Information Systems-RIS, etc....) is also very important factor when considering implementation of Cloud computing.

Several of the perceived advantages and drawbacks of Cloud computing are specific to type of services that need to be supported. Issues that are relevant to one service type may not be seen as important for other service types. This is extremely relevant to possible implementations of Cloud technology in Healthcare regarding legal, privacy, security and interoperability requirements for some of the service types and availability and ease of access for the other types. [5] All stated issues, challenges and critical factors to overcoming those challenges are examined during the course of a project “Modern technology development for Health Information System in ME”-*HISTech*, and presented in this paper by the project partners. Overall objective of the project is to create and prototype innovative model of technology enhanced health information system in ME, on the bases of comprehensive market oriented research, current EU trends and practices

Bojana Tošić, Institute of Modern Technology, Cetinjski put 38/2, 81000 Podgorica, Montenegro (phone: 382-20-648177; e-mail: [bojana.tosic@imtm.me](mailto:bojana.tosic@imtm.me)).

Prof. Dr. Dragan Đurić, Institute of Modern Technology, Cetinjski put 38/2, 81000 Podgorica, Montenegro (phone: 382-20-648177; e-mail: [office@imtm.me](mailto:office@imtm.me)).

Radovan Sekulić, MG Soft d.o.o. Mitra Bakica bb fah109, 81000 Podgorica, Montenegro (phone: 382-67-260520; e-mail: [radovan.sekulic@mg-soft.co.me](mailto:radovan.sekulic@mg-soft.co.me)).

Jelena Simićević, MG Soft d.o.o. Mitra Bakica bb fah109, 81000 Podgorica, Montenegro (phone: 382-67-260580; e-mail: [jelena.simicevic@mg-soft.co.me](mailto:jelena.simicevic@mg-soft.co.me)).

in the field of cloud computing and eHealth and existing Integral Healthcare Information System in ME.

This paper is focused on making analyses of key challenges and issues moving towards development of cloud based solutions in health care system in ME. The paper is organized as follows, Section II gives overview of EU eHealth trends and strategies, Section III describes Integral Healthcare IS in ME, while Section IV presents eHealth cloud computing challenges with focus on ME. Section V concludes the paper with summary findings and directions for the future work

## II. EU eHEALTH TRENDS AND NATIONAL STRATEGIES

e-Health is recognized as emerging commercial ICT model that provides opportunities to solve some of the current limitations facing traditional Healthcare ICT solutions.

European Commission (EC) in the adopted strategy “Unleashing the Potential of Cloud Computing in Europe” confirms that “The Commission aims at enabling and facilitating faster adoption of cloud computing throughout all sectors of the economy which can cut ICT costs, and when combined with new digital business practices, can boost productivity, growth and jobs”. [1] To this end, the EC in the Report of the progress for 2014 noticed that “the next two years, will be developed and put into place will lay the foundation for Europe to become a world cloud computing powerhouse“, and makes a prediction that a rapid take-off phase (2014-2020) could bring the achievement in using of publicly available cloud computing offerings to 38% compound annual growth rate (around double the rate that would be achieved if the decisive policy steps are not implemented). [1]

The EC strongly recommends prior analysis of the overall policy, regulatory and technology landscapes and a wide consultation of stakeholders, to be undertaken to identify what needs to be done to achieve the goal, and calls upon industry to cooperate closely on the development and adoption of common standards and interoperability measures.

ICT society in ME also recognized potentials and importance of moving to clouds, but any concrete steps have not yet been taken. This is why the World Bank defined “Cloud computing: platform for development and innovations” in their Economic Report No.73826-ME related to ME preparation for Prosperity. [6] National policy related to ICT Healthcare is implemented by Ministry of Health of ME in strategy document ‘Healthcare politics in ME to 2020’. National Healthcare policy states that: “integral health information system shall be one of the primary priorities in health, in order to provide quality support for better management, improving quality of work and technological development.” [2]

## III. INTEGRAL HEALTHCARE IS OF ME

Integral Healthcare Information System - HIS of ME is a cornerstone IS that connects primary healthcare centers,

hospitals, pharmacies, public health laboratories and all relevant healthcare institutions to facilitate the secure electronic exchange of clinical information. [7]

It’s implemented in all Primary Health Centers in ME and in seven general hospitals from the secondary level of healthcare.

The patient is placed at the center of the system with evidence considering all the treatments, services and drugs that were provided by the system. This implies standardization and support of the processes “end-to-end” from the registration in Health Insurance Fund through the choice of a physician and getting all needed services and medicines from health care providers and pharmacies.

Fundamentals for developing Integral HIS are: (i) Database of insures (over 95% of ME citizens); (ii) Database of employers (obligate payers); (iii) Database of doctors, pharmacists, dentists; (iv) Database of diagnosis (International Classification of Diseases, revision 10 – ICD10); (v) Database International Classification of Primary Healthcare, revision 2 (ICPC2) for primary healthcare; (vi) Database of medicaments (Anatomical, Therapeutic, Chemical -ATC, classification system); (vii) Database of healthcare institutions; (viii) Large number of code books and parameters.

Development of Integral HIS started from the early 2000, with development of mentioned fundamental databases and codebooks, IS for pharmacies in 2004, HIS at primary health level in 2009 followed by HIS on secondary level of healthcare in 2010, Emergency Medical Service Institute Information System in 2011, Agency for Drugs and Medical Devices Information System in 2012, Institute of Public Health Information System in 2012 and National Transfusion Medicine Information System in 2014.

The essence of integrity of the system is reflected in the fact that it is implemented in a whole country (see Fig.1) and in the fact that the data once entered into the system, becomes available in a proper way to all segments of the system.

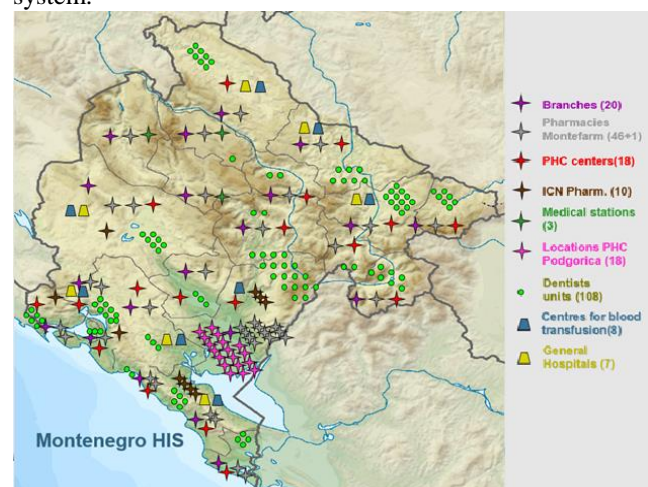


Fig.1. Implementation of Integral HIS of ME on country level

For example, health insurance data of any insured person generated in the Health Insurance Fund database becomes available to all other subjects in the system without need for rewriting it. This approach reduces

possibility for mistakes and makes easy collection of information in the system [7] Approach to develop Integral HIS on a country level has been recognized as right approach to overcome information related problems in healthcare identified in the most developing countries. [3]

Having in mind suggestions and recommendations of EC, as well as current level of development of integrated healthcare information system at national level, it is a right moment to make appropriate analyses of potentials for going one step further integrating Cloud computing technologies into ME Healthcare.

#### IV. EHEALTH CLOUD COMPUTING CHALLENGES

Main challenges in the use of Cloud technology in Healthcare include:

- Technical challenges: (i) Security and privacy; (ii) Performance and service quality; (iii) Reliability and availability; (iv) Integration and interoperability
- Legal and Policy challenges

##### A. *Security and privacy*

The main challenge in implementation of cloud technology in Healthcare, is how it addresses the security and privacy concerns. The fact that the sensitive and confidential medical and personal data about patients will reside outside the traditional infrastructures (on-premise deployments) with firewalls raises serious concerns.

Various attacks and hacking on cloud infrastructure would affect multiple clients, even if only one client is targeted. Also, administration interfaces of public cloud providers are Internet accessible and provide access to larger sets of resources (than traditional hosting providers) and therefore pose an increased risk (especially when combined with remote access and Web browser vulnerabilities). [4] Advanced encryption including key management is one of the basic mechanisms that cloud service providers use to guard against data loss and theft. Also, it is important to point out that a weak key management procedure may cause loss of encryption keys, exposing secret keys or passwords to malicious parties, or unauthorized use for authentication. [5]

Most of the mentioned risks are mitigated by cloud service providers though using encrypted file systems, data loss software, security applications and security hardware for tracking of suspicious behavior across servers. In most cases, it improves security because cloud providers are in the position to allocate huge resources to solving security issues that many clients cannot afford.

In order to fully address these issues it is important to understand and examine service-level agreements (SLAs) of the specific cloud service providers and determine what are security governance policies and procedures could be best suited to provide adequate security and privacy for confidential personal and medical data in Healthcare.

##### B. *Performance and service quality*

Performance is often one of the most significant factors when considering moving services into Cloud.

Performance mostly relies on the type of service that is being used. In many cases, performance can be higher in the cloud because there is more available capacity and scalability. In other cases (mostly present when moving database servers to cloud), performance may be lower than on a traditional server.

Performance is mostly related with bandwidth that is available. Transferring big and complex data over the network requires sufficient bandwidth. This means that costs of renting cloud can be higher depending of rented bandwidth, in some cases even higher than possible savings on procurement of hardware for traditional on-premise infrastructure.

Performance of the cloud infrastructure can be affected by the number of users, usage load and implemented environment. If performance is the main determining factor in moving of some service to cloud, a 'hybrid' cloud can often be the right hosting solution. It provides clients with the scalability and cost efficiencies of cloud technology and the performance of dedicated servers. A hybrid cloud solution provides users with the ability to quickly allocate capacity for busy time periods and reduce it when user demand is lower, keeping IT costs to a minimum by only paying for the services that are currently used.

In order to address performance requirements of Cloud computing applications for chosen healthcare services, it is needed to determine number of users, and application load at peak working hours. Also, as we have mentioned when discussing cloud security, it is important to examine SLAs regarding number of users and other factors that can affect performance of Healthcare cloud services.

##### C. *Reliability and availability*

Reliability of cloud systems is expressed in the system in "nines". Two nines is the same as 99%, which is equal to 100% to 99% during 365 days or 3.65 days of downtime per year, which is typical for non-redundant hardware (time of reloading of operating system and time to restore backups after a failure is included). Three nines is up to 8 hours of downtime, four nines is up to 52 minutes and the best cloud reliability offered on market is five nines or 7 minutes of downtime per year.

When choosing an appropriate cloud technology solution for chosen Healthcare service, it is important to examine the SLA of the cloud service provider as well as the reliability of the supported technology. But if our needs for uptime exceed that which the cloud service provider can offer, there are techniques for improving it, most of which is achieved by doubling the amount of computing nodes in cloud application. There is a saying that each additional nine of uptime doubles the cost. This is explained by the fact that backup systems have to be put in place to take over if the primary servers fail. These solutions require additional design and configuration effort for the cloud application, and they should be highly tested to make sure they will work when the failures occur.

##### D. *Integration and interoperability*

Integration and interoperability is second most

important challenge, after security and privacy, in implementation of Cloud technology. This is especially important for implementation of cloud technology in Healthcare. Large number of healthcare services require significant exchange of health related information in real time. Databases of diseases, medicines and large number of codebooks are used in various healthcare services. This means that during the development and implementation of new cloud services, integration into already existing data sources in healthcare is essential. [8]

Integration of stand-alone SaaS (software as a service) cloud applications can be very difficult, without migration of all data that is needed into cloud. This can be difficult because the used data is generated and updated from already existing modules of the system.

The adoption of PaaS (Platform as a Service) and IaaS (Infrastructure as a Service), other cloud computing service models offer more options for integration. While cloud is used more and more, it will not replace all traditional hosting or on-premise deployments.

Interoperability with other software solutions (ERP, LIS, RIS, etc...) is also very important factor when considering implementation of Cloud computing.

In the hybrid world that connects on-premise data centers and the cloud, full integration and interoperability can be achieved by careful design and planning of healthcare services which are best suited for cloud implementation.

#### *E. Legal and policy challenges*

The use of cloud technology presents many legal issues. Data jurisdiction and privacy issues are major concerns.

Having in mind that healthcare information contains confidential medical and personal data of ME citizens, Montenegrin Law on Protection of Personal Data (Official Gazette of ME, 79/08, 70/09 and 44/12) applies. We will also have to take into account that the data in the Cloud can be stored on various international locations, where different laws on personal data apply.

Montenegrin Law on Protection of Personal Data specifies conditions for personal data transfer and storing on locations in other countries, in section 4 Transfer of personal data from ME (articles 41 and 42).

As it is stated in Article 41 of the Law, data can be transferred to another country only with approval of supervisory authority. In this case supervisory authority is the ME Agency for Personal Data Protection. According to Article 42, item 6, this approval is not necessary if data is transferred to the Member States of the European Union and European Economic Area or countries which are included on the European Union list of countries with an adequate level of personal data protection.

Physical locations of data that is stored on the Cloud could be widely distributed across multiple international jurisdictions, each of which may have different laws about data security and privacy, usage, and intellectual property. For Cloud data centers in European Union and European Economic Area, the EU Directive 95/46/EC applies. EU data protection directive is substantially more restrictive

than data protection rules from other countries (USA, Canada). The Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism (PATRIOT) Act gives the United States government the right to request data if it declares conditions as being an emergency or necessary to homeland security and fighting terrorism. Similarly, the Canadian Personal Information Protection and Electronic Documents Act (PIPEDA) sets limiting factors to organizations to collect, use, or disclose personal information in the course of commercial activities. However, a cloud service provider may, without notice to a client, move the client's information from one jurisdiction to another jurisdiction. Data stored in the cloud may have more than one legal locations at the same time. For all locations separate legal consequences apply, depending on the jurisdiction. [5]

#### V. CONCLUSION

This paper presents preliminary research results on identifying key challenges in adopting cloud based solutions for health care system with special focus on current situation in ME and directions how to overcome the challenges. Although cloud computing is on its way to becoming increasingly powerful tool for extending IT resources and accessibility of IT services in all areas, it will not replace all traditional hosting or on-premise deployments. There will always be situations where security requirements, flexibility, performance or legal requirements will limit cloud implementations.

The future work will be focused on comprehensive design and implementation of cloud based services for health care system in ME, its verification and evaluation, as well as on developing innovative solutions enhancing security, flexibility and other concerns on cloud.

#### ACKNOWLEDGEMENT

This study is a part of the research project "Modern technology development for Health Information System in ME" financed by the EU Assistance Funds (Ref. No: EuropeAid/136938/ID/ACT/ME).

#### REFERENCES

- [1] "Unleashing the Potential of Cloud Computing in Europe", EU Commission, Brussels, September 2012
- [2] "Zdravstvene politike u Crnoj Gori do 2020. godine " GoM Podgorica, 2003
- [3] "From innovation to implementation – eHealth in the WHO European region," WHO 2016
- [4] "Cloud computing – Benefits, risks and recommendations for information security", European Network and Information Security Agency, December 2012
- [5] Alex Mu-Hsing Kuo, PhD, "Opportunities and Challenges of Cloud Computing to Improve Health Care Services", September 2011
- [6] "ME: Preparing for Prosperity", Poverty Reduction and Economic Management Unit Europe and Central Asia Region, December 2012
- [7] Denis Rekovic, "Integral Health Information System, Good Practice case of ME" unpublished Podgorica, 2010.
- [8] Omri Erel, "Cloud Integration Challenges and Solutions", Integration zone, February 2015.

# Osvrt na kontinuirani monitoring glikemije kod pacijenata sa dijabetesom

Jurica A. Žarković M.App, Elektrotehnički fakultet, Univerzitet Crne Gore

**Sadržaj** — U radu se vrši pregled novih načina praćenja nivoa glikemije u krvi kod pacijenata oboljelih od dijabetesa, predlaže upotreba daljinskog monitoringa za kontinuirano praćenje zdravstvenih parametara sa osvrtom na kontinuirani monitoring glikemije, uz primjenu bazične tehnologije koja će nam pomoći u uvođenju proaktivne i pristupačne zdravstvene zaštite pacijenata oboljelih od *diabetes mellitus*-a, naročito djece. Rad sugeriše nova istraživanja u domenu medicinske elektronike, otvara nove smjernice kod sprječavanja dječje hipoglikemije, omogućava pacijentu da pažljivo prati promjene nivoa glikemije sa akcentom na neinvazivne medicinske metode i pri tome pruži povratne informacije koje će pomoći pri održavanju optimalnog nivoa glikemije u krvi. Ukoliko su senzori integrisani u telemedicinski sistem, onda takvi sistemi mogu čak i upozoriti medicinsko osoblje ukoliko nastupe alarmantne promjene na pacijentu.

**Gljučne riječi** — Glikemija, kontinuirano mjerenje, nivo šećera/glikemije u krvi, tip 1.

## I. O DIJABETESU

**D**IJABETES ili **Šećerna bolest** (*lat. diabetes mellitus*), je poremećaj nivoa glukoze u krvi uslijed nepravilnog funkcionisanja žlijezde gušterače (pankreas), koji se zbiva kada gušterača prestane potpuno ili djelimično proizvoditi hormon inzulin ili proizvedeni inzulin nije djelotvoran u organizmu. U tom slučaju ćelije, odnosno organizam ne koristi glukozu (prost šećer) koja je neophodna hrana za mozak i mišiće, jer nije u stanju da je koristi [1].

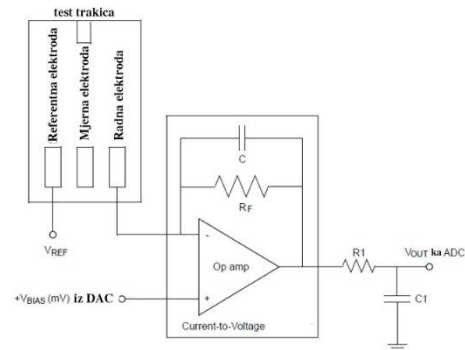
## II. POSTOJEĆI SISTEMI KONTROLE

Svi dijabetičari, a naročito oni koji su zavisni o inzulinu, tokom dana moraju više puta kontrolisati nivo glikemije u krvi. Jedan od najaktuelnijih načina kontrole šećera iziskuje invazivnu metodu koja se zasniva na svakodnevnom „bockanju“ kože kako bi se iz kapilarne krvi, električnim putem, izmjerio nivo šećera.

Više puta dnevno dijabetičari bockaju prste da bi dobili kapljicu krvi koja se zatim ukapava na plastičnu trakicu stavljenu u glukometar - portabilni, mali uređaj, koji im zatim pokazuje nivo glikemije u krvi [2].

Ovakav metod uzorkovanja je jako neugodan, ne

predstavlja vrijednosti podataka u *live* modu, tj. u realnom vremenu, ne posjeduje povratni sistem informisanja (*feedback*), ne preduzima akcije alarmiranja niti nam pomaže u sprječavanju neželjenih posljedica.



Sl.1-Princip rada trakice glukometra

Uzorak krvi se aplikuje na test trakicu glukometra, gdje se zatim odvija reakcija između glukoze u krvi i enzima (reagensa) na samoj trakici. Tokom ove hemijske reakcije dolazi do oslobađanja elektrona. Protok elektrona odgovara protoku struje kroz elektrode trakice (radna i referentna elektroda). Ovaj tok struje će se mijenjati u zavisnosti od koncentracije glukoze u krvi pacijenta. Struja se konvertuje pomoću strujno-naponskog konvertora u napon (*current-to-voltage converter*) koji se zatim analogno digitalnim konvertorom (ADC) prevodi u oblik spreman za prikaz na displeju i memorisanje [3].



Sl.2-Uobičajeni način invazivnog mjerenja

Godine 1970. prvi put je u kliničkoj praksi počela implementacija vještačkog pankreasa. Ovaj sistem je bio dizajniran da kontinuirano mjeri glikemiju tokom 24 sata na svakih sat vremena i ubrizgava odgovarajuću količinu inzulina u krv shodno izmjerenim vrijednostima glikemije. Sam sistem je bio jako velik, pacijent je bio „vezan“ za krevet, što je činilo ovaj sistem jako nepraktičnim. Međutim, već tada su primijećene velike oscilacije glikemije tokom jednog sata što je uslovljavalo stvaranjem novih istraživanja kontinuiranog mjerenja glikemije [4].

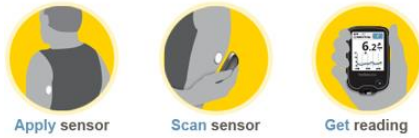
Jurica A. Žarković, M.App. 81000 Podgorica, Dušana D. Mugoše 47, Montenegro. Phone: +382-67-189350, [juricaz@t-com.me](mailto:juricaz@t-com.me).

Ovaj Rad posvećujem svojoj voljenoj kćerki **Josipi J. Žarković**, koja se svakodnevno bori sa težim odlikom dijabetesa...



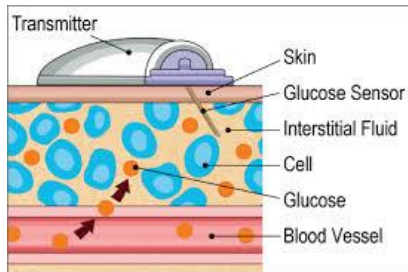
### III. MODERNA RJEŠENJA

Moderna rješenja su prije svega zasnovana na minimalno invazivnim i neinvazivnim metodama kontinuiranog uzorkovanja nivoa glikemije u krvi. S toga se danas se u medicinskoj praksi svakodnevno koriste tri minimalno invazivne metode: enzimске transkutane elektrode, mikro dijaliza, obrnuta jontoforeza.



Sl.3 - CGMS (Continuous Glucose Monitoring System) [5].

Jedan od najzastupljenijih aparata u Crnog Gori, kao i jedna od najviše korištenih tehnologija kontinuiranog mjerenja glikemije zasnovana je na upotrebi minimalno invazivne metode. Ovaj sistem koristi CGMS (Continuous glucose monitoring system) kojim se bolesniku na abdominalni zid postavlja subkutano sonda s kojom se mjeri glikemija putem glukometra. Glukometar i sonda komuniciraju bežično. Senzor uz pomoć fleksibilnih elektroda uronjenih u subkutano tkivo iz intersticijske tečnosti (vanćelijske) elektrohemijomskom metodom mjeri protok struje koji je u zavisnosti od nivoa glikemije. Ovakav sistem kasni u odnosu na mjerenje glikemije iz krvi 10-15 minuta, s toga zahtjeva kalibraciju [5].



Sl.4 - enzimatska subkutana elektroda [6]

Iako su ovi sistemi vrhunac tehnoloških dostignuća koji se kod nas koristi, važno je napomenuti da je situacija u razvijenim zemljama mnogo drugačija. Aparati sa ovim modelom uzorkovanja glikemije su i dalje prilično zastupljeni u upotrebi, ali je evidentna potreba za daljim razvojem i usavršavanjem potpuno neinvazivnih rješenja koja će omogućiti bolji kvalitet kontrole glikemije, a tim i bolji život pacijenata [6], [7].

Do sada se nova i poboljšana rješenja zasnivaju na upotrebi nekoliko inventivnih tehnologija mjerenja:

- Ultrazvučna tehnologija – Ova tehnologija zasniva se na korištenju donjeg spektra ultrazvuka čime se izvlači međućelijska tečnost koja se zatim koristi za određivanje glikemije u krvi putem provodljivosti i uz pomoć reagensa [8].



Sl.5-Primjer ultrazvučnog skeniranja glikemije

- Elektromagnetna tehnologija – Ovaj sistem je zasnovan na korištenju struje i radio talasa pri određivanju međućelijske impendanse, koja pak, ovisi o nivou glukoze u krvi. Frekvencije koje se koriste su od 2,4-2,9MHz, mada, u zavisnosti od temperature medijuma na kome se koriste, potrebno je podesiti optimalnu frekvenciju. Međutim, druga studija koja je vršena na svinjskoj krvi, pokazala je da je optimalna frekvencija 7,77GHz na 25°C [9].



Sl.6 – Primjer EM mjerenja glikemije [9].

- Optička metoda (spektroskopija) - Ovaj sistem mjerenja zasniva se na izračunavanju razlike u fluorescentnoj rezonanci između pacijenta i mjerača. Istraživanjima je utvrđeno da se nivo glukoze može utvrditi fluorescentnom rezonansom suza, čime se otvara jedan novi pravac u istraživanju uzorkovanja glikemije. [10].



Sl.7 – Optičko (spektroskopsko) mjerenje glikemije [10]

### IV. PREDNOSTI I MANE

Naravno, svaki od ovih sistema ima svoje prednosti i mane. Prednosti starijih invazivnih sistema, kao što je uzorkovanje nivoa glikemije uz upotrebu glukometra i trakica, je njihova veća preciznost i brzina, jer su u direktnom kontaktu sa kapilarnom krvi i glukozom u njoj. S aspekta elektronike ovakvi mjerači su izuzetno jednostavni za konstrukciju, upotrebu, energetski vrlo efikasni i jeftini.

Mane ovih sistema su bolni osjećaj, povećana mogućnost infekcije i oštećenje tkiva. Mimo medicinskih nedostataka, postoje i tehnički: trakice koje se mijenjaju pri svakom mjerenju (i do 7-8 puta dnevno), preosjetljivost aparata, tj. mjerne trakice na temperaturu (pogotovo nisku) i vlagu. Sve ovo povećava troškove upotrebe ovog

jednostavnog aparata, što čini ovaj način mjerenja ekonomski skupim.

Novе metode ne zahtijevaju toliko česte promjene senzora. Isti senzori uzorkovanje rade mnogo duže, što ih za pacijenta čini ekonomski mnogo isplativijim. Kako je način uzorkovanja neinvazivan, mogućnost infekcije i oštećenja tkiva skoro da ne postoji, kao ni bol.

Nažalost, ove metode su tek u razvoju, pa je njihova komercijalizacija tek započela. Mane novih sistema su njihova preosjetljivost na influentnost okoline (atmosferske, temperaturne, elektrostatičke, radio inercijenske), zatim izračunavanje glikemije duže traje i njihova preciznost je manja.

## V. PREDLOŽENO RJEŠENJE

Gore navedeni sistemi su ograničeni samo na lokalno praćenje nivoa glikemije. Predloženi sljedeći korak u istraživanju je povezivanje ovih sistema sa inzulinskom pumpom, uređajem koji sam dozira i ubrizgava inzulin u pacijenta. Time u suštini, stvara se *vještački pankreas*, uređaj koji u realnom vremenu sam mjeri glikemiju, dozira i ubrizgava inzulin na osnovu uzorkovanih podataka, dok pacijent biva obaviješten o svemu putem malog displeja. Postojeća rješenja vještačkog pankreasa nemaju potpunu *live* obradu podataka niti su umreženi u veći sistem kontinuiranog i daljinskog monitoringa vitalnih parametara. S toga, ne samo da bi inzulinska pumpa bila vođena softverom i podacima iz neinvazivnog mjernog senzora glikemije, već bi bila dio jednog malog „nosivog“ sistema koji bi kontinuirano odašiljao informacije o nivoima glikemije i stanju pacijenta putem GSM mreže (mobilnog interneta) ka medicinskom serveru [11].

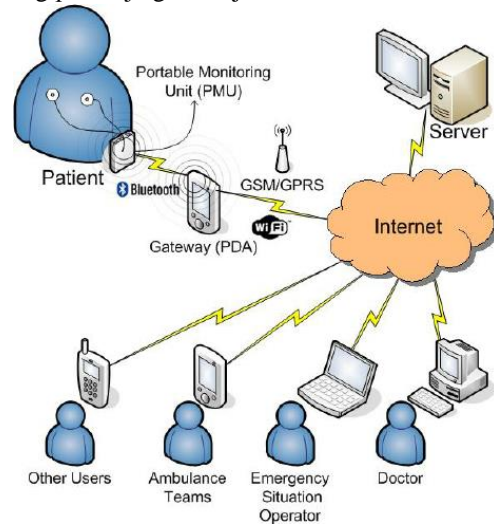


Sl.8 - Primjer „vještačkog pankreasa“ [5]

Dalji tok realizacije bi podrazumijevao povezivanje ovog „tele-pankreasa“ u savremeni telemedicinski sistem koji bi 24/7 vršio nadzor, kontrolu i podršku pacijentu.

Savremeni telemedicinski sistem nadzora je prijedlog moderne sveprisutne medicinske zaštite koji nadgleda i kontrolira zdravstvene parametre pacijenata. Arhitektura na slici 9, koja je primijenjena u ovom sistemu je jedna od mogućih rješenja telemedicinskog sistema. Sistem se sastoji od više senzorskih čvorova, koji prate: kretanje tijela, srčanu aktivnost, temperaturu, kao i obaveznog

neinvazivnog gluko-senzora, zatim ličnog servera instaliranog na mobilnom *smart* telefonu ili ličnom računaru *notebook*-u i na kraju medicinskog servera putem kojeg tim ljekara sa svojih operaterskih stanica (računara) prate pacijente. U slučaju pada glikemije kod djeteta, automatski bi se oglosio alarm, kako kod pacijenta tako i dalje u sistemu medicinskog servera. Na ovaj način, pacijenti sa dijabetesom, imaju nov sistem kontinuiranog i daljinskog praćenja glikemije.



Sl.9 – Telemedicinski nadzor pacijenta [11]

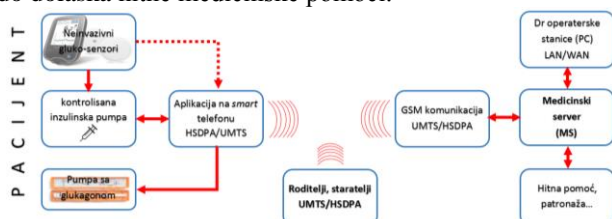
## VI. STUDIJA SLUČAJA

Ovdje je predstavljena studija slučaja kao ilustracija upotrjebljivosti sistema za daljinski nadzor pacijenta oboljelog od teškog oblika dijabetesa, sa kontinuiranim monitoringom glikemije. Predstavljani pacijent nije fiktivni, već osmogodišnja djevojčica Josipa sa dijagnozom dijabetesa *tip1*, koja je ujedno i predstavnik zajedničkih problema koji bi zadesili pojedince tokom života sa dijabetesom. Prijedlog ovog rješenja je zapravo u domenu prevencije i lijeka u situacijama teške hipoglikemije, naročito kod djece. Nakon sagledavanja problema kod klasičnih konvencionalnih mjerenja glikemije, prednosti koje pruža kontinuirani sistem daljinskog monitoringa su vrlo evidentne.

Djevojčici je dijagnostifikovan teži oblik dijabetesa, *tip1*-inzulin zavisni dijabetičar, što znači da ovi pacijenti nemaju ispravan rad pankreasa, da im pankreas ne proizvodi hormon inzulin. Kako bi regulisali nivo šećera/glikemije u krvi, primjenjuje se ubrizgavanje inzulina u tijelo pacijenta. Svaki dan prije jela, Josipa mjeri glikemiju ubadanjem prstića, uzimanjem kapljice krvi na test trakicu glukometra i shodno trenutnom nivou glikemije i planiranom obroku, djevojčici se ubrizgava određeni broj jedinica inzulina. Međutim, određene psihofizičke predispozicije osmogodišnjeg djeteta, dovode do naglog pada ili porasta glikemije, pa tu imamo dva opasna stanja; hipoglikemiju-izuzetno opasno stanje u kom mozak ostaje bez hrane (prostit šećera) i hiperglikemiju-gdje su nivoi glikemije jako visoki što rezultira da cijeli organizam pati mučeći se da toliku količinu energije izbaci i potroši. Ovdje razmatramo opasniji i češći primjer-**hipoglikemiju**: tokom dana, djevojčici je naglo pao nivo glikemije u krvi dok se igrala sa svojim drugaricama. U

tom momentu, mozak brzo pada u stanje nemogućnosti obrade osnovnih funkcija jer ostaje bez hrane, tj. „goriva“. Ona pada u nesvijest koju prate brojne konvulzije i prateće gušenje zbog gutanja jezika uslijed nesvjesticke. Ako se brzo ne tretira, hipoglikemija može dovesti do smrti!

Predloženi sistem kontinuiranog i daljinskog monitoringa pacijenata sa dijabetesom, pruža rješenje u ovom slučaju, slika 10. Josipa, opremljena sa neinvazivnim senzorima, insulinskom pumpom i GSM smart uređajem, koji kontinuirano prate vitalne statistike, povezuje se putem mobilnog interneta na medicinski server. Endokrinolog zadužen za praćenje Josipinog zdravlja, sa svoje operaterske konzole u svakom trenutku može očitati Josipine medicinske parametre, naročito, u ovom slučaju, nivo glikemije. Takođe, roditelji mogu biti obaviješteni putem SMS poruka o kritičnom nivou glikemije kod Josipe, pa već u kriznim situacijama i hitnim pozivom i paljenjem ambijentalnog mikrofona i zvučnika na smart telefonu pacijentice kako bi uspostavili hitni kontakt. Kao treći, krajnji ishod, medicinsko osoblje, ili sam sistem, putem medicinskog servera može izdati komandu kojom bi se ubrizgao glukagon (povisuje nivo glukoze u krvi) sa specijalnom pumpom koja može biti integrisana i u inzulinskoj pumpi, i na taj način spasio život pacijentice do dolaska hitne medicinske pomoći.



Sl.10 – Blok šema predloženog rješenja. Lijevo: gluko-senzor prosljeđuje informaciju inzulinskoj pumpi i smart telefonu. Aplikacija sa telefona vrši obradu podataka, komunicira sa med. serverom i daje komandu inzulinskoj pumpi o količini insulina. Desno: Medicinski server uz kontrolou ljekara. Glukagonska pumpa u slučaju hipoglikemije biva aktivirana daljinski. Roditelji se informišu/alarmiraju putem GSM mreže o stanju glikemije.

## VII. ZAKLJUČAK

Sistemi koji prate vitalne funkcije, obećavaju sveprisutan ali ipak ekonomsko povoljan zdravstveni monitoring. Smatram da će implementacijom „telepankreasaa“, malog, nosivog sistema, u telemedicinski sistem omogućiti drastičnu promjenu načina života pacijenata oboljelih od težeg oblika dijabetesa, tipa 1. na isti način na koji je i inzulini spasio milione života dijabetičara.

Ova promjena prema proaktivnijoj, preventivnoj zdravstvenoj zaštiti ne samo da će poboljšati kvalitet života dijabetičara, već će takođe smanjiti zdravstvene troškove, koji su u ovoj grani medicine zaista veliki. Razvoj bežičnih uređaja i najnovija dostignuća u minijaturizaciji senzora dokazuju tehničku fleksibilnost sveprisutnih zdravstvenih monitoring sistema. Nove tehnologije imaju potencijal da ponude širok spektar povoljnosti za dijabetičare, medicinsko osoblje i društvo, kroz kontinuirano daljinsko praćenje pacijenata, rano otkrivanje abnormalnih stanja, nadzor rehabilitacije i

potencijalna saznanja kroz ekstrakciju podataka iz prikupljenih informacija.

## ZAHVALNICA

Ovim putem zahvaljujem se prof. dr Radovanu Stojanoviću sa Univerziteta Crne Gore-ETF, kao i mom Dalmatincu, dragom prijatelju, Marinu Čavelišu, inženjeru u Zavodu za meteorologiju i seizmologiju Crne Gore.

## LITERATURA

- [1] Dr sci Sandra Stanković, pedijatar endokrinolog - private communication
- [2] Dr Darko Stanković - private communication
- [3] <http://engineering.mit.edu/>
- [4] Napredak u kontroli glikemije - Mr. sc. Manja Prašek, dr. med.
- [5] Inzulinske pumpe i kontinuirano mjerenje glukoze - Manja Prašek, Ana Jakir Sveučilišna klinika za dijabetes, endokrinologiju i bolesti metabolizma Vuk Vrhovac, Zagreb
- [6] Continuous subcutaneous insulin infusion: a comprehensive review of insulin pump therapy-Lenhard MJ, Reeves GD.
- [7] <http://www.freestylelibre.co.uk/> - aktuelni proizvod u Crnoj Gori
- [8] Noninvasive Ultrasonic Glucose Sensing a Lightweight Cymbal Transducer Array and Biosensors-Eun-Joo Park, M.S., Jacob Werner, V.M.D., Joshua Beebe, Samantha Chan, Nadine Barrie Smith, Ph.D, 2009.
- [9] Electromagnetic apparatus and method for the reduction of serum glucose levels, US patent: 4428366 A
- [10] Transcutaneous glucose measurement using near-infrared spectroscopy during hypoglycemia.-I Gabriely, R Wozniak, M Mevorach, J Kaplan, Y Aharon and H Shamon
- [11] Primjena WBN tehnologije u kontinuiranom praćenju medicinskih parametara.-J. Žarković, R. Stojanović, 2010.
- [12] <https://www.researchgate.net/>
- [13] Slike – lične fotografije kćerke dijabetičara (izbor i modifikacija) - J.A. Žarković

## ABSTRACT

This paper is a review of new modes of monitoring of blood glucose levels in diabetic patients, suggests the use of remote continuous monitoring of medical parameters with focus on continuous monitoring of blood glucose, with the use of basic technology that will help us in introducing proactive and affordable health care for patients suffering from diabetes mellitus, especially children. The work suggests new research in the field of medical electronics, opens new guidelines to prevent children's hypoglycemia, allows the patient to closely monitor changes in blood glucose levels with a focus on non-invasive medical methods and thereby provide feedback that will help maintain an optimal level of glycemia. If the sensors are integrated into the telemedicine system, then such systems can even alert medical personnel in case of alarming changes in the patient.

Keywords - glycaemia, continuous measurement, blood sugar / blood glucose levels, type 1

## REVIEW OF THE METHODS OF CONTINUOUS MONITORING OF BLOOD GLUCOSE IN PATIENTS WITH DIABETES

Jurica A. Žarković, M.App

# Retinal Blood Vessel Segmentation: Dijkstra Forest and Multi-Scale Line Detection Methods

Jelena Badnjar, Nataša Popović, Miroslav Radunović, and Tomo Popović, *Senior Member, IEEE*

**Abstract** — Use of information technologies to process medical images is becoming a standard of patient care. Blood vessel segmentation, or the extraction of a blood vessel image from a raw image, is a type of image processing commonly performed in magnetic resonance angiography, and computerized tomography angiography. This is one of the crucial and necessary steps for evaluation of vascular complications related to hypertension, diabetes, heart disease, and cancer-related vascular pathology. This paper focuses on the use of automated computational methods of blood vessel distribution analysis in retinal images. In this paper, we compare two methods for retinal vessel segmentation: Dijkstra forest based automatic vessel segmentation and the multi-scale line detection method. We used these two methods to segment raw retinal images from the STARE public database. To compare the two segmentation methods we calculated the F1 score for each. The results suggest that the multi-scale line detection is superior to Dijkstra forest based method.

**Keywords** — Automatic retinal segmentation, Dijkstra forest based, multi-scale line detection, blood vessel segmentation.

## I. INTRODUCTION

**M**ORPHOLOGICAL characteristics of blood vessels such as length, width, tortuosity, branching pattern and angles are used for evaluation of various cardiovascular and ophthalmologic diseases [1], [2]. Therefore, blood vessel segmentation, or the extraction of a blood vessel image from a raw image, is one of the crucial and necessary steps for studying and diagnosing vascular complications related to hypertension, diabetes, heart disease, and cancer-related vascular pathology. This type of image processing is commonly performed in magnetic resonance angiography and computerized tomography angiography [3], [4]. More specifically, changes in the morphology of retinal blood vessels are known to be associated with retinopathy of prematurity, as well as complications of diabetes and hypertension [5], [6]. Vessel segmentation performed manually by experts

is labor intensive, slow, and often subjective, which results in great expert-to-expert variability of the results [5]. Therefore, a great deal of research has been done in order to develop a reliable, quick, simple and objective automated segmentation method that could eventually replace manual vessel extraction.

Most of the segmentation algorithms could be divided into two general groups: path-based and region-based methods [7]-[9]. Region based techniques are based on the selection of the seed point as the initial step. This selection must be correct in order to have a successful segmentation process. If the initial seed point selection misses the region of interest, the segmentation results will be incorrect. Techniques of this type assign each pixel in the image to the group of vessel and non-vessel pixels. On the other hand, the path-based methods focus on tracking individual vessels, without the classification of pixels.

Most of the automated segmentation techniques published so far have been developed for the analysis of high quality retinal images. However, in real life, the quality of retinal images is variable. This variability can be caused by the use of different cameras, use images captured from an indirect ophthalmoscope, different operators taking the image, or sometimes the images can contain motion or dust artifacts and reflections. Very often the camera settings used for imaging are not recorded so they cannot be reliably replicated by another research group. In addition, retinal images in some databases were captured following pupil dilatation, while in some databases pupil dilatation was not done, which greatly affects the size of the field of view [5], [10]. Therefore, segmentation algorithms that have been developed to perform well for one database, very often do not perform well when used on a different database of retinal images.

The goal of the research we have presented here was to compare two methods for automated retinal vessel segmentation: Dijkstra forest based automatic vessel segmentation and multi-scale line detection method. They represent some of the most robust methods published so far because they are relatively resistant to the variability in quality of raw retinal images [5], [11].

## II. MATERIALS AND METHODS

*Raw images of retina.* The raw images were obtained from STARE (Structured Analysis of Retina) database. In addition to raw images, this database contains two sets of manually segmented images, each by a different expert [12]-[14].

---

Jelena Badnjar, Faculty of Information Systems and Technologies, University of Donja Gorica, Oktoih 1, 81000 Podgorica, Montenegro (e-mail: jelena.badnjar@udg.edu.me).

Nataša Popović, Faculty of Medicine, University of Montenegro, Kruševac bb, 81000 Podgorica, Montenegro (e-mail: npopovic@ac.me)

Miroslav Radunović, Faculty of Medicine, University of Montenegro, Kruševac bb, 81000 Podgorica, Montenegro (e-mail: miroslavr@ac.me)

Tomo Popović, Faculty of Information Systems and Technologies, University of Donja Gorica, Oktoih 1, 81000 Podgorica, Montenegro (phone: +382-20-410-777; e-mail: tomo.popovic@udg.edu.me).

*Binary mask application.* The mask used for separation of the retina region from the background is produced by the author's algorithm using the input retina image from the STARE database and generating an adequate black and white mask [12]. The mask image must be equal to the size of the input image (700×605 resolution) and clearly separate retina region from the background. A number of different masks is used in order to select the best one that clearly separates edges of retina and helps clearing unnecessary artifacts on the segmented image.

*Dijkstra based segmentation.* The first considered method for automatic segmentation extracts the vascular structure using Dijkstra's shortest path algorithm. For the comparison, we used the exploratory Dijkstra vessel segmentation algorithm implementation of Estrada et al. [5] with its original default configuration parameters. The implementation is a hybrid method that extends the path-based methodology into region-based segmentation for detecting retinal vessels. First, the images are preprocessed by Laplace-of-Gaussian (LoG) - Gabor filters, and directional local-contrast filters and this part of the segmentation eliminates artifacts and increases the contrast of the image. Then, each retinal image is represented as a graph of nodes  $G=(V,E)$ , where each pixel is a node and links connecting the pixels are arcs. The method defines and considers *cost* of traversing the arc connecting two neighboring pixels in order to determine if the pixel belongs to vessel network. The position of every pixel is represented in a two-dimensional matrix, where the cost of each pixel is determined using the green channel intensity and inverted response of the LoG filter, followed by the Gabor filter [17]. The LoG filter gives a higher contrast to the image, while the Gabor filter improves the image elongated structures.

The exploratory Dijkstra vessel segmentation algorithm starts from a single pixel node and then progressively explores the rest of the nodes in a way that every unvisited node has a higher minimum path cost when compared with all of the previously visited nodes. A lower cost means that the pixel belongs to a blood vessel or its boundaries [5], [18]. The exploratory algorithm focuses on the path with a non-negative cost. There may be multiple paths with the same starting and end points, but the algorithm selects the path with the minimum cost. One of the advantages of this method is exclusion of any manual intervention [5], [15], [16].

*Multi-scale line detection.* The second method used in the comparison evaluation discussed in this paper is the multi-scale line detection method proposed by Nguyen et al. [11]. This is a path-based method, which is especially robust for segmentation of closely positioned and small vessels, as well as at bifurcation and crossover regions.

The line detector is based on the inverted green channel of the retinal image and the use of window segments around each pixel. It is based on a combination

of line detectors where a basic line detector uses a set of straight lines rotated around the pixel in the center of the window in order to detect if that pixel belongs to the vessel network (Fig. 1) [19]. The line with the highest average gray level is identified and selected. Then, the pixel in the center is characterized by its *line strength* calculated as the difference between average gray level of the selected line and the average gray level of the whole window. The line detector depends on the window size  $W$ . The window size was set to  $15 \times 15$  pixels, which is twice the size of a typical vessel width in the image. In general, the window size is chosen so that the number of vessel and background pixels inside the window is nearly equal.

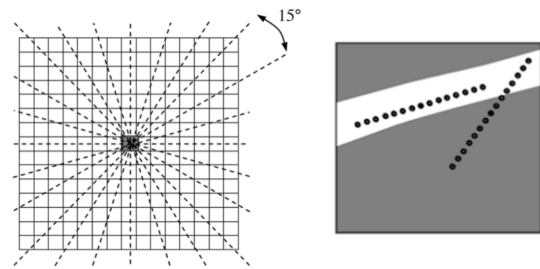


Fig. 1. Basic line detector: a) using twelve lines within 15x15 window; b) assessing alignment between lines and a vessel [19].

In [11], the authors identified drawbacks of the basic line detection such as false detection of a background pixel between two vessels, at the corner of a vessel crossover point, or near a strong vessel. They proposed a multi-scale line detection, which works with different line lengths  $L$ , where  $1 \leq L \leq W$ , in order to overcome these drawbacks. In this method, by varying the lengths of line  $L$ , the method implements line detectors at different scales. The biggest advantage of changing the length of line detector, with relation to decreasing size of windows  $W$ , is finding more accurate vessel boundaries. This change leads to better contrast between vessels and background leading to more precise segmentation results [11]. Since the resulted image may contain other artifacts and unnecessary lines, the post-processing of the segmented image includes filtering in order to produce more clear and correct image.

*Recall, precision, and F1 score.* For each method, manual and automatic segmentation, the pixels that were identified as a part of blood vessels are assigned a value of 1. The pixels identified as non-vessel pixels are assigned a value of 0. True negative ( $tn$ ) means that both methods assigned value of zero to a certain pixel. True positive ( $tp$ ) means that both methods assigned value of one to a certain pixel. False positive ( $fp$ ) is a mismatch where analysis of a pixel by automated method produced value of one and the ground truth, manual segmentation, had a value of a zero. False negative ( $fn$ ) is the inverse mismatch, where the automated method identified pixel as a non-vessel, with a value of zero, while the manual method identified the same pixel as a vessel. Recall is

defined as a percent of pixels that are correctly identified as a vessel by the automatic method out of all pixels that should have been identified because they were identified as a blood vessel by the manual method. It can be calculated by formula (3). Precision is defined as the percent of pixels that are correctly identified by an automatic method as a blood vessel out of all pixels identified as a blood vessel by the automatic method. It can be calculated by equation (2).

In cases of data processing where the number of truly negative findings is much larger than the number of truly positive findings, F1 score is usually considered to be the best measure of the performance of the detection of positive findings. The F1 score is calculated by using equation (1). The performance of the test is better if the value of F1 score is closer to 1 (Fig. 2) [5], [20].

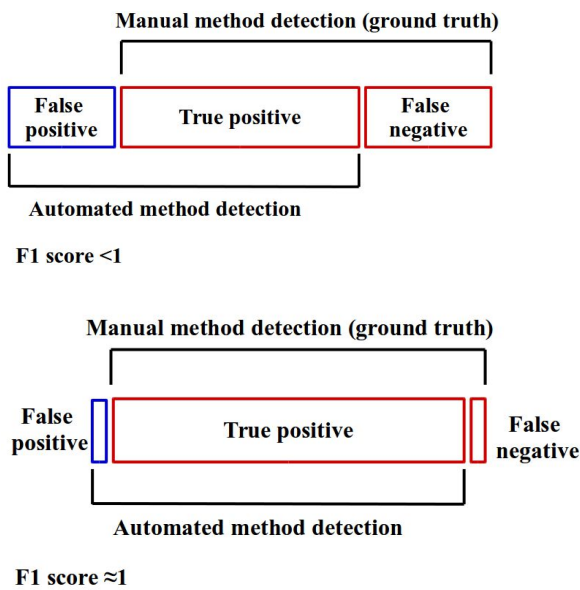


Fig. 2. F1 score calculation: greater overlap in positively detected pixels representing a blood vessel between the manual method and the automated method results in F1 score with a value closer to 1.

$$F1\ score = 2 \frac{precision \cdot recall}{precision + recall} \quad (1)$$

$$precision = \frac{tp}{tp + fp} \quad (2)$$

$$recall = \frac{tp}{tp + fn} \quad (3)$$

### III. EXPERIMENTAL RESULTS

*Visual inspection.* Images manually segmented by V. Kouznetsova (VK) are more detailed than the images segmented by A. Hoover (AH) [14], [21]. Visual inspection of the automatically segmented images shows that the multi-scale line detection yields more detailed segmented images, while the exploratory Dijkstra based

segmentation method sometimes resulted in images with vessel interruptions in places where they should clearly be connected. While multi-scale method produces images with less vessel interruptions, in some images it falsely identified edges of the optic disk as a blood vessel network structure (Fig. 3).

*F1 score results.* F1 scores for the Dijkstra forest based method and multi-scale line detection are shown in Table 1. Additionally, we calculated the F1 scores for manually segmented images by both VK and AH when compared to each other, which were 0.8597 and 0.8642 respectively.

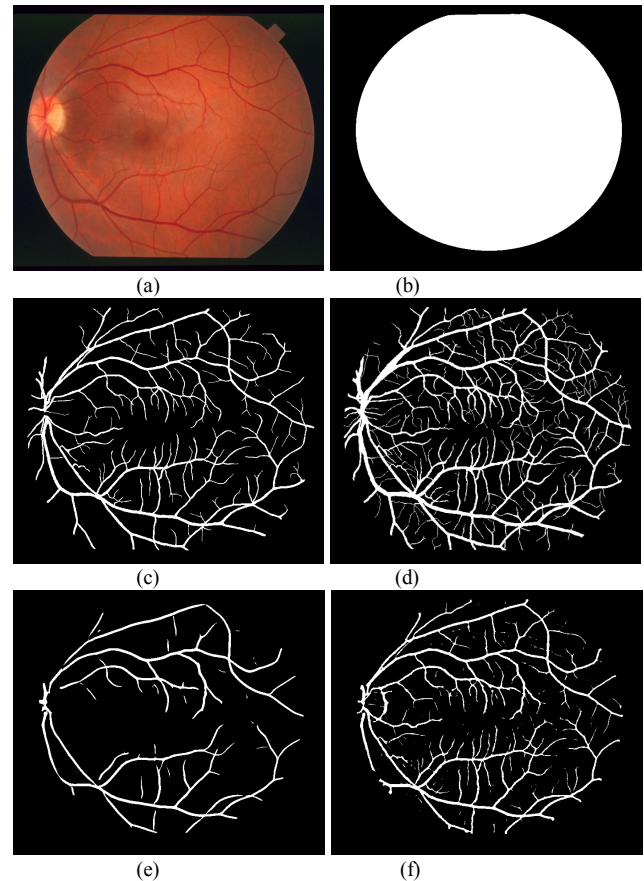


Fig. 3. Segmentation of a representative image from STARE database: a) Raw image of the retina. b) Mask used in both algorithms. c) Manually segmented by AH [21]. d) Manually segmented by VK [21]. e) Image segmented by the Dijkstra based method. f) Image segmented by the multi-scale line detection method.

TABLE 1: COMPARISON OF F1 SCORES

Segmentation methods	Ground truth	F1 score
Dijkstra forest based	AH set	0.6215
Dijkstra forest based	VK set	0.5680
Multi-scale line detection	AH set	0.7318
Multi scale line detection	VK set	0.7093
Manually segmented by VK	AH set	0.8597
Manually segmented by AH	VK set	0.8642

## IV. DISCUSSION

It can be seen in Table 1 that the multi-scale line detection exhibited a better F1 score. For both methods, the scores vary depending on the selection of ground truth. The Dijkstra method yielded images of blood vessels with less detail, while the multi-scale line detection generated more detailed images making them more similar to the ground truth images. This was reflected in the F1 scores. Also, it should be noted that the multi-scale method displayed a much better performance as it was around 150 times faster than the Dijkstra algorithm.

When compared to other published segmentation methods, these two methods have better scores than matched filters and line detection methods, but scores comparable to the wavelet based and centerline detection methods (see Table 2).

A better F1 score could possibly be achieved by adding additional filters that would remove optic disc artefacts or by adding an additional mask to remove an artefact originating from the edge of the retinal image. Also, both methods allow for additional settings tuning, which could possibly improve the results.

Moreover, when we calculated the F1 score for one manually segmented set compared to the other as the ground truth, the values were not close to 1 (see Table 1). This suggests that the F1 score alone might not be the best absolute measure of accuracy of an automatic segmentation method.

TABLE 2: F1 SCORES FROM SEGMENTATION METHODS IN THE PUBLISHED LITERATURE

<i>Segmentation methods</i>	<i>F1 score</i>
Dijkstra forest based	0.6215
Multi-scale line detection	0.7318
Matched filters [9], [22]	0.4755
Centerline detection [19], [23]	0.7168
Line detection [13]	0.7041
Wavelet based [24], [25]	0.7341

## V. CONCLUSION

The multi-scale line detection exhibited better performance and a better F1 score than the Dijkstra method. The results are promising, but more research is needed to evaluate these algorithms with different settings and filters, as well as to check their performance with different databases of retinal images.

## REFERENCES

- [1] H.C. Han, "Twisted blood vessels: symptoms, etiology and biomechanical mechanisms", *Journal of Vascular Research*, Karger, 2012.
- [2] N. Popović, T. Popović, "Fractal analysis of digital images of retina," *XXI međunarodni naučno-stručni skup Informacione tehnologije*, pp. 125-128, 2016.
- [3] B. Wasan, A. Cerutti, S. Ford, R. Marsh, "Vascular network changes in the retina with age and hypertension," *Journal of Hypertension*, Vol. 13, Issue 12, December 1995.
- [4] E.J. Sussman, W.G. Tsiaras, K.A. Soper, "Diagnosis of diabetic eye disease," *The Journal of the American Medical Association*, Vol. 23, June 1982.
- [5] R. Estrada, C. Tomasi, M. T. Cabrera, David K. Wallace, S. F. Freedman, S. Farsiu, "Exploratory Dijkstra forest based automatic vessel segmentation: applications in video indirect ophthalmoscopy," *Biomedical Optics Express*, Vol. 3, Issue 2, pp. 327-339, 2012.
- [6] A. Gole, A. L. Ells, X. Katz, G. Holmstrom, A. R. Fielder, A. Capone, Jr, J. T. Flynn, W. G. Good, J. M. Holmes, J. A. McNamara, E. A. Palmer, G. Quinn, E., M. J. Shapiro, M. G. J. Trese, D. K. Wallace, "The international classification of retinopathy of prematurity revisited," *The international classification of retinopathy of prematurity revisited*, Vol. 123, pp.991-999, July 2005.
- [7] C. Kirbas, F. Quek, "A review of vessel extraction techniques and algorithms," *ACM Computer Surveys*, Vol. 36, pp. 81-121, June 2004.
- [8] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, M. Goldbaum, "Detection of blood vessels in retinal images using two-dimensional matched filters," *IEEE Transactions on Medical Imaging*, Vol. 8, pp. 263-269, 1989.
- [9] M. M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A.R. Rudnicka, C.G. Owen, S.A. Barman, "Blood vessel segmentation methodologies in retinal images - A survey," *Computer methods and programs in biomedicine*, vol. 108, no. 1, pp. 407-433, October 2012.
- [10] T. Kauppi, V. Kalesnykiene, J.C. Kamarainen, L. Lensu, I. Sorri, H. Uusitalo, H. Kalvianen, and J. Pietila, "DIARETDB0: Evaluation Database and Methodology for Diabetic Retinopathy Algorithms." *Technical report*. Lappeenranta University of Technology, Lappeenranta, Finland, 2006.
- [11] Uyen T. V. Nguyen, Aluddin Bhuiyan, Laurence A. F. Park, Kotagiri Ramamohanarao, "An effective retinal blood vessel segmentation method using multi-scale line detection," *Pattern recognition*, Vol. 46, pp. 703-715, June 2012.
- [12] The Structured Analysis of the Retina (STARE) project, [www.parl.clemson.edu/stare/probing](http://www.parl.clemson.edu/stare/probing), September 4, 2005.
- [13] M. Law and A. Chung, "Three dimensional curvilinear structure detection using optimally oriented flux," in *ECCV*, 2008.
- [14] A. Hoover, M. Goldbaum, "Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels", *IEEE Transactions on Medical Imaging*, vol. 22 no. 8, pp. 951-958, August 2003.
- [15] A. S. Sajadi, S. H. Sabzpoushan, "A new seeded region growing technique for retinal blood vessels extraction," *Journal of medical signals and sensors*, Vol. 3, June 2014.
- [16] O. Dalmau, T. Alarcon, "MFCA: Matched filters with cellular automata for retinal vessel detection," *Advances in artificial intelligence*, Vol. 7094, pp. 504-514, 2011.
- [17] Q. Li, J. You, L. Zhang, and P. Bhattacharya, "Automated retinal vessel segmentation using Gabor filters and scale multiplication," *Proceedings of System, Man and Cybernetics IEEE*, pp. 3521-3527, 2006.
- [18] Jun Zhang, Jinglu Hu, "Renal Biopsy Image segmentation based on 2-D Otsu method with histogram analyses," *Medical imaging technology*, Vol. 27, No. 3, May 2009.
- [19] E. Ricci, R. Perfetti, "Retinal blood vessel segmentation using line operators and support vector classification," *IEEE Transactions on Medical Imaging*, Vol. 26, pp. 1357-1365, 2007.
- [20] Uyen T., R. Kotagiri, L. A. F. Park, A. Bhuiyan, "A Quantitative Measure for Retinal Blood Vessel Segmentation Evaluation," *International Journal of Computer Vision and Signal Processing*, January 2012.
- [21] A. Hoover, V. Kouznetsova, M. Goldbaum, "Locating Blood Vessels in Retinal Images by Piece-wise Threshold Probing of a Matched Filter Response", *IEEE Transactions on Medical Imaging*, vol. 19 no. 3, pp. 203-210, March 2000.
- [22] B. Zhang, L. Zhang, L. Zhang, F. Karray, "Retina vessel extraction by matched filter with first-order derivative of Gaussian," *Computer in Biology and Medicine* vol. 40, no. 4, pp. 438-445, April 2010.
- [23] M. M. Fraz, S. A. Barmana, P. Remagnino, A. Hoppe, A. Basitb, B. Uyyanonvarac, A. R. Rudnickad, C. G. Owend, "An approach to localize the retinal blood vessels using bit planes and centerline detection," *Computer methods and programs in biomedicine*, vol. 108, no. 2 pp. 600-616, November 2012.
- [24] A. Fathi, A. R. Naghsh-Nilchi, "Automatic wavelet-based retinal blood vessels segmentation and vessel diameter estimation," *Biomedical Signal Processing and Control*, vol. 8, no. 1, pp. 71-80, January 2013.
- [25] Soares, J.V., Leandro, J.J., Cesar Jr, R.M., Jelinek, H.F., Cree, M.J., "Retinal vessel segmentation using the 2-d gabor wavelet and supervised classification," *IEEE T-MI* Vol. 9, Issue 25, pp. 1214-1222, 2006.

# Retinal blood vessels segmentation using neural networks and features based on line operators

Milija Bajčeta, Petar Sekulić, Slobodan Djukanović, *Member, IEEE*

**Abstract** — In this paper, we present extraction of the blood vessels from eye-fundus digital image by using neural networks. Vessel extraction from retina images is the first and primary step in automatic detection and analysis of several diseases, such as diabetes, hypertension, arteriosclerosis and choroidal neovascularization. The proposed method uses feature extraction from preprocessed retinal image for the classification of the observed pixel into one of two classes: blood vessel or background classes. System was trained on test images from the DRIVE database. System was tested using test images from the DRIVE and STARE databases.

**Keywords** — Neural network, blood vessel segmentation, feature extraction, line detection, retinal image.

## I. INTRODUCTION

THE blood vessels and structures of the eye-fundus show the first signs of the progress or presence of several diseases. Therefore, detection of blood vessels is the first part of the system for eye disease detection, whereas the second part is determining the characteristics that indicate the medical condition of the patient.

The first problem, detection of blood vessels, can be resolved in two ways: a) manual segmentation and b) automatic segmentation. Manual detection of blood vessels of the retina fundus and their extraction from medical images is a tedious and long process that requires a lot of skills and training. Numerous algorithms have been proposed in the literature for blood vessel detection in digital retinal images [1]-[7]. The two most commonly used algorithm categories are:

- a) Pattern classification and machine learning,
- b) Matched filtering.

The algorithms in pattern classification and machine learning category [8], [9] can be separated into two groups: a) supervised and b) unsupervised algorithms. In both groups, pixel classification is performed pixel by pixel based on the features obtained from the input data, in our case retinal image. The observed pixel is put into one of two classes: blood vessel or background classes. In supervised group, which are more often used, the decision on the pixel class is made using initial knowledge. The system has to be trained in order to acquire the initial knowledge. On the other hand, unsupervised algorithms do not require initial knowledge of the system, i.e. they are

capable of using features without initial knowledge when making a decision on the input data class.

In the matched filtering algorithms, the image is filtered with a mask [2], which is designed to represent the characteristics. The matched filter is based on the following properties: a) vessels usually have a limited curvature and may be approximated by piecewise linear segments, b) vessel's width decreases as they move radially outward from the optic disk and c) cross-sectional pixel intensity profile of these line segments approximates a Gaussian curve [2].

This paper considers blood vessel detection in digital retinal images using neural networks. The proposed method consists of four steps: a) preprocessing, b) feature extraction, c) neural network and d) postprocessing. It is trained on the DRIVE database. The size of the images in this database is 565×584 pixels and 45° is a field-of-view (FOV). The DRIVE database contains 40 eye-fundus images, 20 test images and 20 training images. The test images are manually segmented by two experts. Manual images are used for checking the accuracy of the method. The training set also includes the set of manual images made by the first expert. The proposed method is also tested on training images of the STARE database, since only the training images in this database have corresponding manual images. The size of the images in STARE database is 700×605 pixels with FOV of 35°.

In Section II, the concept of neural networks is described. Section III describes the preprocessing, feature extraction, neural network for blood vessel segmentation and postprocessing. In Section IV, the experimental results are presented. Section V concludes the paper.

## II. NEURAL NETWORKS

The inventor of the first neurocomputer, Dr. Robert Hecht-Nielsen, defines a neural network as "...a computing system made up of a number of simple, highly interconnected processing elements, which process information by their dynamic state response to external inputs." In other words, it represents a system that performs parallel processing of data in order to simulate the operation of the human brain. The system should make the right decision based on their own knowledge acquired by learning.

The human brain is made up of large numbers of neurons, that work parallel in the processing of data. Similarly, artificial neural networks or just neural networks, represent a simplified mathematical model of the human brain. The idea originated in 1940 when McCulloch and Pitts (Massachusetts Institute of Technology)

This work has been supported by the Ministry of Science of Montenegro and the HERIC project through the BIO-ICT Centre of Excellence (Contract No. 01-1001).

Milija Bajčeta, Petar Sekulić and Slobodan Djukanović are with Faculty of Electrical Engineering, University of Montenegro. Address: Džordža Vašingtona bb, 81000 Podgorica, Montenegro. Phone: +382 20 245 839, fax: +382 20 245 873. E-mails: {milijab@ac.me, petars@ac.me, slobdj@ac.me}. Web: www.etf.ac.me.



published a mathematical model of neural networks [10].

The basic element of each neural network is a neuron. Neuron is the basic unit of data processing in the neural network. Thus, the neuron define the rules to obtain an output based on  $N$  inputs. The way of obtaining the output based on the input is divided into two parts (Fig. 1):

- a) network function and
- b) activation function.

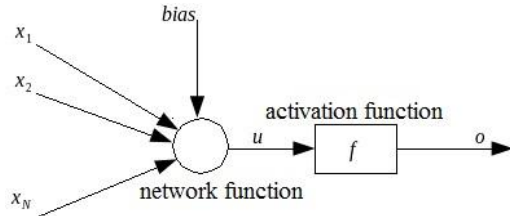


Fig. 1. Neuron.

The network function contains the knowledge of system and it performs a combination of  $N$  input values. There are several types of network function: linear, linear II order and product function. Because of its simplicity, the most widely used type of network function is linear [11], which will be used in this paper:

$$u(n) = \sum_{i=1}^N w_i x_i(n) + \Theta, \quad (1)$$

where  $w_i$  is values of the coefficients,  $x_i(n)$  is input data and  $\Theta$  represents the neuron bias.

Activation function transforms the output of function network to the desired system output. Two most common activation functions are unipolar threshold function and unipolar sigmoid [11]. In our method, unipolar sigmoid is used as activation function:

$$f(u) = \frac{1}{1 + e^{-u}}. \quad (2)$$

Knowledge of the system is contained in the coefficients of the system  $w_i$ , which are obtained by learning (training) of the system.

Depending on whether it is a known system output, there are two ways of training: a) *supervised*, when learning is performed with pairs of input-output, b) *unsupervised*, when learning is performed without the knowledge of the output. In the case of supervised learning, it is necessary to minimize the square error:

$$E_k = \frac{1}{2} (d_k - o_k)^2, \quad (3)$$

where  $d_k$  and  $o_k$  represent the desired output and true output of the system, respectively. The cycle of training ends when all data available for the training are used. The whole cycle is called an *epoch*. When training the system, it is necessary to repeat this cycle several times. In the case that the result is not satisfactory, it is necessary to reinitialize the network coefficients, or change the structure of neural network, or change the features.

In the proposed method, we used the supervised learning and the least mean squares (LMS) algorithm to modify the neuron coefficients in order to minimize (3) [11].

In an ideal case of unsupervised learning, the number of output neurons is equal to the number of data categories (classes). Each neuron recognizes one category after a certain time of training. In the case that two neurons

recognize one class, it is necessary to eliminate one neuron, whereas in the case that class is unrecognized, it is necessary to add a neuron.

The neural network is composed of a number of neurons that are connected to each other and with the inputs data in order to produce the output data. Depending on whether the neural network contains closed contour or not, we distinguish between cyclical and acyclic neural networks. A special category of the acyclic neural networks are layer networks, where the neurons can be divided into disjunctive subsets (layers). The output data from one layer are the input data to other layer. Neurons from the first layer get the information from the network input, whereas the neurons from the last layer produce output information from the network. In our method, we used layer networks.

### III. RETINAL BLOOD VESSEL SEGMENTATION BASED ON NEURAL NETWORKS

In this section, we describe the proposed method. Our method performs the detection and extraction of blood vessels from digital retinal images using the neural network. The method is comprised of four steps:

- a) preprocessing
- b) feature extraction,
- c) neural network in digital image processing,
- d) postprocessing.

#### A. Preprocessing

Digital image of eye fundus is a RGB image, Fig. 3a). RGB color model consists of three channels: red (R), green (G) and blue (B) [9]. The best contrast between the blood vessels and the background gives the green channel [3]. As a first feature for detection of blood vessels used the inverted green channel  $I_g$ , Fig. 3b).

In the first step, we used the morphological (MO) opening to prepare the inverted green channel for the extraction of other feature. The MO is used to remove noise from binary images and degradation noise from non-binary images [9]. In our case, we used the MO to remove the optical disk and reduce the background influence from the inverted green channel. The MO with structural element "square", with a width of 9 pixels, has been used, which perceives blood vessels as noise and is obtained experimentally as the optimal element. This way, we can estimate the image background  $I_m$  (Fig. 3c). Then, we calculate the difference  $D$  between  $I_g$  and  $I_m$  (Fig. 3d) in order to remove the optical disk and the background from the image.

#### B. Feature extraction

In this step, we obtain the other two features from new preprocessed image  $D$ :

- a) Basic line detector [1].
- b) Orthogonal line detector [1].

Basic line detector represents the probability that the observed pixel belongs to a line on the image. This detector is calculated for each pixel in the image. The basic line detector calculates the average value of the line for 12 different orientations. We take a mask of  $15 \times 15$  pixels around the observed pixel (observed pixel is the central pixel in mask), then we rotate the line with an angle of  $15^\circ$  relative to the central pixel in the mask. In this way,

we obtain 12 lines with different orientations. The length of the line is 15 pixels. After calculating average of 12 lines in the mask, we determine the line with the largest average for the observed pixel, and subtract the average value of the mask from this value. Fig. 3e) presents the result of the basic line detector, second feature. With this detector, we have two errors:

- errors on border pixels between a blood vessel and background inside FOV,
- errors on the FOV border.

The first error is reduced using orthogonal line detector. The second error is a consequence of the fact that the pixels outside the FOV are similar to the blood vessel pixels. It occurs when the considered mask includes out-of-the-FOV pixels. To resolve this problem it is necessary to replace the out-of-the-FOV pixels with the average preprocessed level of the remaining pixels in the mask.

Orthogonal line detector resolve the problem between the boundary pixels, i.e. pixels between the blood vessels and the background. This feature is calculated as the basic line detector. When we find a line with the largest average value, we consider an orthogonal three-pixel line centered on the midpoint of the line with the largest average value [1]. Again, the average value of the mask is subtracted from this value. The same mask is used as in the basic line detector. Fig. 3f) presents the result of orthogonal line detector, third feature.

To make the system converge more quickly, the features are normalized as:

$$\bar{x}_i = (x_i - \mu_i) / \sigma_i, \quad (4)$$

where  $x_i$  is the  $i$ -th feature,  $i \in \{1,2,3\}$ , and  $\mu_i$  and  $\sigma_i$  are the corresponding mean value and standard deviation, respectively.

### C. Neural network in digital image processing

The structure of the system is represented in Fig. 2. Neural network with one hidden layer is used. Input layer has three neurons (3 features) and in output layer we have one neuron. As already mentioned, linear function is used as network function and unipolar sigmoid is used as activation function.

Let  $W$  and  $V$  represent the vectors of neuron coefficients in the hidden layer and the output neuron, respectively. Then, the output of the neural network in the  $k$ -th step is

$$o_k = f(VO_k) = f(Vf(WX_k)), \quad (5)$$

where  $X_k$  and  $O_k$  represent vectors of input features and output of the hidden layer in the  $k$ -th step, respectively.

To minimize the error (3), the LMS algorithm is used to modify coefficients of neurons as follows:

$$V_{k+1} = V_k - \mu \frac{\partial E_k}{\partial V}, \quad (6)$$

$$V_{k+1} = V_k + \mu(d_k - o_k)o_k(1 - o_k)O_k^T. \quad (7)$$

$$W_{k+1} = W_k - \mu \frac{\partial E_k}{\partial W}, \quad (8)$$

$$W_{k+1} = W_k + \mu(d_k - o_k)o_k(1 - o_k)V_k[O_k * (1 - O_k)]X_k^T. \quad (9)$$

Fig. 4 represents the convergence of the LMS algorithm. The system training is carried out through 500 epochs. In each iteration, we update network coefficients depending

on the error (3). The number of iterations in one epoch equals the number of training data.

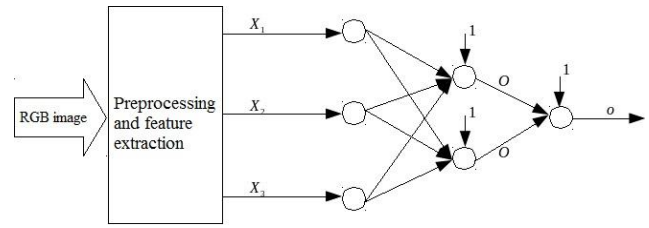


Fig. 2. Structure of the system.

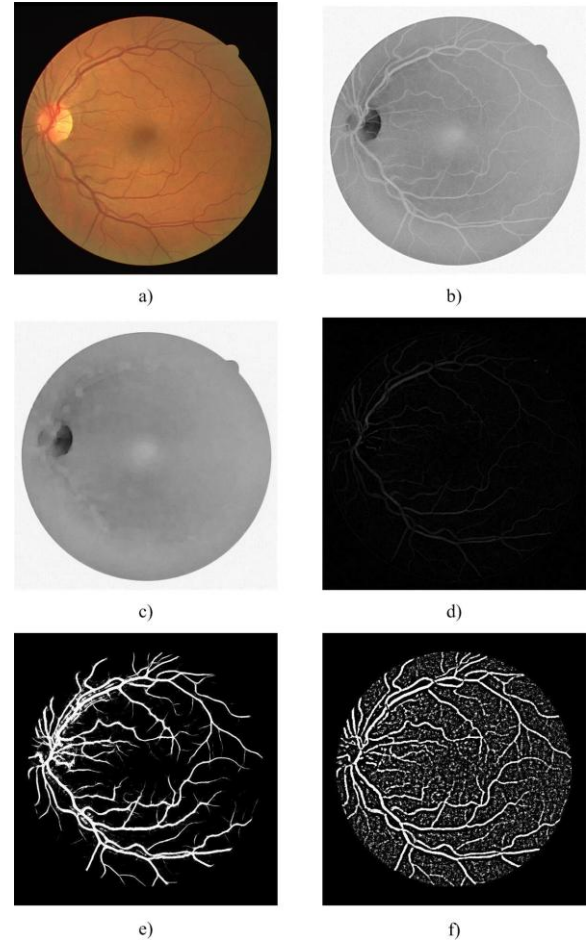


Fig. 3. a) RGB DRIVE image, b) Inverted green channel Ig, c) Image background Im, d) Difference D between Ig and Im, e) Basic line detector, f) Orthogonal line detector.

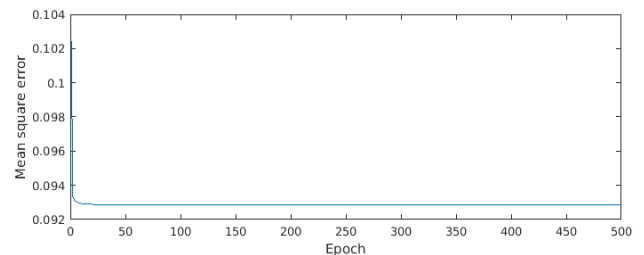


Fig. 4. Convergence of the LMS algorithm.

### D. Postprocessing

In the postprocessing step, we reclassify the incorrectly classified pixels. Here we distinguish two kinds of errors:

- errors in the pixels of blood vessels,
- errors in the background pixels.

In the first case, if the observed pixel represents a pixel

of the blood vessel, and all of its neighboring pixels represent background pixels, then the observed pixel is re-classified as a background pixel.

In the second case, we consider 3-pixel lines (directions  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  and  $135^\circ$ ) in a  $3 \times 3$  square mask. If the central pixel is classified as a background pixel, and the other two line pixels represent blood vessel pixels, then the observed pixel is re-classified as a blood vessel pixel.

#### IV. EXPERIMENTS AND RESULTS

The system is trained using training images from the DRIVE database and tested on test images from the DRIVE database and training images from the STARE database. Randomly extracted 20000 manually segmented pixels (10000 blood vessels pixels and 10000 background pixels) from the available training set of 20 images are used for training. The system accuracy evaluated on the test images classified by the first and the second experts are 94.1% and 94.6%, respectively. For the STARE database, the accuracies are 93.6% and 92%, respectively.

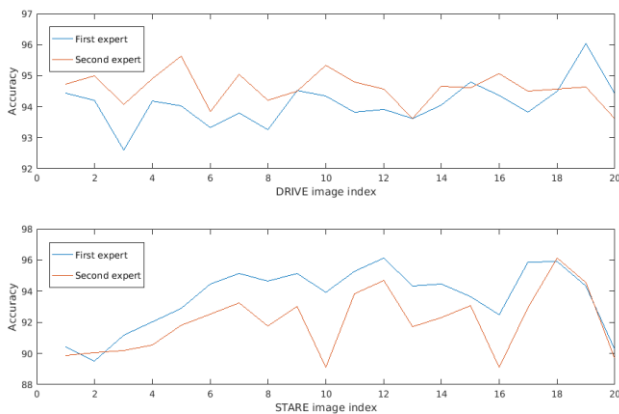


Fig. 5. Accuracy for all training images from DRIVE (top) and STARE database (bottom).

In Fig. 5, the accuracy of the proposed method with respect to image index, has been given, for both databases. Fig. 6 depicts the blood vessels segmentation of one image from the DRIVE database performed by both experts and the proposed method.

In Table 1, we provide performance comparison of the proposed method with the other methods in this field. For the DRIVE database, the proposed method outperforms the other ones, except the method [5], which provides higher accuracy. On other hand, the proposed method is outperformed by the other methods for the STARE database, since several STARE images have diabetic retinopathy (DR) [12] and line detector recognises DR as blood vessels. Note also that “---” in Table 1 denotes a value not provided in the corresponding paper.

#### V. CONCLUSION

We proposed a method for extraction of blood vessels from eye-fundus digital image by using neural networks. The features are obtained by using line detectors that operate on a preprocessed image. Neural network with hidden layer is used for pixel classification. Finally, in order to eliminate obvious errors, pixel reclassification has been performed based on the value of adjacent pixels.

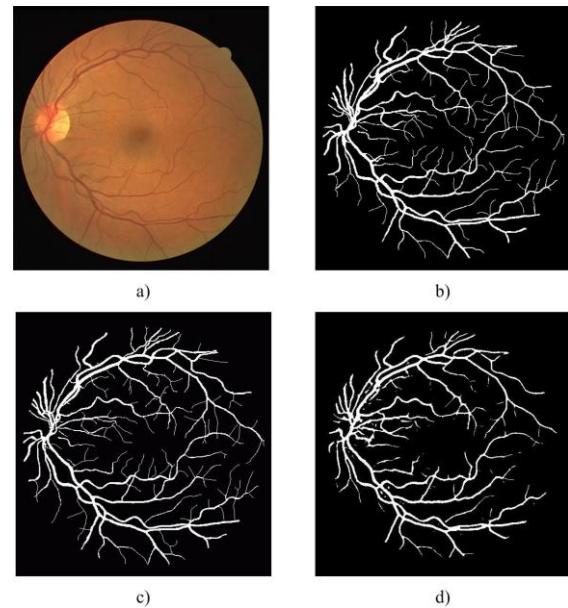


Fig. 6. a) RGB image DRIVE database. b) First expert. c) Second expert. d) Proposed method.

Table 1. Accuracy comparison of several methods

Method	DRIVE	STARE
Marin et al. [3]	0.944	0.952
Cinsdikici and Aydin [4]	0.929	---
Fathi and Naghsh-Nilchi [5]	0.958	0.959
Fraz et al. [2]	0.943	0.944
Proposed	0.946	0.936

#### REFERENCES

- [1] E. Ricci and R. Perfetti, “Retinal Blood Vessel Segmentation Using Line Operators and Support Vector Classification,” *IEEE Trans. On Med. Imag.*, vol. 26, no. 10, pp. 1357–1365, October 2007.
- [2] M. M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A.R. Rudnicka, C.G. Owen, and S.A. Barman, “Blood vessel segmentation methodologies in retinal images - A survey,” *Computer Methods and Programs In BioMedicine*, vol. 108, no. 1, pp. 407-433, October 2012.
- [3] D. Marin, A. Aquino, M. E. Gegúndez Arias, and J. M. Bravo, “A New Supervised Method for Blood Vessel Segmentation in Retinal Images by Using Gray Level and Moment Invariants Based Features,” *IEEE Trans. On Med. Imag.*, vol. 30, no. 1, pp. 146–158, January 2011.
- [4] M. G. Cinsdikici and D. Aydin, “Detection of blood vessels in ophthalmoscope images using MF/ant (matched filter/ant colony) algorithm,” *Computer Methods and Programs In BioMedicine*, vol. 96, no. 2, pp. 85-95, November 2009.
- [5] A. Fathi and A. R. Naghsh-Nilchi, “Automatic wavelet-based retinal blood vessels segmentation and vessel diameter estimation,” *Biomed. Sig. Proc. and Cont.*, vol. 8, no. 1, pp. 71-80, January 2013.
- [6] M. M. Fraz et al. “An approach to localize the retinal blood vessels using bit planes and centerline detection,” *Computer methods and programs in biomedicine*, vol. 108, no. 2 pp. 600-616, November 2012.
- [7] M. Bajčeta, P. Sekulić, S. Djukanović, T. Popović, and V. Popović-Bugarin, “Retinal blood vessels segmentation using ant colony optimization,” *Neural Networks and Applications (NEUREL)*, November 2016.
- [8] R. O. Duda, P. E. Hart, and D. G. Stork. “Pattern classification, 2nd Edition,” Wiley-Interscience, November 2000.
- [9] R. C. Gonzalez and R. E. Woods, “Digital Image Processing (3rd Edition),” Pearson, August 2007.
- [10] W. S. McCulloch and W. Pitts, “A logical calculus of the ideas immanent in nervous activity,” *The bulletin of mathematical biophysics*, vol. 5, no.4 pp. 115-133, 1943.
- [11] Lj. Stanković, “Digital signal processing,” Amazon.com Company, 2015.
- [12] H. R. Taylor and J. E. Keefe, “World blindness: a 21st century perspective,” *British Journal of Ophthalmology*, vol. 85, no. 3 pp. 261-266, 2001.

# Improving the Security of Access to Network Resources Using the 802.1x Standard in Wired and Wireless Environments

S. Kovačić, E. Đulić, and A. Šehidić

*Abstract – Unauthorized access to network resources in a typical TCP/IP networks can be achieved in a simple way. When a user (attacker) accesses the local area network, he may initiate, knowingly or unknowingly, the attack on the servers, eavesdrop the network, activate various malicious programs or Denial of Service (Dos) attacks. 802.1x allows us to solve this problem in both the wireless and the wired environment. In the case of using 802.1x for access to the edge networks the network administrator can easily detect unauthorized access and for authorization is used a central authentication server. In order to use the standard 802.1x client and switch must have support for this standard.*

**Key words — IEEE 802.1x, Security, Network.**

## I. INTRODUCTION

WE are relying solely on the physical security of network access, to limit the possibility of eavesdropping the switched network environment. However, an unauthorized user can use the Address Resolution Protocol (ARP) redirect to trick switch and allow eavesdropping attack. In this way, an unauthorized user can access to computer networks of the institutions and attack network resources that are located in their corporate network, such as printers, internal web servers and other devices. Securing DHCP servers can partially solve this problem, but it does not mean that an unauthorized user cannot manually assign an unused IP address and execute the attack. Implementation of 802.1x in the network environment allows us to solve the above problems. Each user must perform authorization to access the network.

IEEE 802.1x is a standard that is defined by the Institute of Electrical and Electronics Engineers (IEEE), which is designed to give us a port-based network access [1].

Port-based access control to network resources allows the system administrator to restrict unauthorized users access to LAN services. IEEE 802.1X standard defines the

architecture, protocols and functional elements that are used for authentication between the client and the switch port to which it is physically connected [1], [2].

## II. ELEMENTS 802.1X ENVIRONMENT

Because standard 802.1x provides Layer 2 (L2) access control using validation user or device which attempting to access the physical port, as they usually switch, access point or other network edge device. Basic 802.1x mechanism consists of three components, namely: **supplicant, authenticator, and authentication server.**

**Supplicant** is a device client which trying to access network resources. Supplicant may be a desktop computer, Notebook, tablet, smartphone or other device that has a need for access to network resources.

**Authenticator** represents a device that cuts supplicant request for access to network resources, it is frequently switch, access point or other network edge device.

**Authentication server** compares the client (supplicant) ID with credentials that are stored in the database. If the credentials and supplements ID match, the client gets access to network resources.

**Port Access Entry (PAE)** refers to a method of authentication algorithms and the associated physical port to the network devices. PAE represents the 802.1x logical component of the supplicant and authenticator that exchange EAP messages.

**Extensible Authentication Protocol (EAP)** is used between the client and the authenticator. In depending on the type of media that is used to transfers EAP messages used other types of encapsulation. There are some methods of the encapsulation:

- **EAP over LAN (EAPOL):** This method of encapsulation is used for data transmission over LAN media such as Ethernet, FDDI or Token Ring.
- **EAP over Wireless (EAPOW):** This explains how the packets are encapsulated when over wireless networks [3]-[4].

802.1x supports three types of authentication modes: **single, single-secure or multiple.** Single authentication mode allows only the first client to send a request for authentication, while the single-secure mode only one user can be authenticated and access to network resources other users cannot be authenticated until the first user completes its network session. Multiple mode is most often used when we need to on a physical port to connect more than

S. Kovačić, University Information Centre, "Džemal Bijedić" University of Mostaru, University Campus bb, 88104 Mostar, Bosnia and Herzegovina (phone: 387-36-281040; e-mail: salko@unmo.ba).

E. Đulić, Faculty of Information Technologies, "Džemal Bijedić" University of Mostaru, University Campus bb, 88104 Mostar, Bosnia and Herzegovina (phone: 387-62-006391, e-mail: eminao@fit.ba).

A. Šehidić, Faculty of Information Technologies, "Džemal Bijedić" University of Mostaru, University Campus bb, 88104 Mostar, Bosnia and Herzegovina (phone: 387-36-281164, e-mail: admir@fit.ba).

one client device, but each device authenticates individually.

As an authentication server commonly used RADIUS. This server provides secure authentication and a simple way can be connected with other identity databases within the organization, such as LDAP and Active Directory and use already existing user credentials.

III. AUTHENTICATION PROCESS WITH 802.1X

Client and switch or access point must support 802.1x, and this standard has to be turned on (enabled) before the start of the authentication process. Before any traffic on the network client must have successfully completed the process of authentication, including DHCP traffic.

Before authentication, only the messages that are acceptable from the client is EAP messages and forwarded to the authentication server. The authenticator's PEA is set to uncontrolled stat, as shown in Figure 1: Before 802.1x authentication, in this status all other network services are disabled.

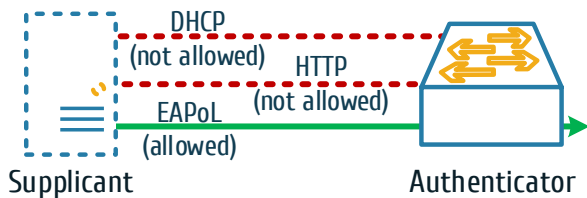


Figure 1: Before 802.1x authentication

After 802.1x client is turned on, it will execute the sending EAP messages to the switch or access points in the wireless environment. The network device (switch or access point) to forward client request to the authentication server. Authentication server shall verify access data that is received and send the response switch, which shall decide whether the port from which the message is addressed remain uncontrolled stat (access denied) or shall change in the controlled stat (access granted). If the authentication server with the message "access granted", then the port it is connected to the client to be changed to "controlled stat" and only after that shall enable network services as shown in Figure 2: After 802.1x authentication, otherwise the port will remain "uncontrolled stat", depending on the network equipment can be disabled.

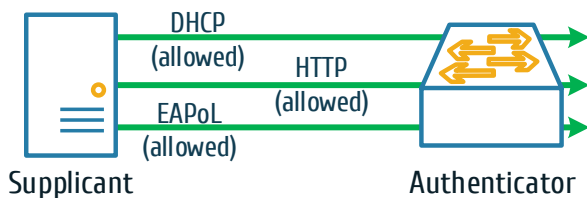


Figure 2: After 802.1x authentication

However, if during boot-up, the client does not receive an EAP-request or identity frame from the authenticator, the client can initiate authentication process by sending an EAPoL-Start frame. This frame prompts the switch or

access point to request identity from the supplicant.

When the supplicant supplies its identity, the authenticator directly exchanges EAPoL to the supplicant until authentication succeeds or fails. In case that the authentication is successful, the port becomes authorized. If the authentication fails, the port becomes unauthorized. When the client does not need network any more, it sends EAPoL-Logoff packet to terminate its 802.1x session. In this case the port state will become unauthorized. The following in Figure 3: 802.1x Authentication message flow, shows the EAPoL exchange ping-pong chart.

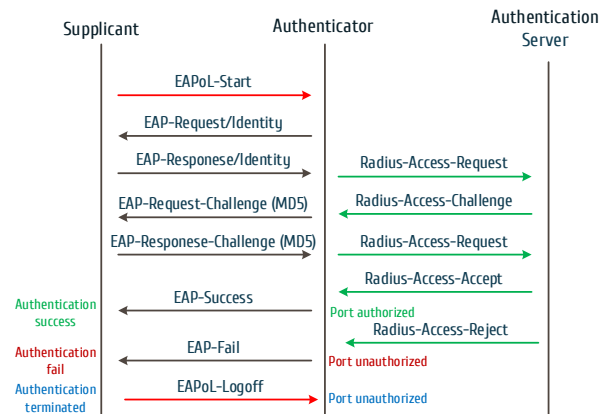


Figure 3: 802.1x Authentication message flow

IV. EAP MESSAGE FORMAT

EAP over LAN (EAPoL) is an encapsulation technique used to protect communications between the client and the authenticator (switch or access point). EAPoL protects those communications that occur only before authentication.

EAP methods define message format used for communication between supplicant and authentication server. Some of them, such as LEAP are not used commonly due their weakness. But there are some other strong EAP methods such as EAP-TLS, PEAP, and EAP-TTLS. All these methods have a different way in order to protect the credentials sent from the supplicant to the authentication.

Protected Extensible Authentication Protocol (PEAP) and variants of them are widely used. One of them named as PEAP-MSCHAPv2 is popular with Windows clients. EAP-TLS is used both client and server certificates, which were presented and approved at any time of communications. However, because certificates price for individuals and small organizations, EAP-TLS is not used frequently. EAP-TTLS creates a secure, encrypted tunnel through which the switch passes the EAP messages. In this case, the client-side certificates are optional, which has made this messages format very popular.

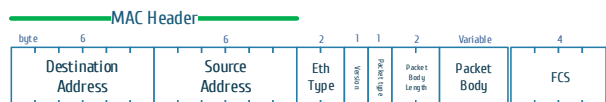


Figure 4: EAP message frame

## V. BENEFITS OF 802.1X IMPLEMENTATION

802.1X is an IEEE standard originally published over many years ago and as a result it is supported almost today's network infrastructure devices like managed switches, wireless access points and controllers, etc.

Implementing of 802.1x standards offers many advantages in combination with the supported network technology, but there are some limitations or disadvantages.

Some of advantages are:

**VLAN Assignment:** This feature allows authorized users to join a pre-configured VLAN. This option allowing you to maintain the name-to-VLAN links to the server authentication. In this case, the authenticity using the 802.1x standard ports assigned to a VLAN based on user identity.

**Port Security:** This option is used to set security port, and only specific the Media Access Control (MAC) address was enabled on a port, while all other addresses do not have access to the network. This eliminates the security risk that additional users have access to the switch without authentication.

**Guest VLAN:** This feature allows users who try to connect to the network without 802.1x that their device are assigned to Guest VLAN.

**ACL Assignment:** Access Control List (ACL) enables dynamic assignment of access control for each interface-based 802.1x user authentication. This option allows certain users to be denied access to some of the network segments, such as servers, specific protocols or applications.

**Encryption of Wireless Keys:** IEEE standard 802.1x enables encrypted communications over wireless networks using dynamic coding. In this case, the authentication server is responsible for providing keys on both sides Authenticator and supplicant to take advantage of the option dynamic WPA/WPA2 security [5].

**Strong Authentication:** This standard (802.1x) use EAPoL for advanced security mechanisms for user authentication. EAPoL uses a different mechanism to authenticate, which allow more flexible choice for the organization to implement the best solution for their environment.

**Secure Access Control:** Network security is further enhanced with 802.1x authentication because it forces it to happen before you are allowed to access the network. All ports on switch or network access points on wireless can be configured to remain in the unauthorized state until successful authentication is completed. This helps to ensure that only authorized users and devices are allowed to access network resources.

## VI. HOW 802.1X CAN HELP

IEEE 802.1x standard for port-based authentication is typically used in large networks, both wired and wireless, and is commonly used in the campus or enterprise networks.

### A. Simplified management

This protocol allows the use of certificates and/or user access credentials to authenticate access to network resources. This method is much easier to expand than other authentication methods such as pre-shared keys (PSK) wireless network.

IEEE 802.1x standard allows the use of certificates (EAP-TLS) or user access credentials (EAP-MSCHAPv2). This is a good omen for system administrators, as it easily can be managed from one place.

This standard can use the existing infrastructure like the Microsoft domain environment with users, passwords, groups which are located in Active Directory. In addition, it is possible to use open source services like OpenLDAP and FreeRADIUS in a network environment that provides user authentication process when accessing network resources. 802.1x is based on the mentioned database and that simplifies the deployment and administration of the system.

### B. Easier to use

From a user perspective, this standard makes it easy to use, because when accessing network resources, the user is prompted to enter user credentials. This usually happens only once and never again you are not bothered until you change your password or certificate.

The 802.1x authentication process on the user side is performed once, and after that the rest of the process occurs in the background, unlike captive portals that logged in once every several times.

When using authentication with certificates, a device that authenticates presents itself to the authentication server, client and server present their certificates for verification devices. The entire process is invisible to the end user and allows the use of devices with the smooth process

### C. Supported device

Almost all of the device on the market that have the ability to connect to wireless with support for 802.1x standard. Some printers and other devices can be excluded from using these types of authentication, but most recently supported 802.1x standard.

Support for 802.1x can be found on all Windows OS, but MacOS this standard supports from their eighth generation. Most of the known Android and iOS mobile phones, tablets and other smart devices have native support for the IEEE 802.1x standard.

### D. Scalability

In this case, we have possibility to manage users, certificates, devices and other profiles from one place, and simplified and automatically assign rolls to groups.

### E. Security

As we speak, 802.1x authentication keys are exchanged between the device and the server to verify, and all the keys are individual and not shared like PSXs.

In case, if you are using 802.1x authentication in conjunction with AES encryption ensures keys are well

protected from being hacked. If your wireless network supports CoA is possible to block access to the network in the case of a change their credentials or certificates after having logged. The user will be automatically disconnected from the network.

The attributes that are used during authentication, which come from the authentication server can be used to join certain rules to the user, such as Student, Staff, Guest and the like. This allows us to rules and properly wireless solution can be applied network access rules. This is commonly known as Role Based Access Control (RBAC).

#### F. Federation

Wireless scalability is a challenge and decentralized environments where many groups and departments run their own network or manage their accounts.

Centralizing services to solve the problem is not always possible, and often can be undesirable. Even highly decentralized groups can develop a scalable, secure wireless network infrastructure via 802.1x and Radius. The only additional components required in addition to those required for standard 802.1x is a trust relationship between the RADIUS server and the core management of the relationship of trust and routing application for authorization.

The resulting collection of loosely connected network is often called a federation. Federated network consisting of several networks of countries that share some level of trust, but the network member retains its administrative control.

### VII. ATTACKS IN 802.1X ENVIRONMENT

The availability of wireless networks is much higher to attacker then access to wired networks. For this reason, wireless networks have more possibilities to attack. IEEE 802.1x standard provides us a strong authentication protocol, but because of its disadvantages in practice is much more used in wireless environment than wired networks [6], [7].

Some of attacks in wireless environment:

**802.1x RADIUS Cracking:** This attack method attempts to brute force to discover the secret key for 802.1x access request to authentication server, to use evil twin access point. Evil Twin Access Point is a method of presenting as one authorized access point. This access point is lure for the users.

**802.1x EAP Replay:** In this case, the hacker performs recording 802.1x Extensible Authentication Protocols (e.g. EAP identity, successful and unsuccessful attempts to access) for later iterations. This method uses a wireless captures and injection tools between client and access point.

**802.1x RADIUS Replay:** In this case, attacker trying captures authentication server Access-Accept or Reject messages for later replay.

**802.1X Password Guessing:** This attack used a captured identity, repeatedly trying 802.1x authentication to try to hit a user's password.

**802.1x EAP-Start Flood:** This attack use flooding an access point with EAP-Start messages to consume resources or crash the target.

**802.1x EAP-of-Death:** This attack sends incorrect 802.1x EAP Identity response that causes a crash on some access points.

### VIII. CONCLUSION

In today's enterprise networks, security is the biggest challenge for all companies, cloud operators and other institutions. One of the big problems in building computer systems represents limitation unauthorized access to network resources, and avoiding errors in the implementation of computer systems and network resources. IEEE 802.1x standard is preferred as a unique way to protect access to resources and can be used for the wired and wireless networks.

As we talk about, 802.1x can solve problems access to the edge networks and network administrator can easily detect unauthorized access and for authorization from a central authentication server. In order to use the standard 802.1x client and switch must have support for this standard.

This standard can use the existing infrastructure like the Microsoft domain environment with users, passwords, groups which are located in Active Directory or other directory services such as LDAP or RADIUS.

Wireless networks have more possibilities to attack. IEEE 802.1x standard provides us a strong authentication protocol, but because of its disadvantages in practice is much more used in wireless environment than wired networks.

### REFERENCES

- [1] IEEE 802.1x Standard, Port-Based Network Access Control, <http://standards.ieee.org/getieee802/download/802.1X-2010.pdf>
- [2] 802.1X and NAC: Best Practices for Effective Network Access Control
- [3] M. Younus, "Wired 802.1x security", SANS Institute, 2000.-2005.
- [4] P. Ding, J. Holliday, A. Celik, Improving the security of wireless LANs by managing 802.1x disassociation, 2004
- [5] N. Hermaduant, I. Riadi, Automation framework for rogue access point mitigation in ieee 802.1X-based WLAN, 2016
- [6] S. H. Ahmadpanah, A. J. Chashmi, A. Ameri, WLAN Security and Applied Research in 802.1x Protocol, 2016
- [7] S. Behal, K. Kumar, Characterization and Comparison of DDoS Attack Tools and Trac Generators - A Review, 2016

# The use of information technology in the healthcare systems, case of Serbia

Armin Lucević, Faculty of Organizational Sciences and Faculty of Medicine, University of Belgrade,  
Lejla Taraniš, Faculty of Medical Sciences, University of Kragujevac

**Abstract** — This paper analyses Integrated Health Information System - IHIS that is financed by European Commission (through Instrument for Pre-Accession Assistance - IPA) and implemented in Serbian healthcare system. We observed previous IT related projects implemented in Serbia, and on those basis reviewed IHIS. As "going live" phases started recently, not all the effects can be perceived.

**Key words** — Health care system, IHIS, IPA, Serbia

## I. INTRODUCTION

IT is impossible to imagine modern life without information technology. Technology has entered every segment of life, and the information is available to us within one's reach. The technological revolution has not outflanked neither the public nor the health sector, so nowadays, the computer has become a basic tool in each health institution.

There are a lot of advantages: faster information approach, expenditure reduction, more precisely planning, making decisions in the right time, increase of efficiency and development.

This research paper is organized in a manner that Chapter II and Chapter III are covering brief overview of IT projects implemented in Serbian Healthcare system and details about Integrated Health Information System (EU-IHIS). Chapter IV is dedicated to Implementation of Integrated Health Information System.

Research is based on the statistical data collected from World Bank, official project documentation and reports of the representative agencies hired by European Commission for evaluation purposes.

## II. DELIVERED PROJECTS IN SERBIAN HEALTHCARE SYSTEM

Healthcare systems have been under fire of the technological innovations that took place in early 90s, especially those, more intense, after 2000. Using

Armin Lucević (author for contacts), Faculty of Organizational Sciences and Faculty of Medicine, University of Belgrade, Doktora Subotića 8, 11000 Beograd, Serbia (cell phone: +36/31/323-8346; email: armin.lucevic@gmail.com).

Lejla Taraniš, Faculty of Medical Sciences, University of Kragujevac, Svetozara Markovića 69, 34 000 Kragujevac, Serbia (phone: +381/34/306 800; email: taranis.lejla@gmail.com).

information system in the health services, leads to the term "health information system." The World Health Organization has defined health information system as the following:

*"The health information system provides the underpinnings for decision-making and has four key functions: data generation, compilation, analysis and synthesis, and communication and use. The health information system collects data from the health sector and other relevant sectors, analyses the data and ensures their overall quality, relevance and timeliness, and converts data into information for health-related decision-making" [1].*

From this definition, we can conclude that health information system is the base for making decisions. There are four main components: *generating the information, collecting, analysis and usage*. It should be noted that the health information system does not collect data only from the health sector, but also from other sectors which may be relevant. After that, the data is processed, checked for their quality and relevance, and ultimately converted into usable information that will be used in planning and decision-making.

The world health organization has divided the usage of data generated through the health information system in four levels [2]:

1. Individual level
2. The level of health care institutions
3. Population level
4. The level of public health

The commencement of the introduction of information technology in the health system of Serbia occurred in the beginning of the 80s. Then the management of some of the financially stronger health institutions provided computers in facilities that they led. The idea was to create an IT team that will enable the process of collecting and using data more efficiently. First information systems were used in finance, human resources, as well as for monitoring the purchase and consumption of remedies.

Due to the crisis in the 90s the process of informatization of health care was diminishing.

The next big wave of the introduction of information systems in the health system of Serbia began after 2000s. Then with the help of the World Bank, EU funds and other donors, several significant projects were completed:



1. Serbia Health Project [3] - Serbia Health project aimed at capacity building for tenable health system, oriented on performances where the quality and efficiency are rewarded. Through this project, it should be provided approachable and efficient health care insurance. Serbia Health project was funded by the World Bank in the amount of 23.48 million dollars. The project lasted 9 years (May 22, 2003 - March 1, 2012). One part of the funds was spent on the construction of the Central Information System.

2. Serbia Health Additional Financing [4] - Serbia Health Additional Financing was funded by the World Bank with \$ 13.5 million. The project started on March 19, 2009 as additional financial support to Serbia Health Project. The project aimed reforming the payment system in hospitals and improving the quality of health care. The money was spent on the purchase of computer equipment and software. This project is supported by institutional reforms as well as reforms of management, with the aim of increasing efficiency and building health system that is financially tenable.

3. Delivery of Improved Local Services Project [5] - Provision of improved services at the local level was initiated on 18 March 2008 and was completed on 1 March 2015. It was funded using the World Bank loan of 46.4 million dollars. The main objective of the project is to build institutional capacity able to provide better health care services, education and social protection at the local level. Through the construction of decentralized environment, it is planned to increase the efficiency of the system.

4. The implementation of hospital information system [6] - was funded through IPA 2008 with a budget of 7.5 million euros. The aim of the project is the integration of data basis of health facilities and the establishment of an advanced hospital information system in 10 hospitals, covering the needs of 30% of the population of Serbia.

### III. INTEGRATED HEALTH INFORMATION SYSTEM

Integrated health information system [7] was launched in 2012. Funded by the European Union through the IPA funds with a budget of 2.5 million euros. The project is implemented by the Ministry of Health of the Republic of Serbia, the European Union and by the Regional Office of the World Health Organization for Europe - Office in Serbia, with administrative support from the United Nations Office for Project Services (UNOPS).

The goal of the project is the implementation and merging information systems of 19 hospitals into a single IT system of electronic medical records. This kind of integration should lead to increased efficiency and better quality of services that will be provided to the patient. From the standpoint of profitability, hospitals will have more market-oriented model of functioning. With implemented IT solutions, the internal communication and access to information will be improved and double data entry will be avoided, which recently was an everyday

occurrence. With the help of IHIS, tracking and invoicing will be simplified, while the management of the hospitals will have a better insight into the costs and will be able to make strategic decisions easier.

The Ministry of Health of the Republic of Serbia has managed to develop the source code of the information system during the implementation of previous projects that were related to the computerization of the health system. For the implementation, maintenance and further development of software, two companies are in charge, ComTrade (for HIS1) and Bit Project (for HIS1), which are originally Serbian.

The requirement to get started was that in 19 selected hospitals infrastructure already existed in terms of computer (LAN) network; and that hospitals are equipped with computers.

This requirement is taken as a given, because the previous project funded through IPA 2008, computer equipment should have been purchased, and computer network capable and advanced (if there was any need for it).

The benefits that were obtained using the health information system:

- Data entry is in the place where it is created
- All data inputs are centralized in one place
- Information about the patient, doctor etc. are easy to find
- Medical history of each patient is easy reachable
- Simplified monitoring of the medicines and the equipment consumption
- Paper is slowly excluded from usage
- Higher level of service that is provided to the patient
- Increased efficiency of the hospital (and management)
- Better quality of the accessed information
- Reduction of costs

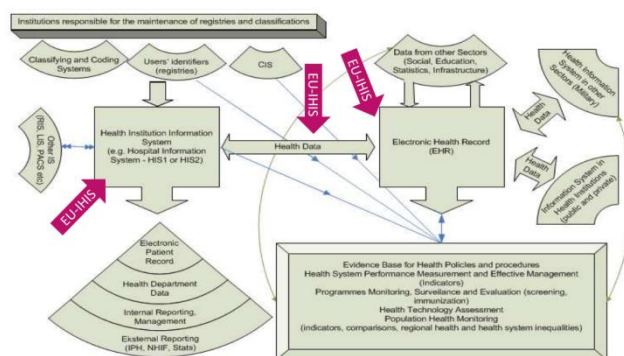


FIGURE 1. DATA FLOW - INTEGRATED HEALTH INFORMATION SYSTEM [8]

Figure 1. is showing data flow in IHIS. We can distinguish two main units: Health Institution Information System (HIS1 or HIS 2) and Electronic Health Record that are interconnected with health data. Other smaller units are connected to Health Institution Information System and Electronic Health Record.

Change that integrated information health care system has brought is great and profound. It permeates every part of the health system and its full implementation to be compromised if it is not properly planned.

One of the most important roles in the implementation of the new system is subordination to the change management. Change management as "*a set of principles, techniques that apply to employees in order to implement significant organizational changes*"[9]. This is part of management that is focused on the implementation process of the new systems so that the process of transition would be simple and with the least negative effects. The implementation of an integrated health information system consists of two factors, human and technological. Technology refers to the equipment, software and other pre-conditions for the use of new technologies. Human factor is the one that implements the changes in real time and from which success of a project depends mostly. In the case of health systems, there are great challenges for human factor.

We have identified couple of factors that can help with project implementation:

**Leadership** – Individuals who are leading projects should be in coordination with management of the medical institutions where new system will be installed. The flow of the information should be well developed in order to avoid misunderstandings. Managers who are leading hospitals should be professionals who's computer literacy is higher than average. With this type of skills, they will be able to adopt new processes and become an example and support for other employees during the implementation process.

Average age of the employees in Serbian hospitals is quite high and level of their computer literacy is low; in pre-implementation process, employees will have to go through different training (not only for IHIS) in order to increase their level of computer literacy.

Management and leadership should be prepared and ready to see the benefits of IHIS projects on long run, as negative comments and critics will come directly from individuals who will benefit from it (patients and employees of different medical institutions).

**Usability** – is one of the biggest challenges that management will face during IHIS implementation. It will be very difficult to show benefits on mid / long run of the project. Especially to advocate how IHIS will increase efficiently, cut the costs and provide better care for the patients, and as the most important, to enable employees to finish their work easier and faster.

Implementation of the new process – Implementation of IHIS project is planned to last for 3 years, which is sufficient for new processes and routines to be adopted by the management and medical staff. The transition process will be "the most painful", as in one moment the data entry will be generated through information system and as a hard copy. It is important that medical staff is aware of the transition process duration and to be clearly stated that

double data entry is just temporary until full migration to information system.

**Equipment** – Through previous IPA projects, computer equipment should be already delivered and installed in the medical institution where IHIS will be implemented. It is also taken by default that computer network is set and infrastructure is ready to be utilized. By the project design, personnel whose work is closely connected with computer usage (finance, procurement, planning, etc.), should have their own device (desktops / laptops), and should not share the unit with other colleagues (only in special circumstances).

**Training and support** – as it is already mentioned, for successful implementation of any IT project, it is important that human factor is very well prepared. In the stage of the development where Serbian healthcare system is right now, it is too complicated to organize one unified training for all employees (medical staff). Medical personnel should be separated by the level of computer literacy and then groups for trainings can be created. This process should be integrated part of IHIS planning process. Personnel whose computer literacy is very low, or don't have it at all, should receive training including general knowledge of computers. On the other hand, personnel whose computer literacy is on higher level (usually younger generation), should have a training designed only for software that will be used. With this type of diversification, all employee groups will be covered and properly trained.

#### IV. IMPLEMENTATION OF INTEGRATED HEALTH INFORMATION SYSTEM

It was planned that implementation of IHIS would take three years (from 2012 to 2015), but not everything was done as per the schedule. Here are some of the identified problems that postponed implementation: lack of equipment (hardware) in medical institutions, lack of training for employees, as well as the synchronization of the old and new systems.

Web site *MojDoktor*: [www.mojdoktor.gov.rs](http://www.mojdoktor.gov.rs) was launched before medical institution started to use new system. *MojDoktor* portal is designed for chosen doctors to make appointments for his/her patients directly to specialist doctor or to certain diagnostic exams. Additional feature is a calendar that shows free time slots (and location) for specific type of examination (mammography, gastroenterologist, etc.). During the scheduling process, the chosen doctor will be able to assign priority for patients. Patients who need emergency treatment will have higher priority.

Depending on the urgency and the condition of the patient, there are five types of referrals:

1. Referral letter with date
2. Referral letter without date
3. Priority referral
4. Emergency referral
5. Control referral (with date)

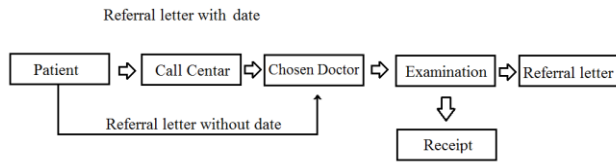


FIGURE 2. THE PROCESS OF MAKING APPOINTMENTS

The process of making an appointment in IHIS is shown at Figure 2. Call Center can be skipped if patient is directly visiting chosen doctor. Depending on examination results, patient gets e-receipt for medicine or referral letter for special examination. At the pharmacy, it is enough that patient shows the personal identification number or the social security number and the prescribed medicine will be provided. For chronic patients, it won't be necessary to visit the chosen doctor for prescription; they will be able to go directly to the pharmacy and medicine will be provided.

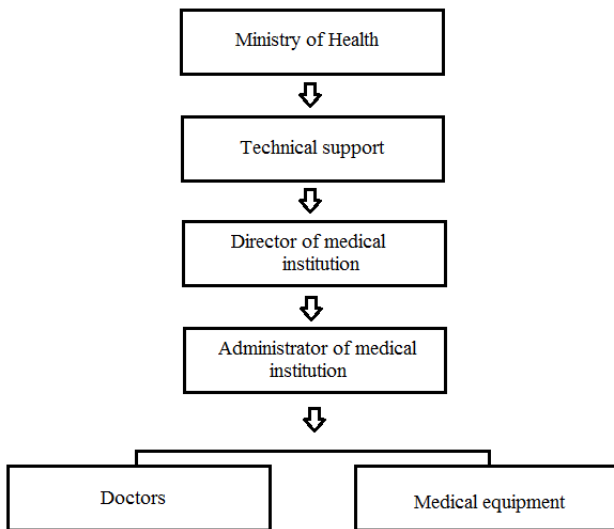


FIGURE 3. HIERARCHICAL RESPONSIBILITIES

Structure of hierarchical responsibilities is shown on Figure 3. Basically, on top of the pyramid is Ministry of Health of Serbia, followed by technical support of IHIS and directors of medical institutions. Administrators of medical institutions are responsible for managing examination schedule (free slots) and utilization of medical equipment.

The first institutions where IHIS was introduced were Health Centre Niš and the Clinical Center Niš. The main reason behind is that that medical institutions in Niš have taken part in similar projects (implementation of information systems), and employees already went through different trainings as well (they already have high level of computer literacy).

Integrated Health Information System was officially launched on 21. March 2016.

- April 18, 2016 IHIS was put into operation in the Clinical Centre in Kragujevac.

- May 5, 2016 IHIS was put into operation in the Belgrade region.

- May 25, 2016 IHIS was put into operation in the Novi Sad region.

V. CONCLUSION

Implementation of information technology in medical science will have big perspective in the future. More and more professionals will work in IT sector that is connected to health care, and medical professionals will rely on the diagnostics technology and more accurate medical treatments. Doctors will have the access to fill in disease history in order to provide better care for the patient.

Making an appointment or medical examination will be easier, and utilization of medical equipment will be optimized.

With implementation of information technology in healthcare systems, potential risk of pharmacological fraud will decrease, as doctors will have quick access to medicines that patient is using and with checking the e-record the best possible therapy will be prescribed.

REFERENCES

[1] World Health Organization (2008). Health Information Systems. [http://www.who.int/healthinfo/systems/WHO\\_MBHSS\\_2010\\_section3\\_web.pdf](http://www.who.int/healthinfo/systems/WHO_MBHSS_2010_section3_web.pdf) (Accessed: 20.12.2016.)

[2] World Health Organization (2008). Health Information Systems. [http://www.who.int/healthinfo/statistics/toolkit\\_hss/EN\\_PDF\\_Toolkit\\_HSS\\_InformationSystems.pdf](http://www.who.int/healthinfo/statistics/toolkit_hss/EN_PDF_Toolkit_HSS_InformationSystems.pdf) (Accessed: 20.12.2016.)

[3] World Bank (2012). Health Project Serbia. <http://www.worldbank.org/projects/P077675/health-project-serbia?lang=en> (Accessed: 30.11.2016.)

[4] World Bank (2011). Serbia Health Additional Financing. <http://www.worldbank.org/projects/P110593/serbia-health-additional-financing?lang=en> (Accessed: 20.12.2016.)

[5] World Bank (2015). Delivery of Improved Local Services Project. <http://www.worldbank.org/projects/P096823/delivery-improved-local-services-project?lang=en> (Accessed: 30.11.2016.)

[6] Government of the Republic of Serbia, National Program for Integration of the Republic of Serbia into the European Union 2007 – 2013. [http://www.seio.gov.rs/upload/documents/medjunarodna\\_pomoc/eu\\_inst\\_rument/ipa\\_2008.pdf](http://www.seio.gov.rs/upload/documents/medjunarodna_pomoc/eu_inst_rument/ipa_2008.pdf) (Accessed: 30.11.2016.)

[7] Web page: Integrated Health Information System (EU-IHIS). <http://eu-ihis.rs/> (Accessed: 20.12.2016.)

[8] Ministry of Health - Republic of Serbia, EU, WHO, UNOPS (2015). Integrated Health Information System: EU-IHIS Project Overview. [http://www.eu-ihis.rs/docs/Docs/TechDocs/EU-IHIS\\_Project%20Overview\\_Final\\_EN.pdf](http://www.eu-ihis.rs/docs/Docs/TechDocs/EU-IHIS_Project%20Overview_Final_EN.pdf) (Accessed: 20.12.2016.)

[9] Linda A. Anderson, Dean Anderson (2010). The Change Leader's Roadmap and Beyond Change Management. John Wiley & Sons Inc.

# Projektovanje spregnutih mikrotrakastih vodova

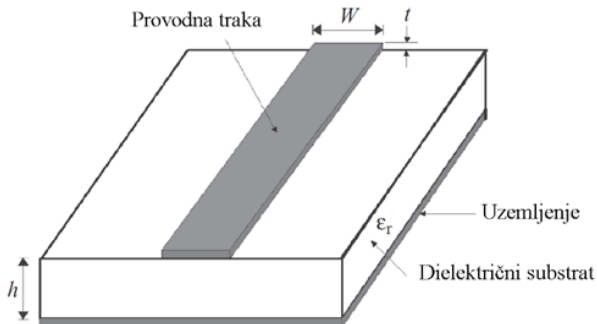
Arsenije Maliković, prof. dr Dragan Filipović, prof. dr Budimir Lutovac

**Sadržaj** — Predmet ovog istraživanja su mikrotrakasti spregnuti vodovi koji imaju značajnu primjenu u mikrotalasnim filtrima propusnika opsega. U radu ćemo pokazati da opisana metoda u sintezi spregnutih mikrotrakastih vodova daje dobre rezultate u projektovanju filtara sa bočnim rezonatorima. Provjera dobijenih rezultata biće predstavljena u NI AWR Desing Enviroment programskom paketu.

**Ključne riječi** — karakteristična impedansa, mikrotaladni filter, NI AWR, propusni opseg, spregnuti mikrotrakasti vodovi.

## I. UVOD

Mikrotrakasti vod je jedan od najpopularnijih tipova planarnih vodova, primarno iz razloga jednostavnog smanjivanja i integrisanja sa pasivnim i aktivnim mikrotalasnim kolima. Geometrija mikrotrakastog voda prikazana je na Sl. 1. Provodnik širine  $W$  i debljine  $t$  je odštampan na tankom, uzemljenom dielektričnom substratu debljine  $h$  i relativne permeabilnosti  $\epsilon_r$ .



Sl.1. Mikrotrakasti vod

Provodne karakteristike mikrotrakastih vodova se opisuju sa dva parametra, efikasnom dielektričnom konstantom  $\epsilon_{re}$  i karakterističnom impedansom  $Z_c$ . Za veoma tanke provodne trake ( $t \rightarrow 0$ ), sledeće jednačine za karakterističnu impedansu  $Z_c$  daju grešku do 1%:

$$\begin{aligned} \text{za } (W/h) \geq 1 \\ Z_c &= \frac{120\pi}{\sqrt{\epsilon_{re}} \frac{W}{h} + 1.393 + 0.677 \ln\left(\frac{W}{h} + 1.444\right)} \\ \text{za } (W/h) \leq 1 \end{aligned}$$

Arsenije Maliković, Elektrotehnički fakultet u Podgorici, Univerzitet Crne Gore, Džordža Vašingtona bb, 81000 Podgorica, Crna Gora (telefon: 382-67-610100; e-mail: arsenijemali@ac.me).

Prof. dr Dragan Filipović, Elektrotehnički fakultet u Podgorici, Univerzitet Crne Gore, Džordža Vašingtona bb, 81000 Podgorica, Crna Gora; (e-mail: draganf@ac.me).

Prof. dr Budimir Lutovac, Elektrotehnički fakultet u Podgorici, Univerzitet Crne Gore, Džordža Vašingtona bb, 81000 Podgorica, Crna Gora; (e-mail: budo@ac.me)

$$Z_c = \frac{60}{\sqrt{\epsilon_{re}}} \ln\left(\frac{8h}{W} + 0.25 \frac{W}{h}\right)$$

Dok ekvivalentnu dielektričnu konstantu dobijamo na sledeći način:

$$\text{za } (W/h) \geq 1$$

$$\epsilon_{re} = \frac{\epsilon_r + 1}{2} + \frac{\epsilon_r - 1}{2} \frac{1}{\sqrt{1 + 12 \frac{h}{W}}}$$

$$\text{za } (W/h) \leq 1$$

$$\epsilon_{re} = \frac{\epsilon_r + 1}{2} + \frac{\epsilon_r - 1}{2} \left( \frac{1}{\sqrt{1 + 12 \frac{h}{W}}} + 0.04 \left(1 - \frac{W}{h}\right)^2 \right)$$

Kada je određena ekvivalentna dielektrična konstanta, talasna dužina mikrotrakastog voda se može prikazati izrazom:

$$\lambda_g = \frac{\lambda_0}{\sqrt{\epsilon_{re}}}$$

$\lambda_0$  je talasna dužina u vakumu, i zavisi od frekvencije  $f$ :

$$\lambda_0 = \frac{c}{f}$$

$c$  je brzina svetlosti u vakumu ( $c \approx 3.0 \times 10^8$  m/s).

U projektovanju mikrotrakastih kola bitnu ulogu zauzima sinteza ( $W/h$ ). Sledeći aproksimativni izrazi za ( $W/h$ ) u zavisnosti od  $Z_c$  i  $\epsilon_r$  daju grešku manju od 1%:

$$\text{za } (W/h) \geq 2$$

$$\frac{W}{h} = \frac{2}{\pi} \left\{ (B - 1) - \ln(2B - 1) + \frac{\epsilon_r - 1}{2\epsilon_r} \left[ \ln(B - 1) + 0.39 - \frac{0.61}{\epsilon_r} \right] \right\}$$

gdje je :

$$B = \frac{60\pi^2}{Z_c \sqrt{\epsilon_r}}$$

$$\text{za } (W/h) \leq 2$$

$$\frac{W}{h} = \frac{8e^A}{e^{2A} - 2}$$

gdje je :

$$A = \frac{Z_c}{60} \left\{ \frac{\epsilon_r + 1}{2} \right\} + \frac{\epsilon_r - 1}{\epsilon_r + 1} \left\{ 0.23 + \frac{0.11}{\epsilon_r} \right\}$$

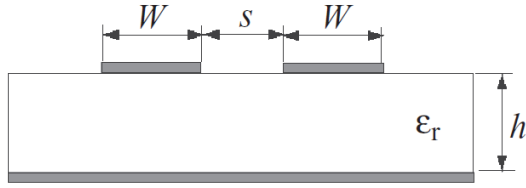
U mikrotalasnoj tehnici veliki značaj igraju spregnuti vodovi. U okviru ovog rada baziraćemo se na spregnute mikrotrakaste vodove koji svoju veliku primjenu nalaze u filtrima.

Kada su dva provodna voda u neposrednoj blizini snaga sistema se povećava zbog interakcije elektromagnetnih polja. Spregnuti vodovi obično imaju dva provodnika, a mogu imati i više. Na Sl. 2 je prikazan spregnuti mikrotrakasti vod širine provodnih traka  $W$  na međusobnom rastojanju  $s$ .

Spregnuti vodovi podržavaju dva različita režima

prostiranja talasa i zbog toga se implementiraju u usmjerenim spojnicama, hibridima i filtrima.

Spregnuti vodovi su simetrični, njihove provodne trake su jednake širine i isto su udaljene od uzemljenja, što pojednostavljuje analizu.

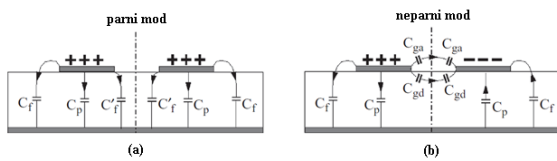


Sl. 2. Spregnuti mikrotrakasti vod.

## II. PROJEKTOVANJE SPREGNUTIH MIKROTRAKASTIH VODOVA

### A. Teorija

Spregnuti mikrotrakasti vod podržava prostiranje dva moda, parnog i neparnog. Struje su kod parnog moda u provodnim trakama istih amplituda i istog smjera, dok su kod neparnog istih amplituda ali suprotnog smjera. Na Sl. 3 su prikazani modovi i kapacitivnosti koje nastaju pri datom modu.



Sl. 3. Kapacitivnosti parnog (a) neparnog (b) moda.

Sa Sl. 3 se mogu označiti kapacitivnosti parnog i neparnog moda, odnosno  $C_e$  i  $C_o$ , kao:

$$\begin{aligned} C_e &= C_p + C_f + C'_f \\ C_o &= C_p + C_f + C_{dg} + C_{ga} \end{aligned}$$

Gdje su:

$C_p$  je kapacitivnost pločastog kondenzatora između provodne trake i uzemljenja.

$C_f$  je ivična kapacitivnost nespregnute provodne trake.

$C'_f$  je kapacitivnost nespregnute provodne trake uz prisustvo druge.

$C_{ga}$  je ivična kapacitivnost duž sprege u vazduhu.

$C_{gd}$  je ivična kapacitivnost duž sprege u dielektriku.

Karakteristične impedanse parnog i neparnog moda  $Z_{ce}$  i  $Z_{co}$  se mogu izraziti iz kapacitivnosti.

$$Z_{ce} = (c\sqrt{C_e^a C_e})^{-1}$$

$$Z_{co} = (c\sqrt{C_o^a C_o})^{-1}$$

gdje su  $C_e^a$  i  $C_o^a$  kapacitivnosti parnog i neparnog moda za spregnute mikrotrakaste vodove u kojima je umjesto dielektrika vazduh.

Efektivna dielektrička konstanta sprege je aritmetička sredina ekvivalentnih dielektričnih konstanta parnog i neparnog moda. Izražena je jednačinom:

$$\sqrt{\varepsilon_{re}} = \frac{cC_e Z_{ce} + cC_o Z_{co}}{2}$$

Izrazi za efektivnu dielektričnu konstantu i karakteristične impedanse spregnutog mikrotrakastog voda mogu se izračunati iz sledećih jednačina:

$$\varepsilon_{re}^e = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2} \left(1 + \frac{10}{v}\right)^{a_e b_e}$$

sa

$$v = \frac{u(20 + g^2)}{10 + g^2} + g e^{-g}$$

$$a_e = 1 + \frac{1}{49} \ln \left[ \frac{v^4 + \left(\frac{v}{52}\right)^2}{v^4 + 0.432} \right] + \frac{1}{18.7} \ln \left[ 1 + \left(\frac{v}{18.1}\right)^3 \right]$$

$$b_e = 0.564 \left( \frac{\varepsilon_r - 0.9}{\varepsilon_r + 3} \right)^{0.053}$$

Gdje je  $u = W/h$  i  $g = s/h$ . Greška u  $\varepsilon_{re}^e$  je 0.7% za opsege  $0.1 \leq u \leq 10$ ,  $0.1 \leq g \leq 10$ , i  $1 \leq \varepsilon_r \leq 18$ .

$$\varepsilon_{re}^o = \varepsilon_{re} + [0.5(\varepsilon_r + 1) - \varepsilon_r + a_o] e^{-c_o g^{d_o}}$$

sa

$$a_o = 0.7287[\varepsilon_{re} - 0.5(\varepsilon_r + 1)][1 - e^{-0.179u}]$$

$$b_o = \frac{0.747\varepsilon_r}{0.15 + \varepsilon_r}$$

$$c_o = b_o - (b_o - 0.207)e^{-0.414u}$$

$$d_o = 0.593 + 0.694e^{-0.562u}$$

Gdje je  $\varepsilon_{re}$  efektivna dielektrična konstanta mikrotrakastog voda debljine  $W$ . Greška u  $\varepsilon_{re}^o$  je reda 0.5%.

Karakteristične impedanse parnog i neparnog moda date su sa preciznošću 0.6% u opsegu  $0.1 \leq u \leq 10$ ,  $0.1 \leq g \leq 10$ , i  $1 \leq \varepsilon_r \leq 18$ :

$$Z_{ce} = Z_c \frac{\sqrt{\varepsilon_{re}/\varepsilon_{re}^e}}{1 - Q_4 \sqrt{\varepsilon_{re}} \cdot Z_c / 377}$$

$Z_c$  je karakteristična impedansa mikrotrakasnog voda debljine  $W$ , i

$$Q_1 = 0.8695u^{0.194}$$

$$Q_2 = 1 + 0.07519g + 0.189g^{2.31}$$

$$Q_3 = 0.1975 + \left[ 16.6 + \left(\frac{8.4}{g}\right)^6 \right]^{-0.387} + \frac{1}{241} \left[ \frac{g^{10}}{1 + \left(\frac{g}{3.4}\right)^{10}} \right]$$

$$Q_4 = \frac{2Q_1}{Q_2} \frac{1}{u^{Q_3} e^{-q} + [2 - e^{-q}]u^{-Q_3}}$$

$$Z_{co} = Z_c \frac{\sqrt{\varepsilon_{re}/\varepsilon_{re}^o}}{1 - Q_{10} \sqrt{\varepsilon_{re}} \cdot Z_c / 377}$$

sa

$$Q_5 = 1.794 + 1.14 \ln \left[ 1 + \frac{0.638}{g + 0.517g^{2.43}} \right]$$

$$Q_6 = 0.2305 + \frac{1}{281.3} \ln \left[ \frac{g^{10}}{1 + \left(\frac{g}{5.8}\right)^{10}} \right] + \frac{1}{5.1} \ln(1 + 0.598g^{1.154})$$

$$Q_7 = \frac{10 + 190g^2}{1 + 82.3g^3}$$

$$Q_8 = e \left[ -6.5 - 0.95 \ln g - \left(\frac{g}{0.15}\right)^5 \right]$$

$$Q_9 = \ln(Q_7) \cdot \left( Q_8 + \frac{1}{16.5} \right)$$

$$Q_{10} = Q_4 - \frac{Q_5}{Q_2} e^{\left[ \frac{Q_6 \ln u}{u^{Q_9}} \right]}$$

### B. Sinteza

Pri projektovanju sinteza ima veći značaj u odnosu na analizu mikrotalasnih kola. Kako odrediti fizičke osobine spregnutog mikrotrakastog voda? U procesu sinteze,  $Z_{ce}$  i  $Z_{co}$  za spegnute mikrotrakasti vod su poznate i neophodno je odrediti  $W/h$  i  $s/h$ . Prvi korak je pronaći za pojedinačne linije odnose  $(W/h)_e$  i  $(W/h)_o$  koji odgovaraju impedansama  $Z_{ce}/2$  i  $Z_{co}/2$ , koristeći jednačinu za karakterističnu impedansu iz uvoda.

Sljedeće jednačine za  $W/h$  i  $s/h$  za spegnute mikrotrakasti vod su dobijene simultanim rješenjem:

$$\left(\frac{W}{h}\right)_e = \frac{2}{\pi} \cosh^{-1} \left( \frac{2p - g + 1}{g + 1} \right)$$

za  $\epsilon_r \leq 6$

$$\left(\frac{W}{h}\right)_o = \frac{2}{\pi} \cosh^{-1} \left( \frac{2p - g - 1}{g + 1} \right) + \frac{4}{\pi \left(1 + \frac{\epsilon_r}{2}\right)} \cosh^{-1} \left( 1 + 2 \frac{W/h}{s/h} \right)$$

za  $\epsilon_r \geq 6$

$$\left(\frac{W}{h}\right)_o = \frac{2}{\pi} \cosh^{-1} \left( \frac{2p - g - 1}{g + 1} \right) + \frac{1}{\pi} \cosh^{-1} \left( 1 + 2 \frac{W/h}{s/h} \right)$$

Gdje je

$$g = \cosh \left[ \frac{1}{2} \pi \frac{s}{h} \right]$$

$$p = \cosh \left[ \pi \frac{W}{h} + \frac{1}{2} \pi \frac{s}{h} \right]$$

Rješavanje ovih jednačina pojednostavljeno je zanemarivanjem drugog člana u izrazu za vrijednosti  $\epsilon_r \leq 6$ . Vrijednost  $s/h$  se onda lako može dobiti iz:

$$\frac{s}{h} = \frac{2}{\pi} \cosh^{-1} \left\{ \frac{\cosh \left[ \frac{1}{2\pi} \left( \frac{W}{h} \right)_e \right] + \cosh \left[ \frac{1}{2\pi} \left( \frac{W}{h} \right)_o \right] - 2}{\cosh \left[ \frac{1}{2\pi} \left( \frac{W}{h} \right)_o \right] - \cosh \left[ \frac{1}{2\pi} \left( \frac{W}{h} \right)_e \right]} \right\}$$

Izrazimo  $W/h$  iz jednačina za parametar  $p$ , a  $p$  iz jednačine za  $(W/h)_e$ . Dobijamo:

$$\frac{W}{h} = \frac{1}{\pi} \cosh^{-1} \left( \frac{g + 1}{2} \cosh \frac{\pi}{2} \left( \frac{W}{h} \right)_e + \frac{g - 1}{2} \right) - \frac{1}{2} \frac{s}{h}$$

### III. TESTIRANJE SINTEZE

#### A. Primjer

Prilikom projektovanja mikrotalnog filtra sa bočnim rezonatorima dobijene su vrijednosti karakterističnih impedansi parnog i neparnog moda,  $Z_{ce}$  i  $Z_{co}$  za spegnute mikrotrakaste vodove.

Odrediti vrijednosti  $(W/h)$  i  $(s/h)$ . Uporediti sa rezultatima dobijenim AWR softverom.

Da bi se izbjegla greška u zaokruživanju, proces sinteze je implementiran u programskom jeziku C#, prikazan na Sl. 4. Za mikrotrakasti filter petog reda centralne učestanosti  $f_0 = 1\text{GHz}$  i propusnog opsega  $\Delta f = 300\text{MHz}$ , debljine substrata  $h = 0.8128\text{mm}$  dielektrične permeabilnosti  $\epsilon_r = 2$  dobijene su karakteristične impedanse:

parnog moda:

$$Z_{ce1} = Z_{ce6} = 82.94 \Omega$$

$$Z_{ce2} = Z_{ce5} = 61.16 \Omega$$

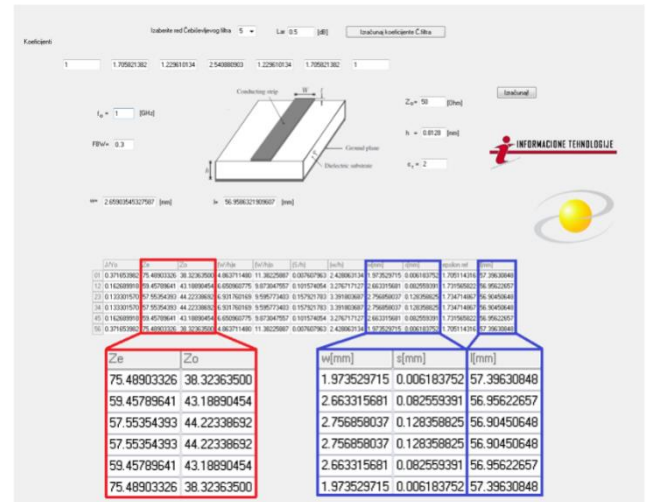
$$Z_{ce3} = Z_{ce4} = 58.18 \Omega$$

neparnog moda:

$$Z_{co1} = Z_{co6} = 37.61 \Omega$$

$$Z_{co2} = Z_{co5} = 42.37 \Omega$$

$$Z_{co3} = Z_{co4} = 43.87 \Omega$$



Sl. 4. Primjena u procesu računanja parametara spregnutih mikrotrakastih vodova koristeći proces opisan u sintezi

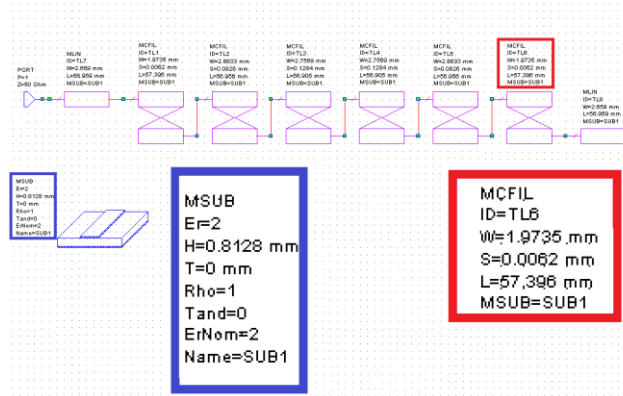
Sa Sl. 4 se vidi da za navedene karakteristične impedanse parnog i neparnog moda dobijamo parametre spregnutih mikrotrakastih vodova označeni plavim pravougaonikom.

Da li dobijeni parametri zadovoljavaju plavne uslove?

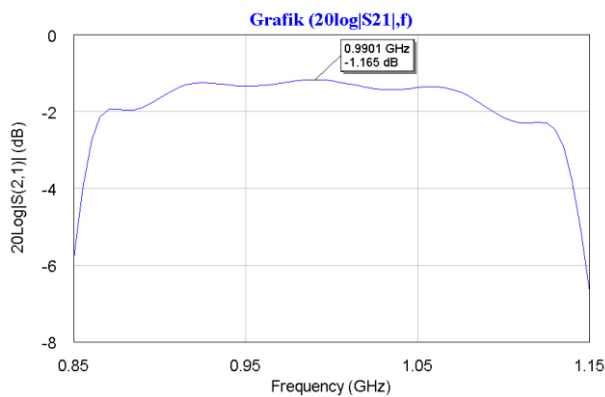
Uporediti sa rezultatima dobijenim u alatu NI AWR Design Environment. Kolika je greška?

### B. NI AWR Desing Enviroment

NI AWR Design Enviroment je softversko rješenje za projektovanje mikrotalasnih kola. Velike kompanije poput Nokia, Siemens, Ericsson, itd. primjenjuju NI AWR svakodnevno u njihovom poslovanju zbog pouzdanosti i praktičnosti prilikom njene eksploatacije.



Sl.5. Projektovano kolo u NI AWR softveru



Sl. 6. Frekvencijski odziv mikrotrakastog filtra

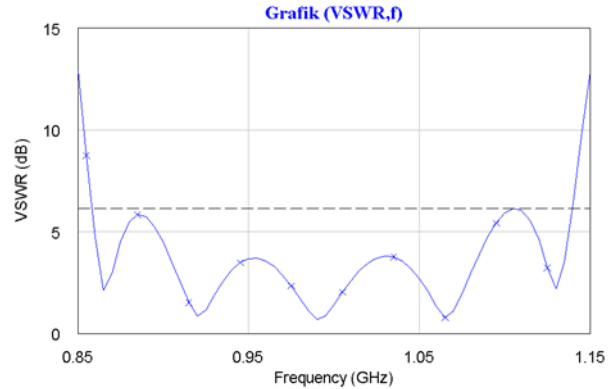
Na Sl. 5 prikazan je projektovan mikrotrakasti filter sa parametrima dobijenim sintezom koji se nalaze na Sl.4. Prepostavili smo idealne uslove da debljina provodne trake  $t \rightarrow 0$  i da nema gubitaka u dielektriku. Frekvencijski odziv našeg kola je prikazan na Sl. 6. Filter unosi slabljenje 1.165 dB. Centralna frekvencija je pomjerena na 0.9901GHz, greška je 1%. Grafik VSWR (voltage standing wave ratio) u zavisnosti od frekvencije je prikazan na Sl. 7. Sa grafika se mogu pročitati granične

vrijednosti propusnog opsega i odrediti širina propusnog opsega:

$$\Delta f = f_2 - f_1 = 1.13GHz - 0.85GHz = 280MHz$$

Greška propusnog opsega je 6.7%.

Metoda opisana u sintezi daje dobre rezultate sa zanemarljivom greškom u odnosu na pojednostavljeno računanje.



Sl. 7. VSWR mikrotrakastog filtra

### LITERATURA

- [1] Sina Akhtarzad, "The Design of Coupled Microstrip Lines", *IEEE Trans. on Microwave Theory and Techniques*, vol. mtt-23, no. 6, June 1975
- [2] Ralph Levy, "Design of Microwave Filters", *IEEE Trans. on Microwave Theory and Techniques*, vol. 50, no. 3, March 2002
- [3] D.M. Pozar, "Microwave engineering" - New York: Wiley, 2012.
- [4] R.E. Collin, "Foundations for Microwave Engineering", Wiley, 1992
- [5] M. Ramadan, W.F. Westgate, "Impedance of coupled microstrip transmission lines", *Microwave J.*, vol. 14, pp. 30-35, July 1971

### ABSTRACT

The subject of the research is coupled microstrip lines which have significant application in bandpass microwave filters. In this paper, we showed that specific methode in synthesis of coupled microstrip lines gives good results in the design of parallel-coupled resonator filters. Verification of the obtained result will be performed in NI AWR Desing Enviroment software.

### DESIGN OF COUPLED MICROSTRIP LINES

Arsenije Maliković, prof. dr Dragan Filipović, prof. dr Budimir Lutovac

# Potencijal, električno polje i podužna kapacitivnost tankog linijskog provodnika u uglu između dvije uzemljene provodne poluravni

Dragan Filipović, Vladan Durković

**Sadržaj** — U radu je dobijena prosta formula za potencijal tankog linijskog provodnika između dvije uzemljene poluravni koje zaklapaju proizvoljan ugao. Na osnovu toga, određena je podužna kapacitivnost provodnika, što je ilustrovano sa više primjera.

**Gljučne reči** — potencijal, električno polje, podužna kapacitivnost, tanki linijski provodnik, uzemljena poluravan.

## I. UVOD

Poznato je [1] da se problem naelektrisanog tankog linijskog provodnika u uglu između dvije uzemljene poluravni može riješiti analitički metodom ogledanja, samo ako ugao ima vrijednost  $\alpha = \pi/n$ ,  $n=1,2,3...$  Međutim, i u tom slučaju metod ima ograničenja, jer je broj likova koji zamjenjuju indukovana naelektrisanja na uzemljenim poluravnima  $2n-1$ , pa su za veće vrijednosti  $n$  izračunavanja potencijala i polja veoma glomazna.

Za proizvoljnu vrijednost ugla između uzemljenih poluravni može se primjeniti moćan metod razdvajanja promjenljivih u Laplasovoj jednačini, koji dovodi do rješenja za potencijal u vidu beskonačnog reda [2], ili integrala na beskonačnom intervalu [3].

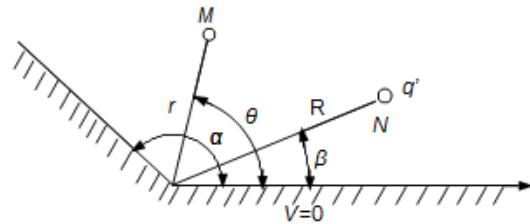
U ovom radu izabran je pristup iz [2], ali je potencijal nađen u zatvorenom obliku, u vidu proste formule. To dalje omogućava da se dobiju proste formule za električno polje i podužnu kapacitivnost linijskog provodnika.

## II. POTENCIJAL TANKOG LINIJSKOG PROVODNIKA U UGLU IZMEĐU DVIJE UZEMLJENE PROVODNE POLURAVNI

Geometrija problema je prikazana na sl. 1. Poluravni koji su na potencijalu  $V=0$  zaklapaju ugao  $\alpha$ , a paralelno njima se proteže tanki linijski provodnik naelektrisan podužnom gustinom naelektrisanja  $q'$ . Njegov položaj definisan je polarnim koordinatama  $R$  i  $\beta$ .

Dragan Filipović, Elektrotehnički fakultet u Podgorici, Univerzitet Crne Gore, Džordža Vašingtona bb, 81000 Podgorica, Crna Gora (e-mail: [draganf@ac.me](mailto:draganf@ac.me)).

Vladan Durković (autor za kontakte), Elektrotehnički fakultet u Podgorici, Univerzitet Crne Gore, Džordža Vašingtona bb, 81000 Podgorica, Crna Gora (e-mail: [vladan.d@ac.me](mailto:vladan.d@ac.me)), Elektrotehnički fakultet u Beogradu, Univerzitet u Beogradu, Bulevar kralja Aleksandra 73, 11120 Beograd, Srbija.



Sl. 1. Geometrija problema.

Potencijal u prostoru između poluravni se određuje iz uslova da zadovoljava Laplasovu jednačinu (osim u tački N, gdje je beskonačan):

$$\Delta V(r, \theta) = \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial V}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 V}{\partial \theta^2} = 0 \quad (1)$$

i granične uslove:

$$V = 0 \text{ za } \theta = 0 \text{ i } \theta = \alpha \quad (2)$$

Metodom razdvajanja promjenljivih može se dobiti rješenje za potencijal u obliku [2]:

$$V(r, \theta) = \frac{q'}{\pi \epsilon_0} \sum_{n=1}^{\infty} \frac{1}{n} \left( \frac{r}{R} \right)^{\frac{n\pi}{\alpha}} \sin \frac{n\pi\beta}{\alpha} \sin \frac{n\pi\theta}{\alpha}, \quad r < R \quad (3)$$

(za  $r > R$  treba na desnoj strani (3) umjesto  $r/R$  staviti  $R/r$ ). Neposredno se provjerava da potencijal (3) zadovoljava (1) i (2).

Beskonačan red na desnoj strani (3) može se prosumurirati u zatvorenom obliku korišćenjem formule [4]:

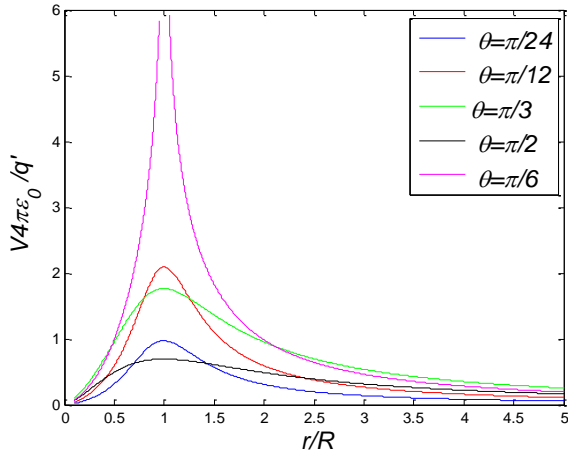
$$\begin{aligned} & \sum_{n=1}^{\infty} \frac{A^n}{n} \sin nx \sin ny = \\ & = \frac{1}{4} \ln \frac{4A \sin^2 \frac{x+y}{2} + (A-1)^2}{4A \sin^2 \frac{x-y}{2} + (A-1)^2} \end{aligned}$$

tako da formula (3) postaje:

$$\begin{aligned} V(r, \theta) = & \frac{q'}{4\pi\epsilon_0} \times \\ & \times \ln \frac{4 \left( \frac{r}{R} \right)^{\frac{\pi}{\alpha}} \sin^2 \frac{\pi}{2\alpha} (\beta + \theta) + \left[ \left( \frac{r}{R} \right)^{\frac{\pi}{\alpha}} - 1 \right]^2}{4 \left( \frac{r}{R} \right)^{\frac{\pi}{\alpha}} \sin^2 \frac{\pi}{2\alpha} (\beta - \theta) + \left[ \left( \frac{r}{R} \right)^{\frac{\pi}{\alpha}} - 1 \right]^2}, \quad r < R \quad (4) \end{aligned}$$



Na sl.2. prikazan je potencijal  $V(r/R) \cdot 4\pi\epsilon_0/q'$ , za  $\alpha=2\pi/3$ ,  $\beta=\pi/6$  i različite vrijednosti ugla  $\theta$ . Krive su dobijene primjenom formule (4).



Sl. 2. Raspodjela potencijala za različite vrijednosti ugla  $\theta$ .

Sa sl.2 se može uočiti da za bilo koju vrijednost ugla  $\theta$  maksimalna vrijednost potencijala se dobija za  $r/R=1$ .

### III. ELEKTRIČNO POLJE

Komponente električnog polja  $\vec{E} = -\text{grad}V$  mogu se dobiti iz (4) korišćenjem formule za gradijent u cilindričnom koordinatnom sistemu [1]:

$$\begin{aligned} E_r(r, \theta) &= -\frac{\partial V}{\partial r} \\ E_\theta(r, \theta) &= -\frac{1}{r} \frac{\partial V}{\partial \theta} \end{aligned} \quad (5)$$

Alternativno, umjesto (4), može se koristiti (3).

### IV. PODUŽNA KAPACITIVNOST PROVODNIKA

Da bi smo odredili podužnu kapacitivnost provodnika moramo uzeti da ima određenu debljinu. Neka je poluprečnik provodnika  $a$  ( $a \ll R$ ). Tada, potencijal provodnika dobijamo iz (4) stavljajući da je  $r=R-a$  i  $\theta=\beta$ :

$$\begin{aligned} V &\equiv V(R-a, \beta) = \frac{q'}{4\pi\epsilon_0} \times \\ &\times \ln \frac{4 \left( \frac{R-a}{R} \right)^{\pi/\alpha} \sin^2 \frac{\pi\beta}{\alpha} + \left[ \left( \frac{R-a}{R} \right)^{\pi/\alpha} - 1 \right]^2}{\left[ \left( \frac{R-a}{R} \right)^{\pi/\alpha} - 1 \right]^2} \end{aligned} \quad (6)$$

odakle je podužna kapacitivnost provodnika:

$$\begin{aligned} C' &= \frac{q'}{V} = \\ &= \frac{4\pi\epsilon_0}{\ln \frac{4 \left( \frac{R-a}{R} \right)^{\pi/\alpha} \sin^2 \frac{\pi\beta}{\alpha} + \left[ \left( \frac{R-a}{R} \right)^{\pi/\alpha} - 1 \right]^2}{\left[ \left( \frac{R-a}{R} \right)^{\pi/\alpha} - 1 \right]^2}} \end{aligned}$$

S obzirom da je  $a \ll R$  možemo pisati:

$$\left( \frac{R-a}{R} \right)^{\pi/\alpha} = \left( 1 - \frac{a}{R} \right)^{\pi/\alpha} \approx 1 - \frac{\pi a}{\alpha R}$$

pa se (6) svodi na:

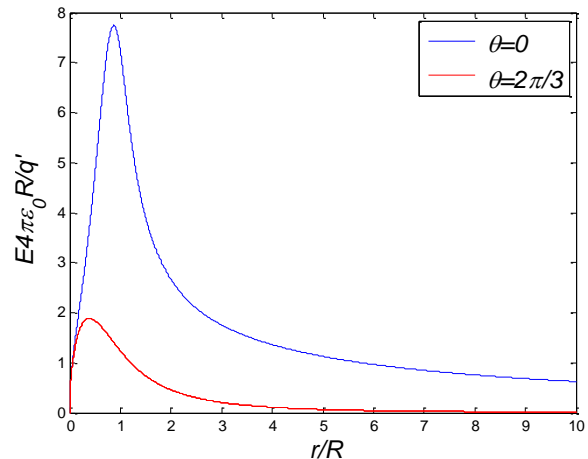
$$C' \approx \frac{4\pi\epsilon_0}{\ln \frac{4 \sin^2 \frac{\pi\beta}{\alpha} + \left( \frac{\pi a}{\alpha R} \right)^2}{\left( \frac{\pi a}{\alpha R} \right)^2}}$$

Konačno, ako ugao  $\beta$  nije previše blizu 0 ili  $\alpha$ , može se drugi član u brojiocu logaritma zanemariti, pa je:

$$C' \approx \frac{2\pi\epsilon_0}{\ln \left( \frac{2\alpha R \sin \frac{\pi\beta}{\alpha}}{\pi a} \right)} \quad (7)$$

### V. PRIMJERI IZRAČUNAVANJA POLJA I PODUŽNE KAPACITIVNOSTI

Na sl. 3 su prikazane jačine električnog polja  $E(r/R) \cdot 4\pi\epsilon_0 R/q'$  na poluravnima  $\theta=0$  i  $\theta=\alpha$ . Krive su dobijene primjenom formula (4) i (5). Uzeto da je:  $\alpha=2\pi/3$ ,  $\beta=\pi/6$ .



Sl. 3. Jačina električnog polja na poluravnima  $\theta=0$  i  $\theta=\alpha$ .

Indukovane površinske gustine naelektrisanja na poluravnima  $\theta=0$  i  $\theta=\alpha$  mogu se dobiti na osnovu vrijednosti za  $E_\theta(\theta=0)$  i  $E_\theta(\theta=\alpha)$  [1]:

$$\sigma_{in}(\theta=0, \alpha) = \pm \epsilon_0 E_\theta(\theta=0, \alpha)$$

gdje se znak plus uzima za  $\theta=0$  a znak minus za  $\theta=\alpha$ , tako da je:

$$q'_{in}(\theta=0, \alpha) = \pm \epsilon_0 \int_0^\infty E_\theta(\theta=0, \alpha) dr$$

Poslije elementarne integracije dobija se:

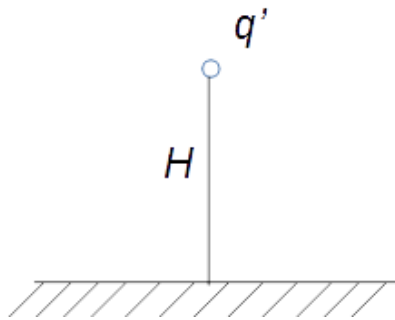
$$q'_{in}(\theta=0) = -q' \left( 1 - \frac{\beta}{\alpha} \right)$$

$$q'_{in}(\theta=\alpha) = -q' \cdot \frac{\beta}{\alpha}$$

Ukupna indukovana podužna količina naelektrisanja na poluravnima  $\theta=0$  i  $\theta=\alpha$  je:

$$q'_{in} = q'_{in}(\theta = 0) + q'_{in}(\theta = \alpha) = -q'$$

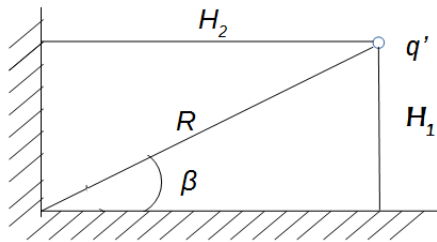
Oredimo primjenom formule (7) podužnu kapacitivnost provodnika za nekoliko specijalnih slučajeva (sl. 4-6) koji se mogu tretirati elementarno pomoću metode ogledanja [1].



Sl. 4.

Za slučaj na sl.4  $\alpha = \pi$ ,  $H = R \sin \beta$ , i iz (7) dobijamo:

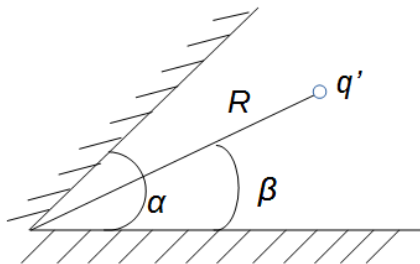
$$C' \approx \frac{2\pi\epsilon_0}{\ln \frac{2H}{a}} \frac{H_2}{H_1}$$



Sl. 5.

Za slučaj sa sl.5  $\alpha = \pi/2$ ,  $H_1 = R \sin \beta$ ,  $H_2 = R \cos \beta$ , iz (7) dobijamo:

$$C' \approx \frac{2\pi\epsilon_0}{\ln \left( \frac{2H_1 \cdot H_2}{a \cdot \sqrt{H_1^2 + H_2^2}} \right)}$$



Sl.6.

Konačno, za slučaj sa sl.6,  $\alpha = \pi/3$ , i npr. za  $\beta = \pi/6$ , dobijamo:

$$C' \approx \frac{2\pi\epsilon_0}{\ln \left( \frac{2R}{3a} \right)}$$

## VI. ZAKLJUČAK

U radu je izvedena prosta formula za potencijal i podužnu kapacitivnost tankog linijskog provodnika između dvije uzemljene poluravnine. Za razliku od metode ogledanja koji se može primijeniti samo za pojedine vrijednosti ugla između poluravnine, opisani metod važi za proizvoljan ugao. Dalja istraživanja obuhvatiće važan praktičan slučaj sistema tankih provodnika (vodova).

## LITERATURA

- [1] B. D. Popović, *Introductory Engineering Electromagnetics*, Addison-Wesley, 1971, ch. 4.
- [2] W. K. H. Panofsky, M. Phillips, *Classical Electricity and Magnetism*, 2<sup>nd</sup> edition, Addison-Wesley, 1962, pp. 77-78.
- [3] W. R. Smythe, *Static and Dynamic Electricity*, 3<sup>rd</sup> edition, McGraw Hill, 1968, pp. 70-72.
- [4] А. П. Прудников, Ю. А. Брычков, О. И. Маричев, *Интегралы и ряды*, Наука, Москва, 1981, стр. 743.

## ABSTRACT

In this paper we have derived a simple formula for the potential of a thin line charged conductor between two ground halfplanes at an arbitrary angle. Next, the conductor capacitance per unit length is determined, followed by illustrative examples.

## POTENTIAL, ELECTRIC FIELD AND CAPACITANCE OF A THIN LINE CONDUCTOR IN AN ANGLE BETWEEN TWO GROUND HALFPLANES

Dragan Filipović, Vladan Durković

# Određivanje širine i divergencije laserskog snopa

Teodora S. Veljović, Boban P. Bondžulić, Ljubiša D. Tomić, Bojan Č. Milanović, Rade M. Pavlović

**Sadržaj** — Ovaj rad sadrži osnovna teorijska razmatranja o prostiranju laserskog snopa kroz slobodan prostor. Kroz eksperimentalni deo istraživanja izvršena je procena minimalne  $1/e$  širine laserskog snopa i divergencije laserskog snopa. U eksperimentalnom delu rada analizirana je i zavisnost greške procene širine snopa od izbora preseka snopa i rastojanja između lasera i prijemnika sa fotodiodom.

**Ključne reči** — divergencija snopa, Gausov snop, laser, prostiranje, širina snopa.

## I. UVOD

OPTIČKI telekomunikacioni sistem za prenos signala u slobodnom prostoru sastoji se od laserskog izvora, sistema za kolimaciju zračenja, odnosno predajne antene, prijemne antene i prijemnika.

Osnovni parametri koji utiču na prenos optičkog signala su: rastojanje između predajnika i prijemnika, digitalni protok sistema i kvalitet prenosa okarakterisan verovatnoćom greške (kod digitalnih sistema) ili odnosom signal/šum (kod analognih sistema).

U cilju što boljeg prenosa signala kroz slobodan prostor, potrebno je najpre izvršiti izbor predajnika i prijemnika. Na prijemu se koriste različite vrste PIN ili lavinskih fotodioda, dok se na predaji obično nalaze laser, ili laserska dioda, čiji izbor treba biti usklađen sa prozorima dobre transmisije u atmosferi [1].

Obično se koristi ugljo-dioksidni laser ili neodijumski laser, čije su talasne dužine  $10.6 \mu\text{m}$  i  $1.06 \mu\text{m}$ , respektivno. Ukoliko se primenjuje laser veće snage, kao na primer ugljo-dioksidni, snaga predajnika može biti znatno veća od vata, pa se takvi predajnici koriste za međusatelitske komunikacije. Laseri oba pomenuta tipa zahtevaju primenu spoljnog modulatora, koji se dodaje u rezonator ili je van njega. Laserski snop se najčeće karakteriše veličinom struka, koja se modifikuje sistemom sočiva radi daljeg sužavanja divergencije snopa, tako da se velika rastojanja mogu savladati relativno malim snagama.

Teodora S. Veljović (autor za kontakte), Vojna akademija, Univerzitet odbrane u Beogradu, Generala Pavla Jurišića Šturma 33, 11000 Beograd, Srbija (e-mail: [veljovic.teodora94@gmail.com](mailto:veljovic.teodora94@gmail.com)).

Boban P. Bondžulić, Vojna akademija, Univerzitet odbrane u Beogradu, Generala Pavla Jurišića Šturma 33, 11000 Beograd, Srbija (e-mail: [bondzulici@yahoo.com](mailto:bondzulici@yahoo.com)).

Ljubiša D. Tomić, Vojnotehnički institut, Ratka Resanovića 1, 11030 Beograd, Srbija (e-mail: [ljubisa.tomic@gmail.com](mailto:ljubisa.tomic@gmail.com)).

Bojan Č. Milanović, Vojna akademija, Univerzitet odbrane u Beogradu, Generala Pavla Jurišića Šturma 33, 11000 Beograd, Srbija (e-mail: [bojan.milanovic@va.mod.gov.rs](mailto:bojan.milanovic@va.mod.gov.rs)).

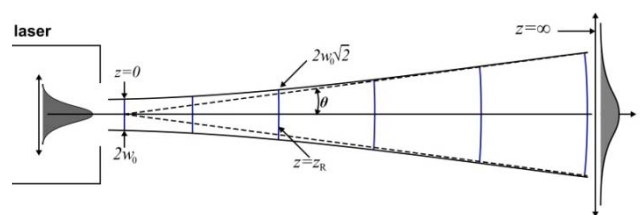
Rade M. Pavlović, Vojnotehnički institut, Ratka Resanovića 1, 11030 Beograd, Srbija (e-mail: [rade\\_pav@yahoo.com](mailto:rade_pav@yahoo.com)).

Optički bežični linkovi se sve više koriste u urbanim sredinama, gde se zahteva velika brzina prenosa, a gde nije opravdano postavljanje optičkih kablova. Prednost optičkog linka bez vlakna u odnosu na radio i mikrotalasne linkove je velika brzina prenosa, reda Gb/s, a nedostaci su ograničen domet i raspoloživost optičkog linka u toku dana, meseca i godine [2].

U prvom delu ovog rada data su teorijska razmatranja o prostiranju laserskog zračenja u slobodnom prostoru. Nakon toga, u eksperimentalnom delu rada analizirano je prostiranje zračenja helijum-neonskog lasera kroz slobodan prostor, tj. mereni su  $1/e$  širina snopa i divergencija snopa, poređeni sa kataloškim vrednostima i određivane greške merenja.

## II. TEORIJSKA RAZMATRANJA

Osnovne karakteristike laserskog zračenja su velika koncentracija energije oko ose snopa i mali prostorni ugao u kome je koncentrisano ovo zračenje. Analiza prostiranja laserskog zračenja u vakuumu svodi se na proračun polja zračenja otvora, koji je osvetljen elektromagnetnim talasom sa osno simetričnom raspodelom amplituda električnog polja po Gausovom zakonu [3], [4]. Zračenje iz ovakvih lasera skraćeno se naziva Gausov snop [5]. Prostiranje Gausovog snopa prikazano je na Sl. 1, gde je sa  $w_0$  označena  $1/e$  širina snopa, dok je sa  $\theta$  označena divergencija snopa.



Sl. 1. Gausov snop.

Širina snopa  $w_z$  na rastojanju  $z$  od mesta minimalne širine snopa  $w_0$  data je izrazom [5], [6]:

$$w_z = w_0 \left( 1 + \left( \frac{\lambda z}{\pi w_0^2} \right)^2 \right)^{1/2} \quad (1)$$

gde je  $\lambda$  talasna dužina.

Kako je na velikim rastojanjima od mesta minimalne širine snopa  $\lambda z / (\pi w_0^2) \gg 1$ , približna vrednost za širinu snopa je:

$$w_z \approx \frac{\lambda z}{\pi w_0} \quad (2)$$

što znači da se širina snopa  $w_z$  linearno menja sa  $z$ .

U dalekoj zoni definiše se ugao Gausovog snopa  $\theta$  (ugao širenja snopa ili divergencija snopa):

$$\operatorname{tg} \theta \approx \theta = \frac{w_z}{z} = \frac{\lambda}{\pi w_0} \quad (3)$$

Kod tipičnog helijum-neonskog (He-Ne) lasera minimalna širina snopa je oko 0.5 mm, tako da je ugao širenja snopa reda jednog minuta. Širina  $w_z$  na 1 km udaljenosti od mesta minimalne širine  $w_0$  za ovaj laser je oko 40 cm [6].

Za razliku od prostiranja svetlosti, Gausov snop ne divergira linearno. Blizu izlaza lasera ugao divergencije je veoma mali, dok se ugao divergencije približava asimptotskom limitu (3) na velikim rastojanjima od izvora laserskog zračenja. Rejljevo rastojanje,  $z_R$ , se definiše kao daljina na kojoj se minimalna širina snopa uveća  $\sqrt{2}$  puta, i ono iznosi:

$$z_R = \frac{\pi w_0^2}{\lambda} \quad (4)$$

Taladni front Gausovog snopa na mestu minimalne širine ( $z=0$ ) je ravan a na nekom rastojanju  $z$  je zakrivljen. Poluprečnik krivine talasnog fronta je [5], [6]:

$$R(z) = z \left( 1 + \left( \frac{\pi w_0^2}{\lambda z} \right)^2 \right) \quad (5)$$

Za dovoljno veliko rastojanje  $z$ , poluprečnik krivine Gausovog snopa postaje jednak  $z$ . Takođe, na  $z=\infty$ , taladni front je ravan ( $R(\infty)=\infty$ ). Sa prostiranjem od mesta minimalne širine, zakrivljenost talasnog fronta raste do maksimalne vrednosti, a nakon toga počinje da opada. Rejljevo rastojanje koje se smatra granicom između divergencije na bliskim i srednjim rastojanjima, je daljina na kojoj taladni front dostiže maksimalnu zakrivljenost. Divergencija snopa koja se navodi u specifikacijama lasera odnosi se na divergenciju na velikim rastojanjima i mora se meriti na daljinama koje su mnogo veće od  $z_R$  (obično se uzima da je  $z > 10 z_R$ ). Ovo je jako bitna napomena jer je određivanje veličine laserskog spota i drugih parametara u jednom optičkom lancu netačno na bliskim i srednjim rastojanjima. Za lasere, daljina daleke zone je reda nekoliko metara.

Preuređivanjem jednačine (1) može se pisati:

$$w_0^4 - w_z^2 w_0^2 + \left( \frac{\lambda z}{\pi} \right)^2 = 0 \quad (6)$$

Ako se ova kvadratna jednačina reši po  $w_0^2$  dobija se sledeći izraz za širinu snopa:

$$w_0 = \sqrt{\frac{w_z^2 \pm \sqrt{w_z^4 - 4K}}{2}} \quad (7)$$

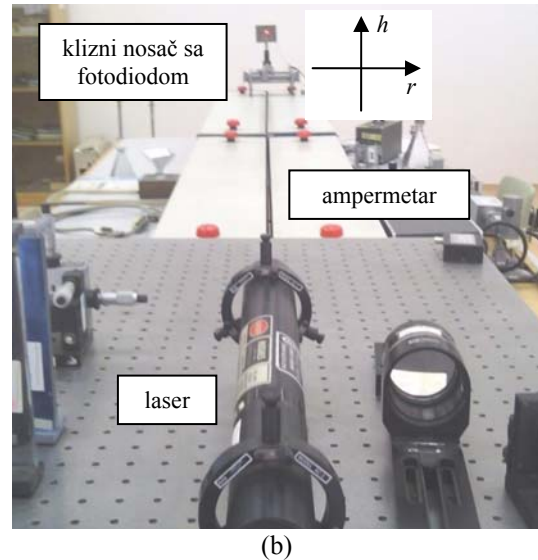
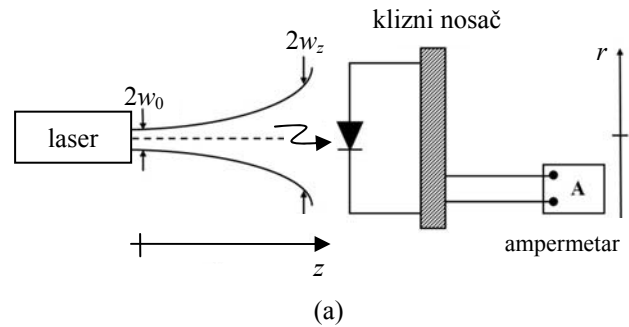
gde je  $K=(\lambda z/\pi)^2$ . Znak ispred korena bira se tako da  $w_0$  bude manje od  $w_z$ . Dakle, ako se zna talasna dužina i meri širina snopa na rastojanju  $z$ , može se odrediti širina snopa u struku (minimalna širina snopa).

### III. ODREĐIVANJE ŠIRINE I DIVERGENCIJE LASERSKOG SNOPIA

U eksperimentu se mere vrednosti struje u funkciji rastojanja od ose snopa. Eksperiment se izvodi korišćenjem:

- He-Ne lasera kao predajnika (talasne dužine  $\lambda=632.8$  nm),
- PIN fotodiode kao prijemnika (osetljivosti 0.4 A/W na  $\lambda=632.8$  nm i aktivne površine 5 mm<sup>2</sup>) i
- ampermetra.

Blok šema sistema za merenje prikazana je na Sl. 2(a), dok je merni sistem prikazan na Sl. 2(b). Merenja su sprovedena u Laboratoriji za optoelektroniku Vojne akademije Univerziteta odbrane u Beogradu, a prema eksperimentalnoj postavci opisanoj u [7].



Sl. 2. (a) blok šema mernog sistema i (b) eksperimentalna postavka mernog sistema.

Eksperiment se realizuje postavljanjem kliznog nosača sa fotodiodom na rastojanje  $z=9$  m od lasera. Pomeranjem fotodiode za 5 mm u oba smera po horizontalnoj osi, pri čemu je korak 0.5 mm, i pomeranjem fotodiode u oba smera po vertikalnoj osi, pri čemu je korak 1 mm, očitavaju se vrednosti struje pomoću kojih je moguće odrediti 1/e širinu snopa. Širina snopa određuje se za svaki horizontalni profil (presek), tj. za različite vertikalne pozicije fotodiode, nakon čega se mogu uporediti i greške pri merenju 1/e širine snopa. Za obradu rezultata koristi se programski paket Matlab®.

Kako je struja fotodiode proporcionalna srednjoj snazi optičkog signala koji je pobuđuje, merenjem struje koju daje fotodiode meri se u stvari srednja optička snaga. Apsolutna kalibracija snage nije potrebna jer je dovoljno da se snimi relativna promena optičke snage, koja je prema jednačini polja laserskog zračenja [8]–[10]:

$$P(r) = P_{\max} \cdot e^{-2\left(\frac{r}{w_z}\right)^2} = k \cdot I(r) \quad (8)$$

gde su:

- $P(r)$  – srednja optička snaga,
- $P_{\max}$  – maksimalna vrednost snage (na osi snopa),
- $I(r)$  – struja fotodiode,
- $k$  – konstanta proporcionalnosti i
- $r$  – rastojanje od ose snopa.

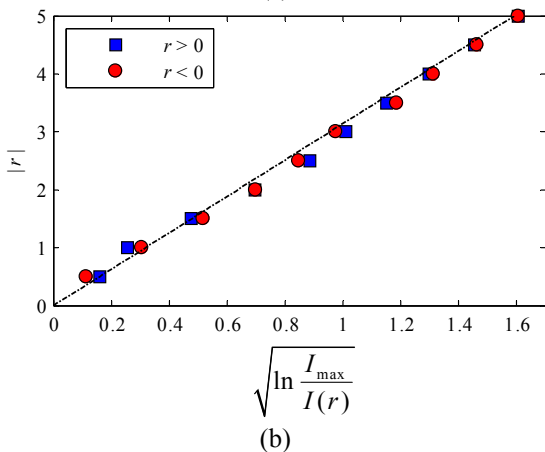
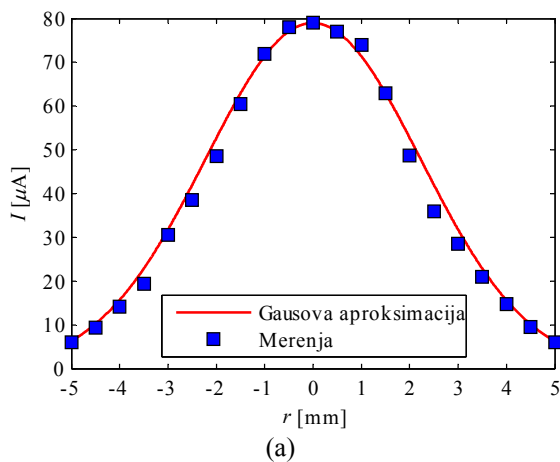
Za merenje 1/e širine laserskog snopa  $w_z$  na rastojanju  $z$  prema jednačini (8) bi bilo dovoljno odrediti  $P_{\max}$  i rastojanje od ose  $r$  za koje je  $P(r)=P_{\max}e^{-2}$ , jer onda važi da je  $r=w_z$ . Tačnost sa kojom se određuje  $w_z$  na ovaj način bi bila mala, pa je bolje iskoristiti više vrednosti  $P(r)$ .

Sa obzirom na logaritamski način promene  $P(r)$  pogodno je preurediti relaciju (8) na sledeći način:

$$r = \frac{w_z}{\sqrt{2}} \left( \ln \frac{P_{\max}}{P(r)} \right)^{1/2} = C_1 \left( \ln \frac{I_{\max}}{I(r)} \right)^{1/2} \quad (9)$$

gde je  $P_{\max}=P(0)$ , odnosno  $I_{\max}=I(0)$ . Ako se nacrtaja dijagram zavisnosti  $r$  u funkciji od  $(\ln I_{\max}/I(r))^{1/2}$  sve tačke treba da leže na pravoj liniji koja prolazi kroz koordinatni početak. Nagib ove prave daje  $w_z/\sqrt{2}$ . Širina struka snopa se nalazi iz poznatog  $z$  i izmerenog  $w_z$  na osnovu (7).

Na Sl. 3 ilustrovan je postupak određivanja 1/e širine snopa  $w_z$ .



Sl. 3. (a) rezultati merenja i njihova aproksimacija i (b) preslikavanje rezultata merenja u linearni zakon promene.

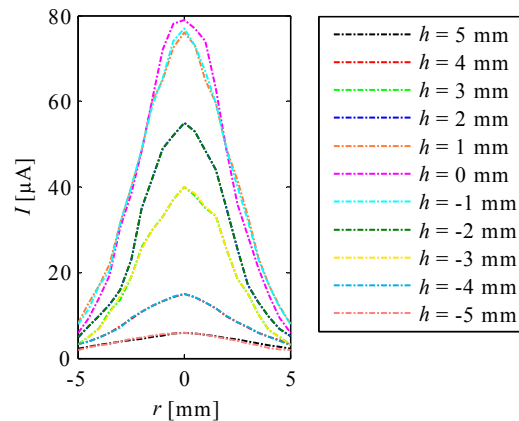
Na Sl. 3(a) prikazane su vrednosti struje za različita horizontalna rastojanja prijemnika od centra ose snopa ( $h=r=0$  mm). Sl. 3(b) ilustruje preslikavanje dobijenih vrednosti struje iz logaritamskog u linearni zakon promene. Iz nagiba dobijene prave određuje se 1/e širina

snopa, koja u datom primeru iznosi  $w_z=4.4465$  mm. Korišćenjem dobijene vrednosti  $w_z$  rezultati merenja se mogu aproksimirati po Gausovom zakonu – Sl. 3(a).

Na osnovu 1/e širine snopa na rastojanju  $z$ ,  $w_z$ , korišćenjem (7) dobija se procena 1/e širine struka snopa lasera  $w_0$ , koja u datom slučaju iznosi  $w_0=0.4094$  mm. Kako je kataloška vrednost 1/e širine 0.415 mm, relativna greška procene širine struka snopa lasera iznosi 1.34 %.

Korišćenjem rastojanja  $z$  i širine snopa  $w_z$ , prema relaciji (3) dobija se da je divergencija snopa 0.494 mrad. Kataloška vrednost divergencije snopa korišćenog lasera je 0.5 mrad.

Postupak određivanja 1/e širine snopa ponovljen je i za druge preseke, tj. za različita vertikalna rastojanja prijemnika od centra ose snopa (ukupno 11 preseka). Vrednosti struje za svaki od preseka prikazane su na Sl. 4. Sa Sl. 4 se vidi da je maksimalna vrednost struje u centru ose snopa (oko 80 µA). Takođe, uočava se da sa udaljavanjem preseka od centra ose snopa vrednosti struje opadaju. Dodatno se uočava sličan trend vrednosti struja preseka na jednakim rastojanjima od centra ose snopa, tj. za jednake pomake po horizontalnoj osi na gornju i na donju stranu. Vrednosti struja za svaki od preseka mogu se dobro aproksimirati Gausovom krivom.



Sl. 4. Vrednosti struje za različite preseke.

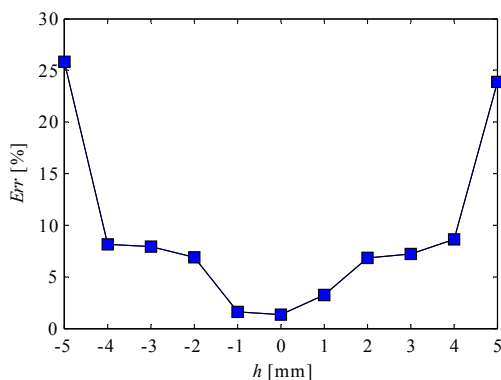
Za svaki od preseka određena je širina snopa, nakon čega je korišćenjem jednačine (7) izvršena procena 1/e širine struka snopa lasera  $w_0$ . Procene minimalne širine snopa i relativne greške procene za svaki od preseka date su u Tabeli 1.

TABELA 1: PROCENA ŠIRINE SNOPA ZA RAZLIČITE PRESEKE.

$h$ [mm]	-1	-2	-3	-4	-5
$w_0$ [mm]	0.4217	0.4435	0.4479	0.3811	0.3077
Err [%]	1.6	6.87	7.92	8.2	25.8

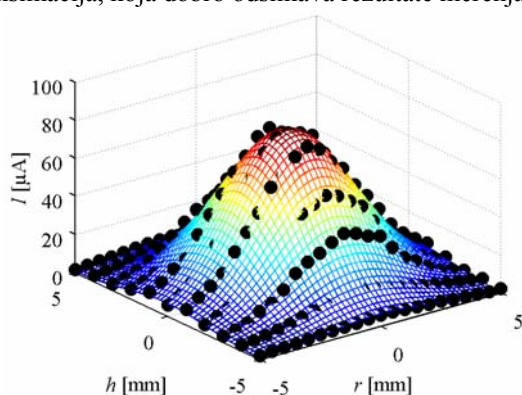
$h$ [mm]	1	2	3	4	5
$w_0$ [mm]	0.4285	0.4434	0.4450	0.3791	0.3150
Err [%]	3.26	6.85	7.23	8.65	23.9

Iz Tabele 1 se uočava da procena širine snopa značajno zavisi od korišćenog preseka, a što utiče na grešku procene. Vrednost greške procene dodatno je prikazana na Sl. 5, gde se uočava da relativna greška raste sa povećanjem rastojanja od centra ose snopa.



Sl. 5. Relativne greške određivanja širine laserskog snopa.

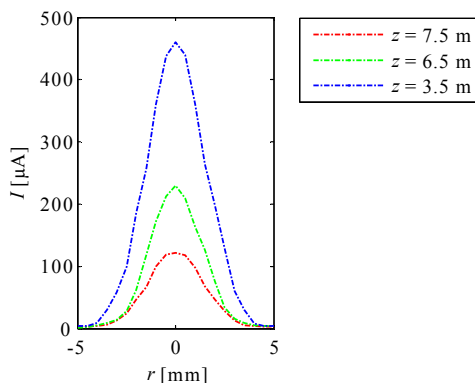
Zavisnost dobijenih vrednosti struje od horizontalnih i vertikalnih rastojanja od centra ose snopa ilustrovane su kroz trodimenzionalni prikaz na Sl. 6. Pored dobijenih vrednosti struje, data je i prostorna Gausova aproksimacija, koja dobro odslikava rezultate merenja.



Sl. 6. Prostorna Gausova aproksimacija rezultata merenja.

#### IV. UTICAJ RASTOJANJA PREDAJNIKA I PRIJEMNIKA NA PROCENU ŠIRINE SNOPA

Postupak određivanja  $1/e$  širine snopa ponovljen je za različita rastojanja između predajnika i prijemnika: 7.5 m, 6.5 m i 3.5 m. Za sva rastojanja prvo je izvršeno određivanje centra ose snopa (maksimalna vrednost struje), nakon čega su vršena pomeranja prijemnika po horizontalnoj osi za 5 mm u obe strane, sa korakom od 0.5 mm. Vrednosti struje merene su za svaki korak, i one su prikazane na Sl. 7.



Sl. 7. Vrednosti struje za različita rastojanja predajnika i prijemnika.

Sa Sl. 7 se uočava da se rezultati merenja mogu dobro aproksimirati Gausovim krivama. Takođe, uočava se da vrednosti struje rastu sa približavanjem prijemnika izvoru

zračenja. Na osnovu rezultata merenja izvršene su procene širine snopa i one su date u Tabeli 2.

TABELA 2: PROCENA ŠIRINE SNOPA ZA RAZLIČITA RASTOJANJA PREDAJNIKA I PRIJEMNIKA.

$z$ [m]	9	7.5	6.5	3.5
$w_0$ [mm]	0.4094	0.4769	0.4860	0.2359
Err [%]	1.34	14.9	17.11	43.15

Iz Tabele 2 se uočava da greška procene značajno raste sa približavanjem prijemnika izvoru zračenja. Jedan od uzroka ovakvih rezultata je što se merenja ne sprovedu u zoni dalekog polja. Naime, Rejljevo rastojanje za korišćeni izvor zračenja iznosi  $z_R=0.86$  m, a merenja je potrebno izvoditi na rastojanjima većim od  $10 \cdot z_R=8.6$  m.

#### V. ZAKLJUČAK

U radu su data teorijska razmatranja i eksperimentalni rezultati pri prostiranju laserskog snopa u slobodnom prostoru. Pokazano je da laserski snop ima oblik Gausove krive. Takođe, pokazano je da prilikom određivanja širine snopa treba voditi računa o rastojanju između predajnika i prijemnika, tj. merenja je potrebno sprovoditi u zoni dalekog polja. Za pouzdano određivanje širine snopa dodatno je potrebno odrediti centar ose snopa. Korišćenjem preseka snopa koji su dalji od centra ose snopa relativna greška određivanja širine snopa raste.

#### LITERATURA

- [1] Grupa autora, *Studija: Elektronski rat – Stanje i perspektive*. Univerzitet u Beogradu, Elektrotehnički fakultet, Beograd, 2000.
- [2] Ž. Barbarić, "Ograničenja u primeni optičkog linka u otvorenom prostoru", 15. *Telekomunikacioni forum – TELFOR*, Srbija, Beograd, 2007, Zbornik radova, str. 393-396.
- [3] O. Bouchet, et al., *Free-space optics – Propagation and communication*. ISTE Ltd, London, UK, 2006.
- [4] B.E.A. Saleh and M.C. Teich, *Fundamentals of photonics*. John Wiley & Sons, Inc., Hoboken, New Jersey, USA, 2007.
- [5] A. Marinčić i Ž. Barbarić, *Optoelektronika i optoelektronski sistemi*. Centar vojnih škola Vojске Jugoslavije, Vojnotehnička akademija, Beograd, Srbija, 1996.
- [6] A. Marinčić, *Laseri i optoelektronika*. NIRO Tehnička knjiga, Beograd, Srbija, 1987.
- [7] A. Marinčić i Ž. Barbarić, *Laboratorijske vežbe iz optoelektronike*. Vojnotehnička akademija, Beograd, Srbija, 1991.
- [8] A. Marinčić, *Osnovi optoelektronskih komunikacija*. IRO Građevinska knjiga, Beograd, Srbija, 1986.
- [9] M. Novakov, "Analiza optičkog prijemnika i merenje osnovnih karakteristika," Diplomski rad, Elektrotehnički fakultet, Univerzitet u Beogradu, 1991.
- [10] M. Tasić, "Realizacija laboratorijskih vežbi iz optoelektronike," Završni rad, Vojna akademija, Univerzitet odbrane u Beogradu, 2013.

#### ABSTRACT

This paper presents basic theoretical considerations of free space laser beam propagation. Through the experimental part of the research carried out the assessment of the minimum width of the laser beam and the divergence of the laser beam. In the experimental part of the work was analyzed and the dependence of error of assessment width of the beam from the selected beam profile and the distance between the laser and the receiver with a photodiode.

#### LASER BEAM WIDTH AND DIVERGENCE MEASURING

Teodora Veljović, Boban Bondžulić, Ljubiša Tomić, Bojan Milanović, Rade Pavlović

# CPW napajana mikrotrakasta monopol antena zasnovana na modifikovanom Sierpinski fraktalu

Luka Lazović, Vesna Rubežić, i Ana Jovanović

**Sadržaj** — U ovom radu je prikazana CPW (Co-Planar Waveguide) napajana fraktalna monopol antena zasnovana na modifikaciji Sierpinski trougaonog fraktala. Korišćenje fraktalnih geometrija za dizajn antena omogućava projektovanje električno malih antena koje zrače na više rezonantnih učestanosti. U cilju ispitivanja performansi i potvrđivanja teorijskih razmatranja izvršene su simulacije raspodjele struje po površini antene, parametara rasijanja i dijagrama zračenja u softverskom paketu AWR. Eksperimentalnim mjerenjima parametara rasijanja potvrđena su teorijska razmatranja i rezultati simulacija.

Pokazano je da se upotrebom fraktalnih geometrija mogu projektovati male antene koje zrače iste dijagrame zračenja na više različitih frekvencija uz očuvanje željenih karakteristika.

**Ključne riječi** — Antena, CPW napajanje, fraktali

## I. UVOD

UBRZANI razvoj bežičnih komunikacionih sistema umnogome zavisi od karakteristika i dizajna samih antena. Komunikacioni sistemi narednih generacija koriste tehnologije kao što su MIMO (Multiple Input Multiple Output) zasnovane na upotrebi antenskih nizova. Jedan mobilni uređaj treba da sadrži antene za 4G, Wi-Fi, GPS i Bluetooth. Zbog toga je pravac razvoja antena je usmjeren ka dizajnu jedne antene koja će zračiti na sve četiri frekvencije, a koja će biti kompaktna i kao takva biti sastavni dio antenskog niza za MIMO sistem i koja se može lako implementirati na elektronskoj ploči mobilnog uređaja.

Antene štampane na elektronskoj ploči je veoma lako izraditi, imaju planarnu geometriju i ne zauzimaju veliki prostor [1]. Mikrotrakaste antene se sastoje od dvije međusobno paralelne provodne površine između kojih se nalazi dielektrik. CPW (Coplanar waveguide) napajanje pojednostavljuje izradu antene jer su trake za napajanje, antena i uzemljena površ na istoj strani dielektrika [2].

Luka Lazović, Elektrotehnički fakultet u Podgorici, Univerzitet Crne Gore, Džordža Vašingtona bb, 81000 Podgorica, Crna Gora (e-mail: lukal@ac.me).

Vesna Rubežić, Elektrotehnički fakultet u Podgorici, Univerzitet Crne Gore, Džordža Vašingtona bb, 81000 Podgorica, Crna Gora (e-mail: vesnar@ac.me).

Ana Jovanović, Elektrotehnički fakultet u Podgorici, Univerzitet Crne Gore, Džordža Vašingtona bb, 81000 Podgorica, Crna Gora (e-mail: anaj@ac.me).

Ovakva unipolarna struktura dodatno pojednostavljuje izradu i jednostavnost implementacije. Međutim, najveći nedostatak ovih antena je uskopojasnost.

Jedan od načina za rješavanje problema uskopojasnosti jeste korišćenje složenih fraktalnih geometrija za izradu antena. Naime, fraktali se mogu definisati kao geometrijski oblici čija je Hausdorff-ova dimenzija veća od topološke dimenzije ili da su fraktalni oblici sastavljeni od oblika koji su na neki način slični cjelokupnom geometrijskom obliku [3]. Kompleksne fraktalne strukture nastaju replikacijom osnovnog oblika.

Još u vrijeme Hercovog istraživanja u oblasti antena, dizajn malih električnih antena je bio centar interesovanja. Konkretno, prvi antenski element koristeći koncept fraktala razvio je Cohen 1997. godine [4]. On je pokazao da ovaj koncept može biti iskorišćen za značajno smanjenje veličine antene bez narušavanja performansi tj. da su fraktalne antene predstavljaju antene malih električnih dužina [5], [6].

Fraktalne antene rezultirale su spajanjem dvije discipline, elektromagnetne teorije i geometrije. Počevši od ranih log periodičnih i spiralnih antena razvijenih šezdesetih godina od strane Carrel Mayes-a, preko radova Benoit Mandelbrot-a [3] vezanih za fraktalne geometrije, fraktalne antene su prirodni slijed događaja u redukciji dimenzija antena i njihovim višefrekvencijskim radom. Isto tako, geometrija fraktala je intenzivnije proučavana sedamdesetih godina prošlog vijeka. Devedesetih godina je uočeno da se ovaj geometrijski oblik može koristiti u dizajnu frekvencijski selektivnih površina i širokopojasnih antena jednako dobro kao i novoj konfiguraciji antenskih nizova [7].

Neka istraživanja od strane Puentea 1998. godine [8] pokazala su da fraktalne antene imaju odlične širokopojasne karakteristike. Pregled ranih početaka rada ovih antena sumirao je Werner 1999. godine [6]. Fraktalne antene se pokazuju veoma korisnim i na velikim frekvencijama u Terahertz sensing aplikacijama [9].

Jednu od najpoznatijih fraktalnih struktura razvio je 1916. godine matematičar Waclaw Sierpinski [3]. Puente 1998. godine [8] pokazuje višekanalne karakteristike fraktala proučavajući ponašanje Sierpinski monopol antene. Sierpinski monopol je električno mala antena koja pokazuje jednostavno ponašanje na nekoliko rezonantnih učestanosti zadržavajući isti dijagram zračenja.

U radu je predložena mikrotrakasta fraktalna monopol

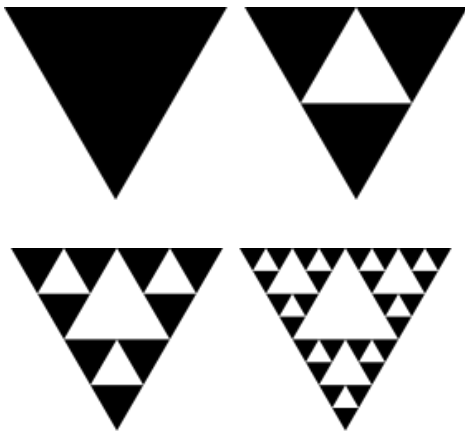
antena čija se geometrija zasniva na modifikaciji Sierpinski trougla. Ova antena se, za razliku od prethodnih verzija, napaja pomoću CPW linije za napajanje.

## II. GEOMETRIJA PREDLOŽENE ANTENE

Fraktalna antena se zasniva na korišćenju samo-sličnih oblika u cilju maksimizacije električne dužine antene. Korišćenje ovih samo-sličnih oblika, tj. fraktala, omogućava dizajn antena koje rade na više frekvencija istovremeno. Ove antene imaju prednost da zrače slične dijagrame zračenja u više frekventnih opsega bez degradacije karakteristika zračenja.

Najviše proučavana antena iz ove grupe je Sierpinski fraktalna antena [8]. U idealnom slučaju, Sierpinski dipol antena može imati beskonačno mnogo iteracija i može zračiti u beskonačno frekventnih opsega. U realnom slučaju, usled ograničenja u tehnologiji izrade, ove antene imaju najviše 6 iteracija [10]. Naime, broj fraktalnih iteracija antene određuje broj rezonantnih frekvencija [11]. Povećanje broja rezonantnih frekvencija se može postići i korišćenjem rekonfigurabilnih antena [12].

Sierpinski trougao je osnovna fraktalna geometrija korišćena u ovom radu. Konstrukcija fraktala počinje sa jednakostraničnim trouglom. Zatim se odrede tačke polovljenja stranica te se od početnog trougla oduzme trougao koji nastaje spajanjem tačaka polovljenja. Na taj način ostaju tri jednakostranična trougla dvostruko manjih dužina stranica od početnog trougla. Sa svakim novim trouglom ponavlja se postupak. Princip formiranja Sierpinski trougla je prikazan na Sl. 1.



Sl. 1. Formiranje Sierpinski fraktala

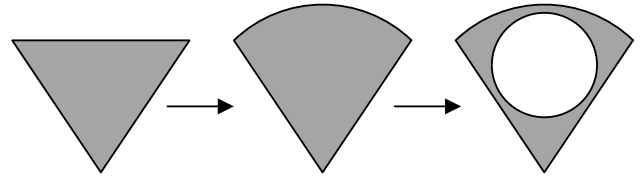
Fraktalna dimenzija Sierpinski trougla je:

$$d = \frac{\log 3}{\log 2} = 1.585 \quad (1)$$

U ovom radu se predlaže modifikovana Sierpinski monopol antena. Predložena antena koristi modifikovanu Sierpinski strukturu kao zračeci element koji se napaja sa CPW (Co-Planar Waveguide) [5] linijom za napajanje za

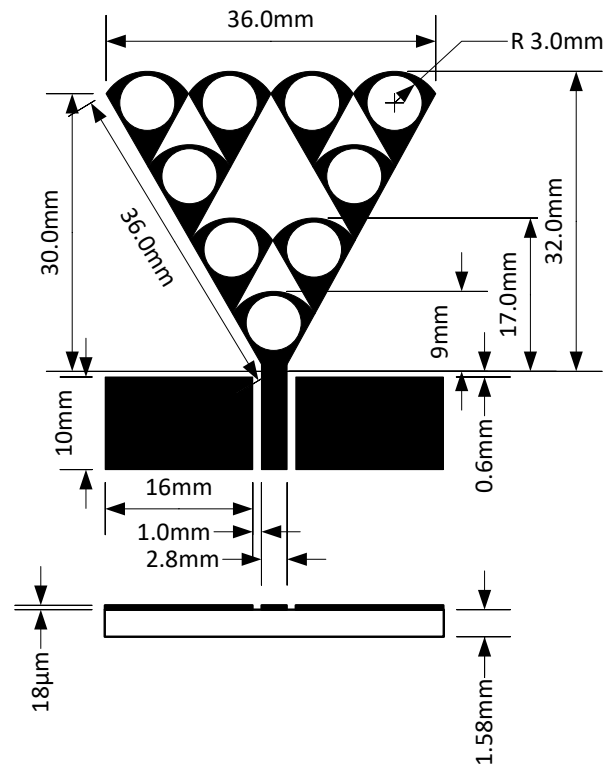
razliku od prethodno predloženih konvencionalnih patch antena. Ovaj metod napajanja je izabran zbog boljeg prilagođenja impedanse voda sa impedansom antene, manjih gubitaka zračenja i jednostavnije izrade.

Ideja za ovu geometriju antene proističe iz rada [5] u kojem se koristi niz od dvije patch fraktalne antene zasnovane na modifikaciji Sierpinskog trougla. Modifikacija Sierpinski trougla prikazana je na Sl. 2.



Sl. 2. Modifikacija fraktalne geometrije Sierpinski trougla [5]

Zračeci fraktalni element se napaja sa mikrotrakastim vodom širine 2.8mm i prilagođena je za napajanje napojnim vodom impedanse  $50\Omega$ . Antena je izrađena na štampanoj ploči od vitroplasta debljine 1.58mm, relativne dielektrične konstante  $\epsilon_r=4.4$ . Bakarna metalizacija je debljine 18  $\mu\text{m}$ . Dimenzije i oblik predložene geometrije prikazane su na Sl. 3. Izrađena antena je prikazana na Sl. 4.



Sl. 3. Predložena monopol antena zasnovana na modifikovanom Sierpinski fraktalu



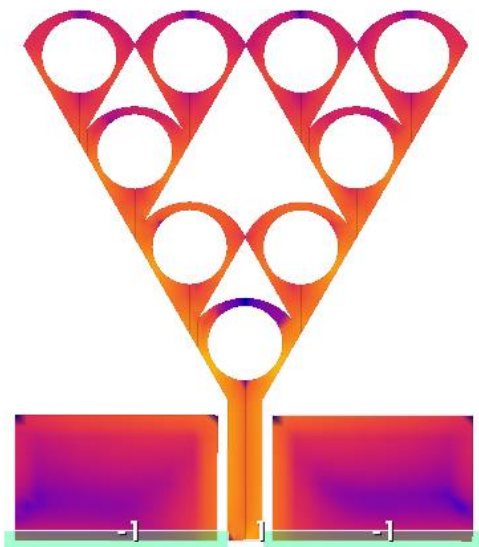


Sl. 4. Izgled izrađene antene

### III. REZULTATI SIMULACIJA I MJERENJA

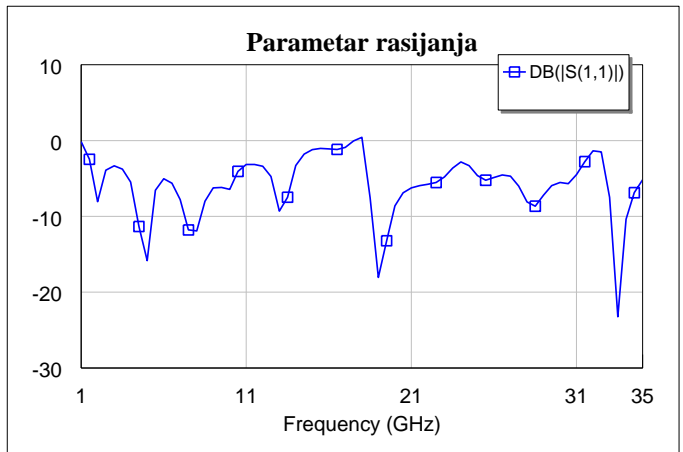
U cilju analize performansi antene izvršeno je modelovanje antene, uz brojne simulacije korišćenjem softverskog paketa AWR. Simulirana je raspodjela struje, parametri rasijanja i dijagram zračenja, dok su eksperimentalno potvrđeni parametri rasijanja.

Raspodjelu struje po površini jedino je moguće odrediti numeričkim metodama. U ovom slučaju, raspodjela struje na anteni određena je numeričkom analizom zasnovanom na metodu momenata, na rezonantnoj učestanosti antene 1.8 GHz. Na Sl. 5. su prikazani simulacioni rezultati raspodjele struje po površini antene.



Sl. 5. Raspodjela struje po površini predložene antene

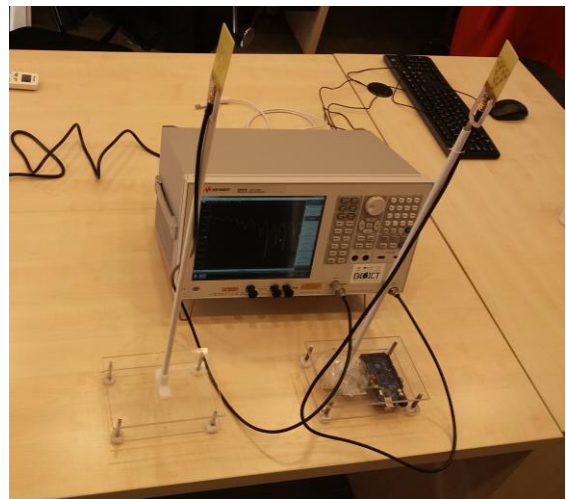
Rezultati simulacije parametara rasijanja predložene antena su prikazani na Sl. 6.



Sl. 6. Parametri rasijanja.

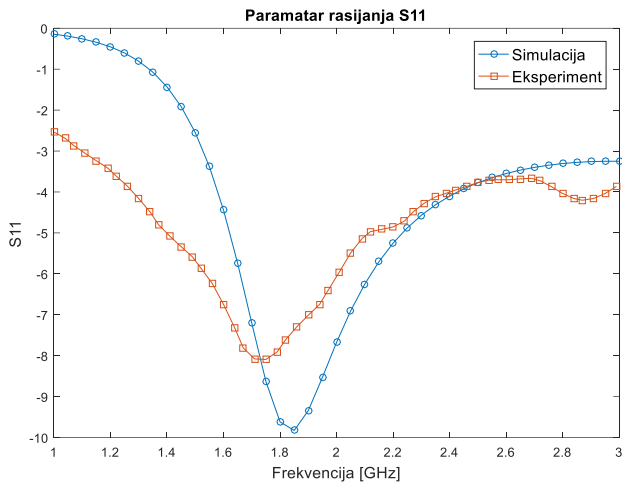
Na osnovu rezultata simulacija parametara rasijanja  $S_{11}$  može se vidjeti da antena ima više rezonantnih učestanosti što je potvrdilo teorijska razmatranja i opravdalo razlog korišćenja fraktalnih geometrija.

U cilju eksperimentalnog određivanja parametara rasijanja realizovane antene korišćen je Network Analyzer Keysight E5061B-3L5. Eksperimentalna postavka za mjerenje  $S_{11}$  parametra u opsegu od 1GHz do 3GHz je prikazana na Sl. 7.



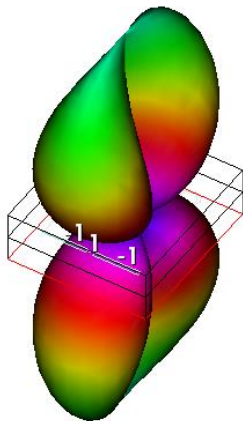
Sl. 7. Eksperimentalna postavka za mjerenje parametara rasijanja

Rezultati eksperimentalnih mjerenja su prikazana na Sl. 8.



Sl. 8. Uporedni rezultati mjerenja i simulacija

Na osnovu uporednih rezultata mjerenja i simulacija prikazanih na Sl.8. može se uočiti da su mjerenja potvrdila rezultate simulacija u opsegu od 1GHz do 3GHz. Rezonantna učestanost antene koja se može vidjeti sa grafika je 1.8 GHz. Na Sl. 9 je prikazan simulirani dijagram zračenja.



Sl. 9. Dijagram zračenja predložene antene

#### IV. ZAKLJUČAK

Teorija fraktalne elektromagnetike predstavlja relativno novo polje istraživanja koje kombinuje elektromagnetiku i geometriju fraktala. Istraživanja u ovoj oblasti su dovela do razvoja novih antena sa poboljšanim širokopojasnim karakteristikama i smanjenim dimenzijama. Moderni komunikacioni sistemi dovode istraživanje u oblasti antena i antenskih sistema u prvi plan uslovljavajući razvoj malih antena koje će da rade na više rezonantnih učestanosti sa što moguće boljim širokopojasnim karakteristikama i jednostavnijom izradom.

U radu je predložena štampana antena sa CPW napajanjem zasnovana na bazi modifikovanog Sierpinski fraktala. Izvršene su simulacije raspodjele struje po površini antene, parametara rasijanja i dijagrama zračenja. Rezultati simulacija sprovedenih u softverskom paketu AWR pokazuju da antena ima širokopojasne karakteristike

i da zrači na više rezonantnih učestanosti.

Takođe su izvršena i eksperimentalna mjerenja parametara rasijanja koja su potvrdila rezultate simulacija.

#### LITERATURA

- [1] M. Kahrizi, T. K. Sarkar, and Z. A. Maricevic, "Analysis of a wide radiating slot in the ground plane of a microstrip line," *IEEE Trans. Microw. Theory Tech.*, vol. 41, no. 1, pp. 29–37, Jan. 1993.
- [2] H.-D. Chen and H.-T. Chen, "A CPW-fed dual-frequency monopole antenna," *IEEE Trans. Antennas Propag.*, vol. 52, no. 4, pp. 978–982, Apr. 2004.
- [3] N. Cohen, "Fractal antenna applications in wireless telecommunications," 1997, pp. 43–49.
- [4] Y. K. Choukiker, S. K. Behera, and S. K. Sharma, "Two and four-element wideband Sectoral fractal array antennas with omni-directional radiation patterns," in *Applied electromagnetics conference (AEMC)*, 2013 IEEE, 2013, pp. 1–2.
- [5] D. H. Wqrner and S. Ganguly, "An overview of fractal antenna engineering research," *IEEE Antennas Propag. Mag.*, vol. 45, no. 1, pp. 38–57, Feb. 2003.
- [6] Wojciech J. Krzysztofik, "Fractal Geometry in Electromagnetics Applications - from Antenna to Metamaterials," *Microw. Rev.*, Dec. 2013.
- [7] D. H. Werner, R. L. Haupt, and P. L. Werner, "Fractal antenna engineering: the theory and design of fractal antenna arrays," *IEEE Antennas Propag. Mag.*, vol. 41, no. 5, pp. 37–58, Oct. 1999.
- [8] C. Puente-Baliarda, J. Romeu, R. Pous, and A. Cardama, "On the behavior of the Sierpinski multiband fractal antenna," *IEEE Trans. Antennas Propag.*, vol. 46, no. 4, pp. 517–524, Apr. 1998.
- [9] Y. Ma, H.-W. Zhang, Y. Li, Y. Wang, and W. Lai, "Terahertz sensing application by using fractal geometries of split-ring resonators," *Prog. Electromagn. Res.*, vol. 138, pp. 407–419, 2013.
- [10] D. E. Anagnostou *et al.*, "Design, Fabrication, and Measurements of an RF-MEMS-Based Self-Similar Reconfigurable Antenna," *IEEE Trans. Antennas Propag.*, vol. 54, no. 2, pp. 422–432, Feb. 2006.
- [11] J. Romeu and J. Soler, "Generalized Sierpinski fractal multiband antenna," *IEEE Trans. Antennas Propag.*, vol. 49, no. 8, pp. 1237–1239, Aug. 2001.
- [12] A. Patnaik, D. Anagnostou, C. G. Christodoulou, and J. C. Lyke, "Neurocomputational analysis of a multiband reconfigurable planar antenna," *IEEE Trans. Antennas Propag.*, vol. 53, no. 11, pp. 3453–3458, Nov. 2005.

#### ABSTRACT

A fractal monopole antenna based on the modification of the Sierpinski fractal with CPW (Co-Planar Waveguide) feeder is presented in this paper. Using the fractal geometry for antenna design allows the design of electrically small antennas that radiate on multiple resonant frequencies. In order to test the performance and to confirm the theoretical considerations, the simulations of current distribution on the surface of the antenna, scattering parameters and the radiation pattern was performed in the AWR software package. Experimental measurements of scattering parameters has confirmed the theoretical considerations and the results of the simulation.

It is shown that electrically small antennas, that radiate the same radiation patterns at different frequencies while maintaining the desired characteristics, can be designed using fractal geometry.

#### CPW FEEDED MICROSTRIP MONOPOL ANTENNA BASED ON MODIFIED SIERPINSKI FRACTAL

Luka Lazović, Vesna Rubežić and Ana Jovanović

# Physical-chemical parameters and phytoplankton in Boka Kotorska Bay

Ana Perošević, Ivana Milašević, BIO-ICT Centre of Excellence in Bioinformatics, Montenegro  
 Danijela Joksimović, Dragana Drakulović, Institute of Marine Biology, University of Montenegro  
 Dijana Đurović, Institute of Public Health of Montenegro  
 Slavka Stanković, Faculty of Technology and Metallurgy, University of Belgrade

**Abstract** — Coastal areas of Adriatic Sea, including Boka Kotorska Bay, are exposed to pollution due to industrialization and urbanization, which has led to increased concentrations of various pollutants. These substances can have different hazardous effects on aquatic organisms. Among most harmful pollutants are certainly heavy metals. In this study contents of three essential, Cu, Zn and Cr, and three toxic metals, Hg, Pb and Cd, were determined in mussel samples collected in Boka Kotorska Bay, Montenegro, in order to evaluate their concentrations in different locations and seasons, as well as impact of temperature and salinity on those concentrations. The obtained results were compared with their maximum allowable concentrations (MAC), and also with results obtained in *Mytilus galloprovincialis* in other areas of Adriatic Sea. Besides that, in order to assess levels of pollution in Boka Kotorska Bay, also total organic carbon (TOC) contents were determined in sediment samples, as well as nutrient contents in seawater as one of fundamental parameters in oceanographic research. Considering possible presence of potentially toxic and toxic phytoplankton species in investigated area, analysis of phytoplankton composition was also done.

**Keywords** — heavy metals, marine environment, nutrient contents, phytoplankton, total organic carbon

## I. INTRODUCTION

**M**ARINE environment is under the strong impact of human activities all over the world. As a result of industrialization, the coastal areas of seas and oceans are exposed to pollution with various chemical substances,

A. Perošević, BIO-ICT Centre of Excellence in Bioinformatics, Džordža Vašingtona bb, 81000 Podgorica, Montenegro (tel: +38239828942, e-mail: [anap@ac.me](mailto:anap@ac.me))

D. Joksimović, Institute of Marine Biology, University of Montenegro, Dobrota bb, 85330 Kotor, Montenegro (tel: +38263204933, e-mail: [daniijela.j@ac.me](mailto:daniijela.j@ac.me))

D. Drakulović, Institute of Marine Biology, University of Montenegro, Dobrota bb, 85330 Kotor, Montenegro (tel: +8263204931, e-mail: [dragana.dr@ac.me](mailto:dragana.dr@ac.me))

D. Đurović, Institute of Public Health of Montenegro, Džona Džeksona bb, 81000 Podgorica, Montenegro (tel: +38267266459, e-mail: [dijana.djurovic@ijzcg.me](mailto:dijana.djurovic@ijzcg.me))

I. Milašević, BIO-ICT Centre of Excellence in Bioinformatics, Džordža Vašingtona bb, 81000 Podgorica, Montenegro (+38268581234, e-mail: [ivana.mi@ac.me](mailto:ivana.mi@ac.me))

S. Stanković, Faculty of Technology and Metallurgy, University of Belgrade, Karnegejeva 4, 11000 Beograd (e-mail: [slavka@tmf.bg.ac.rs](mailto:slavka@tmf.bg.ac.rs))

including heavy metals [1]. During the last decades this has led to increased concentrations of heavy metals in marine environment. Their inputs include urban run-off, industrial effluents, anti-fouling paints for boats, mining operations and atmospheric depositions [2]. Due to anthropogenic and natural emissions and their easy assimilation and bioaccumulation in marine organisms, heavy metals have long been recognized among the most common pollutants of marine ecosystems [1].

Metals are probably the oldest toxins known to human. They differ from other toxic substances because they are neither created nor destroyed by humans [3]. Once released into the environment, they do not disappear, but accumulate in soils, sediments and biota. Although, there are two categories of metals, essential and nonessential, all metals are potential toxins. Essential metals, i.e. micronutrients in living organisms usually occur in trace amounts that are precisely defined for each species. Both, their deficiency and high excess, badly affect living organisms. On the other hand, non-essential metals have no known biological function and are very toxic even at low concentrations. They, also directly or indirectly, throughout air, water, and food (plants, animals) get into human bodies [4].

Since they are accumulated in sediments, concentrations of metals in sediments are often several times higher than in water column and therefore, sediments represent a secondary source of these contaminants to the water column and to living organisms. Trace metals can be released from sediments to the overlying water via natural or anthropogenic processes, consequently causing potential danger to the ecosystem [5]. However, total concentrations of metals in sediments usually do not represent biologically available fraction. Because of that and since many marine organisms accumulate high levels of metals from their environment, living organisms are often used as bioindicators of marine pollution. Mussels *Mytilus galloprovincialis* are well known as good indicator of marine pollution in coastal areas, since they accumulate a wide range of metals in their soft tissues [6].

Basal concentrations of metals in the organisms can be influenced by several environmental factors, including temperature, salinity, pH, organic carbon, food availability, dissolved oxygen, sediment grain size and hydrologic features of the system [7]. Concentrations of metals in mussels can depend also on presence of nutrients, as well as geological characteristics, and therefore demonstrate geographical differences [8].

Because of increasingly more people living in coastal areas, there are more nutrients entering coastal waters from wastewater treatment facilities, runoff from land in urban areas during rains, and from farming. All of these factors can lead to increased nutrient pollution, the process where too many nutrients are added to water and can act like fertilizer, causing excessive growth of algae. Severe algal growth blocks light that is needed for plants, such as seagrasses, to grow. When the algae and seagrass die, they decay. In the process of decay, the oxygen in the water is used up and this leads to low levels of dissolved oxygen in the water. This, in turn, can lead to death of many aquatic animals [9].

An important parameter in environmental evaluation is also content of total organic carbon (TOC). It is used as an indicator of pollution and eutrophication rate. Sources of organic carbon include organic matter from overland runoff and shoreline erosion, and primary productivity within the bays, all of which eventually settle to the bay bottom and are incorporated into the sediment. So, TOC content is proportional to organic matter, which is a primary source of food for benthic organisms. However, too much organic matter can lead to the depletion of oxygen in the sediment and overlying water, which can have a harmful effect on the benthic and fish communities [10].

Determination of phytoplankton assemblages in coastal areas is also very important since these areas are characterized by high spatial and temporal variability of environmental parameters. Mostly it is because of increased impact of human activities on the functioning of coastal ecosystems [11].

Because of that, the aim of this work was to assess pollution of Boka Kotorska Bay environment, based on chemical parameters, such as heavy metal contents, nutrient contents, total organic carbon, as well as phytoplankton composition.

## II. MATERIAL AND METHODS

Sampling sites are shown in Fig. 1. Sampling was conducted from fall 2014 to spring 2015 at three locations: IMB, fish and mussel farm COGImar, both in Kotor Bay, and one location close-by the open sea, at the entrance of the Bay, near the Žanjice beach.

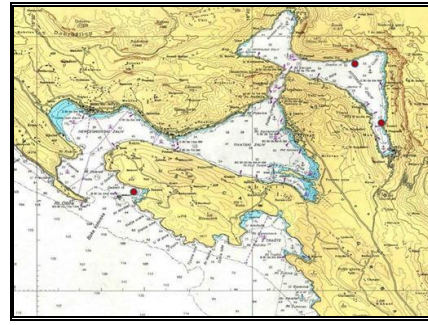


Fig.1. Map of investigated positions

Samples of marine water for analyses of physic-chemical parameters and for phytoplankton analyses were taken every month from four depths: at the surface, 2 m and 4m depth and above the bottom. Samples of mussels for heavy metals determination were taken every three months. Samples were taken by Niskin bottle of 5l.

### A. Analysis of seawater

In water samples further parameters were determined: presence of phosphates, nitrates, nitrites, silicates, total nitrogen and total phosphorus (nutrients content), as well as physical and chemical parameters of seawater (temperature, salinity, transparency...). Nutrients content was determined by UV/VIS spectrophotometer (Perkin Elmer and Analytik Jena SPECORD PLUS).

### B. Analysis of heavy metals in mussels

Mussels of similar shell length were collected at the same time as the sediment samples. They were first cleaned and rinsed with deionized water and then dissected by removing the byssus and shells. All samples were frozen, freeze-dried and then digested with HNO<sub>3</sub> and H<sub>2</sub>O<sub>2</sub> in closed vessel microwave digestion system under high temperature and pressure (Milestone, Ethos 1). For determination of heavy metals inductively coupled plasma-optical emission spectrometer (ICP-OES, Spectro Arcos) was used and also Direct Mercury Analyzer (Milestone, DMA-80), for determination of Hg. The accuracy of the applied analytical procedures was tested using certified reference materials, NIST 2976 and IAEA 158 (Mussel homogenate and Marine sediment).

### C. Analysis of TOC in sediments

Sediment samples for TOC determination were packed in polypropylene boxes until transferring to the laboratory. Those samples were sieved, homogenized and dried at 105°C. After that, TOC content in sediments was determined by TOC/TN analyzer (TOC/TN Analytik Jena Multi N/C 3100), using cc HCl.

### D. Analyses of phytoplankton

Analyses of phytoplankton were performed on Leica inverted microscope according Utermöhl (1958) [12]. Samples were put in chambers of 25 ml and after 24 h of

sedimentation, enumerated at magnification 100, 200 and 630x following Utermöhl (1958).

### III. RESULTS AND DISCUSSION

Concentrations of metal contents as well as temperature and salinity values of seawater measured at depths of mussel growth are given in Table 1. The highest heavy metals contents in mussel samples were mostly at Žanjice, especially in the fall, except for the Pb and Cd in a few samples. Žanjice is a beach near the open sea, but with a lot of cottages and restaurants. Discharge of their wastewaters is especially increased during summer tourist season, when there is also a problem of recreational activities. Regarding seasonal variations of metals in mussels, the general conclusion is that metal contents were decreased in winter compared to fall samples and in spring they were increased along with increasing of temperature and salinity in marine water.

Table 1. Trace elements in mussels (*Mytilus galloprovincialis*) samples (mg/kg) and temperature and salinity of seawater at depth of mussel growth

Season	Autumn			Winter			Spring		
	2014			2015					
	IMB	COGI	Žanjice	IMB	COGI	Žanjice	IMB	COGI	Žanjice
Cu	7.00	8.02	8.04	1.04	1.04	0.88	2.52	5.39	5.83
Zn	105	118	240	8.66	9.12	19.9	131	224	226
Cr	0.72	0.74	1.58	0.14	0.24	0.44	0.80	0.86	1.10
Pb	2.02	1.66	1.26	0.28	0.30	0.26	1.49	1.98	1.92
Cd	1.30	1.94	2.62	0.10	0.12	0.26	1.20	1.77	1.55
Hg	0.202	0.221	0.307	0.403	0.226	0.229	0.127	0.275	0.207
Temperature	16.8	18.5	19.1	12.4	11.6	13.8	19.6	20.2	21.5
Salinity	25.4	31.7	37.6	21.3	13.8	11.3	29.2	30.9	36.4

Comparing the obtained results with maximum allowable concentrations (MAC) for certain metals, it was found that the metal contents in mussels from Boka Kotorska Bay were mostly lower than permitted limits for *M. galloprovincialis* [13] Only Zn at COGImar in spring (224 mg/kg) and at Žanjice in both, fall (240 mg/kg) and spring (226 mg/kg), and also Hg at Žanjice in the fall (0.307 mg/kg), at COGImar in the spring (0.275 mg/kg) and IMB (0.403 mg/kg) were found to be above limits, but these results are within the range of values commonly found in other areas of Adriatic Sea (Table 2).

Table 2. The concentration of heavy metals (mg/kg) in mussels from the Adriatic coast

Location	Cu	Zn	Cr	Pb	Cd	Hg
SE Adriatic, Albania <sup>[14]</sup>	4.61–28.9	59.8–245	1.44–12.5	1.39–5.69	0.27–0.77	0.08–0.42
W Istrian coast, N Adriatic, Croatia <sup>[1]</sup>	1.25–12.4	10.0–32.3	/	0.46–0.78	0.11–0.30	0.00–0.04
E Adriatic, Croatia <sup>[15]</sup>	1.98–11.0	49.4–418	0.41–4.61	0.24–3.69	0.39–2.40	0.08–0.28
SW Adriatic, Italy <sup>[2]</sup>	4.66–19.2	/	0.96–9.46	0.37–3.25	0.38–1.84	0.10–0.81
MAC <sup>[13]</sup>	10	200	2.5	3.2	3.7	0.23
SE Adriatic, Montenegro (This study)	0.88–8.04	8.66–240	0.14–1.58	0.26–2.02	0.10–2.62	0.13–0.40

Regarding TOC contents in sediments from Boka Kotorska Bay we can say that the highest contents were found at COGImar. TOC contents from Žanjice were in the range from 4.99 g/kg to 8.58 g/kg, from IMB: 7.73 g/kg to 12.6 g/kg, and the results from COGImar were in the range from 12.0 g/kg to 14.0 g/kg. Generally, the sediments from anoxic areas contain larger amount of organic matter as a consequence of increased productivity which could lead to oxygen depletion. Besides that, high organic carbon is considered a sign of frequent algal blooms in the overlying water column, the blooms being a result of increased nutrient (nitrogen and phosphorus) loadings into the system. Thus, on the COGIMAR we found two main sources of the TOC content in the superficial sediments: inflows of fresh water and biological inputs since it is a fish farm and has a daily food input in the form of organic matter.

Nutrient contents were in the range of: 3.18 µg/L - 220 µg/L for nitrates, 0.33 µg/L - 7.76 µg/L for nitrites, 0.41 µg/L - 23.5 µg/L for phosphates and 10.3 µg/L - 649 µg/L for silicates. The maximum measured values are within the range of reference values for seawater.

Regarding phytoplankton, maximum ( $2.98 \times 10^5$  cells/l) and minimum ( $1.4 \times 10^3$  cells/l) of microplankton was noticed in November 2014. Diatoms was also the highest in November 2014 and abundance of dinoflagelates was maximal in November ( $1.37 \times 10^4$  cells/l). It was recorded 100 taxa during investigation period and of these taxa it was found nine toxic and potentially toxic species. From diatoms, it was recorded *Pseudo-nitzschia* spp and from dinoflagellates: *D. acuta*, *D. caudata*, *D. fortii*, *D. tripos*, *Lingulodinium polyedra*, *Phalacrocoma rotundatum* and *Prorocentrum cordatum*.

Diatoms were the most abundant group of phytoplankton. Similar domination of diatoms was noticed

in Boka Kotorska Bay by Drakulovic et al. 2012 [16]. Potentially toxic diatoms *Pseudo-nitzschia*, which is important as these species produce domoic acid, was the most frequent. It is generally frequent species in Adriatic Sea [17]. From toxic dinoflagellates, *Prorocentrum cordatum* was recorded in higher abundance. The highest growth of phytoplankton was noticed in autumn period when supply of nutrients was enough to provide its development.

#### IV. CONCLUSION

In common with other coastline areas, Montenegrin coast is also under a great impact of anthropogenic factors and the activities on the shore. It receives a heavy influx of sewages, industrial effluents, domestic and agricultural wastes, all containing various hazardous chemicals, which can cause deleterious effects on aquatic organisms.

Different studies have shown the importance of analyzing multiple stressors in marine environment. However, published reports on the combined effects of multiple stressors, their interaction and seasonal fluctuations, which are common in marine environment and can especially affect metal contents in mussels, are still limited. Because of that and considering the fact that phytoplankton growth is affected by interaction of nutrients uptake, light and temperature, very important step in these analyses could be the use of "smart buoy" systems, which provide real-time data, especially about physical-chemical parameters of water, such as temperature, salinity, pH and some nutrient contents, as well as historical information for specific sites, including weather conditions. These systems are representative examples how the new ICT solutions can improve ecological monitoring of seawater parameters, since it enables automated and online ecological monitoring of seawater parameters.

Considering interactions of all these parameters, for further data processing, the use of different softwares for statistical data analysis is preferable.

#### ACKNOWLEDGEMENT

This work has been supported by the Ministry of Science of Montenegro and HERIC project through the BIO-ICT Centre of Excellence (Contract No. 01-1001).

#### REFERENCES

- [1] Bilandžić, N., Sedak, M., Čalopek, B., Džafić, N., Mišetić-Ostojić, D., Potočnjak, D. (2015) Metal Content in Four Shellfish Species from the Istrian Coast of Croatia. *Bull Environ Contam Toxicol* 95:611–617
- [2] Spada, L., Annicchiarico, C., Cardellicchio, N., Giandomenico, S., Di Leo, A. (2013) Heavy metals monitoring in the mussel *Mytilus galloprovincialis* from the Apulian coast (Southern Italy). *Mediterr Mar Sci* 14:99–108
- [3] Klaassen, D.C. (2001) Casarett and Doull's Toxicology: The basic science of poisons, Sixth edition, McGraw-Hill, New York, USA
- [4] Stanković, S., Kalaba, P., Stanković, A.R. (2014) Biota as toxic metal indicators, *Environ Chem Lett* 12:63–84
- [5] Joksimović, D., Castelli, A., Mitrić, M., Martinović, R., Perošević, A., Nikolić, M., Stanković, S. (2016) Metal pollution and ecotoxicology of the Boka Kotorska Bay. In Joksimović, A. et al. (eds.), *The Boka Kotorska Bay Environment*, Hdb Env Chem, Springer, Heidelberg, DOI 10.1007/698\_2016\_40
- [6] Jović, M., Stanković, S. (2014a) Determination of marine pollution by comparative analysis of metal pollution indices. *Arch Biol Sci, Belgrade*, 66 (3):1205-1215
- [7] Boening, D.W. (1999) An evaluation of bivalves as biomonitors of heavy metals pollution in marine waters. *Environmental Monitoring and Assessment* 55:459–470
- [8] Fattorini, D., Notti, A., Di Mento, R., Cicero, A.M., Gabellini, M., Russo, A., Regoli, F. (2008) Seasonal, spatial and inter-annual variations of trace metals in mussels from the Adriatic sea: A regional gradient for arsenic and implications for monitoring the impact of off-shore activities, *Chemosphere* 72:1524–1533
- [9] <http://oceanservice.noaa.gov/facts/nutpollution.html>
- [10] Wells, D. (2004) Total organic carbon in Maryland Coastal Bays sediments: Status of a regulator of chemical and biological processes. *Marylands coastal Bays: Ecosystem Health Assessment*, Maryland Department of Natural Resources, Tide water Ecosystem Assessment, Annapolis, 380p.
- [11] Cloern, J.E. (1999) The relative importance of light and nutrient limitation of phytoplankton growth: A simple index of coastal ecosystem sensitivity to nutrient enrichment. *Aquat Ecol* 33 (1): 3–16
- [12] Utermöhl, H. (1958) Zur Vervollkommung der quantitativen Phytoplankton Methodik, *Mitt. Int. Ver.Theor. Angew. Limnol.* 9: 1–38
- [13] Cantilo, A. Y. (1997) *World Mussel Watch Data*. Silver Spring, 1-209.
- [14] Çullaj, A., Lazo, P., Duka, S. (2006) Heavy metals and metallothionein levels in mussel samples from the Albanian seacoast. *MAP/Med Pol. In Biological effects monitoring programme, MAP Technical Reports Series No. 166*, Italy: Alessandria. pp. 141–151.
- [15] Kljaković-Gašpić, Z., Ujević, I., Zvonarić, T., Barić, A. (2007) Biomonitoring of trace metals (Cu, Cd, Cr, Hg, Pb, Zn) in Mali Ston Bay (eastern Adriatic) using the Mediterranean blue mussel (1998–2005). *Acta Adriat.* 48, 73–88.
- [16] Drakulović, D., Pestorić, B., Cvijan, M., Krivokapić, S., Vuksanović, N. (2012) Distribution of phytoplankton community in Kotor Bay (south-eastern Adriatic Sea). *Cent Eur J Biol* 7(3): 470-486.
- [17] Ujević, I., Ninčević-Gladan, Ž., Roje, R., Skejic, S., Arapov, J., Marasović, I. (2010) Domoic Acid - A New Toxin in the Croatian Adriatic Shellfish Toxin Profile. *Molecules* 15:6835-6849

# Condition of cultured sea bass and sea bream and occurrence of their spawning in the cages in Boka Kotorska Bay

Mandić Milica, *Institute of marine biology*, Gvozdrenović Slađana, *BIO-ICT Centre of Excellence in Bioinformatics*, Joksimović Aleksandar, *Institute of marine biology*, and Pešić Ana, *Institute of marine biology*

**Abstract** — In this investigation total of 250 individuals of Gilthead sea bream (*Sparus aurata* Linnaeus, 1758) and 245 individuals of European sea bass (*Dicentrarchus labrax* Linnaeus, 1758) were sampled on the fish farm “COGIMAR” located in Orahovac (Boka Kotorska Bay). Samples were taken from September 2015 up to January 2016. Concurrently, sampling of ichthyoplankton was performed on 4 locations – of which 2 are located on the same farm and the other 2 are located on the another fish farm which is also located in the area of Boka Kotorska Bay. The aim of this study is to determine the state of cultured fishes, the occurrence of their spawning in the cages and to determine whether it can lead to ecological consequences on wild populations of sea bass and sea bream.

**Key words** — Boka Kotorska Bay, condition factor, European sea bass, Gilthead sea bream, ichthyoplankton composition, Length-weight relationship

## I. INTRODUCTION

Development of cage fish farming in Montenegro, although the environmental conditions are very favorable, is at a very low level. One of the main reasons is the lack of defined areas for cultivation, and spatial planning of the coastal part of the Montenegrin coast, which would be the basis for definite defining the location and provision of safe investments in this field. Currently there are only two fish farms in the area of Boka Kotorska Bay, which were the subject of this research. In addition to the known negative impact that fish farming has on the environment and benthic biocenoses in the immediate vicinity of the farm, there is another, not enough scientifically explored, impact of spawning of cultured organisms and its possible ecological stimulus on wild fish populations. Available

M. Mandić, Institute of marine biology, University of Montenegro, Dobrota bb, 85330 Kotor, Montenegro (phone: +382 32 303 140; e-mail: [mamilica@ac.me](mailto:mamilica@ac.me))

S. Gvozdrenović, BIO-ICT Centre of Excellence in Bioinformatics, Faculty of Electrical Engineering, University of Montenegro, Cetinjska br. 2, 81000 Podgorica, Montenegro (phone: +382 32 334 569; e-mail: [sladjanag@ac.me](mailto:sladjanag@ac.me))

A. Joksimović, Institute of marine biology, University of Montenegro, Dobrota bb, 85330 Kotor, Montenegro (phone: +382 32 334 569; e-mail: [acojo@ac.me](mailto:acojo@ac.me))

A. Pešić, Institute of marine biology, University of Montenegro, Dobrota bb, 85330 Kotor, Montenegro (phone: +382 32 334 569; e-mail: [pesica@ac.me](mailto:pesica@ac.me))

publications which have dealt with this issue pointed out the importance of intensive research on the effect of cage farming from different aspects [1], [2]. Fish farms may affect the presence, abundance, residence times and diet of fishes in a given area [3], [4] but despite the obvious potential for attraction, little is known about the ecological effects on wild fish assemblages [5]. Gilthead sea bream and European sea bass are the most cultivated species in aquaculture at Mediterranean, and only two fish species cultivated in marine aquaculture in Montenegro. *S. aurata* is protandric hermaphrodite species, maturing first as male (during the first or second year of age, at length 20-30 cm) and after the second or third year of age, at length 33-40 cm, as female. Spawning of this species happens generally from October to December, with sequenced spawning during the whole period [6], [7], [8], [9]. European sea bass is a gonochoristic species with a polygenic sex determining system (Figure 1), where temperature is known to influence the sex ratio [10].



Fig. 1. Sea bass gonads showing sexual dimorphism

## II. MATERIAL AND METHODS

Sampling of Gilthead sea bream (*Sparus aurata*, Linnaeus 1758) and European sea bass (*Dicentrarchus labrax*, Linnaeus 1758) from fish farm COGIMAR started in September 2015, and continued until January 2016. Each month 50 individuals were sampled, except November when 45 individuals of European sea bass were sampled instead of 50. The length was measured to the nearest mm, and the total body weight (using a electronic balance) to the nearest 1 g. Length-weight relationship was determined for the entire sample, according to the formula  $W=a \times L^b$ . Parameters  $a$  and  $b$  were estimated using ordinary least-square regression after transforming the data in nature logarithms ( $\ln W = \ln a + b \times \ln TL$ ). Condition factor,  $Ka$  was calculated according to the formula  $Ka = 100(W/L^3)$ ,

where  $W$  is the whole body wet weight in grams and  $L$  is the length in centimeters; the factor 100 is used to bring  $Ka$  close to a value of one [11]. Additionally condition factor by [12] was calculated according to the formula  $Ka = W/aL^b$ , where  $a$  and  $b$  are the coefficient and exponent of the length–weight relationship, respectively.

Samples for qualitative and quantitative composition of ichthyoplankton were taken using PairOVET (modified CalVet) plankton net at 4 stations – of which 2 are located on the COGIMAR farm and the other 2 are located on the another fish farm which is also located in the area of Boka Kotorska Bay (Figure 2). Diameters of net cylinders were 25 cm each, and the total mouth area was 0.098 m<sup>2</sup>, with mesh size of 0.160 mm. Net was towed vertically at a speed of 0.5 m\*sec<sup>-1</sup>, from the depth of 5 m above the bottom. Plankton material was preserved immediately on board in 2.5% solution of buffered formaldehyde and processed in the laboratory.



Fig. 2. Locations of fish farms with geographical coordinates of sampling stations

### III. RESULTS

A total 250 individuals of Gilthead sea bream (*S. aurata*) and 245 individuals of European sea bass (*D. labrax*) from COGIMAR farm were processed. Total length of *S. aurata* ranged from 24 to 33.1 cm, with mean value  $28.09 \pm 1.54$  cm (Figure 3), while weight ranged from 95 to 515 g, with mean value  $344.80 \pm 51.64$  g. Total length of *D. labrax* ranged from 20.8 to 43.5 cm, with mean value  $32.19 \pm 3.40$  cm (Figure 3), while weight ranged from 38.5 to 880 g, with mean value  $368.31 \pm 133.87$  g.

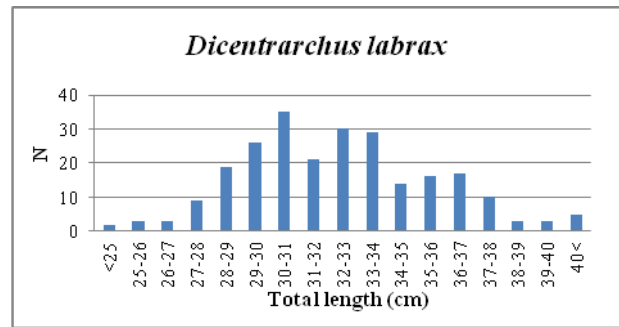
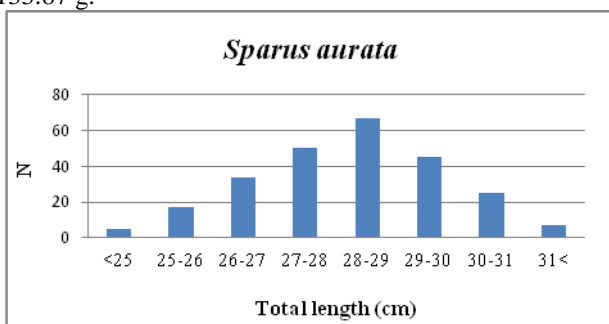


Fig. 3. Total length frequency distribution of *S. aurata* and *D. labrax*

Length-weight relationships of *S. aurata* and *D. labrax* are presented in Figure 4. The highest value of parameter  $b$  for *S. aurata* is recorded in December,  $b = 3.1207$ , and the lowest in October,  $b = 2.0272$ . The highest value of parameter  $b$  for *D. labrax* is recorded in November,  $b = 3.3140$ , and the lowest in October,  $b = 2.4765$ . The length-weight relationship parameters for the total sample of *S. aurata* and *D. labrax* also, show that the parameter  $b$  (slope) has a value lower than the ideal isometric value of 3, implying a faster growth in length compared to the gain in weight, or, alternatively, the body assuming a more elongated shape with growth.

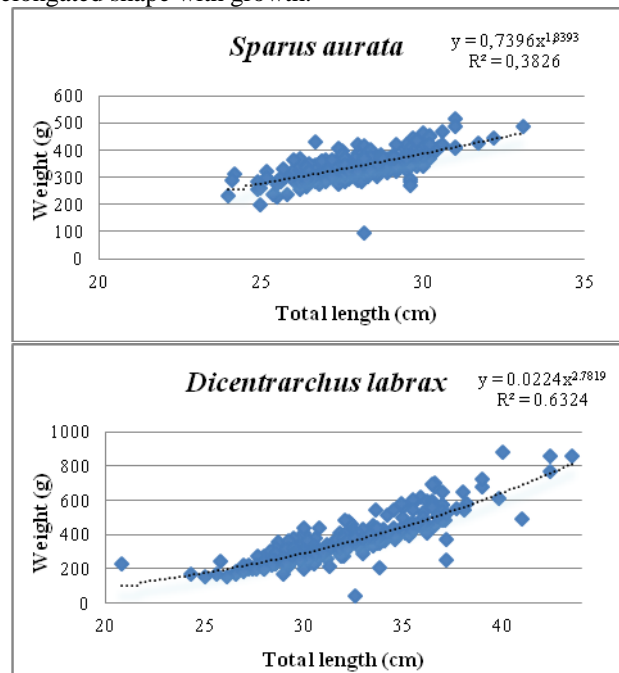


Fig. 4. Length-weight relationship of *S. aurata* and *D. labrax*

Condition factor of *S. aurata* and *D. labrax* by month is presented in Figure 5. Highest value of condition factor for *S. aurata*, according to Fulton's method was reported in September,  $Ka = 1.7837$ , the lowest value was recorded in January,  $Ka = 1.3950$ , while average value was  $Ka = 1.5571 \pm 0.15$ . Highest value of condition factor for *S. aurata* according to Le Cren's method was reported in January,  $Ka = 1.0152$ , the lowest value was recorded in December,  $Ka = 0.9997$ , while average value was  $Ka = 1.000 \pm 0.006$ . Highest value of condition factor for *D. labrax*, according to Fulton's method was reported in September,  $Ka = 1.3048$ , the lowest value was recorded in



January,  $K_a = 0.9531$ , while average value was  $K_a = 1.0754 \pm 0.14$ . Highest value of condition factor for *D. labrax* according to Le Cren's method was reported in September,  $K_a = 0.8372$ , the lowest value was recorded in December,  $K_a = 0.6427$ , while average value was  $K_a = 0.7354 \pm 0.07$ .

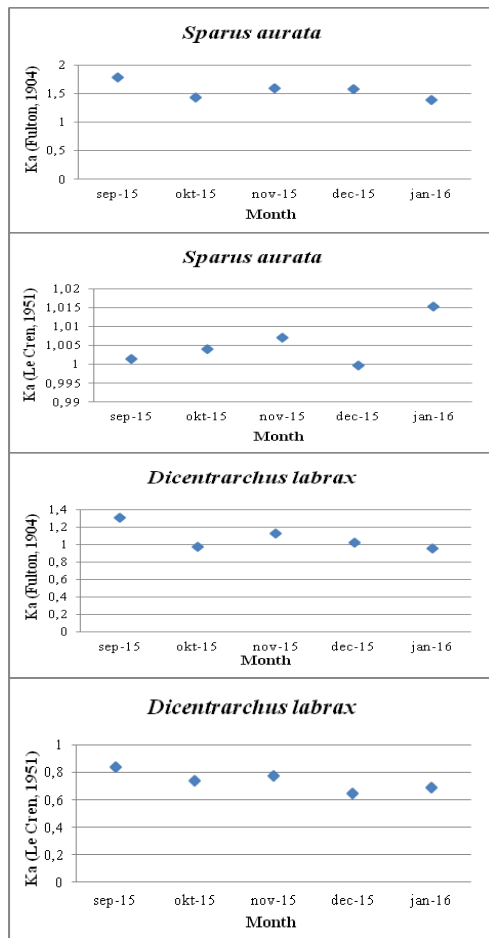


Fig. 5. Condition factor of *S. aurata* and *D. labrax* by month according to Fulton (1904) and Le Cren (1951)

Analyses of qualitative and quantitative composition of ichthyoplankton showed that during the studied period 7 species of fishes spawned in the immediate vicinity of fish farms – *Engraulis encrasicolus* (Linnaeus, 1758), *Serranus hepatus* (Linnaeus, 1758), *Lythognathus mormyrus*, *Sardina pilchardus* (Walbaum, 1792), *Arnoglossus laterna* (Walbaum, 1792), *S. aurata* and *D. labrax*. Mean abundance of fish eggs and larvae ranged from 10,19-30,57 eggs/larvae per m<sup>2</sup> of sea surface.

Spawning of sea bass and sea bream has been observed only in Stoliv farm during December 2014, January and February 2015. During September, October and November 2014, dominance belongs to spawning of anchovy on both farms; while in December 2014 it was noticed only spawning of sea bream in Stoliv farm. In January 2015 it was observed spawning of *A. laterna* and sea bass in the Stoliv farm, with numerous early developmental stages of shellfish. February 2015 was characterized by a continuation of sea bream spawning on the Stoliv farm, while the farm COGIMAR was negative to the early life

stages of fishes during December 2014, January and February 2015.

#### IV. DISCUSSION

Parameters of length-weight relationship are indicators of fish growth, faster growth in length or in weight. When parameter  $b$  (slope) has a value lower than the ideal isometric value of 3, implies a faster growth in length compared to the gain in weight, or, alternatively, the body assuming a more elongated shape with growth. The highest value of parameter  $b$  for *S. aurata* is recorded in December,  $b = 3.1207$ , and the lowest in October,  $b = 2.0272$ , which corresponds to beginning of spawning of this species [13]. The highest value of parameter  $b$  for *D. labrax* is recorded in November,  $b = 3.3140$ , and the lowest in October,  $b = 2.4765$ . Similar results for parameter  $b$  of length-weight relationship for Gilthead sea bream was reported for some other parts of Mediterranean:  $b = 2.736$  [14],  $b = 2.835$  [15],  $b = 2.985$  [16], while for Adriatic Sea reported value is  $b = 3.052$  [17]. For Adriatic area reported value of  $b$  parameters for European sea bass are  $b = 3.065$  [18] and  $b = 3.146$  [19], while for North – West Aegean Sea are reported values  $b = 3.158$  and  $b = 3.200$  [20].

Condition factor is a measure of the condition of fish. If  $K_a < 1$  fish is in a poor condition, if  $K_a > 1.4$  a fish is in good to excellent conditions. For both species the lowest values of condition factor were reported for January (*S. aurata* -  $K_a = 1.3950$ , *D. labrax* -  $K_a = 0.9531$ ), when both species are at the end of spawning season and reserves are used for the development and maturation of the gonads, rather than to the growth and development of individuals, for what would normally be used. Earlier reported results of condition factor for *S. aurata* shows similar values:  $K_a = 1.29 - 1.49$  [21], or much higher values  $K_a = 1.72 - 2.59$  [22]. For *D. labrax* reported values for  $K_a$  are  $K_a = 1.0713 - 1.5862$  [23].

Analysis of qualitative and quantitative composition of ichthyoplankton showed that several economically important species spawns near the fish farms. Although it is known that the position of fish eggs and larvae in seawater depends mostly on the movement of water (streams, waves, etc.), it is very likely that the adult specimens are attracted to spawn in the vicinity of farms due to the availability of food. Actually, fish farms play a role such as artificial reefs attracting certain types of fishes, which are aggregated in their vicinity [5]. It is interesting to note the finding of *A. laterna* larvae in January 2015, given that this species spawn during spring and summer months.

Although it was noticed occurrence of sea bream and sea bass spawning, it cannot be said with certainty that it derived from cultured organisms. It is probably the spawning of individuals belonging to wild populations, primarily because of the relatively small total length of cultured fishes and definite answer can be obtained by further genetic analysis. Recent research has found that the so-called 'escape through spawning' may have significant negative environmental effects [2], which indicates the need of intensifying similar studies, focusing on analysis of larger specimens and genetic analysis of ichthyoplankton

in order to protect wild populations from cultured organisms and to protect natural resources.

In the frame of BIO-ICT project several investigations are developing with aim of improvement of farming process and increasing of quality of monitoring of ecological state in farms vicinity. SEMaR Buoy system developed in the frame of this project, will be mounted at every fish farm in Boka Kotorska Bay, and this system will enable continuous monitoring of physical parameters of sea water which will provide useful information's about time of spawning and incubation time of fish eggs and larvae. Temperature data gathered from SEMaR Buoy system will be used for optimization of amount of food spread at cages which will decrease pressure of nutrients on the environment and improve quality of fish products. Through BIO-ICT project investigations of sea currents in the farms vicinity is conducted. Sea currents influence distribution of eggs and larvae and distribution of food spreading in cages. Based on a collected data it will be possible to predict areas of farm where "dropping" of food will have best usage by fish, and decrease accumulation of uneaten food bellow of cages. Introduction of ICT systems in farming process will improve quality of environment and production.

#### ACKNOWLEDGEMENT

This work has been supported by the Ministry of Science of Montenegro and the HERIC project through the BIO-ICT Centre of Excellence (Contract No. 01-1001).

#### REFERENCES

- [1] I. Karakassis, M. Tsapakis, E. Hatziyanni, K. N. Papadopoulou, and W. Plaiti, "Impact of cage farming of fish on the seabed in three Mediterranean coastal areas," *Journal of Marine Science*, vol. 57, pp. 1462–1471, 2000.
- [2] S. Somarakis, M. Pavlidis, C. Saapoglou, C. Tsigenopoulos, and T. Dempster, "Evidence for 'escape through spawning' in large gilthead sea bream *Sparus aurata* reared in commercial sea-cages," *Aquaculture Environment Interact.*, vol. 3, pp. 135–152, 2013
- [3] D. N. Carss, "Concentrations of wild and escaped fishes immediately adjacent to fish farm cages," *Aquaculture*, vol. 90, pp. 29–40, 1990
- [4] A. Bjordal, A. B. Skar, "Tagging of saithe (*Pollachius virens* L.) at a Norwegian fish farm: preliminary results on migration," *ICES Counc. Meet. Pap.*, pp. 35, 1992
- [5] T. Dempster, P. Sanchez-Jerez, J. T. Bayle-Sempere, F. Giménez Casualdero, and C. Valle, "Attraction of wild fish to sea-cage fish farms in the south-western Mediterranean Sea: spatial and short-term variability," *Marine Ecological Progress Series*, vol. 242, pp. 237–252, 2002
- [6] M. L. Bauchot, J. C. Hureau, and J. C. Miguel, "Sparidae," in W. Fischer, G. Bianchi, and W.B. Scott, (eds.) FAO species identification sheets for fishery purposes. Eastern Central Atlantic, FAO, Rome, vol. 4., 1981
- [7] C. D. Buxton, and P. A. Garratt, "Alternative reproductive styles in seabreams (Pisces: Sparidae)," *Environ. Biol. Fish.*, vol. 28, no. 1-4, pp. 113-124, 1990
- [8] Pešić, A., Đurović, M., Joksimović, A., Marković, O., Ikica, Z. (2015). Condition of fish, sea bream (*Sparus aurata*) and sea bass (*Dicentrarchus labrax*) from fish farm in Boka Kotorska Bay (South-East Adriatic). VII International Conference "Water & Fish". Book of abstracts. 419- 424.
- [9] A. Pešić, A. Joksimović, Z. Ikica, M. Nikolić, S. Gvozdenović, „Condition of Sea Bass (*Dicentrarchus labrax*) from fish farm in Bokakotorska Bay," 45<sup>th</sup> Annual Conference of the Serbian Water Pollution Control Society "WATER 2016", pp. 445-449, 2016
- [10] M. Vandeputte, M. Dupont-Nivet, H. Chavanne, and B. Chatain, "A Polygenic Hypothesis for Sex Determination in the European Sea Bass *Dicentrarchus labrax*," *Genetics*, vol. 176, no. 2, 1049–1057, 2007
- [11] T. W. Fulton, "The rate of growth of fishes," 22<sup>nd</sup> Annual Report of the Fishery Board of Scotland, vol. 3, pp.141-241, 1904
- [12] E. D. Le Cren, "The length-weight relationship and seasonal cycle in gonad weight and condition in the perch (*Perca fluviatilis*)," *J. Anim. Ecol.*, vol. 20, pp. 201-219, 1951
- [13] M. L. Whitehead, J. C. Bauchot, J. Hureau Nielsen and E. Tortonese, "Fishes of the north-eastern Atlantic and the Mediterranean," UNESCO, Paris, vol. 2, 1986
- [14] T. Ceyhan, O. Akyol, and M. Erdem, "Length-weight relationships of fishes from Gökova Bay, Turkey (Aegean Sea)," *Turk. J. Zool.*, vol. 33, pp. 69-72, 2009
- [15] L. Sangun, E. Akamca, and M. Akar, "Weight-length relationships for 39 fish species from the North-Eastern Mediterranean coast of Turkey," *Turk. J. Fish. Aquat. Sci.*, vol. 7, pp. 37-40, 2007
- [16] E. Cicek, D. Avsar, H. Yeldan, and M. Ozutok, "Length-weight relationships for 31 teleost fishes caught by bottom trawl net in the Babadillimani Bight (northeastern Mediterranean)," *J. Appl. Ichthyol.*, vol. 22, pp. 290-292, 2006
- [17] M. Kraljević, and J. Dulčić, "Age and growth of gilt-head sea bream (*Sparus aurata* L.) in the Mirna estuary, northern Adriatic," *Fish. Res.*, vol. 31, no. 3, pp. 49-255, 1997
- [18] J. Dulčić, and B. Glamuzina, "Length-weight relationships for selected fish species from three eastern Adriatic estuarine systems (Croatia)," *J. Appl. Ichthyol.*, vol. 22, pp. 254-256, 2006
- [19] J. Dulčić, and M. Kraljević, "Weight-length relationship for 40 fish species in the eastern Adriatic (Croatian waters)," *Fish. Res.*, vol. 28, no. 3, pp. 243-251, 1996
- [20] E. T. Koutrakis, and A. C. Tsikliras, "Length-weight relationships of fishes from three northern Aegean estuarine systems (Greece)," *J. Appl. Ichthyol.*, vol. 19, pp. 258-260, 2003
- [21] S. Deguara, K. Jauncey, J. Feord, and J. López J. "Growth and feed utilization of gilthead sea bream, *Sparus aurata*, fed diets with supplementary enzymes," in J. Brufau, and A. Tacon, (eds.) Feed manufacturing in the Mediterranean region: Recent advances in research and technology Zaragoza: CIHEAM Cahiers Options Méditerranéennes, no. 37, pp. 195-215, 1999
- [22] C. Faggio, G. Piccione, S. Marafioti, F. Arfuso, G. Fortino, and F. Fazio, "Metabolic Response to Monthly Variations of *Sparus aurata* Reared in Mediterranean On-Shore Tanks," *Turkish Journal of Fisheries and Aquatic Sciences*, vol. 14, pp. 567-574, 2014
- [23] D. Erguden, and C. Turan, "Growth properties of Sea bass (*Dicentrarchus labrax* (L., 1758), Perciformes: Moronidae) live in Iskenderun Bay," *Pakistan Journal of Biological Science*, vol. 8, no. 11, pp. 1584-1587, 2005

# Valve movements in bivalves as a behavioral biomarker

Marko Nikolić, Slađana Gvozdenović, BIO-ICT, Faculty of Electrical Engineering, Podgorica, Montenegro

Sergey Kholodkevich, Tatiana Kuznetsova, Anton Kurakin, SRCES RAS, Sankt Petersburg, Russia

**Abstract** — Bio-monitoring based on behavioral parameters has been used for a long time. In combination with ecological monitoring it can give a complete picture of ecosystem state e.g. marine ecosystem. Behavior biomarkers, such as valve movements, have proved to be very useful in environmental and aquatic pollution monitoring. The most used organisms in marine ecosystem monitoring are mussels. The measurements of mussel valve movements (valvometry) using remote sensing technologies were first implemented during 1989 as a warning tool for aquatic pollution. Since that time this method has been in use in biomonitoring studies. But, most importantly, it represent a non-invasive method for investigated organisms. To date, a lot of different studies on bivalves have been done based on this behavioral biomarker and they showed that it is a very useful tool in bio-monitoring studies.

**Key words** — bio-monitoring, mussels, valve movement

## I. INTRODUCTION

Today, biological sensors became a very important tool in monitoring the quality of aquatic ecosystems, and bivalves present one of the most widely used organisms in bio-monitoring programs because of their cosmopolitan distribution and ability to accumulate various pollutants, primarily heavy metals and bio-toxins. There are 75 different bivalve species, marine and freshwater, which are in high use in bio-monitoring studies [1]. The most used are: *Dreissena polymorpha*, *Anodonta cygnea*, *Unio pictorum*, *Mytilus edulis*, *M. galloprovincialis*, *M. californicus*, *Crassostrea gigas*, *Ostrea edulis*. Not only bivalves, but the entire mollusk phylum is included in bio-monitoring studies today. Mollusks are one of the most varied animal phyla and a very important component of

aquatic ecosystems. Because of their abundance and presence in different ecosystem types, they have a very important ecological role [2].

A lot of characteristics make mollusks good indicator organisms:

- Abundant and widely distributed, inhabit all ecosystem types, which can lead to the possibility of research in wide geographical areas. Such species are Mediterranean mussel (*Mytilus galloprovincialis*) and blue mussel (*Mytilus edulis*);
- Key links in ecosystem and because of that, any negative effects of pollutants in mussels will be also reflected on the whole ecosystem;
- Most of Gastropods and Bivalves are sessile or low mobility organisms, so pollutant concentrations in their tissues represent a realistic image of the environment;
- They have different reproductive abilities as well as life cycles, mainly with long life cycles especially marine bivalves. According to [3] more than 40% of marine bivalves have life cycles longer than 14 years;
- Mollusks are mainly large animals, easy for manipulation and keeping in aquariums;
- Pollutants from environment can be absorbed not only through gastrointestinal system in mollusks but also through integument and respiratory system;
- Compared with other invertebrates and vertebrates, mollusks have a lower ability of pollutant metabolism and excretion;
- Most are highly sensitive to pollutants, and a lot of them are endangered;
- Morphology, anatomy and physiology of mollusks are well known.

## II. VALVE GAPE AS A BEHAVIORAL BIOMARKER IN BIVALVES

Bivalves are protected by shell from external impacts, and every bivalve has its own rhythm of valve movements, forming a weak low-and-high tidal [4], [5], [6] and/or circadian rhythm [7], [8]. Because of food intake as well as oxygen, mussels keep their shells open more than 70% of the time [9]. Wide open and motionless valves are characteristic of dead animals. Sometimes their shells are

Marko Nikolić, BIO-ICT, Faculty of Electrical Engineering, University of Montenegro, Džordža Vašingtona bb, 81000 Podgorica, Montenegro (phone: 382-67-530582; e-mail: markon@ac.me).

Slađana Gvozdenović, BIO-ICT, Faculty of Electrical Engineering, University of Montenegro, Džordža Vašingtona bb, 81000 Podgorica, Montenegro (phone: 382-69-286485; e-mail: sladjanag@ac.me).

Sergey Kholodkevich, Scientific-Research Centre for Ecological Safety RAS (SRCES RAS) 18, Korpusnaya Street, Sankt Petersburg, Russia (e-mail: kholodkevich@mail.ru).

Tatiana Kuznetsova, Scientific-Research Centre for Ecological Safety RAS (SRCES RAS) 18, Korpusnaya Street, Sankt Petersburg, Russia (e-mail: kuznetsova\_tv@bk.ru).

Anton Kurakin, Scientific-Research Centre for Ecological Safety RAS (SRCES RAS) 18, Korpusnaya Street, Sankt Petersburg, Russia (e-mail: rebel55@mail.ru).

closed but in particular the closing of shells for a long time is behavioral reaction to stress conditions [10], [11], [12], so the behavioral response of mussels to toxicants may be measured by valve gape monitoring, [13], [14]. Monitoring and the whole system is based on the fact that if a group of bivalve mollusks keeps their shells closed for a long time, the system (equipment for valve movements registration) generates the alarm signal which indicates that some of environment parameters is not in “well range” [15], [16].

The measurements of mussel valve movements (valvometry) using remote sensing technologies were first implemented by [16] as a tool for early warning alarms for aquatic pollution. Electromagnetic sensors, which are attached to both sides of the mussels, monitor gaping behavior. Various types of movements are recorded in response to contaminants, including the closure of shell valves for a long period to avoid contaminants, decrease in the average distance between valves over a certain timeframe, increased activity (e.g., rapid opening and a higher closing frequency than that observed under normal conditions), and gaping, which indicates that the mussels are no longer alive (no further movement and 100% opening of valves) [17].

Valve movement in bivalves, especially in mussels, has been used in a lot of different studies. The day-night rhythm in shell movement in *Anodonta anatine* and *Unio timudis* was studied by [18]. Both species showed a diurnal rhythm of valve gaping, and the state of openness of the shells was highest during night and lowest in the morning. Detection of noxious dinoflagellate *Heterocapsa circularisquama* by valve movement monitoring in *Pinctada fucata* was studied by [19]. In filtered sea water, valve movement activity of pearl oysters, as displayed on the output screen of the detection system, showed sporadic thorn-like spikes, indicating the shell closing for a short time, with a flat waveform indicating that the shell was open. The frequency of these spikes increased as the number of *H. circularisquama* cells in the filtered seawater increased. However, the frequency of the spikes almost immediately decreased when the water was changed to filtered seawater with no *H. circularisquama*.

Effluents of phosphate industry and valve movement response in *Mytilus galloprovincialis* to different heavy metals was studied by [20]. In a 1 h exposure experiments, all heavy metals induced a decrease in the time of normal opening and the appearance of sequences of stress behavior, including enhanced valve adductions and complete closure at high concentrations.

Investigation of day-night behavior in *Pinna nobilis* based on valve movement was implemented by [21]. The records have shown that far from being all day open, as previously suggested, *P. nobilis* follows marked circadian and circalunar rhythms of gaping activity. Individuals usually close the valves during night, sometimes for more than 12 h. The repetition of patterns observed made it possible to distinguish between two main behaviors: (1) night-closing related to sunset and (2) night-opening related to the Moon visible in the sky with the disc illuminated more than 50%. Another two less common trends were also observed: (3) day-closing and (4) night-opening when no Moon was visible.

Valve movement response of *Corbicula fluminea* following exposure to waterborne arsenic was investigated by [22]. Their findings revealed that *C. fluminea* experiences a valve opening in the absence of arsenic predominantly in the morning hours (03:00–08:00) with a mean daily opening/closing period of 21.32 (95% CI: 20.58–22.05) h. Amplification of daily activity occurred in the presence of arsenic. Behavioral toxicity assays revealed arsenic detection thresholds of 0.60 (95% CI: 0.53–0.66) mg l<sup>-1</sup> and 0.35 (95% CI: 0.30–0.40) mg l<sup>-1</sup> for response times of 60 and 300 min.

Impact of pesticides on valve movements in *Mytilus galloprovincialis* was studied by [23]. The results showed that insecticide (cypermethrin) impaired the valve activity in a concentration-dependent manner, leading to a reduction in the time of normal opening. The lowest effect concentration was determined as 100 µg/l cypermethrin. Prolonged closures of mussels were also observed for the exposures to 400 and 800 µg/l.

The valve movement response of three freshwater mussels *Corbicula fluminea*, *Hyriopsis cumingii* and *Anodonta woodiana* exposed to copper showed the mean amplitude of valve opening was significantly decreased in *C. fluminea* when exposed to 0.1 and 1 mg/l, and decreased for *H. cumingii* and *A. woodiana* after exposure to 10 mg/l [24].

### III. EQUIPMENT FOR VALVE GAPE MEASUREMENTS

Equipment for monitor the opening and closing of the shell consists of 16 motion sensors that rely on shell (Figure 1). Magnetic Hall effect sensors are used to determine the distance between the valves of bivalve mollusks. The sensors are connected via cable to the microprocessor in which the transformed data obtained and transmitted to the computer (Figure 2). On the computer to read graphics with time and the degree of openness or closure shells (Figure 3). The module configuration includes:

- Measurement module (FM) including: frame device (FD) equipped with 16 valve gape sensors (GS), water temperature sensor (TS), connect cables (CC1).
- Microprocessor module (MM) including: electronic controller unit (EC) for primary measurements, processing and data transmission via RS232 port, connecting cable (CC2).
- Power-and-data cable (PDC) with AC adapter and standard RS-232 9-pin connector.
- Software (S).
- Micrometric device (MD) to calibrate valve gape sensors [25].

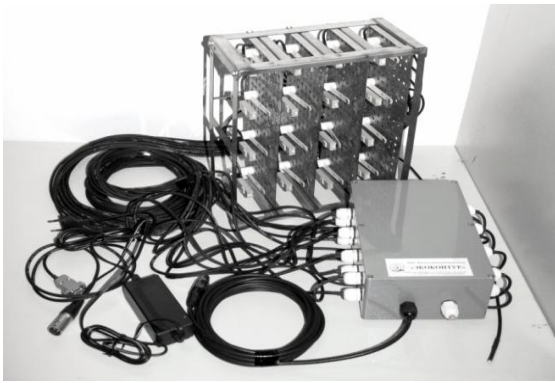


Fig. 1. 16- Channel module for mollusks valve gap measurements

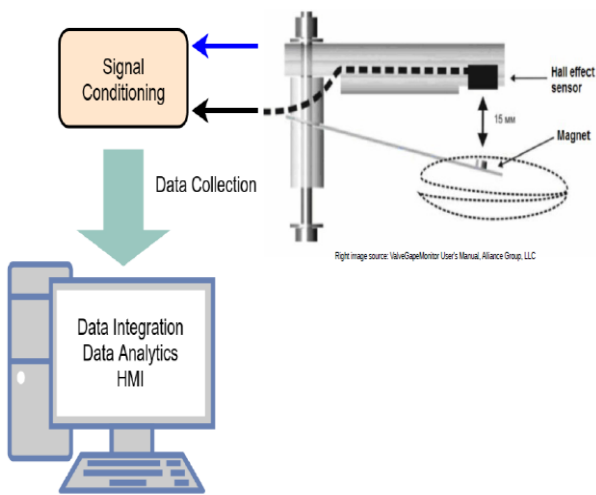


Fig. 2. Connection of sensor with computer through microprocessor

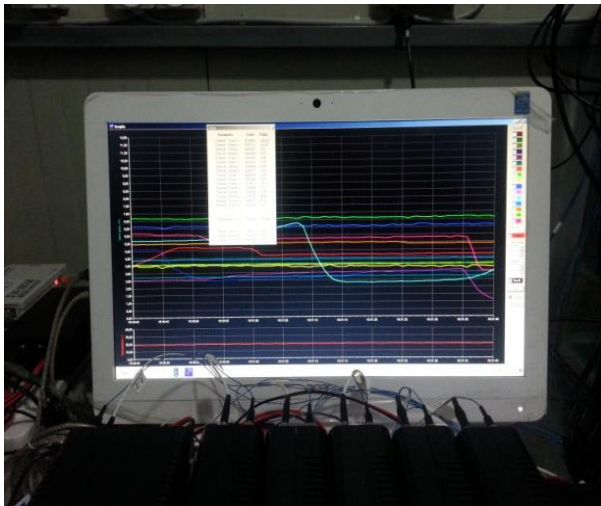


Fig. 3. Numerical and graphical visualization of the measured data

The module provides aquatic (seawater and freshwater bodies) environmental assessment in real-time automated mode based on the bivalve mollusks responses to the water contamination. It can be used both for laboratory research and for ensuring environmental safety of water bodies by the early detection of toxic pollution as a result of natural or technogenic disasters. The module sensors measure

behavioral responses of the bivalve mollusks and hydro-physical parameters followed by a software-based processing of the obtained information (Figure 4). The module can be used for measurements in free running mode or telemetry mode featuring the following functions:

- automatic measurement of water temperature;
- automatic measurement of the valve gapes of bivalve mollusks;
- initial data processing in real time;
- numerical and graphical visualization of the measured data;
- initial and processed data storage;
- data transmission [25].



Fig. 4. The measurement of mussels valve movements

#### IV. CONCLUSIONS

Bio-monitoring studies based on behavior biomarkers are very important and useful in assessing the state of the environment and as an early warning system for presence of potential pollutants which can lead to preventive actions and preventing potential accidents.

#### ACKNOWLEDGMENT

This work has been supported by the Ministry of Science of Montenegro and HERIC project through the BIO-ICT Centre of Excellence (Contract No. 01-1001).

#### REFERENCES

- [1] T. P. O'connor, A. Y. Cantillo, and G. G. Lauenstein, "Monitoring of temporal trends in chemical contamination by NOAA National Status and Trends Mussel Project," In: K. J. M. (Ed.): Biomonitoring of coastal waters and estuaries," CRC Press, Boca Raton, 1994, pp. 29-50.
- [2] J. Oehlmann, and U. Schulte-Oehlmann, "Molluscs as bioindicators, Chapter 17," In: B. A. Markert *et al.* (Ed.): "Bioindicators and Biomonitoring," Elsevier Science, Amsterdam, 2003, pp. 577-635.
- [3] J. Heller, "Longevity in mollusks," *Malac.*, vol. 31, pp. 259-295, 1990.
- [4] W. S. Kim, H. T. Huh, J. H. Lee, "Endogenous Circatidal Rhythm in the Manila Clam, *Ruditapes philippinarum* (Bivalvia: Veneridae)," *Mar. Biol.*, vol. 134, no. 1, pp. 107-112, 1999.
- [5] R. W. Langton, and P. A. Gabbott, "The Tidal Rhythm of Extracellular Digestion and the Response to Feeding in *Ostrea edulis*," *Mar. Biol.*, vol. 24, no. 2, pp. 181-187, 1974.

- [6] K. P. Rao, "Tidal rhythmicity of rate of water propulsion in *Mytilus*, and its modifiability by transplantation, *Biol. Bull.*, vol. 106, pp. 353–359, 1954.
- [7] V. F. Gnyubkin, "Dynamics of the mobility of the Black Sea Mussel, *Monitoring prirodnykh i tekhnogenykh sred (Monitoring of natural and technogenic media): Materialy vseukrainskoy nauchn. konferentsii (Materials of All-Ukrainian Scientific Conference)*," Simferopol: DIAJPI, 2008, pp. 159–162.
- [8] L. N. Slatina, "Circadian respiration rhythms in a mussel of the Black Sea," *Biologiya morya*, no. 1, pp. 63–71, 1991.
- [9] K. J. M. Kramer, and E. M. Foekema, "The "Musselmonitor" as biological early warning system," In: F. M. Butterworth, A. Gunatilaka, and M. E. Gensebatt (Ed.): "Biomonitoring and biomarkers as indicators of environmental change 2: A handbook," New York: Springer, 2001, pp. 59–87.
- [10] D. de Zwart, K. J. M. Kramer, and H. A. Jenner, "Practical experiences with the biological early warning system musselmonitor," *Environ. Toxicol. Water.*, vol. 10, pp. 237–47, 1995.
- [11] J. Borchering, "Ten years of practical experience with the Dreissena-Monitor, a biological early warning system for continuous water quality monitoring," *Hydrobiologia*, vol. 556, pp. 417–26, 2006.
- [12] V. F. Gnyubkin, "An early warning system for aquatic environment state monitoring based on an analysis of mussel valve movement," *Russ. J. Mar. Biol.*, vol. 35, pp. 431–436, 2009.
- [13] A. R. Manley, and J. Davenport, "Behavioural responses of some marine bivalves to heightened seawater copper concentrations," *Bull. Environ. Contam. Toxicol.*, vol. 22, pp. 739–44, 1979.
- [14] W. Slooff, D. de Zwart, and J. M. Marquenie, "Detection limits of a biological monitoring system for chemical water pollution based on mussel activity," *Bull. Environ. Contam. Toxicol.*, vol. 30, pp. 400–5, 1983.
- [15] A. N. Krainyukova, A. V. Ryazanov, and V. V. Emeljanenko, "Method of biotesting by the reaction of closure of valves of bivalve Mollusks, *Metody biotestirovaniya vod (Methods of Biotesting of Water) Chernogolovk*," Institute of Chemical Physics of AN SSSR, 1988, pp. 57–60.
- [16] K. J. M. Kramer, H. A. Jenner, and D. de Zwart, "The valve movement response of mussels: a tool in biological monitoring," *Hydrobiologia*, vol. 189, pp. 433–43, 1989.
- [17] K. J. M. Kramer, and E. M. Foekema, "The Musselmonitor as biological early warning system," In: F. M. Butterworth, A. Gunatilaka, and M. E. Gensebatt (Ed.): "Biomonitoring and biomarkers as indicators of environmental change 2: a handbook," Springer, New York, 2001, pp. 59–87.
- [18] E. Villi, and H. Mikko, "Valve movement of Anodonta anatina and Unio timidis (Bivalvia, Unionidae) in a eutrophic lake," *Ann. Zool. Fennici*, vol. 33, pp. 257–261, 1994.
- [19] N. Kiyohito, H. Tsuneo, G. Jyoji, Y. Hiroyasu, and O. J. Seok, "Detection of the shellfish killer *Heterocapsa circularisquama* (Dinophyceae) by measuring bivalve valve activity with a Hall element sensor," *Aquaculture*, vol. 255, pp. 395–401, 2006.
- [20] A. F. Mohamed, M. Abdelfattah, O. Abdelkader, B. Ali, M. Abdelmalek, and J. C. Pihan, "Valve movement response of the mussel *Mytilus galloprovincialis* to metals (Cu, Hg, Cd and Zn) and phosphate industry effluents from Moroccan Atlantic coast," *Ecotoxicology*, vol. 15, pp. 477–486, 2006.
- [21] J. R. G. March, M. A. S. Solsona, and A. M. G. Carrascosa, "Shell gaping behaviour of *Pinna nobilis* L., 1758: circadian and circalunar rhythms revealed by in situ monitoring," *Mar. Biol.*, vol. 15, pp. 689–698, 2008.
- [22] C. M. Liao, S. F. Jau, C. M. Lin, L. J. Jou, C. W. Liu, V. H. C. Liao, and F. J. Chang, "Valve movement response of the freshwater clam *Corbicula fluminea* following exposure to waterborne arsenic," *Ecotoxicology*, vol. 18, pp. 567–576, 2009.
- [23] M. A. Ayad, M. A. Fdil, and A. Mouabab, "Effects of cypermethrin (Pyrethroid Insecticide) on the valve activity behavior, byssal thread formation, and survival in air of the marine mussel *Mytilus galloprovincialis*" *Arch. Environ. Contam. Toxicol.*, vol. 60, pp. 462–470, 2011.
- [24] H. Liu, X. Chen, I. J. Kang, Y. Shimasaki, Y. Oshima, and J. Yang, "The valve movement response of three freshwater mussels *Corbicula fluminea* Muller 1774, *Hyriopsis cumingii* Lea 1852, and *Anodonta woodiana* Lea 1834 exposed to copper," *Hydrobiologia*, 2015.
- [25] ALLIANCE Group LLC, "16 - Channel Module for Mollusks' Valve Gape Measurements, Valve Gape Monitor," Moscow, 2015.

# Real-Time Measurements Analysis of Seawater Physical Parameters with SEMaR Buoy System

Nikola Bulatović, Nikola Žarić, *Member, IEEE*, Ana Pešić, Mirko Đurović, Trifun Savić, Aleksandar Joksimović

**Abstract** — Seawater physical and chemical parameters represent the base of hydrographical and biological investigation of any marine environment. These parameters influence and determine distribution and biology of all living beings in the seas and oceans, which have a great impact on human health too. In order to provide continuous monitoring and measurements of seawater physical parameters we developed a modular Smart Environmental Monitoring and Reporting system. In this paper we will present the analysis of the on-field measurements' results collected during three months operating period of the SEMaR system at COGImar fish farm.

**Keywords** — Internet of Things, real-time measurements, seawater physical parameters, smart buoy.

## I. INTRODUCTION

**W**ATER is the foundation of life. The clean water resource is essential for quality food production and keeping all the living beings and environment healthy. What we are currently facing as a mankind is the challenge to efficiently conserve, manage, and distribute water we have. Clean and safe water is scarce in the developing world. Today, nearly 1 billion people don't have access to it [1]. The water crisis is the number one global risk based on impact to society (as a measure of devastation) [2]. This is why a significant attention of researchers and community has been recently devoted to the analysis and permanent monitoring of physical and

This work has been supported by the Ministry of Science of Montenegro and the HERIC project through the BIO-ICT Centre of Excellence (Contract No. 01-1001).

Corresponding Nikola Bulatović is with the Faculty of Electrical Engineering, University of Montenegro, Džordža Vašingtona bb, 81000 Podgorica, Montenegro (phone: 382-67-418264; e-mail: nbulatovic@ac.me).

Nikola Žarić is with the Faculty of Electrical Engineering, University of Montenegro, Džordža Vašingtona bb, 81000 Podgorica, Montenegro (e-mail: zari@ac.me).

Trifun Savić is with the Faculty of Electrical Engineering, University of Montenegro, Džordža Vašingtona bb, 81000 Podgorica, Montenegro (e-mail: trifuns@ac.me).

Ana Pešić is with the Institute of Marine Biology, University of Montenegro, Dobrota bb, 85331 Kotor, Montenegro (e-mail: pesica@ac.me).

Mirko Đurović is with the Institute of Marine Biology, University of Montenegro, Dobrota bb, 85331 Kotor, Montenegro (e-mail: mirko.djur@ac.me).

Aleksandar Joksimović is with the Institute of Marine Biology, University of Montenegro, Dobrota bb, 85331 Kotor, Montenegro (e-mail: acjo@ac.me).

chemical parameters, related to clean water maintenance and mariculture and agriculture food production.

Seawater physical and chemical parameters represent the base for every investigation of marine environment, hydrographical and biological. Parameters such as temperature and salinity influence horizontal and vertical distribution of water masses, density of water, distribution of dissolved particles, etc. while together with pH and oxygen influence and determine distribution and biology of all living beings in the seas and oceans, which has an important impact on human health too.

New ICT solutions, integrating the core concepts of Internet of Things (IoT) [3]-[5], could be the solution for efficient and continuous environmental biomonitoring. One such platform is developed within the BIO-ICT Center of Excellence in Montenegro. The purpose of the platform is to enable accurate and reliable real-time monitoring of numerous environmental parameters. After the measured data are stored in data center, it is used for making analytics, presentation, visualization and data mining purposes. The goal is to make a large global network of interconnected environmental IoT platforms which will include the processing of unprecedented amount of data. As a result of novel and promising data mining techniques, the processed data is used for finding the hidden patterns which could enable us to improve the clean water management and conservation.

## II. ABOUT BIO-MONITORING

In the frame of various national and international projects and monitoring programs, Institute of marine biology, University of Montenegro (UoM-IBMK), conducts marine environment monitoring. The most important parameters of each ecological monitoring are temperature, salinity, pH and oxygen. For example, water quality and fishing sea biomonitoring is regularly conducted for the purpose of public health and mariculture protection. This monitoring is currently partially aligned with EU regulations Directive 2006/113/EC, Regulations (EC) No 853/2004 and (EC) No 854/2004 and in the next period will be fully harmonized. Every 15 days, in the period from April to October, as it is prescribed in a Law on water (Official Gazette of Montenegro 27/07) and a Law on environmental protection (Official Gazette of Montenegro

48/08, 40/10, 40/11), UoM-IBMK conducts continuous monitoring of public beaches in Montenegro (more than 100 beaches) in order to establish water sanitary quality and to prevent any negative impact on human health [6]-[10].

Thus, for the purposes of monitoring schemes of different projects, number of samplings for determination of physical parameters can reach hundreds during one month. Even with this high number of samples, there are gaps in data series. Data are scattered and collected from separate locations, different depths, specific time of day or week, etc. The costs of continuous biomonitoring, in financial and time consumption terms, are high. Financial costs include equipment and chemical solutions needed for laboratory analysis of environmental parameters, purchase and maintenance of sensor probes and travel expenses for repetitive visits to the sampling sites. Excessive time consumption is expressed in visiting all the sampling sites and delayed parameters measuring and evaluation in the laboratory. Inaccuracy of “off-site” measured parameters caused by delay of transportation and scheduling of analysis can be a significant problem.

In order to improve the accuracy and reliability of water physical parameters datasets, there is a need for continuous real-time monitoring with novel ICT solutions deployed into the monitoring process.

#### A. Related works

Currently there are numerous solutions for real-time water quality monitoring. These solutions are usually smart buoy-based platforms that collect information from various measuring probes and transfer them (through GSM/GPRS communication) to remote databases and data centers for further processing and research [11]. In some areas of the world, this kind of platforms are used for measuring marine parameters such as temperature and illuminance, as a very indicative parameter of red tide phenomenon [12], for oil spill early detection in order to avoid water pollution [13], or for wave height measurement based on accelerometer data [14]. They are also used for coastal marine monitoring as a resource for touristic sector [15]. In order to process huge amount of measured data, new communication algorithms are constantly emerging [16].

The existing IoT solutions for maritime parameters measuring have some disadvantages that should be pointed out: high cost of platforms and sensor probes; platforms are sealed in the waterproof box and cannot be used for any kind of software or hardware editing; problematic integration of different sensor probes with the processing unit. The main idea was to design the system that would outperform the existing solutions in terms of cost, ease of use and integration of diverse sensor probes.

### III. SEMAR PLATFORM

SEMaR system represents an example how the new ICT solutions, using the technological concepts of Internet of Things, can improve ecological monitoring of seawater parameters worldwide. The IoT steps, incorporated into the system, include: data measuring, data transfer, data

processing, data analysis and presentation, and finally decision-making upon the processed data. The aim of the system is to enable automated and online ecological monitoring and processing of seawater parameters, avoiding direct human involvement in water sampling and later manual processing of measured parameters.

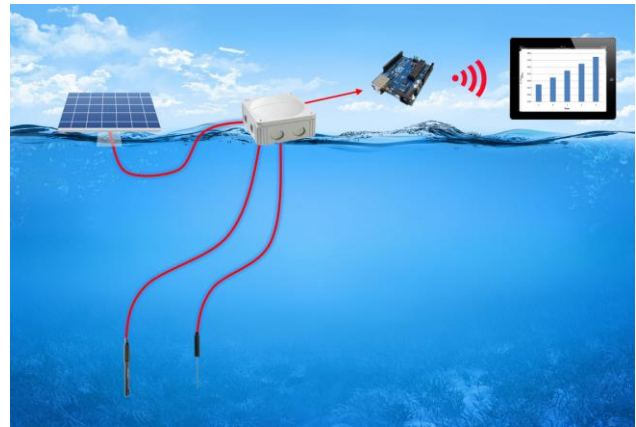


Fig. 1. SEMaR system schematic overview.

The core architecture of the system is comprised of three important structural parts: the sensing unit (sensor probes) for measuring the water and air bio-parameters, the processing and communication unit with the microcontroller platform and the cloud-based data visualization platform (Fig. 1).

The sensing unit of the platform consists of the sensor probes used in seawater ecological monitoring. The parameters of interest are water temperature, pH level, dissolved oxygen level and salinity. These are the most indicative parameters of water pollution and can be used as an early alarm. Two water temperature probes (at different depths of 1 and 2 meters) and pH level low-consumption probe are integrated in the prototype of SEMaR system in the testing phase (Fig. 2). Air conditions can also influence the water parameters so the air temperature and humidity sensor probes are also integrated.



Fig. 2. SEMaR system prototype.

The processing unit consists of the microcontroller and communication module. The Arduino microcontroller platform is the central part of the system. Its main purpose is to collect the sensor probes' measured data, process it and send it to communication module. The



communication SIM900 module is in charge of processed data wireless transmission into the data center. The Real Time Clock (RTC) module integration enables the system to easily configure the time intervals for data measuring and transferring. The favorable light conditions on the sea level were important in choosing the solar power unit. It makes the power independent and “green energy” solution.

The cloud based data visualization application BlueLeaf is developed for the purposes of data storage, presentation, detailed analytics and decision-making [17].

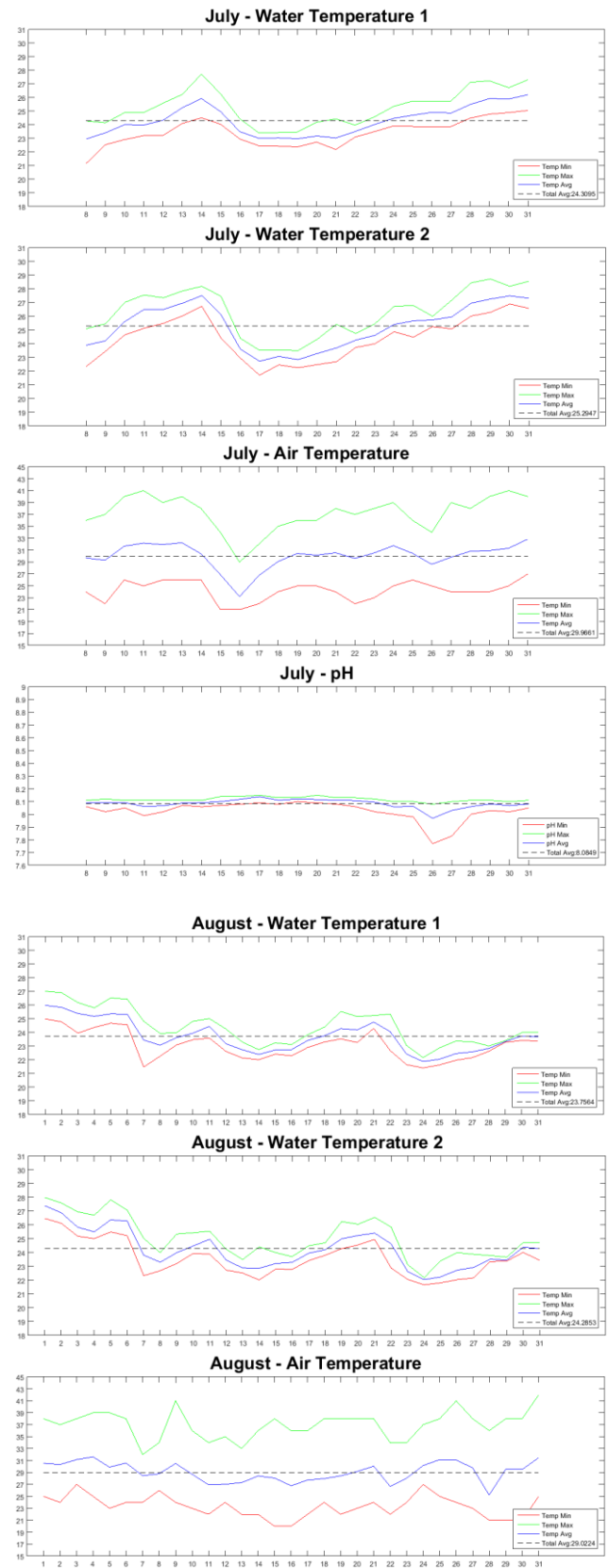
#### IV. DATA ANALYSIS

Seawater temperature, air temperature and pH level were monitored at COGImar fish farm during summer months of 2016 (Fig. 3). Seawater temperatures data are collected from depths of 1 and 2 meters. Temperature measuring at different depths, stratum, provides us with useful information on stratification of water column, depth of thermocline, which is influencing distribution of marine organisms (mainly phytoplankton and zooplankton) and development of early stages of marine organisms (eggs and larvae). Thermal stratification of water column also distinguishes movement of water masses and currents.

Seawater temperature at 1 and 2 meter depths shows same monthly trend, since the thermocline during summer months is on a deeper stratum, and upper parts of water column are well mixed and with uniform temperature. In July and August seawater temperature is stable and varies between 21°C and 28°C, while in the end of September temperature is dropping to 18°C. This decrease of temperature corresponds to start of rainy period, when high amount of fresh water inflows to marine ecosystem through precipitation and underwater springs which are numerous in the area of COGImar fish farm. Daily variations and night/day temperature variations are insignificant, ranging between 1°C and 3°C, although air temperature are very high, especially during July and August when it goes over 40°C during day. In September daily variations of seawater temperatures are higher, almost 10°C, since the variations of air temperature are also high. Obtained air temperature is higher than seawater temperature during entire examined period, which is expected. Sea is a stable system and characteristic of seawater is that its temperature increases and decreases very slowly due to very high specific heat capacity. This water characteristic and its thermal stability play important role for all living beings in marine ecosystem providing them stable living environment.

Seawater pH level is typically limited to a range between 7.5 and 8.4, averaging about 8.2, which makes oceans slightly basic [18]. During 3 months period of monitoring in Bokakotorska Bay, at COGImar fish farm, pH values are constant and stable around value of 8.1. The variations in July and August are ranging from 7.8 and 8.2, which totally correspond to the normal limits of pH values in the sea. Small variations of pH value, during

these two months, can be explained by high precipitation in some days and summer showers.



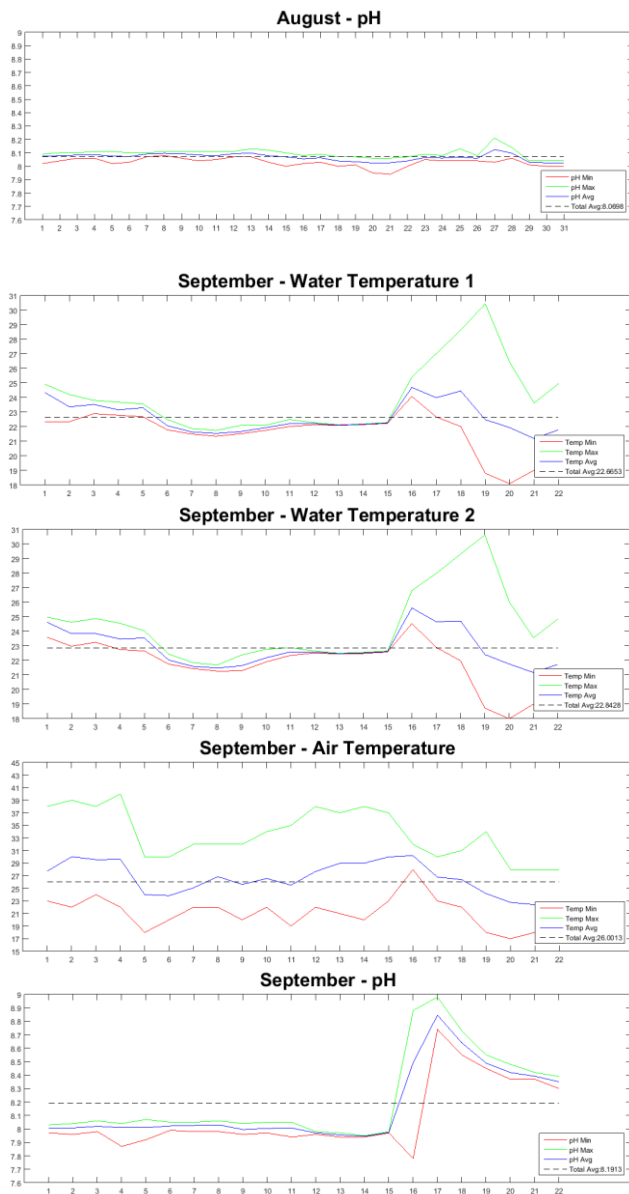


Fig. 3. Water temperature at depths of 1 and 2 meters, air temperature and pH level data for July, August and September 2016.

We may notice that there are certain discontinuities and somewhat higher variations in the monitored parameters in the mid of September which corresponds to heavy rainy days and big floods. Therefore, these variations are not a consequence of any pollution, but natural response to the high inflow of fresh water. Finally, from the presented analysis one may conclude that the sea ecosystem in Bokakotorska Bay is quite stable and that variations of the monitored parameters are in the expected limits.

## V. CONCLUSION

The system presented in this paper represents an efficient solution for permanent ecological monitoring of seawater parameters. It will significantly contribute to the efforts to provide healthier environment, better food quality and safety, and healthier life of people in Montenegro and worldwide. Data analytics, conducted

during summer months of 2016, showed some real trends and variations of seawater and air parameters that were mainly a consequence of weather conditions and changes. Future plans in the SEMaR system development include small scale deployment on several Montenegrin coastal areas which will produce larger sets of data to be processed and analyzed. This will also comprise further development of an alarm system in cases of possible water pollution which will make the platform more oriented as a support in decision making.

## REFERENCES

- [1] World Health Organization and UNICEF Joint Monitoring Programme (JMP). (2015)
- [2] World Economic Forum. (2015)
- [3] D. Uckelmann, M. Harrison, and F. Michahelles, "An Architectural Approach Towards the Future Internet of Things", Springer Berlin Heidelberg, March 2011.
- [4] L. Yan, Y. Zhang, L. T. Yang, and H. Ning, "The Internet of Things: From RFID to the Next-Generation Pervasive Networked Systems", CRC Press, March 2008.
- [5] B.B.P. Rao, P. Saluia, N. Sharma, A. Mittal, S.V. Sharma, "Cloud computing for Internet of Things & sensing based applications," *6<sup>th</sup> International Conference on Sensing Technology (ICST)*, pp.374-380, 2012
- [6] D. Regner, N. Vuksanović, M. Dutina, D. Joksimović, and B. Stjepčević, "Monitoring of the sea water quality in the coastal Montenegrin area", *Proceedings 28th Conference "WATER POLLUTION CONTROL 1999"*, Soko Banja, Serbia, pp: 319-325, 1999.
- [7] D. Regner, N. Vuksanović, B. Stjepčević, and D. Joksimović, "Sea water quality of the Montenegrin Coastal sea throughout bathing season" *In 33th Conference "WATER 2004"*, Serbia, pp: 409-414, 2004.
- [8] S. Krivokapić, B. Pestorić, S. Bosak, G. Kušpilić, and C. Wexels Riser, "Trophic state of Bokakotorska Bay - south-eastern adriatic sea," *Fresenius Environmental Bulletin*, vol 20, no 8, pp:1960-1969, 2011.
- [9] Directive 2008/56/ec Establishing a framework for community action in the field of marine environmental policy (marine strategy framework directive)
- [10] Z. Kljajić, S. Marković, J. Žmukić, M. Mitrić, A. Kasteli, B. Pestorić, and D. Drakulović, "Montenegrin coastal sea monitoring", *Proceedings of 41st Conference WATER 2012*, Divčibare, Serbia, pp:207-214, 2012.
- [11] A.H.Mohd Akmal Helmi, M.M. Hafiz, M.S.B. Shah Rizam, "Mobile buoy for real time monitoring and assessment of water quality," *IEEE Conference on In Systems, Process and Control (ICSPC)*, 2014, pp.19-23.
- [12] S. Hiroshi, H. Miyake, H. Tokimune, A. Hamasaki, and Y. Matsuyama. "A study of the Red-tide Monitoring System Using Drifting Buoy and Wireless Networks." *In the 25<sup>th</sup> International Offshore and Polar Engineering Conference*, 2015.
- [13] M. Davide, G. Pieri, O. Salvetti, M. Tampucci, C. Domenici, and A. Tonacci. "Sensorized buoy for oil spill early detection." *In proc of IEEE OCEANS 2015*, Genova, Italy, pp. 1-5., 2015.
- [14] D.-K. Chung, M.-J. Lim, Y.-M. Kwon, and J.-S. Park. "A Study on an Improved Wave Height Measurement Algorithm using Accelerometer in the Buoy." *Information*, vol. 18, no. 4, 2015. pp: 1331-1336.
- [15] C. Albaladejo Pérez, H. Navarro Hellin, F. Soto Valles, A. Iborra García, and R. Torres. "Low cost buoy for monitoring recreational areas." *Instrumentation Viewpoint* 18, 2015, pp: 61-61.
- [16] G. Gediminas, D. Drungilas, A. Andziulis, D. Dzemydiene, M. Voznak, M. Kurmis, and S. Jakovlev, "Advanced approach of multiagent based buoy communication." *The Scientific World Journal*, 2015.
- [17] BlueLeaf cloud based web application. Available: <http://www.blueleaf.ac.me>
- [18] R. Chester, and T. Jickells, "Marine Geochemistry" *Blackwell Publishing*, 2012.

# Anomaly Detection and Preprocessing of soil database of Montenegro

Elma Hot, *Faculty of Electrical Engineering*, Balša Bajagić, *Biotechnical Faculty*

**Abstract** — A preprocessing of soil database of Montenegro is done and presented in this paper. Anomaly detection and cleaning the data is one of the first steps of data mining analysis. Hence, preprocessing of database is done using pedological knowledge and different approach of anomaly detection based on clustering. K-means clustering method is used. Results of simulation will be presented and analyzed.

**Keywords** — Data mining, preprocessing, anomaly detection, soil database, clustering, K-means

## I. INTRODUCTION

A data mining (DM) represents a set of techniques which have a purpose of discovering useful information in the big data. Discovery of information is generally performed by finding and recognizing patterns as well as establishing a relationship between data. The main focus of data mining is to find the most important information in the big data [1].

Data cleaning is the first step of data mining. It is the process of detecting and correcting mistakes, anomalies, and identifying incomplete, incorrect or irrelevant parts of the data and then replacing, modifying, or deleting [2].

Anomaly detection (AD) is also known as outlier detection, it aims to detect items or events which have unexpected value or they are different than an expected pattern or other items in a dataset [3]. It has application in many research areas like: fraud detection, activity monitoring, network performance, fault diagnosis, structural defect detection, detecting novelties in images, motion segmentation, time-series monitoring, medical condition monitoring, pharmaceutical research.

Three categories of anomaly detection techniques exist [4]:

- Unsupervised AD;
- Semi-supervised AD;
- Supervised AD.

This work has been supported by the Ministry of Science of Montenegro and the HERIC project through the BIO-ICT Centre of Excellence (Contract No. 01-1001).

Elma Hot (corresponding author), Faculty of Electrical Engineering, University of Montenegro, Džordža Vašingtona bb, 81000 Podgorica, Montenegro (phone number: 382-67-677525; e mail: elma@ac.me).

Balša Bajagić, Biotechnical Faculty, University of Montenegro, Džordža Vašingtona bb, 81000 Podgorica, Montenegro; (e mail: balsab@ac.me)

Cluster-based detection of outlier is used in simulations. One approach of AD using clustering is detecting a significantly smaller cluster, which may present cluster of outliers.

Second approach of AD using clustering is analyzing distance between data samples and centroid of its cluster after KM clusterization. Based on the distance some of data samples are declared as outliers [5].

The aim of this paper is preprocessing of soil database of Montenegro using unsupervised anomaly detection and removing the mistakes.

The main goal of whole research is to prepare database for data mining and define soil types by using clusterization algorithms and make pedological map based on results.

The paper is organized as follows. In Section II detection of anomalies using K-means clustering is presented; in Section III soil database of Montenegro is reviewed. Performance of analyzed algorithms in the case of clustering data samples of soil are analyzed in Section IV through simulation results. Conclusion is drawn in Section V.

## II. CLUSTERING BASED ANOMALY DETECTION

Unsupervised anomaly detection is the process of finding outliers without prior need for training. Focus of this paper is clustering based anomaly detection.

K-means (KM) clustering of data is grouping of data samples in clusters. KM aims to make clusters which contain data that are more similar to each other (based on a selected properties) than to those in other clusters. Clusters are defined with the central points, called the centroids. Results of the KM method are exclusive clusters. Hence, KM clustering is also called hard clustering [6]. KM is widely used partitioning method.

The KM algorithm is sensitive to outliers. In Fig. 1. is shown the effect of outliers to result of clustering. Second cluster which is presented with red color is much bigger that it should be, because of outlier. If outliers are removed (red data sample) result of KM clustering is as expected (Fig. 2.). Thus, outliers have to be removed before clustering of soil data.



Fig. 1. Example of KM clustering with outlier



Fig. 2. Example of KM clustering without outlier

#### A. AD based on clustering and appearance of small or sparse cluster

The first method is AD based on clustering and appearance of small and sparse cluster. This method of AD assumes that normal data samples belong to large, dense clusters, while anomalies belong to small or sparse clusters. Hence, if clustered data contains one significantly smaller cluster, that cluster may present anomalies [7], [9].

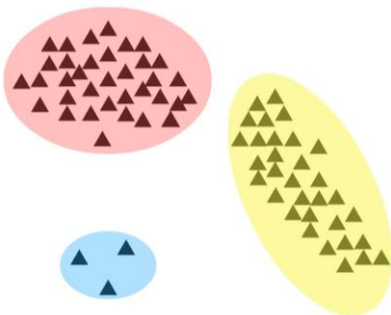


Fig. 3. Example of AD based on clustering, smallest cluster presented with blue color presents anomalies

This method involves analysis of resulted clusters. This analysis is comparing number of cluster members and visualization of clusters. One example of this method is shown in Fig. 3. The smallest blue cluster presents cluster of anomalies.

#### B. AD based on clustering and distance between data samples and its centroids

AD based on clustering and distance between data samples and its centroids is second proposed approach. KM can be applied on training data. The result of clustering in two clusters of training data will be a cluster of normal data and cluster of anomalies. Based on that for each new data sample can be inspected is that sample anomaly. After calculating and comparing distance between new data sample and two centroids, inspected

data sample will be declared as anomaly or normal data (Fig. 4.).

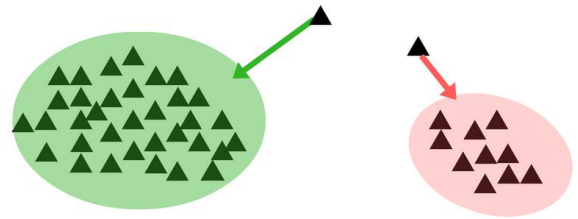


Fig. 4. Example of A. AD based on clustering and distance between data samples and its centroids

If there is no any training data, this approach can be used but with a modification. After clustering data into defined number of cluster, the distance between each data sample and its centroid will be calculated. Distance threshold should be defined. And data samples which are more distanced from its centroid then threshold are declared as an outlier (Fig. 5.). This method will be used in simulation, because there is no training soil data.

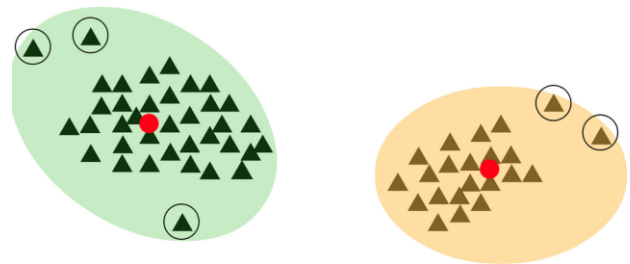


Fig. 5. Example of A. AD based on clustering and distance between data samples and its centroids, using threshold, detected outliers are signed with black circle

### III. PREPROCESSING OF THE DATABASE

The soil database of Montenegro was made during the period of 1958-1988. Data was collected and saved in hard copy version and main result of process is soil map in scale of 1:50 000. Digitalization of the database is result of researchers of BIO-ICT Centre of Excellence [10].

The database contains physical and chemical characteristic of soil. It contains more than 22 000 rows and more than 200 information per each row.

During the translation from hard copy version to *Excel* files some mistakes were made. Preprocessing of database was necessary before using various data mining tools for analysis and visualization.

#### A. AD using pedological knowledge

First step of preprocessing of database and cleaning data was done using a pedological knowledge. Distribution of value characteristics were analyzed, extremes are found etc. (Fig. 6, Fig. 7.). For determination of soil textural classes, the relative proportions of sand (coarse and fine),

silt and clay are used (Fig. 8.). According to methodology, sum of all fractions is 100%, total sand is sum of coarse and fine sand, while total clay is sum of silt and clay. The distributions of values of these soil fractions were analyzed.

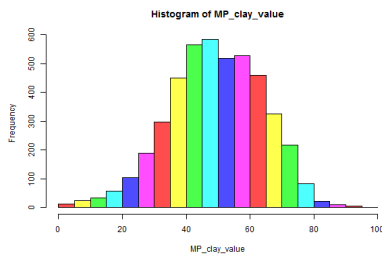


Fig. 6. Histogram of MP\_clay\_value

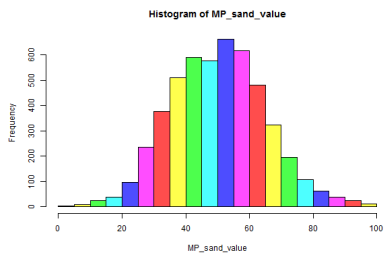


Fig. 7. Histogram of MP\_sand\_value

Based on mentioned rules for soil fractions, mistakes in database are detected.

Rule1:

$$MP\_sand\_value + MP\_clay\_value = 100$$

Rule2:

$$MP\_size2000\_0250um\_value + MP\_size0250\_0020um\_value = MP\_sand\_value$$

Rule3:

$$MP\_size0020\_0002um\_value + MP\_size0002um\_lt\_value = MP\_clay\_value$$

Rule4:

$$MP\_size2000\_0250um\_value + MP\_size0250\_0020um\_value + MP\_size0020\_0002um\_value + MP\_size0002um\_lt\_value = 100$$

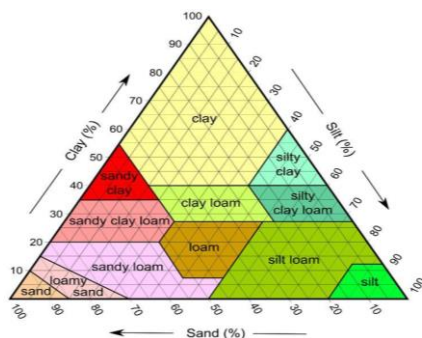


Fig. 8. Soil Texture Triangle

Visualization of detected anomalies mistakes is performed. In Fig. 9- Fig. 12 deviation of these four rules is presented, respectively. The X-axis is the index of data samples and the Y-axis presents the extent to which the data samples deviates from the rules 1, 2, 3 and 4, respectively. On these graphs data samples with zero value are without mistake.

Visualizing allows better overview of detected anomalies. These analysis and visualization is done using R programming language.

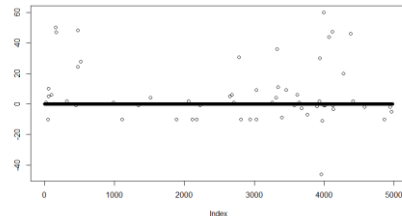


Fig. 9. Deviation from Rule1

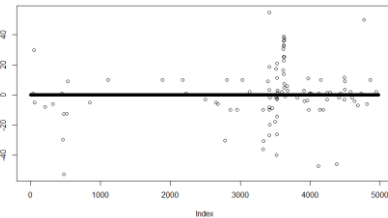


Fig. 10. Deviation from Rule2

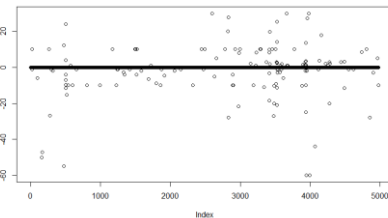


Fig. 11. Deviation from Rule3

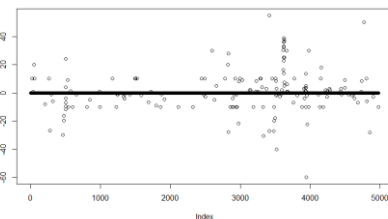


Fig. 12. Deviation from Rule4

Arrays, which contain number of profiles with deviation, were made. Using these arrays detected mistakes were corrected.

After detailed investigation of detected mistakes, a conclusion is that most of the mistakes are made during translation data from hard copy version to Excel files.

### B. Clustering based AD

1) AD based on clustering and appearance of small or sparse cluster

Results of one example of this method of AD for soil database are presented in Fig. 13. Three mechanical-physical characteristics of soil are chosen for clustering. After KM clustering in three clusters results are visually presented in 3D.

Green cluster in which is signed with red circle is detected as a cluster of outliers and it is used for more detail analysis.

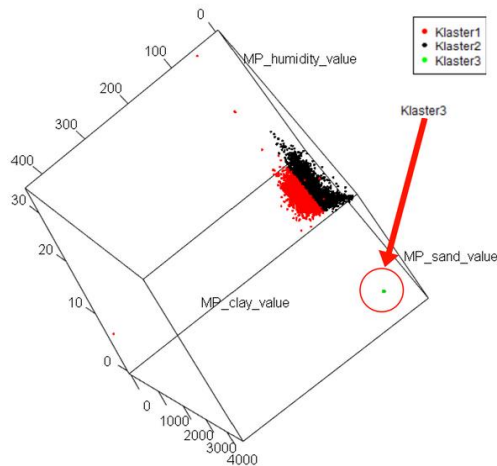


Fig. 13. Anomaly definition using KM clustering - the smallest cluster (green color) presents anomalies

## 2) AD based on clustering and distance between data samples and its centroids

KM clustering of soil data is done in five clusters based on three mechanical-physical characteristics. After that distance between each data point and its cluster centroid is calculated (Fig. 14.). Also, an average distance for each centroid is calculated.

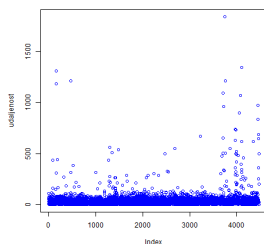


Fig. 14. Distance each data point from centroid of its cluster

One example of this AD approach is shown in Fig. 4. In this example smaller distance is chosen as a threshold, so anomalies are numerous and more visible on this graph. Data samples colored with yellow are detected anomalies. Each cluster is presented with different color, while the purple points are centroids.

## IV. CONCLUSION

A preprocessing of the database is performed using different methods and algorithms. Proposed methods are analyzing database based on pedological rules and two different approaches using clustering.

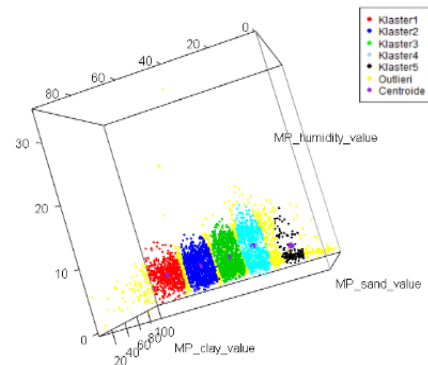


Fig. 15. Anomaly definition using KM clustering - yellow points presents outliers

After detection of all mistakes, anomalies and outliers, they were replaced or modified based on excretory data - hard copy version of pedological data. Almost all of detected anomalies are made in translation from hard copy version to Excel files (i.e. 8004 instead of 80.04). Some of the detected anomalies are data samples with unusual combination of the value of parameters. Experts from Biotechnical faculty are investigating now whether the detected samples are really anomalies.

Digital pedological soil database is ready for data mining tools and visualization.

## REFERENCES

- [1] I. Witten, E. Frank, M. Hall, "Data Mining: Practical Machine Learning Tools and Techniques," The third edition, Morgan Kaufmann Publishers, Januar 2011.
- [2] X. Wu, V. Kumar, J. Ross Quinlan, J. Ghosh, Q. Yang, H. Motoda, G. J. McLachlan, A. Ng, B. Liu, P. S. Yu, Z. Zhou, M. Steinbach, D. J. Hand, D. Steinberg, "Top 10 algorithms in data mining," Published online: 4 December 2007, Springer-Verlag, London Limited, 2007.
- [3] Hodge, V.J, J. Austin, "A Survey of Outlier Detection Methodologies," *Artificial Intelligence Review*. 22 (2): 85–126. doi:10.1007/s10462-004-4304-y., 2004
- [4] V. Chandola, A. Banerjee, V. Kumar, "Anomaly detection: A survey," *ACM Computing Surveys* 41 (3): 1–58. doi:10.1145/1541880.1541882, 2009
- [5] G. Munz, S. Li, G. Carle, "Traffic Anomaly Detection Using K-Means Clustering," GI/ITG Workshop MMBnet, doi 10.1.1.323.6870, 2007
- [6] E. Hot, V. Popović-Bugarin, "Soil data clustering by using K-means and fuzzy K-means algorithm," *23rd Telecommunications Forum TELFOR 2015*, Belgrade, November 2016.
- [7] V.J. Hodge, J. Austin, "A survey of outlier detection methodologies," *Artificial Intelligence Review*, pp. 85-126. ISSN 1573-7462, 2004
- [8] H. Zengyou, X. Xiaofei, D. Shengchun, "Discovering cluster-based local outliers," *Pattern Recognition Letters* 24 (9–10): 1641–1650. doi:10.1016/S0167-8655(03)00003-5, Elsevier Science, December 2002
- [9] S. Ramaswam, R Rastogi, K. Shim, "Efficient Algorithms for Mining Outliers from Large Data Sets," *Proceedings of the 2000 ACM SIGMOD international conference on Management of data – SIGMOD* p. 427. doi:10.1145/342009.335437. ISBN 1-58113-217-4.
- [10] E. Salković, "Digitizing the pedological data of Montenegro," *Information Technology IT 2015 Zabljak*, Montenegro, February 2015.

# Method for Measuring Released Amount of Water in Smart Irrigation System

Trifun Savić, Balša Bajagić, Mirko Knežević, Milutin Radonjić, *Member, IEEE*

**Abstract** — Method for measuring released amount of water in smart irrigation system is presented in this paper. The calculation of optimal amount of water, which is used as input parameter for smart irrigation algorithm, is also described. In order to develop reliable signal detection from the Hall sensor mounted on the water meter, we decided to use software interrupt on Libelium Wasmote open source platform. Proposed solution enables reliable measurement, thus optimizing the irrigation process and providing optimal water saturation of the agricultural field.

**Keywords** — Agriculture, irrigation, software interrupt, Wasmote

## I. INTRODUCTION

CONSTANT growth of the human population, food demands and demands for energy, together with the competition between different water consumers, increase pressure on water resources [1]. In addition, effects of global climate change and other environmental and economic issues are causing further complications related to water resource management. These problems pose a great obstacle to the development of sustainable agriculture. Therefore, different strategies and actions are needed in order to preserve water resources worldwide.

Agriculture is one of the most dominant consumer of water resources globally. Agricultural production system depends mainly on two types of agriculture: rainfed agriculture and irrigated agriculture. Rainfed agriculture is the most dominant one. It depends on the rainfall and it is practiced at highland areas and in the dry and humid tropics. Eighty percent of the current world cultivated area are rainfed [2]. Weather conditions and different soil characteristics often cause inadequate soil moisture

availability, and techniques to improve water availability (such as water harvesting) are expensive. On the other hand, irrigated agriculture is far more productive and it is developed rapidly in the past years. However, most irrigated farming systems are performing well below their potential, and there is considerable scope for improving agricultural productivity in irrigated agriculture.

Unless water-use efficiency is increased, greater agricultural production will require increased irrigation [3]. The usage of state of the art ICT technologies, in developing automated and smart irrigation systems, allows autonomous irrigation control, thus providing optimal water consumption and the reduction of human labor. Automated irrigation systems provide easy and flexible control of the irrigation process, and they are considered as the first step in smart irrigation system development [4]. In such systems, it is essential to have precise information about the amount of water that is released in the agricultural field in order to provide the maximum water efficiency according to the needs of plants.

In this paper we present the solution for measuring the released amount of water on the agricultural fields. System is developed on the Libelium Wasmote platform, and water meter is equipped with Hall sensor for digital water-flow measurement. Detection of the signal generated by the Hall sensor is implemented using software interrupt.

The paper is organized as follows. Method for calculation of optimal amount of water for irrigation is given in section II. In section III we present part of the algorithm for smart irrigation. Measurement of the released amount of water using software interrupt is described in section IV. Paper is concluded in section V.

## II. CALCULATION OF OPTIMAL AMOUNT OF WATER FOR IRRIGATION

Amount of water that should be released to the agricultural field is calculated by soil science specialists based on the soil characteristics and the type of plant that is grown on that field. It is determined prior to installation of the system. This amount of water has to be released when the soil dries up to a certain level. This amount of water represents irrigation norm, and it is the volume that is given to the plant in one irrigation.

Calculation of irrigation norms largely depends on the soil and its characteristics, and knowledge of the crop, which is in the plantation [5].

To properly calculate the norm of irrigation it is necessary to:

This work has been supported by the Ministry of Science of Montenegro and the HERIC project through the BIO-ICT Centre of Excellence (Contract No. 01-1001).

Corresponding Trifun Savić is with the University of Montenegro, Faculty of Electrical Engineering, Bul. Džordža Vašingtona bb, 81000 Podgorica, Montenegro (phone: 382-20-245839; e mail: trifuns@ac.me). Balša Bajagić is with the University of Montenegro, Biotechnical Faculty, Bul. Džordža Vašingtona bb, 81000 Podgorica, Montenegro (phone: 382-20-245839; e mail: bals6666@gmail.com). Mirko Knežević is with the University of Montenegro, Biotechnical Faculty, Bul. Džordža Vašingtona bb, 81000 Podgorica, Montenegro (phone: 382-20-245839; e mail: mirkok@ac.me). Milutin Radonjić is with the University of Montenegro, Faculty of Electrical Engineering, Bul. Džordža Vašingtona bb, 81000 Podgorica, Montenegro (phone: 382-20-245839; e mail: m.radonjic@ieee.org).

- Take a soil sample and determine its mechanical composition.

- Determine soil water properties.

Based on mechanical analysis we can calculate:

- *FC* (field capacity) - represents the quantity of water which remains in the soil, after the excess gravitational water drain, or almost completely stopped the movement of water in this way.

- *WP* (wilting point) - refers to the quantity of water in the soil in which the plants permanently vein.

- Saturation (maximum water capacity) - the maximum amount of water in soil, i.e. porosity.

- *TAW* (total available water) - represents usable water for the plant in the soil, or water that the plant can use for their physiological needs (the difference in the water content of the *FC* and *WP*).

- *RAW* (readily available water) - the lower limit of reduction in soil water content.

- Depletion - the recommended water consumption.

#### A. Basic steps before calculating the irrigation norm

First step is to take soil samples at certain soil moisture tension levels, measured by appropriate sensor (for example: 33 cbar, 80 cbar, 160 cbar and 200 cbar). Using laboratory methods we determine the weight of the moisture content in the samples ( $T_c$ ). At the same time we take soil samples, using Kopecký cylinder [6], and determine the bulk density ( $\rho = M_s / V_t$ ). After these steps we calculate the volumetric moisture content by multiplying the weight of moisture content and density of the soil ( $W_{m(\%)} = T_c * \rho$ ). Calculation of water volume depends on the surface plot and the depth of the root system. Soil moisture characteristic curves are depicted in Fig. 1.

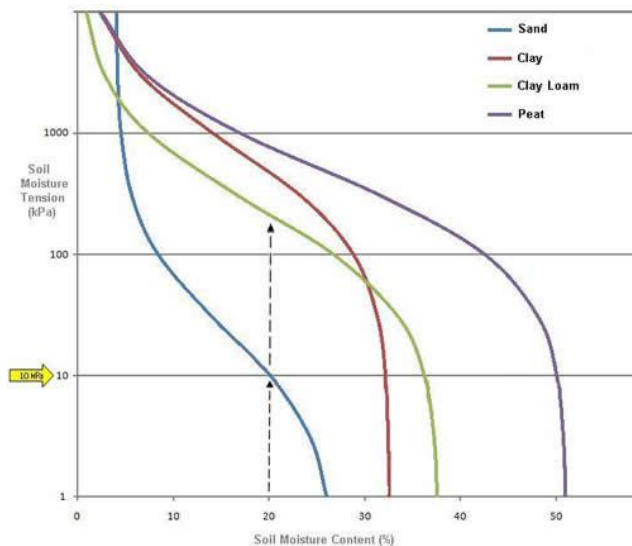


Fig. 1 Soil moisture characteristic curves [7]

#### B. Calculation of the irrigation norm

Calculation of total available water is given with:

$$TAW = FC - WP \quad (1)$$

Determination of the limits allowed for water consumption, i.e. depletion is presented with:

$$Depletion = \eta * TAW \quad (2)$$

or:

$$TAW = Depletion + RAW \quad (3)$$

where coefficient  $\eta$  for different crops can varies from 0.3 to 0.6. The water amount for irrigation is calculated according to the following equation:

$$W = 100 * d * \rho * (FC - RAW) \left( \frac{m^3}{ha} \right) \quad (4)$$

where  $d$  is plant's root depth. Relationship between soil moisture and crop stress [8] is shown in Fig. 2.

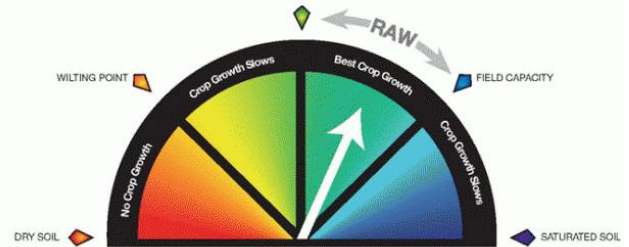


Fig. 2. Soil moisture and crop stress relationship [8]

### III. ALGORITHM FOR SMART IRRIGATION

Development of smart irrigation system is based on multi-disciplinary approach, since creation of high-quality irrigation algorithm requires knowledge of experts in the field of soil science. The algorithm aims to determine the optimal irrigation period, based on the different environmental and soil parameters, and the type of plant.

System is designed to independently make decisions on switching on and off solenoid valves for irrigation. The most important input parameter of the algorithm, besides the released amount of water, is soil moisture tension. We decided to collect this data using Wireless Sensor Network (WSN). WSNs are networks of battery-powered sensor nodes interconnected through wireless medium and are typically deployed to serve a specific application purpose [9]-[10]. The number of nodes typically depends on the surface area that is irrigated. Each node is equipped with soil moisture tension sensor and communication module. Nodes exchange measured values and transfer them to the main control unit (microprocessor platform). Ambient conditions, temperature and humidity are obtained by the weather station, located in the agricultural field. Based on the obtained data, by an algorithm for smart irrigation, main controller unit determines the moments for switching on and off solenoid valves that regulate the water flow. The system also provides the communication between the main controller unit and the cloud platform [11] used for monitoring and analysis of the system.

One of the algorithm input parameters, the optimal amount of water  $W$  to be released, has to be assigned. This parameter is calculated for the particular agricultural field, where the smart irrigation system will be installed. The other input parameter is soil moisture tension limit  $S$ . This parameter indicates the extent to which the soil should be allowed to dry off, before irrigation starts. Measured amount of released water and measured soil moisture tension are labeled as:  $Wm$  and  $Sm$ , respectively. The measured value of the soil moisture tension  $Sm$  is



permanently compared with the parameter  $S$ . When the measured value reaches the level of dryness  $S$ , irrigation starts and water is released, through the water meter, to the agricultural field. The amount of released water  $W_m$  is permanently measured and compared with the calculated optimal amount of water  $W$ . When the amount of released water reaches optimal amount, irrigation ends. We reset the measured amount of water and wait again until the measured soil moisture tension reaches the maximal level of dryness  $S$ . The flowchart of this part of algorithm for smart irrigation is depicted in Fig. 3.

In order to have an optimal water saturation of the agricultural field, it is obvious that the measurement of the released amount of water must be accurate. Therefore, special attention has to be paid to detection of signal from the water meter, as described in the next section.

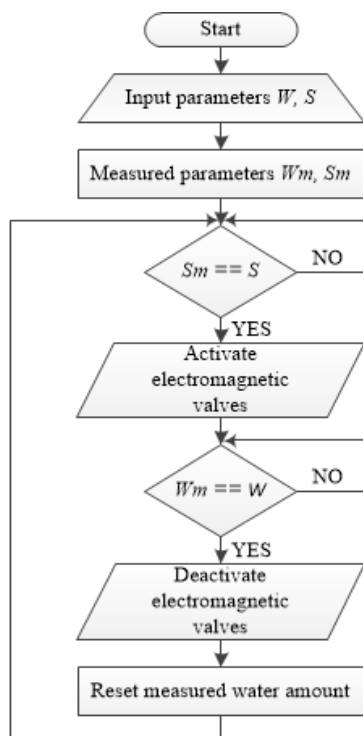


Fig. 3. Portion of the algorithm for smart irrigation

#### IV. MEASURING AMOUNT OF RELEASED WATER USING SOFTWARE INTERRUPT

Information on the amount of water that is released through the valves is obtained from the water meter, by means of the Hall sensor [12]. Water meters equipped with the Hall sensor are also known as pulse water meters (see Fig. 4.). On the water meter used in our system there are four needles for showing water flow in  $m^3$  on four scales:  $0.0001m^3$ ,  $0.001m^3$ ,  $0.01m^3$  and  $0.1m^3$ . A small magnet is located on the needle that we want to observe. During one rotation of the needle, Hall sensor is activated and pulse signal is generated. This magnet can be mounted on any of the four needles. Depending on the position of the sensor, i.e. which needle sensor observes, each pulse signal represents certain amount of released water. In our testbed

implementation, single impulse can represent: 0.1L, 1L, 10L or 100L of released water.

The sensor circuit is open collector and supplied with 3.3V from Waspote board. The output signal from the sensor is connected to input port with a pull-up resistor. This means that unexcited sensor gives a high output signal. When magnet at the appropriate needle passes below the sensor, the output signal drops to zero. Thus, the control unit detects a falling edge signal when there is a full rotation of the proper needle. Detection of this signal has to be reliable and it is therefore straightforward to be performed by an interrupt. However, appropriate hardware interrupt pins, available on the Waspote platform, were already used to operate GPRS module. Thus, we used software interrupt to detect falling edge of the Hall sensor output signal.



Fig. 4. Water meter with Hall sensor

Since one rotation of the needle generates only one pulse signal, which represents a quite large amount of water, it is very important to detect each falling edge. The most reliable solution is to detect signal using interrupt routine. Since we couldn't use hardware interrupt, software interrupt was logical choice. For this purpose, interrupt is generated using built in microcontroller's timer/counter. This software interrupt, also known as timer interrupt, allows us to perform a task at very specifically timed intervals regardless of what's happening in the main program. We can immediately interrupt the normal sequence of events taking place in the loop() function at precisely timed intervals, and execute an appropriate sequence of commands. Once these commands are executed, the program continues where it was interrupted in the loop() function.

Libelium Waspote open source microcontroller platform is based on ATmega1281 microcontroller. It has two 8-bit timer/counters and four 16-bit timer/counters. Each timer has a counter that is incremented on each tick of the timer's clock. For this purpose, Clear Timer on Compare Match (CTC) timer interrupt was initiated, which is the most commonly used timer interrupt. CTC timer interrupts are triggered when the counter reaches a specified value, stored in the compare match register. Once a timer/counter reaches this value it will reset to zero on the next tick of the timer's clock, continuing to count up to the compare match value again. We can define the frequency of timer interrupts by setting the compare match

value and timer's clock period.

The maximum frequency of the timer/counter clock for ATmega1281 is 16 MHz. Using 16-bit timer/counter with the maximum compare match register value (65536), interrupts will occur every 4 ms. In order to increase this period we have to use prescaler. A prescaler controls the clock of timer according to the following equation:

$$f_t = \frac{f_c}{p} \quad (5)$$

where  $f_t$  is the timer clock frequency,  $f_c$  is Waspote's clock frequency and  $p$  is prescaler. The prescaler can be 1, 8, 64, 256 or 1024. Therefore, we can lower timer clock frequency up to 1024 times.

Interrupt frequency  $f_i$  can be calculated as:

$$f_i = \frac{f_c}{p * (c + 1)} \quad (6)$$

where  $c$  is the content of compare match register. To set up desired frequency of interrupt, we calculate compare match register content as:

$$c = \frac{f_c}{p * f_i} - 1 \quad (7)$$

This value has to be less than 256 for 8-bit timers and less than 65536 for 16-bit timers.

For the purpose of this application we used 16-bit timer/counter1, with interrupt frequency of 4 Hz, i.e. 250 ms period and 256 prescaler. The following code is used for setting timer/counter1 interrupt used in this application:

```
noInterrupts(); // disable all interrupts
TCCR1A = 0; // set entire TCCR1A register to 0
TCCR1B = 0; // set entire TCCR1B register to 0
TCNT1 = 0; // initialize counter value to 0
// set compare match register for 4Hz increments
OCR1A = 15624; // (16*10^6) / (4*256) - 1
TCCR1B |= (1 << WGM12); // turn on CTC mode
TCCR1B |= (1 << CS12); // Set CS12 bit for 256 prescaler
TIMSK1 |= (1 << OCIE1A); // enable timer compare interrupt
interrupts(); // enable all interrupts
```

The detection of the falling edge of the Hall sensor signal during timer interrupt is implemented with the following code:

```
// timer compare interrupt service routine
ISR(TIMER1_COMPA_vect){
    //checks if the signal from the water meter dropped to zero
    if(digitalRead(VODOMJER) == 0) {
        if(vodomjer_staro == 1) { // to count only once per signal
            br_rotacija++; // signal counter
        }
        vodomjer_staro=0;
    }
    else {
        vodomjer_staro=1;
    }
}
```

It is important to keep interrupt routine as short as possible. This is especially important if we are interrupting frequently.

This method for measuring the amount of released water in the agricultural field is implemented and tested on the

prototype of the smart irrigation system, shown in Fig. 5.



Fig. 5. Prototype of the smart irrigation system

## V. CONCLUSION

In this paper, we present method for measuring amount of water released during the irrigation process. Also, the method for calculating amount of water needed for irrigation is explained. The main goal was to develop reliable solution for measuring released water with minimum efforts. This measurement is performed using software interrupt implemented on Libelium Waspote open source platform. Our approach is verified by implementation on the prototype of the smart irrigation system, which is installed on the experimental field. Proposed solution allows users to optimize irrigation process and provides optimal water saturation of the agricultural field.

## REFERENCES

- [1] Karimov, Akmal, et al. "A water accounting procedure to determine the water savings potential of the Fergana Valley." *Agricultural water management* 108 (2012): 61-72.
- [2] FAO. 2011. The state of the world's land and water resources for food and agriculture (SOLAW) – Managing systems at risk. Food and Agriculture Organization of the United Nations, Rome and Earthscan, London.
- [3] Tilman, David, et al. "Agricultural sustainability and intensive production practices." *Nature* 418.6898 (2002): 671-677.
- [4] Savic T., Radonjic M., "Proposal of Solution for Automated Irrigation System", Proc. of 24th Telecommunication Forum TELFOR 2016, Belgrade, Serbia, November 2016.
- [5] Resulović, H., Čustović, H., (2002) *Pedologija opšti dio* (Knjiga I), Sarajevo.
- [6] Dugalić, G., Gajić, B., (2005) *Pedologija praktikum*, Čačak
- [7] <http://mea.com.au/soil-plants-climate/soil-moisture-monitoring/learning-centre/comparing-soil-moisture-tension-and-soil-moisture-content>
- [8] [http://hydrogold.org/education/soil\\_moisture\\_concepts.html](http://hydrogold.org/education/soil_moisture_concepts.html)
- [9] Adamala, S., Raghuvanshi, N.S., Mishra, A., 2014. Development of surface irrigation systems design and evaluation software (SIDES). *Comput. Electron. Agric.* 100, 100–109.
- [10] Akyildiz, I.F., Kasimoglu, I.H., 2004. Wireless sensor and actor networks: research challenges. *Ad Hoc Netw.* 2 (4), 351–367.
- [11] <http://blueleaf.ac.me>
- [12] Ramsden, Edward. *Hall-effect sensors: theory and application*. Newnes, 2011.

# Measurements Calibration Analysis of Novel IoT Platforms

Nikola Bulatović, Tanja Vukićević, Rabina Šabotić, Vanja Asanović

**Abstract** — In this paper, we have presented the results of novel IoT sensors calibration. The concept of Internet of Things (IoT) is based on the global network of connected devices, which communicate without human intervention, with the aim to sense our environment and make appropriate decisions upon measured data. However, the Internet of Things introduces new risks. The unprecedented amount of generated data has to be reliable and accurate to make appropriate decisions. Hence, there is a need for accurate measurement methods to support and improve security and reliability of IoT products. Calibration results will be used to increase the accuracy of climate measurements needed to understand climate change.

**Keywords** — Internet of Things platforms, measurement calibration.

## I. INTRODUCTION

TODAY we are facing incredible technological improvements and developments. The global network of Internet of Things (IoT) nodes is developed with the aim to cover all aspects of humanity through sensing and measuring environmental parameters [1]-[3]. The unprecedented amount of data is being generated, but those data have to be reliable and accurate to make fair and valid decisions. Hence, there is a need for accurate measurement methods which support and improve security and reliability of IoT products. The concept of measurement has deep roots in the human culture. Measurement shapes scientific theories and represents the bridge between the empirical world of phenomena and the abstract world of concepts and knowledge [4]. Accurate measurement implies the existence of standards of measurement. The evaluation of uncertainties is inherent in the measurement process. If the instruments used were not calibrated there will be doubt regarding the accuracy and reliability of all

measurements. It is well known that calibration is operation that, under specified conditions, in a first step, establishes a relation between the quantity values with measurement uncertainties provided by measurement standards and corresponding indications with associated measurement uncertainties and, in a second step, uses this information to establish a relation for obtaining a measurement result from an indication [4]. Understanding calibration is not complete without understanding traceability which provides the confidence and convenience of calibration results. Traceability is defined as the property of a measurement result whereby the result can be related to a reference through a documented unbroken chain of calibrations, each contributing to the measurement uncertainty [5]. Traceability ensures that a measurement result can be confidently linked to the internationally accepted measurement references.

New ICT solutions, integrating the core concepts of Internet of Things (IoT), are used for efficient and continuous environmental biomonitoring. Several of these platforms are deployed within the BIO-ICT Center of Excellence in Montenegro. The purpose of the platforms is to enable accurate and reliable real-time monitoring of various environmental parameters in agricultural and maritime areas. Therefore, we have calibrated the sensors for measuring air temperature (air sensors), sensors for measuring land temperature and sensors for measuring relative air humidity operating within Davis Vantage Pro2 weather station and Libelium Smart Agriculture. The calibration results showed that after only one year of exploitation, the differences between measured and actual values could have a significant influence on actions performed by corresponding actuators.

In the Section II, the overview of novel IoT concepts and platforms is presented. Calibration methods are described in Section III, while the obtained results are shown in Section IV. Finally, conclusions are given in Section V.

## II. IOT AND BIO-MONITORING PLATFORMS

The foundation concept of IoT platforms is based on the following principles: data measurement, data transfer, data processing, data analysis and presentation and finally decision making upon the processed data. All the mentioned steps are focused on data. That is why data

Corresponding Nikola Bulatović is with the Faculty of Electrical Engineering, University of Montenegro, Džordža Vašingtona bb, 81000 Podgorica, Montenegro (phone: 382-67-418264; e-mail: nbulatovic@ac.me).

Tanja Vukićević is with the Bureau of Metrology, Kralja Nikole 2, 81000 Podgorica, Montenegro (e-mail: tanja.vukicevic@metrologija.gov.me).

Rabina Šabotić is with the Bureau of Metrology, Kralja Nikole 2, 81000 Podgorica, Montenegro (e-mail: rabina.sabotic@metrologija.gov.me).

Vanja Asanović is with the Bureau of Metrology, Kralja Nikole 2, 81000 Podgorica, Montenegro (e-mail: vanja.asanovic@metrologija.gov.me).

mining is one of the most promising research fields today. New data mining algorithms are evolving every day to search for hidden patterns in large data clusters. Hidden data patterns are used for numerous applications such as: prediction of customer behavior, earthquakes early detection, monitoring of vehicles and pedestrian levels to optimize driving and walking route, control of CO<sub>2</sub> emissions of factories, control real time leakages and wastes in the sea, monitoring of soil moisture and trunk diameter in vineyards to control the amount of sugar in grapes and grapevine health [1]-[3], [6]-[7].

The IoT platforms for environmental parameters measuring have some disadvantages. These are the high cost of the platforms and sensor probes, inaccuracy of measurements with time passing and low sensor probes durability and environmental resistance. Platforms are usually sealed in a “black box”, disabling any internal architecture or code editing. The integration of different manufacturers’ sensor probes with the platform’s processing unit can also represent a major issue. The designing the system as the main idea of the novel IoT platforms is focused on outperform the existing solutions regarding cost, power consumption, ease of use, as well as the integration of different sensor probes. With all these new trends in IoT development, data measuring accuracy has to be maintained at high precision levels.

Davis Vantage Pro2 weather station is one of the most reliable and accurate IoT platforms designed for measuring, monitoring and managing weather data [8]. The platform consists of the integrated sensing unit equipped with sensor probes and the communication console for real-time wireless data transfer to the desired cloud application as shown in Fig. 1. The sensing unit is customizable with various sensor probes for measuring weather data such as inside and outside temperature and humidity, barometric pressure, dew point, rainfall, wind speed and direction, UV and solar radiation.

It can be used for designing automatic irrigation system based on plant water demand measured with evapotranspiration (ET) sensor probe. The soil moisture probes are also used to improve irrigation decision making. The platform is used for measuring and forecasting pesticide risk and danger levels to specific crops. The platform enables weather forecast for maximum harvest yields. A significant number of farmers rely on Davis weather stations to help them prevent from crop damage during frost events in the form of an early notification and alarm message.

Davis Vantage Pro2 weather station’s temperature and humidity sensor probes measure air temperature in the range from -40 °C to 65 °C and air humidity in the range from 1 % rh to 100 % rh. Air temperature sensor probe has a resolution of 0,1 °C with the accuracy of  $\pm 0,3$  °C. Air humidity sensor probes have a resolution of 1% rh with the accuracy of  $\pm 2$  % rh and drift of  $\pm 0,5$  % rh per year.

Libelium Smart Agriculture platform designed to be the user-friendly and portable unit is shown in Fig. 2 [9]. Very simple integration of versatile sensor probes and low

power consumption is what makes this platform so popular.

Libelium Smart Agriculture platform is widely used in different agriculture areas around the world such as:

- precision farming to control irrigation and improve fertilization strategies on corn crops
- rain forest monitoring for climate change control
- monitoring weather conditions to prevent pest in olives
- smart agriculture production of organic food.

Libelium platform’s integrated temperature sensor probe has a measurement range from -40 °C to 123 °C with a measurement resolution of 0,04 °C (minimum) and 0,01 °C (typical). The accuracy of air temperature measurement is  $\pm 0,4$  °C (in the range from 0 °C to 70 °C) and  $\pm 4$  °C (in the range from -40 °C to 125 °C).

Air relative humidity sensor probe has a measurement range from 0 % rh to 100 % rh with a measurement resolution of 0,4 % rh (minimum) and 0,05 % rh (typical). The accuracy of air humidity measurement is  $\pm 1,8$  % rh.



Fig. 1. Davis Vantage Pro2 weather station.



Fig. 2. Libelium Smart Agriculture sensing platform.

### III. CALIBRATION METHODS

Calibrations of sensors for measuring air temperature (air sensors), sensors for measuring land temperature and sensors for measuring relative air humidity operating within Davis Vantage Pro2 weather station and Libelium Smart Agriculture, were performed in the Laboratory for Temperature of the Bureau of Metrology. The Laboratory is responsible for realization, preservation and maintenance of the national standard of thermodynamic temperature, the Standard Platinum Resistance Thermometer (SPRT) for temperatures in the range from  $-80\text{ }^{\circ}\text{C}$  to  $650\text{ }^{\circ}\text{C}$  and thermocouple (TC) in the temperature range from  $650\text{ }^{\circ}\text{C}$  to  $1200\text{ }^{\circ}\text{C}$ . Also, the Laboratory is responsible for realization, preservation, and maintenance of the national standard of relative humidity dewpoint meter (DP) in the temperature range from  $10\text{ }^{\circ}\text{C}$  to  $70\text{ }^{\circ}\text{C}$  and relative humidity from 10 % rh to 90 % rh. These standards have the best metrological properties in Montenegro and, in addition to other equipment, serve as the basis for determining metrological traceability in the field of temperature and relative humidity.

The above-specified sensors were calibrated following internal calibration procedures, accredited by the MEST ISO/IEC 17025 standard. The competence of the Bureau of Metrology for performing calibrations of temperature measuring instruments, following the requirements of the MEST EN ISO/IEC 17025 standard, was approved by Accreditation Body of Montenegro (ATCG) as well as Croatian Accreditation Agency (HAA) signatory to the EA MLA agreement. Furthermore, the Bureau of Metrology proved its competence for performing certain calibrations of temperature measuring instruments/standards in the indicated scope and with the shown measurement uncertainties through the publishing of calibration and measurement capabilities (CMC) in the field of contact thermometry in the BIPM KCDB.

Climate chamber manufactured by KAMBIĆ as shown in Fig. 3, type: KK-190 CHULT, capacity: 190 L, temperature range from  $-75\text{ }^{\circ}\text{C}$  to  $150\text{ }^{\circ}\text{C}$ , humidity range from 10 % rh to 90 % rh, resolution  $0,1\text{ }^{\circ}\text{C}/\%$  rh was used for the calibration of sensors for temperature measurement (air sensors) as a thermostatic medium. Working Standard Platinum Resistance Thermometer Pt100, manufactured by Fluke Hart Scientific, type 5615, manufacturer number: 842198 was used as a reference standard. Sensors for air temperature measurement were calibrated as air sensors in the climatic chamber using direct comparison with Laboratory's working temperature measurement standard. Sensors and working standard were set in the thermostatic medium – climatic chamber, whose temperature was measured with the working standard and read on the Hart 1560 System instrument, of serial number B05228. Calibration was performed in 7 calibration temperature points in the range from  $-10\text{ }^{\circ}\text{C}$  to  $50\text{ }^{\circ}\text{C}$ .

Two sensors for temperature situated within the Libelium Smart Agriculture platform were calibrated as digital indication thermometers. In the calibration of these sensors, the following thermostatic media were used:

alcohol bath from  $-80\text{ }^{\circ}\text{C}$  to  $+20\text{ }^{\circ}\text{C}$ , ice point at  $0,01\text{ }^{\circ}\text{C}$  and water bath from  $5\text{ }^{\circ}\text{C}$  to  $80\text{ }^{\circ}\text{C}$ . Working Standard Platinum Resistance Thermometer Pt100, manufactured by: Fluke Hart Scientific, type 5615, manufacturer number: 842198 was used as a reference standard. Thermometers were calibrated using direct comparison with the Laboratory's working temperature measurement standard in thermostatic media in 7 calibration temperature points, range from  $-10\text{ }^{\circ}\text{C}$  to  $50\text{ }^{\circ}\text{C}$ .



Fig. 3. Climate chamber type: KK-190 CHULT.

During the calibration of sensors for measuring relative air humidity, a climatic chamber was used as a thermostatic medium, while precise dewpoint meter, manufactured by Michell Instrument, type: dewpoint sensor, manufacturer number: 14483ST2 was used as a working standard. Sensors for measuring relative air humidity (relative humidity measuring instruments) were calibrated using direct comparison in the generator of relative humidity - climatic chamber, whereby the measured dewpoint temperature and generator's ambient temperature, are both in accordance to ITS-90. Based on these temperatures, the relative humidity in the relative humidity generator was calculated. Sensors were calibrated in calibration points 30 % rh, 50 % rh and 70 % rh at the temperature of  $22\text{ }^{\circ}\text{C}$ .

The importance of calibrated and accurate data, obtained through the environmental measurements, is of great value and can be indicative in the final step of decision making upon the measured data.

### IV. RESULTS

The output of the calibration process applied to Davis Vantage Pro2 and Libelium platforms gave the results which can be seen in Fig. 4. Three Davis Weather stations, located at different agricultural fields in Montenegro (BTF, Spuž, Zeta), were calibrated regarding outside air temperature. The biggest variations can be seen at the weather station located in an agricultural field in Spuž. Temperature sensor probe's corrections vary from  $-0,2\text{ }^{\circ}\text{C}$  to  $1,1\text{ }^{\circ}\text{C}$  which can have a significant influence on actions performed upon the inaccurate measurements. Other two weather stations, located at BTF's agricultural field and in Zeta, have lower variations of air temperature correction.

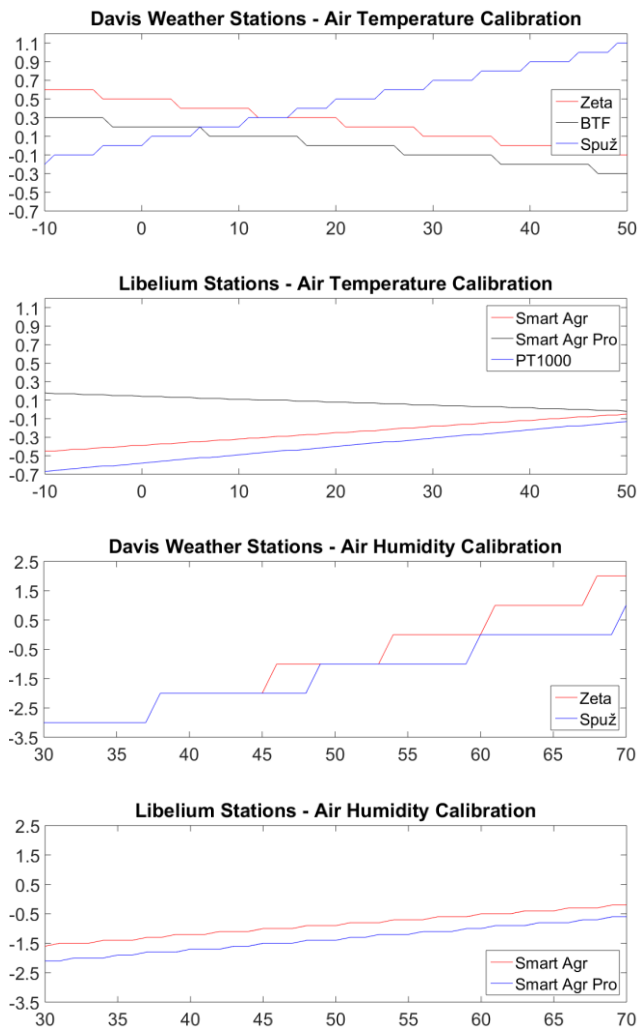


Fig. 4. Calibration corrections for air temperature and humidity sensor probes in three Davis Vantage Pro2 weather stations (located in Zeta, Biotechnical Faculty's agricultural field and Spuž) and two Libelium Smart Agriculture platforms. Air temperature corrections (y-axis) are expressed in degrees Celsius in the range of reference values (x-axis) from -10 °C to 50 °C. Air humidity corrections (y-axis) are expressed in percentages in the range of reference values (x-axis) from 30 % rh to 70 % rh.

Two Libelium platforms showed better results, regarding air temperature measurements accuracy. Maximal variations of both Libelium platforms' temperature sensor probes is lower than 0,5 °C which should not have a significant influence on final decision-making system. Regarding air relative humidity measurements obtained from both platforms, it can be seen that the air humidity measurement accuracy for Davis weather stations varies between -3 % rh and +2 % rh. The accuracy is a little bit better in Libelium platforms with the maximal correction of 2 % rh with measurement uncertainty of 3,3 % rh.

The measurement standards used in described calibrations were calibrated in laboratories of an NMI or in laboratories of the EURAMET Designated Institutes whose CMC was published in the BIPM KCDB or who are accredited by accreditation body signatory to the EA MLA agreement. Hence, an unbroken chain of traceability of

metrological results of measurements carried out in the laboratory to the International System of Units (SI) is achieved. These calibration results are traceable to international standards which represent the realization of the physical quantity unit in agreement with the International System of Units (SI).

The measurement uncertainty was evaluated and expressed according to [10] and [11]. Expanded measurement uncertainty is given as standard measurement multiplied by coverage factor  $k=2$  that, when a normal distribution is applied, corresponds to the desired confidence level equaling to approximately 95%.

## V. CONCLUSION

In this paper we have presented the results of calibration of sensors for temperature and humidity in Davis Vantage Pro2 weather station and Libelium Smart Agriculture platform. The overall goal of calibration of sensors is the use of precise metrological weather stations guaranteeing accurate and traceable measurements of temperature and humidity, which was achieved based on calibration results presented in this paper. Hence it is necessary to perform calibration of equipment in regular periods of time. Since the weather stations operate in a wide range of environmental conditions it represents a challenge to develop system for in situ calibration of instruments, including those operating in extreme conditions. In situ calibrations will insure consistency and reliability of all measurements.

## ACKNOWLEDGMENT

This work has been supported by the Ministry of Science of Montenegro and the HERIC project through the BIO-ICT Centre of Excellence (Contract No. 01-1001).

## REFERENCES

- [1] R. Soderbery, "How Many Things Are Currently Connected To The 'Internet of Things' (IoT)", 2013.
- [2] D. Uckelmann, M. Harrison, and F. Michahelles, "An Architectural Approach Towards the Future Internet of Things", Springer Berlin Heidelberg, March 2011.
- [3] L. Yan, Y. Zhang, L. T. Yang, and H. Ning, "The Internet of Things: From RFID to the Next-Generation Pervasive Networked Systems", CRC Press, March 2008.
- [4] S.Salicone, "Measurement Uncertainty - An Approach via the Mathematical Theory of Evidence," Springer Science+Business Media, LLC2007
- [5] International vocabulary of metrology – Basic and general concepts and associated terms (VIM), 3<sup>rd</sup> ed., JCGM 200: 2012
- [6] B.B.P. Rao, P. Saluia, N. Sharma, A. Mittal, S.V. Sharma, "Cloud computing for Internet of Things & sensing based applications," 6<sup>th</sup> International Conference on Sensing Technology (ICST), pp.374-380, 2012
- [7] X. Wu, X. Zhu, G.Q. Wu, and W. Ding, "Data mining with big data," *IEEE Transactions on knowledge and data engineering*, 26(1), pp.97-107, 2014.
- [8] Davis Vantage Pro2 Weather Station: <http://www.davisnet.com/solution/vantage-pro2/>
- [9] Libelium plug & sense platforms: <http://www.libelium.com/products/plug-sense/>
- [10] ISO/IEC GUIDE 98-3:2008 "Uncertainty of measurement – Part 3: Guide to the expression of uncertainty in measurement (GUM:1995)"
- [11] EA 4/02 M:2013 "Evaluation of the Uncertainty of Measurement In Calibration"

# Implementation and performance evaluation of OpenFlow Data-Centre network testbed

Gojko Gogic, Slavica Tomovic, *Student Member, IEEE*, and Igor Radusinovic, *Member, IEEE*

**Abstract** — This paper presents results of performance evaluation of OpenFlow data center network testbed, developed as part of BIO-ICT research and innovation platform. We show how a large-scale SDN (Software Defined Networking) testbed could be implemented in small scale environment with small number of hardware-based OpenFlow switches and open-source solutions of SDN controllers. Performances of several controller solutions have been compared via Cbench benchmarking tool, and multiple diverse experiments have been conducted in order to investigate impact of the control plane parameters on the network performance. The obtained results indicate that Floodlight controller outperforms ONOS and POX controllers in terms of throughput and processing latency.

**Key words** — Data center, SDN, OpenFlow, testbed.

## I. INTRODUCTION

SOFTWARE Defined Networking is a new networking paradigm that changes traditional network architecture by decoupling the control and management logic from data forwarding elements. Instead, intelligence of SDN network is localized on external controller, which maintains global view of the network state and makes routing and other control decisions. SDN is commonly associated with OpenFlow protocol, a southbound API which defines rules in communication between OpenFlow switches and the controller. Using OpenFlow protocol, the controller can pro-actively or reactively instruct OpenFlow switches how to identify and serve different traffic flows in the network. When OpenFlow switch receives the instructions for specific traffic flow, it is able to handle packets belonging to the flow without further interaction with the controller until validation time for the instructions expires [1].

Although OpenFlow was originally proposed for campus and wide-area networks, there are many quantified arguments that OpenFlow applications are able to increase quality of service and resource utilization in data center (DC) networks [2], [3]. By making control plane

programmable, SDN/OpenFlow enables innovation in network management. However, in order to be production ready, OpenFlow applications need to pass thorough evaluation in experimental environment. The challenge here is to create a large-scale testbed in most efficient manner. The data plane of the testbed may involve hardware and software solution of OpenFlow switches. While open-source solutions are cost-efficient and convenient for educational purposes, they cannot reach performance of the specialized hardware switching devices. Thus, the testbed should be designed by taking into account tradeoff between the implementation expenditures and the expected performance.

In this paper, we present implementation of DC network testbed developed as a part of the BIO-ICT research and innovation platform. We exploited separation of the control plane from the data plane in SDN architecture to logically partition hardware-based OpenFlow switches into a number of virtual OpenFlow switches. From the controller's perspective, each of the virtual switches is considered as independent physical device. In this way, it is possible to create a full network of high-performance DC switches from a single physical device with sufficient number of network interfaces. In order to further increase cost-efficiency of our testbed, we use open-source SDN controllers as well. Considering that controller scalability is one of the main challenges in SDN networks, in this paper we present performance evaluation of three popular open-source controllers: ONOS, Floodlight and POX. As a benchmarking tool we used Cbench software [4]. It can emulate OpenFlow networks with thousands of switches and hosts. In addition, we investigated how the control plane parameters impact network performance by conducting a range of experiments on the testbed: from ping, over TCP transfer of different data sizes, to UDP transfer with different data rates [5]. We found that controller-switch delay degrades throughput more significantly when larger data sizes are transferred. Also, we showed that Java based controllers (Floodlight and ONOS) outperform Python-based POX.

The rest of this paper is organized as follows. In Section II we present our SDN testbed and brief background on the analyzed OpenFlow controllers. Section III explains the experimental methodology we used. Experimental results with corresponding discussion are given in Section IV. The paper concludes in Section V.

## II. SDN TESTBED

In this section we present a cost-efficient design of

This work has been supported by the Montenegrin Ministry of Science and the HERIC project through the BIO-ICT Centre of Excellence (Contract No. 01-1001).

G. Gogic is with the Faculty of Electrical Engineering, University of Montenegro, Džordža Vasiingtona bb., 81000 Podgorica, Montenegro (e-mail: gojkog@ac.me).

S. Tomovic is with the Faculty of Electrical Engineering, University of Montenegro, Džordža Vasiingtona bb., 81000 Podgorica, Montenegro (e-mail: slavicat@ac.me).

I. Radusinovic is with the Faculty of Electrical Engineering, University of Montenegro, Džordža Vasiingtona bb., 81000 Podgorica, Montenegro (e-mail: igorr@ac.me).

OpenFlow DC network testbed. To implement data center topology with 20 network nodes (Fig. 1), we have used only two hardware-based OpenFlow switches: Pica8 3295 and HP Procurve 6600, as well as four software-based OpenFlow switches.

Software solutions of OpenFlow switches are implemented by running Open vSwitch (OVS) software on HP Z620 Workstations with Centos 6 Operating System. These switches are colored in blue in Fig. 1.

Pica8 is white box L3 switch that can run OVS mode in order to support OpenFlow. Like other Linux-based machines with OVS installed, Pica8 could be logically partitioned into multiple independent OpenFlow switches with dedicated physical interfaces. For our purpose, Pica8 is virtualized into twelve virtual switches (colored in orange in Fig. 1). These virtualized switches are managed in the same way as OVS switches, but with the possibility to store the flow entries in the hardware flow tables.

HP ProCurve 6600 switches cannot be virtualized in the same way as Pica8. While white box switches allow creation of multiple virtual switches and assignment of physical ports to them, in order to create multiple virtual OpenFlow switches from HP ProCurve 6600, separate VLAN (Virtual Local Area Network) has to be configured for each virtual switch. OpenFlow virtual switches are independent and have their own configuration and connection towards the OpenFlow controller. In our testbed, HP ProCurve is logically separated into four OpenFlow virtual switches, colored in green on Fig. 1.

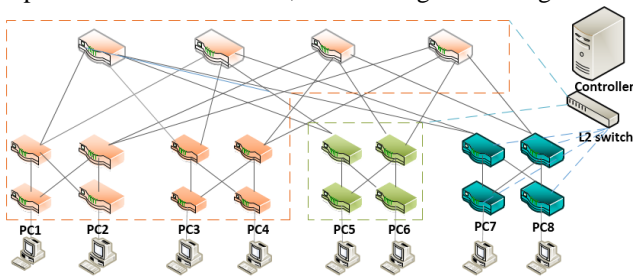


Fig. 1. SDN Data Center topology

The intelligence of SDN network resides on the controller device, which instructs all OpenFlow switches in its area of responsibility. Thus, it is necessary to run controller on machine with respectable performances. The PC that is dedicated for this purpose in our testbed is HP Z620 Workstation with Ubuntu 16.04 OS. It's provided with 16GB of main memory and Intel Xeon E5-2640 CPU with 24 physical cores and 2 threads per core. On this machine we installed open-source controllers: Floodlight, ONOS and POX.

ONOS is Java-based open-source SDN controller developed at Stanford University. ONOS is designed with distributed architecture in order to meet the needs of service providers for scalability, high availability and performance. The system offers REST API, CLI and an extensible, dynamic web-based GUI. It also supports multi-threading.

Floodlight is multi-threaded Java-based SDN controller. It is considered as one of the enterprise class OpenFlow

controllers that is easy to use, build and run. It has been supported by developers all around the world, and Big Switch Networks Company.

POX is single-threaded user-friendly SDN controller, whose components are written in Python. It is widely used in education and research as a learning and prototyping tool. The components of POX controller are programmed to implement different networking functions.

### III. EXPERIMENTAL METHODOLOGY

In this section we explain experimental methodology we used to evaluate performance of our testbed.

Considering that processing packets at high rates with minimum latency is a key requirement for any SDN controller, we used throughput and processing latency as key indicators of SDN controller's performance. To compare performance of ONOS, Floodlight and POX we used Cbench benchmarking tool. Cbench emulates OpenFlow switches which communicate with controller. As input arguments it takes a number of switches that should be emulated, number of hosts per switch and the controller's address. It supports two working modes: latency and throughput mode. In latency mode, it sends PACKET\_IN messages to the controller. These messages are used by OpenFlow protocol to inform the controller when a packet received by a switch does not match any entry in the flow table. As a response, controller generates FLOW\_MOD message, which installs a new entry in the flow table. If a controller cannot make routing decision, (e.g. because packet destination is unknown), it will just instruct switch to drop or flood a packet via PACKET\_OUT message, depending on the control application used. When Cbench works in latency mode, it measures the time taken to handle a single packet. The next PACKET\_IN message is generated only after a corresponding response is received from the controller. In throughput mode, the emulated switches send as many PACKET\_IN messages as possible to the controller, making sure that the controller always has messages to process [6]. The result of both, throughput and latency tests, are expressed in number of received responses per second. In order to examine how multi-threading impacts controllers' performances, we conducted multiple experiments where controllers have been run with different number of threads (using *taskset* Linux command).

While Cbench is very compelling tool for testing controller scalability, it just emulates bunch of independent OpenFlow switches, which are not connected in any specific network topology. Thus, in order to get indication regarding the performance of our testbed (Fig. 1), we conducted several additional experiments. Our first experiment involved running ping between hosts reachable by 1, 3 and 5 hops. The purpose of this experiment was to contemplate the effect of the switch-controller communication on ping RTT.

In second set of experiments we measured the impact of controller-switch RTT delay on transfer time of TCP flow. For this purpose, we installed simple client-server



application on PC1 and PC8. At the beginning of the experiment we ensure that flow tables of OpenFlow switches are empty. Then we run the client-server application that measures time from the moment when TCP connection is established until the connection is closed [7].

In third set of experiments we generated UDP traffic and measured packet loss during the first second of UDP transfer. Traffic is generated between hosts PC1 and PC8 via Iperf software. We varied the rate of UDP traffic.

Each of the experiment explained above is repeated under conditions of increased controller-switch delay of 5ms and 10ms. The additional propagation delay between the controller and switches is emulated by using Linux *tc* command. This allows one to observe how controller-switch delay degrades quality of the network service.

At the beginning of each experiment, all ARP caches and *flow tables* in the network were empty.

#### IV. THE RESULTS

Fig. 2(a) shows the results of Cbench latency test when ONOS, Floodlight (FDL) and POX controllers were run with single CPU thread, for a different number of emulated switches. The results indicate that Floodlight and ONOS controllers process incoming PACKET\_IN messages significantly faster than POX controller. It should be noted that all controllers were running simple L2 learning control script for the sake of fairness and simplicity. From the figure, one can observe that ONOS performs slightly better than Floodlight when number of switches is smaller than 64. For larger number of switches Floodlight outperforms ONOS. Difference in performance between ONOS and Floodlight on one side, and POX on the other side, are most obvious from the Fig. 2(b), which shows results of the Cbench latency test when 5 CPU threads were used to run controller software. Since POX does not support multi-threading, its performance did not improve at all. Latency performance of ONOS and Floodlight increased drastically.

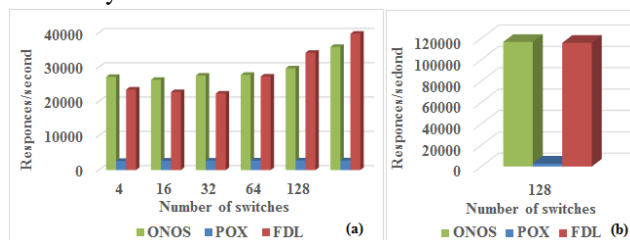


Fig. 2. Results of the latency tests: (a) Single-threaded; (b) Multi-threaded

The results of the Cbench throughput test are shown in Fig. 3. One can see that for less than 128 switches, ONOS and Floodlight performed almost the same. In scenario with 128 switches, there is a sudden drop in throughput of Floodlight controller. Again, the worst results are achieved with POX. Thus, one can conclude that POX is not suitable for large-scale production environments. The results from Fig. 3(b) go in line with this claim. It can be observed how Floodlight and ONOS can exploit good hardware capabilities of the host machine (multiple

threads) to speed up processing.

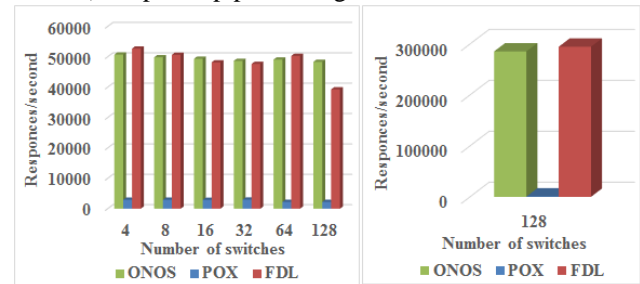


Fig. 3. Results of the Cbench throughput test: (a) Single-threaded; (b) Multi-threaded

In the rest of the section we discuss results of the experiments performed on the DC network testbed. Fig. 4 shows the results of the ping experiment for different values of switch-controller delay. The results present RTT measurements for the first (Fig. 4a) and second (Fig. 4b) ICMP Echo Request/Reply transfers. Ping was run between PC5 and PC6 (one-hop route), between PC7 and PC8 (three hop route) and between PC1 and PC8 (five hop route). The results from Fig. 4(a) refer to RTT of the first ICMP Echo Request/Reply transfer in cases with and without additional delay on controller-switch link. Since ARP tables of hosts and flow tables of switches were empty at the beginning of the experiment, RTT for first ICMP request is approximately equal to three times controller-switch RTT delay increased by packet processing time. This is because switch seeks instructions from the controller three times during the transfer period of first ICMP Echo Request. Firstly, access switch encapsulates ARP request from the first host in PACKET\_IN message. The controller responds by flooding ARP packet, because location of destination is still unknown. Then, access switch of the destination host sends ARP reply in the form of PACKET\_IN. Now location of the packet destination is known, thus, controller installs flow rules for ARP packets. The third interaction with controller happens when ICMP request reaches the access switch of the sending host. Controller installs routing rules as a response [5].

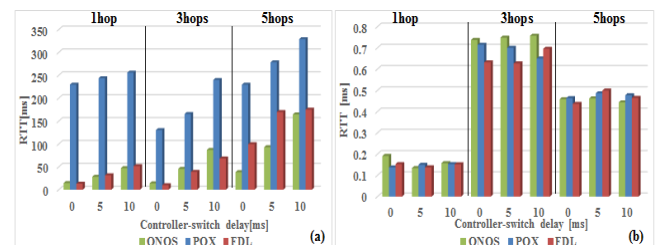


Fig. 4. Comparison of the average RTT for: (a) first ICMP packet, (b) second ICMP packet

Fig. 4(a) just confirms the conclusion of our previous Cbench experiments that POX controller introduces the highest processing latency. The results from Fig. 4(b) are more interesting because they give indication regarding the data plane performance. In this case switches have no need to communicate with the controller (the route already exists), so RTT depends only on delay introduced by the

data plane. It can be noted that delay in 3-hop experiments is larger than in 5-hop experiments. This could be attributed to the fact that 3-hop route goes over 3 software OpenFlow switches (OVS), which have more limited performance than Pica8 and HP switches.

Fig. 5 shows results for the client-server experiments explained in Section III. The server process was running on PC8, while PC1 was used as a client. As a performance indicator we used transfer time of the data file. In experiments we varied size of the file and initial delay on control links. In this testing scenario, initial interaction with the controller slows down the procedure of TCP handshaking and route setup. In case of Floodlight and ONOS controllers, the additional propagation delay on control links was the main cause of the increased transfer time of TCP flow. On the other side, transfer time in case of POX controller was large even when propagation delay between controller and switch is negligible [5].

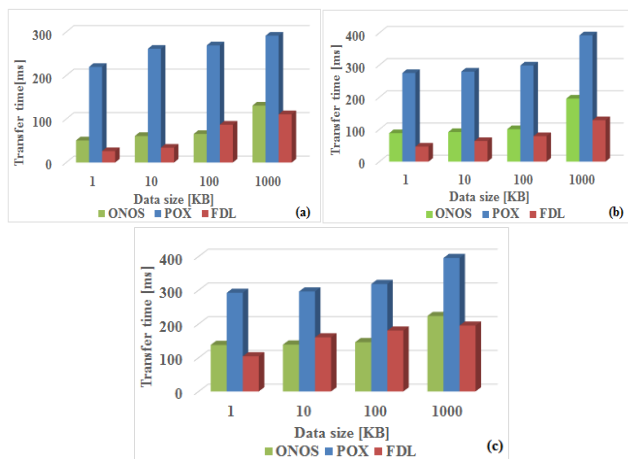


Fig. 5. Average TCP transfer time for different file sizes in experiments with added controller-switch delay of (a) 0 ms, (b) 5 ms and (c) 10 ms

The results of UDP experiment are presented in Fig. 6. The graphs show packet loss rate during the first second of UDP transfer in function of the data rate and added controller-switch delay.

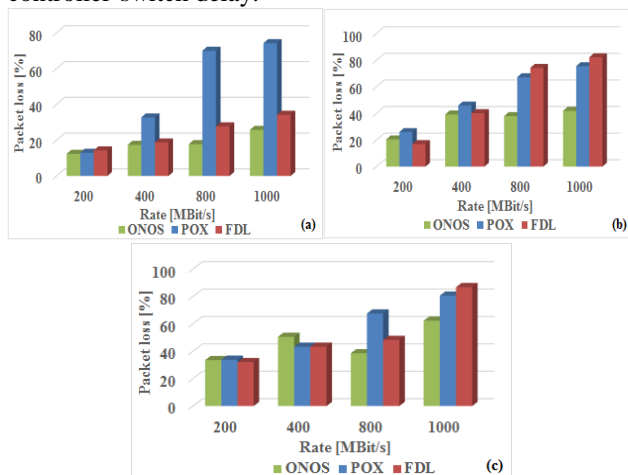


Fig. 6. Packet loss rate during the first second of UDP transfer for different data rates in experiments with added controller-switch delay of (a) 0 ms, (b) 5ms and (c) 10 ms.

Considering that UDP doesn't use handshaking

procedure, when first packet of UDP flow enters the network, there are no routes installed and packet is directed towards the controller. The actual problem is that not only the first packet is directed to the controller, but also all subsequent packets until the flow rule is installed. In case of large and high-speed UDP flows, packets can overburden the controller and get lost. The results from Fig. 6 show that packet loss rate increases with the UDP data rate. For data rates higher than 800 Mb/s the network gets unstable, with very high level of packet loss rate. This network bottlenecks are not only software switches but also HP ProCurve switch, which can't support data rate over 850 Mb/s.

## V. CONCLUSIONS

In this paper, we explained cost-efficient implementation of DC OpenFlow network testbed and presented results of its performance evaluation. The data plane of our testbed consists of two hardware and four software OpenFlow switches. We have logically partitioned hardware switches in 16 virtual OpenFlow switches, thereby creating compact and high performance testbed connected in large-scale DC network topology. In order to determine the most suitable OpenFlow controller, we tested three open-source solutions: ONOS, Floodlight and POX. Performances of the controllers in terms of throughput and processing latency have been tested via Cbench software, which allows emulation of thousands OpenFlow switches and host connected. The results show that ONOS and Floodlight outperform POX controller significantly. In scenarios with large number of emulated OpenFlow switches, Floodlight introduces the lowest processing delay, but cannot cope with large amount of switch queries as good as ONOS. More complete insight into the testbed performance is obtained through ping, TCP transfer and UDP transfer experiments. In our future work, we will use the testbed for experimental evaluation of traffic engineering applications.

## REFERENCES

- [1] S. Tomovic, K. Yoshigoe, I. Maljevic, and I. Radusinovic, „Software-Defined Fog Network Architecture for IoT.,” *Wirel. Pers. Commun.*, Vol. 92, No1, pp. 181-196, January 2017.
- [2] A. R. Curtis, J. C. Mogul, J. Tourrilhes, P. Yalagandula, P. Sharma, S. Banerjee “DevoFlow: Scaling Flow Management for High-Performance Networks,” SIGCOMM, pp. 254-265, August 2011.
- [3] M. Al-Fares, S. Radhakrishnan, B. Raghavan, N Huang, and A. Vahdat, “Hedera: dynamic flow scheduling for data center networks,” In Proc. of the 7th USENIX conference on Networked systems design and implementation, CA, USA, 2010.
- [4] C. Laissaoui, N. Idboufker, R. Ellassali, K. El Baamrani, „A measurement of the response times of various OpenFlow/SDN controllers with Cbench, AICCSA 2015, pp. 1-2.
- [5] H. D. Vu and J. But, “How RTT Between the Control and Data Plane on a SDN Network Impacts on the Perceived Performance”, International Telecommunication Networks and Applications Conference (ITNAC), pp. 1-4, 2015.
- [6] O. Salman Imad H. Elhadj Ayman Kayssi Ali Chehab, “SDN Controllers: A Comparative Study,” Lebanon: American University of Beirut, pp. 1-6, 2016.
- [7] A. Shalimov, D. Zuikov, D. Zimarina, V. Pashkov, R. Smeliansky, “Advanced Study of SDN/OpenFlow controllers,” Proc. of the 9th Central & Eastern European Soft. Eng. Conference, 2013, pp. 3-9.

# A Brief Overview of IoT Architecture and Web Platforms

Božidar Škrbić, Slobodan Đukanović, *Member, IEEE*

**Abstract** — Internet of Things (IoT) is an increasingly popular topic nowadays. However, up to date, there have not been created any definitive standards when it comes to the architecture of an entire IoT infrastructure. Several reference models have been proposed making it a significant step towards making the first IoT standard. One of those reference models is presented in this paper, which briefly overviews some of the most popular IoT web platforms, as well as the web platform created within the BIO-ICT project.

**Key words** — Analysis, Data, Internet, Monitoring, Overview, Platform, Things, Web

## I. INTRODUCTION

OVER the years, the number of devices that have an access to the Internet has increased dramatically, with an estimation of around 8.7 billion devices by the year 2012 [1]. Due to the ever decreasing size and complexity of hardware needed for establishing connection with the global network, Internet is no longer limited to PCs only. Today, hardware dedicated to the communication with other devices over the Internet can be found in a variety of devices ranging from smartphones to household appliances and cars. The diversity of the devices is predicted to keep expanding in the future.

What made the expansion and popularity of Internet of Things possible in the first place are Wi-Fi and mobile (WAN) networks. The needlessness of cords or any other form of physical connection to enable Internet access for a device, as well as network modularity in the sense of dynamic increase and decrease of number of devices present on it without requiring prior configuration of the network (plug-and-play principle) are crucial for moving internet access from traditional PCs to the range of other devices.

The constant development of mobile network, primarily 3G and 4G broadens the network territorial coverage, and allows the implementation of Internet of Things (IoT) in a wider geographical area, thus increasing the possibilities of the IoT incorporation. Today, most applications belong to the projects regarding city and environment monitoring,

This work has been supported by the Ministry of Science of Montenegro and the HERIC project through the BIO-ICT Centre of Excellence (Contract No. 01-1001).

Božidar Škrbić and Slobodan Đukanović are with Faculty of Electrical Engineering, University of Montenegro. Address: Džordža Vašingtona bb, 81000 Podgorica, Montenegro. Phone: {+382 69 202 977, +382 67 266 109}. E-mails: {bozidar.skrbic@gmail.com, slobdj@ac.me}. Web: www.etf.ac.me.

healthcare, security, smart homes, farming and agriculture and industry, among others [2]. Generally speaking, the versatility and low costs of IoT makes it popular among both enterprises and individuals throughout the world.

## II. IOT ARCHITECTURE

The IoT concept assumes collecting and sending various forms of data, in large amounts from one machine to another, and eventually displaying the data to a human. Therefore, it is important for it to have a strictly defined and standardized architecture which would allow machine-to-machine interaction as well as machine-to-human, and the other way around. Currently, this has not been covered by the standards yet. However, there are several proposed reference models. A model given by CISCO [3] defines seven architectural layers, listed bottom to top:

- physical devices (things)
- communication units
- cloud/fog computing
- data accumulation and storage
- data abstraction
- applications
- collaboration and processes

It should be noted that the reference model does not necessarily imply that all the levels are implemented on physically separate hardware or places on different locations. It is often the case that several levels are all incorporated on a single device, depending on the way IoT is applied within a certain project.

The bottommost layer comprises of end-point devices which, as already stated, could be of various types and purposes. The diversity grows daily, and there are, basically, no rules when it comes to size, application or location as long as a device has an access to the Internet.

Communication units are hardware components dedicated to establishing connection between layer one devices or between layer one devices and data centres (belonging to the third layer) and reliable data transfer across the network. Generally, all GPRS shields and microcontroller communication modules would fall under this layer. The IoT uses standard TCP/IP architecture for transferring the data. When it comes to application layer protocols, there are multiple choices. However, among the most common are standard HTTP (or, more often, its secure version HTTPS), MQTT (MQ Telemetry Transport) and CoAP (Constrained Application Protocol).

Level three and four are in charge for data acceptance and aggregation, filtering, processing (level three) and data storage and accumulation (level four). Basically, data

flows are first analysed and processed at level three, and then the moving data is converted to stationary data at level four. Due to the data accumulation, further analysis and processing does not have to occur at network speed.

Data storages do not necessarily have to share the same location and could be physically separated. Even the data which belongs to the same context can be stored in different formats. Hence prior to passing the data to the application level, it is necessary to do certain pre-processing, which is done at the data abstraction level (level five). The data from various sources (i.e. data accumulations) are synchronized and converted to a format appropriate for further use in the application level.

Application level (level six) is the level which has the task of interpreting the data. The term “application” is not strictly defined by the IoT Reference Model. Therefore, it does not place any constraints in terms of application appearance, functionality and complexity. There exists a variety of applications currently being used for various purposes, including monitoring, remote device controlling, security etc. Furthermore, the way the applications are developed varies as well. They can run on different platforms, devices, operating systems, they can be developed using different tools and programming languages, and so on.

The whole purpose of the IoT concept is, in the end, facilitating everyday human activities. Thus, the Reference Model has to include people and processes done by the people, since all the data acquired by an IoT network mean nothing by themselves unless they trigger a human performed action by giving the right data, at the right moment, in the best way possible. Having this in mind, it is only fair to abstract these actions as a separate level in the IoT architecture stack, above the application level. Therefore, this level would comprise of all the interpersonal communication, collaboration and work done upon gathering the valuable information.

As always, an important aspect of every system is security, and such is the case with IoT as well. Since security in general is an elaborate topic, going into every detail would fill out a considerable number of pages. It suffices to say here that it is highly recommended to implement a certain level of security on each layer of the reference model. It is necessary to secure all physical devices and all processes as well as the communication routes thus creating an extremely secure system spanning across all the levels of the model.

### III. REPRESENTATIVE PLATFORMS

With the increased popularity of Internet of Things, the number of available web platforms became larger than ever. ThingSpeak, Microsoft Azure IoT Suite, Amazon Web Services IoT, Kaa, Predix, Carriot, ThingWorx, Aeris etc. are only some of the existing solutions. Here, the focus is placed only on the first three.

#### A. ThingSpeak™

ThingSpeak is an open source web application dedicated to IoT. It is available as a free service for non-commercial use, however it is required to buy one of licenses for using

it in larger projects or commercial applications. It was launched in 2010 by the ioBridge company [4]. It also offers an API for integrating ThingSpeak into other applications.

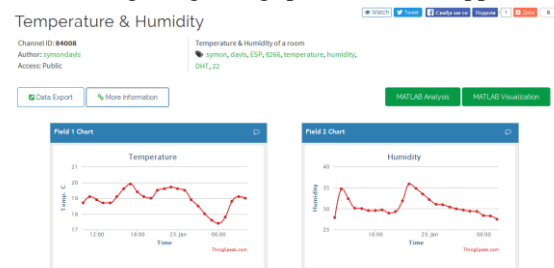


Fig. 1. ThingSpeak data charting

ThingSpeak has full support from Mathworks and has integrated MATLAB thus allowing users to use it for analysing and visualising the uploaded data. Even ThingSpeak documentation was made as a part of MathWorks documentation [5].

As stated on their homepage [6], core ThingSpeak functionalities are data collection, data analysis and action triggering. It supports various types of hardware for acquiring the data from the sensors, most notably Arduino, Raspberry PI, which collect the data from the sensors and send it over the internet using one of the application level protocols. Until recently, only HTTP was supported. However, as of December 2016, it also supports MQTT.

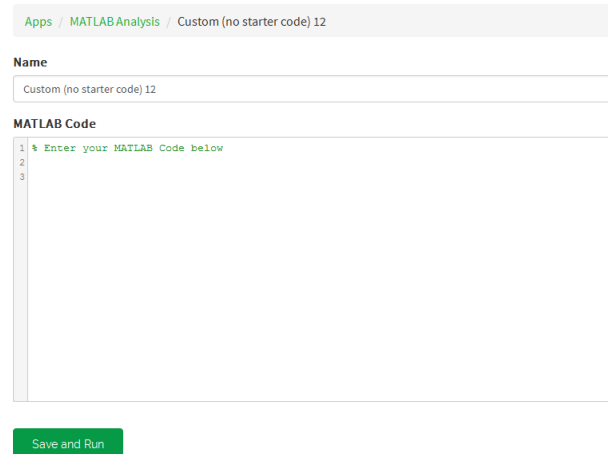


Fig. 2. ThingSpeak MATLAB editor

The hardware which collects and sends the data is abstracted in the form of a user channel, and a single set of data acquired from a single sensor on that hardware is represented by a field. Thus, when creating a channel, it is first given a name, and then all the fields are listed, with the limit of eight fields. Upon its creation, the user is provided with unique read and write keys. Write key is used as a unique identifier which enables data writing to that channel to anyone in possession of that code, through a HTTP Post request. Read key is the same kind of identifier, the only difference being that it is used for accessing the data of a private node without logging to the owner’s account (possibly from an application, for example). For security purposes, ThingSpeak allows recreation of the keys for specified channel at any point of time, if the owner feels his data may be compromised [5].

Once the channel has been created, a remote hardware

may be configured to send its data periodically which are charted on the webpage with charts being constantly updated with new data (Fig. 1). Apart from charting, the data may be subjected to MATLAB analysis either by using one of the prewritten scripts (there are around dozen of such scripts available), or by writing your own directly in a web editor, using MATLAB syntax (Fig. 2). MATLAB can also be used for data visualisation and plotting graph using all the available built in MATLAB functions [5].

ThingSpeak has a large online community. There are numerous tutorials on the Web with projects incorporating ThingSpeak or its API. There is also a forum available, where people discuss topics regarding ThingSpeak or IoT in general [7]. A user may also connect his Twitter account to ThingSpeak account and use it to receive warnings when, for instance, value of certain data jump over a predefined threshold [5].

### B. Microsoft Azure IoT Suite

Microsoft Azure is a cloud computing service developed by Microsoft in 2010. As a part of this service an IoT Suite has been established. In comparison to ThingSpeak, Azure IoT Suite is somewhat more complicated and possesses functionalities that are not present in other Web platforms. Being a part of Azure cloud computing service, Azure IoT Suite has all its computing capabilities at its disposal. Therefore, apart from basic data charting and monitoring, the users of Azure IoT Suite are able to employ the entire computing capabilities of Azure cloud computing service, including the built-in machine learning system. That way they can perform extremely advanced analysis on their data and even configure an action to be triggered when certain condition is fulfilled (i.e. when an anomaly has been detected by the system). Also, Azure IoT Suite supports two-way communication between the platform and the devices, thus allowing device remote configuring and control. For that purpose it relies on several protocols: HTTP, MQTT and AMQP (Advanced Message Queuing Protocol) [8].

Azure IoT Suite provides its service through two so-called solutions: remote monitoring (Fig. 3) and predictive maintenance. As the name suggests, remote monitoring is dedicated to those who have a remote hardware with several sensors attached and in terms of usage is similar to the service ThingSpeak offers. It is dedicated to the people who wish to have an insight on the measured parameters of a distanced hardware and sensors. Among other options, it provides data charting (here named data telemetry), geographic mapping of the devices and the ability to define a rule for alarm triggering. Sending commands to the devices (such as sending a ping to a device to check if it is responding) is possible as well. A neat feature, especially for the newbies in the world of IoT device simulation where a user does not necessarily need to connect physical hardware to the platform when performing tests, and thus makes it easy to do a proof of concept prior to building a whole sensor infrastructure [8].

Predictive maintenance solution is basically remote monitoring backed with more advanced algorithms. It

combines device telemetry with machine learning algorithms present in Microsoft Azure cloud service. A user is able to quickly configure a device and analyze its data in real time. Furthermore, the IoT platform could also predict failures in certain parts of the monitored system, sending timely reports to the people responsible for the system, or even predict an approximate moment when the failure will happen. Also, the platform can be programmed to make the system react in a certain way if the failure happens (for instance trigger an alarm or turn on back-up system) [8].

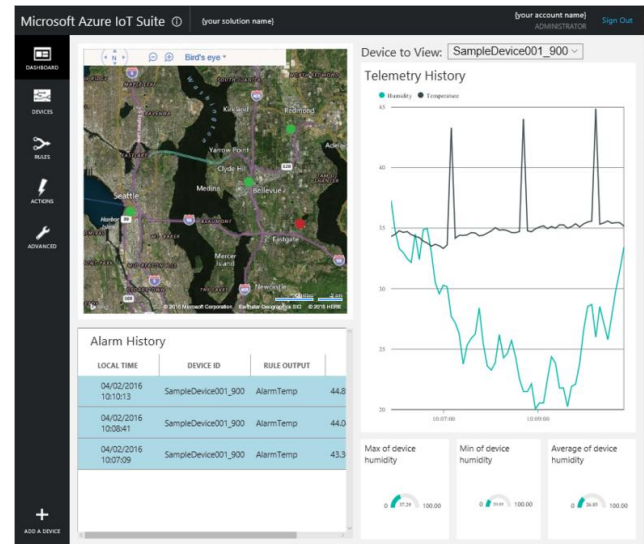


Fig. 3. Microsoft Azure IoT Interface

Overall, Azure IoT Suite is somewhat more developed than ThingSpeak, but also proportionally more expensive. With its functionality, it is perhaps better suited for the enterprises, even though it could be equally as useful for individuals.

### C. Amazon Web Service IoT

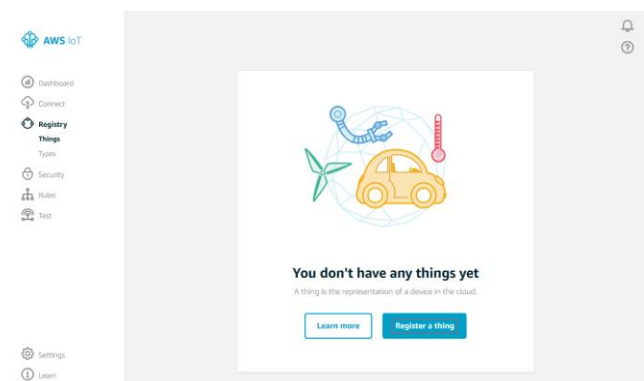


Fig. 4. AWS IoT Interface

Amazon Web Service IoT (AWS IoT) is a platform that serves as an extension of Amazon Web Service, a cloud computing service, similar to Microsoft Azure. Just as Azure IoT Suite, AWS IoT is also able to use all the underlying cloud services. Therefore, the functionalities of these two platforms are similar to a great extent (i.e. data charting, real-time data analysis and processing, possibility of using machine learning algorithms etc.). A two-way

communication is available on this platform as well, thus it can be used to remotely configure and control the devices. The communication relies on HTTP and MQTT protocols [9].

AWS IoT platform sees a device as a device gateway through which a bi-directional communication is established. The record of all devices is kept in a so called device registry where user can add new “things” or modify the existing (Fig. 4). Furthermore, AWS IoT has a Rules Engine which allows continuous processing of data as they are being sent by devices. Also, developers of other applications (mobile or web applications regardless) may use an existing API to access the devices connected to AWS IoT, or the data these devices have sent [9].

More or less, it shares the same functionalities with Microsoft Azure IoT and has an interface and device management system that closely resembles Azure’s remote monitoring solution (Fig. 4). However, AWS IoT documentation does not explicitly mention data prediction using machine learning. However, since it does have AWS machine learning algorithms at disposal, and it is able to send instructions to the devices, it is safe to assume that such a functionality can be implemented in AWS IoT as well [9].

#### IV. BLUELEAF

BlueLeaf is an IoT platform developed within BIO-ICT project for the purpose of collecting, storing, processing and visualizing the data gathered from various devices, regardless of their structure provided that they send the data using a supported application-level protocol [10], [11]. At the moment, BlueLeaf operates only using the HTTP protocol, but the MQTT protocol will be implemented soon.



Fig. 5. BlueLeaf data charting

The platform contains numerous neat and practical functionalities, such as real time data charting (Fig. 5), node mapping using Google maps, data calibration etc. Users are also allowed to create their own accounts for the purpose of better node management. The platform also allows users to download the data from a node in form of a JSON file or an XLS/CSV table.

Currently, development is aimed at integrating Octave, R and Python programming languages with the BlueLeaf platform. At the moment, the users can choose from a variety of predefined analytic tools (data plotting,

histogram plotting, calculating data parameters such as mean, maximum, minimum, median value or standard deviation, calculating data correlation from a single node and performing data clusterization) as well as uploading custom script (Fig. 6). At the moment, only Octave and R language have been integrated, but in the upcoming time Python will be added as well.

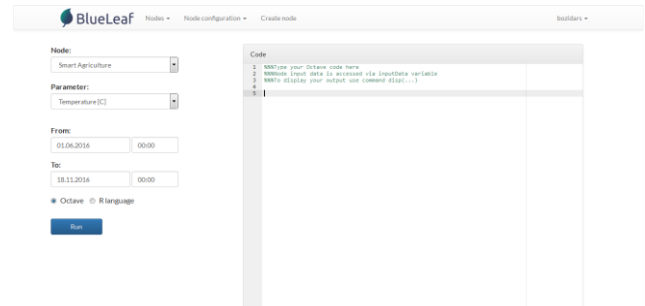


Fig. 6. BlueLeaf code editor

Having this functionality is what makes BlueLeaf stand out from most available IoT platforms. Additionally, it is planned to make the platform exportable, so it can serve as a standalone project which can be installed on other servers. That way, a potential user would not have to necessarily rely on the BIO-ICT data centre.

BlueLeaf is currently being used actively in BIO-ICT centre of excellence for soil and water monitoring. However, it is still being developed and currently not available for commercial use.

#### CONCLUSION

The paper presents a brief overview of general IoT architecture, as well as most popular web platforms and a new solution created within the BIO-ICT project. It could serve as a good starting point to all those who wish to do more research in the field of IoT.

#### REFERENCES

- [1] R. Soderbery, “How Many Things Are Currently Connected To The ‘Internet of Things’ (IoT):” <http://www.forbes.com/sites/quora/2013/01/07/how-many-things-are-currently-connected-to-the-internet-of-things-iot/>
- [2] “50 Sensor Applications for a Smarter World”: [http://www.libelium.com/resources/top\\_50\\_iot\\_sensor\\_applications\\_ranking/](http://www.libelium.com/resources/top_50_iot_sensor_applications_ranking/)
- [3] “The Internet of Things Reference Model”: [http://cdn.iotwf.com/resources/71/IoT\\_Reference\\_Model\\_White\\_Paper\\_June\\_4\\_2014.pdf](http://cdn.iotwf.com/resources/71/IoT_Reference_Model_White_Paper_June_4_2014.pdf)
- [4] ioBridge blog: <http://blog.iobridge.com/2014/12/thingspeak-plus-third-party-hardware-an-alternative-to-iobridge-hardware/>
- [5] ThingSpeak documentation: <https://www.mathworks.com/help/thingspeak/>
- [6] ThingSpeak IoT Web Platform Homepage: <https://thingspeak.com/>
- [7] ThingSpeak Community Forum: <http://community.thingspeak.com/forum/>
- [8] Microsoft Azure IoT Suite Documentation: <https://opbuildstorageprod.blob.core.windows.net/output-pdf-files/en-us/Azure.azure-documents/live/iot-suite.pdf>
- [9] Amazon Web Services IoT Documentation: <https://aws.amazon.com/documentation/iot/>
- [10] “A Private IoT Cloud Platform for Precision Agriculture and Ecological Monitoring”, Milija Bajčeta, Petar Sekulić, Božo Krstajić, Slobodan Đukanović and Tomo Popović, ICETAN 2016
- [11] “An IoT Cloud Solution based on open source tools”, Tomo Popović, Milutin Radonjić, Žarko Zečević, Božo Krstajić, IT’16

# Cloud Based Solution for Automatic Image Mosaicking and Georeferencing

Žarko Zečević, *Member, IEEE*, Tomo Popović, *Senior Member, IEEE*, Božo Krstajić, *Member, IEEE*

**Abstract** — Unmanned air vehicles (UAVs) provide an economical means of imaging large areas at far lower cost than satellites. Also, in some applications such as precision agriculture, the satellite imagery does not provide satisfactory spatial and spectral resolution. Since UAVs fly at low attitudes, in order to cover a large terrain, thousands of pictures should be taken. To provide a unified view of the area, these images should be aligned into a one composite panorama. A cloud application for automatic image stitching and georeferencing, which use several open source image and digital terrain model (DTM) processing tools, is described here. The presented solution is compared to the commercial cloud and desktop software solutions in the sense of image processing time.

**Keywords** — cloud, mosaicking, drone, georeferencing, orthomosaic, drone, Unmanned air vehicle (UAV), precision agriculture, mapping.

## I. INTRODUCTION

Accurate and timely information on crop conditions obtained during the growing season is very important for precision agriculture and crop management applications, [1, 2]. Aerial color and color-infrared photography provides spatial and spectrally derived terrain parameters that can be used for crop classification and mapping [3], crop water stress detection [4], crop forecasting and yield prediction, crop status and condition, soil erosion detection [5, 6], nutrient efficiency and photosynthetic pigment content [7], etc.

Current satellite-based imagery has limited applications in agriculture due to low spatial and spectral resolution and the large revisit time. Since precision farming requires detailed information about terrain parameters, a spatial resolution of a few meters is required and some specific applications may need higher resolutions [8]. Therefore, the images acquired by aircraft have been more widely used in both research and commercial operations in agriculture. Typically, the manned aircraft are equipped with large-format cameras, which are expensive, compared to other types of imagery [8].

Unmanned air vehicles (UAVs) are practical and have

Ž. Zečević is with the Faculty of Electrical Engineering, University of Montenegro, Džordža Vašingtona bb, 81000 Podgorica, Montenegro (e-mail: zarkoz@ac.me).

T. Popović, Faculty of Information Systems and Technologies, University of Donja Gorica, Oktoih 1, 81000 Podgorica, Montenegro (e-mail: tomo.popovic@udg.edu.me).

B. Krstajić is with the Faculty of Electrical Engineering, University of Montenegro, Džordža Vašingtona bb, 81000 Podgorica, Montenegro (e-mail: bozok@ac.me).

potentially lower cost, but also have limited payload capacity, so the light-weight cameras are required. They are cheap enough for more frequent image updates, which enable close monitoring of crop development. If UAV flies close enough to the ground it can provide more detailed photos, even with a resolution of 2cm/pixel [8].

Because a small UAV flies at a low altitude-typically under 100 meters, the camera has a limited field of view. To cover a reasonable sized area, the camera must take thousands of photographs. To provide a unified view of the area, these photographs must be aligned and combined into a one composite image. The process of joining two or more independent images to single image is called mosaicking, and the resulting image is called a mosaic [9]. Finally, a process of assigning a real-world geographical coordinates to each pixel of mosaic is called georeferencing [9-11].

There are a number of commercial desktop tools that can be used for professional aerial mapping applications in precision agriculture [12, 13]. Standalone photogrammetry software is expensive and can require considerable processor and storage resources. In recent years cloud computing gains traction and many mapping applications are moving to the cloud [13]. For the ordinary user, the cloud-based services are simpler to use, cheaper and also can provide all the necessary information to make timely decisions. It frees them from buying expensive photogrammetry software and computing and storage resources.

In this paper a cloud application for automatic image mosaicking and georeferencing, which will be referred to as DroneMapper, is described. Several open source tools for image and digital terrain model (DTM) processing are used here. DroneMapper lets users to upload aerial photos, stitch the images and host created maps. The presented solution is compared to the commercial cloud and desktop software solutions in the sense of image processing time.

The paper is organized as follows. After an introduction, an overview of existing standalone and cloud software products for image mosaicking is given. In the third section structure of the proposed cloud application is described. Finally, some image processing results and conclusion are given.

## II. OVERVIEW OF AUTO-MOSAICKING SOFTWARE

The simplest way to create a mosaic from aerial imagery is to use photo-stitching software, which assembles a series of overlapping aerial images into a one-composite image [13]. When flying a UAV, latitude should be considered.

Higher altitude results in lower resolution, but it allows the UAV to photograph a larger area. Higher-altitude imagery produce less relief displacement, and also can help at reducing the distortion found in images of buildings and other objects on the ground. On the other side, the lower-altitude photography increases the ground sample distance (GSD) and the mapping quality, but also the time required to photograph and processes specified area [11]. The UAV flying path should be carefully planned. In order to obtain good mosaicking results, the images should be overlapped by minimum 20% [11].

There are a number of commercial photogrammetry software products that can be used for professional aerial mapping applications. Two most popular are Pix4D [14] and Agisoft PhotoScan [15]. Pix4D mapper includes the ability to generate 3D textured models, digital surface models (DSM) and NDVI maps with greater accuracy compared to other mapping software. Pix4D AG is the special drone mapping version of Pix4D, designed for precision agriculture. Aerial crop analysis, generating index maps and automated processing templates are some of available features in Pix4D AG [14]. The second popular software Agisoft PhotoScan, includes features as photogrammetric triangulation, 3D model generation and texturing, dense point cloud generation, etc. Also it has built tools to precisely measure distance, area and volume of objects [15].

Standalone photogrammetry software is expensive and can require considerable processor and storage resources, especially in the cases when processing a large number of high-definition images. In recent years cloud computing gains traction and many conventional mapping applications are moving to the cloud [16]. Maps Made Easy (MME) is a cloud application that lets users upload aerial photos, stitch the images and host created maps. Maps Made Easy supports NDVI processing and provides some advanced measurement tools such as a measuring of the object volumes. MME uses standard methods for georeferencing and works with existing GIS solutions [16].

Open-source software is another possibility for aerial imagery processing [13]. MapKnitter from Public Lab is free and open source tool for manual combining and positioning images in geographic space into a composite image map [17]. Another open source tool OpenDroneMap uses GPS data to automatically create orthophotos and texture models from drone imagery [18]. Visual Software from Motion (VisualSFM) is a GUI application for 3D reconstruction using structure from motion (SFM). [19].

### III. SYSTEM DESCRIPTION

A global outline of the cloud application for automatic mosaicking and georeferencing is shown in Figure 1. As it can be noticed, the DroneMapper consists of two main parts: the web interface and the server side that is placed in the cloud. By using the web interface, user can upload photos acquired by UAV, start map processing or preview already created maps. UAV images are uploaded to the

cloud storage, where they are processed by different open source tools in order to get mosaicked map or GeoTiff file.

The database consists of two tables – the first one contains basic user information, while the other one contains information about created maps.

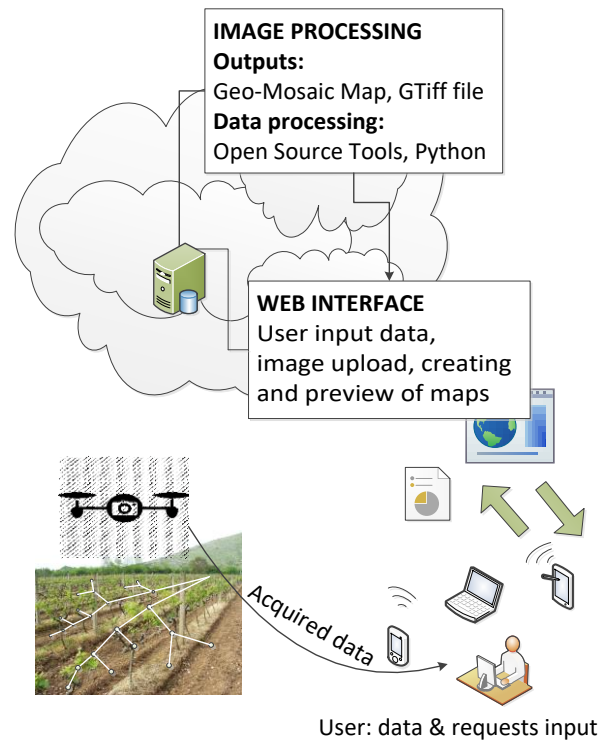


Fig 1. A global outline of DroneMapper

The web based interface for image upload and map preview is designed as multi-user. It is written in PHP, by using the Laravel framework. To create a new map, the user should specify a name of map, its description, type of visibility (public or private) and finally upload images (Figure 2). Information about maps is stored in MySQL database. Considering that the application is not intended for a large number of users per second, the relational database represents a good choice. All images uploaded by users are stored at the cloud storage for seven day. After one week, all images are deleted by background shell script and only tiff files and map tiles are kept.

Dashboard consists of two main parts: maps that are already created, and currently processing maps or maps that are in queue. When user creates a new map and uploads images, a processing job is sent into a queue. When a map processing is started, an e-mail notification is sent to user. Also, the user is notified when a map processing is finished. Currently a cloud server can process one map in time. In future version, the number of jobs served from the queue can be optimized, since the information about the number of processed images, processing time, number of cores, etc. is stored in database.

Finally, maps are visualized by mobile-friendly, open-source Leflet.js library. Beside ability to preview created maps, the user can download georeferenced orthomosaic in GeoTiff format and use it with GIS software.



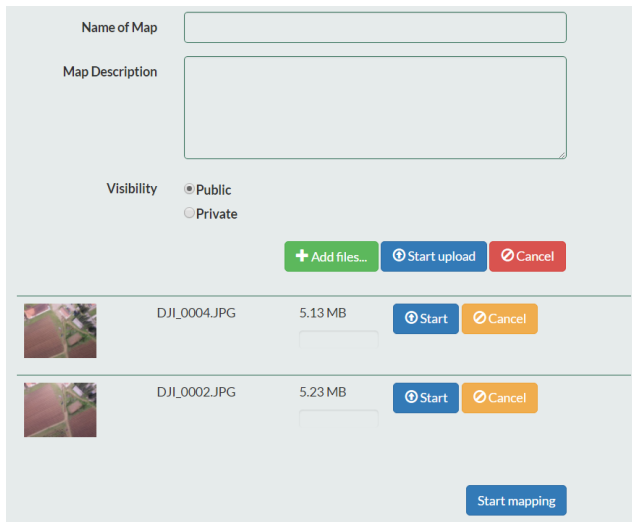


Figure 2. Image uploading interface

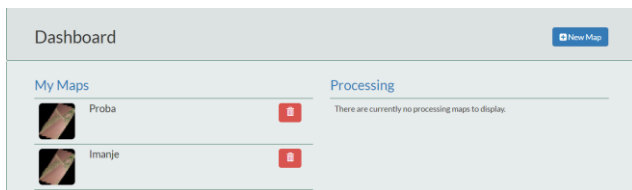


Figure 3. Dashboard of developed application

OpenDroneMap (ODM) is an open source tool for processing aerial drone imagery, [18]. To produce Point Clouds, Digital Surface Models, Textured Digital Surface Models and Orthorectified Imagery, ODM uses several images processing libraries such as an OpenSfM, PMVS, bundler tools, etc [19]. It consists of five modules: georeferencing, meshing, orthophoto, slam and texturing.

The OpenDroneMap can be run from the command line, which allows the one to write scripts that can be called threw web interface, on the user demand. ODM has many input parameters that can optimized in order to obtained more quality image. The input parameters should be set to the values that provide best compromise between image quality and processing time.

Finally, when georeferenced mosaic is generated, a *gdal* library is used to create map tiles [20]. Gdal generates a directory with small tiles and metadata, following the OSGeo Tile Map Service Specification [20].

At the moment, apache server and image processing tools are installed on the same virtual machine. A process control system supervisor is installed in order to keep the server from overloading [21]. In future version a separate installation of the web server and image processing tools will be considered.

The presented application can be applied in precision agriculture. For, example, it can be integrated with already developed expert system and a web service for grapevine monitoring and forecasting a disease [22, 23].

#### IV. RESULTS AND DISCUSSION

The images are acquired by DJI Phantom 4 which is equipped with 12.6Mpx camera. The test images are taken at attitudes 50m and 100m, respectively. Processing

timings are compared for Agisoft, MapsMadeEasy and described web application. Agisoft software was executed on a Windows PC with a dual quad-core CPU and 8GB of memory, whereas the described application was executed on the server with 24 cores and 6GB of RAM.

In the first example, 180 images were used. Figure 4 shows the results of processing. The resulting mosaic was rendered at 5 cm per pixel and is shown overlaid a Google Earth. It can be seen that resulting mosaic have a good overlap with Google Earth. Figure 5 shows the resulting mosaic at zoom level 2. It can be concluded that image quality is on a satisfactory level.



Fig 4. Resulting mosaick at zoom level 12

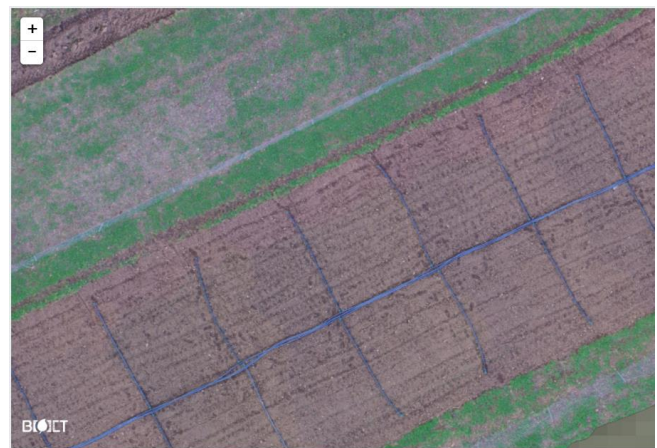


Fig 5. Resulting mosaick at zoom level 4

TABLE I. COMPARISON OF SOFTWARE TIMINGS

Software	DroneMapper	MapsMadeEasy	PhotoScan
Number of images	180	180	180
Processing time	35min	5h 10min	8h 40min

Table I shows the software timings for the same dataset of 180 images. It can be seen, that the DroneMapper has the shortest processing time – 35min, while the MapsMadeEasy and PhotoScan took 5h 10min and 8h 40min for map processing, respectively.

In the second example, 65 images are used. The resulting mosaic was rendered at 10 cm per pixel and is

shown overlaid a Google Earth. From Fig. 6 it can be seen that in this example the resulting mosaic also have a good overlap with Google Earth

Table II shows the software timings for the same image dataset. It can be seen that in this example the DroneMapper also has shorter processing time compared to the MapsMadeEasy and PhotoScan software.



Fig 6. Resulting mosaick

TABLE II. COMPARISON OF SOFTWARE TIMINGS

Software	DroneMapper	MapsMadeEasy	PhotoScan
Number of images	65	65	65
Processing time	10min 52 s	1h 10min	3h 20min

## V. CONCLUSION

In this paper DroneMapper - a cloud application for automatic image mosaicking and georeferencing is described. Several open source tools are used for image processing and mapping. The DroneMapper produces high quality maps of photographed area with accurately assigned geographic coordinates. It is shown that developed application requires shorter processing time compared to the MapsMadeEasy and PhotoScan software.

## ACKNOWLEDGMENT

This work has been supported by the Ministry of Science of Montenegro and the HERIC project through the BIO-ICT Centre of Excellence (Contract No. 01-1001).

## REFERENCES

[1] J. A. J. Berni, P. J. Zarco-Tejada, L. Suarez and E. Fereres, "Thermal and Narrowband Multispectral Remote Sensing for Vegetation Monitoring From an Unmanned Aerial Vehicle," in

*IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 3, pp. 722-738, March 2009.

[2] E. R. Hunt, W. D. Hively, S. J. Fujikawa, D. S. Linden, C. S. T. Daughtry, "Acquisition of NIR-Green-Blue digital photographs from unmanned aircraft for crop monitoring", *Remote Sens*, vol. 2: pp. 290-305, 2010.

[3] H. Erol and F. Akdeniz, "A multispectral classification algorithm for classifying parcels in an agricultural region," *Int. J. Remote Sens.*, vol. 17, no. 17, pp. 3357-3371, Nov. 1996.

[4] P. Boissard, J. G. Pointel, and P. Huet, "Reflectance, green leaf area index and ear hydric status of wheat from anthesis until maturity," *Int. J. Remote Sens.*, vol. 14, no. 14, pp. 2713-2729, Sep. 1993.

[5] M. S. Rasmussen, "Assessment of millet yields and production in northern Burkina-Faso using integrated NDVI from the AVHRR," *Int. J. Remote Sens.*, vol. 13, no. 18, pp. 3431-3442, Dec. 1992.

[6] J. G. P. W. Clevers, C. Buker, H. J. C. van Leeuwen, and B. A. M. Bouman, "A framework for monitoring crop growth by combining directional and spectral remote sensing information," *Remote Sens. Environ.*, vol. 50, no. 2, pp. 161-170, Nov. 1994.

[7] M. L. Adams, W. D. Philpot, and W. A. Norvell, "Yellowness index: An application of spectral second derivatives to estimate chlorosis of leaves in stressed vegetation," *Int. J. Remote Sens.*, vol. 20, no. 18, pp. 3663-3675, Dec. 1999.

[8] C. Yang, J. H. Everitt, J. M. Bradford, "Comparison of QuickBird Satellite Imagery and Airborne Imagery for Mapping Grain Sorghum Yield Patterns," in *Precision agriculture*, vol. 7, no. 1, March, 2006.

[9] Z. Li and V. Isler, "Large Scale Image Mosaic Construction for Agricultural Applications," in *IEEE Robotics and Automation Letters*, vol. 1, no. 1, pp. 295-302, Jan. 2016.

[10] J. Navia, I. Mondragon, D. Patino and J. Colorado, "Multispectral mapping in agriculture: Terrain mosaic using an autonomous quadcopter UAV," *2016 International Conference on Unmanned Aircraft Systems (ICUAS)*, pp. 1351-1358, Arlington, VA, 2016.

[11] M. D. Pritt, "Fast orthorectified mosaics of thousands of aerial photographs from small UAVs," *2014 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, pp. 1-8, Washington, DC, 2014.

[12] J. W. Gross, "A Comparison of Orthomosaic Software for Use with Ultra High Resolution Imagery of a Wetland Environment", *IMAGIN's 25th Annual Conference*, Great Wolf Lodge Traverse City, Michigan, 2016.

[13] U. Niethammer, S. Rothmund, U. Schwaderer, J. Zeman, M. Joswig, "Open source image-processing tools for low-cost UAV-based landslide investigations", *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XXXVIII-1/C22, 2011, Zurich.

[14] Pix4D - Drone Mapping Software for Desktop + Cloud + Mobile: <https://pix4d.com/>.

[15] Agisoft PhotoScan: <http://www.agisoft.com/>.

[16] MapsMadeEasy: <https://www.mapsmadeeasy.com/>.

[17] Public Lab: MapKnitter: <https://mapknitter.org/>.

[18] Open Source Toolkit for processing Civilian Drone Imagery: <http://opendronemap.github.io/odm/>.

[19] VisualSFM : A Visual Structure from Motion System: <http://ccwu.me/vsfm/>.

[20] GDAL: GDAL - Geospatial Data Abstraction Library: <http://www.gdal.org/>.

[21] Supervisor: A Process Control System: <http://supervisord.org/>.

[22] S. Vujović, M. Brajović, V. Popović, Bugarin, N. Latinović, J. Latinović, M. Bajčeta, "A web service for grapevine monitoring and forecasting a disease", *Information Technology IT'16*, pp. 185-188, Zabljak, 2016.

[23] M. Brajović, S. Vujović, V. Popović, Bugarin, S. Đukanović, M. Knežević, A. Topalović, "Soil analysis database and the expert system for recommendations of fertilization in agriculture", *Information Technology IT'16*, pp. 189-193, Zabljak, 2016.

# Online promocija poljoprivrednih proizvoda: studija slučaja bioportal.me

Nikola Žarić, *Member, IEEE*, Boban Melović, i Miomir Jovanović

**Sadržaj** – Osnovno istraživačko pitanje, kojim se bavi ovaj rad, jeste sagledavanje mogućnosti online marketinga u promociji poljoprivrednih proizvoda, što se pozitivno reflektuje na ukupni nivo konkurentnosti. Autori polaze od hipoteze da se *adekvatne promotivne aktivnosti poljoprivrednih proizvoda putem interneta nalaze u direktnoj pozitivnoj korelaciji sa percepcijom potrošača, te jačanjem konkurentnosti i ekonomskim efektima koji se na toj osnovi mogu generisati*. Teorijski stavovi su analizirani kroz studiju slučaja bioportal.me. Zaključak do kojeg autori dolaze potvrđuje da je online promocija jedna od ključnih determinanti konkurentnosti u oblasti poljoprivrede, te da se ista u Crnoj Gori koristi u nedovoljnoj mjeri.

**Ključne riječi** – marketing, online promocija, poljoprivreda, bioportal.me, ICT.

## I. UVOD

Korišćenje mogućnosti interneta u razvijanju koncepta online marketinga podrazumijeva prihvatanje specifičnosti koje donosi integrisana marketing komunikacija, u odnosu na tradicionalni oblik komunikacije. Internet kao marketinški kanal je interaktivan, dostupan i može se integrisati sa svim elementima marketinga. U tom smislu, upotreba online marketinga je važno strategijsko sredstvo za svako preduzeće, a poseban značaj može imati u oblasti poljoprivredne proizvodnje. Može se reći da online marketing u oblasti poljoprivredne proizvodnje dovodi do poboljšanja prodaje, veće dostupnosti informacija o proizvodima i proizvođačima, optimalnije isporuke proizvoda i efikasnijeg i efektivnijeg građenje odnosa sa kupcima. Kada govorimo o Crnoj Gori, iskustvo pokazuje da je oglašavanje na internetu i dalje u razvoju, ali da je u pitanju tržište koje ima potencijal, iz razloga što, za relativno mali iznos novca, oglas, odnosno poruka, može stići do velikog broja potencijalnih potrošača. Shodno navedenom, primjena online marketinga, iako za sada nije na zadovoljavajućem nivou u Crnoj Gori, dobija sve više na značaju i od strane domaćih kompanija, kako u opštem smislu, tako i od strane onih koje djeluju u oblasti poljoprivredne proizvodnje. Da potencijal za primjenu,

*Rad je podržan od strane Ministarstva nauke Crne Gore i HERIC projekta kroz BIO-ICT Centar Izvrsnosti (Uvovor broj. 01-1001).*

Rad je dio projekta: Jačanje konkurentnosti kroz podsticaj razvoja organske poljoprivrede - komparativna studija između Crne Gore i Srbije.

Nikola Žarić (autor za kontakte), Elektrotehnički fakultet, Univerzitet Crne Gore, Džordža Vašingtona bb, 81000 Podgorica, Crna Gora (e-mail: zarić@ac.me).

Boban Melović, Ekonomski fakultet, Univerzitet Crne Gore, ul. Jovana Tomaševića 37, 81000 Podgorica, Crna Gora, (e-mail: bobanm@ac.me).

Miomir Jovanović, Biotehnički fakultet, Univerzitet Crne Gore, Džordža Vašingtona bb, 81000 Podgorica, Crna Gora (e-mail: miomirj@ac.me).

odnosno tržište, postoji potvrđuju i podaci koji pokazuju da je u Crnoj Gori u 2016. godini bilo 388,057 korisnika interneta, što čini 62% ukupne populacije [1]. Sa druge strane, postoje i infrastrukturni uslovi. Naime, rezultati istraživanja Monstata iz 2016. godine pokazuju da 99,1% preduzeća, koja u svom radu imaju upotrebu kompjutera, imaju pristup internetu [2]. Od preduzeća koja imaju pristup internetu, 74,6% preduzeća ima veb sajt. U anketi koju je sproveo Centar Izvrsnosti BIO-ICT krajem 2015. godine 98% poljoprivrednih proizvođača izjasnilo se da ima pristup Internetu.

Kvalitetan proizvod je jedan od najboljih načina da se savremenom potrošaču ukaže na kvalitet i specifičnost određenog proizvoda (područja) i da se isti proda. Međutim, samo kvalitetan proizvod nije dovoljan za uspjeh na tržištu, ukoliko nije praćen odgovarajućom informativno-promotivnom podrškom [3]. Shodno navedenom, u radu je data studija slučaja bioportal.me - portala kojem je namjera da omogući crnogorskim proizvođačima poljoprivrednih proizvoda i hrane iz mora da na jednom mjestu imaju integrisanu organizovanu online promociju, te da na taj način predstave ponudu potencijalnim kupcima ne samo na tržištu Crne Gore, već i šire.

## II. AKTUELNO STANJE CRNOGORSKE POLJOPRIVREDE – ONLINE PROMOCIJA KAO RAZVOJNA ŠANSА

Primjena online marketinga i promotivnih aktivnosti, pored njihove nesumnjivo bitne uloge u sektoru proizvodnje i usluga, sve veću zastupljenost ima i u poljoprivredi. Ovo iz razloga što u razvijenim zemljama poljoprivreda postaje sve više integrisan sistem na području proizvodnje hrane, zaštite životne sredine i izbalansiranog regionalnog razvoja. Crnogorska poljoprivreda nalazi se u veoma intenzivnom procesu promjena. U poređenju sa više zemalja u Evropskoj Uniji (EU), poljoprivreda u Crnoj Gori ima znatno veći ekonomski značaj, s obzirom na njeno učešće u bruto domaćem proizvodu [4]. Kada se primarnoj proizvodnji doda prerađivački sektor, proizilazi da se značajan dio BDP stvara u sektoru proizvodnje hrane i pića. Poljoprivredne površine zauzimaju 38% (518.067 ha) ukupne teritorije Crne Gore [5]. Fond poljoprivrednog zemljišta od 0,82 ha po stanovniku predstavlja važan resurs budući da je Crna Gora po ovom pokazatelju pri samom vrhu evropskih zemalja [6]. Sa druge strane, Crna Gora može ponuditi autohtone sorte hrane i pića, ali i tradicionalne i ekološki uzgojene namirnice, što je savremenim potrošačima itekako zanimljivo, jer je toga u visoko razvijenim zemljama sve manje [7]. Međutim, kao jedan od ključnih problema izgradnje prepoznatljivih

proizvoda (brendova) u oblasti poljoprivrede javlja se nedostatak finansijskih sredstava, pogotovo dijela koji se odnosi na marketing budžet. Polazeći od navedenog, kreiranje online marketing kampanja je logičan strategijski izbor poljoprivrednih proizvođača u Crnoj Gori.

### III. KREIRANJE ONLINE MARKETING (PROMOTIVNE) KAMPANJE – RELEVANTNOST ZA POLJOPRIVREDU

Pojam “internet marketinška kampanja”, odnosno “online marketing kampanja” predstavlja sveobuhvatan proces analize, istraživanja i kreiranja strategije online nastupa u odnosu na zadate ciljeve. Kvalitetna strategija internet nastupa može dovesti do odličnih rezultata, koji su uglavnom znatno uspješniji u odnosu na tradicionalni marketing [8]. S obzirom da online strategija nastupa podrazumijeva integraciju svih vidova internet oglašavanja, kako bi se došlo do ciljane grupe, nezavisno od toga gdje provode vrijeme na internetu, logičnim se nameće sinergijski efekat jedne ovakve internet kampanje. Naime, ujedinjena strategija internet nastupa podrazumijeva kombinovanje veb sajta, društvenih mreža, *Google* servisa i *Newsletter* alata, u onoj mjeri koja se odnosi na očekivanja ciljane grupe, koja se utvrđuje istraživanjem. Nakon istraživanja ciljane grupe, određuju se njihova očekivanja i potrebe, na osnovu kojih se, zatim, kreiraju strategije online nastupa koje će im pomoći da se opredijele za datu ponudu.

Marketing u poljoprivredi obuhvata sve aktivnosti usmjerene ka definisanju marketing miksa, koji pored klasičnog modela “4P” (*product, price, place and promotion*), obuhvata i model “3C” [9]:

- *consumerisam* - obezbjeđenje dugoročnih interesa potrošača,
- *control* - kontrolu i racionalno korišćenje raspoloživih resursa i
- *conservation* - očuvanje čovjekove okoline.

Komunikacija sa potrošačima poljoprivrednih proizvoda postaje sve važnija i sve veći broj proizvođača shvata da, bez sistemskog, planskog i udruženog pristupa oglašavanju, postoje veliki rizici u plasmanu poljoprivrednih proizvoda. S tim u vezi, promocija poljoprivrednih potencijala (proizvoda) podrazumijeva proces komuniciranja između prodavca i potencijalnog kupca sa ciljem da se utiče na njegove stavove i ponašanje. Funkcija promocije i u oblasti poljoprivrede je da informiše, nagovori (ubijedi) i utiče na potrošačevu odluku o kupovini.

Promocija u poljoprivredi nije, niti smije biti homogena aktivnost. Naprotiv, kao što je i sam marketing miks heterogen, tako je promotivna aktivnost kompleksna i sastavljena od različitih elemenata koji marketing menadžerima stoje na raspolaganju. Za realizaciju izabrane promotivne strategije postoje određene mogućnosti i njihova kombinacija. Zato se i može govoriti o promocijnom miksu, kao mješavini (kombinaciji) različitih elemenata promocije, pažljivo dizajniranih za pojedine ciljane segmente, odnosno tržišta.

Bez sumnje, online promocija u oblasti poljoprivrede, kao i oglašavanje na internetu predstavljaju interaktivni direktni marketing. Ovakav marketing uključuje i kreativni i tehnički aspekt interneta, koji sadrži dizajn, propagandu (reklamu) i tradicionalni marketing. Sa druge strane,

detaljna i precizna segmentacija tržišta, odnosno njegova fragmentacija, koju omogućava online oglašavanje, a čiji je rezultat “jedna osoba - jedan segment”, dovodi do mogućnosti transformacije sa masovnog na mikromarketing, što je, takođe, od velikog značaja ponuđačima poljoprivrednih proizvoda u Crnoj Gori.

U dijelu koji slijedi dajemo kraće objašnjenje odabranih oblika online promocije, koji se odnose na veb sajt, društvene mreže i *Google AdWords* oglašavanje, akcentujući relevantnost za oblast poljoprivrede.

**Veb sajt - funkcionalna platforma komunikacije.** Veb sajt najčešće predstavlja primarni izvor potrošačevog iskustva i smatra se najvažnijim komunikacionim elementom u digitalnom okruženju. Kada govorimo o ponudi poljoprivrednih proizvoda, veb sajt može sadržati virtualni prikaz proizvoda, promotivni materijal, katalog cijena, a u uslovima elektronske trgovine predstavlja i prodajno-distributivno mjesto. U kontekstu digitalne ekonomije, veb sajt objedinjuje sve elemente tradicionalnog “4P” modela marketing miksa. Ponuđači poljoprivrednih proizvoda u Crnoj Gori treba da imaju u vidu da veb sajt treba da bude dizajniran tako da privlači što je moguće veći broj posjetilaca koji će posjetiti sajt, ostati na njemu što duže vremena i često ga ponovo posjećivati. Ovi uslovi će se postići ukoliko se veb sajt stalno ažurira, kako bi u svako vrijeme bio aktuelan i privlačan. Neophodno je da sajt sadrži što više korisnih informacija i interaktivnih alata, koji omogućavaju korisnicima da pronađu željeni proizvod. U većini slučajeva veb sajtovi u Crnoj Gori, čiji sadržaj tangira oblast poljoprivrede, zapravo nemaju dvosmjernu komunikaciju, a kupci na sajtu nemaju mogućnost da sarađuju sa preduzećem, niti da iskažu svoje zadovoljstvo ili nezadovoljstvo njihovim proizvodima ili uslugama. Primjer dobre prakse koji omogućava pravazilaženje ovih nedostataka je *bioportal.me*, o kome će biti više riječi u poglavlju IV.

**Društvene mreže kao oblik online promocije.** Fenomen društvenih medija i mreža je zadnjih godina ostavio dubok trag na cjelokupno društvo. Danas preovlađuje stav da društvene mreže predstavljaju standard i “nešto bez čega se ne može”. Kada govorimo o mogućnosti upotrebe društvenih mreža u oblasti poljoprivrede u Crnoj Gori, kompanije mogu da promovišu svoje proizvode preko društvenih mreža i da stupe u direktni kontakt sa potrošačima, da komuniciraju sa njima i saznaju reakcije vezane za njihove proizvode. Isto tako i potrošači mogu da se informišu o ponudi poljoprivrednih proizvođača i da dobiju direktan odgovor. Za ponuđače poljoprivrednih proizvoda je vrlo važno da na pravi način iskoriste mogućnosti koje im društvene mreže pružaju. Drugim riječima, jako je bitno plasirati propagandnu poruku preko medija koji ciljna grupa prati, a s obzirom da na društvenim rečama postoji veoma veliki broj korisnika, to one predstavljaju pravo mjesto za plasiranje određene poruke.

Postoje različiti oblici oglašavanja na društvenim mrežama, kao što su: *Facebook* i *Instagram* kampanje, *Google AdWords*. Osim navedenog, kada govorimo o internet oglašavanju, postoje i različite varijacije *banner* oglašavanja, *e-mail* oglašavanje, oglašavanje putem ključnih riječi, te razni oblici kontekstualnog oglašavanja,

koji zbog dubine istraživanja nisu predmet analize u ovom radu.

Jedna od brojnih prednosti oglašavanja putem pomenutih društvenih mreža jeste mjerljivost. Napredni alati koje ove mreže pružaju omogućavaju da se sve može pratiti, odnosno da se može izmjeriti svaki aspekt kampanje. Naime, ubrzo nakon pokretanja kampanje za neki poljoprivredni proizvod oglasi se počinju prikazivati, pa će potencijalni kupci veoma brzo saznati za ponudu, što dovodi do boljih rezultata. Takođe, jedna od prednosti je i *pay per click*, odnosno plaćanje isključivo ukoliko korisnik posjeti link, odnosno sajt. Kao i kod *Facebook*-a, *Google AdWords* nudi mogućnost preciznog targetiranja ciljane grupe, gdje se može odabrati način, vrijeme i mjesto prikazivanja oglasa. Nadalje, pravovremenost je veoma bitna prednost društvenih mreža, jer se oglasi prikazuju korisnicima onda kada oni vrše pretragu za dati ili sličan proizvod, odnosno uslugu. Dakle, oglašivaču se naplaćuje isključivo onda kada neko klikne na oglas i posjeti veb sajt, pa možemo konstatovati da je ovo posebno važno poljoprivrednim proizvođačima u Crnoj Gori, koji, najčešće, nemaju zavidan budžet za oglašavanje.

#### IV. STUDIJA SLUČAJA - BIOPORTAL.ME

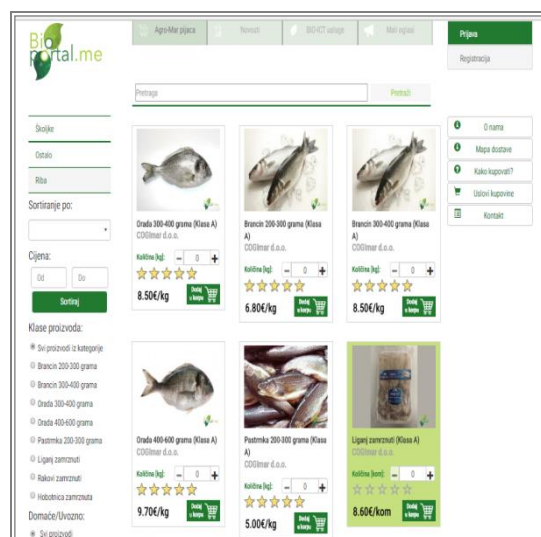
U cilju utvrđivanja potreba poljoprivrednih proizvođača u Crnoj Gori i njihove spremnosti da koriste nove tehnologije kako u proizvodnji, tako i u procesu prodaje i promocije svojih proizvoda Centar Izvrsnosti u bioinformatici – BIO-ICT sproveo je anketu u kojoj je učestvovao 41 poljoprivredni proizvođač. Rezultati ove ankete pokazuju da je čak 76% proizvođača zainteresovano da vrši prodaju svojih proizvoda putem interneta, ali da nemaju resurse da samostalno organizuju prodaju na taj način. Anketom je pokazano da 63% proizvođača koristi makar neki vid informaciono-komunikacionih tehnologija u svom poslovanju ili informisanju o proizvodnji, da *e-mail* adresu ima 78%, dok čak njih 98% ima pristup internetu. Za edukaciju u oblasti poljoprivredne proizvodnje putem interneta zainteresovano je čak 85% proizvođača.

Podaci dobijeni putem ankete ukazuju na izuzetnu zainteresovanost i spremnost proizvođača za korišćenjem novih tehnologija i digitalnog marketinga u plasmanu svojih proizvoda i unapređenju proizvodnje uopšte. Međutim, niko od proizvođača nije samostalno realizovao proces prodaje i intenzivne promocije svojih proizvoda putem interneta. Razlog leži u nedovoljnim resursima kojim raspolažu, a i u kompleksnosti čitavog procesa digitalnog marketignog prezentovanog u prethodnom poglavlju.

Polazeći od prethodno navedenog, autori smatraju da holistički pristup organizaciji online promocije i prodaje predstavlja optimalniju strategiju za nastup na tržištu. Navedeno podrazumijeva kreiranje jedinstvene IT i makretinške podrške, koja bi predstavljala osnov da pojedinačni proizvođači putem takve platforme mogu predstaviti ponudu i "ličnu kartu" svog gazdinstva i uređivati je prema sopstvenim potrebama. Na taj način moguće je ostvariti veći nivo efikasnosti, što za posljedicu ima pozitivan uticaj na kreiranje jedinstvenog brenda u oblasti crnogorske poljoprivrede, što, nadalje, omogućava

bolji nastup kako na domaćem, tako i na inostranom tržištu.

U kontekstu prethodno navedenog, kao efikasnu platformu za jedinstvenu promociju i online prodaju poljoprivrednih proizvoda autori navode veb portal BIOPORTAL.ME ([www.biportal.me](http://www.biportal.me)) razvijen od strane Centra Izvrsnosti u bioinformatici – BIO-ICT. Platforma je realizovana kao multifunkcionalni alat koji omogućava proizvođačima poljoprivrednih proizvoda i hrane iz mora da na jednom mjestu, brzo i efikasno predstavljaju svoju ponudu svim kupcima koji imaju pristup Internetu. Na ovaj način značajno se ubrzava i pojednostavljuje plasman proizvoda. Takođe, kupac na jednom mjestu ima uvid u ponudu i mogućnost odabira i kupovine željenih proizvoda. Kreiran je sistem lojalnosti i sistem ocjenjivanja proizvoda, što omogućava proizvođačima da na bazi kvaliteta svojih proizvoda izgrade reputaciju pouzdanih i provjerenih ponuđača, a sa druge strane omogućava kupcima da na bazi prethodnih iskustava lakše donesu odluku o kupovini. Sistem lojalnost nagrađuje vjerne kupce kroz različite vidove promocija, popusta, gratis proizvoda.



Slika 1. Biportal.me platforma

Pored prodaje proizvoda platforma nudi i niz drugih servisa koji će svim akterima u procesu poljoprivrede omogućiti pristup različitim informacijama i servisima na jednom mjestu. Ovi servisi su: Novosti, BIO-ICT usluge i Oglasi.

- Novosti predstavljaju dio portala na kojem proizvođači, ali i svi ostali zainteresovani mogu na jednom mjestu da pronađu informacije o procesu proizvodnje, tehnološkim trendovima, promociji zdravlja, projektima, događajima i sajmovima. Pored toga, IT infrastruktura portala omogućava proizvođačima koji su uključeni u ovaj proces da kroz autorske članke, izjave, intervju prezentuju svoje poljoprivredno gazdinstvo i proizvode.

- BIO-ICT usluge predstavljaju segment platforme putem kojeg će se proizvođačima pružati različite vrste usluga iz oblasti rada BIO-ICT centra, a to su između ostalih: online edukacija; konsultantske usluge iz oblasti proizvodnje i zaštite bilja, stočarstva, uzgoja ribe; opis dostupnih laboratorijskih analiza, razvoj softverskih i hardverskih rješenja, online pedološka mapa sa analizom

zemljišta u Crnoj Gori, ekspertni sistemi za pametno navodnjavanje, prskanje, đubrenje, itd.

- Oglasi kao dio portala namijenjen besplatnom oglašavanju ponude i potražnje u oblastima: pružanja usluga u poljoprivredi, mehanizacije, opreme i djelova, poljoprivrednog zemljišta, sadnog materijala i zaštitnih sredstava i đubriva.

Prethodno navedeno predstavlja objedinjeni pristup u pružanju servisa, usluga, edukacije proizvođača, kao i opštih informacija koje mogu biti od koristi kako proizvođačima tako i kupcima. Namjera je da se na ovaj način na jednom mjestu obezbijedi širok spektar informacija, što će uticati na stravanje navika svih učesnika u ovom procesu, da na jednom mjestu imaju pristup različitim segmentima vezanim za poljoprivredu.



Slika 2. Biportal.me – Novosti, projekti, događaji, sajmovi

U trenutku pisanja ovog rada portal je u fazi testiranja, a njegova puna upotreba očekuje se u prvom kvartalu 2017. godine. Iskustva dobijena tokom testiranja, kako od strane kupaca (samo odabrani testni kupci), tako i od strane proizvođača izuzetno doprinose stalnom unapređenju vizuelnih i funkcionalnih performansi portala, što je još jedan od benefita korišćenja alata za digitalnu promociju.

U testnom periodu nije rađena javna online promocija, a plan promocije u inicijalnoj fazi bi obuhvatio i otvaranje profila na društvenim mrežama, kao i Google AdWords oglašavanje. U sledećoj fazi plan obuhvata i upotrebu e-mail marketinga i newsletter-a. Cilj upotrebe newsletter-a ne bi bio samo za potrebe oglašavanja, već bi njegova primarna svrha bila izgradnja dugoročnih odnosa sa kupcima i/ili onima koji su se prijavili na newsletter listu biportal.me. Svakako da bi to indirektno uticalo i na rast prodaje, Fokus newsletter-a biće i na pružanju relevantnog i korisnog sadržaja iz oblasti poljoprivrede, onima koji ga primaju (savjeti, događaji, sajmovi ...).

## V. ZAKLJUČAK

Online marketing kao način komunikacije omogućava ponuđačima poljoprivrednih proizvoda u Crnoj Gori da ostvare personalizovan odnos sa svojim kupcima i da na osnovu toga kreiraju diferencirane marketing poruke, prilagođene pojedinačnim, specifičnim segmentima korisnika na internetu. Zadatak online promocije u oblasti poljoprivrede je kreirati strategiju pozicioniranja koja će osvojiti potrošača, što je u konačnom moguće ako se

osmišljena poruka uspješno prenese ciljnoj grupi, preko pravih kanala komunikacije. U tom smislu, od posebnog značaja je upotreba veb sajta, društvenih mreža, Google oglasa, banera i sl. Ponuđači poljoprivrednih proizvoda u Crnoj Gori ne koriste u dovoljnoj mjeri potencijale koje nudi online oglašavanje. U skladu sa navedenim, u radu je data studija slučaja biportal.me koji ima za cilj integrisano organizovanje online promocije, sa višestrukim prednostima kako za proizvođače, tako i za kupce. U krajnjem, online promocija je jedna od ključnih determinanti konkurentnosti, zbog čega poljoprivredni proizvođači moraju koristiti šanse koje im online oglašavanje nudi. U tom pravcu treba da budu razvijana i buduća istraživanja iz predmetne problematike analizirane u ovom radu.

## LITERATURA

- [1] Internet Users, dostupno na: [www.InternetLiveStats.com](http://www.InternetLiveStats.com), datum pristupa 24.01.2017. godine
- [2] Vlada Crne Gore, Zavod za statistiku Crne Gore, Upotreba ICT u preduzećima u Crnoj Gori 2016, dostupno na [www.monstat.org](http://www.monstat.org), datum pristupa 23.01.2017. godine
- [3] Mihailović B., *Marketing*, CPI, Podgorica, 2012, str. 256.
- [4] Ministarstvo poljoprivrede, šumarstva i vodoprivrede (2006), *Crnogorska poljoprivreda i Evropska Unija – Strategija razvoja proizvodnje hrane i ruralnih područja*, Podgorica, Montenegro.
- [5] Monstat (2016), Opšti statistički podaci o strukturi poljoprivrede u Crnoj Gori, dostupno na: [www.monstat.org](http://www.monstat.org), datum pristupa 23.01.2017. godine
- [6] Fabris, N., Pejović, I. (2012), "Montenegrin agriculture: diagnosis and policy recommendations", *Economics of Agriculture*, Vol. 59(4): 657-673.
- [7] Melović B., Mitrović S., Milisavljević S., Pejanović R., Čelić Đ. (2012) „Research of consumption and competitiveness of homemade products for manufacturing improvement: A case study from Montenegro”, *African Journal of Agricultural Research*, Vol. 7(26), p. 3758.
- [8] Kesić, T. „Integrirana marketinška komunikacija“, Kratis, Zagreb, 2013, str. 22.
- [9] Melović, B. (2009), „Marketing menadžment u funkciji kreiranja brenda – primjer Crne Gore”, Ekonomski fakultet Podgorica, Montenegro, 332.

## ABSTRACT

The basic research question this paper is dealing with is considering the possibility of using online marketing in the promotion of agricultural products, which will have the positive impact to the overall competitiveness level. The starting assumption is that adequate promotion activities related to agricultural products over the Internet are in direct positive correlation with consumer perception and strengthening the competitiveness and economic effects. Theoretical attitudes are analyzed through the case study biportal.me. It is confirmed that online promotion is one of the key determinants of competitiveness in agriculture, and that the same is inadequately used in Montenegro.

**ONLINE PROMOTION OF AGRICULTURAL PRODUCTS: CASE STUDIES BIOPORTAL.ME**  
Nikola Žarić, Member, IEEE, Boban Melović, and  
Miomir Jovanović

# Retinal blood vessels segmentation using support vector machine and modified line detector

Petar Sekulić, Milija Bajčeta, Slobodan Djukanović, *Member, IEEE*

**Abstract** — In this paper, we present blood vessel segmentation approach, which can be used in digital retinal image analysis to extract the retinal image vessels. Support vector machine (SVM) is used to segment the vessels. Training of SVM is performed by using of features that are obtained by feature extraction step. In the feature extraction step, two features are obtained, which are used by SVM to perform classification. The method is evaluated on fundus retinal images of the DRIVE dataset. The proposed method is compared to several existing methods in the field, exhibiting similar or better performance.

**Keywords** — Support vector machine, retinal image, blood vessel segmentation, feature extraction.

## I. INTRODUCTION

Automated segmentation of retinal blood vessels enables ophthalmologists to perform vision screening exams on a large number of people for early detection of several diseases including diabetes, hypertension, arteriosclerosis and choroidal neovascularization. This is a non-invasive method that can prevent the occurrence of blindness as well as other undesired consequences of the mentioned diseases.

The previous work on the topic of blood vessel segmentation can be classified into three main categories [1], [2]:

- a) unsupervised learning,
- b) supervised learning,
- c) pathological retina based methods.

Unsupervised learning approaches [3] attempt to find inherent patterns of blood vessels in retinal images that can be used to determine whether a particular pixel belongs to the blood vessel or not. In this approach, the manual data are used only for calculating accuracy metrics.

Supervised methods [3] learn the rule for vessel extraction by means of an algorithm which extracts knowledge based on a training set of manually processed and segmented images. A very significant approach is

This work has been supported by the Ministry of Science of Montenegro and the HERIC project through the BIO-ICT Centre of Excellence (Contract No. 01-1001).

Petar Sekulić, Milija Bajčeta and Slobodan Djukanović are with Faculty of Electrical Engineering, University of Montenegro. Address: Džordža Vašingtona bb, 81000 Podgorica, Montenegro. Phone: +382 20 245 839, fax: +382 20 245 873. E-mails: {petars@ac.me, milijab@ac.me, slobdj@ac.me}. Web: www.etf.ac.me.

based on the application of line operators and support vector machine (SVM) for pixel classification [4]. The proposed method belongs to this category.

Pathological retinal based methods [5]-[7] are the only category in which pathology in retinal images is considered during vessel segmentation process. These methods produce better results on the pathological retinal images than most of the existing algorithms. However, these algorithms cannot achieve the performance of algorithms from the previous two categories on retinal images where there are no pathologies.

This paper presents an effective automated algorithm that is able to detect blood vessels of pathological as well as healthy person's retinal images. We propose to use SVM with characteristic line detector for retinal blood vessel segmentation. SVM is trained by a matrix of features which has two columns, of which the green channel of a retinal color image is the first feature, whereas the second feature is new line detector (continual line detector). The proposed line operator overcomes some disadvantages of the line detectors proposed in [4].

The proposed method is tested on the DRIVE<sup>1</sup> database, which comprises 40 retinal images with a 45 field-of-view (FOV). The size of images, cropped around the FOV, is 565×584 pixels. The DRIVE database is made up of two sets of images: a training set and test set, both sets of have 20 pictures. The test set provides the corresponding FOV masks for the images, which are circular (approximated diameter of 540 pixels) and two manual segmentations generated by two different specialists for each image.

Section II describes the SVM. Proposed methodology about this work is described in section III. Section IV discusses the results obtained. Section V gives the conclusion of the work.

## II. SUPPORT VECTOR MACHINE

SVM is a supervised classification technique in Machine Learning. The basic version of SVM takes the datasets at the input and for each of them determines which of two possible classes belongs to, which makes it a non-probabilistic binary linear classifier. For a given set of training samples, all of which belong to one of two possible classes, SVM represents patterns as points in space, mapped in that way to make samples of various classes separated with as wider gap (margin) between them as

<sup>1</sup> Available at <http://www.isi.uu.nl/Research/Databases/DRIVE/>

possible.

Besides linear classification, support vectors can successfully perform a nonlinear classification using kernel methods (Fig. 1), implicitly mapping their inputs into multidimensional spaces.

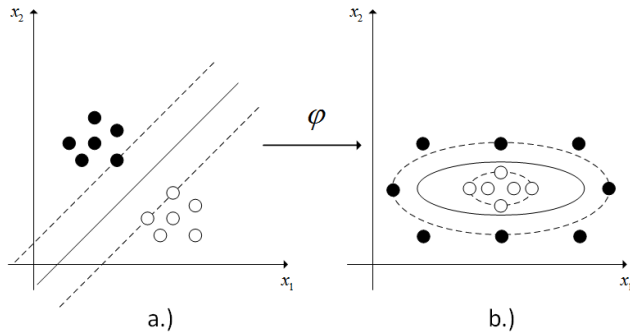


Fig. 1 An example of classification by SVM a) Linearly separable case, b) Linearly inseparable case (using kernel methods).

The minimum distance between training examples and decision boundary represents margin. The function of SVM is to maximize the margins for the classification of input data, creating as more as possible wider gap between different classes. Input data, which are used for the training of SVM, are denoted as  $(x^{(i)}, y^{(i)})$ ,  $i \in \{1, 2, \dots, m\}$ .  $x^{(i)}$  is a vector of features, whereas  $y^{(i)}$  is a scalar that provides information about the input data class  $y^{(i)} \in \{+1, -1\}$ . There are two types of margins:

a) functional

$$\gamma^{(i)} = y^{(i)}(w^T \cdot x + b) \quad (1)$$

b) geometric

$$\gamma^{(i)} = y^{(i)} \left( \frac{w^T \cdot x + b}{\|w\|} \right) \quad (2)$$

The geometric margin is the distance between the decision boundary and the specific training data. Provided that  $\|w\|=1$ , the geometric margin is equal to the functional. Now that we have introduced concepts of functional and geometric margins, from their calculation (1) and (2), it can be concluded that the functional margin carries information whether certain input example is correctly classified, whereas the geometric margin also includes information about the distance of the data from the decision boundary.

The SVM decision boundary satisfies the equation  $w^T \cdot x + b = 0$ . From the geometric margin and the decision boundary equations, we can conclude that the side on which the input example will be placed in regards to the decision boundary depends on the part  $y^{(i)}(w^T \cdot x + b)$  of (2). Knowing that  $y^{(i)}$  changes the sign depending on which class of the input data the example belongs to,  $w^T \cdot x + b$  must be a positive number in order to perform a

correct classification.

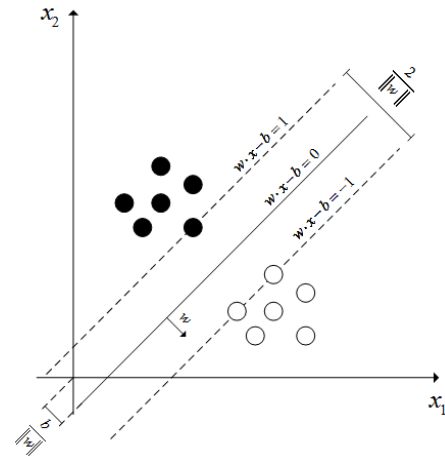


Fig. 2 Maximum-margin hyperplane and margins for an SVM trained with samples from two classes. Samples on the margin are called the support vectors.

The task of SVM is to perform maximization of geometric margin so that the functional margin of any input example is greater than or equal to the geometric margin of support vectors  $\gamma$ , i.e.

$$\max_{\gamma, w, b} \gamma \quad (3)$$

so that

$$y^{(i)}(w^T \cdot x + b) \geq \gamma, i = 1, \dots, m \quad (4)$$

$$\|w\| = 1 \quad (5)$$

However, the  $\|w\|=1$  constraint is cumbersome (non-convex), and this problem certainly is not in any format that we can plug into standard optimization software to solve. To avoid this condition, the previous maximization can be rewritten as:

$$\min_{\gamma, w, b} \frac{1}{2} \|w\|^2 \quad (6)$$

so that

$$y^{(i)}(w^T \cdot x + b) \geq 1, i = 1, \dots, m \quad (7)$$

Sequential minimal optimization (SMO) [9] is an algorithm for solving the optimization problem that arises during the training of SVM. This algorithm uses the Lagrangian dual problem to solve the optimization. When the Karush–Kuhn–Tucker (KKT) [3] conditions are fulfilled, the primary optimization problem defined by relations (6) and (7) is identical to the dual one:

$$\max_{\alpha} W(\alpha) = \max_{\alpha} \left[ \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i,j=1}^m y^{(i)} y^{(j)} \alpha_i \alpha_j \langle x^{(i)} x^{(j)} \rangle \right] \quad (8)$$

so that

$$\alpha_i \geq 0, i = 1, \dots, m \quad (9)$$

$$\sum_{i=1}^m \alpha_i y^{(i)} = 0. \quad (10)$$

The SMO algorithm offers an efficient way for solving the dual problem. If in every iteration of the algorithm only one Lagrangian multiplier can be varied, while the other



ones are fixed, the algorithm would not converge. That is why in every algorithm iteration two Lagrange multipliers change until the algorithm converges.

### III. PROPOSED METHOD

This section describes the proposed retinal blood vessel segmentation method based on the SVM algorithm. The first step in detection of blood vessels is feature extraction from RGB retinal image (Fig. 3a)). During the extraction of the blood vessels, only the green channel of RGB images is used since it makes the best contrast between the blood vessels and the background [4], [10].

Input data of SVM are characteristics that are obtained by digital image processing [8] of the retinal image. Features that are suggested are:

- a) inverse green channel,
- b) continual green channel.

When calculating the inverse green channel, it is important to note that the green channel is scaled in the range [0,1]. The inverse green channel value is calculated when subtracting scaled green channel value from 1. The inverse channel of one of the images from DRIVE base is shown in Fig. 3b).

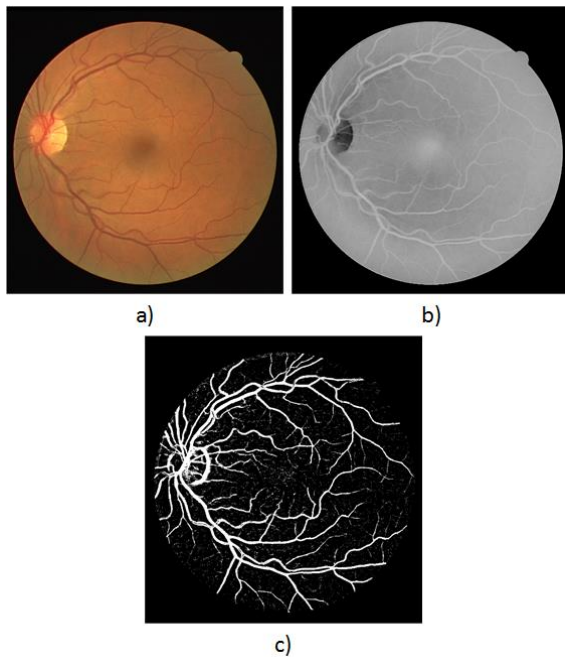


Fig. 3 a) Original RGB image. b) Inverse green channel. c) Continual line detector.

The continual line detector is obtained when we move through the image with window and calculate the average value of the line for 12 different orientations. The length of the line is 15 pixels [4]. The line is rotated at an angle of  $15^\circ$  relative to the central pixel of the window (Fig. 4b)). The optimal size of the square window is found to be  $15 \times 15$  [4]. After the line with the largest average gray level is found, the direction of pixels is determined by the direction of the strongest lines (one of the twelve angles). The continual line detector is calculated by summing pixels that are on the

line with a maximum average value in the mask  $15 \times 15$ , but only in the case when the direction of pixels corresponds to the direction of the central pixel (Fig. 4a)). The final value of the second feature was obtained when the average value of the sum was subtracted from the average value of the whole mask. The second feature of one of the images from DRIVE database is shown in Fig. 3a).

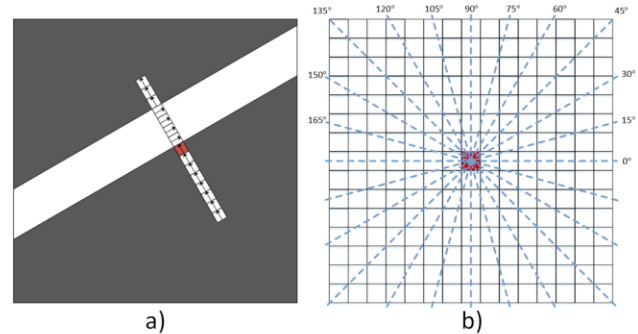


Fig. 4 a) Basic line detector (blood vessel shown in white). b) Twelve orientations to determine the strongest line.

After the features were calculated, SVM is trained. During the training of SVM, the Gaussian kernel [3] was used. The optimal parameters, that are determined on the training set, are  $C = 0.1$  and  $\sigma = 3$ . We used 20000 pixels (1000 pixels per image) that are randomly chosen from the training set. Classification performance is then evaluated on the 20 images of the test set.

### IV. RESULTS

The accuracy of the proposed classifier is tested on a test set of the DRIVE database. The test set contains sixty images, twenty of which are original images, twenty are binary images representing the first expert's classification of vessels, and the last twenty are binary images representing the second expert's classification of vessels. The accuracy of the system evaluated on test images classified by the first and the second experts are 93.7% and 94.3%, respectively. The other metrics used to evaluate the performance of the method is an  $F_1$  score [10], [12], [13]. Fig. 5 shows one of the images from DRIVE database, more precisely Fig. 5a) shows the RGB image while the Fig. 5b) and Fig. 5c) show manual segmentation of the first and second expert, respectively. The result of the proposed methods is shown in Fig. 5d). In Fig. 6, we compare the accuracy (top plot) and  $F_1$  score (bottom plot) for two experts. Clearly, better results are obtained when comparison is done with the second expert's results.

In Table 1, we provide performance comparisons of the proposed method with the other methods in this field. The comparison is made with respect to precision, recall,  $F_1$  score and accuracy, averaged over all test images in DRIVE database.

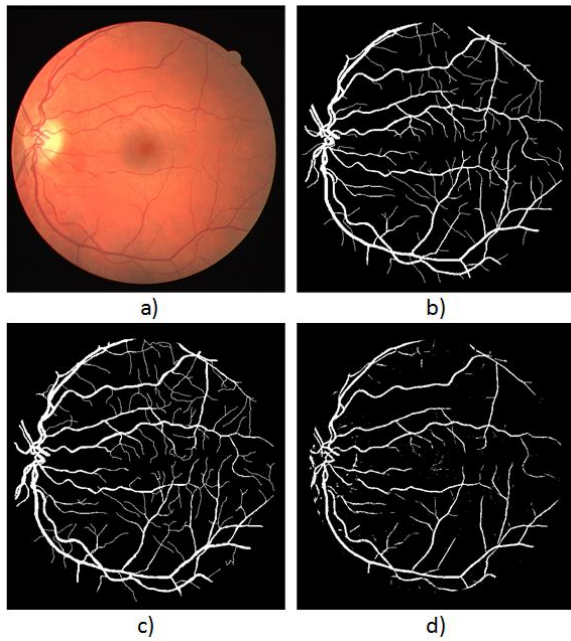


Fig. 5 Example of segmentation for one image from DRIVE database. a) Original retinal image. b) First expert's segmentation. c) Second expert's segmentation. d) Segmentation obtained by the proposed algorithm.

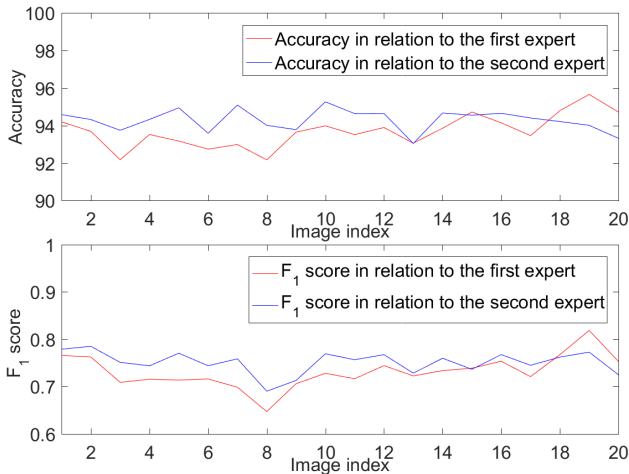


Fig. 6 Accuracy and F1 score for all test images from DRIVE database.

TABLE 1: COMPARISON OF SEVERAL METHODS IN TERMS OF PRECISION, RECALL, F1 SCORE AND ACCURACY

Method	Precision P	Recall R	F1 score	Accuracy
Cinsdikici and Aydin [11]	---	---	---	0.929
Fathi and Naghsh-Nilchi [12]	0.7559	0.7768	0.7662	0.958
Fraz et al. [13]	0.815	0.7152	0.7618	0.943
Asad et al. [14]	---	0.7388	---	0.902
Proposed	0:8234	0:6988	0:7560	0.943

## V. CONCLUSION

In this paper, a system for retinal blood vessels segmentation is trained with the help of features obtained

by digital image processing and SVM. The green channel of RGB image is used because it provides the biggest contrast between the blood vessels and the background. Features matrix is contained in two columns: green channel and continual line detector. Future research will address finding new features which will enable better accuracy and extend the system to enable detect whether the disease is present and possible categorization of diseases.

## REFERENCES

- [1] M. M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A.R. Rudnicka, C.G. Owen, and S.A. Barman, "Blood vessel segmentation methodologies in retinal images - A survey", *Computer Methods and Programs In BioMedicine*, vol. 108, no. 1, pp. 407-433, October 2012.
- [2] V. M. Saffarzadeh, A. Osareh, and B. Shadgar. "Vessel segmentation in retinal images using multi-scale line operator and K-means clustering", *Journal of medical signals and sensors*, vol. 4, no. 2, pp.122-129, April 2014.
- [3] R. O. Duda, P. E. Hart, and D. G. Stork. "Pattern classification, 2<sup>nd</sup> Edition", New York, 2001.
- [4] E. Ricci and R. Perfetti, "Retinal Blood Vessel Segmentation Using Line Operators and Support Vector Classification", *IEEE Transactions On Medical Imaging*, vol. 26, no. 10, pp. 1357–1365, October 2007.
- [5] B. S. Y. Lam, Y. Gao, A. W. C. Liew, "General retinal vessel segmentation using regularization-based multiconcavity modeling", *IEEE Transactions Medical Imaging*, vol. 29, no. 7, pp. 1369-1381, Jul 2010.
- [6] A. M. Mendonca, A. Campilho, "Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction", *IEEE Transactions Medical Imaging*, vol. 25, no. 9, pp. 1200-1213, Septembre 2006.
- [7] B. S. Y. Lam, H. Yan, "A novel vessel segmentation algorithm for pathological retina images based on the divergence of vector fields", *IEEE Transactions Medical Imaging*, vol. 27, no. 2, pp. 237--246, February 2008.
- [8] R. C. Gonzalez and R. E. Woods, "Digital Image Processing (3rd Edition)", Pearson, August 2007.
- [9] J. C. Platt, "12 fast training of support vector machines using sequential minimal optimization", *Advances in kernel methods*, pp. 185-208, 1999.
- [10] D. Marín, A. Aquino, M. E. Gegúndez Arias, and J. M. Bravo, "A New Supervised Method for Blood Vessel Segmentation in Retinal Images by Using Gray Level and Moment Invariants Based Features", *IEEE Transactions On Medical Imaging*, vol. 30, no. 1, pp. 146–158, January 2011.
- [11] M. G. Cinsdikici and D. Aydin, "Detection of blood vessels in ophthalmoscope images using MF/ant (matched filter/ant colony) algorithm", *Computer Methods and Programs In BioMedicine*, vol. 96, no. 2, pp. 85-95, November 2009.
- [12] A. Fathi and A. R. Naghsh-Nilchi, "Automatic wavelet-based retinal blood vessels segmentation and vessel diameter estimation", *Biomedical Signal Processing and Control*, vol. 8, no. 1, pp. 71-80, January 2013.
- [13] M. M. Fraz, S. A. Barmana, P. Remagnino, A. Hoppe, A. Basitb, B. Uyyanonvarac, A. R. Rudnickad, and C. G. Owend, "An approach to localize the retinal blood vessels using bit planes and centerline detection", *Computer methods and programs in biomedicine*, vol. 108, no. 2 pp. 600-616, November 2012.
- [14] A. H. Asad, A. T. Azar, and A. E. Hassaanien. "Ant colony-based system for retinal blood vessels segmentation," *Proceedings of Seventh International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA 2012)*, Springer India, pp. 441-452, 20

# Komparacija računarskog modela molekula sa eksperimentalnim rezultatima

Miljan Bigović, Žarko Zečević, Luka Filipović i Božo Krstajić

**Sadržaj** — U ovom radu je izvršena komparacija molekulskog modela, dobijenog primjenom odgovarajućih programskih paketa, i eksperimentalnih rezultata, na osnovu kojih je utvrđena egzaktna trodimenzionalna struktura ispitivanog organskog molekula. Primjenom strategije organske sinteze, sintetisan je molekul, za koji je urađena rendgenska strukturalna analiza i dobijen je precizan raspored svih atoma u prostoru. Nakon toga je isti molekul pomoću različitih hemijskih softverskih alata geometrijski optimizovan, konvertovan u odgovarajući standardni hemijski format i vizuelizovan. Na ovaj način je pokazano da je raspored svih atoma identičan eksperimentalno dobijenim vrijednostima. Na osnovu takvog rezultata, mogu se nastaviti molekulske simulacije i ispitivanje interakcija odabranog molekula sa receptorskim proteinima sa velikom sigurnošću da će rezultati moći biti i praktično interpretirani, s obzirom na veliku tačnost u slaganju eksperimentalnih i računarskih metoda.

**Ključne riječi** — Enol-karbonat, interakcije, modelovanje, molekulaska dinamika, NAMD, organski molekul, protein, rendgenska strukturalna analiza

## I. UVOD

SVRHA ovog rada je ispitivanje primjene simulacija molekulske dinamike u cilju rješavanja određenih problema koji su aktuelni u nekim prirodnim naukama, poput hemije, farmakologije, farmaceutske hemije, biohemije i sl. U prvom redu, tu se podrazumijeva ispitivanje raznih vrsta međumolekulskih nekovalentnih interakcija (koordinacije, Van der Waals-ove interakcije, vodonične veze, stvaranje jonskih mostova) i kovalentnih interakcija između malih organskih molekula (koji su sintetisani u laboratoriji, dok u prirodi ne postoje) i receptorskih mjesta u organizmu (proteina ćelijske membrane, koji su vrlo često takozvani „target-centri“ za dejstvo lijekova i biološki aktivnih hemijskih supstanci). Upravo na postojanju i proučavanju ovakvih interakcija bazira se moderna farmaceutska hemija, bioorganska

Miljan Bigović, Prirodno-matematički fakultet u Podgorici, Univerzitet Crne Gore, Džordža Vašingtona bb, 81000 Podgorica, Crna Gora (telefon: 382-20-245204; e-mail: miljan@ac.me).

Žarko Zečević (autor za kontakte), Elektrotehnički fakultet u Podgorici, Univerzitet Crne Gore, Džordža Vašingtona bb, 81000 Podgorica, Crna Gora (telefon: 382-67-224009; e-mail: zarkoz@ac.me).

Luka Filipović, Centar Informatičnog sistema Univerziteta Crne Gore, Džordža Vašingtona bb, 81000 Podgorica, Crna Gora (telefon: 382-69-668800; e-mail: lukaf@ac.me).

Božo Krstajić, Elektrotehnički fakultet u Podgorici, Univerzitet Crne Gore, Džordža Vašingtona bb, 81000 Podgorica, Crna Gora (telefon: 382-67-313 222; e-mail: bozok@ac.me).

hemija, kao i nauke koje se bave dizajniranjem lijekova. Međutim, te interakcije su se uglavnom proučavale ili na izolovanim ćelijskim kulturama i tkivima (*in vitro*) ili na eksperimentalnim životinjama pa i samim dobrovoljcima (*in vivo*). Sama sinteza novih i često kompleksnih organskih molekula, koji ispoljavaju neko fiziološko dejstvo, predstavlja skup i dugotrajan proces. Cijene postupka sinteze, hemikalija, opreme, prečišćavanja dobijenih prekursora na kraju diktiraju i cijenu finalnog proizvoda. Postojanje računarskog modela, koji bi mogao predvidjeti i precizno izračunati pomenute interakcije organskih molekula sa proteinima (koji su veoma dobro definisani i za većinu je određena trodimenzionalna struktura i koji se nalaze u dostupnim proteinskim bazama podataka - [1]) svakako da bi dalo ogroman doprinos rješavanju ove složene problematike, koja je u prirodnim i biomedicinskim naukama veoma aktuelna. Tom značaju posebno doprinosi i činjenica da proces razvoja novog lijeka u prosjeku traje od 12 do 15 godina i da košta između 600 i 800 miliona američkih dolara. Od 20 000 odabranih jedinjenja, koja se na samom startu razmatraju kao kandidati za novi lijek, tek njih desetak stigne do faze kliničkih ispitivanja, a u prosjeku samo jedno završi na polici apoteke [2]. Korišćenjem računarskih simulacija, visoka cijena troškova koja je uključena u cio proces razvoja novog lijeka se može drastično smanjiti, a cjelokupan proces ubrzati. Na taj način, daje se značajna prednost upotrebi računarskih modela u odnosu na tradicionalan način sinteze novih molekula, za kojima slijede ispitivanja njihove biološke aktivnosti.

Našim radom namjeravamo da pokažemo da je moguće, korišćenjem moćne računarske infrastrukture i odgovarajućih aplikacija, kreirati dobre modele molekula, koji se kasnije mogu koristiti u simulacijama koje daju za rezultat validne osobine predmetnih molekula u interakciji sa različitim organskim jedinjenjima. Na ovaj način se, uz primjenu savremenih računarskih resursa, omogućava brže i efikasnije kreiranje novih jedinjenja koja mogu biti prilagođena namjeni raznim grupama korisnika ili samim jedinkama.

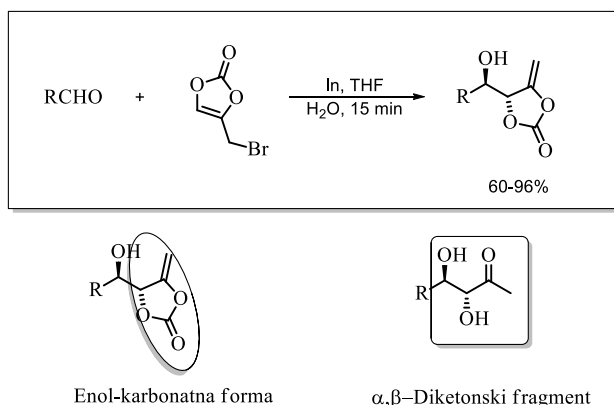
Za ispitivanje interakcija malih organskih molekula sa proteinima ćelijske membrane su korišćeni softverski paketi NAMD (NAnoscale Molecular Dynamics, [3]) i GROMACS (GROningen MACHine for Chemical Simulations, [4]), koji obuhvataju proučavanje molekulske dinamike datih sistema ligand-receptor (ligand je zapravo sintetisani mali organski molekul, dok receptor predstavlja sam protein). Ove aplikacije iziskuju enormne računarske resurse pa je istraživanje realizovano na infrastrukturi

(300000+ CPU jezgara i 11 PB skladišta) raspoloživoj u okviru H2020 projekta „Virtual Research Environment (VRE) in Southeast Europe and the Eastern Mediterranean (SEEM) – VI-SEEM“ koji okuplja 16 partnera iz različitih država navedenog regiona [5].

Ova tematika je veoma aktuelna u savremenom ispitivanju jedinjenja koja se razmatraju kao potencijalni novi lijekovi, a istraživanja su zasnovana upravo na primjeni različitih kompjuterskih simulacija i programa. Dostupan je veliki broj publikacija koje svjedoče o veoma uspješnoj primjeni ispitivanja interakcija u raznim oblastima nauke [7]-[9].

## II. MODEL MOLEKULA (LIGANDA) NAD KOJIM SE RADE MOLEKULSKA SIMULACIJE

Sintetisli smo ukupno 17 strukturno potpuno novih i do tada u hemiji nepoznatih molekula, čija je zajednička strukturna odlika postojanje takozvane enol-karbonatne strukturne jedinice [10]. Ta jedinica se može jednostavnim sintetičkim transformacijama prevesti u  $\alpha,\beta$ -dihidroksi-ketonsku jedinicu, koja je čest strukturni motiv mnogih prirodnih proizvoda i fiziološki aktivnih jedinjenja [11]-[13]. Sinteze enol-karbonata smo izvodili u reakciji alilovanja aldehida pod veoma blagim reakcionim uslovima u vodenoj sredini<sup>1</sup> i uz veoma dobre prinose očekivanih proizvoda, slika 1.

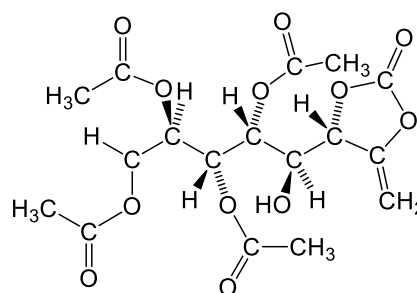


Sl. 1. Prikaz reakcije dobijanja enol-karbonata (na slici su istaknute enol-karbonatna i dihidroketonska forma)

Kompletna spektroskopska dokumentacija za svako novosintetisano jedinjenje (IC, <sup>1</sup>H NMR, <sup>13</sup>C NMR, MS) je dostupna [10]. Prisustvo polioksigenovanog fragmenta (oblika sa tri atoma kiseonika) je svakako interesantno jer pruža mogućnosti dalje funkcionalizacije derivata, što je u dizajniranju lijekova od izuzetne važnosti, a samo po sebi ima značaja i u organskoj sintezi [14].

Prvi zadatak simulacija je potvrda da će odgovarajući program podržati realnu trodimenzionalnu strukturu prvog malog organskog molekula, koji smo namjeravali da ispitamo – to je (2R,3R,4R,5S)-5-hidroksi-5-((S)-5-

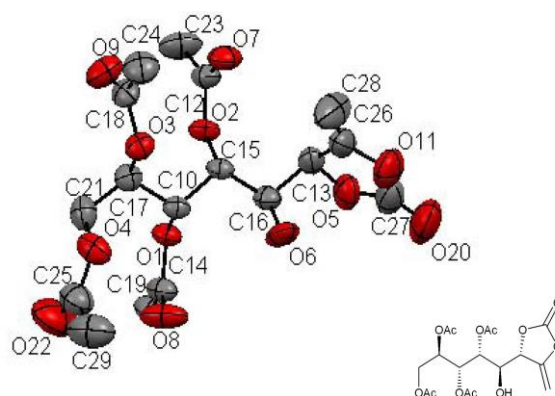
metilen-2-okso-1,3-dioksolan-4-il)pentan-1,2,3,4-tetraacetat, čija je strukturna formula prikazana na slici 2.



- Hemijska veza orijentisana prema posmatracu (ispred ravni hartije)
- Hemijska veza orijentisana suprotno posmatracu (iza ravni hartije)
- Hemijska veza koja se nalazi u ravni hartije

Sl. 2. Strukturna formula (2R,3R,4R,5S)-5-hidroksi-5-((S)-5-metilen-2-okso-1,3-dioksolan-4-il)pentan-1,2,3,4-tetraacetata

Ovo jedinjenje je sintetisano prema standardnom eksperimentalnom postupku i dobijeno je u vidu bijelih igličastih kristala. Kristali su pažljivo prečišćeni metodom prekrizalizacije iz smješe organskih rastvarača u cilju dobijanja reprezentativnog monokristala. Takav kvalitetan monokristal je poslužio da se uradi rendgenska strukturna analiza (difrakciona analiza), koja spada u veoma moćne, precizne i reprezentativne tehnike analize organskih molekula. Na slici 3 prikazan je ORTEP<sup>2</sup>-dijagram (2R,3R,4R,5S)-5-hidroksi-5-((S)-5-metilen-2-okso-1,3-dioksolan-4-il)pentan-1,2,3,4-tetraacetata.



Sl. 3. ORTEP-dijagram (2R,3R,4R,5S)-5-hidroksi-5-((S)-5-metilen-2-okso-1,3-dioksolan-4-il)pentan-1,2,3,4-tetraacetata (crvenom bojom su predstavljeni atomi kiseonika, a crnom atomi ugljenika; atomi vodonika su, zbog bolje preglednosti, uklonjeni sa strukture)

Kao što se na slici 3 vidi, difrakciona analiza daje uvid u tačan prostorni raspored svih atoma i veza kod datog molekula.

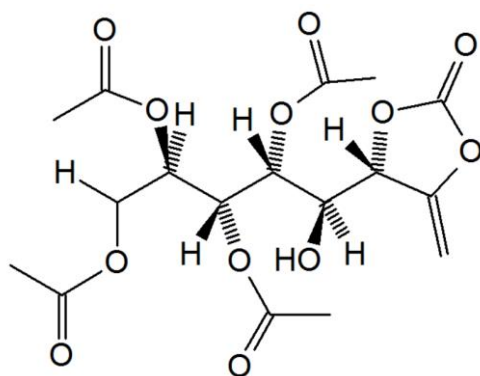
<sup>2</sup> ORTEP - Oak Ridge Thermal-Ellipsoid Plot Program

<sup>1</sup> Primjena vode kao rastvarača u hemijskoj sintezi je veoma poželjna kako iz ekoloških i zdravstvenih razloga (izbjegavanje korišćenja toksičnih, zapaljivih, kancerogenih, mutagenih organskih rastvarača) tako i iz ekonomskih (voda je daleko jeftinija nego organski rastvarači).

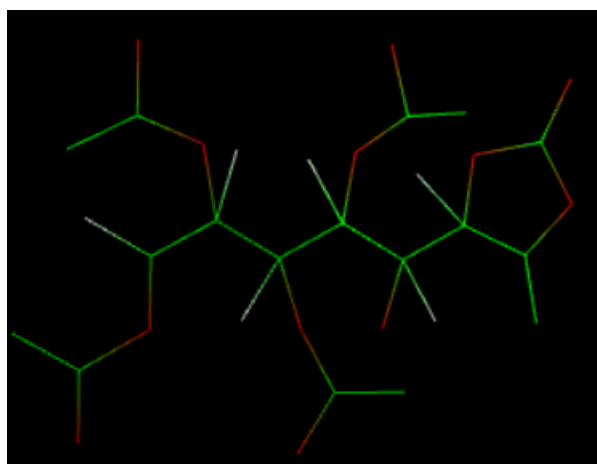
### III. MODELOVANJE MOLEKULA POMOĆU HEMIJSKIH SOFTVERSKIH ALATA

Program VegaZZ omogućava pripremu ulaznih fajlova za programski paket NAMD, na račun čega se generiše topologija molekula i to na veoma jednostavan način i bez upotrebe kompleksnih softverskih sistema.

Enol-karbonat, (2R,3R,4R,5S)-5-hidroksi-5-((S)-5-metilen-2-okso-1,3-dioksolan-4-il)pentan-1,2,3,4-tetraacetat je najprije skiciran korišćenjem programskog paketa ChemScetch, koji se rutinski koristi za crtanje strukturnih formula organskih i neorganskih molekula. ChemScetch ima mogućnost čuvanja nacrtanih molekula u standardnim hemijskim formatima, kao što je MDL molfile. MDL molfile sadrži sve informacije o atomima, vezama i njihovim tipovima, kao i koordinatama atoma u molekulu. Na slici 4 je prikazana struktura posmatranog molekula u ChemScetch-u, dok je na slici 5 data slika molekula u programu VegaZZ, sačuvana u MDL molfile formatu.



Sl. 4. Izgled (2R,3R,4R,5S)-5-hidroksi-5-((S)-5-metilen-2-okso-1,3-dioksolan-4-il)pentan-1,2,3,4-tetraacetata u programu ChemScetch

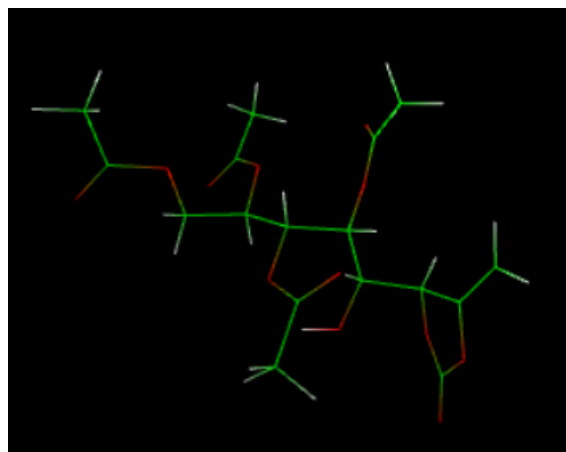


Sl. 5. Izgled (2R,3R,4R,5S)-5-hidroksi-5-((S)-5-metilen-2-okso-1,3-dioksolan-4-il)pentan-1,2,3,4-tetraacetata u program VegaZZ (MOL format)

Sa slike 5 se može uočiti da posmatrani molekul ima planarnu strukturu, te je potrebno izvršiti optimizaciju njegove geometrije u cilju dobijanja optimalnog rasporeda

atoma u 3D-prostoru, kako bi se molekul pripremio za dalje simulacije sa odgovarajućim proteinima. Ovo je moguće odraditi pomoću open source programa OpenBabel, [15]. Pored geometrijske optimizacije, OpenBabel pruža mogućnost konverzije molekula iz jednog standardnog hemijskog formata u drugi.

Na slici 6 je prikazan posmatrani molekul nakon geometrijske optimizacije i konvertovanja u standardni format PDB (protein data bank).



Sl. 6. Izgled (2R,3R,4R,5S)-5-hidroksi-5-((S)-5-metilen-2-okso-1,3-dioksolan-4-il)pentan-1,2,3,4-tetraacetata u program VegaZZ (PDB format)

Može se zaključiti da je upotrebom pomenutih softverskih alata generisana struktura koja, prema prostornom rasporedu atoma i atomskih grupa, u potpunosti odgovara rezultatu koji je dobijen eksperimentalnim mjerenjem pomoću difrakcione analize (struktura, prikazana na slici 3). Na osnovu toga se može zaključiti da ovaj model daje tačnu strukturu i da je isti potpuno korektan za dalje simulacije. Na osnovu ovog modela se mogu dobiti validni rezultati simulacija što bi imalo izuzetan značaj: ukoliko bi se ispostavilo da ovaj molekul zaista ispoljava posebne vrste interakcija sa receptorskim proteinima ćelijske membrane (to jest, da posjeduje fiziološku aktivnost), onda bi se moglo krenuti u laboratorijsku sintezu organskog molekula i eksperimentalno ispitivanje njegovih interakcija. Ako tome dodamo i činjenicu da ovakav vid aktivnosti za navedeno jedinjenje, pa i cijelu grupu njegovih homologa, ne postoji, značaj dobijenih rezultata postao bi i veći. Implementacija bi se svakako našla u aktuelnoj strategiji razvoja novih lijekova (farmaceutska hemija), aditiva u hrani (hemija hrane), raznih kofaktora (vitaminologija, kozmetologija), sredstava u borbi protiv bolesti i štetoina biljaka (agronomija, agrohemija) i mnogim drugim oblastima. Na osnovu interakcija, koje bi bile evidentne u rezultatima simulacija, mogli bismo pouzdano naći i novo polje primjene klase molekula, u koju spada i ispitivani. Sa druge strane, izmjenama strukturnih fragmenata već postojećih molekula dobili bi se novi (koji nisu sintetisani) i nad njima sprovele identične simulacije. Ukoliko bi se ispostavilo da su novodobijeni rezultati još bolji, postojao bi osnov za laboratorijsku sintezu i takvih

molekula, a potom i njihovih daljih ispitivanja u biološkim testovima.

#### IV. ZAKLJUČAK

Postojanje računarskog modela, koji bi mogao predvidjeti postojanje nekovalentnih i kovalentnih interakcija sa poznatim i strukturno dobro definisanim proteinima dalo bi ogroman doprinos iz najmanje dva razloga – proučavane bi interakcije organskih molekula koji prethodno uopšte nijesu sintetisani u laboratoriji (čime se značajno skraćuje i vrijeme istraživanja, ali i cijena samog procesa) i bili bi dostupni podaci o mogućim novim strukturnim jedinicama samog molekula, koje bi bile odgovorne za dodatne ili eventualno još jače interakcije sa receptorima (proteinima) ćelijske membrane. Na taj način bi se i same farmakološke osobine ispitivanog molekula daleko poboljšale. Organskom sintezom bi se, dalje, takvi molekuli mogli sintetisati u laboratoriji u količinama koje bi bile dovoljne za preklinička i klinička ispitivanja, a na osnovu molekulskih proračuna.

Razumijevanje ovih interakcija, njihovo detaljno opisivanje i kreiranje molekula koji su strukturno bliski ispitivanim predstavljaju budućnost ovog polja računarske hemije i molekulskog modelovanja.

#### ZAHVALNICA

Ovaj rad je realizovan u sklopu aktivnosti na projektu „Virtual Research Environment (VRE) in Southeast Europe and the Eastern Mediterranean (SEEM) – VI-SEEM“ na čijoj su infrastrukturi rađene simulacije.

#### LITERATURA

- [1] Proteinske banke podataka: RCSB Protein Data Bank: <http://www.rcsb.org/pdb/home/home.do>
- [2] Ž. Čeković i saradnici "Molekuli u tajnama života i svetu oko nas", 1. Izdanje, Zavod za udžbenike, Beograd, 2009, strane 139-149.
- [3] M. Nelson, W. Humphrey, A. Gursoy, A. Dalke, L. Kalhe, R. D. Skeel, K. Schulten, „A Parallel, Object-Oriented Molecular Dynamics Program”, Theoretical Biophysics Group University of Illinois and Beckman Institute 405 North Matthews Urbana, IL 61801 June 10, 1996.
- [4] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J. Berendsen. "GROMACS: fast, flexible, and free". *J Comput Chem.* Vol.16, 2005., 1701–18.
- [5] <https://vi-seem.eu>
- [6] E. Athanasiadis, Z. Cournia, G. Spyrou, „ChemBioServer: a web-based pipeline for filtering, clustering and visualization of chemical compounds used in drug discovery”, *Bioinformatics* Vol. 28, No. 22, November 2012, pp. 3002-3003.
- [7] E. Lionta, G. Spyrou, D. K. Vassilatis, Z. Cournia, „ Structure-based virtual screening for drug discovery: principles, applications and recent advances“ , *Curr. Top. Med. Chem.*, Vol.14, No. 16, Oct. 2014, pp. 1923-1938.
- [8] M. Mansha, U. U. Kumari, Z. Cournia, N. Ullah, „ Pyrazole-based potent inhibitors of GGT1: Synthesis, biological evaluation, and molecular docking studies”, *Eur. J. Med. Chem.*, Vol. 124, Nov. 2016, pp. 666-676.
- [9] Y. Wang, P. Gkeka, J. E. Fuchs, K. R. Liedl, Z. Cournia, „ DPPC-cholesterol phase diagram using coarse-grained Molecular Dynamics simulations”, *Biochim. Biophys. Acta*, Vol. 11, Nov. 2016, pp. 2846-2857.
- [10] M. Bigovic, V. Maslak, Z. Tokic-Vujosevic, V. Divjakovic, R. N. Saicic, „A useful synthetic equivalent of a hydroxyacetone enolate”, *Organic Letters*, vol. 13, July 2011, pp. 4720-4723.

- [11] S.S.V. Ramasastry, H. Zhang, F. Tanaka, C.F. Barbas, „Direct Catalytic Asymmetric Synthesis of anti-1,2-Amino Alcohols and syn-1,2-Diols through Organocatalytic anti-Mannich and syn-Aldol Reactions”, *J. Am. Chem. Soc.* Vol. 129, Jul. 2007, pp. 288–289.
- [12] M. M. Market, M. Mulzer, B. Schetter, R. Mahrwald, „ Amine-Catalyzed Direct Aldol Addition”, *J. Am. Chem. Soc.* Vol. 129, Nov. 2007, pp. 7258–7259.
- [13] J. Li, S. Luo, J. P. Cheng, „ Chiral Primary–Tertiary Diamine Catalysts Derived From Natural Amino Acids for syn-Aldol Reactions of Hydroxy Ketones”, *J. Org. Chem.* Vol. 74, Avg. 2009, pp. 1747-1750.
- [14] M. Bigovic, S. Skaro, V. Maslak, R. N. Saicic, „Expanding the scope of the indium-promoted allylation reaction: 4-(bromomethyl)-1,3-dioxol-2-one as a synthetic equivalent of a 3-aryhydroxyacetone enolate”, *Tetrahedron Letters*, No. 54, September 2013, pp. 6624-6626.
- [15] Open Babel: The Open Source Chemistry Toolbox, [http://openbabel.org/wiki/Main\\_Page](http://openbabel.org/wiki/Main_Page).

#### ABSTRACT

This paper is based on a comparison of the molecular model obtained by appropriate software tools, and experimental results on the basis of which the exact three-dimensional structure of organic molecules tested. Using a strategy of organic synthesis, it was synthesized molecule, for which we recorded diffraction analysis and obtained exact arrangement of atoms in space. Then, by using the different chemical software tools, the same molecule was geometrically optimized into standard chemical format and visualized. In this way, it has been shown that the arrangement of all the atoms is identical to results obtained by experiments. Taking into account the high accuracy in setting experimental and computational methods, it is possible to continue with molecular simulation and testing the selected molecules interaction with receptor proteins, with high confidence that the results can be interpreted in practice.

#### COMPARISON OF COMPUTER MODEL OF MOLECULE WITH EXPERIMENTAL RESULTS

Miljan R. Bigovic, Zarko Zecevic, Luka Filipovic and Bozo Krstajic

# Key Performance Indicators Visualization through “BIMP” application

Bojan Đaković, *CyberCo., Podgorica*, Budimir Lutovac, *Faculty of Electrical Engineering, Podgorica*

**Abstract** — This paper presents the “Business Intelligence and Performance Measurement” application, which main goal is to visualize key performance indicators of banking sales sector. BIMP is developed by authors of this paper with the aim to provide quick and simplified presentation of qualitative and quantitative performances on various levels, as well as comparative data, whether it is a comparison of entities within organization, or accomplished goals in relation to predefined goal.

**Keywords** — Business intelligence, Key Performance indicators, Management Information Systems, Performance measurement, Visualization of business activities .

## I. INTRODUCTION

**B**USINESS process automation, electronic data interchange, as well as availability of enormous amount of archived operational data, allows effortless business, but it also enables trend following and measuring performances of a certain business aspect. Special group of IS (Information System) is BI (Business intelligence). BI requires three main categories of technology - data warehouses, analytical tools, and reporting tools. [1]

Focal points of this assignment are some of the segments of the system, precisely, measuring business performances and visualization of key performance indicators for the business objectives

Modern trends in economy, as well as in other scientific fields, promote overview of available data by checking so-called key indicators (KPI - Key Performance Indicators) which represent qualitative or quantitative status display that includes singular indicators and their weighting coefficient. On the microeconomic level, there is a standard set of KPI, divided by fields of business, and also there is an infinite number of the same KPI derived from the primary group and adapted to specific needs of the system that is being observed.

This assignment will present an applicative solution for a segment of banking described above.

Bojan Đaković is with the Cyber Co , Studentska 6/13, 81000 Podgorica, Montenegro (phone: 382-68-040359; e-mail: djakovic@t-com.me).

Budimir Lutovac is with the Faculty of Electrical Engineering, University of Montenegro, Džordža Vašingtona bb, 81000 Podgorica, Montenegro (phone: 382-20-245839; e-mail: budo@ac.me).

## II. DATA AGGREGATION AND ITS IMPLEMENTATION

Data, which are used for testing the application, are from the core system from one of the Montenegrin banks, taken for a six-month period. It is clear that the basic issue when generating a report concerns system performances for such amount of data. Considering that every table includes up to ten fields for data classification, and that generating reports itself impacts work performances of the bank system, there is a demand for data aggregation on a periodical level, depending on the intended report type

Non-efficient aggregation data is one of the core problems with reports, and since about 90% of bank reports include aggregated data, it creates a demand for a proactive work to collect these data and further improve performances in generating reports and data analysis.

Architecture of the aggregator should have a number of features that should be common in the most cases in the implementation process of this kind of system. Successful aggregator’s architecture also implies a one-way connection to the main base. So there is a whole bunch of features that an aggregator should have in order that the solution becomes effective. Based on previous knowledge and experience in particular case those are:

**Flexibility:** Flexible aggregator architecture should enable import of a new data series. In this particular case, a core system is enabled to form the tables or modify them, if such a demand occurs. Both databases (core and BIMP) runs on the same server, this is established by enabling roles on BIMP database.

**Performances:** Although the summary tables contain the data, performances of the core system and its inner structure may also affect the speed of generating a report. All the tables in the aggregator’s database are carefully indexed, whereat a referential integrity was taken into account and foreign keys were formed where necessary. Instead of going to the cash memory in the *Views*, the data is dynamically created on application level, where the necessary calculations are also done.

**Scalability:** Aggregated data refer to a six-month period, and upon initial data aggregation , transfer for that whole period is completed. A feature that every aggregator should have is scalability. The number of data must grow and increase in time. Each and every store procedure transfers data incrementally, so that aggregation time is shortened, while the applicative solution takes a subset referring to a chosen time period. This is done by creating daily and monthly scheduled jobs.

**Effective resource utilization:** An effective aggregator is considered to be the one which has none or insignificant resource increase. Efficiency in resource utilization is achieved by shifting the time of aggregation and data transfer to the late night hours when the data are not used in current daily activities, thus avoiding the need to increase processor performances. Furthermore, aggregated data consume very small amount of memory and may be saved in the existing storage.

Finally, we should also mention some facts that contribute to creating of an aggregator of the data and forming the categories of the data subsets. In the case where the whole data set should be used to work with, that is if the queries are generated in the core system itself, the speed could not be even approximately close. Besides, that would require an investment in the increase of hardware performances, incessant re-indexing of production data, forming summary tables within core database. It would also cause safety risks because of the applicative solution access to the core database.

### III. KEY PERFORMANCE INDICATORS

Every company has an option to assemble and present a large number of measurable data to its management. However, if that task is executed unsystematically, the result would be clogging from a numerous reports, making it hard to recognize the relevant ones from the irrelevant. In these cases, a clear distinction should be made among relevant indicators and those without essential importance. *Key performance indicators* are those indicators that focus on the aspects of organizational performance that are the most critical for the current and future success of the organization. [2].

In order to identify indicators that could be considered as the key ones, it is vital to have a predefined objective and a strategy to which measuring of the mentioned indicators contributes. Second step is to choose specific set of KPI and adjusted to the special needs. In this case, Bank as a customer, formed project team to determine those KPIs that applies to strategic objective, which was a change in organizational culture of the employees in the branch network sector. Team consisted of two banking experts, one HR expert and author of this paper, based on their skills and previous knowledge on this matter. Previous used and desired report set were also considered during identification.

It was planned that the most of the KPI, if possible, were observed from the individual point of view within its organizational unit (a branch), as well as at the level of the whole bank. Some values were pondered for a more objective presentation, i.e. a better insight into the work circumstances of the employee. On that level, the same indicators were derived at a branch level, and from the same data, the managers reports are generated later, summing all the data in one place, ranking the same, etc.

Graphic presentation and evaluation method for some of the KPI, will be described in detail through a review and instruction for the user interface of applicative solution.

Short description of KPI which are defined as relevant, and subsequently implemented in this project, follows:

**Portfolio credit rating:** This indicator refers to the quality of a complete portfolio at the level of a manager of an organizational unit. Beside quantitative insight, reflected in a number of loans that makes a manager's portfolio, his monetary equivalent, an indicator provides qualitative data, in the form of quality of distributed funds.

**Loan production rate:** The indicator that displays invested funds until the certain date in the current year at the branch level. Except just showing number and amount of the distributed funds, the data are also compared by number and amount via arithmetic mean residuals, and quality of work is checked by measuring individual performances and it's appearance within value inside or outside the standard deviation derived in comparison to average values of all the employees in a branch.

**Efficiency of operations:** As Previously described indicator, this indicator treats a subject of a number of executed transactions of a bank teller at the level of a branch or a bank, with a possibility to present only chosen types of operations.

**Card coverage ratio:** Indicator checks a number of active debit and credit cards in relation to a number of active clients. Instead of a total number of branch clients and total number of the cards, only the active clients of an organizational unit are being verified, which are those who had at least one transaction in the last three months from the day of review. Percentage of coverage, which is a ratio between these two values, is calculated. This method provides a fair pondered review, because it also takes into account a number of available clients.

**E-banking coverage ratio:** Identically as in previous case, a number of users of e-banking services in relation to a number of active branch clients is verified.

**Deposit production rate:** As in Loan production rate paragraph, this indicator points out a problem of the deposits taken. Quantitative data of a deposit (number and amount) are summed up, compared at the branch level, and also work efficiency is verified at this segment.

**Branch growth ratio:** The indicator shows the information about participation of some branches in the growth of active clients in regard to total number of clients, as well as a coverage percentage, arithmetic mean residues and standard deviations.

**Customer involvement rate:** This one indicates a number of active products of active clients, ratio of these values and comparative data at the bank level. It also gives the information of a number of clients and active products.

**Attrition rate:** This indicator is defined to compare two periods. In this particular case it compares a current month to the previous, offering quantitative data about a number of active clients and products, comparing those data and showing inflow/outflow for the clients as well as for products.

### IV. APPLICATIVE SOLUTION



Application *Business Intelligence and Performance Measurement* is developed for the purposes of Prva Banka CG and it is intended for measuring performances and tracking KPI in banks.

It represents a tool for the managers and employees that enable them to have a quick and precise overview of the bank activities results at different levels and in different time periods. BI provides the ability to analyze business information in order to support and improve management decision making across a broad range of business activities. [3].

BIMP application is developed in Delphi 10.1 Development environment. Most of calculating functions and procedures, as well as user interface features and security issues are programmed by authors of this paper. Beside standard set of components, it also uses TMS Smooth Control library for visualization. In the background is SQL Server 2008 database with tables which represents aggregated data from core database. It does not include views and stored procedures and they are generated dynamically according to user interaction. User roles are integrated with domain users, but it provides standalone security also if needed.

Application's user interface is divided in four segments (Fig. 1), the last of which includes the forms for defining objectives, while the other three constitute groups of reports thematically divided, and those are: Executive reports and KPIs (intended for the bank management), Sales Quality KPIs (which visualise qualitative aspect of business) and Sales Quantity KPIs (provide quantitative business review).



Fig. 1. Main menu of BIMP application

At the level of individual forms, each of which contains one or more indicators, user interface will be presented, as well as the given parts of the code, used to calculate the desirable data, and afterwards be presented to the end user.

#### A. Qualitative KPIs

Owing to lack of space here, only some of the representative forms will be displayed. Therefore, to present qualitative KPI we will use Portfolio credit rating.

In addition to transfer of the basic data, there are daily transfers of single credit parties and calculations of a delay for every single party. As it can be seen on the Figure 2, taken from that transaction table, a summary table is created by a simple grouping query. The content of the

query changes depending on a user interaction with a form.

Scrolling through this table, the application event starts additional procedures that bring tabular or graphic displays, such as: a graphic portfolio display based on loan categories, a timeline of a portfolio (monthly increase or decrease), a timeline of a quality of portfolio itself, a graphic interpretation of the value of a class portfolio, as well as standard deviation based on percentage of reserved funds at the bank level and manager's position in the interpretation.

A portfolio of a manager consists of all the invested funds through loans. Every loan party initiates reserving of the funds that serve as collaterals in the case where client does not pay annuities regularly, so that a percentage of the reserved funds depends on a number of days of the delay and then is classified. In the final, a percentage of completely reserved funds in relation to invested funds define the quality of the portfolio itself, at the level of a manager or a bank. This form of a display enables us to not to get into a detailed analysis, but to direct our time and attention to portfolios which quality deviates from the desirable objective, with a simple scrolling and browsing, observing just a graphic division of class portfolios. On the other hand, timelines can reveal a great deal about quality of work in individual cases. The timeline of a portfolio also displays a direct connection of an investment with its quality, that is, if invested funds have downgraded the quality of a portfolio. Finally, the indicator with a pointer shows if a specific manager work in relation to his colleagues. Blue section in the circle is formed by calculating the center of the circle by taking an ideal percentage of the reservation (0) and a dislocation obtained by calculating standard deviation percentage of the reserved funds. The position of a pointer is determined as a dislocation from a center in percentage value of reserved funds.

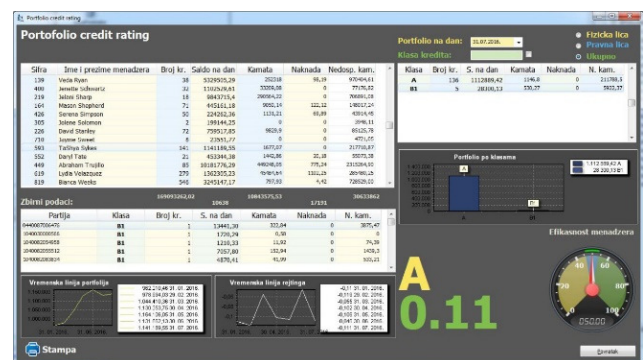


Fig. 2. Portfolio credit rating visualization

#### B. Quantitative KPIs

Quality of performed work can be quite different from its quantity. In order to illustrate that, we will use another form that is engaged with loans.

A portfolio is made of delivered loans. A form Loan production rate (Figure 3.) deals with utterly different dimension of this field. Displayed data provide summary data of invested funds at the branch level. Graphic display

enables to see individual contribution of employees, whether it is amount or number of loans.

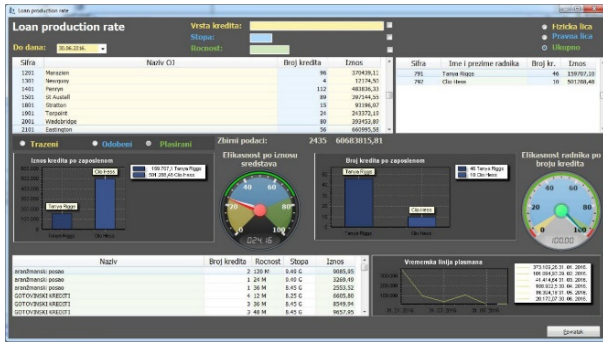


Fig. 3. Loan production rate visualization

Graph gives a comparative review of employees in a branch, while indicators with a pointer express efficiency of an employee in a branch, considering standard deviation. Timeline presents investing funds monthly throughout a current year, thus enabling trend following and engagement of an employee, or even a branch. In a similar way, the other sales segments are interpreted, so Efficiency of operations form provides an insight into a work of a bank teller and a number of executed transactions in the chosen time period. Unlike the previous form, an employee's position is visualized here, not only at the branch level, but at the bank level as well. Furthermore, a dynamic selection of observed transaction categories is enabled, so that a series of necessary reports could be done with one form (Figure 4.).

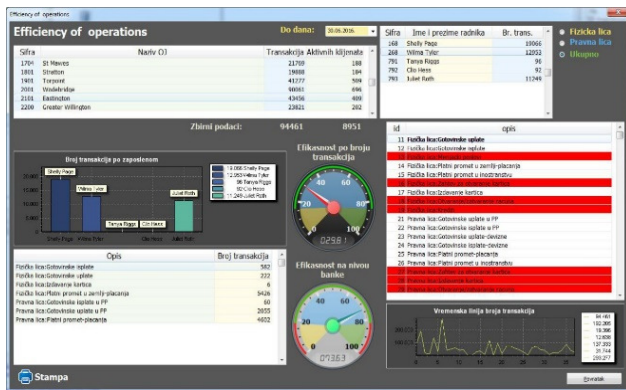


Fig. 4. Efficiency of operations form

C. Management reports

For the purposes of management, who are the main group of users, every single indicator should be in the reports that in a short time give global overview.

Depending on its goal, more kinds of these reports is developed. For example, **Branch overview** gives an overview of every relevant segment of the previously defined and individually described forms at the branch level. (Figure 5.)

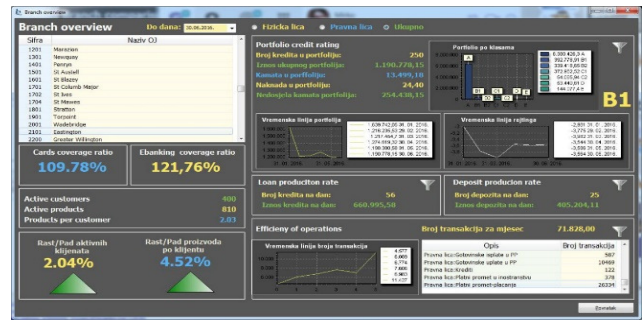


Fig. 5. Branch overview form

As one of the dominant goals of applicative solution is measuring realization of the objectives, as well as comparing performances of one or more entities with similar features (Figure 6.).

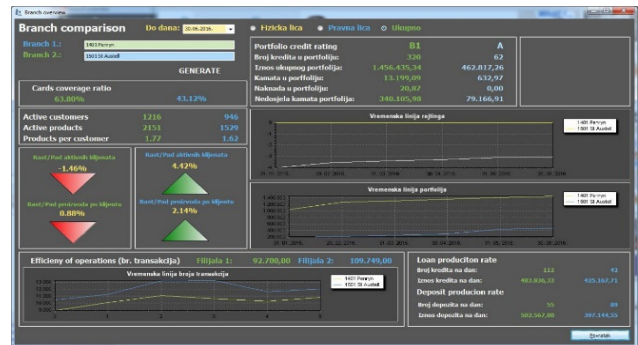


Fig. 6. Branch comparison form

Next to numeric display, the comparative data are presented in a graph as well, for a better comparison, and in a form that tracks down a realization of a predefined objective, the data are shown as visualized percentage of every category.

V. CONCLUSION

A solution presented in this paper is created with a clear purpose and intent to measure one segment of the business. Its focus restricts a potential to use a whole spectra of possibilities that this displaying method provides, either a total capacity of its implementation. The aim of this paper is to present a method how this technique can be applied in practice

Applicative solution may be used as groundwork for more projects in this field. Certainly, the closest one, based on a form and content, is visualization of the other aspects of business, in banking, as well as in any other field of economy.

REFERENCES

- [1] B Avzine, Z Cui, DD Nuck, B Majed, "Real time business for adaptive performance," in *IEEE Computer Society*, (CEC/EEE 06), 2006, pp. 1–8.
- [2] David Parmenter, "Developing, implementing and using winning KPIs" in *Wiley*, New Jersey 2010, pp. 3–42.
- [3] Elbashir Mohamed, Collier Phillip, Davern Michael, "Measuring the effects of business intelligence systems" in *International journal of Accounting IS*, Melbourne, 2008, pp. 135–153.

# Kvalitativna komparativna analiza Vodopadnog, Spiralnog i Scrum modela razvoja softvera

Ivan Mraković, *Fakultet za Informacione tehnologije, Univerzitet „Mediteran“, Podgorica*

**Sadržaj** — U IT praksi postoji čitav niz razvojnih modela softverskih projekata čija zastupljenost varira. Model koji će se koristiti mora biti takav da u konačnom potpuno zadovolji klijenta. Cilj ovog rada jeste stvaranje kvalitativne komparativne analize tri referentna modela: Vodopadnog, Spiralnog i Scrum. Svaki od njih ima svoje prednosti i nedostatke koji će biti posebno istaknuti. Komparacijom pomenutih modela dolazi se do zaključka koji od njih je najprikladniji za razvoj različitih softverskih projekata.

**Gljučne riječi** — model razvoja softvera, Scrum model, softverski projekat, Spiralni model, Vodopadni model

## I. UVOD

**P**RAKSA razvoja i upravljanja softverskim projektima poznaje niz metoda koje se primjenjuju u manjoj ili većoj mjeri. Svaka organizacija je u potrazi za metodom koja će joj donijeti najveće prednosti što je moguće brže i po nižoj cijeni. Međutim, ne postoji univerzalni model koji može zadovoljiti potrebe različitih softverskih projekata. To se posebno odnosi na velike, komplikovane projekte kao i na one čiji su zahtjevi loše specifikovani.

U procesu razvoja softvera izdvaja se nekoliko razvojnih faza od kojih svaka ima svoj cilj i krajnji rezultat a koje zajedno čine jedinstvenu cjelinu. Razvojne faze definisane su i raspoređene kako bi se problematika stvaranja softvera razložila na više manjih i jednostavnijih problema. U literalnim izvorima sreću se različiti modeli razvoja softvera i to tradicionalni, agilni i dr. Zajednički i osnovni koncept svih modela baziran je na prikupljanju zahtjeva, dizajniranju, implementaciji, testiranju, stavljanju u funkciju i održavanju gotovog sistema. Način i redoslijed izvođenja ovih faza čini metode softverskog razvoja jedinstvenim.

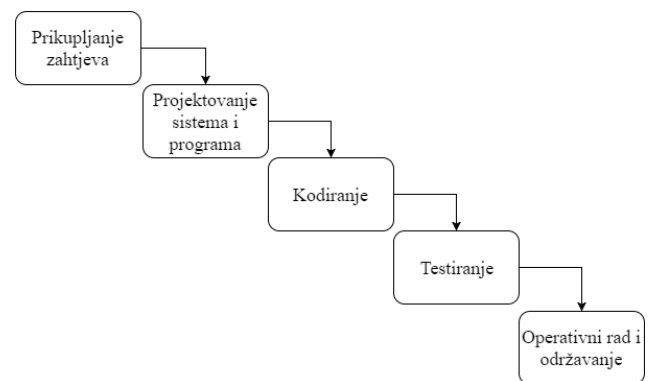
U ovom radu, opisana su tri modela razvoja softvera, čime se bavi naredno poglavlje. U trećem poglavlju dat je komparativni prikaz opisanih modela nakon čega su predstavljani zaključci.

## II. MODELI PROCESA RAZVOJA SOFTVERA

### A. Model Vodopada

Model Vodopada je praktično najstariji oblik modela razvoja životnog ciklusa softvera ali se i danas primjenjuje širom IT industrije [1-2]. Objavljivanjem svog istraživanja

u [1], Royce je otvorio novu eru softverskog inženjerstva koja pokazuje kompleksnost aktivnosti i međusobne fuzije članova razvojnog tima i značaj prepreka koje se javljaju tokom razvoja, a koje se moraju uspješno savladavati da uspjeh softverskog projekta ne bi izostao. Na sl. 1 predstavljene su razvojne faze ovog modela.



Sl. 1. Razvojne faze Vodopadnog modela

Na kraju svake faze pravi se presjek učinjenog sa ciljem utvrđivanja da li treba nastaviti razvoj projekta ili ne. Svaka naredna faza započinje tek po zaključivanju prethodne tako da ne postoji njihovo preklapanje.

Model Vodopada karakteriše potreba za detaljnošću što je ujedno prednost i nedostatak. Prednost je u tome što su korisnički zahtjevi objedinjeni i dokumentovani nakon faze prikupljanja istih, kao i to što svaku narednu fazu prati čitav set dokumenata. Međutim, za formulisanje dokumentacije potrebno je izdvojiti mnogo vremena što se ne može okarakterisati kao pozitivna strana modela.

Prikupljanjem zahtjeva postavljaju se jasno definisani ciljevi koji se u međuvremenu, dok se razvoj projekta kompletira, mogu pokazati nepotpunim i nedovoljno dobrim. Uz to, sa početkom izrade softverskog projekta iščezava povratna informacija i interakcija sa klijentima i naručiocima. Završetkom faze prikupljanja softverskih zahtjeva praktično prestaje mogućnost njihove dopune ili izmjene jer bi takav postupak u mnogome zakomplikovao i prolongirao završetak projekta.

U cilju otklanjanja sumnje da ne postoji dovoljno empirijske potpore za brojne teorijske tvrdnje o ovom modelu, u [3] je objavljeno istraživanje koje potvrđuje literalne navode.

Istražujući literaturu, primjenom kvalitativne komparativne analize napravljen je presjek prednosti i nedostataka Vodopadnog modela predstavljenih u tabeli 1.

TABELA 1: PREDNOSTI I NEDOSTACI VODOPADNOG MODELA.

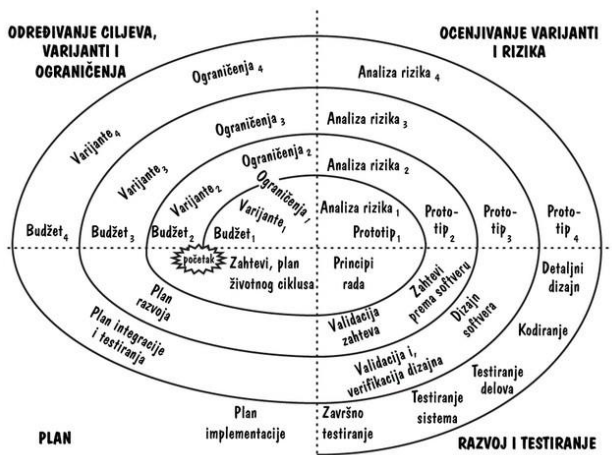
<b>Prednosti</b>
Zahtjevi sakupljeni na početku projekta [4], [24]
Jednostavan i lako razumljiv model [4]-[6]
Strogo definisan i jasan proces [2], [6]
Testiranje rezultata na kraju svake faze [1-6], [24]
Svaki korak je trajno dokumentovan [1-6], [24]
<b>Nedostaci</b>
Mnogo napora da se kompletira obimna dokumentacija koja prati svaku fazu razvoja [3-4], [24]
Faze iteracije i dizajniranja traju predugo [3], [5], [24]
Limitirana mogućnost izmjene zahtjeva [3-5], [24]
Velika mogućnost da krajnji proizvod ne odgovara korisnicima uslijed izmjene zahtjeva [3]
Korisnici nemaju mogućnost da daju povratnu informaciju dok sistem ne bude kompletiran [3-6]
Uloge u razvojnom timu oštro podijeljene što se može negativno odraziti na uspješnost projekta [3], [6]

Iako tradicionalan i ograničavajuć, model Vodopada je poslužio kao osnova mnogim drugim modelima softverskog razvoja upravo zbog svojih velikih nedostataka. Na temeljima Vodopada razvio se i Spiralni model kao oblik iterativnog razvoja.

### B. Spiralni model

Koncipiranje Spiralnog modela [7] bio je revolucionaran korak s obzirom da model povezuje Vodopad sa iterativnim i prototipskim modelom razvoja.

Kvalitet koji donosi Spiralni model osiguran je sprovođenjem niza aktivnosti poput analize rizika, razvoja prototipa, planova, validacije i verifikacije, integracije i testiranja. Svaka sledeća iteracija započinje tek kada je prethodna okarakterisana zadovoljavajuće. Model čine četiri razvojne faze prikazane na sl. 2.



Sl. 2. Razvojne faze Spiralnog modela [8]

Prolaskom kroz razvojne faze završava se iteracija – spirala. Na osnovu stavova klijenata i korisnika, proces razvoja softvera ulazi u narednu iteraciju kako bi se implementirale sugestije i želje klijenata koji po završetku prve iteraciju imaju jasniju viziju razvoja. Ponavljanje iteracija teče sve dok se ne zadovolje ciljevi projekta.

U [9] se Spiralni model definiše kao “meta-model” iz

razloga dobijanja prototipa na kraju svake iteracije. U [4] se kao glavna razlika Spiralnog i ostalih modela navodi eksplicitno prepoznavanje komponente rizika. Hronološki posmatrano, to je prvi model razvoja baziran na procjeni rizika, što zahtjeva ekspertsko znanje [10]. Prije započinjanja razvoja potrebno je detaljno analizirati faktor rizika što osigurava da u projekat ulaze samo izvodljivi softverski zahtjevi [11].

Pregled najvažnijih prednosti i nedostataka ovog modela dat je u tabeli 2.

TABELA 2: PREDNOSTI I NEDOSTACI SPIRALNOG MODELA.

<b>Prednosti</b>
Svaka faza se po završetku analizira što smanjuje mogućnost grešaka u projektu [7], [9]
Velika pokrivenost dokumentacijom [9]
Zahtjevi se lako mijenjaju tokom projekta [9], [12]
Dobro kontrolisan pristupni model, posebno sa aspekta komponente rizika [4], [7], [9], [13]
Pogodan za velike, rizične i kompleksne projekte [7], [9], [12-14]
<b>Nedostaci</b>
Stroga pravila i procedure razvoja [7], [9]
Obimna dokumentacija [9]
Proces procjene i pregleda učinjenog zahtjeva ekspertsko znanje [5], [7], [9], [13]
Nije prikladan za male i nisko-rizične projekte [9], [12-14]
Veoma skup model za kratkotrajne projekte [8], [14]
Cijena primjene često je visoka [8], [12]

Ovaj model praktičan je kada se primjenjuje za razvoj riskantnih projekata kod kojih su softverski zahtjevi komplikovani te kada cijena izrade igra važnu ulogu.

### C. Scrum model

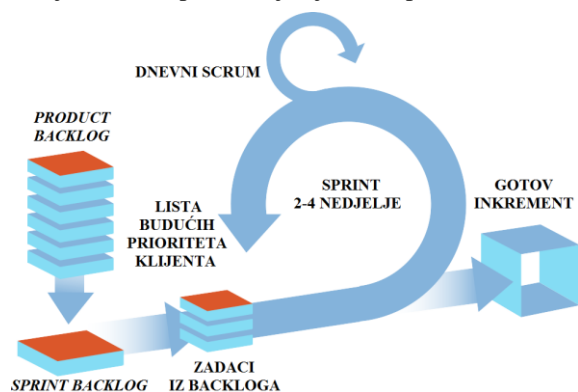
Učešće tradicionalnih metoda je sve rjeđe jer korisnici često ne znaju do detalja šta im je potrebno pa stoga pristupaju izmjeni zahtjeva tokom projektnog razvoja. Problem postaje preozbiljan onog trenutka kada korisnici shvate da softver nije ono što su očekivali tek nakon njegovog kompletiranja. Iz potrebe da se nadomjesti takvi nedostaci nastale su Agilne metode softverskog razvoja.

Agilne metode baziraju se na kratkim iteracijama kao fazama u procesu izgradnje softvera, organizovanim timovima i jednostavnijim dizajnom koji se odlikuje inkrementalnošću sa ciljem stvaranja radne verzije softvera na kraju svake iteracije. U literaturi se sreću brojne Agilne metode od kojih su najznačajnije Scrum, XP, RAD i FDD [15].

Rezultati istraživanja [16] pokazali su da među kompanijama koje primjenjuju Agilne metode razvoja, Scrum koristi 52% kompanija, metodu Scrum+XP 14%, dok ostalih 34% koriste druge Agilne metode kumulativno.

Scrum je kao agilna metoda predstavljen 1995. godine. Razvojne faze u okviru Scrum modela se preklapaju, a projekat se izvodi od strane multi-funkcionalnog tima kroz niz iteracija.

Scrum je okvir koji je namijenjen razvoju kompleksnih softverskih proizvoda. Sastavni djelovi Scrum-a su razvojni timovi, događaji, artefakti i pravila koja ih povezuju. Na sl. 3 predstavljen je koncept Scrum modela.



Sl. 3. Koncept Scrum modela [17]

Scrum tim ima zadatak da iznese glavni epitet metode a to je brzo i dinamičko prilagođavanje promjenama korisničkih zahtjeva. Scrum je jednostavan okvir koji se često opisuje kao "lak za naučiti ali težak za usavršiti" [18].

Izrada proizvoda dešava se u vidu tzv. Sprintova u okviru kojih se vrši implementacija tačno definisanih funkcionalnosti. Tokom trajanja Sprintsa, ne postoji mogućnost izmjene postavljenih zahtjeva. Sprints obično traju do 30 dana nakon čega se vrši prezentovanje dostignutih funkcionalnosti i postavljanje novih zahtjeva.

U tabeli 3 objedinjene su prednosti i nedostaci ovog modela.

TABELA 3: PREDNOSTI I NEDOSTACI SCRUM MODELA.

<b>Prednosti</b>
Agilnost [18-19], [21], [24]
Zahtjevi se dodaju u bilo kojem trenutku [13], [19], [22], [24]
Mogućnost povratka u prethodnu fazu [18], [21], [24]
Timovi fleksibilni [4], [21-22], [24]
Visoka uključenost korisnika što dovodi do smanjenja rizika od neuspjeha projekta [4], [13], [22], [24]
<b>Nedostaci</b>
Nema autoritativnog lidera jer Scrum Master ima sugestivnu ulogu [18], [20], [22], [24]
Nedovoljna korelacija sa korisnicima ili korisnička indiferentnost na kraju Sprintsa ugrožava projekat [13], [20], [22], [24]
Ukoliko članovi razvojnog tima ne saraduju kvalitetno ili se ne slažu dolazi do ugroženosti projekta [16], [20-21], [24]
Članovi razvojnog tima moraju biti multi-funkcionalni [18], [20], [22], [24]
Geografska ili vremenska razdvojenost članova razvojnog tima ugrožava uspješnost projekta [16], [20-21], [24]

### III. KVALITATIVNA KOMPARACIJA

Prezentovani modeli razvoja softvera prilično se razlikuju. Svaki ima za cilj stvaranje projekta koji će zadovoljiti korisnike i naručioce. Komparativnom kvalitativnom analizom sličnosti i razlika iznesenih u literaturi, došlo se do objedinjavanja rezultata predstavljenih u tabeli 4.

TABELA 4: UPOREDNI PREGLED MODELA RAZVOJA SOFTVERA.

<b>Model / Karakteristike</b>	<b>Vodopad</b>	<b>Spiralni</b>	<b>Scrum</b>
Kompleksnost	Mala	Srednja	Mala
Fleksibilnost	Mala	Srednja	Visoka
Uključenost korisnika u proces razvoja	Samo na početku	Da	Da
Mogućnost izmjene zahtjeva u bilo kojoj fazi	Ne	Da	Da
Razvojne faze	Sekven- cijalne	Iterativne	Iterativne
Dokumentacija	Obimna	Obimna	Sažeta
Rizik od neuspjeha	Visok	Nizak	Nizak
Primjenjivo na projekte čiji je rizik	Nizak	Visok	Visok
Vrijeme razvoja	Dugo	Kratko	Kratko
Cijena	Niska	Visoka	Niska
Uspješnost	Mala	Srednja	Visoka

Za razliku od ostalih, tradicionalni model Vodopada najčešće se opisuje kao jednostavan model zbog toga što ne postoji preklapanje niti vraćanje na prethodne faze razvojnog ciklusa. Njegova ograničena fleksibilnost, posebno zbog nemogućnosti izmjene i dopune korisničkih zahtjeva u kasnijim fazama, nadomještena je u modelu Spirale a još više u Scrum modelu.

Iterativna koncepcija razvojnih faza karakteriše modele Spirale i Scruma dok je kod Vodopadnog modela riječ o sekvencijalnim fazama.

Spiralni model je prvi u nizu modela koji se bave procjenom rizika. Takva osobina pravi drastičnu razliku u odnosu na model Vodopada gdje je rizik od neuspjeha jako izražen. Po istom kriterijumu, rizik koji sa sobom nose Spiralni i Scrum modeli je poprilično nizak.

Scrum prednjači ispred ostalih razvojnih metoda zbog svoje prilagodljivosti promjenama i česte interakcije sa korisnicima koji postaju aktivni učesnici razvojnog tima. Promjene korisničkih zahtjeva su dobrodošle u Scrum-u koji u cjelosti počiva na principima agilnosti. Suprotno tome, model Vodopada svoje principe zasniva na seriji uzastopnih koraka kako bi došao do finalnog rezultata, dok promjena korisničkih zahtjeva u nekoj od razvojnih faza može stvarati brojne poteškoće.

U [23] je dato poređenje agilnih modela, kakav je Scrum, sa tradicionalnim modelima:

- Pojedinci i interakcije naspram procesa i alata;
- Primjenljiv softver naspram detaljne dokumentacije;
- Saradnja sa klijentima naspram ugovornih aranžmana;
- Reakcija na promjene naspram pridržavanja plana.

Scrum je prikladan za realizaciju projekata nezavisno od njihove dužine i obima. U slučaju potrebe razvoja velikih projekata, dolazi do formiranja „Scrum of Scrums“ koji stvara više manjih timova kako bi se razvoj nesmetano odvijao.

Zbog svoje iterativne prirode i fleksibilnosti kao i mogućnosti izmjene zahtjeva u bilo kojem trenutku, Spiralni i Scrum modeli se mogu primijeniti na visokorizičnim projektima. Njihovo vrijeme razvoja je znatno kraće nego kod Vodopadnog modela. U slučaju pravilne implementacije i uigranog razvojnog tima, dobre saradnje sa korisnicima kao i dobro definisanih zahtjeva, cijena Vodopadnog kao i Scrum modela može biti relativno niska. Uz to, Scrum je više predvidiv od ostalih modela kako po pitanju cijene izrade projekta tako i po pitanju trajanja, s obzirom da se po završetku Sprints sagledava učinjeno kao i ono što treba učiniti u narednom Sprintu.

#### IV. ZAKLJUČAK

Razvoj softverskog projekta može biti riskantan i skup, posebno ukoliko se vrši česta izmjena korisničkih i softverskih zahtjeva. Veći projekat istovremeno znači povećan rizik, veći broj članova tima i veću potrebu za interakcijom među njima. Kod malih i kratkih projekata razlike među modelima praktično nisu od velikog značaja. Međutim, sa povećanjem obima razvojnog projekta razlike su izraženije. Model Vodopada može biti prikladan za jednostavne projekte minimalnog rizika kod kojih su zahtjevi jasno definisani. Spiralni i Scrum modeli mogu biti idealno rješenje kod rizičnih ili projekata gdje na početku ne postoji mnogo definisanih zahtjeva, te onda kada se u kratkom periodu želi doći do opipljivog rješenja. Posebno treba istaći prednosti Scruma koji se lako prilagođavanja izmjenama u bilo kojoj fazi razvoja, a uz to zadržava jednostavnost i nisku stopu rizika implementacije. Poređenjem modela razvoja softvera jasno je uočljiv hronološki napredak koji kulminira sa Scrum metodom. Komparacija predstavljenih modela ima za cilj da na osnovu jasno definisanih kriterijuma, prednosti i nedostataka, olakša izbor modela softverskog razvoja.

#### LITERATURA

- [1] W.W. Royce, "Managing the development of large software systems", in *Proc. IEEE WESCON*, Los Angeles, 1970, pp. 1-9
- [2] P.A. Laplante and C.J. Neill. (2004, Feb). The Demise of the Waterfall Model Is Imminent and Other Urban Myths. *QUEUE*. [Online]. 1(10), pp. 10-15. Available: <http://queue.acm.org/detail.cfm?id=971573>
- [3] K. Petersen, C. Wohlin and D. Baca, "The Waterfall Model in Large-Scale Development", *Lecture Notes in Business Information Processing*, Berlin: Springer, pp. 386-400 [PROFES, Oulu, 2009]
- [4] I. Sommerville, *Software Engineering*, 9<sup>th</sup> ed., USA: Pearson, 2011.
- [5] L.S. Pfleeger and M.J. Atlee, *Software engineering theory and practice*, preveli S. Jovanić, J. Gonda i R. Janković, Beograd: Računarski Fakultet i CET, 2006

- [6] M.D.Shah and C.Dhiman. (2016, Apr). Software Process Models Outline. *IJCSITS*. [Online]. 6(2), pp. 127-132. Available: <http://ijcsits.org/papers/vol6no22016/24vol6no2.pdf>
- [7] B.W. Boehm. (1988, May). A Spiral Model of Software Development and Enhancement. *Computer*. 21(5), pp. 61–72.
- [8] Internet izvor: <http://www.link-university.com/lekcija/Fazni-razvoj-i-spiralni-model/2620>, [pristupljeno 17.12.2016]
- [9] A. Mathur and A. Acharya. (2015, Nov). A Comparative Study on Utilization of Scrum and Spiral Software Development Methodologies: A Review. *IJERT*. [Online]. 4(11), pp. 90-94. Available: <https://www.ijert.org/view-pdf/14309/a-comparative-study-on-utilization-of-scrum-and-spiral-software-development-methodologies-a-review>
- [10] L.L. Ray. (2013, Aug). Security considerations for the spiral development model. *International Journal of Information Management*. 33(4), pp. 684– 686
- [11] S.I. Hashmi and J. Baik, "Software Quality Assurance in XP and Spiral - A Comparative Study", in *Proc. ICCSA*, Kuala Lumpur, 2007, pp. 367-374
- [12] M. A. Ben-Zahia and I. Jaluta, "Criteria for Selecting Software Development Models", in *Proc. GSCIT*, Sousse, 2014, pp. 1-6
- [13] R.C.G. Dhanajayan i S.A. Pillai. (2016, Aug). SLMBC: spiral life cycle model-based Bayesian classification. *Soft Computing*. 20(8), pp.1-13
- [14] R. Arora and N. Arora. (2016, Feb). Analysis of SDLC Models. *IJCET*. [Online]. 6(1), pp. 268-272, Available: <http://inpressco.com/wp-content/uploads/2016/02/Paper48268-272.pdf>
- [15] P. Bourque and R.E. Fairley, „Guide to the Software Engineering Body of Knowledge, version 3.0“, IEEE Computer Society, 2014, Available: [www.swebok.org](http://www.swebok.org)
- [16] N.K. Rad and F. Turley. (2014). *EXIN Agile Scrum Foundation Workbook*. [Online]. Available: <https://www.scribd.com/doc/263185194/EXIN-Agile-Scrum-Training-Manual>
- [17] S. Lancaster. (2013). *60 Minute:Scrum*. [Online]. Available: <https://www.scribd.com/read/211052847/60-Minute-Scrum#>
- [18] Internet izvor: <http://www.kaleida.co.uk/our-approach/>, [pristupljeno 28.12.2016]
- [19] B.L. Romano and A.D. da Silva, „A case study within a small enterprise“, in *Proc. ITNG*, Las Vegas, 2015, pp. 774-776
- [20] M. Tanner and A. Mackinnon. (2013, Jul). Sources of Disturbances Experienced During a Scrum Sprint. *Electronic Journal of Information Systems Evaluation*. [Online]. 18(1), pp. 3-18, Available: <http://www.ejise.com/volume18/issue1/p3>
- [21] J. Sutherland. (2010). *Scrum handbook*. [Online]. Available: <https://www.scribd.com/document/261003399/Scrum-Handbook>
- [22] P. Hundermark. (2015). *Do better Scrum*. [Online]. Available: <https://www.scribd.com/doc/294903971/Do-Better-Scrum>
- [23] Internet izvor: <http://agilemanifesto.org/iso/sr/manifesto.html>, [pristupljeno 22.12.2016]
- [24] G.V. Waardenburg and H.V. Vliet. (2013, Dec). When agile meets the enterprise. *Information and Software Technology*. 55(12), p. 2164.

#### ABSTRACT

There is a whole set of software development models in practice which usage vary. Model in use has to be shaped in such a manner to completely fulfill customers' expectations. The purpose of this article is to create qualitative comparative analysis of three reference models: Waterfall, Spiral and Scrum. Each of them has their own advantages and disadvantages which are emphasized. Conclusion is drawn by comparing mentioned models in a way to state which model is optimal for development of different software project.

#### QUALITATIVE COMPARATIVE ANALYSIS OF WATERFALL, SPIRAL AND SCRUM SOFTWARE DEVELOPMENT MODELS

Ivan Mraković

# Visokotehnoški (sajber) kriminal i anatomija bezbjednosnih rizika u informaciono-komunikacionim sistemima

Branko Džakula, *Elektrotehnički fakultet, Univerzitet Crne Gore, Student Member, IEEE*

**Sadržaj** — Rapidan razvoj informacionih tehnologija i interneta, uslovio je da informacija postaje jedan od ključnih resursa potražnje korisnika i faktor opstanka organizacija. Ovaj rad se fokusira na kompletno upravljanje informaciono-bezbjednosnim rizicima od procjene i analize, preko prioritizacije do same mitigacije rizika na primjerima bezbjednosnih rizika korištenjem online servisa. Upravljanje informaciono-bezbjednosnim rizicima predstavlja imperativ za organizacije sa velikim brojem korisnika putem interneta i najbolji pristup identifikacije i rješavanja svih bezbjednosnih prijetnji koje su gotovo konstantno prisutne posebno u današnjem svijetu velike ekspanzije visokotehnoškog (sajber) kriminala.

**Ključne riječi** — *bezbjednosni rizik; informaciona bezbjednost; onlajn servisi; cloud; visokotehnoški kriminal;*

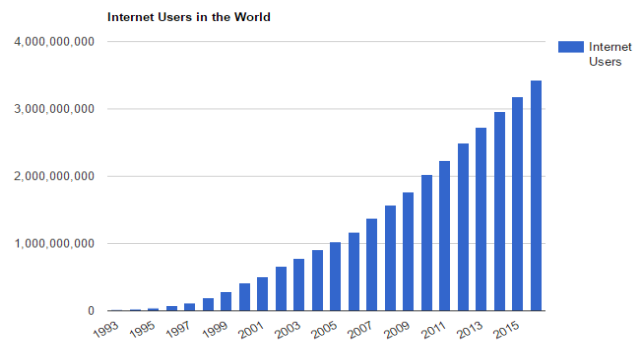
## I. UVOD

INTERNET tehnologije polako ali sigurno postaju dio društva u svakodnevnom prikupljanju i razmjeni informacija. Trend rasta internet korisnika u poslednjih 15 godina je više nego fascinirajući podatak. U svom istraživanju iz 2004. godine, Mark Andrisen (Marc Andreessen), pionir u oblasti komercijalnog Interneta, objavio je da u tom trenutku u svijetu ima oko 900 miliona korisnika interneta i da očekuje da ih do kraja 2020 bude blizu 3 milijarde. Danas se može zaključiti da je ovo predviđanje bilo krajnje skromno jer trenutno u svijetu ima oko 3,5 miliona korisnika po poslednjim podacima "Internet Live Stats" [1]. Trend rasta korisnika u svijetu se može jasno vidjeti na Slici 1. dok su internet korisnici u Crnoj Gori prikazani u Tabeli 1.

Internet je omogućio i slobodu rada na daljinu što su mnoge organizacije kao i pojedinci rado primijenili u svojim okruženjima u cilju efikasnijeg poslovanja, komfora u radu kao i manjih troškova. Ovakva praksa sve češćeg rada na daljinu preko Interneta uslovljava da informacije postaju jedan od ključnih resursa i faktora poslovanja, a komfor i efikasnost koji pruža Internet za sobom donosi niz informaciono-bezbjednosnih rizika koji predstavljaju ozbiljnu prijetnju po bezbjednost informacija, njenu povjerljivost, integritet i dostupnost.


Branko Džakula (autor za kontakte), Zgrade solidarnosti B-III stan 6, 85343 Herceg Novi, Crna Gora (telefon +382-69-010396, e-mail: branko@dzakula.com).

Informaciono-bezbjednosni rizik predstavlja osobu ili situaciju koja može potencijalno da ugrozi bezbjednost informacija u svim oblicima. Kroz ovaj rad će se koristiti termin "bezbjednosni rizik" koji će predstavljati informaciono-bezbjednosni rizik, a ne rizik vezan za fizičku bezbjednost organizacije, pojedinca ili informacionog sistema.



Sl. 1. "Internet Live Stats" – Trend rasta broja Internet korisnika u svijetu od 1993. god. do danas

TABELA 1. "INTERNET LIVE STATS" – BROJ INTERNET KORISNIKA U CRNOJ GORI

 <b>MONTENEGRO</b>	
Stanovništvo: 676,872	
Glavni Grad: Podgorica – stanovništvo: 136,473	
388,057 Internet korisnika (1. Jul 2016) – Penetracija 57.33% od ukupnog broja stanovništva	
320,000 Facebook korisnika 1. Jul 2016 – Penetracija 47.28% od ukupnog broja stanovništva	

Informacije se generalno definišu kao „podaci koji imaju vrijednost“ i koji trebaju biti adekvatno zaštićeni. Informaciona bezbjednost podrazumijeva očuvanje:

- povjerljivosti – da je informacija dostupna samo licima koja su ovlašćena za pristup toj informaciji;
- integriteta – zaštita postojanja, tačnosti i kompletnosti informacije kao i procesnih metoda;
- raspoloživosti – da autorizovani korisnici imaju mogućnost pristupa informaciji i pripadajućim sredstvima kada se usluga zahtjeva;

## II. VISOKOTEHNOLOŠKI KRIMINAL

Visokotehnoški kriminal, po definiciji Ministarstva unutrašnjih poslova Crne Gore, obuhvata skup krivičnih djela gde se kao objekat izvršenja i kao sredstvo za izvršenje krivičnog djela javljaju računari, računarske mreže, računarski podaci, kao i njihovi produkti u materijalnom i elektronskom obliku.

### A. Najčešći oblici kompjuterskog kriminala u Crnoj Gori

Prema nedavnom članku portala "Analitika" [2] na temu visokotehnoškog (sajber) kriminala u Crnoj Gori, prisutna su sva krivična djela iz oblasti visokotehnoškog kriminala. Kao krivična djela, koja su najviše prisutna navode se djela iz oblasti krađa identiteta, računarskih prevara, krađa e-mail naloga, kripto loker ucjene (najčešće distribuirane e-mail porukama), narušavanje prava intelektualne svojine, a tek onda dolaze krivična djela dječije pornografije, sajber špijunaže i klasičnih djela iz oblasti kompjuterskog kriminala kao što su (DDoS (Distributet Denial Of Service) napadi, SQL inekcije i sl.). Članak navodi da se procjenjuje da je šteta pričinjena vršenjem ovih krivičnih djela u Crnoj Gori u prethodnih nekoliko godina preko 1.500.000 eura, što je samo pričinjena šteta prijavljena od strane crnogorskih kompanija i građana. Sumnja se da je stvarna pričinjena šteta mnogo veća od zvanične.

### B. Legislativa Evropske unije u oblasti kompjuterskog kriminala

Kako je rastao broj korisnika interneta u Crnoj Gori, što se jasno vidi statistikom prikazanom u sekciji I, tako su se prilagođavali i crnogorski zakoni vezani za krivična djela „visokotehnoškog kriminala“: 2004. godine u Krivični zakonik implementirana su krivična djela iz oblasti računarskog kriminala. 2014. godine zakonom su obuhvaćena sva krivična djela iz oblasti visokotehnoškog kriminala koja su u potpunosti usklađena sa legislativom Evropske unije.

### C. Institucije u nadležnosti za borbu protiv visokotehnoškog kriminala

Država Crna Gora je išla u skladu sa novim trendovima iz oblasti informaciono-komunikacionih tehnologija i formirala Ministarstvo za informaciono društvo u okviru kojeg je formiran CIRT tim (CIRT – Computer Incident Response Team odnosno tim za odgovor na računarske incidente) koji je zadužen za koordinaciju i pomoć državnim organima kako bi se smanjio rizik od računarskih incidenata, kao i podizanje svijesti o sajber prijetnjama i sajber bezbjednosti. U svijetu trenutno postoji više od 250 CIRT timova.

Takođe, u MUP-u Crne Gore u Upravi policije, 2015. godine formirana je Jedinica za borbu protiv visokotehnoškog kriminala, koja se nalazi u sklopu Odsjeka za borbu protiv organizovanog kriminala i korupcije.

### D. NATO i projekti za borbu protiv visokotehnoškog kriminala

Visokotehnoški kriminal postaje sve organizovaniji, učestaliji i ciljno orijentisaniji. Nije nepoznato da organizovane kriminalne grupe koriste usluge profesionalnih hakera kako bi vršila određena krivična djela i na taj način „prala“ svoj nelegalno stečeni novac kao i vršila „sajber“ špijunažu i sl.

Država Crna Gora je dobila poziv za pristup NATO alijansi, a prema tvrdnjama stručnjaka samim pristupom imala bi veću mogućnost prevencije, otkrivanja, odbrane od „sajber“ napada kao i saniranju posljedica. Kao primjer se navodi NATO centar za taktičke odgovore (CDMA – Sajber Defence Management Authority), koji je smješten u Briselu, a bavi se koordinacijom „sajber“ odbrane članica NATO alijanse i posjeduje najmodernije i najnaprednije tehnologije za otkrivanje „sajber“ napada.

### E. Situacija u svijetu

Situacija u svijetu po pitanju sajber napada je na zabrinjavajućem nivou. Za najadekvatniji prikaz situacije se pobrinula kompanija NORSE, poznata po besplatnom online alatu Norse map [3] koji prikazuje sajber napade u svijetu u realnom vremenu, prikazujući i detalje svakog napada počevši od izvora i cilja napada (država), tipa napada (broj porta i tip servisa) do krajnje detaljne vizuelne reprezentacije ovih podataka (Slika 2.).

Posmatranjem Norse mape samo nekoliko minuta može se zaključiti koliko je ozbiljna situacija po pitanju visokotehnoškog kriminala i informacione bezbjednosti na globalnom nivou. Postoje mnoge statistike na temu sajber kriminala u svijetu i po do sada prikazanim statistikama u ovom radu jednostavno je zaključiti da je situacija u drugim državama, tj. prisutnost sajber kriminala veća zbog veće online prisutnosti korisnika. Podaci koji naviše zabrinjavaju javnost se svode na porast krađe identiteta gdje je tokom 2015-te godine broj ukradenih identiteta korisnika iznosio zapanjujućih 232 miliona, a trend rasta koji je primijećen iz godine u godinu dokazuje da 2016-ta neće prikazivati bolje rezultate.



Sl. 2. "Norse map" – Mapa visokotehnoških napada sa prikazom napada uživo na globalnom nivou



## III. UPRAVLJANJE BEZBJEDNOSNIM RIZICIMA

Rizik, sam po sebi, predstavlja potencijal da se dobije ili izgubi nešto od vrijednosti (zdravlje, socijalni status, materijalno bogatstvo, i sl.). Srodnije ovom radu, rizik u informacionoj bezbjednosti predstavlja potencijalni bespovratan gubitak informacija (pad sistema usljed hardverskih problema ili sajber napada), otuđivanje i/ili curenje osjetljivih informacija (lični podaci, finansijski podaci, korisnički kredencijali, itd.) i kompromitovanje informacija i povreda integriteta istih u vidu da korisnik raspolaže lažnim informacijama. Rizik u informacionoj bezbjednosti se može predstaviti uz pomoć dva osnovna parametra:

- **Posljedica** - Ovaj parametar se izračunava kroz potencijalni nivo gubitaka po organizaciju u slučaju da se rizik realizuje i najčešće se predstavlja sledećim modelom:

1	Umjeren Gubitak
2	Srednji Gubitak
3	Veliki Gubitak
4	Kritičan Gubitak

- **Vjerovatnoća** – Ovaj parametar predstavlja procijenjenu vjerovatnoću da se identifikovani rizik izvrši i najčešće je predstavljen sledećim modelom:





1	0% - 10%
2	10% - 30%
3	30% - 80%
4	80% - 100%

Međuzavisnost ovih parametara je očigledna i koristi se prilikom procjene rizika da se jasno definiše nivo rizika kao i prioritet u mitigaciji ili ublažavanju istog. Kombinacijom ova dva parametra dobija se matrica procjene rizika koja je predstavljena u Tabeli 2.

TABELA 2. MATRICA PROCJENE RIZIKA PO MODELU NAJBOLJE PRAKSE [7]

		Posljedica			
		Umjeren gubitak	Srednji Gubitak	Veliki Gubitak	Kritičan Gubitak
Vjerovatnoća	0% - 10%	1	2	3	4
	10% - 30%	2	4	6	8
	30% - 80%	3	6	9	12
	80% - 100%	4	8	12	16

Legenda:

Mali rizik:		Srednji rizik:		Veliki rizik:		Izrazito veliki rizik:	
-------------	---	----------------	---	---------------	---	------------------------	---

Proces identifikacije rizika, na osnovu clauzule 6.1.2. internacionalnog standarda ISO 27001:2013 [8], sačinjen je iz pet elemenata koji predstavljaju vodič ka kvalitetnom izvođenju aktivnosti identifikacije rizika i daje slobodu organizaciji da odluči najadekvatniji pristup koji zadovoljava zahtjeve ovih 5 elemenata:

- 1) Definisati način na koji će se identifikovati rizici koji mogu dovesti do povrede povjerljivosti, integriteta i/ili dostupnosti podataka.
- 2) Definisati način na koji će se identifikovati nosioci rizika
- 3) Definisati kriterijume po kojima će se procjenjivati vjerovatnoća i posljedice realizacije rizika
- 4) Definisati kako će se rizik proračunavati
- 5) Definisati kriterijume po kojima će se rizik prihvatati

Metode upravljanja, identifikacije i procjene bezbjednosnih rizika mogu biti sprovedene internom i/ili eksternom revizijom. Interna revizija podrazumijeva primjenu metoda od strane zaposlenih iste kompanije nad kojom se vrši revizija dok eksterna revizija podrazumijeva angažovanje trećih lica za primjenu istih metoda. Preporuka je primijenjivati oba pristupa radi dobijanja što preciznijih, objektivnijih i transparentnijih podataka. Metode koje se primijenjuju revizijom informacionih sistema u fazi procjene rizika su:

#### A. Intervju zaposlenih

Ovaj metod procjene rizika podrazumijeva prikupljanje informacija iz prve ruke od zaposlenih koji su angažovani na implementaciji, održavanju i/ili upravljanju određenim informacionim sistemima od velike važnosti. Intervju je najčešće koncipiran postavljanjem niza precizno definisanih pitanja kako bi se razotkrile potencijalne bezbjednosne prijetnje po informacioni sistem. Pitanja se fokusiraju na razotkrivanje načina upravljanja korisničkim nalozima, kreiranju lozinki, odobravanja pristupa sistemu, revizija IT arhitekture sistema, implementiranim bezbjednosnim kontrolama, konfiguracije sistema, pratećih procedura i papirologije o datom sistemu, kao i poznatih bezbjednosnih propusta kojih su zaposleni već svjesni.

#### B. Vulnerability Test (Test ranjivosti)

Testovi ranjivosti se primijenjuju korištenjem niza dostupnih alata u zavisnosti od karakteristika sistema. Alati poput Nessus-a [4] se koriste za testiranje internet aplikacija i internet stranica na sigurnosne propuste poređenjem primijenjenih konfiguracija na sistemu sa konfiguracijom najbolje prakse bezbjednosti kao što su OWASP [5] (Open Web Application Security Project) Top 10. Za druge informacione sisteme se koriste kombinacije raznih alata u zavisnosti od kompleksnosti sistema i mogu znatno da variraju (operativni sistemi, baze podataka, mrežni uređaji, itd.). Center for Information Security poznatiji kao CIS [6] posjeduje niz alata za bezbjednosnu provjeru sistema na ranjivosti kao i najprimijenjenije dokumentacije za bezbjednosnu konfiguraciju velikog

broja operativnih platformi, baza podataka i mrežnih uređaja.

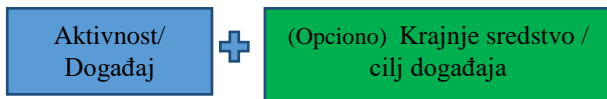
### C. Penetration Test (Penetracijski test)

Ovi testovi se vrše od strane specijalno obučeni osoba za etičko hakovanje. Penetracijski test podrazumijeva onemogućavanje pristupa ciljanom informacionom sistemu etičkom hakeru sa ciljem da uspije da pridobije neautorizovani pristup sistemu simulirajući visokotehnoški napad i time otkrivajući ranjivosti koje bi potencijalni haker mogao da iskoristi u budućnosti da na isti način pridobije pristup sistemu i time ga kompromituje. Nerijetko se angažuju treća lica za ovu vrstu testiranja jer se rijeđe sprovodi, obično u "go-live" fazi implemetacije sistema, tj. puštanja novog sistema u produkciju.

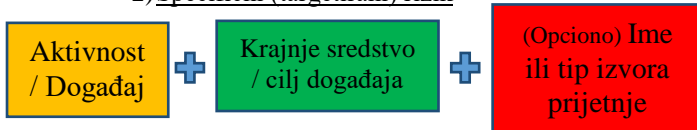
### IV. REGISTAR (KATALOG) RIZIKA

Nakon prikupljenih podataka tokom faza identifikacije i procjene rizika potrebno je i adekvatno imenovati rizik. Za ovu aktivnost postoje dva generalno prihvatljiva pristupa:

#### 1) Generalni ili globalni rizik



#### 2) Specifični (targetirani) rizik



Uprošćen primjer imenovanja rizika u registru (katalogu) rizika za scenario nedostatka implementiranog antivirus rješenja je prikazan u Tabeli 3.

TABELA 3. MATRICA PROCJENE RIZIKA PO MODELU NAJBOLJE

Rizik	Opis	Vjerovatnoća	Mitigacija
Povreda povjerljivosti, integriteta i/ili dostupnosti korisničkih podataka zbog nedostatka implementiranog antivirus rješenja	Nedostatak adekvatne antivirus zaštite sprečava mogućnost identifikacije i sprečavanja hakerskog napada na korisnički računar koristeći se malicioznom softverom	4	Implementacija minimum osnovnog (freeware) antivirus rješenja sa redovnim ažuriranjem definicija
		Posljedica	
		4	
		Ocjena	
		16	

U korporativnim okruženjima registar rizika sačinjava i niz dodatnih podataka o svakom riziku poput vlasnika rizika, procjena finansijskih i/ili reputacionih posljedica, detaljan mitigacioni plan, datume identifikacije, revizije i

mitigacije rizika i opcione druge podatke u zavisnosti od tipa rizika i industrije u kojoj kompanija posluje.

### V. ZAKLJUČAK

Zaključak većina istraživanja na temu informacione bezbjednosti jeste da 100%-tna zaštita ne postoji. Konstantan rad na unapređenju bezbjednosti informacionih sistema i podizanju svijesti o informacionoj bezbjednosti jesu ključne aktivnosti za umanjene vjerovatnoće da se identifikovani rizici izvrše. Rad na spremnosti svake organizacije da reaguje na događaj koji ugrožava bezbjednost informacija je ključan za omogućavanje nesmetanog poslovanja kao i sprečavanje velikih finansijskih gubitaka. Za krajnjeg korisnika je od izuzetne važnosti da poznaje osnovne prijetnje koje vladaju internetom i da se svakom i svakodnevnom korišćenju internet servisa pristupa sa visokom dozom obrazivosti.

### LITERATURA

- [1] InternetLiveStats.com Online Sources, (2016, December, 25<sup>th</sup>). [Website]. Available: <http://www.internetlivestats.com/internet-users/>.
- [2] J. Backović, Portal Analitika. (2016, December, 2<sup>nd</sup>) [Web news portal]. Available: <http://portalanalitika.me/clanak/224774/sajber-kriminal-u-crnoj-gori-krada-indentiteta-ucjene-djecja-pornografija>
- [3] Norse Corporation, Online Sources, (2016, December, 14<sup>th</sup>). [Website]. Available: <http://map.norsecorp.com/#/>.
- [4] Tenable Network Security, Online Sources, (2016, December, 14<sup>th</sup>). [Website]. Available: <https://www.tenable.com/products>.
- [5] Open Web Application Security Project, Online Sources, (2016, December, 14<sup>th</sup>) [Website]. Available: [https://www.owasp.org/index.php/Main\\_Page](https://www.owasp.org/index.php/Main_Page)
- [6] Center for Information Security CIS, Online Sources, (2016, December, 14<sup>th</sup>) [Website]. Available: <https://benchmarks.cisecurity.org/>
- [7] Jakeman Business Solutions, Publication, What's right with risk matrices?, Online Sources, (2017, January 31<sup>st</sup>) [Website] Available: <http://www.jakeman.com.au/media/whats-right-with-risk-matrices>
- [8] ISO, Online Sources (2017, January 31<sup>st</sup>) [Website] Available: <http://www.iso.org/iso/iso27001>

### ABSTRACT

The rapid development of information technology and the Internet, has caused that information is becoming a key resource user demand and factor of organizations survival. This work focuses on the complete management of information security risks of assessment and analysis, prioritization over themselves to mitigate risk to the examples of security risks using the online service. The management of information security risks is imperative for organizations with a large number of users over the Internet and the best approach to identifying and addressing any security threats that are almost constantly present especially in today's world of the great expansion of high-tech (cyber) crime.

### CYBER CRIME AND ANATOMY OF SECURITY RISKS IN INFORMATION-COMMUNICATION SYSTEMS

Branko Džakula, Faculty of Electrical Engineering, University of Montenegro, IEEE Student Member

# Home Automation System Control Based on Speech Commands

Katerina Cekova, Jasna Cekova, *Faculty of Computer Science, UGD - Shtip, R.Macedonia;*  
Saso Gelev, *Faculty of Electrical Engineering, UGD - Shtip, R.Macedonia*

**Abstract** — Home automation is one of the major growing industries that can change the way people live. Home automation gives the user the comfortable and easily for using the home devices to control automatically. Typical wireless home automation system allows controlling house appliances from an Android application. Also, home automation helps handicapped and old people which will enable them to control home appliances.

The developed system can be integrated into a single portable unit and allows the user to control appliances like lights, fans, electronic doors, electrical equipment's etc. The main goal of this system is to control home appliances by using voice commands.

The proposed system can recognize the voice commands, convert them into the required data format, and send the data through the Wi-fi on the microcontroller. This paper presents a design and prototype implementation of a home automation system that uses Wi-fi technology. This system is designed to be low cost and allowing different devices to be controlled.

**Keywords** — Android Application, Arduino Uno, Home Automation, Voice control, Wi-fi module.

## I. INTRODUCTION

With advancement of automation technology, life is getting simpler and easier in all aspects. The concept of remote control of house devices over the internet from anywhere, any time in the world today can be a reality, which has changed the way people live.

A smart home is equipped with lighting, heating, and electronic devices that can be controlled remotely by smartphone or via the internet [1]. An internet based home automation system focuses on controlling home electronic devices from inside or outside the home.

Wireless communication based home automation system has high momentum for the last couple of years. Wireless communication decreases the complexity associated with the installation and maintenance compared with wire communication. Bluetooth and Wi-Fi, are the popular choice for the home automation systems [2]. Wireless network based smart home systems have become very popular as they provide comfort, security, and safety.

The goal of this project is to operate home devices smartly through an Android application using IoT(Internet Of Things) [3]. IoT is the network of "things" or physical objects which include electronics, software, sensors, actuators and network connectivity. All these things collect and transfer data between themselves. IoT has increased

significantly in the last few years since it has added a new dimension to the world of information and communication technologies.

Voice controlled home automation systems have drawn considerable attention in the recent years [4]. Also, the purpose of this project is to develop an Android application which controls home appliances via human voice. Google Voice search software will be used as voice capture device. This is a project to produce a system that can control any electrical appliances by using user's voice. User's voice is an input to activate the electrical appliances through this system. Speech recognition software converts spoken words to machine readable input.

This system needs graphical user interface which provided by Android software. The Android application for recognize voice will send data to the Arduino Uno microcontroller [5] to activate or deactivate the electrical appliances that user wants to use.

The system is portable and constructed in a way that is easy to install, configure, run, and maintain. Android mobile phone as a remote control will make the system more user-friendly and portable.

## II. RELATED WORK

Shirsath N. S, Dhole P. S, Mohire N. P, Naik S. C & Ratnaparkhi N.S [6] proposes a home automation system that employs the integration of multi-touch mobile devices, cloud networking, wireless communication, and power-line communication to provide the user with remote control of various lights and appliances within their home. This system uses a consolidation of a mobile phone application, handheld wireless remote, and PC based program to provide a means of the user interface to the consumer.

Javale, D., Mohsin, M., Nandanwar, S. and Shingate, M. [7] The objective of this paper is to assist handicapped/old aged people. It gives a basic idea of how to control various home appliances and provide a security using Android phone. The design consists of Android phone with home automation application, Arduino Mega ADK. User can interact with the android phone and send control signal to the Arduino ADK which in turn will control other embedded devices/sensors.

Basil Hamed [8] propose to design and implement a control and monitor system for the smart house. Smart house system consists of many systems that controlled by LabVIEW software as the main controlling system.

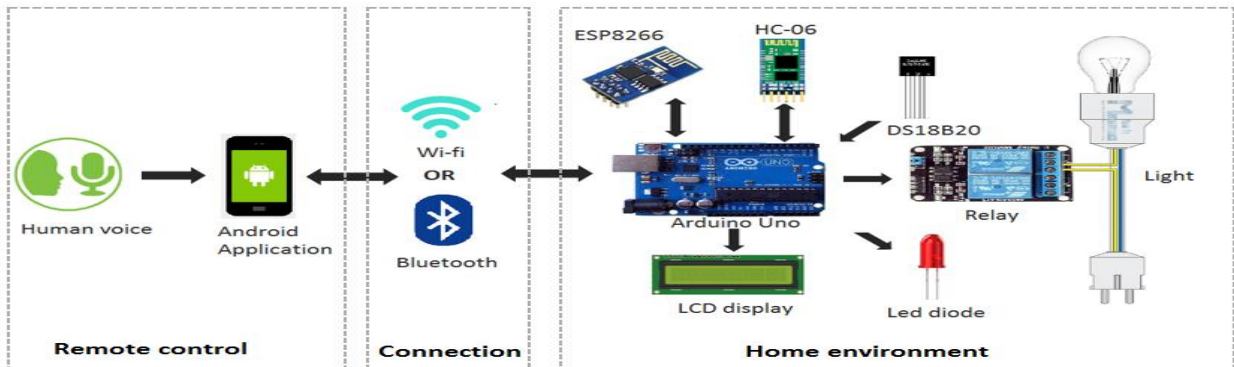


Fig. 1. System architecture of home automation system

## I. PURPOSE OF SYSTEM

The purpose of this paper is to design and implement a home automation system using IoT that is capable of controlling and automating most of the house appliances through voice commands. The proposed system has a great flexibility by using Wi-fi technology. This will decrease the deployment cost and will increase the ability of upgrading. Also, the user can control the house appliances via Bluetooth.

## II. SYSTEM ARCHITECTURE

The system architecture of the project is shown above in Fig. 1. The system mainly comprises two parts: the remote control and home environment.

Android application is used to recognize voice which is installed to Android phone. Spoken words convert into control commands and send them to the Arduino microcontroller via Wi-fi or via Bluetooth. The microcontroller is connected to the appliances through relays. The home environment is controlled depending on the received command.

### A. Hardware design

The hardware consists of the following components:

- Arduino Uno,
- ESP8266 Wi-fi module,
- HC-06 Bluetooth module,
- LCD display,
- DS18B20,
- relay.

The control algorithm runs on Arduino Uno. The Arduino Uno is a microcontroller board based on the ATmega32. It has 14 digital input/output, 6 analogs, ceramic resonator, USB connector, power connector, ICSP, and reset button. Arduino Uno is connected to the ESP8266 module and HC-06 module by I/O points.

The ESP8266 is now one of the leading platforms for the Internet of Things. The ESP8266 Wi-fi module is a self-contained SOC with integrated TCP/IP protocol stack that can give any microcontroller access to your Wi-fi network. Its parameters can be set by AT command to connect with Wi-fi networks.

The HC-06 module is suitable where wireless data transmission is needed in slave mode. The HC-06 module can reach a range of up to 9 meters.

Relay is used as a switch. A relay is an electromagnetic switch operated by a relatively small electric current that can turn on or off a much larger electric current. Relays are used where it is necessary to control a circuit by a separate low-power signal, or where several circuits must be controlled by one signal. The microcontroller controls the appliances through relays.

The DS18B20 Digital Thermometer provides 9 to 12-bit readings which indicate the temperature. DS18B20 is used to get the temperature from home. The LCD screen is an electronic display module. A 16x2 LCD display is a very basic module and is very commonly used in various devices and circuits. The temperature and IP address for Wi-Fi connection is displayed on the LCD.

### B. Software design

Android operating system is developed for smartphones and tablets. Android is an open-source operating system which means that anyone can use it in their phones free. Android is the most widely used mobile operating system by the people nowadays. Android application is designed by a user-friendly interface. To control the home environment user has to install the Android application.

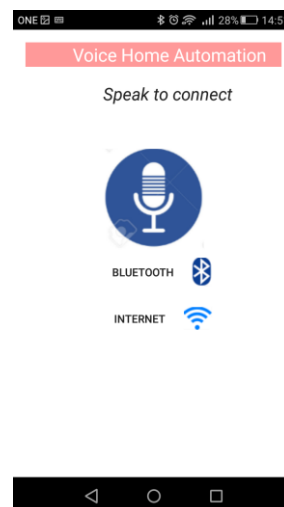


Fig. 2. First Android user interface

Fig.2 illustrate the first Android user interface for communication with the microcontroller. Users can choice Wi-Fi or Bluetooth communication.

For W-fi communication, the user must enter the IP address and port given on LCD display. After establish connection users can choice to control home system by voice commands or with a click on the button to turn on/off the appliance. On Fig. 3 is shown user control interface. The application is designed in Android version 4.2 with API level 17.

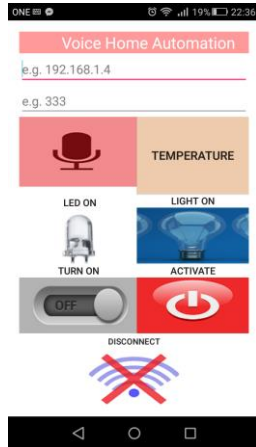


Fig. 3. Control interface

C. Activity diagram

In Fig.4 is shown activity diagram for Android application to understand the overall working procedure and properly which will be helpful to operate the project.

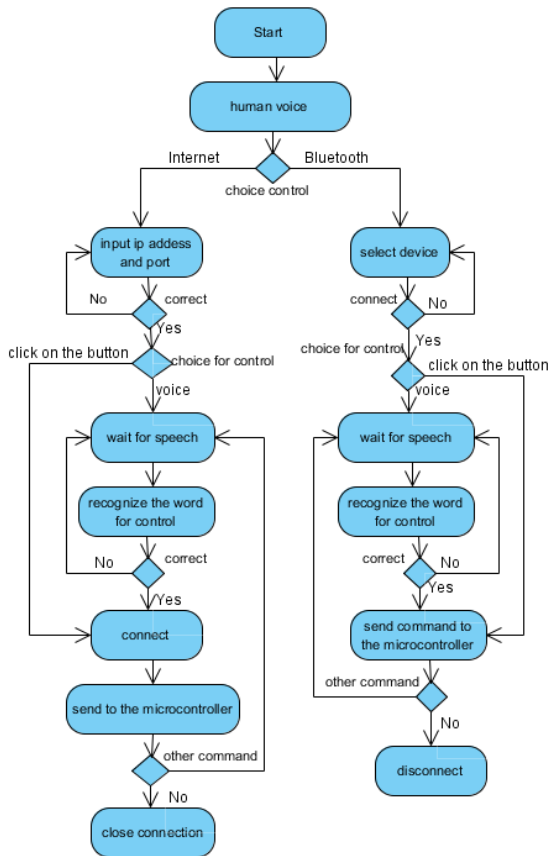


Fig. 4. Activity diagram for Android application

The user can easily manage system locally or remotely home automation system, through voice commands and with a click on the buttons. The user will just have to spoke the command if is chosen voice control. The received voice input, the application will convert this input into byte format and transmit this data to the Arduino microcontroller via Wi-fi module or via Bluetooth.

Activity diagram for Arduino Uno is shown in Fig. 5. after initialization the microcontroller wait for a command. If the command is correct the user command is executed. The activity diagrams show the work-flow of the system.

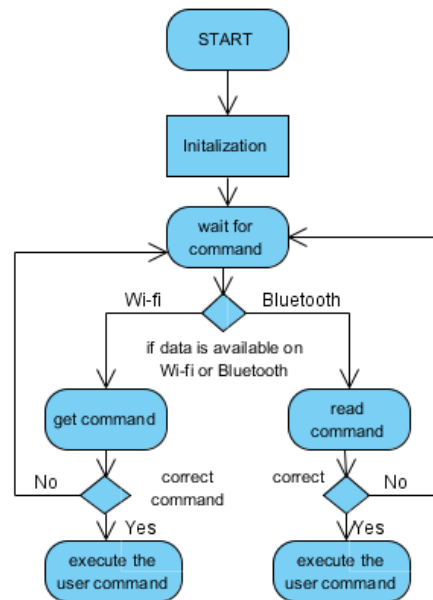


Fig. 5. Activity diagram for the Arduino Uno

D. Test and Results

The proposed system successfully is tested in the home. Android application is tested on a smartphone with operating system version 5.0 (Lollipop). The system is tested with different appliances and Wi-Fi or Bluetooth communication. The Android application gives the information about the temperature of the house. Also gives the status of the controller electrical appliances which we can control remotely. The completed home automation system layout is shown in Fig. 6.

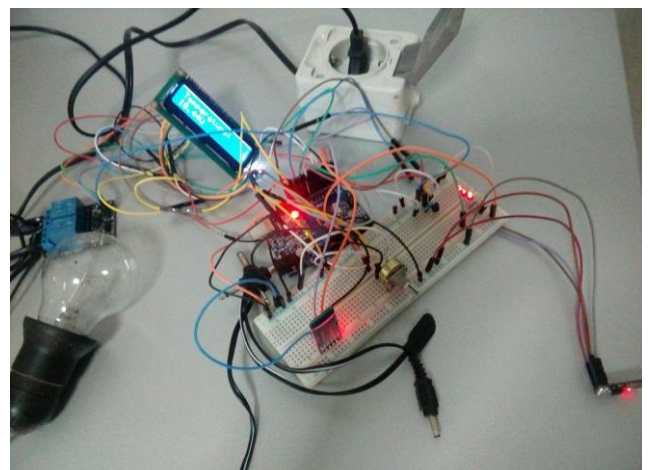


Fig. 6. Home automation system layout

The accuracy of the recognition can be affected by background noise, speed of the speaker, and the clarity of the spoken accent. To get the correct result system was tested in an environment without background noise by five different users. They received instructions for working procedure of the system, and each user perform each command 20 times to determine the percentage of successfully recognized and executed commands. The percentage of success recognized commands is higher in light on, activate and deactivate than the led on and led off. The results (in %) of successfully recognized and executed commands are presented in Table 1.

TABLE 1: TEST OF RECOGNIZED AND EXECUTED COMMANDS

Command	Recognized %	Executed %
Temperature	99	100
Led on	89	100
Led off	85	100
Light on	100	100
Light off	91	100
Turn on	96	100
Turn off	94	100
Activate	100	100
Deactivate	100	100
Disconnect	98	100

From the performed tests for recognition and execution of commands for control is notice different time for execution. As soon as the system receives the voice command from the Android application first checks if the command is correct and changes the status of appliances. To change the status some devices takes a little bit more time to respond to different voice commands respectively. Response time for appliances is almost the same using Wi-fi and Bluetooth connection. Results of the average response time of appliances for both connection together are shown in Fig.7.

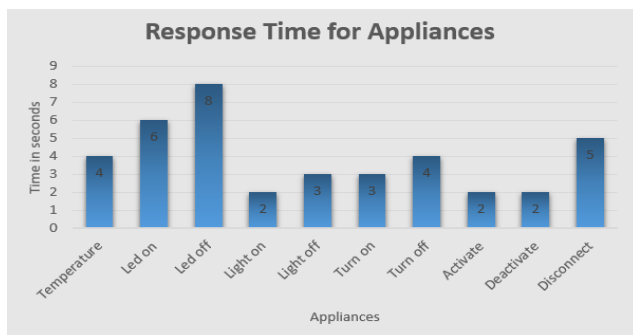


Fig. 7. Response time for Appliances

### III. DISCUSSION

The performance test of the project shows the rate of success of the project. This system is controlled in real time and is easy to use. The important feature of the system is the user interface. This determines how the user will interact with the system. The interface is simple to use, the user can simply with voice command or with touch on the icon turn on/off the appliances.

To control home appliances via Bluetooth the user must be in 9 meters from the microcontroller. So the control cannot be achieved from outside this range. With Wi-fi the home appliances can control from anywhere. This project is very helpful for the old people who will not capable of doing works properly.

### IV. CONCLUSION AND FUTURE WORK

The home automation using Internet of Things has been experimentally proven to work excellent by connecting simple appliances and the appliances were successfully controlled remotely through internet. This technology makes life easier for the user and saves energy by utilizing devices. The user will have complete control over the appliances of their homes.

This project is based on Android and Arduino platform. We used speech recognition system to implement this work. Recognized commands are used to control the designed home automation system.

The approach discussed in the paper is novel and has achieved the target to control home appliances remotely using the Wi-fi technology. They indicate the real-time switches status. The system is designed in a user-friendly interface. Any Android based mobile can be used to control the home appliances. This low-cost home automation system was successfully designed, implemented and tested.

Using this system as a framework, the system can be expanded to include various other options which could include home security feature like capturing the photo of a person moving around the house and storing it in the cloud.

### REFERENCES

- [1] Palaniappan, Satish, Naveen Hariharan, Naren T. Kesh, and S. Vidhyalakshimi. "Home Automation Systems-A Study." *International Journal of Computer Applications* 116, no. 11, 2015.
- [2] Pandya, Bhavik, Mihir Mehta, and Nilesh Jain. "Android Based Home Automation System Using Bluetooth & Voice Command." *International Research Journal of Engineering and Technology (IRJET)* Volume: 03, 2016.
- [3] Vinay Sagar, K. N., and S. M. Kusuma. "Home Automation Using Internet of Things." *International Research Journal of Engineering and Technology (IRJET)* Volume: 02, 2015.
- [4] Akhter, Sonia, Md Faisal Arif, Md Nur-Amin, Nafiz Mustafiz, Roman Khan, Shah Waliullah, and Khalid Hossain. "Voice Controlled Home Appliances: The use of Android Phone." *Journal of Modern Science and Technology* 4, no. 1, 2016.
- [5] David, Nathan, Abafor Chima, Aronu Ugochukwu, and Edoga Obinna. "Design of a Home Automation System Using Arduino." *International Journal of Scientific And Engineering Research* 6, no. 6 (2015).
- [6] Javale, Deepali, Mohd Mohsin, Shreerang Nandanwar, and Mayur Shingate. "Home automation and security system using android adk." *International journal of electronics communication and computer technology (IJECCCT)* 3, no. 2, 2013, 382-385.
- [7] Srisath, N. S., P. S. Dhole, N. P. Mohire, S. C. Naik, and N. S. Ratnaparkhi. "Home Automation using Cloud Network and Mobile Devices." *ITSI Transaction On Electrical and Electronics Engineering* 1, 2013, 785-791.
- [8] Hamed, Basil. "Design & implementation of smart house control using LabVIEW." *International Journal of Soft Computing and Engineering (IJSCE)* 1, no. 6 (2012): 2231-2307.

# Maloletnička računarska pornografija u Republici Makedoniji

Darko Naumovski, Saso Gelev, Vlatko Cingoski and Biljana Petrevska

**Abstract** — U Republici Makedoniji u prošlom periodu sve više i više se pojavljuju slučajevi pornografije sa maloletnicima. Podizanje kulturnog, socijalnog i obrazovnog nivoa građana, stvorice uslove da se otvoreno priča o temama o kojima je donedavno bilo nemoguće pričati, da se prepoznaju ove kriminogene pojave i pravovremeno da se traži i obezbedi zaštita nadležnih institucija. Ovakvi slučajevi kriminala usmereni protiv polove slobode i morala, opravdano izazivaju reakcije u javnosti, najviše radi toga što najčešće kao objekat napada su maloletna deca ispod 14 godina. Ova dečja pornografija obuhvata širok dijapazon posledica, ne samo zdravstvenih, već i socijalnih, koja se neposredno negativno odražavaju na kompletno stanje maloletnika. U ovom radu prezentovaćemo trenutno stanje u Republici Makedoniji. U tom cilju napravljena je anketa povezana sa računarskom pornografijom koja je sprovedena u Ministarstvu unutrašnjih poslova i sudstvu u Makedoniji.

**Cljučne reči** — dečja pornografija, dete, internet, maloletnik, međunarodni propisi.

## I. UVOD

Najmlađa populacijaje je najranjivija kategorija i pod velikim uticajem masovne elektronske komunikacije putem interneta, i samim tim predstavlja potencijalna žrtva računarske maloletničke pornografije i mogućnost da deca budu umešana i seksualno eksploatirana.

Današnja mlada populacija sve više komunicira sa ostatkom sveta računarom, pomoću socijalnih mreža na internetu. Računarska dečja pornografija je kriminalno ponašanje koje se ne prepoznaje, vrlo retko se prijavljuje, mnogo teško ga je dokazati i naročito je teško sprečiti ga. Postoji velika "tamna brojka" računarske dečje pornografije. Računarska dečja pornografija ulazi u kategoriju seksualne zloupotrebe deteta putem interneta. Kao pojavni oblik seksualne zloupotrebe, računarska dečja pornografija može dovesti do seksualnog zlostavljanja, tj podvođenje na maloletničku prostituciju, seks turizam, trgovina ljudima i seksualno iskorištavanje maloletnika.

Darko Naumovski, Ministry of Defence of the Republic of Macedonia (tel:389-70-354947;e-mail:darko\_naumovski\_buci@yahoo.com).

Saso Gelev, Faculty of Electrical Engineering, University „Goce Delcev“, str. „22 October“ bb, Radovish 2420, Macedonia (tel: 389-32-550033; e-mail: saso.gelev@ugd.edu.mk).

Vlatko Cingoski, Faculty of Electrical Engineering, University „Goce Delcev“, str. „22 October“ bb, Radovish 2420, Macedonia (tel: 389-32-550-033; e-mail: vlatko.cingoski@ugd.edu.mk).

Biljana Petrevska, Faculty of Tourism and Business Logistics, University „Goce Delcev“, str. „Krstev Misirkov“ No.10-A , Stip 2000, Macedonia (tel: 389-32-550365; e-mail: biljana.petrevska@ugd.edu.mk).

Dečja pornografija je veliki međunarodni problem. Internet komunikacija nameće nove norme, pravila ponašanja, nove mogućnosti, ali i velike opasnosti od zloupotrebe. Globalne računarske mreže stvorile su odlične uslove za razvoj novih oblika kriminaliteta. Deca su konstantno izložena nasilnim sadržajima, nepristojnim jezikom, rasističkim, političkim i seksualnim sadržajem koji su neodgovarajući za njihov uzrast. Postoje i slučajevi kada maloletnici objavljuju seksualni sadržaj u kojima su i oni uključeni ili njihovi prijatelji, a dosta često se objavljuju i osobne podatke.

Nejveći deo pornografskog sadržaja prikazuje nasilni polov snošaj odraslih nad decom ili iznuđeni seksualni kontakt među decom. Naročito karakteristične su t.z. „montirane fotografije“. To su iskonstruisane fotografije urađene tehnološki sofisticiranom tehnologijom, upotrebljavajući digitalnu rekonstrukciju. Uvećane usluge koje su dostupne na internetu kao i ogromni napredak u digitalnoj obradi slika i sve snažniji grafički programi dostupni na internetu, doveli su do velikih promena u obimu, prirodi i dostupnosti dečje pornografije. Naročito velika opasnost u distribuciji dečjih pornografskih sadržaja je virtuelna fotografija. Upotrebom digitalnog grafičkog softvera mogu se kombinovati dve ili više fotografija u jednu fotografiju ili se urade takve promene kako bi se dobila potpuno nova fotografija. Ovaj proces, poznat kao „metamorfoza“, daje mogućnosti nepornografske fotografije da budu pretvorene i zloupotrebljene kao pornografski sadržaj.

## II. DEČJA RAČUNARSKA PORNOGRAFIJA U MEĐUNARODNIM DOKUMENTIMA

Regulisanje problematike unapređenja prava deteta, u globalnim okvirima, najveći progres je ostvario 1989 godine, kada je Generalna skupština organizacije ujedinjenih nacija usvojila Konvenciju o pravima dece. Ova Konvencija, prihvaćena je i ratifikovana od strane Skupštine Republike Makedonije 1993 godine i predstavlja nadgradnju Deklaracije o pravima deteta iz 1959 godine. Konvenciju o pravima deteta u rekordnom roku je prihvaćena od potrebnog broja zemalja članica Ujedinjenih nacija i stupila na snagu godinu dana nakon njezinog usvajanja. Ovom konvencijom su utvrđena fundamentalna prava svake individue na uzrastu ispod 18 godina. Ova prava su sastavni deo međunarodnog prava. Dečja pornografija je uređena članom 34 ove Konvencije.

Nakon ove Konvencije idu brojna međunarodna dokumenta. Značajno je pomenuti Deklaraciju i Agendu o akcijama usvojenu na Svetskom kongresu protiv komercijalne seksualne eksploatacije dece, održan u

Stokholmu u periodu od 27 do 31 augusta 1996 godine. Iz ovih razloga biće organizovane i mnogobrojne konferencije na kojima će biti donešene zaključci koji će predstavljati osnovu za dalju legislaciju i preciznu inkriminaciju dečje pornografije na Internetu i drugim razvijenim tehnologijama.

Jedna od tih je Međunarodna konferencija o borbi protiv dečje pornografije na Internetu, održana u Beču 1999 godine na kojoj je donešen zaključak koji poziva na široku inkriminaciju proizvodnje, distribucije, izvoza, prenosa, uvoza, namernog posedovanja i oglašavanja dečje pornografije i u kojoj je naglašeno značenje bliže saradnje i partnerstva između vlada i internet industrije.

U kontekstu Konvencije o pravima deteta biće dodan Fakultativni protokol ka Konvenciji o pravima deteta koji se odnosi na prodaju dece, dečje prostitucije i dečje pornografije usvojen 25.05.2000 godine u Njujorku. Ovaj protokol je potpisan od strane R. Makedonije 17.07.2001 godine, a ratifikovan 17.10.2013 godine kada je i stupio na snagu.

Kao relevantni međunarodni dokument iz ove oblasti napomenućemo i Konvenciju o kibernetском kriminalitetu (Convention on Cybercrime, ETS nbr. 185) koju je R. Makedonija potpisala 15.09.2004 godine, a stupila na snagu 01.01.2005 godine.

Drugi relevantni međunarodni instrumenti i programi na ovom polju koji su od izuzetnog značaja su:

- Akcioni plan donešen na Trećem samitu šefova država i vlada Saveta Evrope ( Varšava, 16-17 Maj 2005 godine) koji poziva na razradu mera za sprečavanje seksualne eksploatacije dece;
- Preporuka broj P(91)11 Komiteta Ministra koja se odnosi na seksualnoj eksploataciji, pornografiji i prostituciji i trgovija decom i mlađim punoletnicima;
- Preporuka Rec ( 2001)16 za zaštitu deteta od seksualne eksploatacije, i Konvenciju Saveta Evrope za borbu protiv trgovijom ljudima (ETC BR.197);
- Europska Konvencija o Ostvarivanju prava deteta (1996, ETC br.160);
- Protokol za prevenciju, borbu i kažnjavanje radi trgovanjem ljudima, naročito žene, i dece koji nadopunjuje Konvenciju Ujedinjenih nacija protiv transnacionalnog organizovanog kriminala, kao i Konvencije Međunarodne organizacije rada koja se odnosi na zabranu i hitno reagovanje radi eliminacije najlošijeg oblika dečjeg rada;
- Okvirna odluka Saveta Evropske Unije o borbi protiv seksualne eksploatacije dece i dečje pornografije (2004/68/JHA), Okvirna odluka Saveta Evropske Unije o učešću žrtava i krivičnim postupcima (2001/220/JHA) i Ramkovna odluka Saveta Evropske Unije za borbu protiv trgovijom ljudima (2002/629/JHA);
- Stokholmska deklaracija i agenda o postupcima donošena na Prvom svetskom kongresu protiv komercijalne seksualne eksploatacije dece( 27-31 August 1996), Globalna obaveza od Jokohame

usvojena na Drugom Svetskom kongresu protiv komercijalne seksualne eksploatacije dece (17-20 Decembar 2001);

- Obaveze i Plan za akciju iz Budimpešte usvojeni na Pripremnoj konferenciji za Drugi Svetski kongres protiv komercijalne seksualne eksploatacije dece ( 20-21 Novembar 2001);
- Rezolucija EC-27/2 Generalne skupštine Ujedinjenih nacija „Svet po meri deteta“ i trigodišnja programa „Izgradnja Evrope za i sa decom“ usvojeni nakon Trećeg samita i utvrđeni na Konferenciji u Monaku ( 4-5 April 2006 );

### III. ZAKONSKA LEGISLATIVA O MALOLETNIČKOJ PORNOGRAFIJI U REPUBLICI MAKEDONIJI

U Republici Makedoniji maloletnička pornografija je regulisana Krivičnim zakonom RM i to sledećim poglavljima.

#### A) Prikazivanje pornografskih materijala maloletnicima regulisano je članom 136, 137, 139

(1) Svako koji će maloletniku koji nije napunio 14 godina prodati, prikazati javnim izlaganjem ili će na neki drugi način učiniti dostupnim fotografije, audiovizuelne i druge predmete sa pornografskim sadržajem, kazniće se zatvorom od šest meseci do tri godine.

(2) Ako je ovo delo urađeno putem sredstava javnog informisanja, prestupnik će biti kažnjen zatvorom od tri do pet godina.

(3) Kaznom stava (2) će biti kažnjen i onaj koji će zlostaviti maloletno lice za izradu audiovizuelnih slika ili drugih predmeta sa pornografskim sadržajem ili za pornografsko predstavljanje.

(4) Predmeti iz stava (1), (2) i (3) će biti oduzeti.

#### B) Postupci nakon krivičnih dela iz oblasti seksualnog zlostavljanja maloletnih lica ili dece

Krivični zakon Republike Makedonije (Sl. cesnik RM br. 37/96, 80/99, 4/02, 43/03, 19/04, 81/05, 60/06, 73/06, 7/08, 139/08, 114/09 i 51/11) u Glavi XIX („Krivična dela protiv polove slobode i polovog morala“) predviđa sledeća krivična dela iz oblasti seksualnog zlostavljanja dece:

##### Član 122

Pod dečjom pornografijom podrazumevamo svaki pornografski materijal koji vizuelno prikazuje polove akcije sa maloletnicima ili zrelije lice koje izgleda kao maloletnik, ili realne fotografije koje prikazuju očigledne polove akcije sa maloletnicima ili prikazuju maloletnike u očiglednim seksualnim položajima.

#### B) Proizvodnja i distribucija dečje pornografije - Član 193-a

(1) Svako koji proizvodi dečju pornografiju radi distribucije ili je prenosi ili daje ili na neki drugi način je pravi dostupnu, kazniće se zatvorom od najmanje pet godina.



(2) Svako koji nabavlja dečju pornografiju za sebe ili za drugog ili poseduje dečju pornografiju, kazniće se zatvorom od pet do osam godina.

(3) Ako je delo iz stava (1) i (2) ovog člana urađeno putem računarskih sistema ili sredstvima masovne komunikacije, prestupnik će biti kažnjen zatvorom od najmanje osam godina.

(4) Ako je ovo delo uradilo pravno lice, kazniće se paričnom kaznom.

#### IV. ISTRAŽIVANJE SPROVEDENO U REPUBLICI MAKEDONIJI

Sprovedena je anketa u sudovima, javnom tužilaštvu i policiji u Republici Makedoniji. Ispitanici su trebali da odgovore na sledeća pitanja:

1. Koji je profil žrtava dečje pornografije?
2. Dali je jedno isto dete dva ili više puta bilo žrtva i dali se isti prestupnik ponovo javlja u istoj ulozi?
3. Dali su preuzete mere za zaštitu deteta kao svedoka dečje pornografije?
4. Koje kategorije deteta su naročito izložene riziku?
5. Sa kojim službama-organizacijama ste sarađivali?
6. Jeste li učestvovali u programu za prevenciju?
7. Imate li nekakvih preporuka i predloga za poboljšanje stanja i prevazilaženja problema dečje pornografije?

Anketa je realizovana u periodu maj-decembar 2015 godine. Na anketi su odgovorili više od 100 ispitanika koji su direktno uključeni u ovoj problematici.

Tabela o broju ispitanika

Policija						Sudstvo					
Muškarci			Žene			Muškarci			Žene		
20 do 32 godina	30 do 45 godina	Iznad 45 godina	20 do 32 godina	30 do 45 godina	Iznad 45 godina	20 do 32 godina	30 do 45 godina	Iznad 45 godina	20 do 32 godina	30 do 45 godina	Iznad 45 godina
civili	Uniformisana lica	civili	Uniformisana lica	civili	Uniformisana lica	administracija	Sudije /zamenici	administracija	Sudije /zamenici	administracija	Sudije /zamenici
3	7	3	9	5	12	3	3	2	3	1	2
10	12	17	6	5	3	2	0	2	2	10	7
39			14			25			31		
53						56					
UKUPNO anketirano 109 ispitanika											

Ako pogledate tabelu videćete da su ispitanici oba pola i različitog doba. Pokušali smo da obuhvatimo sve moguće kategorije kojih ima u policiji i sudovima.

Bilo je raznoraznih opisnih odgovora, a bilo je i pitanja na koja ispitanici nisu odgovarali.

Na osnovu analize dobijenih odgovora na postavljena pitanja u anketi u sudovima i policiji došli smo do sledećih konstatacija:

- Ako se analizira profil dece kao žrtve dečje pornografije najčešće žrtve su ženska deca, ali nije zanemarliv ni broj muške dece. Na osnovu odgovora svih anketiranih organa kao žrtve najčešće se javljaju deca na uzrastu od 14 do 16 godina, a manji broj i na manjoj uzrastu. Najčešće žrtve su makedonske ili romske nacionalnosti sa osnovnim obrazovanjem i potiču od siromašnih, disfunkcionalnih, konfliktnih i socijalno ugroženih porodica, a manji deo i od porodica gde postoji porodično nasilje;
- Kategorija dece koji su najizloženi riziku da postanu žrtve nasilja su deca bez roditeljske nege

ili zanemarena deca, deca sa ulice, deca koji koriste drogu, siromašna deca i deca sa vaspitno-socijalnim problemima koji se nalaze izvan obrazovnog sistema;

- Što se tiče saradnje i koordinacije između institucija, može se zaključiti da dobra saradnja postoji samo između nekoliko institucija (centra socijalnog rada, policije, pravosudni organi), i skoro i da nema saradnje sa obrazovnim i zdravstvenim ustanovama i drugih adekvatnih organa i organizacija, kao i sa lokalnom samoupravom, što je neophodno za uspešno rešavanje ovakvih pojava i osiguranje pomoći i zaštite dece;
- Nadležni organi i institucije nedovoljno učestvuju u programu prevencije koji imaju najznačajniju ulogu u sprečavanju i smanjenju dečje pornografije. Stručna lica koja rade sa decom nisu dovoljno obučena i osposobljena o pravima dece i prepoznavanju svih povreda i zlostavljanja, što je neophodno potrebno za pravilnu i efikasnu prevenciju, zaštitu i rehabilitaciju;

- Deca, žrtve retko sami prijavljuju slučaj. Najčešće to rade druga lica ili se za slučaj sazna na osnovu drugih saznanja. Ovo upućuje na potrebu edukacije dece o njihovim pravima, prepoznavanje zlostavljanja i njihovo ohrabivanje da sami prijave povredu i zlostavljanje nadležnim organima, i gde i kako mogu da prijave ovakve slučajeve;
- Kao osnovnu pretpostavku za prevenciju je edukacija dece putem obrazovnog sistema, ali i edukacija roditelja, službenih osoba i cele javnosti putem obuke i edukativnih tribina, raznih medijskih programa i kampanja. Također, kao preventivna mera su navedene i strožije kazne prestupnicima, ali i praćenje i rad sa njima da se slučaj ne ponovi;
- Ne postoji dovoljan broj predviđenih mehanizama i načina za rehabilitaciju žrtava, a i ona koja postoje nedovoljno se sprovode;
- Svi organi i institucije nemaju dovoljno finansija, kadrovskih i tehničkih uslova za preduzimanje preventivnih, zaštitnih ili rehabilitacijskih mera;
- Neophodno su potrebna češća istraživanja da se utvrde uzroci koji dovode do ove pojave. Također, potrebno je na osnovu dobijenih podataka preduzeti mere za eliminaciju ili smanjenje uzroka koje dovode do ovakva zlostavljanja i iskorišćavanje dece.

#### V. ZAKLJUČAK

Seksualna zloupotreba maloletnika je tema koja je dosta eksploatisana zadnjih godina u svetu. Sve su veća i intenzivna proučavanja i istraživanja u svetu, što je dobar početak u podizanju stepena spoznaje o ovoj ozbiljnoj pojavi. Podatke koje dobijamo istraživanjem rijetko možemo uporediti sa drugim podacima, ali naši napori na ovom planu ne trebaju stati.

Prvo mora da se postigne konsenzus o tome što sve obuhvata seksualna zloupotreba dece kao deo šireg pojma o seksualnoj eksploataciji i maltretiranju, pa zatim precizno pristupiti istraživanjima i analizi.

Mračna brojka ove pojave je stvarno velika prije svega radi njihove prirode i radi involviranosti dece kao ranjive grupe, međutim u trenutku ne možemo se osloniti samo na informacije koje su neproverene i koje se svakodnevno plasiraju medijumima.

U rešavanju ove vrste kriminala Republika Makedonija preuzima više aktivnosti u oblasti ratifikacije međunarodnih dokumenata, na zakonodavnom nivou, preuzima mere na jačanju kapaciteta nadležnih institucija i naročito na dizajniranju sveobuhvatnih obuka sudija i javnih tužitelja. U prošlom periodu u praksi u Republici Makedoniji je primećen značajan porast dečje pornografije.

Važno je istaći da je u zadnje dve godine primećena pojava krivičnih dela „Proizvodnja i distribucija dečje pornografije“ putem računarskih sistema, za što je predviđena kazna zatvora od najmanje 8 godina. Svi ovi predmeti su procesuirani ispred nadležnih sudova.

#### DODATAK

Može se se приметити да у раду није приказана ниједна фотографија, слика којом би била илустрована малолетна рачунарска порнографија. Ако би ово урадили и ми би били једни од оних који шире малолетничку порнографију. На основу свих законских прописа јавно приказивање таквих фотографија је казниво чак и ако је у циљу науке.

#### LITERATURA

- [1] Кривичен законик на Р.М., („Службен весник на Република Македонија“, бр. 37/1996, 80/1999, 4/2002, 43/2003, 19/2004, 81/2005, 60/2006, 73/2006, 7/2008, 139/2008, 114/2009, 51/2011).
- [2] Кривичиот Законик на Република Македонија (Сл. весник на РМ бр. 37/96, 80/99, 4/02, 43/03, 19/04, 81/05, 60/06, 73/06, 7/08, 139/08, 114/09 и 51/11).
- [3] Child Pornography and Sexual Exploitation: European Forum for Child Welfare Position Statement, 3 (Nov. 1993).
- [4] Sanderson C, Zavodjenje djeteta, (Kako zastititi djecu od seksualnog zlostavljanja- smjernice za roditelje i ucitelje), prevod Zagreb 2005.
- [5] Tory J. Caeti, Sex Crimes, Part 1, Child Pornography, LETIN -16-40-102 Low Enforcment Treining Network, LLC 2009 .
- [6] Anketa sa osobama zapošljena u policiji RM I osobe zapošljene u sudovima u RM.

#### ABSTRACT

In the Republic of Macedonia in the past period, more and more cases we have pornography with maloletnicime. Raising the cultural, social and educational level of citizens, will create conditions for an open discussion about topics that until recently it was impossible to talk, to identify these kriminogene appearance and timely to seek and ensure the protection of relevant institutions. Such cases of crime directed against the poles of freedom and morality, justified cause reactions in the public, mostly because they most often as the target of attack are minor children under 14 years. This child pornography covers a wide range of consequences, not only medical, but also social, which is a direct negative impact on the overall situation of minors. This paper will present the current situation in the Republic of Macedonia. To this end, made the poll associated with computer pornography, which was conducted at the Ministry of Interior and the judiciary in Macedonia

#### JUVENILE COMPUTER PORNOGRAPHY IN MACEDONIA

Darko Naumovski, Saso Gelev, Vlatko Cingoski and Biljana Petrevska

# Developing of the Simulation System to Model the Filling Layer of Spherical Granulated Particles

Vasily N. Kruglov, Andrey V. Chiryshv, Artem V. Kruglov

**Abstract** — The paper is devoted to the issue of the simulation system development to model filling layer of granulated material. The tools for the filling layer simulation with particular force to green ore pellets are the utmost important for mining processing plants. The software suits implemented for system realization are described. The main features and options of the application are reviewed. The algorithm for the filling layer model formation is adduced. The results of the test show that the model of the granulated material reflects the real physical object with high accuracy so it can be used for its further investigation.

**Keywords** — simulation system, non-coherent material, grain-size distribution, density, machine vision system, pellets

## I. INTRODUCTION

One of the principal characteristic in production and processing of the non-coherent material fraction is its distribution over size grade. Grain-size distribution assessment of the crushed, granulated or pelletized particles of non-coherent material is significantly important for mining processing plants to control their technological process. In particular, grain-size distribution of the green ore pellets is main characteristic of their quality. At the moment it is paid close attention to the problem of green pellets geometry measuring [1-8]. Analysis of the bibliography shows that the advanced way to measure this criterion belongs to the machine vision systems. This approach allows on-line controlling of the material grain-size distribution in the workflow. Camera captures images of the top leaf of the granulated material. Then the isolation of each visible particle and its linear size and volume measurement are implemented with help of digital image processing methods. The main part of any contactless measurement system based on machine vision is an image processing and analysis unit which includes an algorithm for the isolation and detection of granulated particles.

It is necessary to estimate the tolerance of the computation unit while developing and integration of such

system. In order to evaluate the precise of the algorithm's grain-size distribution the screen analysis of the of the green ore pellets test sample should be implemented [10, 11]. However, such a test of validity of computational data over against experiment is highly laborious and time-consuming. Moreover, the correctness of the algorithm should be tested over the huge amount of the top leaf images with the given distribution law of the non-coherent material particles. Also it is preferably to have an opportunity to form images with arbitrary distribution law. That is why the simulation system for modelling of the filling layer of spherical granulated particles is likely to be developed. Such a system will allow to form images of the green pellets with target values of their grain-size distribution. The result of the modeling will be applied in computation unit checkout and validation, performance and accuracy comparison of the different algorithms devoted to granulated particles measurement.

## II. DEVELOPMENT OF THE SIMULATION SYSTEM

Modern equipment and software allow to create a physical model of the huge amount of particles interaction with high accuracy and realistic behavior. There are a number of companies who offer software and hardware to simulate complicated physical processes. For example, GDI provides ROCKY software for 3D analysis of behavior of the non-coherent material particles [12]; Pasimodo solution by Inpartik company is suitable for granular medium modeling [13]. Despite the wide latitude of these software packs they are not appropriate for development of the algorithms for machine vision systems as they have no tools to simulate displacement, quantity or type of lighting sources in the scene.

The developed system is based on the PhysX middleware by NVidia which used in modern games and graphic editors. Implementation of these middleware gave an opportunity to create effective framework for filling layer simulation which combines high computation performance and accuracy of the geometry and physical compositions. Interface DirectX is implemented to visualize simulation process. It is specified by the high image processing speed due to v video adapter hardware acceleration. DirectX is one of the most popular tool for the graphics application development. Combining PhysX and DirectX allows to compute the interaction of modeling objects with high accuracy and display the results of the

Vasily N. Kruglov is with the Institute of Radioelectronics and Information Technologies, Ural Federal University, 620004, Yekaterinburg, Russia (+79126971200, v.kruglov@mail.ru)

Andrey V. Chiryshv is with the Institute of Radioelectronics and Information Technologies, Ural Federal University, 620004, Yekaterinburg, Russia (+79538272923, chiryshev@mail.ru)

Corresponding Author Artem V. Kruglov, is with the Institute of Radioelectronics and Information Technologies, Ural Federal University, 620004, Yekaterinburg, Russia (+79089094353, avkruglov@yandex.ru).

simulation in a clear view.

#### A. Principal features of the system

Simulation system has a wide range of input parameters which influence the simulating process. System operates with conditional units so user determines the specific system of units for the parameters of the structure elements while simulation. It is necessary to set the specific probability laws for distribution of the position and object size in each instance of the simulation to obtain the most adequate model. For example, the position of the spherical particle in X-axis direction could be set according to the uniform density

$$f(x) = \begin{cases} \frac{1}{b_x - a_x}, & x \in [a_x, b_x] \\ 0, & x \notin [a_x, b_x] \end{cases} \quad (1)$$

where  $[a_x, b_x]$  – variation interval of X coordinate, or the normal density

$$f(x) = \frac{1}{\sigma_x \sqrt{2\pi}} e^{-\frac{(x - \mu_x)^2}{2\sigma_x^2}} \quad (2)$$

where  $\mu_x$  – mean value,  $\sigma_x$  – dispersion of X coordinate distribution. It is similar to Y coordinate.

The size of the generated spheres is set in conditional units of their diameters which can also be generated according to the uniform, normal or even arbitrarily density. In the latter case the density of the diameter or gran-size distribution is made as histogram of categorization. The developed system allows to create up to 9 categories. Within each category the size of the sphere is set by normal density.

It is extremely important zone in machine vision system to arrange for the artificial lighting of the measuring in order to develop effective algorithm for the isolation and detection of granulated particles. The developed software provide for tools to configure the scene lighting. It assumes generating of the lighting sources and definition of their parameters. Lights are also significant in the visualization of the filling layer simulation. They give an opportunity to set the required contrast of the image, configure shadows and regulate luminance differences on the edge between the visible particles of the filling layer. It is possible to generate up to 8 light sources. Each light is regulated by its position relative to the scene. The light can modulate one of three types of the light flux: spot light, point light and directional light.

The result of the simulation is 3D scene of the location of the model mass of the spherical granulated particles in respect to the static container. The images of observed scene are of the 768\*576 pixels and .bmp format.

#### B. Model formation algorithm

The filling layer model construction is carried out as following.

1) Initialization. Initialization of the model means creation of the external environment – scene. It involves the parameters required for simulation, environment settings and all objects which could be added to the model. What is more, PhysX middleware and DirectX interface are also initialized while creation of the scene. It includes setting the initial parameters and operation mode.

2) Object creation. Each object in the system is described by two features: actor and shape. The former includes whole range of the physical characteristics that are necessary to simulate behavior of the object: size, density and interaction forces. The latter is a representation of the object in the scene. The visual representation of the object is created in parallel with the object itself and used through the whole simulation cycle.

3) The location of the object is set by two components: coordinates of the center of mass and the attitude quaternion which describe the body rotation around the given 3D vector by the specific angle. Objects are generated in the system throughout the simulation cycle.

4) Simulation time-step calculation. Simulation time-step is one of the most significant parameter that has an effect on the model creation accuracy. The frequently used time-step is 1/25 second.

5) Simulation. At this stage PhysX calculates objects' collisions within given amount of time and outputs their coordinates by the end of the time-step.

6) Visualization. The shape of the each object is displayed by the DirectX interface. The location of the camera and influence of the lighting sources are calculated while visualization process.

7) Model destruction. Stages 2-5 are repeated until the model will be stopped. After that all objects are deallocated. System is ready to the new simulation cycle.

### III. RESULTS OF THE MODELING

Two auxiliary experiments were conducted in order to validate the system. Images of the monolayer of the particles were generated (Fig. 1) with adjusting coefficient of 5 pix/mm.

Validity of sphere generating on the images was checked manually and by software-analytic approach. An amount of totally visible spheres is 108 in Fig. 1a and 27 in Fig. 1b. The manual measurement of the spheres' sizes was implemented with help of Paint graphic editor over 70 spheres from Fig. 1a and 27 spheres from Fig. 1b. The results confirmed the accuracy of the sphere generating. In order to fulfill the latter validation approach the algorithm for measuring geometry of the granulated or pelletized particles in the filling layer was programmed [9]. Algorithm output is presented in Fig. 2-3.

One object in Fig. 2a isolated incorrectly, however, its volume is quite small and the extent of an error is acceptable. There are two detection incidents (at the edge of the frame) in Fig. 3a. Absolute error in this case is 1.7% due to the small amount of the spheres in the image.

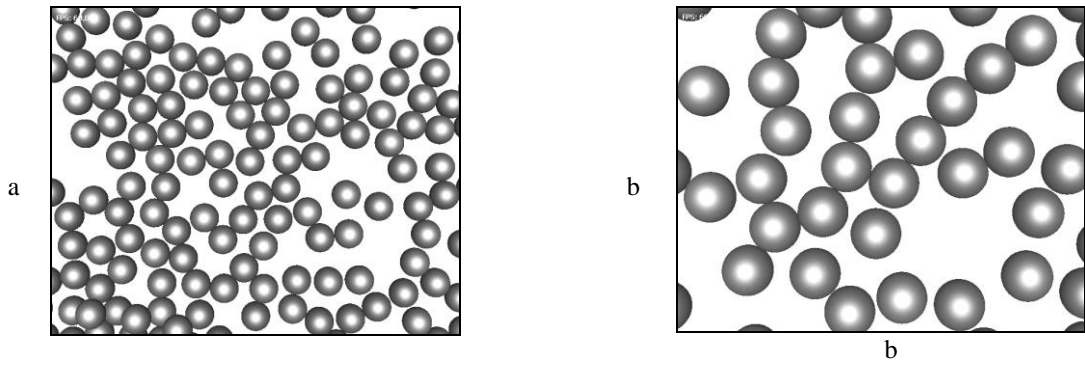


Fig. 1. Images of the monolayer of the spherical particle of the different size: a – 11 mm, b – 19 mm

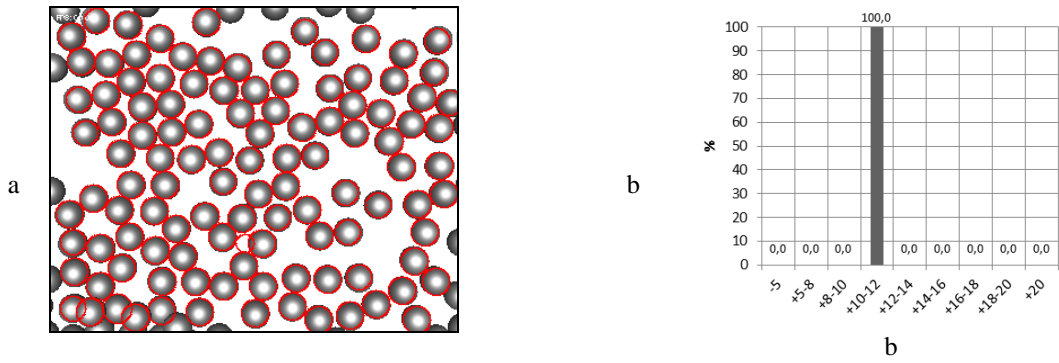


Fig. 2. Sphere isolation algorithm output for spheres of the 11 mm in diameter: a – display of the isolated spheres; b – distribution histogram

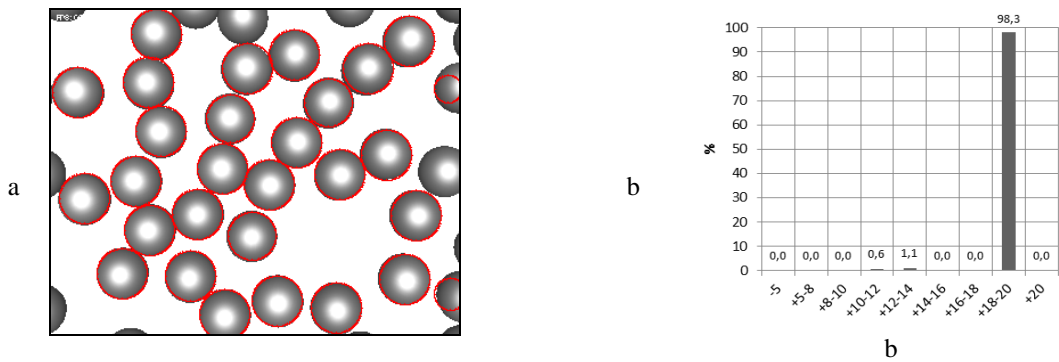


Fig. 3. Sphere isolation algorithm output for spheres of the 19 mm in diameter: a – display of the isolated spheres; b – distribution histogram

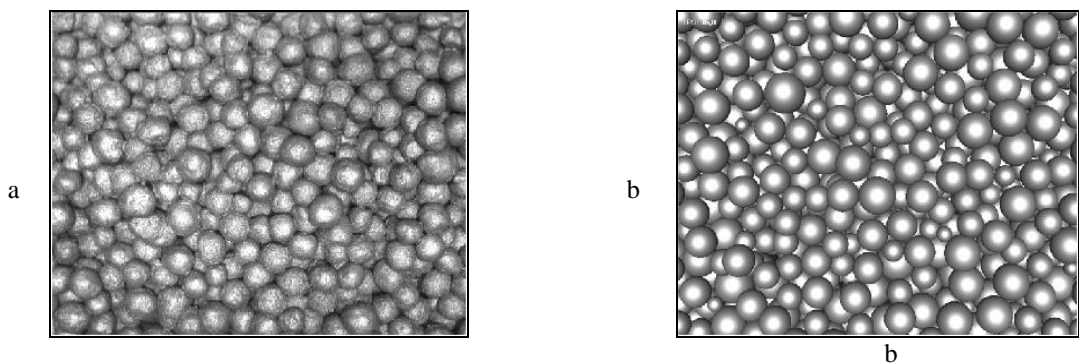


Fig. 4. Image of the sample of green ore pellets (a) and simulated filling layer of the spherical particles with the same grain-size distribution

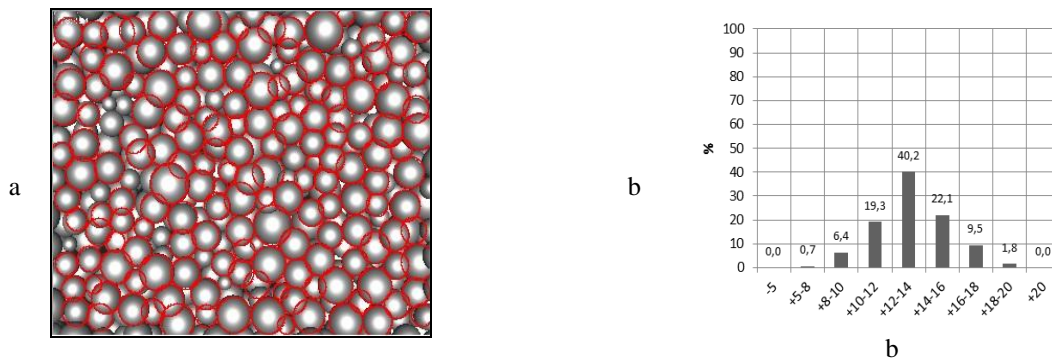


Fig. 5. Output of the algorithm for the isolation and detection of spherical particles with given distribution law: a- display of isolated spheres; b- distribution histogram

In order to validate the generation of the filling layer of the spherical particles by the arbitrarily density the sample of the green ore pellets were taken [10]. Image of this sample is presented in Fig. 4a. The results of the screen analysis [11] of the sample are in the Table 1.

TABLE 1: GRAIN-SIZE DISTRIBUTION IN THE SAMPLE OF GREEN ORE PELLETS

<b>Category, mm</b>	- 5	+5-8	+8-10	+10-12	+12-14
<b>Category mass percentage, %</b>	0	0,2	3,9	24,7	39,9
<b>Category, mm</b>	+14-16	+16-18	+18-20	+20	
<b>Category mass percentage, %</b>	26,8	4,5	0	0	

Obtained data on green ore pellet diameter distribution were input in modeling system as initial conditions. Image of the filling layer simulated by the arbitrary density law introduced in Fig. 4b. This image was processed by algorithm for particles geometry measurement [9]. Output of the algorithm is presented in Fig. 5.

Obtained results show that:

- Mean absolute error of the grain-size distribution of the simulated filling layer of the spherical particles is 2.88% over all categories.
- Margin of absolute error in main category of the simulated particles (+ 12mm-14mm) is 0.3%.

#### IV. CONCLUSION

Developed simulation system allows to create models of the filling layer of spherical particles with arbitrary characteristics of grain-size distribution as well as display the results of the simulation. This tool is applicable to validating algorithms for grain-size distribution measurement. Moreover, the developed software provides definition of the spatial features of the non-coherent material. According to the result of the simulation, the model of the filling layer reflects the real physical processes of the filling layer formation with high accuracy and could be applied for its further investigation. In particular simulation results could be applied for experimental check of such properties of the simulated structure as fractional void volume, closeness of packing, etc. Fractional void volume is a significant indicator for

measurement of the gas permeability of the green ore pellets that is argue problem for mining processing plants. Developing of such system gives the balance of advantage. Firstly, it is an optimization of the gas flow rate required for the roasting. Secondly, it is a reduction in expenses for production tests and laboratory data analysis of incoming pellet flow. Finally, it is an opportunity to forecast the roasted pellets quality.

#### REFERENCES

- [1] Jagdish Lal Raheja, Gaurav Sahu. Pellet Size Distribution Using Circular Hough Transform in Simulink, American Journal of Signal Processing 2012, 2(6): 158-161.
- [2] Dhiraj, J.L. Raheja, Vaishali Singh, Sreekanth Pusapati, Ashutosh Gupta. Design and Implementation of Floating Point based System for Pellet Size Distribution using Circular Hough Transform. International Journal of Computer Applications. Volume 62–No.13, January 2013,p.9-15
- [3] Navdeep Singh, ShaliniSaxena, Dhiraj, J L Raheja. Industrial vision system for distribution and size analysis of pellets. International Journal of Application or Innovation in Engineering & Management. Volume 2, Issue 3, March 2013 Page 7-13
- [4] J. Kaartinen and A. Tolonen. 2008. Utilizing 3D height measurement in particle size analysis. In: Myung Jin Chung, Pradeep Misra, and Hyungbo Shim. Proceedings of the 17th IFAC World Congress (IFAC 2008). Seoul, Korea. 611 July
- [5] T. Andersson, M. J. Thurley, O. Marklund. Pellet Size Estimation Using Spherical Fitting. IEEE Instrumentation and Measurement Technology Conference. Warsaw, Poland, May 1–3, 2007.
- [6] Wang Jie-sheng, Wang Jie-sheng. PCNN Edge Detection of Sintering Pellets Image Based on Hybrid Harmony Search Algorithm, 2009 International Conference on Artificial Intelligence and Computational Intelligence, p. 535 – 539
- [7] Huanyu Liu, Lihong Xu, Dawei Li. Detection and recognition of uneaten fish food pellets in aquaculture using image processing. Conference Paper in Proceedings of SPIE - The International Society for Optical Engineering · October 2014, <http://spiedigitallibrary.org/>
- [8] Sergey Kucheryavski, Kim H. Esbensen, Andrey Bogomolov. Monitoring of pellet coating process with image analysis—a feasibility study. Journal\_of\_Chemometrics, 2010; 24: 472–480
- [9] V.N. Kruglov A method of determining geometry of the pelletized particles and/or granulated material in the filling layer. Patent № 2557330, 2015
- [10] GOST 17495-80 Iron ores, concentrates, agglomerates and pellets. Methods of sampling and preparation of samples for particle size analysis
- [11] GOST 27562-87 Iron ores, concentrates, agglomerates and pellets. Determination of size distribution by sieve analysis
- [12] ROCKY - DEM Particle Simulator. URL: <http://www.rocky-dem.com/>
- [13] Inpartik Simulation Software & Engineering. URL: <http://www.inpartik.de/>

# Novel technique for Authentication & Encryption in Next Generation Networks

Aleksandar Tudzarov<sup>1</sup>, Goce Stefanov<sup>2</sup>, Saso Gelev<sup>3</sup>,  
*Faculty of Electrical Engineering, Goce Delcev University – Stip, R.Macedonia*

**Abstract** — This paper focuses on providing of novel technique for authentication and encryption in Next Generation Mobile Networks. The novel technique uses higher level of authentication methods that offer more flexibility and are based on well developed and proven IT security standards. Implementation of proposed technique will simplify the authentication process and will bring more simplified and manageable authentication environment that will cover fix/mobile convergence and can scale for different level of service needed security. By combining of several protocols that are used in process of network authentication like DIAMETER/EAP methods with PKI as a bases for authentication and by adding new parts in authentication access network architecture like PANA and EAP-SMIL we have proposed new authentication technique that will improve and modernize authentication and encryption process in Next Generation Networks.

**Keywords** — NGN, Authentication, Encryption, PKI, Certificate, DIAMETER, PANA, EAP, IPSEC

## I. INTRODUCTION

One of the main focuses of NGN (Next generation networks) is placed on the convergence of telecom and IT to develop a ubiquitous infrastructure that offers higher capacity to customers and creates new opportunities to interconnect smart objects. According to the requirements in NGN there will be a massive number of devices (e.g. sensors, actuators and cameras) with a wide range of characteristics. Integration of all these heterogeneous devices with different security needs on NGN poses new security challenges towards a secure, reliable and dependable infrastructure. The Authentication and Key-agreement (AKA) procedures for 2G, 3G and 4G generations of mobile network have mostly fulfilled the requirements of each of these generations. 1G established the foundation of mobile networks; 2G increased the voice connectivity capacity to support more users per radio channel; 3G introduced high-speed internet access; 4G provided more data capacity. One of the key expectations for NGN is to be the reference network for the Internet of Things (IoT) connectivity. Although it can be assumed that NGN will utilize a basic authentication such as

4G/LTE protocols, the use cases for NGN brings new requirements that the next generation AKA protocol must support. The collection of connected devices (or "things"), referred in literature as Internet of Things (IoT), will increase substantially and NGN must provide an adequate level of security, which in turn introduce novel security challenges for authentication of these devices. The current SIM cards, while providing strong key protection, are also imposing a hurdle for certain use cases, and is also a non-negligible cost. To maximize the usability of IoT devices, new approach that will enable full authentication and security without the need of hardware SIM (UICC) need to be developed. In massive deployments of devices, the existing AKA protocol is not suitable, as each device have to run the full AKA procedure. The AKA protocol has a central role in the security of mobile networks as it place a fundament for exchange of the parameters needed to form a security context that is agreed by the parties. The protocol provides mutual authentication between device and serving network, and establishes session keys. The authentication protocol used in 4G (EPS-AKA) which is the state of the art authentication protocol in mobile networks is almost identical to its predecessor used in 3G (introduced in the late 90s). A limitation of EPS-AKA is that, for each device that requires network access, the protocol requires massive signaling among the device, the local serving network and the device's remote home network. NGN needs to offer light-weight and flexible authentication protocols.

Furthermore, it can be foreseen that large enterprises already have an existing AAA infrastructure in place. To further minimize the costs with future NGN subscribers the best way is to allow reuse of the pre-existing identities as a basis for NGN access, i.e. a bring your own identity (BYOI) solution. This way of thinking imposes need of close cooperation between different authentication schemes and methods in IT and telecom world as the requirements of both worlds are reaching the convergence point. Today, in the world of fix and wireless telecommunications, techniques and processes for authentication and encryption are developed as a separate stream from main IT security development paths. This separation leads to implementation of different solutions for the same or similar authentication requirements by means of security. Moreover, certain implementations differ from each other so much that they are not compatible whit each other or even not interoperable. In order to achieve more future prove and flexible security in the new information era where telecommunication and information technology are collapsing into single technology domain, we propose implementation of novel technique that will combine most useful algorithms and

<sup>1</sup> Aleksandar Tudzarov, Faculty of Electrical Engineering, Goce Delcev University – Stip, R.Macedonia, E-mail: aleksandar.tudzarov@ugd.edu.mk

<sup>2</sup> Goce Stefanov, Faculty of Electrical Engineering, Goce Delcev University – Stip, R.Macedonia, E-mail: goce.stefanov@ugd.edu.mk

<sup>3</sup> Saso Gelev, Faculty of Electrical Engineering, Goce Delcev University – Stip, R.Macedonia, E-mail: saso.gelev@ugd.edu.mk

protocols to achieve convergent authentication techniques into technology domain by introducing technology agnostic procedures for authentication and encryption.

The basis of authentication in proposed architecture will be X.509 certificate that would be issued to each user and stored on some authentication medium (either eSIM – “embedded SIM” or embedded in some local security store). This authentication assumption will simplify the process for SIM changes (including eSIM which popularity is increasing and claims to be the new standard for all NGN authentication processes) while user changes network by simple secure download of new certificate and installing it in the store – already know procedures and protocols in IP world. These certificates will have to be issued and signed by Operators CA (Certificate Authority). In this way, increased level of security will be achieved.

Second step in the authentication will be using of secure and flexible protocols for exchanging of security messages that can offer authentication in converged fix/mobile environment, and can be flexible enough to offer exchange of security polices as well as basic service policies for proper handling of users payload traffic. This means that with finalizing of authentication procedure access units (Baseband processing modules or fixed network routing entities) can enquire basic routing policies for handling of the user traffic.

Third step of the authentication process (which is not directly related with authentication) should be encryption of the user payload traffic. Having in mind that proposed authentication process is based on using of Certificates; they can also be used for encryption of the traffic in unsecure fix or wireless environment.

Considering the principles set above proposed novel technique assumes using of PANA and EAP-SAML protocols for authentication in correlation with DIAMETER/EAP in the backend communication with authentication server for establishing the authentication ecosystem that will be used in the authentication and later in encryption process of the user communication. PANA refers to (Protocol for Carrying Authentication for Network Access). EAP-SAML refers to (Security Assertion Markup Language) over EAP (Extensible Authentication Protocol). SAML represents XML-based standard for exchanging of authentication and authorization data between security entities; that is between the entity that poses the identity and the one who provides the service. This standard is a product of the OASIS (Organization for the Advancement of Structured Information Standards). The use of this EAP method provides an opportunity for exchange of additional network policies between the authentication server and the authentication client that are needed in phase of establishment of network connectivity.

## II. DETAILED PROCESS FOR AUTHENTICATION AND ENCRPTION

The process of authentication, authorization and exchange of policies when connecting to certain access network begins with the first phase, which is phase of initiation request for connection towards accessing network. In the second stage, a mechanism for authentication and authorization is activated. This process

is accomplished through access network. The first step in achieving this goal is to check the user authentication parameters by the client. This involves verifying the authentication resources that for the proposed architecture based on user certificates involves checking the existence of the user certificate. In next step validity check of the certificate, its private and public keys and the validity of the issuer and its purpose is conducted. Once the validity of user parameters are determined, the flow continues through the process of authentication and authorization between the user and the network. Communication between them is based on PANA protocol where authentication of user preferences is performed through EAP-TLS or EAP-SAML. In this case, function of EAP client is performed by user client module, while function of authenticator is conducted by network module on NGN side where this module will be presented as DIAMETER server as it represents the central registry of user data and their profiles. Procedures for authentication and authorization of access technology in NGN are presented as step 3, 6 and step 9.

After the authentication process ends, network module will initiate the process of exchange of network policies and user parameters as presented in steps 12 and 15.

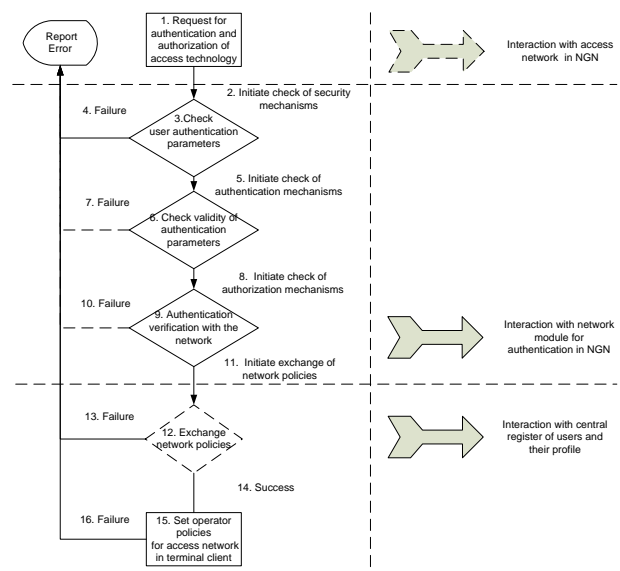


Fig 1. Block diagram for authentication and authorization of access technology of the architecture of PANA protocol

The process of authentication is based on EAP as basis for exchange of authentication in step 9. Authentication with EAP typically begins with a process of negotiation of EAP method. Once EAP authentication method is negotiated and accepted between network entities, a process of exchange of EAP authentication messages between the EAP client and the EAP server is started. When the process of authentication is completed, EAP server sends the EAP message to the EAP client, which confirms or denies the success of the authentication process. Authenticator is informed about the result of the authentication process through the AAA protocol. Based on this information authenticator can provide user access to the requested access network, or to continue to block access. Depending on the chosen EAP method, EAP authentication can be used in the derivation of keys by the EAP Client and the EAP server. Such keys can be



transported through the AAA protocol from the EAP server to the authenticator.

After completing of this procedure and getting derived EAP keys on the side of the terminal (EAP Client) and authenticator, the process of derivation of transport keys to protect access link transmission data can begin. In line with development of new method of authentication, as most suitable candidate for the proposed architecture we can mark EAP-SAML or EAP-TLS protocol, while in the field of transmission techniques for the transmission of EAP the PANA (Protocol for Carrying Authentication for Network Access) can be chosen as a protocol developed for transmission of EAP protocol over "any transport network".

### A. NETWORK AUTHENTICATION OVER ACCESS NETWORKS

PANA is a protocol designed by the IETF as a protocol that should provide transparent network authentication over data link network layer. Its goal is to ensure the establishment of any authentication protocol, over any transport technology (data link level). This goal is achieved by setting the EAP over IP at the transmission level. Along with this basic principle this protocol provides a number of additional and powerful functionalities, such as: separate NAP and ISP authentication, possibility of reuse of local security associations, fast reauthentication, and secure exchange of EAP messages, protocol extensibility by introducing additional protocol messages and so on. These functionalities of PANA protocol makes it suitable for use in procedures for authentication in heterogeneous networks. PANA protocol is set on the last IP link between PANA client (PaC) and PANA authentication agent (PAA). PAA client is set at controller for network access side in this case access network unit (wireless Baseband or fix access xDSL/GPON). This access controller attempts to bridge the AAA sessions between the client and the AAA server using PANA on one side (the client) and DIAMETER on the other side (to the AAA server). PANA platform, despite these network entities also defines other network entity called EP (enforcement point). EP control access in a manner that prohibits access of unauthenticated users to the network resources. EP entity should be placed at key locations in the network architecture (This would be Baseband/DU units or fixed network authentication routers) in order to allow full control of the traffic in both directions.

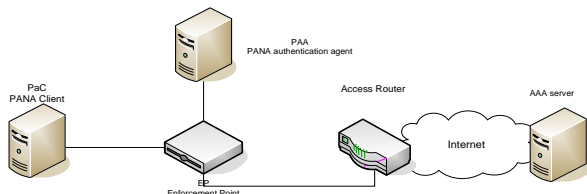


Fig 2. Overview of the architecture of PANA protocol

In most networks, the EP and the authentication agent (PAA) are collocated within a single network entity, but despite this, PANA protocol allows separation of these two functions in the architecture in different entities as shown in Figure 2.

PANA protocol in IP networks is defined as UDP based protocol that operates between two IP-based network entities on the same IP link on UDP port 716. It provides transmission of messages in a specified order as required by EAP specification. Transport message flow is shown in Fig 3. Initiation phase of the protocol consists of exchange of a series of messages with requests and responses. Some of the transmitted messages carry information between the client and the network, while others are used to manage the entire PANA authentication session.

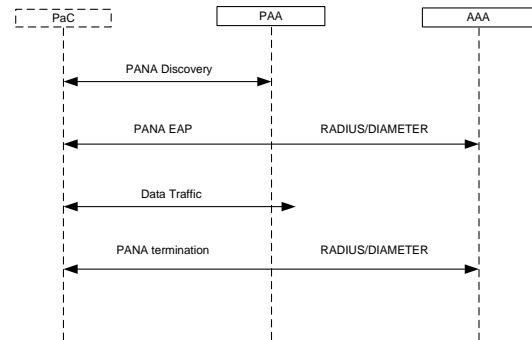


Fig 3. PANA protocol - Message flow

Discovery phase involves two possible scenarios. The first concerns the discovery initiated by PaC client that sends a message to detect PAA on access links, while the second refers to the discovery of PAA. After this process is finished authentication process goes on with exchange of series of PANA authentication messages that represent simple exchange of EAP messages over PANA protocol between the EAP client and EAP server. If authentication is successful, the message instantly imposes a common understanding of the identifier of the device (if the device is identified by its MAC or IP address) and associated level of protection of exchanged packets. Agreed identifier of PaC will be handed to the EP (Enforcement point) for performing of access control in the next phase. Meanwhile the two entities will decide whether to use the basic link-level encryption or IPSec for cryptographic user data protection (if EAP method enables derivation of cryptographic keys). [1], [2], [3]

Received keys are used to generate PANA SA (PANA security associations) that are used to protect the exchange of multiple consecutive PANA messages. In less secure environments it is expected that EAP methods that provide mutual authentication and generate cryptographic keys are chosen in order to protect the communication. [4], [5]

PANA protocol allows IPSec-based access control (which can be bases for encryption in wireless unsecure wireless/fixed transport) in a way that helps the IPSec protocol in the creation of IPSec security associations.

PANA protocol generates cryptographic keys upon completion of the EAP protocol for creating PANA security associations, but they cannot be used directly to create IPSec security associations. This relationship can be used as a basis for "pre-shared secret" for generating dynamic IPSec associations.

This approach leads to the use of IKE protocol for reuse of PANA security associations for IPSec associations. The keys are obtained by derivation of the PANA security association and as such are delivered to IKE protocol. This new IPSec associations are used further to create a

connections between the PaC and PAA, or EP entity for providing of authenticated or encrypted data transport.

Detail exchange message flow for described process above can be seen on Figure 4.

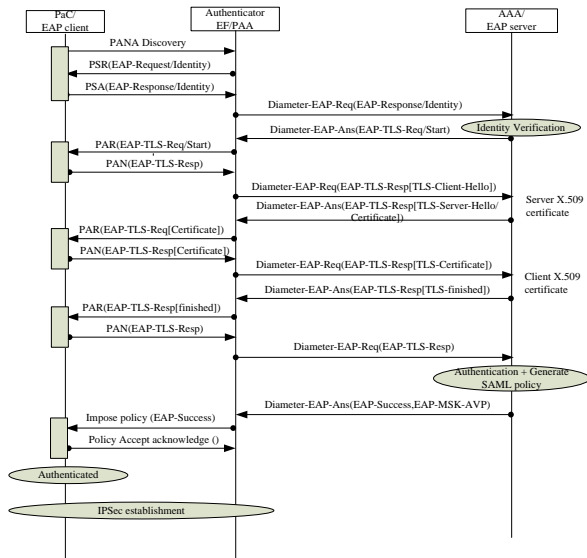


Fig 4. Exchange message flow in EAP-TLS over PANA and DIAMETER

### B. NETWORK AUTHENTICATION ECOSYSTEM FOR NEXT GENERATION NETWORKS

Process of authentication and authorization based on PANA protocol as forefront for formation of authenticated encrypted connections is the basis of the design of the proposed technique for interworking among heterogeneous networks. First step in the process of achieving IP connectivity through access technologies is authentication. For authentication protocol in this case it is best to use EAP and as authentication method on top of it SAML will be used. As a transfer protocol for authentication between client and authenticator in this case PANA protocol can be used where user terminal will be treated as PaC and (Radio Baseband/fixed router – called Authenticator) as PAA. In the background, authenticator should achieve connectivity with certain “user database” server in the cloud, using DIAMETER as authentication protocol. In this way in the authentication process the user terminal acts as EAP client as long as User Database server acts as the EAP server. EAP session is established directly between these two entities. After successful authentication and authorization of user using the PANA protocol as described above and by using of IKE protocol process continues in the stage of formation of the IPsec connections between the two ends of the communication, the user terminal and PAA. Use of EAP-TLS as authentication protocol is due to the concept of security set in the new architecture. The security module by the client and PAA poses public certificates that consist of public and private cryptographic keys. These certificates are issued by Operator. There can be various ways how this can be done, but mainly it is based on setting up a public key infrastructure (PKI).

With the help of PKI each customer is associated with a X.509 certificate that has been issued for its use in this architecture. In the process of user authentication via EAP-TLS/EAP-SAML protocol check of the certificate parameters is performed for the user who in User Database

is connected to the appropriate customer. If its authenticity and validity are determined and also the client that is represented by this certificate is entitled to use the appropriate service it will enable establishment of a link between client and PAA. In this process User Database performs authentication/authorization of access technology in order to determine whether it can be used in the interworking operation. This completes the client's authentication and authorization phase. [6],[7]

### III. CONCLUSION

In the presented paper new novel technique for authentication and transparent encryption process for next generation of networks was presented. As part of the proposed technique, processes for authentication and authorization while connecting to the access technologies were analyzed and innovative techniques are introduced. In this context as the basis of defined authentication process an independent mechanism for authentication and authentication procedures based on certificates was set. Presented process of authentication and authorization are based on PANA protocol as the vanguard of establishing authenticated IPsec tunnel. This technique is fundamental in the design of the secure architecture for interworking in heterogeneous networks. In the next period we plan to conduct in deep analyses of the proposed algorithm (including simulation) to prove the fast convergence time of authentication scheme and to propose standard SAML exchange scheme for the NGN authentication exchange that need to cover all identified NGN security use cases till now and to offer easy extensibility for the future.

### REFERENCES

- [2] Kent, S. and Atkinson, R., “Security Architecture for the Internet Protocol”, RFC 2401, November 1998.
- [3] Oppliger, R., “Security Technologies for the World Wide Web”, Artech House Computer Library, 2000.
- [4] Yuan, R. and Strayer, T., “Virtual Private Networks”, Addison-Wesley, 2001.
- [5] D. Forsberg, Y. Ohba, B. Patil, H. Tschofenig, A. Yegin, “Protocol for Carrying Authentication for Network Access (PANA)”, draft-ietf-panapana- 07 (work in progress), December 2004.
- [6] A Olivereau, A.F. Gomez Skaremta, R.M Lopez, B Weyl, P. Brandao, P Mishra, c. Hauser, “An advanced Authorization Framework for IP-based B3G Systems”, 14<sup>th</sup> IST Mobile & Wireless Communication Summit, Dresden 19-23 June 2005
- [7] Toni Janevski, Aleksandar Tudzarov, Marko Porjazoski, Pero Latkoski, University “Sv. Kiril i Metodij”, Faculty of Electrical Engineering and Information Technologies, “System for analyses of end-to-end quality of data services in cellular networks”
- [7] Rosario Giustolisi, Christian Gehrman, Markus Ahlström, and Simon Holmberg. "A Secure Group-Based AKA Protocol for Machine-Type Communications" 19th Annual International Conference on Information Security and Cryptology. November 30 - December 2, 2016, Seoul, Korea

# Insertovanje vodenog žiga skremblovanog GMSAT algoritmom u DWT-SVD domenu

Zoran S. Veličković, Zoran N. Milivojević, Marko Z. Veličković

**Sadržaj** — U ovom radu je razmatrana zaštita video sadržaja od kopiranja insertovanjem skremblovanog vodenog žiga. Vodeni žig je skremblovan GMSAT algoritmom, a insertovanje je obavljeno u DWT-SVD domenu svakog frejma. Primenom naprednog algoritma za popravku kvaliteta ekstrahovanog vodenog žiga je realizovan odličan SSIM indeks 0.8956, odnosno, NC koeficijent 0.9852 što je u rangu sa najboljim publikovanim rezultatima. Primenjene tehnike obezbeđuju viši kvalitet ekstrahovanog vodenog žiga u poređenju sa tehnikama insertovanja u DCT-SVD domenu.

**Gljučne reči** — DWT, GMSAT algoritam, H.264/AVC kodek, SVD, Vodeni žig.

## I. UVOD

**Z**AŠTITA originalnih multimedijalnih sadržaja od kopiranja i nelegalne distribucije postaje neophodna aktivnost pre njihovog publikovanja na Internetu [1]. Primena standardnih kriptografskih tehnika ne obezbeđuje visok nivo bezbednosti jer se jednom dekriptovan multimedijalni sadržaj može bez odobrenja nelegalno kopirati i distribuirati. Takođe, dodatno vreme potrebno za dekriptovanje sadržaja, kao i česte retransmisije paketa, mogu negativno uticati na nivo korisničkog zadovoljstva QoE (engl. *Quality of Experience*) [2]. U praksi se za zaštitu multimedijalnih sadržaja često koristi koncept zasnovan na insertovanju vodenog žiga u multimedijalni sadržaj [3]. Ovaj koncept zahteva da insertovani vodeni žig ostaje trajno u multimedijalnom sadržaju. Obzirom da se vodeni žig nikada ne odstranjuje iz sadržaja, njegovo insertovanje (umetanje) ne sme da izazove vidljivu degradaciju. Sa druge strane, u procesu ostvarivanja autorskih prava potrebno je ekstrahovati insertovani vodeni žig zadovoljavajućeg kvaliteta [4]. Kompromis između pouzdane ekstrakcije vodenog žiga i izazvanih smetnji u multimedijalnom sadržaju su osnovni zahtevi za dobre algoritme insertovanja. Jačina insertovanja vodenog žiga u multimedijalnom sadržaju se može regulisati faktorom insertovanja. Veći faktor insertovanja obezbeđuje ekstrakciju vodenog žiga višeg kvaliteta, dok je pritom zaštićeni video nižeg kvaliteta (i obratno). Vodeni žig se može smatrati sastavnim delom multimedijalnog sadržaja te na taj način deli sudbinu samog multimedijalnog sadržaja. Pokušaj uklanjanja vodenog žiga iz videa će imati uticaja

kako na kvalitet samog multimedijalnog sadržaja tako i na kvalitet vodenog žiga. U ovom radu je razmatrana zaštita nekodovanog video sadržaja ugradnjom istovetnog vodenog žiga u svaki video frejm. Zbog toga što poznavanje sadržaja vodenog žiga može biti kompromitujuće, vodeni žig se kriptuje pre ugradnje. Za razliku od algoritama gde se koriste hibridne tehnike šifrovanja, u ovom radu je predložena primena generalizovane višestapne Arnoldove transformacije GMSAT (engl. *Generalized MultiStage Arnold Transformation*). GMSAT pripada klasi invertibilnih haotičnih mapa (engl. *Chaotic Maps*) [5]. Primenom GMSAT-a se kriptološki prostor vodenog žiga može proizvoljno uvećati čime se zaštita sadržaja podiže na viši nivo. U [6] je pokazano da ugradnja vodenog žiga niže rezolucije izaziva manje artefakata u zaštićenom video sadržaju. Predložena tehnika ugradnje kriptovanog vodenog žiga niže rezolucije zasnovana je na kombinovanju DWT-a (engl. *Discrete Wavelet Transform*) i SVD-a (engl. *Singular Value Decomposition*). Za dobijanje originalnog vodenog žiga iz skremblovanog (šifrovanog), razvijena je inverzna generalizovana višestapna Arnoldova transformacija [3], [6]. Drugi nivo DWT-a je iskorišćen za usaglašavanje rezolucija za potrebe primene pouzdanog SVD algoritma insertovanja vodenog žiga. Insertovanjem kriptovanog vodenog žiga u sve frejmove, dobija se zaštićeni video sadržaj. Kodovanje zaštićenog videa je standardni proces koji se obavlja pre izlaganja na Internetu. Poznato je da se kodovanjem zanemaruju detalji u frejmovima ali se za uzvrat dobija visok nivo kompresije. Sa stanovišta preživljavanja vodenog žiga, kodovanje se sa može smatrati atakom na sam video sadržaj. U ovom radu je korišćen H.264/AVC koder koji se odlikuje visokim stepenom kompresije i dobrim kvalitetom kodovanog videa. Posledice kodovanja i matematičkih zaokruživanja rezultuju u promenljivoj kvalitetu ekstrahovanih vodenih žigova. U svrhu popravke kvaliteta ekstrahovanog žiga razvijen je napredni algoritam popravke [3].

U drugom poglavlju su prikazane osnove korišćenih matematičkih transformacija. Za kriptovanje vodenog žiga u ovom radu je korišćena generalizovana višestapna Arnoldova transformacija koja se može svrstati u klasu 2D haotičnih mapa GMSAT, dok su za umetanje vodenog žiga korišćene specifičnosti SVD-a i DWT-a. U trećem poglavlju su prezentovani modifikovani pouzdani algoritmi

Zoran S. Veličković, (autor za kontakte), Visoka tehnička škola strukovnih studija Niš, Aleksandra Medvedeva 20, 18000 Niš, Srbija (telefon: 382-67-610100; e-mail: zoran.velickovic@vtsnis.edu.rs).

Zoran N. Milivojević, Visoka tehnička škola strukovnih studija Niš, Aleksandra Medvedeva 20, 18000 Niš, Srbija (e-mail: zoran.milivojevic@vtsnis.edu.rs).

Marko Z. Veličković, Visoka tehnička škola strukovnih studija Niš, Aleksandra Medvedeva 20, 18000 Niš, Srbija (e-mail: marko.velickovic93@yahoo.com).

za ugradnju/ekstrakciju vodenog žiga u/iz zaštićenog videa U četvrtom poglavlju izvršena je evaluacija predloženog algoritma ekstrakcijom vodenog žiga iz kodovanog videa H.264/AVC koderom. Procena kvaliteta ekstrahovanih vodenih žigova je obavljena objektivnim tehnikama zasnovanim na SSIM (engl. *Structural SIMilarity*) indeksu i NC (engl. *Normalized Correlation*) koeficijentu. U petom poglavlju su izvedeni odgovarajući zaključci na bazi sprovedenih ispitivanja

## II. KORIŠĆENE BAZNE TRANSFORMACIONE TEHNIKE

### A. Generalizovana višestapna Arnoldova transformacija GMSAT

U prethodnim radovima autori su predložili višestapnu Arnoldovu transformaciju [4]-[6] za skremblovanje (kriptovanje) sadržaja vodenog žiga. Skremblovanjem se prepoznatljiv sadržaj žiga preuređuje u naizgled besmislen. Osnovna ideja ove transformacije se zasniva na uzastopnoj primeni više različitih Arnoldovih transformacija - etapa ( $I$ ) sa sopstvenim parametrima na vodeni žig. Transformacioni parametri  $i$ -te etape  $a_i$ ,  $b_i$ , broj uzastopnih iteracija etape  $k_i$  predstavljaju ključeve za kriptovanje, dok se perioda Arnoldove transformacije etape  $T_i$  dodatno zahteva za dekriptovanje [4]. Dimenzija kvadratnog vodenog žiga  $N$  je konstanta i jednaka je za sve etape. Ovo ima za posledicu da se višestapna Arnoldova transformacija uvek primenjuje nad kompletnim vodenim žigom. U ovom radu se primenjuje generalizovana višestapna Arnoldova transformacija GMSAT kod koje je dozvoljena varijacija dimenzije kvadratnog vodenog žiga  $N_i$  u svakoj etapi. Tako, u svakoj etapi ( $i$ ) ove transformacije moguće je birati proizvoljnu vrednost dimenzije kvadratnog žiga na koju se primenjuje uz uslov da je  $N_i \leq N$ . Svaka od etapa generalizovane višestapne 2D Arnoldove transformacije ( $i$ ) se može opisati izrazima (1) i (2):

$$\begin{bmatrix} x_{n+1} \\ y_{n+1} \end{bmatrix} = \begin{bmatrix} 1 & b_i \\ a_i & a_i b_i + 1 \end{bmatrix} \begin{bmatrix} x_n \\ y_n \end{bmatrix} \text{ mod } N_i \quad (1)$$

$$N_i \leq N, i \in (1, 2, \dots, I)$$

$$(x, y) \in (0, 1, \dots, N_i - 1) \times (0, 1, \dots, N_i - 1) \subset Z^2 \quad (2)$$

gde  $x_n$ ,  $y_n$  i  $x_{n+1}$ ,  $y_{n+1}$  predstavljaju lokacije piksela slike, a  $a_i$ ,  $b_i$  i  $N_i$  predstavljaju parametre Arnoldove transformacije. Primena generalizovane višestapne Arnoldove transformacije zahteva poznavanje i dodatnih parametara  $k_i$  i  $T_i$  [4]. Skup parametara  $Key_I$  koji određuju generalizovanu višestapnu Arnoldovu transformaciju se može predstaviti izrazom (3):

$$Key_I = f(E_i(a_i, b_i, k_i, N_i, T_i)), i = 1, 2, \dots, I \quad (3)$$

gde  $E_i$  predstavlja  $i$ -tu etapu od  $I$  etapa GMSAT-a. Prilikom skremblovanja (preuređivanja) lokacije piksela na ulaz prve etape  $E_1$  dovodi se originalni vodeni žig, dok se na izlazu iz  $I$ -te etape  $E_I$  dobija transformisani vodeni žig.

### B. Diskretna Wavelet transformacija DWT

Na osnovu svojih karakteristika, DWT [7] se može svrstati u klasu „multi-resolution“ „multi-level“ transformacija što algoritmima za ugradnju vodenog žiga

znatno proširuje mogućnosti primene. U slučaju zaštite videa, vodeni žig se može ugraditi u željeni transformacioni podopseg svakog video frejma. Izborom odgovarajućeg podopsega, ugrađeni vodeni žig se može zaštititi od pojedinih vrsta smetnji i ataka. Na svakom nivou dekompozicije video frejma dobijaju se četiri frekvijska podopsega koji se označavaju sa LL, LH, HL i HH. Podopseg LL je dobijen filtriranjem frejma niskopropusnim filtrom u horizontalnom i vertikalnom pravcu. Ovaj podopseg nosi najveću energiju frejma. Ugradnja vodenog žiga u LL podopsegu može izazvati značajnu degradaciju video frejma, ali istovremeno obezbediti robusnost vodenog žiga na pokušaje njegove ekstrakcije. Frekvijski podopseg HH je dobijen filtriranjem visokopropusnim filtrom u horizontalnom i vertikalnom pravcu i sadrži visokofrekvijske komponente frejma duž dijagonala. Ovaj podopseg uključuje ivice i teksturu frejma. Frekvijski podopsezi HL i LH se dobijaju niskofrekvijskim filtriranjem u jednom pravcu i visokofrekvijskim filtriranjem u drugom pravcu. LH podopseg sadrži informacije o vertikalnim detaljima koji odgovaraju horizontalnim ivicama, dok HL podopseg sadrži informacije o horizontalnim detaljima koji odgovaraju vertikalnim ivicama. Ugradnja vodenog žiga u frekvijske podopsege HL i LH predstavlja kompromis između robusnosti vodenog žiga i perceptualnog kvaliteta zaštićenog videa. Daljom dekompozicijom LL podopsega dobija se sledeći nivo dekompozicije. Nivo dekompozicije zavisi od potrebe aplikacije, a u ovom radu je korišćen drugi nivo dekompozicije sa poznatim *Harr* wavelet filtrom. Na ovaj način su izjednačene rezolucije  $LL_2$  podopsega i žiga.

### C. SVD dekompozicija

SVD dekompozicija se koristi kao alat za unetanje vodenog žiga i bazirana je na teoremi da se pravougaona matrica  $A$  dimenzija  $m \times n$  može razložiti na tri matrice [1]:

$$A = USV^T \quad (4)$$

gde je  $A \in \mathcal{R}^{n \times n}$ ,  $U \in \mathcal{R}^{n \times m}$  i  $V \in \mathcal{R}^{m \times m}$ . Matrice  $U$  i  $V$  su ortogonalne matrice, a kolone ovih matrica se nazivaju levi, odnosno, desni singularni vektori. Matrica,  $S$  je dijagonalna matrica, poznata pod nazivom matrica singularnih vrednosti. Ako je  $r$  rang matrice  $A$ , tada elementi matrice  $S$  zadovoljavaju sledeću relaciju:

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r \geq \sigma_{r+1} = \sigma_{r+2} = \dots = \sigma_n = 0, \quad (5)$$

a matrica  $A$  se može predstaviti na sledeći način:

$$A = \sum_{p=1}^r \sigma_p \mathbf{u}_p \mathbf{v}_p^T, \quad (6)$$

gde  $\mathbf{u}_p$  i  $\mathbf{v}_p$  predstavljaju  $p$ -tu sopstvenu vrednost matrica  $U$  i  $V$ , dok je  $\sigma_p$   $p$ -ta singularna vrednost. Singularni vektori specificiraju geometriju matrice  $A$ , dok singularne vrednosti specificiraju energiju (osvetljaj slike) matrice  $A$ . Ako se matricom  $A$  predstavi jedan video frejma, onda se nizom sličnih matrica može predstaviti video.

Najvažnije karakteristike SVD dekompozicije koje su važne za ovaj rad su invarijantnost na transponovanje, skaliranje, rotaciju i zamenu kolona i vrsta matrice. Ove karakteristike su značajne za očuvanje samog video sadržaja kao i za obezbeđenje otpornosti insertovanog vodenog žiga na geometrijske i druge atake.

## III. MODIFIKOVANI POUZDANI DWT-SVD ALGORITAM

U ovom radu se ugradnja i ekstrakcija vodenog žiga u video frejm u SVD domenu obavlja pouzdanim algoritmom [3]. Pouzdani SVD algoritam rešava problem lažne detekcije žiga koji je svojstven standardnom SVD algoritmu. Detalji algoritma za ugradnju i ekstrakciju kriptovanog vodenog žiga u DWT-SVD domenu su predstavljeni nizom sledećih  $I$  i  $E$  koraka.

## A. Algoritam ugradnje vodenog žiga

**Korak  $I_1$ :** Dekompozicija frejma  $F$  primenom drugog nivoa DWT transformacije:

$$\{\mathbf{F}^k, \mathbf{F}^l\} = DWT_2(\mathbf{F})_{Haar} \quad (7)$$

$$k \in \{LL_2, HL_2, LH_2, HH_2\}$$

$$l \in \{HL_1, LH_1, HH_1\}$$

**Korak  $I_2$ :** SVD dekompozicija podopsega  $F^k$ :

$$\mathbf{F}^k = \mathbf{U}_F^k \cdot \mathbf{S}_F^k \cdot (\mathbf{V}_F^k)^T \quad (8)$$

**Korak  $I_3$ :** Kriptovanje originalnog vodenog žiga  $W'$  (niže rezolucije) primenom generalizovane višestapne Arnoldove transformacije i dobijanje vodenog žiga  $W$  koji se insertuje u svaki frejm.

$$\mathbf{W} = Gen\_Arnold(\mathbf{W}')_{E_i(a_i, b_i, k_i, N_i, T_i)}$$

$$i = 1, 2, \dots, I \quad (9)$$

**Korak  $I_4$ :** SVD dekompozicija kriptovanog vodenog žiga  $W$  i računanje principalne komponente  $A_{wa}$  [8].

$$\mathbf{W} = \mathbf{U}_w \cdot \mathbf{S}_w \cdot \mathbf{V}_w^T = \mathbf{A}_{wa} \cdot \mathbf{V}_w^T; \mathbf{A}_{wa} = \mathbf{U}_w \cdot \mathbf{S}_w \quad (10)$$

**Korak  $I_5$ :** Ugradnja principalne komponente  $A_{wa}$  u dijagonalnoj matrici podopsega  $S_F^k$  sa faktorom insertovanja  $\alpha$ :

$$\mathbf{S}_{F,1}^k = \mathbf{S}_F^k + \alpha \cdot \mathbf{A}_{wa} \quad (11)$$

**Korak  $I_6$ :** Kreiranje modifikovanog podopsega sa ugrađenim vodenim žigom:

$$\mathbf{F}_w^k = \mathbf{U}_F^k \cdot \mathbf{S}_{1,F}^k \cdot (\mathbf{V}_F^k)^T \quad (12)$$

**Korak  $I_7$ :** Zamena originalnih podopsega drugog nivoa frejma sa modifikovanim i primena inverzne diskretne wavelet transformacije  $IDWT_2$  za dobijanje žigovanog frejma.

$$\mathbf{F}_w = IDWT_2(\mathbf{F}_w^k, \mathbf{F}^l)_{Haar} \quad (13)$$

## B. Algoritam ekstrakcije vodenog žiga

Proces ekstrakcije vodenog žiga  $W^*$  iz zaštićenog videa se može obaviti sledećim  $E$  koracima:

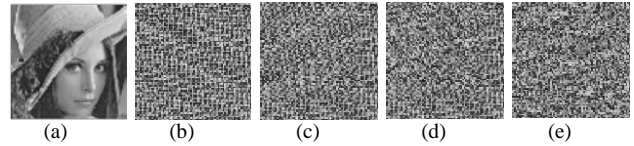
**Korak  $E_1$ :** Dekompozicija originalnog frejma  $F$  primenom drugog nivoa DWT transformacije:

$$\{\mathbf{F}^k, \mathbf{F}^l\} = DWT_2(\mathbf{F})_{haar} \quad (14)$$

$$k \in \{LL_2, HL_2, LH_2, HH_2\}$$

$$l \in \{HL_1, LH_1, HH_1\}$$

**Korak  $E_2$ :** SVD dekompozicija podopsega  $F^k$ :



Sl. 1. Izgled vodenog žiga (a) original, posle (b) prve (c) druge (d) treće (e) četvrte etape GMSAT-a.

$$\mathbf{F}^k = \mathbf{U}_F^k \cdot \mathbf{S}_F^k \cdot (\mathbf{V}_F^k)^T \quad (15)$$

**Korak  $E_3$ :** Dekompozicija potencijalno atakovanog frejma  $F_w^*$  primenom drugog nivoa DWT transformacije:

$$\{\mathbf{F}_w^{*k}, \mathbf{F}_w^{*l}\} = DWT_2(\mathbf{F}_w^*)_{haar} \quad (16)$$

**Korak  $E_4$ :** SVD dekompozicija podopsega  $F_w^{*k}$ :

$$\mathbf{F}_w^{*k} = \mathbf{U}_{F_w}^{*k} \cdot \mathbf{S}_{F_w}^{*k} \cdot (\mathbf{V}_{F_w}^{*k})^T \quad (17)$$

**Korak  $E_5$ :** Kreiranje razlike originalnog ( $F^k$ ) i zaštićenog frejma ( $F_w^{*k}$ ):

$$\mathbf{F}_1^k = \mathbf{F}_w^{*k} - \mathbf{F}^k \quad (18)$$

**Korak  $E_6$ :** Određivanje principalne komponente [7]:

$$\mathbf{A}_{wa}^{*k} = \frac{(\mathbf{U}_F^k)^{-1} \cdot \mathbf{F}_1^k \cdot (\mathbf{V}_F^k)^{-1} \cdot (\mathbf{V}_F^k)^T}{\alpha} \quad (19)$$

**Korak  $E_7$ :** Izračunavanje insertovanog kriptovanog vodenog žiga  $W^{*k}$  se obavlja na sledeći način:

$$\mathbf{W}^{*k} = \mathbf{A}_{wa}^{*k} \cdot \mathbf{V}_w^T \quad (20)$$

**Korak  $E_8$ :** Dekriptovanje originala vodenog žiga  $W^{*k}$  primenom inverzne generalizovane višestapne Arnoldove transformacije [4] i dobijanje originalnog žiga  $W^{*k}$ :

$$\mathbf{W}^{*k} = Inv\_Gen\_Arnold(\mathbf{W}^{*k})_{E_i(a_i, b_i, k_i, N_i, T_i)}$$

$$i = 1, 2, \dots, I \quad (21)$$

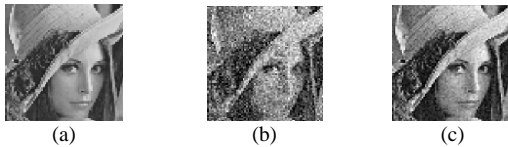
## IV. EVALUACIJA PREDLOŽENOG ALGORITMA

Ne umanjujući značaj prikazanog algoritma, u ovom radu je pretpostavljeni .yuv format za čuvanje nekodovanog video sadržaja. Napomenimo da se kod ovog formata informacije o svakom video frejmu čuvaju u karakterističnim matricama označenim kao  $Y$ ,  $C_b$  i  $C_r$ . U matrici  $Y$  se čuvaju vrednosti osvetljaja za svaki piksel frejma, dok se informacije o boji čuvaju u matricama  $C_b$  i  $C_r$ . Zbog robusnosti, umetanje vodenog žiga u nekodovanom domenu se obavlja upravo u matrici  $Y$ .

Za razliku od prethodnih radova, gde je insertovanje vodenog žiga obavljeno u DCT-SVD domenu, u ovom je insertovanje vodenog žiga obavljeno u DWT-SVD domenu.

## A. Simulacioni rezultati

Kao vodeni žig u eksperimentalnom delu ovog rada korišćen je adaptirani centralni deo poznate slike „Lena.bmp“ u rezoluciji 72×72 piksela. Izgled originalnog vodenog žiga prikazan je na sl. 1(a). Da bi se povećao nivo zaštite, sadržaj ovog vodenog žiga se pre insertovanja skrembluje GMSAT algoritmom. U ovom radu je primenjena 4-etapni GMSAT sa parametrima prikazanim u Tabeli 1. Na osnovu vrednosti parametara  $N_1$  i  $N_4$  može se zaključiti da će se u 1. i 4. etapi primeniti Arnoldova



Sl. 2. Izgled (a) originalnog vodenog žiga (b) vodenog žiga ekstrahovanog iz 50. frejma i (c) vodeni žiga posle primene naprednog algoritma za popravku.

transformacija na ceo vodeni žig dobijen u prethodnim etapama. Skup parametara (22) predstavlja ključ  $Key_4$  za skremblovanje vodenog žiga 4-etapnim GMSAT-om u ovom radu:

$$Key_4 = f(E_1, E_2, E_3, E_4). \quad (22)$$

Na Sl. 1(b) - 1(e) prikazani su dobijeni vodeni žigovi na kraju svake etape GMSAT-a. U svim etapama GMSAT-a se jasno može uočiti prostorna dekorelacija piksela originalnog vodenog žiga. Za vraćanje skremblovanog vodenog žiga u originalni, nužno je posedovati inverzni GMSAT, parametre svih etapa kao i početne uslove svake etape. Ako samo jedan parametar GMSAT-a nije poznat nije moguće dešifrovati skremblovani vodeni žig. Iako se insertovanje vodenog žiga može obaviti u svim DWT podopsezima, u ovom radu je insertovanje skremblovanog vodenog žiga sa Sl. 1(e) obavljeno samo u LL2 DWT podopsegu. Prikazanim DWT-SVD algoritmom insertovanja ugrađen je kriptovan vodeni žig sa faktorom insertovanja  $\alpha=0.05$ . Posle insertovanja vodenih žigova u sve frejmove nekodovanog videa, izvršeno je kodovanje H.264/AVC koderom. Kodovanje zaštićene video sekvence je obavljeno JM referentnim softverom ITU-a (engl. *International Telecommunication Union*) u verziji 18.4 FRExt. Kvalitet kodovanja je definisan skupom FRExt parametara od kojih ključni uticaj imaju sledeći parametri:  $IntraPeriod = 12$ ,  $NumberReferenceFrames = 5$ ,  $NumberBFFrames = 1$ . U procesu dokazivanja vlasništva autorskih prava se iz dekodovanog zaštićenog videa može ekstrahovati ugrađeni vodeni žig. Iz svakog frejma se ekstrahuju vodeni žigovi koji zbog pomenutih razloga imaju promenljiv kvalitet. SSIM indeks ekstrahovanih vodenih žigova se kreće u opsegu od 0.05 do 0.52. Veći SSIM indeks predstavlja bolji kvalitet ekstrahovanog vodenog žiga, a njegova maksimalna vrednost iznosi 1. Da bi se popravio kvalitet ekstrahovanog vodenog žiga primenjen je napredni algoritam za popravku kvaliteta [3]. Broj realizovanih iteracija ovog algoritma je znatno veći u odnosu na broj iteracija prilikom korišćenja DCT-SVD domena za insertovanje vodenog žiga. U ovom slučaju je algoritam popravke ekstrahovanog vodenog žiga načinio 32 iteracije, dok je u radu [3] realizovao samo 12 iteracija. Veći broj iteracija ovog algoritma obezbedio je i veći kvalitet ekstrahovanog vodenog žiga. Na sl. 2(a) prikazan je izgled originalnog vodenog žiga, na sl. 2(b) izgled vodenog žiga ekstrahovanog iz 50. frejma, dok je na sl. 2(c) prikazan izgled popavljenog vodenog žiga iz prvih 50 frejmova zaštićenog videa. Normalizovani kros-korelacioni koeficijent za popavljeni vodeni žig iznosi  $NC=0.9852$ , dok je njegov SSIM indeks 0.8956. Dobijeni rezultati opravdavaju primenu predloženog algoritma jer je u poređenju sa rezultatima publikovanim u [9] ( $NC=0.9824$ ) realizovan nešto veći NC koeficijent.

TABELA 1: PARAMETRI 4-ETAPNOG GMSAT-A.

Parametri	$E_1$	$E_2$	$E_3$	$E_4$
a	2	1	4	3
b	2	1	2	1
N	72	60	50	72
k	9	5	7	7
T	12	60	18	18

## V. ZAKLJUČAK

Zaštita video sadržaja isertovanjem skremblovanog vodenog žiga u DWT-SVD domenu je prikazana u ovom radu. U komparaciji sa rezultatima dobijenim za DCT-SVD domenu, realizovan je veći SSIM indeks ekstrahovanog vodenog žiga za oko 32.3%. Ova značajano povećanje SSIM indeksa se može pripisati nešto većem prosečnom SSIM indeksu ekstrahovanih vodenih žigova, kao i većem broju iteracija algoritma za popravku. U nastavku istraživanja biće ispitane mogućnosti insertovanja vodenih žigova u druge podopsege DWT-a i njihov uticaj kako na kvalitet zaštićenog videa tako i na kvalitet ekstrahovanog vodenog žiga.

## LITERATURA

- [1] I. Cox, M. Miller, J. Bloom, J. Fridrich, T. Kalker, *Digital Watermarking and Steganography*, Elsevier, 2008.
- [2] M. Jevtović, Z. Veličković, "Protokoli prepletenih slojeva", Akademska misao, Beograd, 2012.
- [3] Z. Veličković, Z. Milivojević, M. Veličković, M. Jevtović, "The impact of prediction structures H.264 encoder on the quality of the extracted watermark from the chaos domain", *ETF Jour. of Electrical Engineering*, Vol. 22, pp. 111-121, Podgorica, 2016.
- [4] Z. Veličković, M. Veličković, Z. Milivojević, *Improved Gray-Scale Watermark Encryption Based on Chaotic Maps*, UNITECH 2016, pp. II-145-150, Gabrovo, 2016.
- [5] N. Nikolaidis, A. Tefas, I. Pitas, "Chaotic sequence for digital watermarking", *Advances in Nonlinear Signal and Image Processing*, Ed. S. Marshall and G. L. Sicuranza, Hindawi, 2006.
- [6] Z. Veličković, Z. Milivojević, M. Veličković "The Insertion of the Encrypted Low-Resolution Watermark in the Uncompressed Video", *ICEST 2016*, pp. 191-194, Ohrid, 2016.
- [7] M. Vetterli, J. Kovačević, *Wavelets and Subband Coding*, Prentice Hall, Englewood Cliffs, NJ, 1995.
- [8] C. Jain, S. Arora, P. Panigrahi, "A Reliable SVD based Watermarking Scheme", *Journal CoRR*, vol. abs/0808.0309, 2008.
- [9] R. M. Ibrahim, N. Kader, M. Zorkany, "Video Multiple Watermarking Technique Based on Image Interlacing Using DWT," *The Scientific World Journal*, Hindawi, Vol. 2014.

## ABSTRACT

In this paper, the copy protection of video content by inserting scrambled watermark is considered. A watermark is scrambled by GMSAT algorithm, and insertion was performed in DWT-SVD domain of each frame. By applying advanced algorithms for correcting the quality of the extracted watermark is realized an excellent SSIM index of 0.8956, i.e. NC coefficient of 0.9852 which is at par with the best published results. Applied techniques provide a higher quality of the extracted watermark compared to the techniques of insertion DCT-SVD domain.

## INSERTING THE WATERMARK SCRAMBLED BY GMSAT ALGORITHM IN DWT-SVD DOMAIN

Zoran S. Veličković, Zoran N. Milivojević, and Marko Z. Veličković

# Uporedna analiza mera za dinamičku procenu uspešnosti sjedinjavanja video sekvenci

Rade M. Pavlović, Vladimir S. Petrović, Ljubiša D. Tomić

**Sadržaj** — U ovom radu dati su rezultati objektivnih mera sa različitim pristupima za dinamičko sjedinjavanje video sekvenci. Generalno objektivne mere mogu se podeliti u 4 grupe i iz svake od grupa je uzeta po jedna mera kako bi se videle njihove sličnosti i razlike. Za analizu je uzeta baza od video sekvenci sa različitim scenarijima, kako bi se uzeo u obzir što veći opseg realnih slučajeva. Metode uzete za proces sjedinjavanja su sa različitom računarskom efikasnošću i koriste različite pristupe izbora sjedinjenih koeficijenata.

**Ključne reči** — mera kvaliteta, objektivna mera, video baza, video sjedinjavanje.

## I. UVOD

Osnovna ideja osmatranja spregom više senzora je eliminacija efekata njihovih pojedinih nedostataka što obezbeđuje uspešno funkcionisanje sistema u svim uslovima. Primenom više senzora zasnovanih na merenju različitih fizičkih fenomena, kao na primer toplotnog zračenja (IC senzori) i refleksije svetlosti (kamere u vidnom opsegu), dobija se više realnih informacija o objektima koje posmatramo [1].

Sjedinjavanje slika dobijenih sa više senzora prevazilazi ograničenja pojedinačnih senzora, kombinujući izvorne slike u kompozitnu sliku [2]. U najvećem broju slučajeva televizijskoj kameri pridodata je termovizijska kamera i slike sa oba senzora se kombinuju u jednu sliku. Ovakav pristup prevazilazi probleme obrade veće količine informacija. Dok je televizijska kamera pogodna za rad u dnevnim uslovima, termovizijska kamera se koristi u uslovima slabe osvetljenosti scene.

Za procenu uspešnosti sjedinjavanja i mogućnost primene određene metode u sistemu za sjedinjavanje slika razvijene su metode koje kao izlaz daju odgovarajući skor za testirane metode. Osnovna podela na mere je na subjektivne, koje izvode ljudi kao krajnji korisnici sistema i objektivne, koje se izvršavaju na računaru [3]. Iako je razvijen veći broj objektivnih mera za procenu uspešnosti sjedinjavanja slika i dalje ne postoji mera koja je opšteprihvaćena i pogodna za procenu u svim uslovima sjedinjavanja. Sa druge strane subjektivne mere koje su

pouzdanije u smislu da je čovek krajnji korisnik sistema i da se njegova procena uzima kao relevantna, komplikovane su za izvođenje. One zahtevaju duži period planiranja, strogo kontrolisane uslove i veći broj ispitanika. Pogotovo je složeno ponavljanje mera u slučaju promena parametara na metodama za sjedinjavanje ili na samim sensorima. Sa druge objektivne metode se mogu ponavljati veći broj puta i ne zahtevaju posebne uslove za primenu.

U literaturi postoji više različitih pristupa za objektivnu procenu uspešnosti sjedinjavanja slika. Rezultati mera su u različitim opsezima (u najvećem broju slučajeva između nule i jedan), neke od metoda predstavljaju bolje rezultate ukoliko im je vrednost manja, što dovodi do otežanog poređenja i izbora pogodne mere za procenu implementirane metode ili postojećih metoda. Možda je najpribližnija podela i svrstavanje metoda za procenu uspešnosti sjedinjavanja u 4 osnovne grupe [4], prema njihovom pristupu prilikom procene i to: i) mere zasnovane na teoriji informacija, ii) mere zasnovane na odlikama slika, iii) strukturalna sličnost slika kao mera i iv) mere inspirisane ljudskom percepcijom.

## II. OBJEKTIVNE DINAMIČKE MERE

Od osnovnog značaja je primena sjedinjavanja slika u realnom vremenu kao na primer u vojsci, civilnoj avijaciji ili medicini. To znači da se video sekvenca od 30 i više slika iz različitih senzora sjedinjuje u jednu izlaznu sekvencu svake sekunde [5]. Većina algoritama za procenu uspešnosti sjedinjavanja slika napravljeni su za mirne slike i daju dobre rezultate prilikom sjedinjavanja prostornih informacija iz ulaznih slika u sjedinjenu sliku. Ovo ne znači da se ovi algoritmi ne mogu koristiti za sjedinjavanje video sekvenci ali pri njihovom korišćenju javljaju se problemi. Jedan od njih je vremenska stabilnost zbog promena na sceni zbog koje bi trebalo menjati parametre kroz sekvencu, na primer broj nivoa dekompozicije kod Laplasove [6] ili ROLP piramide [7] ili veličina prozora filtra kod računarski efikasne metode [8], kako bi dobili što bolje rezultate sjedinjavanja. Pored toga računarska zahtevnost predloženih algoritama nije bitna ako se radi o sjedinjavanju mirnih slika ali je veoma važna ako se radi sjedinjavanje u realnom vremenu. Iz ovih razloga uzima se kompromis između računarske zahtevnosti i rezultata koje daju algoritmi za sjedinjavanje.

Prema tome potrebno je da metoda bude pouzdana za proračun uspešnosti sjedinjavanja video sekvenca i za poređenje rezultata različitih metoda sjedinjavanja. Mere

Rade M. Pavlović, (autor za kontakte), Vojnotehnički institut, Ministarstvo odbrane Republike Srbije, Ratka Resanovića 1, 11000 Beograd, Republika Srbija (telefon: 381-11-2051846; e-mail: rade\_pav@yahoo.com).

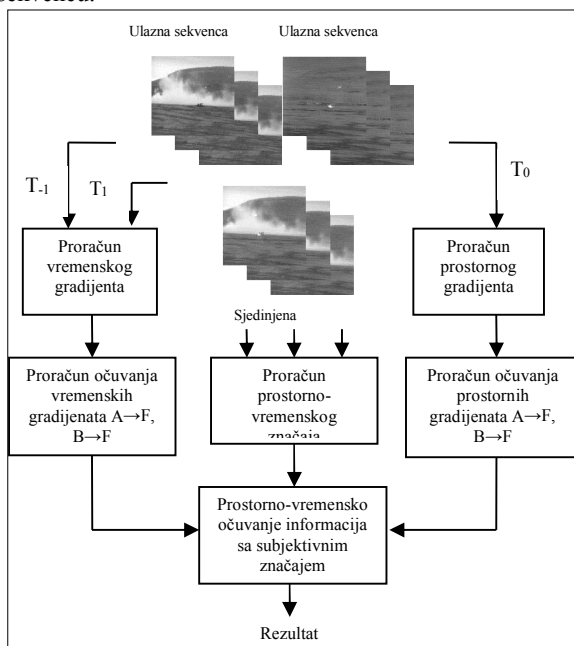
Vladimir S. Petrović, Fakultet tehničkih nauka, Trg Dositeja Obradovića 6, 21000 Novi Sad, Republika Srbija (e-mail: v.petrovic@manchester.ac.uk).

Ljubiša D. Tomić, Vojnotehnički institut, Ministarstvo odbrane Republike Srbije, Ratka Resanovića 1, 11000 Beograd, Republika Srbija (e-mail: ljubisa.tomic@gmail.com)

treba da obezbede promenu parametara algoritama sjedinjavanja u zavisnosti od uslova. Ovo je posebno pogodno za sjedinjavanje sekvenci u realnom vremenu gde se uslovi menjaju posle dužeg vremenskog perioda pa se optimizacija parametara može izvršavati za veći broj frejmova. Većina mera za procenu uspešnosti sjedinjavanja video sekvenci bazirana je na merama za mirne slike uzimajući vremensku komponentu u proračun. Na taj način razvijene su i tri najzastupljenije mere za video sekvenci čija će analiza i rezultati biti predstavljeni u narednom delu.

#### A. Gradijentna mera (DQ)

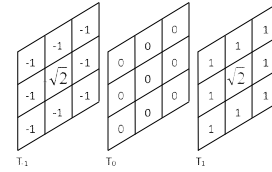
Dinamička, gradijentna mera za proračun uspešnosti sjedinjavanja sekvence slika [5],  $DQ$ , prikazana je na slici 1. Ova mera daje prenesene informacije iz ulaznih sekvenci u sjedinjenu sekvencu. Proračun mere zasnovan je na tri uzastopna frejma sve tri sekvence. Za tekući frejm iz sve tri sekvence koristi se mera za očuvanje ulaznih informacija u sjedinjenoj sekvenci. Ovaj doprinos očuvanju informacija računa se za svaku lokaciju  $m, n$  na slici. Međutim, prethodni i sledeći frejm iz svake sekvence koriste se za proračun vremenskog gradijenta informacija koji obezbeđuje proračun očuvanja ulaznih vremenskih informacija na svakoj lokaciji u sjedinjenu sekvencu. Oba skupa proračuna integrišu se u jedan skup za proračun očuvanja prostorno-vremenskih informacija za svaki piksel na slici. Ovi proračuni množe se težinskim koeficijentima koji su dobijeni kao rezultat subjektivnog značaja za tekući frejm. Individualni rezultati za svaki frejm se onda usrednjavaju da bi se dobila jedinstvena mera za celu sekvencu.



Sl. 1. Realizacija dinamičke mere za procenu uspešnosti sjedinjavanja slika.

Vremenske informacije dobijaju se iz sve tri sekvence koristeći vremenski gradijentni operator, prikazan na slici 2. Ovaj operator je analogan standardnom Sobel ivičnom operatoru koji je korišćen za dobijanje prostornih

informacija za meru. Ovaj operator je simetričan u prostornom domenu.



Sl.2. Gradijentni operator.

#### B. Mera indeksa strukturne sličnosti ( $Q_t$ )

Mera zasnovana na indeksu strukturne sličnosti –  $SSIM$  (structural similarity) za mirne slike može se prilagoditi i za video sekvence [9]. Kao i kod prethodne mere od osnovnog interesa je očuvanje prostornih informacija i vremenske stabilnosti na koje je ljudski vizuelni sistem osetljiv. Osnovni princip mere je sličan kao kod mere za mirne slike za dobijanje prostornog značaja pomoću  $SSIM$  uz odgovarajuće težinske faktore, dok se za dinamički značaj uzima razlika sukcesivnih frejmova uz težinske komponente.

Procena uspešnosti prenesenih prostornih informacija  $S$  (spatial) iz ulaznih videa  $V_a$  i  $V_b$  u sjedinjeni  $V_f$  veličine  $M \times N$ , na osnovu  $SSIM$  indeksa za tekući frejm  $t$ , za svaki piksel  $m, n$  izračunava se prema:

$$Q_{S,t} = \frac{\sum_{m=1}^M \sum_{n=1}^N \lambda_a SSIM(V_{a,t}, V_{f,t}) + \lambda_b SSIM(V_{b,t}, V_{f,t})}{\sum_{m=1}^M \sum_{n=1}^N \lambda_a + \lambda_b} \quad (1)$$

Lokalni težinski faktori  $\lambda_a$  i  $\lambda_b$  određuju lokalni značaj svakog piksela  $m, n$  za frejm ulaznih videa. Regioni sa više energije ili sa većom varijansom sadrže više informacija i samim time privlače i više pažnje od posmatrača.

Vremenski indeks kvaliteta sjedinjavanja  $Q_{T,t}$  računa se na sličan način kao i prostorni, samo što se umesto proračuna  $SSIM$  indeksa za tekući frejm koristi razlika dva frejma  $dV_a$ ,  $dV_b$  i  $dV_f$  za ulazne i sjedinjeni video i težinski koeficijenti  $\xi_a$  i  $\xi_b$  koji odgovaraju promenama u video signalu. Težinski koeficijenti predstavljaju značaj promena u video signalu i slično kao kod prostornih koeficijenata, da bi se povećala otpornost na prisustvo šuma računaju se prema jednačini 2.

$$\xi_{a/b} = 1 + \log\left(1 + \left\| v_{r^{a/b}}(m, n, t) \right\|_2\right), \quad (2)$$

gde  $v_r$  predstavlja brzinu promene promene piksela i računa se iz algoritma za optički tok (Optical flow) [10], dok indeks  $a/b$  označava prvi, odnosno drugi video.

Konačno, kombinujući prostorno i vremensko očuvanje informacija dobijenih pomoću  $SSIM$  indeksa dobija se jedinstvena mera za dinamičko sjedinjavanje video sekvenci za svaki frejm kao proizvod pojedinačnih mera sa različitim doprinosima  $\gamma$ :

$$Q_t = Q_{S,t}^\gamma \cdot Q_{T,t}^{1-\gamma}, \quad (3)$$

gde se koeficijent  $\gamma$  nalazi u opsegu od 0 do 1 i definiše relativni značaj prostorne i vremenske komponente mere.



### C. Mera zasnovana na prostorno-vremenskim frekvencijama (*rSTFe*)

Kako bi uzeli u obzir i vremenske promene na slici, kao nastavak mere za procenu uspešnosti sjedinjavanja slika na osnovu prostornih frekvencija nastala je mera za dinamičko sjedinjavanje video sekvenci koja je zasnovana na prostorno-vremenskom očuvanju frekvencija iz ulaznih u sjedinjenu sekvencu [11]. Za proračun vremenskog očuvanja gradijenata koristimo tri uzastopna frejma iz ulaznih i sjedinjene sekvence i računamo dve vremenske frekvencije za određeni frejm, odnosno prethodnu i narednu vremensku frekvenciju (*BTF* i *ATF*):

$$BTF / ATF = \sqrt{\frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N [I_T(m,n) - I_{T+1}(m,n)]^2} \quad (4)$$

Ukupna prostorno-vremenska frekvencija (STF) za posmatrani frejm u trenutku  $t$  dobija se kao kombinacija prostornih frekvencija u četiri pravca i dve vremenske frekvencije.

Za dobijanje referentne prostorno-vremenske frekvencije, najpre se izračunaju vremenski gradijenti za oba vremenska pravca (prethodni i sledeći) kao maksimalna apsolutna vrednost vremenskih gradijenata. Vremenski gradijenti se dobijaju kao razlika tekućeg frejma i prethodnog odnosno sledećeg frejma. Prethodni vremenski gradijent sa indeksom B za tekući frejm  $t$  iz ulazne video sekvence  $V_a$  računa se kao:

$$Grad^B(V_{a,t}(m,n)) = V_{a,t}(m,n) - V_{a,t-1}(m,n) \quad (5)$$

Analogno proračunu prostorne mere za procenu uspešnosti sjedinjavanja mirnih slika, mera zasnovana na prostorno-vremenskom očuvanju frekvencija za tekući frejm  $t$  dobija se iz relacije:

$$rSTFe(t) = (STF_F(t) - STF_R(t)) / STF_R(t) \quad (6)$$

Ukupna mera za celu sekvencu sa  $N_T$  frejmova dobija se usrednjavanjem apsolutnih vrednosti za svaki frejm pojedinačno:

$$rSTFe = \frac{1}{N_T - 2} \sum_{t=2}^{N_T-1} abs(rSTFe(T)) \quad (7)$$

U ukupnu meru za sekvencu uzima se apsolutna vrednost mere za svaki frejm iz razloga što one mogu uzeti i pozitivnu i negativnu vrednost i samim tim bi dolazilo do potiranja. Vrednost mere 0 predstavlja idealno sjedinjavanje, dok što je vrednost veća to je manja uspešnost.

### D. Mera kroz zajedničke informacije (DMI)

Zajedničke informacije mogu se koristiti za proračun uspešnosti sjedinjavanja sekvence slika. Oliver Rockinger i Thomas Fechner [12] pedložili su meru za sjedinjavanje sekvence slika koja se bazira na proračunu zajedničkih informacija razlika dva frejma ulaznih slika i sjedinjene slike. Jedan od glavnih ciljeva sjedinjavanja sekvence slika je vremenska stabilnost i saglasnost sjedinjene sekvence. Vremenska stabilnost podrazumeva da promene vrednosti nivoa sivog u sjedinjenoj sekvenci mogu biti

prouzrokovane samo promenama nivoa sivog u ulaznim sekvencama ili bar jednoj ulaznoj sekvenci, tj. da ove promene ne smeju biti prouzrokovane samim algoritima za sjedinjavanje. Vremenska postojanost sjedinjene sekvence podrazumeva da promene u ulaznim sekvencama moraju biti predstavljene u sjedinjenoj sekvenci bez kašnjenja ili promene kontrasta. Autori u ovom radu porede razlike između frejmova *IFD* (*inter-frame-differences*) ulaznih sekvenci i sjedinjene sekvence. Razlika između frejmova (*IFD*) između dva sukcesivna frejma sekvence  $V_a(m,n)$  definiše se kao:

$$dV_a(m,n) = V_{a,t}(m,n) - V_{a,t-1}(m,n) \quad (8)$$

Za proračun vremenske stabilnosti i postojanosti, prvo se pronalaze delovi u slikama razlike frejmova sjedinjene sekvence koje se nalaze i na ulaznim sekvencama, odnosno na jednoj ulaznoj sekvenci.

Za proračun uspešnosti sjedinjavanja dve sekvence preko zajedničkih informacija koristi se sledeća jednačina:

$$DMI = H(dV_a, dV_b) + H(dV_f) - H(dV_a, dV_b, dV_f) \quad (9)$$

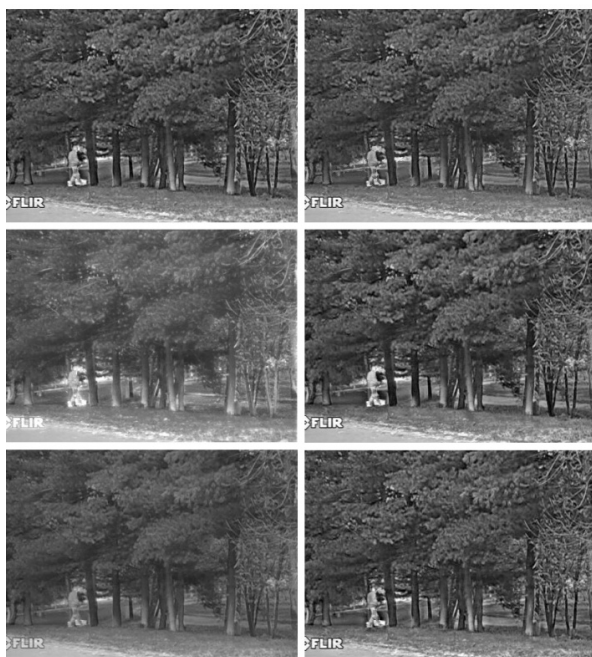
gde su  $H(dV_a, dV_b)$  i  $H(dV_a, dV_b, dV_f)$  združene entropije razlike frejmova ulaznih sekvenci i razlike frejmova ulaznih i sjedinjene sekvence respektivno. Iz jednačine se može videti da ukoliko postoji promena vrednosti nivoa sivog u razlikama frejmova u sjedinjenoj sekvenci koje se ne nalaze u ulaznim sekvencama rezultira povećanje zajedničke entropije  $H(dV_a, dV_b, dV_f)$  koja smanjuje vrednost zajedničkih informacija.

## III. ANALIZA REZULTATA

Za testiranje mera uzeto je ukupno šest metoda za sjedinjavanje video sekvenci koji uzimaju frejm po frejm sjedinjavanje. Kao najjednostavniji način sjedinjavanja, a samim tim i računarski najmanje zahtevna uzeto je aritmetičko sjedinjavanje (AS) sa istim težinskim koeficijentima za oba ulazna seta (0,5 i 0,5). Takođe, računarski efikasna metoda za sjedinjavanje sa dvoveličinskim razlaganjem (DVR) [8] koja koristi samo dva nivoa razlaganja (slika pozadine i slika prvog plana) uzeta je za testiranje mere. Sledeće četiri metode predstavljaju multiveličinske metode za sjedinjavanje slika: Laplasova piramida (LAP) [6], ROLP ili kontrast piramida (ROLP) [7], sjedinjavanje diskretnom vejevlet transformacijom (DWT) [13] i sjedinjavanje vejevlet metodom koja je invarijantna na pomeraj (Shift invariant discrete wavelet SIDWT) [12]. Za sve četiri metode uzete su selekcije apsolutnog maksimuma na višim nivoima i aritmetičko usrednjavanje na najnižem nivou. Sjedinjeni primeri za sve metode iz sekvence va1 (slika 3), prikazani su na slici 4.



Sl. 3. Multisenzorske slike iz sekvence va1.



Sl. 4. Rezultati sjedinjavanja VA1 sekvence sa metodama sledećim redosledom: SIWT, DVR, ROLP, LAP, AS i DWT.

Rezultati mera ispitani su na bazi od 7 video sekvenci sa 6 različitih metoda sjedinjavanja opisanih u prethodnom delu. Pokriven je širok spektar primene sjedinjavanja video signala, tako da će predstavljeni rezultati dati poređenje metoda a i njihovu usaglašenost sa subjektivnim testovima u većem broju realnih uslova. Rezultati su dati u tabeli 1. Takođe, rezultati za meru  $rSTFe$  su normalizovani od 0 do 1, gde 1 predstavlja idealno sjedinjavanje kako bi sve mere imale rezultat "što veći to bolja metoda", što je kod ove mere bilo obrnuto.

TABELA 1: REZULTATI OBJEKTIVNIH MERA.

	AS	DVR	LAP	ROLP	sidwt	DWT
<b>DQ</b>	0.18	0.23	0.24	0.17	0.25	0.176
<b>DMI</b>	9.48	11.9	12.2	11.26	12.44	11.91
<b><math>Q_t</math></b>	0.79	0.86	0.87	0.8	0.87	0.82
<b><math>rSTFe</math></b>	0.41	0.51	0.53	0.47	0.68	0.49

Iz table se vidi da mere imaju različite dinamičke opsege i vrednosti te ih je teško upoređivati bez dodatne obrade. Takođe neke od mera ( $DQ$  i  $rSTFe$ ) su veoma osetljive, tako da razlika u drugoj decimali utiče na određivanje uspešnosti izabrane metode za sjedinjavanje.

U cilju jednostavnijeg poređenja mera za dinamičko sjedinjavanje potrebno je izvršiti rangiranje metoda za sjedinjavanje prema objektivnim metodama. U tabeli 2 su date vrednosti rangova ali sa bodovanjem, tako da najbolji rang dobije 6 bodova, a najlošiji 1 u cilju lakšeg i razumljivijeg prikaza.

Posmatrajući rezultate svi mera i ukupan rezultat zaključuje se da metoda SIDWT u većem broju mera (3 od 4) pokazuje najbolje rezultate, što dovodi do zaključka da su mere poprilično usaglašene. Takođe, mere  $DMI$  i  $Q_T$  daju identične rangove za sve metode uzete u testiranje. Za

dalju analizu potrebno je izvršiti subjektivne testove i poređenje objektivnih i subjektivnih rezultata. Na taj način bi se dobila mera koja je najpribližnija ljudskom vizuelnom sistemu.

TABELA 2: REZULTATI RANGIRANJA OBJEKTIVNIH MERA.

	AS	DVR	LAP	ROL P	SIDWT	DWT
<b>DQ</b>	3	4	5	1	6	2
<b>DMI</b>	1	3	5	2	6	4
<b><math>Q_t</math></b>	1	3	5	2	6	4
<b><math>rSTFe</math></b>	1	4	6	2	5	3
<b>Uk</b>	1	4	5	2	6	3

#### LITERATURA

- [1] V. Petrović, B. Zrnčić, "Multisenzorsko sjedinjavanje informacija za otkrivanje, praćenje i identifikaciju ciljeva", *TELFOR 2001*, Beograd
- [2] S. Rick Blum, L. Zheng, "Multi-sensor image fusion and its applications". CRC press, 2005.
- [3] V. Petrović, C. Xydeas, "Objective evaluation of signal-level image fusion performance," *Optical Engineering* 44(8), August 2005.
- [4] Z. Liu, E. Blasch, Z. Xue, J. Zhao, R. Laganier, W. Wu, "Objective assessment of multiresolution image fusion algorithms for context enhancement in night vision: a comparative study", *Pattern Analysis and Machine Intelligence*, IEEE Transactions on, 34(1), 94-109, 2012.
- [5] V. Petrovic, T. Cootes, R. Pavlovic, "Dynamic image fusion performance evaluation", In *Information Fusion*, 2007 10th International Conference on (pp. 1-7). IEEE.
- [6] P. Burt, E. Adelson, "The Laplacian pyramid as a compact image code", *IEEE Transactions on Communication*, COM-31,1983, str. 532-540
- [7] A. Toet, "Image fusion by a ratio of low-pass pyramid", *Pattern Recognition Letters* 9 North-Holland (1989) str. 245-253,
- [8] V. Petrović, C. Xydeas, "Computationally efficient pixel-level image fusion", *Proceedings of Eurofusion99*, Stanford-upon-Avon, October 1999, str. 177-184
- [9] N. Cvejic, A. Loza, D. Bul, N. Canagarajah, "A Similarity Metric for Assessment of Image Fusion Algorithms," *Int'l J. Signal Processing*, vol. 2, no. 3, pp. 178-182, 2005.
- [10] T. Brox, A. Bruhn, "High accuracy optical flow estimation based on a theory warping", *Proceedings of the 8th European Conference on Computer Vision*, 2004, pp. 25-36.
- [11] R. Pavlović, V. Petrović, M. Stanojević, Đ. Tasić, "Prostorno-vremenske frekvencije kao mera za procenu dinamičkog sjedinjavanja slika", *ETRAN 2016*, Zlatibor, 13-16. Jun 2016. godine
- [12] O. Rockinger, T. Fechner, "Pixel-level image fusion: The case of image sequences", *Proc. SPIE*, Vol. 3374, 1998, str. 378-388
- [13] H. Li, B.S. Manjunath, S.K. Mitra, "Multisensor image fusion using the wavelet transform," *Graphical Models and Image Processing* Vol. 57, No. 3, pp. 235-245, May 1995.

#### ABSTRACT

This paper presents the results of objective measures for video sequences fusion. Objective measures can be divided into 4 groups and from each group a measure is taken for comparison. In order to measure a wide range of real cases, the database of video sequences with different scenarios for the analysis is considered. Methods for the process of fusion are also with different approaches and efficiency.

#### COMPARATIVE STUDY OF DYNAMIC MEASURES FOR VIDEO FUSION

Rade Pavlović, Vladimir Petrović, Ljubiša Tomić

# A Higher Order Time-Frequency Representation With Reduced Cross-Terms

Miloš Brajović, *Student Member, IEEE*, Miloš Daković, *Member, IEEE*,  
Ljubiša Stanković, *Fellow, IEEE*

**Abstract** — A higher order time-frequency (TF) representation, fully concentrated up to the fourth order polynomial phase of the signal is revisited. This representation is based on the first derivative approximation of the phase. We propose a realization based on the concept of windowed frequency convolutions in order to reduce the cross-terms in multi-component signals, following the ideas from the S-method framework. Several numerical examples illustrate and verify the presented theory.

**Keywords** — Digital signal processing, Non-stationary signals, Time-frequency signal analysis

## I. INTRODUCTION

DURING the last few decades time-frequency (TF) signal analysis has been an emerging research field, followed by numerous applications [1]– [8]. Instantaneous frequency (IF) estimation is one of the most important problems analyzed in wide TF analysis literature [1], [5], [6].

The aim to reveal and extract high level signal characteristics such as the IF has resulted in numerous proposed time-frequency representations (TFR), [1]. Many of these representations provide improvements in useful signal content (auto-terms) concentration at or around the IF, comparing with the frequently used Short-time Fourier transform (STFT) and Wigner distribution (WD) [3], [4]. Main problems regarding higher-order representations include high numerical complexity, noise sensitivity, demanding parameters search, and the appearance of undesirable cross-terms in multi-component signals [1], [2].

Following the idea to define a representation based on signal's phase first derivative approximation, a higher order representation being able to ideally concentrate signals having up to the fourth order polynomial phase is proposed in [7]. However, in the case of multi-component signals, this representation produces strong cross-terms, preventing a successful IF estimation. In this paper we present an algorithm for the realization of this TFR with significantly reduced cross-terms.

The paper is organized as follows. In Section II basic theory regarding the analyzed higher order TFR is presented. Its reduced cross-terms realization is introduced in Section III. Numerical results illustrating and confirming the theory are presented in Section IV, whereas the paper ends with concluding remarks.

## II. BACKGROUND THEORY

Starting from the analytic signal definition

$$x(t) = A(t)e^{j\phi(t)}, \quad (1)$$

Authors are with the Faculty of Electrical Engineering, University of Montenegro, Džordža Vasiingtona bb, 81000 Podgorica, Montenegro, e-mail: {milosb, milos, ljubisa}@ac.me

and assuming slow amplitude variations comparing with the instantaneous phase variations  $|A'(t)| \ll |\phi'(t)|$ , the IF is defined as

$$\Omega(t) = \frac{d\phi(t)}{dt}. \quad (2)$$

An ideal TFR can be introduced in the following form:

$$ITF(t, \Omega) = 2\pi|A(t)|^2\delta(\Omega - \phi'(t)), \quad (3)$$

fully concentrating signal energy at the IF. One out of many open topics in the TF signal analysis deals with the development of representations having the form and properties as close as possible to the ITF. This TFR can be considered as the Fourier transform (FT) of function

$$R(t, \tau) = |A(t)|^2 e^{j\phi'(t)\tau}. \quad (4)$$

As the signal phase first derivative can be approximated with [1]:

$$\Omega(t) \approx \frac{\sum_i b_i \phi(t + c_i \tau)}{\tau} = \frac{d\phi(t)}{dt} + O(\phi^{(p)}(\tau)), \quad (5)$$

the general form of function (4) whose FT produces ITF follows

$$R(t, \tau) = \prod_i x^{b_i}(t + c_i \tau). \quad (6)$$

Under the assumption of slow amplitude variations, its influence is neglected in further analysis. General form of TFR based on (6) reads

$$GD(t, \Omega) = \int_{-\infty}^{\infty} \prod_i x^{b_i}(t + c_i \tau) e^{-j\Omega\tau} d\tau. \quad (7)$$

Expanding term  $b_i \phi(t + c_i \tau)$  in Taylor series around  $t$

$$\begin{aligned} b_i \phi(t + c_i \tau) \approx & b_i \phi(t) + b_i \phi'(t) c_i \tau + b_i \phi''(t) \frac{(c_i \tau)^2}{2!} + \\ & + b_i \phi'''(t) \frac{(c_i \tau)^3}{3!} + b_i \phi^{(4)}(t) \frac{(c_i \tau)^4}{4!} + \dots \end{aligned}$$

coefficients  $b_i$  i  $c_i$  follow with the conditions that [1]:

- The sum of coefficients with  $\phi(t)$  is equal to 1, eliminating the signal phase influence;
- The sum of coefficients with  $\phi'(t)$  is equal to 1, as the goal is the first derivative approximation;
- The sum of coefficients with  $\phi^{(n)}(t)$  is equal to 0 up to the desired order.

Either following these conditions, or directly using the well-known first derivative approximation of the form

$$\begin{aligned} \Omega(t) \approx & \frac{\phi(t - \frac{\tau}{6}) - 8\phi(t - \frac{\tau}{12}) + 8\phi(t + \frac{\tau}{12}) - \phi(t + \frac{\tau}{6})}{\tau} \\ = & \frac{d\phi(t)}{dt} + O(\phi^{(5)}(\tau)), \end{aligned}$$

according to (6) we get the function

$$R(t, \tau) = x(t - \frac{\tau}{6})x^{*8}(t - \frac{\tau}{12})x^8(t + \frac{\tau}{12})x^*(t + \frac{\tau}{6}), \quad (8)$$

whose FT is a new TF representation fully concentrating signals up to the fourth order polynomial phase. Its form reads

$$PD(t, \Omega) = \int_{-\infty}^{\infty} x(t - \frac{\tau}{6})x^{*8}(t - \frac{\tau}{12}) \times x^8(t + \frac{\tau}{12})x^*(t + \frac{\tau}{6})e^{-j\Omega\tau} d\tau. \quad (9)$$

The discretization over the time and lag with  $t = m\Delta t$  and  $\tau = n\Delta t$ , with sampling period  $\Delta t = 1/(12f_{\max})$ , and the discretization of frequency  $\omega = \Omega\Delta t$  using  $\omega = \pi k/6N$ , lead to the discrete form of this representation

$$PD(n, k) = 12 \sum_{m=-6N}^{6N-1} x(n-2m)x^{*8}(n-m) \times x^8(n+m)x^*(n+2m)e^{-\frac{j2\pi mk}{N}} \quad (10)$$

suitable for numerical implementations. Maximal frequency in signal spectrum is denoted with  $f_{\max}$ . Note that 6 times more samples are needed for aliasing-free calculation of this representation, compared with the corresponding pseudo-WD with the same window duration.

### III. PROPOSED REALIZATION

#### A. Continuous windowed frequency convolutions

Let us observe the definition (9) of higher order representation  $PD(t, \Omega)$ , assuming, for simplicity, unit symmetric lag window  $w(\tau) = 1$ ,  $-T/2 \leq \tau \leq T/2$ . For each observed time instant  $t$  it can be understood as a frequency domain convolution of the form

$$PD(t, \Omega) = R_{x^2}(t, \Omega) *_{\Omega} R_{x^{16}}(t, \Omega), \quad (11)$$

with the following definitions

$$R_{x^2}(t, \Omega) = FT \left[ x(t - \frac{\tau}{6})x^*(t + \frac{\tau}{6}) \right], \quad (12)$$

$$R_{x^{16}}(t, \Omega) = FT \left[ x^{*8}(t - \frac{\tau}{12})x^8(t + \frac{\tau}{12}) \right]. \quad (13)$$

Operator  $FT[\cdot]$  denotes the Fourier transform over variable  $\tau$ . Furthermore, introducing notation  $S_{n_1}(t, \Omega) = FT \left[ x(t - \frac{\tau}{6}) \right]$  and  $S_{p_1}(t, \Omega) = FT \left[ x^*(t + \frac{\tau}{6}) \right]$  equation (12) can be also represented in form of a frequency domain convolution

$$R_{x^2}(t, \Omega) = S_{n_1}(t, \Omega) *_{\Omega} S_{p_1}(t, \Omega). \quad (14)$$

Following the same approach, introducing

$$W_{x^2}(t, \Omega) = S_{n_2}(t, \Omega) *_{\Omega} S_{p_2}(t, \Omega), \quad (15)$$

with  $S_{n_2}(t, \Omega)$  and  $S_{p_2}(t, \Omega)$  being defined as  $S_{n_2}(t, \Omega) = FT \left[ x^*(t - \frac{\tau}{12}) \right]$  and  $S_{p_2}(t, \Omega) = FT \left[ x(t + \frac{\tau}{12}) \right]$ , we rewrite (13) in terms of frequency convolutions

$$R_{x^{16}}(t, \Omega) = W_{x^2}(t, \Omega) *_{\Omega} W_{x^2}(t, \Omega) *_{\Omega} W_{x^2}(t, \Omega) *_{\Omega} W_{x^2}(t, \Omega). \quad (16)$$

It can be assumed that components  $S_{p_2}(t, \Omega)$  are localized in frequency, such that  $S_{p_2}(t, \Omega)$  centered at any  $\Omega_0$  is spread over a region  $[\Omega_0 - \Omega_L/2, \Omega_0 + \Omega_L/2]$ . This

means that values of  $S_{p_2}(t, \Omega)$  appart from  $\Omega_0$ , that is, outside this region, are not related with value of  $S_{p_2}(t, \Omega_0)$ , for observed instant  $t$ . It should be noted that there is no assumption regarding the exact location of central frequency  $\Omega_0$ . In the multi-component signal case, each component is localized in its own region.

As the symmetric lag window is assumed, the analyzed component is localized within same frequency region  $[\Omega_0 - \Omega_L/2, \Omega_0 + \Omega_L/2]$  in  $S_{n_2}(t, \Omega) = FT \{x^*(t - \frac{\tau}{6})\}$ , also around central frequency  $\Omega_0$ . Please note that the change in the lag sign is compensated by the signal conjugation. Hence, (15) can be further written as:

$$W_{x^2}(t, \Omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{p_2}(t, \varpi) S_{n_2}(t, \Omega - \varpi) d\varpi = \frac{1}{4\pi} \int_{-\Omega_L}^{\Omega_L} S_{p_2}(t, \frac{\Omega}{2} + \frac{\varpi_1}{2}) S_{n_2}(t, \frac{\Omega}{2} - \frac{\varpi_1}{2}) d\varpi_1, \quad (17)$$

where substitution  $\varpi_1 = \Omega/2 - \varpi/2$  is exploited. The integral limits in (17) are changed according to the localization assumption and doubled having in mind the length of the resulting convolution. The signal component in (17) is therefore spread over region  $[2\Omega_0 - \Omega_L, 2\Omega_0 + \Omega_L]$ . Note that this analysis correspond to the well-known S-method [3] principles, leading to a significant cross-terms reduction. The aim is to improve the signal concentration further applying this principle, in order to obtain the higher order representation  $PD(t, \Omega)$ . The next term,  $R_{x^4}(t, \Omega)$  can be calculated as

$$R_{x^4}(t, \Omega) = W_x(t, \Omega) *_{\Omega} W_x(t, \Omega) = \frac{1}{4\pi} \int_{-2\Omega_L}^{2\Omega_L} W_x(t, \frac{\Omega}{2} + \frac{\varpi}{2}) W_x(t, \frac{\Omega}{2} - \frac{\varpi}{2}) d\varpi,$$

with new frequency region where the component is spread being  $[4\Omega_0 - 2\Omega_L, 4\Omega_0 + 2\Omega_L]$ . The same approach is applied until the highest order term is obtained

$$R_{x^{16}}(t, \Omega) = R_{x^8}(t, \Omega) *_{\Omega} R_{x^8}(t, \Omega) = \frac{1}{4\pi} \int_{-8\Omega_L}^{8\Omega_L} R_{x^8}(t, \frac{\Omega}{2} + \frac{\varpi}{2}) R_{x^8}(t, \frac{\Omega}{2} - \frac{\varpi}{2}) d\varpi \quad (18)$$

having region of interest  $[16\Omega_0 - 8\Omega_L, 16\Omega_0 + 8\Omega_L]$ .

In order to calculate (14) for the same analyzed component, frequency regions where the component is spread in  $S_{n_1}(t, \Omega)$  and  $S_{p_1}(t, \Omega)$  have to be related with corresponding region  $[\Omega_0 - \Omega_L/2, \Omega_0 + \Omega_L/2]$  for terms  $S_{p_2}(t, \Omega)$  and  $S_{n_2}(t, \Omega)$ . Based on the definitions of  $S_{n_1}(t, \Omega)$  and  $S_{p_1}(t, \Omega)$  it can be concluded that the components are spread over the same frequency region in these two terms. Let us relate the region of the component in term  $S_{p_1}(t, \Omega) = FT \left[ x^*(t + \frac{\tau}{6}) \right]$  with the frequency region  $[\Omega_0 - \Omega_L/2, \Omega_0 + \Omega_L/2]$  where the component is spread in term  $S_{p_2}(t, \Omega) = FT \left[ x(t + \frac{\tau}{12}) \right]$ . Conjugate operator appearing in definition of  $S_{p_1}(t, \Omega)$  causes the opposite direction of frequency axis compared with  $S_{p_2}(t, \Omega)$ . Having in mind the lags  $\frac{\tau}{12}$  and  $\frac{\tau}{6}$  ratio, the component appearing in  $S_{p_2}(t, \Omega)$  at central frequency  $\Omega_0$  appears at

frequency  $-2\Omega_0$  in term  $S_{p_1}(t, \Omega)$ , having a twice larger bandwidth than the component in  $S_{p_2}(t, \Omega)$ . Consequently, the resulting frequency region for  $S_{n_1}(t, \Omega)$  and  $S_{p_1}(t, \Omega)$  is  $[-2\Omega_0 - \Omega_L, -2\Omega_0 + \Omega_L]$ .

The expression (14) now can be calculated as

$$\begin{aligned} R_{x^2}(t, \Omega) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{p_1}(t, \varpi) S_{n_1}(t, \Omega - \varpi) d\varpi \quad (19) \\ &= \frac{1}{4\pi} \int_{-2\Omega_L}^{2\Omega_L} S_{p_1}(t, \frac{\Omega}{2} + \frac{\varpi}{2}) S_{n_1}(t, \frac{\Omega}{2} - \frac{\varpi}{2}) d\varpi. \end{aligned}$$

The frequency region of interest for  $R_{x^2}(t, \Omega)$  becomes  $[-4\Omega_0 - 2\Omega_L, -4\Omega_0 + 2\Omega_L]$ .

Combining (18) and (19) with (11) the resulting cross-terms free representation can be calculated as follows

$$\begin{aligned} PD(t, \Omega) &= R_{x^2}(t, \Omega) *_{\Omega} R_{x^{16}}(t, \Omega) \quad (20) \\ &= \frac{1}{4\pi} \int_{-8\Omega_L}^{8\Omega_L} R_{x^2}(t, -\frac{\Omega}{4} - \frac{\varpi}{4}) R_{x^{16}}(t, \Omega - \varpi) d\varpi. \end{aligned}$$

### B. Numerical implementation

Let us observe, for a fixed instant  $t$  the samples corresponding to  $x(t - \frac{\tau}{6})$ ,  $x^*(t + \frac{\tau}{6})$ ,  $x^*(t - \frac{\tau}{12})$  and  $x(t + \frac{\tau}{12})$ , obtained by the discretization over  $\tau$ . The use of unit symmetric window  $w(n)$  of length  $N$  is inherently assumed. The procedure for numerical calculation of (9), assuming fixed point  $t$  follows:

**Step 1:** Calculate the set of signals  $x_{n_1}(n)$ ,  $x_{p_1}(n)$ ,  $x_{n_2}(n)$  and  $x_{p_2}(n)$  by sampling  $x(t - \frac{\tau}{6})$ ,  $x^*(t + \frac{\tau}{6})$ ,  $x^*(t - \frac{\tau}{12})$  i  $x(t + \frac{\tau}{12})$  over  $\tau$ , for fixed instant  $t$ . Calculate discrete Fourier transforms:  $S_{n_1}(t, k) = DFT[x_{n_1}(n)]$ ,  $S_{p_1}(t, k) = DFT[x_{p_1}(n)]$ ,  $S_{n_2}(t, k) = DFT[x_{n_2}(n)]$  and  $S_{p_2}(t, k) = DFT[x_{p_2}(n)]$ , for  $-N/2 \leq k \leq N/2 - 1$ .

**Step 2:** Calculate  $W_x(t, k) = DFT[x_{n_2}(n)x_{p_2}(n)]$  as convolution of the form

$$W_x(t, k) = \sum_p S_{p_2}(t, p) S_{n_2}(t, k - p), \quad (21)$$

We assume that  $S_{n_1}(t, k)$  and  $S_{p_1}(t, k)$  are localized in discrete frequency domain, i.e. that the component centered at frequency  $k_0$  is spread over region  $[k_0 - L, k_0 + L]$ . Under the assumption of symmetric window, this component is localized in the same region for both considered terms. This means that for each  $k$  in (21), region  $[k - L, k + L]$  is considered. The limits for  $p$  in (21) are obtained eliminating  $k_0$  from the system of inequalities  $k_0 - L \leq p \leq k_0 + L$  and  $k_0 - L \leq k - p \leq k_0 + L$ :

$$k/2 - L \leq p \leq k/2 + L. \quad (22)$$

Component being centered at  $k_0$  in  $S_{n_1}(t, k)$  and  $S_{p_1}(t, k)$  is centered at  $2k_0$  in resulting  $W_x(t, k)$ , spreading over region  $[2k_0 - 2L, 2k_0 + 2L]$ . The number of frequency points in  $W_x(t, k)$  is  $2N - 1$ .

**Step 3:** Following the previous analysis calculate:

$$R_{x^4}(t, k) = \sum_p W_x(t, p) W_x(t, k - p), \quad (23)$$

$$R_{x^8}(t, k) = \sum_p R_{x^4}(t, p) R_{x^4}(t, k - p), \quad (24)$$

$$R_{x^{16}}(t, k) = \sum_p R_{x^8}(t, p) R_{x^8}(t, k - p). \quad (25)$$

According to the analysis in Step 2, the limits for  $p$  in (23) are:  $k/2 - 2L \leq p \leq k/2 + 2L$ . The signal component in term  $R_{x^4}(t, k)$  corresponding to the component at  $k_0$  in  $S_{n_1}(t, k)$  and  $S_{p_1}(t, k)$ , is spread over region  $[4k_0 - 4L, 4k_0 + 4L]$ . Convolution  $R_{x^4}(t, k)$  is consisted of  $4N - 3$  frequency samples. Similarly, this component is spread over region  $[8k_0 - 8L, 8k_0 + 8L]$  in term  $R_{x^8}(t, k)$ , whereas the limits for  $p$  in the calculation of (24) are given with  $k/2 - 4L \leq p \leq k/2 + 4L$ . The resulting convolution  $R_{x^8}(t, k)$  is consisted of  $8N - 7$  samples. For convolution  $R_{x^{16}}(t, k) = DFT[x_{n_2}^8(n)x_{p_2}^8(n)]$  the limits for  $p$  read  $k/2 - 8L \leq p \leq k/2 + 8L$ . The analyzed component is placed in interval  $[16k_0 - 16L, 16k_0 + 16L]$ , whereas the number of frequency samples is  $16N - 14$ .

**Step 4:** Let us calculate the next convolution

$$R_{x^2}(t, k) = \sum_p S_{n_1}(t, p) S_{p_1}(t, k - p). \quad (26)$$

Following the continuous-time analysis, the component appearing in  $S_{n_1}(t, k)$  appears in the same region and central frequency in  $S_{p_1}(t, k)$ . The component located in terms  $S_{n_2}(t, k)$ , and  $S_{p_2}(t, k)$  at frequency  $k_0$  is spread over region  $[-2k_0 - 4L, -2k_0 + 4L]$  in terms  $S_{n_2}(t, k)$  and  $S_{p_2}(t, k)$ . Convolution (26) is calculated with following  $p$  limits:  $k/2 - 4L \leq p \leq k/2 + 4L$ . The new region of interest is  $[-4k_0 - 8L, -4k_0 + 8L]$ , and the resulting number of points is  $2N - 1$ .

**Step 5:** The resulting TFR is finally obtained as

$$PD(t, k) = \sum_p R_{x^2}(t, p) R_{x^{16}}(t, k - p). \quad (27)$$

It is important to note that the terms order in convolution is crucial as obtained regions for  $R_{x^2}(t, p)$  and  $R_{x^{16}}(t, k - p)$  differ. Following the results presented in previous steps, the resulting component is spread over  $[16k_0 - 16L, 16k_0 + 16L]$ , whereas  $p$  is calculated within limits obtained eliminating the unknown  $k_0$  from inequalities  $-4k_0 - 8L \leq p \leq -4k_0 + 8L$  and  $16k_0 - 16L \leq k - p \leq 16k_0 + 16L$ :

$$-k/3 - \lceil 16L/3 \rceil \leq p \leq -k/3 + \lceil 16L/3 \rceil, \quad (28)$$

where  $\lceil \cdot \rceil$  denotes the rounding to the nearest greater integer. Previous algorithm is presented assuming that samples  $x_{n_1}(n)$ ,  $x_{p_1}(n)$ ,  $x_{n_2}(n)$  i  $x_{p_2}(n)$  are obtained by discretization over  $\tau$ . The calculation of discrete representation  $PD(n, k)$  based on  $x(n)$  assumes a discretization of  $x(t)$  over  $t$  following the sampling theorem. Samples not appearing on the discrete axis  $n$ , that correspond to continuous-time signals  $x(t - \frac{\tau}{6})$ ,  $x^*(t + \frac{\tau}{6})$ ,  $x^*(t - \frac{\tau}{12})$  and  $x(t + \frac{\tau}{12})$ , are obtained by interpolation based on zero-padding in the frequency domain [1].

## IV. NUMERICAL RESULTS

**Example 1:** An FM signal being defined as  $x(t) = \exp(j(20 \sin(6\pi t) + j9 \cos(8\pi t) + j5 \cos(10\pi t)))$  is considered, for  $-1 \text{ s} \leq t \leq 1 \text{ s}$  and sampled with period  $\Delta t = 0.002$ . Four TFRs are calculated:  $STFT(t, \Omega)$ , Pseudo- $WD(t, \Omega)$ , S-method  $SM(t, \Omega)$  with  $Ld = 10$  and  $PD(t, \Omega)$  with  $L = 3$ . All transforms are calculated using a Hanning window of length  $N = 256$  (0.512s). The results shown in Fig. 1 (a)-(d) respectively confirm that  $PD(t, \Omega)$  calculated according to the proposed realization preserves

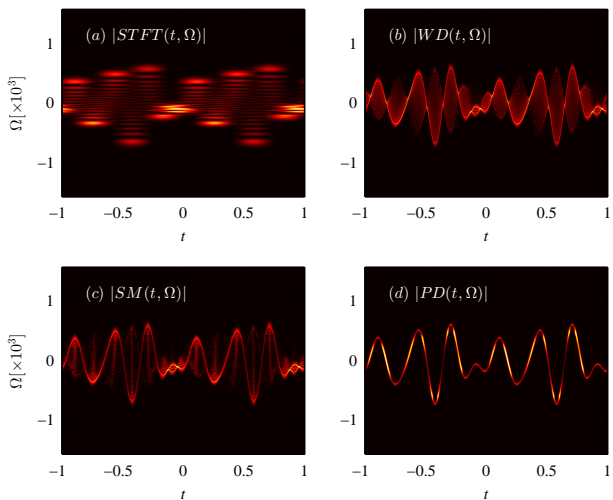


Fig. 1. Comparison of  $PD(t, \Omega)$  with various TF representations in the case of a fast varying FM signal.

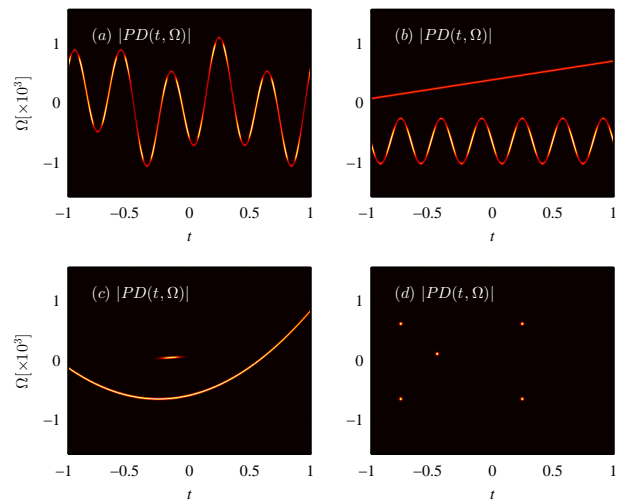


Fig. 2. Representation  $PD(t, \Omega)$  calculated for a mono-component (a) and various multi-component signals (b)-(d).

high concentration, while significantly reducing the inner interferences, when compared with other TFRs.

**Example 2:** Let us observe signals defined for  $-1 \text{ s} \leq t \leq 1 \text{ s}$ , sampled with period  $\Delta t = 0.002$ . Hanning window of width  $0.512 \text{ s}$  ( $N = 256$  samples) is assumed in  $PD(t, \Omega)$  calculations. Following signals are considered: (a) mono-component FM signal of the form  $x_1(t) = \exp(j50 \cos(5\pi t) + j50 \sin(2\pi t))$ ; (b) two-component FM signal consisted of one sinusoidally modulated and one LFM component, i.e.  $x_2(t) = \exp(-j(20 \sin(6\pi t) - 200\pi t)) + \exp(j50\pi t^2 + j100\pi t)$ ; (c) two-component signal consisted of third order polynomial phase signal (PPS) and LFM signal with Gaussian amplitude:  $x_3(t) = \exp(j100\pi t^3 - j200\pi t) + \exp(-10(t - 0.1)^2) \times \exp(j20\pi(t + 0.4)^2)$ ; and (d) five-component signal consisted of stationary signals having Gaussian amplitudes, defined as follows:  $x_4(t) = \sum_{i=1}^5 \exp(-500(t + a_i)^2) \exp(jb_i\pi(t + c_i))$ , with  $a_i = [0.2, -0.5, 0.5, 0.5, -0.5]$ ,  $b_i = [40, 200, 200, -200, 200]$  and  $c_i = [0.2, -0.5, 0.5, 0.5, 0.5]$ , for  $i = 1, \dots, 5$ .

Calculated representations  $PD(n, k)$  are shown in Fig. 2 (a)-(d), where  $L = 3$  is used. One can observe that a high concentration is obtained, and that inner interferences and cross-terms are significantly reduced. The cross-term reduction and a comparison of the proposed realization method with  $PD(n, k)$  calculated by definition is shown in Fig. 3 for the case of a two-component signal.

## V. CONCLUSION

The concept of windowed frequency convolutions, being originally proposed in the S-method theoretical framework, is applied in the realization of a higher order TFR, aiming cross-terms reduction in multi-component signals. The proposed realization method preserves high concentration of the considered representation. Presented results imply that the S-method based implementation approach is general and applicable in higher-order TF analysis. Noise influence on the considered TFR is the part of our further research.

## REFERENCES

- [1] Lj. Stanković, M. Daković, T. Thayaparan, *Time-Frequency Signal Analysis with Applications*, Artech house, 2013.

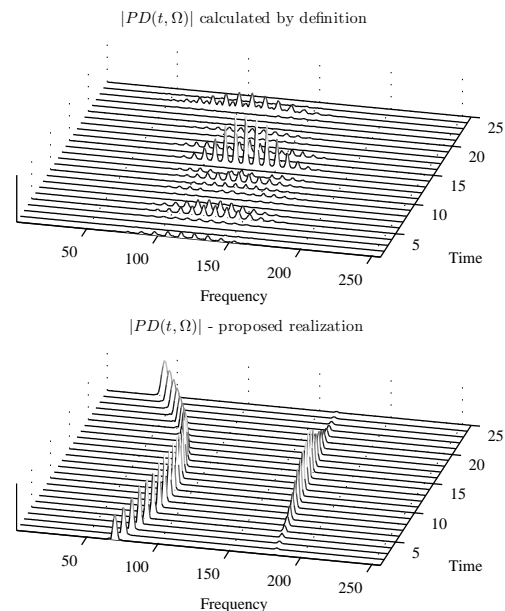


Fig. 3. Cross-terms reduction using the proposed implementation of  $PD(t, \Omega)$  in two-component signal (second subplot). Strong cross-terms completely mask the auto-terms in  $PD(t, \Omega)$  calculated by definition (first subplot).

- [2] B. Boashash, *Time-Frequency Signal Analysis and Processing – A Comprehensive Reference*, Elsevier Science, Oxford, 2003.
- [3] Lj. Stanković, "A method for time-frequency signal analysis," *IEEE Trans. on Signal Processing*, Vol-42, No.1, Jan.1994, pp.225-229.
- [4] Lj. Stanković, M. Daković, and T. Thayaparan, "A Real-Time Time-Frequency Based Instantaneous Frequency Estimator," *Signal Processing*, Volume 93, Issue 5, May 2013, pp.1392-1397
- [5] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal –Part 1: Fundamentals," *Proc. of the IEEE*, vol. 80, no.4, pp.519-538, April 1992.
- [6] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal: A tutorial review-Part 2: algorithms and applications," *Proc. IEEE*, vol. 80, April 1992.
- [7] M. Brajović, M. Daković, and Lj. Stanković, "Vremensko-frekvencijska transformacija višeg reda zasnovana na aproksimaciji prvog izvoda faze," *ETRAN 2012*, Zlatibor, July 2012
- [8] V. C. Chen, F. Li, S.-S. Ho, and H. Wechsler, "Micro-Doppler effect in radar: phenomenon, model, and simulation study," Vol. 42, No. 1, pp. 2-21, *IEEE Trans. on Aero. and El. Systems*, 2006.

# Izdvajanje obeležja radarskih ciljeva obradom spektrograma

Tamara D. Šević, Milenko S. Andrić, Boban P. Bondžulić, Dimitrije M. Bujaković, Dejan S. Ivković

**Sadržaj** — U radu je predstavljen jedan način za izdvajanje obeležja radi klasifikacije realnih signala sa izlaza izviđačkih radara. Obeležja su dobijena obradom spektrograma, kroz projekciju spektrograma na frekvencijsku osu i obradom vremenskih binova spektrograma primenom tehnika digitalne obrade slike. Izvršeno je poređenje pet tehnika za određivanje praga segmentacije spektrograma primenom mere maksimalne separabilnosti klasa. Dobijeni rezultati pokazuju da se normalizacijom spektrograma po maksimalnoj vrednosti i primenom Kapurove metode segmentacije obezbeđuje maksimalna separabilnost analiziranih klasa.

**Ključne reči** — Doplerova frekvencija, izviđački radar, klasifikacija, segmentacija slike, spektrogram.

## I. UVOD

JEDAN od glavnih zadataka u vojnim operacijama je obrada prikupljenih podataka i izdvajanje informacija sa ciljem ostvarivanja misije u nekom periodu vremena. Kompleksnost bojišta zahteva korišćenje senzora koji simultano rade u različitim delovima elektromagnetnog spektra (radar, televizijske i termalne kamere, seizmički senzori, akustični senzori, ...). Korišćenje ovih senzora ima svojih prednosti, ali i nedostataka. Na primer, radar može detektovati ciljeve na većim udaljenostima u poređenju sa optoelektronskim sensorima, a takođe može detektovati ciljeve koji se nalaze razdvojeni od senzora nekom preprekom, kao što je na primer zid. Osim toga, radar može raditi u različitim atmosferskim uslovima (danju, noću, po kiši, snegu i magli). Takođe, radar ima manju rezoluciju u poređenju sa optoelektronskim sensorima i zbog emisije energije se lako detektuje i ometa.

Tamara D. Šević (autor za kontakte), Vojnotehnički institut, Ministarstvo odbrane Republike Srbije, Ratka Resanovića 1, 11000 Beograd, Republika Srbija (telefon: 381-69-27-26-590, e-mail: [tamara.sevic23@gmail.com](mailto:tamara.sevic23@gmail.com)).

Milenko S. Andrić, Vojna akademija, Univerzitet odbrane u Beogradu, Generala Pavla Jurišića Šturma 33, 11000 Beograd, Republika Srbija (telefon: 381-64-11-60-093; e-mail: [andricsmilenko@gmail.com](mailto:andricsmilenko@gmail.com)).

Boban P. Bondžulić, Vojna akademija, Univerzitet odbrane u Beogradu, Generala Pavla Jurišića Šturma 33, 11000 Beograd, Republika Srbija (telefon: 381-64-143-82-79; e-mail: [bondzulici@yahoo.com](mailto:bondzulici@yahoo.com)).

Dimitrije M. Bujaković, Vojna akademija, Univerzitet odbrane u Beogradu, Generala Pavla Jurišića Šturma 33, 11000 Beograd, Republika Srbija (telefon: 381-11-3603-434; e-mail: [dimitrije.bujakovic@va.mod.gov.rs](mailto:dimitrije.bujakovic@va.mod.gov.rs)).

Dejan S. Ivković, Vojnotehnički institut, Ministarstvo odbrane Republike Srbije, Ratka Resanovića 1, 11000 Beograd, Republika Srbija (telefon: 381-64-80-42-060, e-mail: [dejan.ivkovic@vti.vs.rs](mailto:dejan.ivkovic@vti.vs.rs)).

Ključni zadaci korišćenja izviđačkih radara su detekcija i klasifikacija pokretnih ciljeva. U tipičnim izviđačkim radarima detekcija je potpuno automatizovana, dok su klasifikacija i prepoznavanje ponašanja ciljeva i dalje zasnovani na iskustvu operatera. Jedan od najčešće korišćenih izviđačkih radara je impulsno-Doplerovski radar i to zbog njegove jednostavne konstrukcije. Klasifikacija realnih signala kod radara koji rade na ovom principu je zasnovana na analizi jedinstvenog Doplerovog potpisa cilja.

Klasifikacija signala impulsno-Doplerovog radara predstavlja izazov za istraživače. Najčešće korišćena metoda za vremensko-frekvencijsku analizu signala je spektrogram. Spektrogram diskretnog signala dobija se kvadriranjem koeficijenata njegove vremenski zavisne Furijeove transformacije (Short Time Fourier Transform, STFT) [1]. Autori su u ovom radu pokazali da je spektrogram efikasna tehnika za analizu Doplerovog potpisa kretanja čoveka. Istraživanje [2] pokazuje da različite klase ciljeva imaju različitu Doplerovu frekvenciju i da pokretni delovi ciljeva prouzrokuju dodatne modulacije oko centralne Doplerove frekvencije. Ove dodatne modulacije se nazivaju mikro-Doplerova signatura cilja i mogu se koristiti za klasifikaciju radarskih signala, [3]. U [4] je predloženo korišćenje šest obeležja izdvojenih korišćenjem spektrograma za klasifikaciju različitih pokreta jedne osobe (hoda, trči, hoda dok drži štap, puzi, boksuje dok se pomera napred, boksuje dok stoji u mestu, mirno sedi). Širina projekcije spektrograma na frekvencijsku osu je korišćena u [5] kao vektor obeležja za klasifikaciju radarskih ciljeva (osoba, grupa osoba i vozilo). Autori su u [6] koristili višeprozorski S-metod za izdvajanje obeležja kako bi odredili da li osobe nose predmete u jednoj ili obe ruke. Na osnovu Doplerove frekvencije moguće je odrediti brzinu cilja [7].

U ovom radu razmatrana su obeležja dobijena obradom spektrograma. Za analizu spektrograma korišćene su metode za segmentaciju slike zasnovane na histogramu nivoa sivog [8]–[11] i Ojlerovom (engl. *Euler*) broju [12]. Za određivanje najbolje metode za segmentaciju slike korišćena je kriterijumska funkcija koja je opisana u [13].

Ostatak rada je organizovan na sledeći način: drugo poglavlje opisuje korišćenu bazu realnih radarskih signala. U trećem poglavlju je opisano izdvajanje obeležja obradom spektrograma, dok je u četvrtom poglavlju analiziran uticaj metoda za segmentaciju slike na izdvajanje obeležja. Zaključak i smernice daljeg rada se mogu naći u poslednjem, petom poglavlju.

## II. OPIS BAZE RADARSKIH CILJEVA

U radu su analizirani realni radarski signali prikupljeni izviđačkim radarom kratkog dometa. Ovaj radar je impulsno-Doplerovski, koherentan i radi u Ku-opsegu. Kada se radio frekventni signal reflektuje od cilja koji se kreće ka radaru ili od radara, eho signal reflektovan od različitih komponenti cilja će imati Doplerov pomak koji je proporcionalan brzini te komponente.

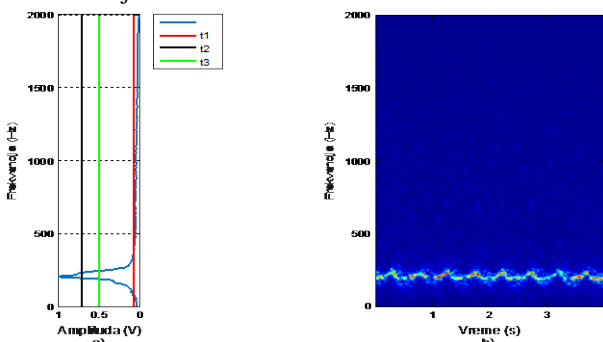
Karakteristike korišćenog radara su: centralna frekvencija – 16.8 GHz, srednja snaga – 5 mW, trajanje impulsa – 14.63  $\mu$ s, frekvencija ponavljanja impulsa 34.18 kHz, rezolucija po daljini – 150 m, rezolucija po elevaciji – 7.5° i rezolucija po azimutu – 5°. Korišćena je monostatička radarska konfiguracija sa paraboličnom antenom (dobitka 32 dB  $\pm$  2 dB) [1].

Tokom procesa prikupljanja podataka, cilj je detektovan i praćen automatski. Rastojanje između radara i cilja bilo je između 100 m i 1000 m, uz obezbeđenu optičku vidljivost između radara i cilja. Kretanje ciljeva je bilo unapred definisano, pri čemu su se ciljevi kretali po asfaltnom putu i u prirodnom ambijentu gde je preovladavala niska vegetacija.

Amplituda primljenog eho signala je u opsegu  $\pm$  1 V, pri čemu je učestanost odabiranja 4 kHz. Napravljena je baza sirovih realnih Doplerovih signala u više različitih situacija, pri čemu je snimljeno najmanje 20 s svakog scenarija. Baza je javno dostupna [14].

## III. IZDVAJANJE OBELEŽJA OBRADOM SPEKTROGRAMA

Centralna Doplerova frekvencija i širina spektra oko centralne Doplerove frekvencije su korišćeni za klasifikaciju radarskih ciljeva u [15]. Ova obeležja su određena iz projekcije spektrograma na frekvencijsku osu, kao što je ilustrovano na Sl. 1. Centralna Doplerova frekvencija se može odrediti i iz Furijeove transformacije analizirane sekvence, dok projekcija spektrograma na frekvencijsku osu predstavlja estimaciju modula ove transformacije.



Sl. 1. a) projekcija spektrograma na frekvencijsku osu, b) spektrogram eho signala osobe koja hoda.

Centralnoj Doplerovoj frekvenciji odgovara maksimalna vrednost projekcije spektrograma, dok je za određivanje širine potrebno uvesti prag. Na Sl. 1 uzete su tri vrednosti praga: prag t1 odgovara srednjoj vrednosti projekcije spektrograma, prag t2=0.707 odgovara opsegu od 3 dB i prag t3=0.5 odgovara opsegu od 6 dB. Uočava se da širina spektra oko centralne Doplerove frekvencije značajno zavisi od vrednosti praga.

Na Sl. 2 prikazan je spektrogram eho signala vozila koje

usporava. Na osnovu projekcije spektrograma na frekvencijsku osu jasno se uočava da je određivanje širine spektra primenom praga nepouzđano. Zbog toga je u ovom radu širina spektra oko centralne Doplerove frekvencije određivana za svaki vremenski bin spektrograma.

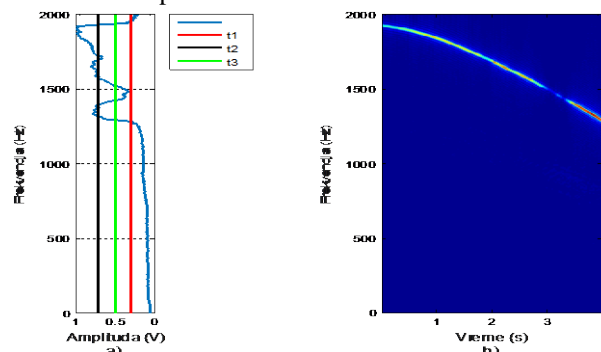
Za određivanje širine spektra oko centralne Doplerove frekvencije potrebno je sprovesti sledeće korake:

1. normalizacija spektrograma (prema maksimalnoj vrednosti ili prema maksimumu svakog vremenskog bina),
2. segmentacija (binarizacija) spektrograma primenom praga određenog prema metodama [8]–[12], i
3. određivanje donje i gornje granične frekvencije u svakom vremenskom binu (njihova razlika predstavlja širinu spektra  $i$ -tog vremenskog bina).

Srednja vrednost dobijenih širina predstavlja širinu spektra ( $W$ ) oko centralne Doplerove frekvencije za dati spektrogram i računa se kao:

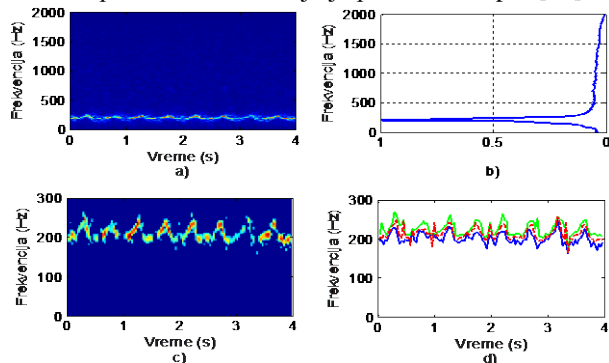
$$W = \frac{1}{M} \sum_{i=1}^M (gg(i) - dg(i)) = \frac{1}{M} \sum_{i=1}^M shir(i), \quad (1)$$

gde  $M$  predstavlja broj vremenskih binova,  $gg(i)$  i  $dg(i)$  predstavljaju gornju i donju graničnu frekvenciju  $i$ -tog vremenskog bina, dok  $shir$  predstavlja njihovu razliku. Na ovaj način očekuje se da će vozilo imati najmanju širinu spektra oko centralne Doplerove učestanosti, bez obzira da li ubrzava ili usporava.



Sl. 2. a) projekcija spektrograma na frekvencijsku osu, b) spektrogram eho signala vozila u pokretu.

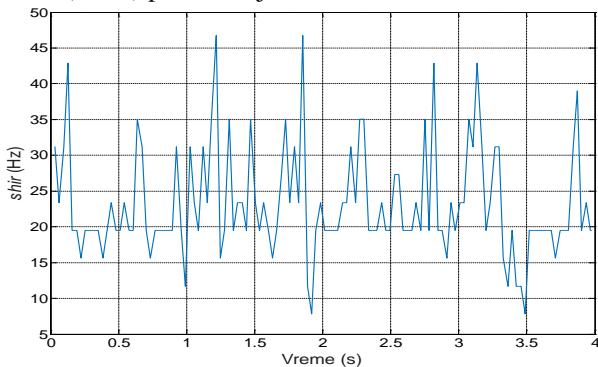
Na Sl. 3 prikazani su rezultati obrade spektrograma eho signala osobe koja hoda. Normalizacija spektrograma izvršena je prema maksimalnoj vrednosti, a za segmentaciju spektrograma koristi se vrednost praga određena pomoću metode koju je predložio Kapur [10].



Sl. 3. a) spektrogram radarskog signala, b) projekcija spektrograma na frekvencijsku osu, c) spektrogram signala nakon primenjenog praga, d) izdvojene frekvencije u svakom vremenskom binu.



Sa Sl. 3d) se vidi da gornja i donja granična frekvencija prate promenu centralne Doplerove frekvencije. Njihova razlika (širina) prikazana je na Sl. 4.



Sl. 4. Promena širine spektra oko centralne Doplerove frekvencije za primer sa Sl. 3.

Sa Sl. 4 se uočava da je širina spektra oko centralne Doplerove frekvencije vremenski promenljiva. Zbog toga je pored srednje vrednosti širine spektra oko centralne Doplerove frekvencije, kao obeležje analizirana i standardna devijacija širine:

$$Std = \sqrt{\frac{1}{M} \sum_{i=1}^M (shir(i) - W)^2}. \quad (2)$$

#### IV. UTICAJ IZBORA METODE ZA SEGMENTACIJU SLIKE NA IZDVAJANJE OBELEŽJA

Za određivanje najbolje metode za segmentaciju spektrograma koristi se kriterijum koji obezbeđuje maksimalnu udaljenost između klasa i minimalnu udaljenost oblika unutar klase [13], [16]. Svaka klasa je opisana na osnovu vektora matematičkog očekivanja  $\mathbf{M}_i$  i kovarijacione matrice  $\Sigma_i$ . Matrica unutarklasnog rasejanja,  $\mathbf{S}_w$ , se definiše kao:

$$\mathbf{S}_w = \sum_{i=1}^L P_i \Sigma_i, \quad (3)$$

a matrica međuklasnog rasipanja,  $\mathbf{S}_B$ , kao:

$$\mathbf{S}_B = \sum_{i=1}^L P_i (\mathbf{M}_i - \mathbf{M}_0)(\mathbf{M}_i - \mathbf{M}_0)^T, \quad (4)$$

gde  $L$  predstavlja broj klasa,  $P_i$  je apriorna verovatnoća pojave klase  $i$  i  $\mathbf{M}_0$  je vektor srednje vrednosti matematičkog očekivanja svih klasa:

$$\mathbf{M}_0 = \sum_{i=1}^L P_i \mathbf{M}_i. \quad (5)$$

U [13] i [16] definisana je kriterijumska funkcija koja obezbeđuje maksimalnu udaljenost između klasa i minimalnu udaljenost obeležja unutar klase:

$$J = tr(\mathbf{S}_w^{-1} \mathbf{S}_B) \quad (6)$$

pri čemu je  $tr(\mathbf{S}_w^{-1} \mathbf{S}_B)$  trag matrice  $\mathbf{S}_w^{-1} \mathbf{S}_B$ . Maksimalna vrednost kriterijumske funkcije obezbeđuje maksimalnu separabilnost razmatranih klasa.

U Tabeli 1 date su vrednosti kriterijumske funkcije  $J$ , za bazu radarskih odraza [14] u kojoj su eho signali svrstani u pet klasa: grupa osoba hoda, grupa osoba trči, osoba hoda,

osoba trči i vozilo. Ove vrednosti su određene ukoliko se pre segmentacije izvrši normalizacija po maksimalnoj vrednosti spektrograma, odnosno izvrši normalizacija po vremenskim binovima, a nakon toga sprovede binarizacija po pet metoda segmentacije. Sa  $J1$  označene su vrednosti kriterijumske funkcije  $J$  ukoliko se eho signal opiše vektorom obeležja čiji su elementi centralna Doplerova frekvencija ( $fd$ ) i širina spektra oko nje ( $W$ ). Sa  $J2$  označene su vrednosti kriterijumske funkcije ukoliko se umesto širine spektra eho signal opiše pomoću standardne devijacije  $Std$  (2).

Na osnovu rezultata iz Tabele 1 uočava se da je normalizacija prema maksimalnoj vrednosti spektrograma bolje rešenje od normalizacije po vremenskim binovima, osim u slučaju primene metode segmentacije koju je predložio Otsu [8]. Ova metoda segmentacije ima najmanje vrednosti kriterijumske funkcije  $J$ , odnosno daje najmanju separabilnost klasa. Najveću separabilnost klasa obezbeđuje binarizacija korišćenjem tehnike koju je predložio Kapur [10]. Poređenjem vrednosti kriterijumskih funkcija  $J1$  i  $J2$  uočava se da je veća separabilnost klasa ostvarena za izabrani vektor obeležja  $[fd W]^T$ .

Na Sl. 5 prikazane su projekcije eho signala radarskih odraza u dvodimenzionalnom prostoru obeležja (centralna Doplerova frekvencija – širina spektra oko nje). Izvršena je normalizacija po maksimalnoj vrednosti spektrograma, a segmentacija je urađena korišćenjem praga određenog po metodama koje su predložili Otsu [8] i Kapur [10]. Markerima ML označene su srednje vrednosti klasa.

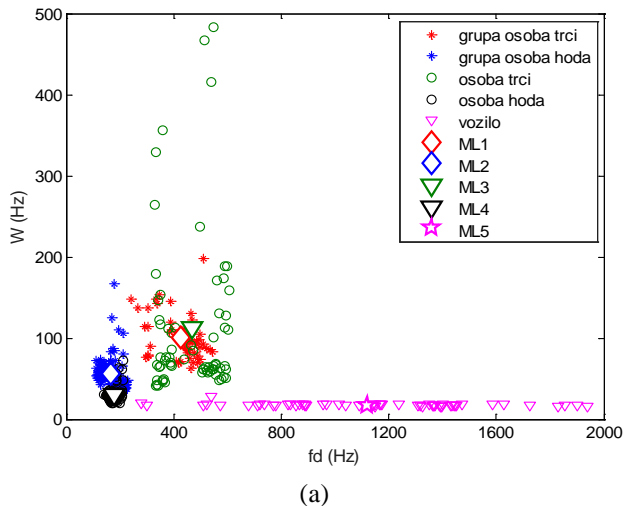
Sa Sl. 5 se uočava da je za oba primenjena metoda segmentacije klasa "vozilo" odvojena od ostalih razmatranih klasa. Vektori obeležja koji potiču od klasa "osoba hoda" i "grupa hoda" su razdvojeni od klasa "osoba trči" i "grupa trči". Bolju separabilnost klasa "osoba hoda" i "osoba trči", odnosno "grupa hoda" i "grupa trči", obezbeđuje primena metode segmentacije koju je predložio Kapur. Na ovaj način, potvrđeni su kvantitativni rezultati dati kroz kriterijumsku funkciju  $J$  u Tabeli 1.

#### V. ZAKLJUČAK

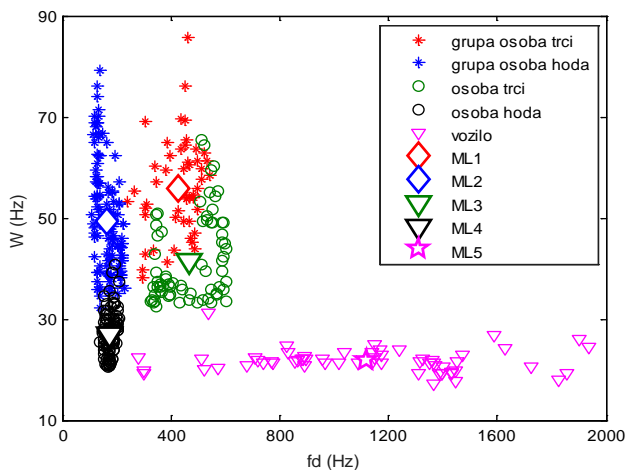
U radu je predložen jedan način za izdvajanje obeležja realnih signala sa izlaza izviđačkih radara. Nakon određivanja spektrograma potrebno je izvršiti njegovu normalizaciju prema maksimalnoj vrednosti, a nakon toga primenom praga sračunatog prema metodi koju je predložio Kapur potrebno je formirati binarnu sliku. Određivanjem gornje i donje granične frekvencije u svakom vremenskom binu binarne slike spektrograma dobija se širina oko centralne Doplerove frekvencije. Posmatranjem eho signala radarskih ciljeva u dvodimenzionalnom prostoru centralna Doplerova frekvencija i širina oko nje, ostvaruje se dobra separabilnost razmatranih klasa. U daljem istraživanju posebna pažnja biće posvećena klasifikaciji radarskih ciljeva na osnovu ovako izdvojenih obeležja, kao i analizi uticaja drugih vremensko-frekvencijskih distribucija (S-metod, Wigner-Ville-ova distribucija) na kvalitet izdvojenih obeležja. Takođe, u daljem radu analiziraće se i uticaj dužine trajanja signala na vrednost kriterijumske funkcije.

TABELA 1: VREDNOSTI KRITERIJUMSKE FUNKCIJE  $J$  ZA KORIŠĆENE METODE SEGMENTACIJE

metode segmentacije	normalizacija prema maksimalnoj vrednosti					sa normalizacijom po binovima				
	<i>Euler</i> [12]	<i>Kapur</i> [10]	<i>Kittler</i> [11]	<i>Tsai</i> [9]	<i>Otsu</i> [8]	<i>Euler</i> [12]	<i>Kapur</i> [10]	<i>Kittler</i> [11]	<i>Tsai</i> [9]	<i>Otsu</i> [8]
$J_1$	54.44	121.29	62.74	102.46	43.24	54.15	103.94	61.303	94.408	54.612
$J_2$	51.94	98.09	49.31	63.921	40.38	38.48	72.761	39.43	55.00	40.85



(a)



(b)

Sl. 5. Projekcije eho signala radarskih ciljeva u prostoru obeležja ukoliko se za segmentaciju koristi: (a) Otsuova metoda i (b) Kapurova metoda segmentacije.

#### LITERATURA

- [1] M. Andrić, D. Bujaković, B. Bondžulić, S. Simić i B. Zrnić, "Analysis of radar Doppler signature from human data," *Radioengineering*, vol. 23, no. 1, pp. 11-19, April 2014.
- [2] A. Aygar, "Doppler radar data processing and classification," Master thesis, Middle East Technical University, August 2008.
- [3] V. Chen, and H. Ling, *The Micro - Doppler Effect in Radar*, Boston, MA, USA: Artech House, 2011.
- [4] Y. Kim, and H. Ling, "Human Activity Classification Based on Micro-Doppler Signatures Using a Support Vector Machine," *IEEE Trans. Geoscience and Remote Sensing*, vol. 47, no. 5, pp. 1328-1337, May 2009.
- [5] J. Zabalza, C. Clemente, G. di Caterina, J. Ren, J. Soraghan, and S. Marshall, "Robust PCA Micro-Doppler Classification Using SVM on Embedded Systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 3, pp. 2304-2310, July 2014.
- [6] I. Orović, S. Stanković, and M. Amin, "A new approach for classification of human gait based on time-frequency feature representations," *Signal Processing*, vol. 91, no. 3, pp. 1448-1456, June 2011.
- [7] M. Hamaidia, "Veštačka inteligencija, primena pri detekciji upada i nadziranju granica," Master rad, Vojna akademija, Univerzitet odbrane u Beogradu, 2015.
- [8] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62-66, January 1979.
- [9] W.-H. Tsai, "Moment-preserving thresholding: A new approach," *Computer Vision, Graphics, and Image Processing*, vol. 29, no. 3, pp. 377-393, 1985.
- [10] J.N. Kapur, P.K. Sahoo, and A.K.C. Wong, "A new method for gray-level picture thresholding using the entropy of the histogram," *Computer Vision, Graphics, and Image Processing*, vol. 29, no. 3, pp. 273-285, 1985.
- [11] J. Kittler, and J. Illingworth, "Minimum error thresholding," *Pattern Recognition*, vol. 19, no. 1, pp. 41-47, 1986.
- [12] P.L. Rosin, and T. Ellis, "Image difference threshold strategies and shadow removal," in *Proc. British Machine Vision Conference*, Birmingham, UK, 1995, pp. 347-356.
- [13] D. Bujaković, M. Andrić, D. Mikluc, and B. Bondžulić, "Parameter order selection of autoregressive model for classification of ground surveillance radar targets," *Scientific Technical Review*, vol. 66, no. 2, pp. 3-9, 2016.
- [14] M. Andrić, B. Bondžulić, and B. Zrnić, *The database of radar echoes from various targets*. Available: <https://onedrive.live.com/?cid=3aaf3e18829259c0&id=3AAF3E18829259C0!105&ithint=folder.&authkey=!ANWj70DmFMULN-s>
- [15] M. Andrić, Ž. Đurović, and B. Zrnić, "Ground surveillance radar target classification based on fuzzy logic approach," in *Proc. IEEE International Conference on Computer as a Tool (EUROCON)*, Belgrade, Serbia, 2005, pp. 1390-1392.
- [16] K. Fukunaga, *Introduction to statistical pattern recognition*, 2nd edition, San Diego, CA, USA: Academic press, 1990.

#### ABSTRACT

This paper presents one approach to the feature extraction in order to classify real signals from the ground surveillance radar audio output. The features are obtained using spectrogram, through spectrogram projection on frequency axis and processing of spectrograms time bins using digital image processing techniques. Five techniques for spectrogram segmentation threshold are compared using classes' maximal separability measure. Obtained results show that using spectrogram normalization by the maximal value and Kapur segmentation method, the maximal separability of analyzed classes is provided.

Key words: Doppler frequency, ground surveillance radar, classification, image segmentation, spectrogram.

#### RADAR TARGETS FEATURE EXTRACTION USING SPECTROGRAM

Tamara Šević, Milenko Andrić, Boban Bondžulić, Dimitrije Bujaković, Dejan Ivković

# Performance of the Serbian Matrix Sentence Test in presence of the Babble Noise

Zoran Milivojević, Dijana Kostić, Zoran Veličković and Darko Brodić

**Abstract** - The first part of paper describes base of Serbian Matrix Sentence Test, which was formed from the words spoken on the Serbian language. This base was used to perform testing intelligibility of words and sentences with normal people and people with impaired hearing. Phonetic structure Serbian MST was compared with the phonetic structure of the novel Bridge on the Drina Ivo Andrić. Second part of paper presents the results of testing the intelligibility of words and sentences by using the Serbian MST at the group of subjects composed from students of College of Applied Technical Sciences of Niš. Intelligibility was tested in the presence of Babble noise, type N8 for SNR =  $-5 \div 5$  dB. The results are presented in tabular and graphic form.

**Key word** — Babble noise, Serbian MST, speech intelligibility.

## I. INTRODUCTION

Speech intelligibility is defined as the percentage of speech units which was correctly understood during a performance of speech intelligibility test [1]. Speech intelligibility depends on the quality of the spoken message, the level of degradation speech on the way of transferring and perceptive system of listener. In human communication speech intelligibility is very important. Everyday conversation usually have a influence of background noise, which have impact on speech intelligibility.

Speech audiometry have a different kind of tests for speech intelligibility such as: logatom tests (monosyllabic word) [2] and sentences test [3]. Also, tests speech intelligibility can be implemented for different conditions: a) superimposed noise like Babble noise, Gaussian noise (noise are expressed by parameter Signal to Noise Ration (SNR), and b) reverberation time (RT) [15]. Logatom intelligibility test is the test which used monosyllabic word without meaning type: CVC, VCV, CCVC (C=constant, V= vocal) [2].

Zoran Milivojević, College of Applied Technical Sciences of Niš, 20. Aleksandra Medvedeva, St, 18000 Niš, Serbia;  
(e-mail: zoran.milivojevic@vtsnis.edu.rs)

Dijana Kostić, College of Applied Technical Sciences of Niš, 20. Aleksandra Medvedeva, St, 18000 Niš, Serbia;  
(e-mail: koricanac@yahoo.com)

Zoran Veličković, College of Applied Technical Sciences of Niš, 20. Aleksandra Medvedeva, St, 18000 Niš, Serbia;  
(e-mail: zoran.velickovic@vtsnis.edu.rs)

Darko Brodić, University of Belgrade, Technical Faculty in Bor, 112 Vojske Jugoslavije St., 19210 Bor, Serbia;  
(e-mail: dbrodic@tf.bor.ac.rs)

There are a two type sentences test for standard measurement speech intelligibility: a) test which using everyday sentences (Plomp type sentences test) [4] and b) test which using sentences with fixed syntactical structure so called matrix sentences [3]. The first type of test first was performed by Kalkow et al [5] and Plomp&Mimpen [4]. Plomp type sentences test consist from short meaningful sentences. The measurements are normally performed in speech-shaped interfering noise. This sentences can be easily memorized and that can have influence on result. This test is provided for few language: German [6], French [7], English [8]. Second type of test is so called matrix tests. The base are included 50 word from which is composed sentences and it was proposed in [3], were is developed for Swedish language. Mutually combination of word, one word from each column, it is possible forming 100000 sentences. The similar test with the same structure already existing in: Polish [10], Russia [11], Spanish [12], US English [13], Swedish [3] language.

Basis on Hagerman's test, the authors are formed base of word (Serbian Matrix Sentences Test) which was tested for different reverberation time [9]. Testing speech intelligibility in the presence of Babble noise SNR = 5, 3, 1, 0, -1, -2, -3, -5 dB is the continuation of research. Testing was performed by choose word from certain type of word, from each column, and formed a sentences (computer is chosen according to a random law). The results can be: a) recognition the whole sentences or b) some word.

This paper is organized as follows. In the Section 2 is a description of the Serbian Matrix Test Sentences. In Section 3 was shown experimental results speech intelligibility using sentences from the Serbian MST in the condition of superimposed Babble noise for different value SNR= 5, 3, 1, 0, -1, -2, -3, -5 dB. Section 4 represents the conclusion.

## II. SERBIAN MATRIX SENTENCES TEST

First base of international matrix of tests was made by Hagerman's for Swedish language. Hagerman defined base of matrix test, which consist ten equal syntactic structure sentences [3]. Each sentence contains: name, verb, number, adjective and object. In this way the base consist 50 independent words and their mutual combinations can be form 100000 sentences.

### A. The base of speech material

The base of speech material was generated from matrix test for the Serbian language. The syntactic structure was respected as well as for the other international matrix test (name, verb, number, adjective and object). In Table 1 was shown sentences from Serbian MST and in Table 2 meaning on English language [9].

Recording a base matrix test was performed at the premises the College of Applied Technical Sciences of Nis, Serbia. The sentences was read a female age 22, with clear diction. Recording was performed for  $f_s=44.100$  kHz, and 16 bps. Speech signal was memorized on Hard drive in form *wav* files [9]. For the purpose of testing computer, by the random law, choose sentence with the same syntactic structure.

TABLE 1: MATRIX TEST SENTENCES FOR SERBIAN LANGUAGE.

Name	Verb	Number	Adjective	Object
Slaviša	čuva	osamsto	skupih	knjiga
Tamara	kupuje	sedamdeset	crvenih	jabuka
Gordana	ima	šesnaest	čistih	stolova
Tomislav	pravi	petnaest	velikih	kuća
Jasmina	želi	deset	jeftinih	rukavica
Miodrag	vidi	devet	različutih	brodova
Marina	daje	osam	žutih	stolica
Dragoslav	prodaje	sedam	novih	fotelja
Miroslav	briše	šest	starih	ormara
Danica	voli	pet	lepih	cvetova

TABLE 2: MATRIX TEST SENTENCES FOR SERBIAN LANGUAGE ON ENGLISH.

Name	Verb	Number	Adjective	Object
Slaviša	keep	eight	expensive	books
Tamara	buys	seventy	red	apple
Gordana	has	sixteen	clean	tables
Tomislav	makes	fifteen	big	houses
Jasmina	wants	ten	cheap	gloves
Miodrag	sees	nine	different	ships
Marina	gives	eight	yellow	chairs
Dragoslav	sales	seven	new	armchair
Miroslav	clears	six	old	closet
Danica	likes	five	nice	flowers

### B. The phonetically structure Serbian MST

Phonetically structure Serbian MST shown graphic ('o') on Fig. 1. In the purpose of comparison, like reference, was analysis phonetic occurrence at novel "The Bridge on Drina" for which is the author Ivo Andrić awarded with the Nobel prize for literature 1961. Novel have a 1329154 phonemes. Percent of occurrence was graphic shown in Fig. 1. ('□'). It can be seen great correlation distribution phonemes, so it can be considered that the Serbian MST have good phonetically structure. From Fig. 1. can be seen greater occurrence (b, g, d, đ, ž, z, dž, p, k, t, ć, š, s, č, f, h, c) [9].

### III. EXPERIMENTAL RESULTS AND ANALYSIS

This section described an experiment in which was tested intelligibility of Serbian MST base in ambient condi-

tions with the Babble noise. For the purpose of experiment were formed: base of speech material (from Serbian MST) and base of Babble noise. MOS test intelligibility were performed, and results are based on correct and incorrect answer: a) whole sentences, b) type of word.

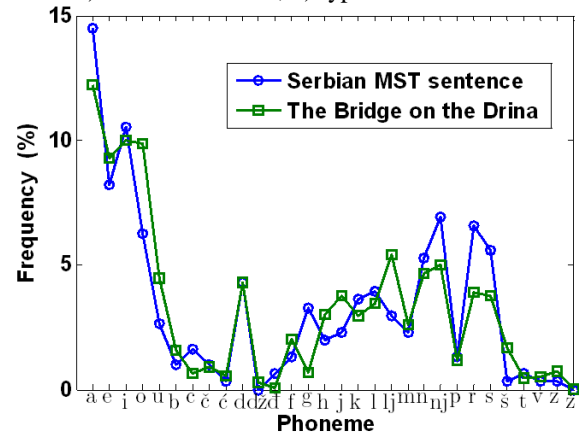


Fig. 1. Phonetic structure Serbian MST.

### A. Experiment

Experiment of intelligibility was performed in the three steps:

- Step 1: Forming base of sentences from Serbian MTS,
- Step 2: Forming base of Babble noise,
- Step 3: Forming test group for measuring intelligibility.

In the Fig. 2. is shown block diagram of experiment were is:  $x$  clean signal, **BN8** babble noise from 8 speakers,  $k$ -the coefficient for the determination SNR,  $y$  noisy signal.

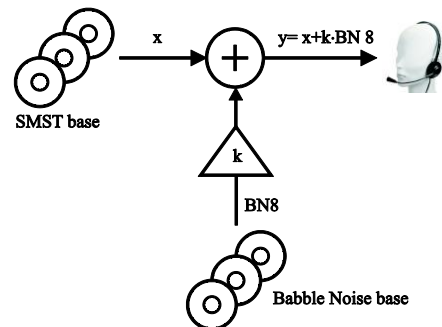


Fig. 2. Block diagram of the signal formed for MOS test intelligibility

The coefficient  $k$ , for the determination SNR can be calculated from equations 1:

$$SNR = 10 \log_{10} \frac{\sigma_x^2}{k \cdot \sigma_{BN8}^2} [dB] \quad (1),$$

where is  $\sigma_x$  variance of signal  $x$ ,  $\sigma_{BN8}$  variance of Babble noise.

The experiment was implemented by placed subjects in an acoustically isolated room where they were listened speech signal using headphone. Speech signal presented sentences generated by computer (random), with defined value SNR and Babble noise. Combinations word which formed sentences is unknown for examiner and examinee.

On the screen were display sentence which represent current speech signal. The screen is visible only for examiner. Examinee listened speech signal and loud spoke what he/she hear. In the case that the answer is incorrect it is considered that: a) incorrect whole sentence, b) incorrect some type of word. At the end of testing it was performed statistical analysis, and success the results intelligibility of sentences and some type of word is calculated in percent .

*B. Base of Babble noise*

Babble noise is characteristic by the number (N) of speaker, from which it produced. The base of Babble noise is generated by recording 8 speaker, 4 man and 4 female which were read text. The Babble noise represents the one of the best noise for masking speech. For the purpose of testing parameter Signal to Noise Ratio have a defined value, SNR={-5, -3, -2, -1, 0, 1, 3, 5} dB.

*C. Test group for measuring intelligibility*

The test group were formed from students of the College of Applied Technical Sciences of Niš, Serbia (14 male and 16 female test subjects). The age of the examinees ranges from 19 to 32 years, with mean age 21. Examinees were acoustically isolated in room and listened speech signal using headphone. After reproduction sentences with appropriate SNR and babble noise the examinee verbally repeated what he/she heard. For each examinee were write down accuracy whole sentences as well as individual types of word. After testing for the results of tests were performed statistical analysis.

*D. Results*

For word scoring, only word repeat correctly is counted, and when is sentences scoring only a whole sentence repeated correctly considered as a answer. Percent intelligibility type of word depending on the SNR (dB) is shown in Table 3 (N-Name, V-Verb, Nu-Number, A-Adjective, O-Object, W.S.-Whole Sentence) and graphically in Fig. 3. Percent intelligibility of word and sentences depending SNR (dB) is shown in Table 4 and graphically in Fig. 4b. Percent intelligibility of word and sentences depending SNR (dB) is shown in Table 3 and graphically in Fig. 4a.

TABLE 3 : INTELLIGIBILITY TYPE OF WORD

SNR [dB]	Intelligibility [%]					
	N	V	Nu	A	O	W.S.
5	40	50	90	90	63.33	20
3	60	60	60	70	50	20
1	50	56.67	60	50	33.33	10
0	53.33	36.67	53.33	46.67	30	6.67
-1	40	20	46.67	36.67	33.33	3.33
-2	40	10	36.67	20	16.67	6.67
-3	46.67	30	33.33	23.33	20	0
-5	26.67	6.67	10	3.33	6.67	0

TABLE 4 : INTELLIGIBILITY OF WORD

Intelligibility (%)	Type of word (%)				
	N	V	Nu	A	O
	71.61	55.48	78.06	68.39	50.97

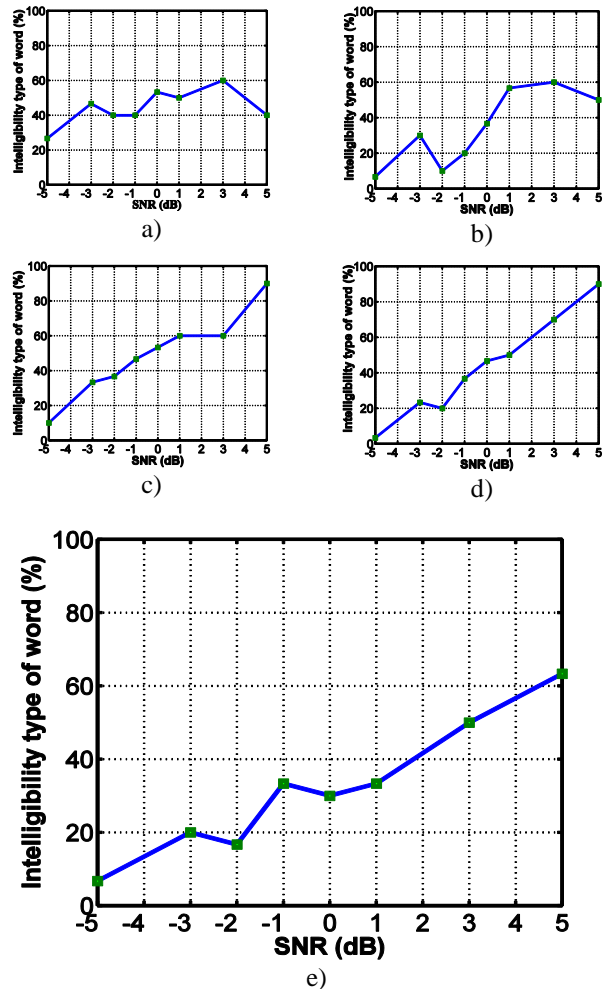


Fig. 3. Percentage of speech intelligibility in the dependence of SNR for: a) Name, b) Verb, c) Number, d) Adjective, e) Object.

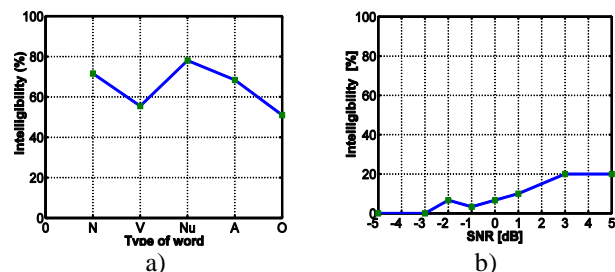


Fig.4. Percentage of speech intelligibility signal in the dependence of SNR for: a) type of word and b) whole sentence.

### E. Analysis of results

Based on the results shown in Table 3 and Fig. 3. it can be concluded that intelligibility of type of word:

a) 'Name' is best for 3 dB (60%) and the worst for -5 dB (26.67%),

b) 'Verb' is best for 3 dB (60%) and the worst for -5dB (6.67%),

c) 'Number' is best for 5 dB (90%) and the worst for -5dB (10%),

d) 'Adjective' is best for 5 dB (90%) and the worst for -5dB (3.33%),

e) 'Object' is best for 5 dB (63.33%) and the worst for -5dB (6.67%).

Based on the results shown in Table 4 and Fig.4a. it can be concluded that intelligibility of word:

a) 'Number' is the best 78.06% and

b) 'Object' the worst for 50.97%.

Based on the results shown in Table 3 and Fig. 4b. it can be concluded that intelligibility of sentence:

a) is best for 5dB and 3 dB (20%) and

b) the worst for -3dB and -5dB (0%).

Comparing results with results given in [16], comes to conclusion that is the best intelligibility on -5 dB is double less for Serbian language than for French language.

Comparing results with results given in [9], comes to conclusion that is the best speech intelligibility, whatever of different test conditions (reverberation time), show type of word 'Number' and 'Adjective'.

### IV. CONCLUSION

The aim of this paper is the evaluation performance (intelligibility of word and sentences) from base of Serbian MST, tested in presence of Babble noise with defined value SNR (-5÷5 dB). Results of testing shown that the percentage success of recognition: a) the type of word goes in range: Name (26.67÷60%), Verb (6.67÷60%), Number (10÷90%), Adjec-

tive (3.33÷90%), Object (6.67÷63.33%), b) the word (no matter which type of word) goes in range: 50.97÷78.06%) and c) the sentences goes in range from 0÷20%. The conclusion, based on the results, is that the best speech intelligibility are accomplished for type of word 'Number' and 'Adjective'.

### LITERATURA

- [1] BCJ. Moore, An instruction to the psychology of hearing. 5<sup>th</sup> Ed.London: *Academic Press*; 2003
- [2] D. Kostić, Z. Milivojević, V. Stojanović, "The Evaluation of Speech Intelligibility in the Orthodox Church on the Basis of MOS Test Intelligibility Logatom Type CCV", *ICEST 2016*, Ohrid, Makedonija, pp. 153-156, 2016.
- [3] B. Hagerman, Sentences for testing speech intelligibility in noise, *Scand Audio*, Vol. 11, pp. 79-87, 1982.
- [4] R. Plomp & A.M. Mimpfen, Improving the reliability of testing the speech reception threshold for sentences. *Audiology*, 18, 43-52, 1979.
- [5] D.N. Kalikow, K.N. Stevens & L.L. Elliott, Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *J Acoust Soc Am*, 61, 1337-51, 1977.
- [6] B. Kollmeier, M. Wesselkamp, Development and evaluation of a sentence test for objective and subjective speech intelligibility assessment. *J. Acoust. soc. Am.*, 102 (4), 1085-1099, 1997.
- [7] H. Luts et al., FIST: a French sentence test for speech intelligibility in noise. *International Journal of Audiology*, 47:373-374, 2008.
- [8] C. Cattermole, Comparison of Speech Recognition in Noise With Sound Field And Earphone Presentation. *MSc thesis, University of Southampton*, 2003
- [9] Z. Milivojević, D. Kostić, Z. Veličković, D. Brodić, "Serbian sentence matrix test for speech intelligibility measurement in different reverberation conditions". *UNITEH Gabrovo*, pp. - ,2016
- [10] E. Ozimek, A. Warzybok, D. Kutzner, Polish sentence matrix test for speech intelligibility measurement in noise, *Int. J. Audio*. 49:444-454, 2010.
- [11] M. Boboshko, A. Warzybok, M. Zokoll, N. Maltseva, RUMatrix test: construction, evaluation and clinical validation. *Otorhinolaryngologia Hungarica*. Vol. 59, N 2., 2013.
- [12] S. Hochmuth, T. Brand, M. Zokoll, F. Zenker, N. Wardenga, B. Kollmeier, A Spanish matrix sentence test for assessing speech reception thresholds in noise. *Int. J. Audiol*. 51(7) 536-544, 2012.
- [13] M. Alison, Evaluation of the American English Matrix Test with cochlear implant recipients, *Independent Studies and Capstones, Paper 709*, Program in Audiology and Communication Sciences, Washington University School of Medicine, 2015, [http://digitalcommons.wustl.edu/pacs\\_capstones/709](http://digitalcommons.wustl.edu/pacs_capstones/709).
- [14] GOST 25902-83. Available: <http://gostexpert.ru/data/files/25902-83/0491c8b1dc63b12979081245338f8f23.pdf>
- [15] Wagener, K.C., Factors influencing sentence intelligibility in noise. Phd thesis. bis-Verlag Oldenburg. ISBN = 3-8142-0897-8, 2004.
- [16] Jansen S., Luts H., Wagener K.C., Kollmeier B., Del Rio M., et al., Comparison of three types of French speech-in-noise tests: a multi-center study. *Int J Audiol*, 51 (3), 164-73, 2012.
- [17] Hagerman B. & Kinnefors C., Efficient adaptive methods for measuring speech reception threshold in quiet and in noise. *Scand Audiol*, 24, 71-77, 1995.
- [18] Kocinski J., Niemiec D., Time-compressed speech intelligibility in different reverberant conditions., *Applied Acoustics*, 113, 58-63, 2016.

# Realizacija mobilnog medicinskog podsetnika integrisanog sa medicinskim informacionim sistemom MEDIS.NET

Dragan S. Janković, Eleonora M. Milić, Aleksandar M. Milenković *Member, IEEE*

**Sadržaj** — U radu je prikazana mobilna aplikacija koja služi kao podsetnik za uzimanje terapije i praćenje parametara značajnih za procenu stanja pacijenta. Ova aplikacija je prilagođena medicinskom informacionom sistemu (MIS) MEDIS.NET sa kojim komunicira i iz kog preuzima potrebne podatke za pojedine pacijente (jednog ili više) kao što su propisani recepti sa medikamentima i terapijom odnosno u koji skladišti unete izmerene podatke značajne za praćenje stanja pacijenta. Aplikacija je razvijena za Android operativni sistem koji je široko zastupljen na našim prostorima. S druge strane medicinski informacioni sistem MEDIS.NET sa kojim aplikacija razmenjuje podatke je komercijalno rešenje razvijeno u Laboratoriji za medicinsku informatiku Elektronskog fakulteta u Nišu i koristi se u svakodnevnom radu u više od 20 zdravstvenih ustanova u Republici Srbiji. Realizovanu mobilnu aplikaciju uz izvesne izmene moguće je koristiti i u sprezi sa proizvoljnim medicinskim informacionim sistemom.

**Ključne reči** — Android, mobilna aplikacija, komunikacija, kolaboracija, medicinski informacioni sistem, MEDIS.NET, pacijent, recept.

## I. UVOD

DANAS, većina ljudi kao jedan od glavnih problema za narušeno zdravlje i lošu psiho-fizičku formu navodi nedostatak vremena. Obaveze, dinamičan i stresan način života vode ka tome da se ljudi nemarno ponašaju prema sebi i svom zdravlju. Veliki problem nastaje kod akutnih pacijenata, ali i kod pacijenata koji boluju od neke hronične bolesti i koji bi morali da uzimaju redovnu propisanu terapiju, a iz nekog razloga su sprečeni ili su zaboravili da popiju propisanu dozu medikamenta. Zbog neadekvatnog uzimanja terapije vreme oporavka pacijenta se značajno produžuje ili se postojeći problem ne ublažava (hronični pacijenti [1, 2]). Upotreba moderne tehnologije može pomoći ljudima u cilju podizanja psihofizičke forme i vođenja zdravijeg i kvalitetnijeg načina života. Mobilne tehnologije pronašle su veliku primenu u ublažavanju

postojećih problema i u medicini [3, 4]. Proučavanjem postojećih rešenja i potreba pacijenata (korisnika), u radu su prikazane funkcionalnosti realizovane mobilne aplikacije koja pomaže pacijentu da na najadekvatniji način primenjuje propisanu terapiju kao i da omogućiti praćenje pojedinih parametara bitnih za zdravlje pacijenta.

U toku mnogih lečenja izuzetno je važno da se terapija uzima na vreme u propisanim vremenskim intervalima sa tačno određenim dozama. Isto tako je često potrebno pratiti neke parametre stanja pacijenta koji su bitni za dalji tok lečenja i procenu da li trenutna terapija daje željene efekte ili ne tj. da li je treba menjati ili ne. Sa druge strane savremeni stil života, pogotovo kod određenih grupacija ljudi kao što su na primer poslovni ljudi, ili pak osobe starije životne dobi dovodi do toga da se zanemaruje trenutno stanje ili pak pravilno uzimanje postojeće terapije. Nekada je uzimanje terapije nestandardno tj. terapija se uzima po određenoj šemi koja se dinamički menja i teško je zapamtiti, pogotovo ako je period uzimanja terapije dugotrajan, kao na primer kod višemesečnog uzimanja antikoagulanata (npr. lek Farin) po šemi koja se u zavisnosti od stanja periodično menja. Veoma često je za praćenje stanja pacijenta potrebno pratiti vrednosti pojedinih parameta kao što su na primer krvni pritisak, puls, temperatura tela, nivo šećera u krvi i slično. Sve ovo navodi na potrebu kreiranja mobilne aplikacije koja bi pomogla u prevazilaženju opisanih problema uzimajući u obzir i činjenicu da su danas pametni telefoni široko rasprostranjeni i korišćeni od strane mnogih, posebno poslovnih ljudi.

Nakon kratkog uvoda dat je pregled postojećih rešenja a zatim i kratak osvrt na realizovanu mobilnu aplikaciju iza čega sledi opis integracije sa medicinskim informacionim sistemom MEDIS.NET i zaključak.

## II. PREGLED POSTOJEĆIH REŠENJA

Aplikacije koje predstavljaju medicinske podsetnike na osnovu svojih funkcionalnosti mogu se svrstati u tri kategorije:

- **SMR** (*eng. Simple Medication Reminders*): aplikacije koje nude osnovne funkcionalnosti pamćenja recepata, obaveštavanja korisnika, podešavanje alarma, biranje tipova obaveštenjaj pratećih zvučnih efekata, odlaganje alarma.
- **AMR** (*eng. Advanced Medication Reminders*): obezbeđuju dodatne opcije kao što su podrška prilikom

ˆ Dragan S. Janković, Elektronski fakultet u Nišu, Univerzitet u Nišu, Aleksandra Medvedeva 14, 18000 Niš, Srbija (telefon: +38118529105; e-mail: dragan.jankovic@elfak.ni.ac.rs).

Eleonora M. Milić, Elektronski fakultet u Nišu, Univerzitet u Nišu, Aleksandra Medvedeva 14, 18000 Niš, Srbija (telefon: +381641459104; e-mail: eleonoramilic@gmail.com).

Aleksandar M. Milenković, Elektronski fakultet u Nišu, Univerzitet u Nišu, Aleksandra Medvedeva 14, 18000 Niš, Srbija (telefon: +38118529443; e-mail: aleksandar.milenkovic@elfak.ni.ac.rs).

promene vremenske zone, zaštita od predoziranja, razni opisi, uputstva, itd.

- *MMA (eng. Medication Management Apps)*: u stvari su AMR aplikacije koje podržavaju više korisničkih naloga. Omogućavaju korisnicima da čuvaju bitne informacije, kontakte odabranih lekara, podsetnike za buduće preglede, itd.

Postoji puno aplikacija koje se baziraju na različitim konceptima i platformama. Korišćenje aplikacija u domenu zdravstvene zaštite i nege svakim danom sve više raste. Neke od njih zahtevaju i dodatni hardver kao što su senzori, RFID tagovi ili detektori pokreta. Mnogi pokušaji da se smanje administrativne greške u lečenju pacijenata bili su fokusirani na razvoj medicinskih dispanzera. Komercijalna rešenja su bili specijalizovani uređaji gde su korisnici morali da unose doze lekova koje treba da konzumiraju u određeno vreme. Naravno, takve operacije su često rezultirale greškama koje je izazivao ljudski faktor. Bilo je i takvih automatizovanih uređaja koji su bili instalirani u domovima starih i bolesnih osoba. U nastavku dati su primeri aplikacija koje su rezultati ranijih istraživanja.

*MyMediHealth* aplikacija [12] je podsetnik namenjen deci. Izvršava se na pametnim telefonima i omogućava konfigurisanje podsetnika za obaveštavanje korisnika o leku koji treba da popije po rasporedu.

*Wedjat* je predstavljao pokušaj integracije podrške u zdravstvenoj nezi i mobilnog računarstva. Aplikacija je zamišljena tako da sprečava pacijente u pravljenju grešaka prouzrokovanih administracijom. Glavne funkcionalnosti su: podešavanje na potreban dnevni unos medikamenata, uputstva vezana za dozvoljeni unos medikamenata i čuvanje spisa o konzumiranim lekovima.

*Med Minder* aplikacija je jednostavno i besplatno rešenje zamišljeno kao planer aplikacija, čija je glavna mana što je potrebno puno manuelnog podešavanja koje odnosi puno vremena.

*Medicine reminder pro* je besplatna aplikacija koja podržava do 15 podsetnika. Korisnik može da ih konfigurise kao repetitivne ili nerekativne alarme. Moguć je bilo koji vremenski interval između dva alarma, sa minimalnim intervalom od 1 sata. U planirano vreme aplikacija bi okidala notifikaciju sa alarmom, vibracijom i LED indikacijom.

*Medical Reminder and Healthcare* Android aplikacija ima za cilj da zvonjavom alarma podseća pacijente kada je vreme da uzmu svoju dozu leka. Pored alarma korisnik dobija notifikacije putem SMS- ili email-a.

*Medication Reminder System* je Android aplikacija takođe namenjena pacijentima. Ona automatski podešava vreme kada pacijent treba da konzumira narednu dozu leka. Taj podatak aplikacija preuzima iz recepta pacijenta i na taj način smanjuje mogućnost greške koju može da napravi pacijent ukoliko sam podešava vreme.

Većina aplikacija na tržištu nude podsetnike koji su bazirani na tajmerima. Njihova glavna svrha je obaveštavanje da je potrebno konzumirati određeni lek u određeno vreme. Podržavaju korišćenje puno

medikamenata i doza istovremeno, sem podsetnika koji su vezani za kontraceptivne pilule koji su limitirani na jednu dnevnu dozu. Neke od aplikacija dozvoljavaju samo jedno obaveštenje za određeni lek dnevno, a ne podržavaju vremenske interval kao što je podešavanje na nekoliko sati ili nekoliko dana. Mali broj aplikacija ima opciju odlaganja alarma za određeni lek, što je loše ako korisnik ne odreaguje odmah, jer može opet zaboraviti da uzme lek. Slaba strana postojećih sistema je i ta što najčešće korisnik mora sam da unese lekove koje konzumira što mu oduzima puno vremena. Ne postoji server koji čuva originalne recepte propisane od strane medicinske ustanove. Nema obaveštenja o tome da će pacijentu nestati lekova za određeni broj dana i da treba da poseti doktora radi daljih konsultacija. Zbog ručnog unošenja receptata postoji velika verovatnoća da se potkrade neka greška koja bi rezultirala velikim problemima u lečenju pacijenta. Pojedine aplikacije imaju podrazumevani zvučni efekat za alarm koji korisnik ne može da promeni. Veći potencijalni problem je taj što bi podsetnici obavestavali korisnika da treba da konzumira lek, ali bez predložene doze i periode propisane od strane lekara, što ponovo može izazvati veliku štetu u lečenju. Veliki nedostatak je to da većina aplikacija ne podržava više jezika, tako da je korisnički interfejs nerazumljiv velikom broju korisnika sa našeg podneblja.



Sl. 1. *PersonalMedicalReminder* – primer preuzetog propisanog leka iz elektronskog kartona pacijenta u MIS-u MEDIS.NET.

### III. KRATAK OSVRT NA REALIZOVANU MOBILNU APLIKACIJU - *PERSONALMEDICALREMINDER*

Kako je Android mobilna platforma najzastupljenija u Republici Srbiji najpre je razvijena Android mobilna aplikacija *PersonalPillsReminder* koja je integrisana sa postojećim MIS-om MEDIS.NET-om [5]. Cilj je bio da se mobilna aplikacija iskorsiti kao podsetnik pacijentu



(korisniku) kada je potrebno uzimati terapiju propisanu od strane lekara i zapamćene u elektronskom kartonu pacijenta (EKP) [6, 7] ali isto tako i za unos nekih karakterističnih parametara (krvni pritisak, puls, nivo šećera u krvi, telesna temperatura, težina itd.) u cilju detaljnijeg monitoringa pacijenta i adekvatnog praćenja dejstva propisane terapije. Najpre, lekar preko EKP-a propisuje odgovarajuću terapiju pacijentu. Pacijent upotrebom mobilne aplikacije preuzima sve propisane recepte. Nakon toga sam korisnik postavlja inicijalno vreme uzimanje prve doze. Nakon toga aplikacija sama kreira notifikacije za svaki medikament i to na osnovu preporuke kako je lekar propisao terapiju na receptu (Sl. 1.). Pod medikamentom ovde podrazumevamo najčešće neku formu leka (tableta, kapsula, karpula, itd.). Zatim se od korisnika zahteva i podešavanje broja tableta/kapsula u pakovanju koje je preuzeo u apoteci ukoliko se taj broj razlikuje od broja na propisanom receptu (broj tableta zavisi od samog proizvođača i najčešće lekari propisuju lekove sa generičkim nazivima lekova). Naravno, korisnik aplikacije ima mogućnost promene vremena kada se notifikacije pojavljuju. U aplikaciji se pamti tačno vreme kada je lek uzet, pod kojim okolnostima je lek uzet, koja je doza leka uzeta (propisana ili je bilo nekih odstupanja).

Sl. 2. Prozor za unos nekih izmerenih parametara.

Korisniku mobilne aplikacije je omogućeno da u svakom trenutku dobije statistički izveštaj za unete parametre u obliku tabele ali i grafikona za određeni izabrani vremenski period. Prilikom unosa vrednosti za neki odabrani parametar, korisnik ima mogućnost i unosa pod kojim okolnostima je vršeno merenje nekog parametra (pre/nakon uzimanja medikamenta, pre/posle jela, pre/nakon neke fizičke aktivnosti, pre/tokom/nakon važnog sastanka, pre spavanja, nakon buđenja, itd.) (Sl. 2.). Parametri koje je korisnik u obavezi da kontroliše, definišu se za svakog pacijenta ponaosob u EKP preko MIS-a MEDIS.NET.

Svedoci smo vremena gde sve više mladih i poslovnih ljudi ima problem sa povišenim krvnim pritiskom (Hipertenzijom) [8]. Hipertenzija (dijagnoza u MKB 10 šifarniku dijagnoza: I10 – Esencijalna (primarna) hipertenzija, I11 – Hipertenzivna bolest srca) je postala

bolest savremenog, užurbanog doba i pogađa na žalost sve više mlađu populaciju, koja živi, neadekvatnim, nekvalitetnim načinom života sa puno stresa i visoko postavljenim ciljevima [9, 10]. U cilju uspostavljanja adekvatne terapije, pacijenti koji boluju od hipertenzije imaju i mogućnost vođenja dnevnika krvnog pritiska u realizovanoj mobilnoj aplikaciji (Sl. 3.). Vođenjem dnevnika krvnog pritiska tokom uzimanja propisane terapije, korisnik dobija obaveštenja kada je potrebno da izmeri krvni pritisak. Svi sakupljeni podaci se periodično sinhronizuju sa EKP, tako da izabrani lekar u svakom trenutku može dobiti ažurne podatke o stanju pacijenta. Ukoliko dođe do nekih nepredviđenih okolnosti pacijent može kontaktirati, pomoću mobilne aplikacije, lekara jednostavnim pisanjem poruke. Izabrani ili dežurni lekar dobiće odgovarajuću notifikaciju u MIS-u MEDIS.NET.

Terapija: Ujutru: Bisoprolol 5mg, Magnezijum 375mg Preko dana: - Uveče: Amlodipin 2.5mg					
Datum	Vreme	Gornji	Donji	Puls/min	Komentar
01.02.2017.	Jutro	139	87	67	
	Veče	137	80	67	
02.02.2017.	Jutro	125	75	57	
	Veče	128	72	57	
03.02.2017.	Jutro	130	80	70	
	Veče	136	78	53	
04.02.2017.	Jutro	120	74	64	
	Veče	111	75	63	
05.02.2017.	Jutro	130	79	68	
	Veče	135	80	60	
06.02.2017.	Jutro	134	76	65	
	Veče	126	80	61	
07.02.2017.	Jutro	132	78	63	
	Veče	140	80	59	
Srednja vrednost 6 dana (bez prvog dana)		128.9	77.3	61.7	

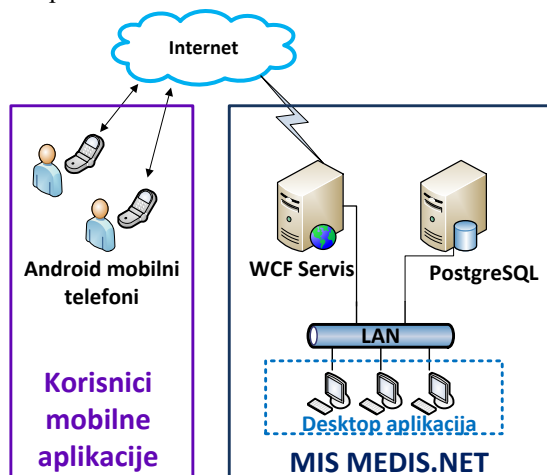
Sl. 3. Primer dnevnika krvnog pritiska.

Svi prikupljeni podaci, upotrebom mobilne aplikacije, dostupni su lekarima u EKP-u. Lekar, podatke može posmatrati na dva načina: u obliku tabela ali i u obliku grafikona.

#### IV. INTEGRACIJA SA MIS-OM MEDIS.NET

Mobilna aplikacija *PersonalMedicalReminder* je u tesnoj vezi sa medicinskim informacionim sistemom MEDIS.NET (Sl. 4.). MEDIS.NET je realizovan kao desktop aplikacija koja je razvijena upotrebom Microsoft-ovih tehnologija. Aplikacije se može izvršavati na Microsoft Windows OS-u ali i na Linux operativnom sistemu pod Mono frejmvorkom. Kao baza podataka koristi se PostgreSQL. Za pristup podacima korićen je Microsoft ORM alat EntityFramework. Na svakom desktop računaru potrebno je instalirati posebnu aplikaciju. Aplikacije sa klijentskih računara komuniciraju sa bazom podataka preko LAN-a. Kako bi bila moguća komunikacija između MIS-a MEDIS.NET i mobilne aplikacije razvijen je WCF servis [11] sa odgovarajućim metodama. Realizovani WCF Servis omogućava dvosmernu komunikaciju.

Propisani recept od strane lekara u elektronskom kartonu pacijenta može se preuzeti ukoliko korisnik poseduje korisnički nalog u MIS MEDIS.NET. Prilikom dolaska kod lekara potrebno je da pacijent dobije PIN kako bi mogao prvi put da pristupi svojim podacima koji se nalaze u elektronskom kartonu preko instalirane mobilne aplikacije. Nakon toga, pacijent može kreirati korisničko ime i lozinku koju će sledeći put koristiti prilikom prijavljivanja na sistem. Mobilna aplikacija je relizovana tako da se može koristiti i za veći broj članova jedne porodice. Sve što je potrebno jeste da se definišu članovi porodice u MIS-u MEDIS.NET.



Sl. 4. Integracija mobilne aplikacije *PersonalMedicalReminder* i MIS-a MEDIS.NET.

#### V. ZAKLJUČAK

U ovom radu prikazana je realizovana Android aplikacija *PersonalPillsReminder* koja je u tesnoj vezi sa MIS-om MEDIS.NET. Realizovanu aplikaciju je moguće koristiti i u sprezi sa nekim drugim MIS-om koji implementira odgovarajuće funkcionalnosti potrebne za rad *PersonalPillsReminder*-a. Realizovana aplikacija se trenutno nalazi u test fazi.

Sledeći koraci u daljem razvoju bi bili da se realizovana aplikacija nakon završenog testiranja distribuira što većem broju korisnika (pacijentima), da se realizuje mobilna aplikacija sa istim funkcionalnostima samo za iOS uređaje, da se implementira i deo funkcionalnosti koji bi se odnosio na zakazivanje pregleda. Dalji razvoj bi išao i u smeru povezivanja različitih senzora, pametnih satova, narukvica koje mogu prikupljati neometano vitalne parametre pacijenta i snimati iste na mobilnom uređaju. Poseban segment daljeg rada čini modul za analizu prikupljenih podataka i davanje odgovarajućih sugestija.

#### LITERATURA

- [1] G. Chiarini, P. Ray, S. Akter, C. Masella, A. Ganz, "mHealth Technologies for Chronic Diseases and Elders: A Systematic Review", IEEE Journal on Selected Areas in Communications. Vol. 31, Issue. 9, September 2013, DOI: 10.1109/JSAC.2013.SUP.0513001, ISSN: 0733-8716.
- [2] M. Vazquez, J. A. Garibaldi, J. I. Nieto, J. de D. Sanchez, "Model for Personalization of Mobile Health Systems for Monitoring Patients with Chronic Disease", IEEE Latin America Transactions, Vol. 14, Iss. 2, Feb. 2016, DOI: 10.1109/TLA.2016.7437247, ISSN: 1548-0992.

- [3] mHealth - New Horizons for Health Through Mobile Technologies, Global Observatory for eHealth series – Volume 3, World Health Organization, ISBN: 978 92 4 156425 0.
- [4] C. Villalonga, H. Pomares, O. Banos, "Analysis of the Innovation Outputs in mHealth for Patient Monitoring", Ubiquitous Computing and Communications and 2016 International Symposium on Cyberspace and Security (IUCC-CSS), IEEE, DOI: 10.1109/IUCC-CSS.2016.025, ISBN: 978-1-5090-5567-8.
- [5] D. Janković, A. Milenković, P. Rajković, T. Stanković, I. Marković, S. Cvetković, D. Vučković, S. Pešić, "Medicinski informacioni sistem MEDIS.NET", Elektronski fakultet Niš - Laboratorija za medicinsku informatiku, 2010., tehničko rešenje [http://www.elfak.ni.ac.rs/phptest/new/html/nauka/tehnicka\\_resenja/resenja/0259.html](http://www.elfak.ni.ac.rs/phptest/new/html/nauka/tehnicka_resenja/resenja/0259.html).
- [6] Goetz, G. D., Kuzel, A. J., Feng, L. B., DeShazo, J. P., & Love, "EHRs in primary care practices: benefits, challenges, and successful strategies". The American journal of managed care, 18(2), e48-54.
- [7] Holroyd-Leduc, J. M., et al, "The impact of the electronic medical record on structure, process, and outcomes within primary care: a systematic review of the evidence", Journal of the American Medical Informatics Association, 18(6), 732-737.
- [8] Arterijska hipertenzija, Nacionalni vodič za lekare u primarnoj zdravstvenoj zaštiti, Ministarstvo zdravlja Republike Srbije, Srpsko lekarsko društvo, ISBN: 86-85313-33-3.
- [9] Hypertension – the 'Silent Killer', Briefing Statement, Faculty of Public Health of the Royal Colleges of Physicians of the United Kingdom.
- [10] M. P. Peršić, S. V. Mikuličić, S. Rački, "Arterial hypertension", Medicina fluminensis 2010, Vol. 46, No. 4, p. 376-389.
- [11] W. Zhang, G. Cheng, "A Service-Oriented Distributed Framework-WCF", Web Information Systems and Mining, 2009, WISM 2009, International Conference on, DOI: 10.1109/WISM.2009.69, ISBN: 978-0-7695-3817-4.
- [12] MyMediHealth - <https://ww2.mc.vanderbilt.edu/mymedihealth/>

#### ABSTRACT

This paper presents a mobile application that serves as a reminder for taking the prescribed therapy and monitoring of parameters relevant to the patient's health condition. This application uses data from the electronic medical records of the patient's from the medical information system MEDIS.NET, which is used in medical institutions for primary health care. Our application communicates with MEDIS.NET and takes all the necessary information for individual patients (one or more), like prescribed recipes with medications and therapy. One more important feature is that application also stores the entered measured data (cardiac pressure, body temperature, etc.) relevant for the monitoring of the patients health condition. The application is developed for the Android operating system which is widely used in our region. On the other side there is medical information system MEDIS.NET, with which application exchanges data, which presents commercially solution developed at the Laboratory of Medical Informatics, Faculty of Electronic Engineering in Nis and it is used in everyday work in more than 20 health care institutions in the Republic of Serbia. Developed mobile application with certain modifications can also be used together with any other medical information system.

#### IMPLEMENTATION OF MOBILE MEDICAL APPOINTMENTS INTEGRATED WITH THE MEDICAL INFORMATION SYSTEM MEDIS.NET

Dragan S. Janković, Eleonora M. Milić, Aleksandar M. Milenković

# The Application of Simulated Intuition in Minimizing the Number of Moves in Guessing the Series of Imagined Objects

Aleksa S. Srdanov, Selena Ž. Vasić, Nada V. Ratković Kovačević, and Dragan M. Milovanović

**Abstract** - Combining logic and randomness could increase the efficiency of the search algorithm, compared to an algorithm based on either logic or randomness. The method is applied on the case of finding a pre-conceived quintuplet from a relatively large set of possible quintuplets. Additional increase in efficiency of the search algorithm is achieved with detecting and using special properties, a certain symmetry in pattern, thus providing the reduction of the number of available cases. In addition, it is shown that the partial overlapping of decomposed actions can significantly contribute to the efficiency of the search algorithm.

**Key words** - Algorithms and models, Artificial Intelligence, Classification, Games, Searching.

## I. INTRODUCTION

THE paper [1] considered the following problem and one example of its solution – the search algorithm that was applied for the game of guessing the five-digit numbers (quintuplets). A human and a computer play the guessing game: each player writes a sequence of five different decimal figures. Then, the players alternatively try to guess the exact sequence of digits of the opponent's quintuplet. In each trial, the player whose number is guessed, announces the number of correct digits and the number of correct positions that are found in that specific attempt. The player's reply has the form  $x:y$ , where  $x$  represents the number of digits the opponent had correctly guessed, and  $y$  represents the number of their correct positions. The goal is to win the game and a winner is the first player who attains a perfect match - the correct sequence of the opponent's quintuplet.

Introducing randomness in the algorithm have proven useful when used for searching or optimization or decision making, to name the few: simulated annealing [2], genetic

Aleksa S. Srdanov, Technical College of Vocational Studies, Požarevac, Nemanjina 2, 12000 Požarevac, Serbia (phone: +381-12-531-667; e-mail: [aleksa.srdanov@vts-pozarevac.edu.rs](mailto:aleksa.srdanov@vts-pozarevac.edu.rs)).

Selena Ž. Vasić, Metropolitan University, Faculty of Information Technology, Tadeuša Koščuška 63, 11000 Belgrade, Serbia, Technical College of Vocational Studies, Požarevac, Nemanjina 2, 12000 Požarevac, Serbia (phone: +381-62-401-858; e-mail: [selena.vasic@vts-pozarevac.edu.rs](mailto:selena.vasic@vts-pozarevac.edu.rs))

Nada V. Ratković Kovačević, Technical College of Vocational Studies, Požarevac, Nemanjina 2, 12000 Požarevac, Serbia (phones: +381-12-531-667; +381-63-8422-044; e-mail: [nada.ratkovic.kovacevic@vts-pozarevac.edu.rs](mailto:nada.ratkovic.kovacevic@vts-pozarevac.edu.rs)).

Dragan M. Milovanović, Technical College of Vocational Studies, Požarevac, Nemanjina 2, 12000 Požarevac, Serbia (phone: +381-12-531-667; e-mail: [dragan.milovanovic@vts-pozarevac.edu.rs](mailto:dragan.milovanovic@vts-pozarevac.edu.rs)).

algorithms [3] - [4], randomized algorithms [5].

The combination of randomness and logic here arise from the fact that a random quintuplet is formed from the set of digits, while the initial set of "candidates" is successively broken into a number of subsets containing fewer elements from which a random selection is further performed. The process ends when the selection is not random any more, but is uniquely determined, because all the decomposed subsets contain a single digit.

In this paper, the similar problem to the one in [1] is considered, with lesser constraint - that the digits in a quintuplet can be repeated. If digits cannot repeat in the initial trial, as was the case in [1], there could be 30240 different quintuplets. If the digits can be repeated, the total number of possible quintuplets is  $10^5 = 100000$ . As is shown in [1], it is possible to find the unknown quintuplet in a maximum of six attempts. With repetitions of digits, the number of moves necessary to find initial quintuplet is at most 8. Although there are more than three times more available/possible candidates than in [1], it is shown here that no more than eight moves suffices to reach the goal.

The fact that some figures may appear several times considerably complicates the search as compared to the approach presented in [1]. In this case, the exact order of five digits can be revealed only after finding both figures that appear as well as their frequency of occurrence in that specific quintuplet. The procedure given in [1], as well as the one presented here, are based on trial and error. In [6] the trial and error was formally modeled as a technique for problem solving where input is unknown.

Here we exploit the benefits of the special cases that assume a certain symmetry in a pattern which allows us to make distinction from the general procedure. It will be shown that using this approach significantly improves the efficiency of the search algorithm. Here, it is possible to distinguish the outcome when the opponent announces the reply 0:0 in the first two moves from the cases when both his responses are different and do not start with zero.

## II. ANALYSIS OF THE PROBLEM.

There are ten digits available to choose from: 0, 1, 2, ..., 9. We can make a random variation of these numbers and denote digits by the letters:  $a, b, c, d, e, f, g, h, i, j$ .

In [1] where a series of five figures were chosen without repetitions, the problem have been decomposed and the search performed using two functions: one to determine the digits and another to determine their positions.

Rearranging the digits to their proper positions was possible only after the figures have been determined. If the process of determining the numbers needed less than five trials, it was necessary to proceed to maximally the fifth move, to determine their order.

In this case, the problem can be decomposed into four separate functional units:

1. The one that will determine the number of digits used;
2. The other that will determine which figures are used;
3. The function that will determine how many times each of the figures is used; and
4. The function that will detect the positions of these figures in a sequence imagined.

Actions are listed in order of application. The advantage of such decomposition lies primarily in the fact that the number of possible candidates considerably decrease when we know the answer to some of the 4 problems identified. For example, the number of all five-digit sequences (quintuplets) with exactly three different decimal digits is  $\binom{10}{3} \cdot 3^5 - \binom{10}{2} \cdot 2^5 - 10 = 27710$ , where the number of those quintuplets having only two different digits, or a single digit, is subtracted from the total number of available sequences. Also, if we know what these three digits are, the number of sequences is reduced to only  $3^5 = 243$  quintuplets, which is less than one percent of a previous set.

On the other hand, while performing some of the four decomposed functions, whenever is possible the following procedure should be performed. In order to determine the quintuplets, some figures will be discovered. In subsequent process of finding the remaining digits, it is possible to determine the number of times the opponent have used already discovered digits. In each subsequent step it is very important never to use the same figure in the same place, which will significantly improve the efficiency of the last function. Exhibited partial overlapping in tasks of the functions described will significantly increase the overall efficiency of the search algorithm.

To find the number of digits that have been used in an unknown quintuplet, there have to be exactly two mandatory questions asked. (There are at least  $\binom{10}{2} = 45$  such questions, and one possible pair is **1. a, b, c, d, e** and **2. f, g, h, i, j.**) After the opponent replies, adding his results to the number of correctly guessed digits from both trials, gives us information on how many different digits the opponent had used, and additionally, how many from each individual quintuplet that had been offered. These two questions complete the task of the first function. It is important to note that all the possible results of applying the first function are always symmetric.

After applying the first function, in order to increase the efficiency of the second function it is recommended to consider each particular result obtained so far. Obviously, if the opponent has used a different number of digits in the

initial sequence, it will require a different number of moves to end the guessing process. There are six cases.

Case 1. If the sum of the digits that have been correctly guessed is either higher than five or equal to zero, in the first reply as well as in the second reply from the first function, one can object that either the opponent did not count properly or he did not follow the rules of the game.

Case 2. If the opponent have used just one digit in his initial quintuplet, then for example, he could give a reply 1:1 to the first guessing attempt, and for the second - his reply could be 0:0, or vice versa. In the first case, the player can make the third move with: *a, a, b, b, c*, in order to try to guess what that one figure might be. Depending on the reply, one could need not more than one more move to determine that digit and thus the game would be finished. The number of different sequences in this case is 10 i.e. ten different quintuplets using one digit can be formed from ten decimal digits. To reach the goal in any of these 10 cases, one would need not more than 4 attempts or moves.

Case 3. Let us assume that the opponent has used two digits to form his initial quintuplet. The total number of possible different pairs is 45 i.e. the number of variations of 2 elements out of 10, with repetition. All available variations are divided into three possible sub-cases, that differ in answers received from the opponent (or from the first function), namely: 1) 1:*x* and 1:*y* (25 different pairs), 2) 2:*z* and 0:0 (10 pairs), and 3) 0:0 and 2:*w* (10 pairs). When these two digits are determined, one can make exactly 30 ( $= 2^5 - 2$ ) different quintuplets out of these. If the number of occurrence of both figures is known, the number of variations can be made considerably smaller. Each of the two digits could be used at least once, and not more than four times by the opponent. There is a symmetry here that can be exploited to good use. As described above, if the frequency of both digits is known, there are exactly 5, 10, 10 and 5 different quintuplets. The total number of quintuplets having two different figures is  $1350 = 45 \cdot 30$ . The opponent's replies which leaves 25 different pairs to explore are less favorable than those after which there are only 20 possible pairs to investigate. In the first sub-case, one would need to ask one question more in order to discover the unknown quintuplet.

Let us consider in detail a particularly complex case. When just one digit is properly guessed in both the first and the second attempt, in each of the two quintuplets that were used in the first and the second move, it is necessary to proceed in the following manner. The third move could be: *c, a, a, b, b*. Opponent should provide an answer about the number of correctly chosen digits, and all possible answers are: 0, 1 or 2. These answers further decrease the set of candidates to explore, providing additional information in order to help player determine how many times each digit appears in the opponent's initial unknown quintuplet and at which positions. Out of all available replies, the worst case response is of the form 1:*x*, causing the number of possible candidates to be narrowed down to 3. It can be concluded that in this case, one unknown digit could be: *a, b* or *c*. If that is the case, the first two, *a* or *b*, can appear exactly once, while the third one, *c*, appears at

least once. The fourth move now should be:  $c, c, b, c, c$ . Furthermore, the most unfavorable reply out of all possible, would be  $1:x$  again. After receiving that one, the player deduces that the figure used is either  $b$  or  $c$  and that it occurs exactly once. The third and the fourth move has to be repeated as previously described, for the quintuplet formed in the second move. Therefore, the second function completes its analysis using no more than six moves. In the seventh move it is possible to discover all necessary information that remained hidden after the fourth and the sixth move.

After finding the digits the opponent has used to form the initial quintuplet, and the number of times each of these figures occurred in the initial quintuplet, the number of possible quintuplets is reduced from 1350 to at most 10. At each position in the player's guess, one or the other figure appears. Demanding that each digit has to appear at a different position throughout entire guessing, provides that using the last function for each position will determine the presence or absence of these figures, without asking any further question. To finish the game with success in this case it is enough to ask seven questions.

Case 4. Let us assume that the player has used three different digits to form the initial quintuplet. It is possible to form 120 different triplets out of ten decimal digits or variations of 3 elements out of 10, with repetition. Further on, each triplet can be used to form  $150 (= 3^5 - 3 \cdot 30 - 3)$  different quintuplets. The total number of quintuplets thus is 18000. The opponent has used three digits when the answers he had been giving to the first two questions were: 1)  $3:x$  and  $0:0$ , or 2)  $2:y$  and  $1:z$ , or 3)  $1:w$  and  $2:q$ , or 4)  $0:0$  and  $3:r$ . This helps reducing the number of digits that can appear in an unknown triplet and the conclusion can be drawn that the number of possible triplets for sub-cases 1) to 4) is: 10, 50, 50 and 10, respectively. Upon determining that the three digits were used, among all possible cases of initial quintuplets, the worst case is either 2) or 3) i.e. the one where the two digits are in one quintuplet and the remaining digit is in another quintuplet. Fig. 1. (on the right) shows the procedure of asking questions in such complex cases.

In Fig. 1, after the third question, the notation  $\binom{abc}{1}$  is used to indicate that the opponent has used just one out of three numbers,  $a, b$  or  $c$ . Not later than the fifth move unraveling begins, both of the digits occurring in the quintuplet as well as the number of their repetitions, if digits are already detected. The whole process requires not more than seven questions. The number of moves to reach the goal is determined by the digits used and the frequency of their occurrence in the initial quintuplet. The eighth question is needed to determine the correct order of digits.

Case 5. Let us assume that the opponent have used four different figures. The number of different quadruplets is 210, and each of them can form  $240 (= 1024 - 4 \cdot 150 - 6 \cdot 30 - 4)$  different quintuplets. The total number of all possible quintuplets is the 50400. Out of main 6 cases, this is the worst one (the most difficult to solve) - when the opponent uses four digits one of which is repeated. The

process of discovering unknown digits is conducted in a manner analogous to the previous four cases. After the first function has been applied, all possible results are symmetrical and are described with pairs of replies: 1)  $4:x$  and  $0:0$ , 2)  $3:y$  and  $1:z$ , 3)  $2:q$  and  $2:w$ , 4)  $1:r$  and  $3:s$ , or 5)  $0:0$  and  $4:m$ . Each of these paired answers helps us conclude that the number of possible quadruplets is: 5, 50, 100, 50, and 5, respectively. Most unfavorable is the sub-case 3), when the two digits appear in each of the two quintuplets. It is already shown (in the case 3. having two digits) that there are up to six trials needed to determine the unknown figures. To find out which of the figures is repeated, two more questions are required. So, if the opponent had used four unknown figures, not more than eight questions are needed to discover initial quintuplet. The procedure is completely analogous to that conducted in the case of using three digits.

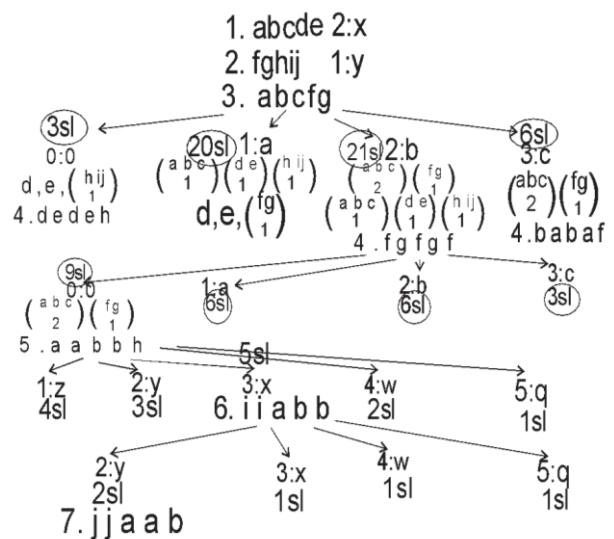


Fig. 1. The part of the search tree for the case with three different digits

Case 6. If the player has used five different digits, then the problem becomes the same as the one presented in [1]. The first move is as described in [1]. The second move of the algorithm given here is not an integral part of the algorithm presented in [1], however it can be used further in the process of determining positions of digits in the quintuplet. To reach the end of the procedure given in [1], it takes a maximum of six moves, which should be incremented by one, to add the second move described here. In the case investigated here, it takes up to seven moves to reach the end of this algorithm i.e. to determine the initial quintuplet formed out of five different digits. The total number of these quintuplets is 30240.

Upon determining the digits and their frequency the opponent has taken, the player forms all variations of five elements out of ten, now without repetition. The number of quintuplets made of two elements with repetitions is 30. However, if the player knows how many times the figure is repeated, then the number of variations is only 5 or 10, depending on whether the two digits in the quintuplet are used 1) 1 and 4, or 2) 2 and 3 times, respectively. This is important fact and it emphasizes the effect of including

functions that determine the frequency of digits used, because to find the final sequence, this procedure requires elimination of substantially smaller number of possible candidates. Once all the digits used are found and their frequencies are discovered, the player starts to determine their positions. For that purpose, all the possible variations of length five are formed with a given number of digits and known frequencies. Using the answers already known to the opponent, all variations that do not correspond to any result are eliminated from the set explored in the following way. Let us assume that one of the answers is 3:1. Out of all available variations, the player retains only those in which exactly one out of three known digits is in its position. Different answers help the player to eliminate different number of possible quintuplets. Number of eliminations depends on the number of different digits as well as on its frequencies. Number of eliminations that follow the answer 3:1 out of all initially possible variations is given in the Table 1. The same answer obtained in the later moves causes that player eliminates smaller number of quintuplets because there are overlaps.

TABLE 1. THE NUMBER OF REMAINING ELIMINATIONS DEPENDING ON THE NUMBER OF DIFFERENT DIGITS WHEN ANSWER IS 3:1.

Answer 3:1	The total number of quintuplets	The number of eliminated quintuplets
<i>a, a, b, b, c</i>	30	18
<i>a, b, c, c, c</i>	20	11
<i>a, a, b, c, d</i>	60	38
<i>a, b, c, d, e</i>	120	78

In paper [1] it was shown that for elimination of all variations, except one that is the perfect match, 4 moves are sufficient, none of which are in form 0:0. It should be emphasized that the procedure described here uses different questions, as compared to those asked in [1], so the starting six questions here can contain several steps getting the reply 0:0. Such steps are totally useless in the proposed analysis. In the approach given in [1], such cases might have happened only in the first step. However, having in mind that 6 moves are used from the beginning and at least two of these cannot be 0:0, to finish the game no more than one more move will be needed.

Summarizing the above mentioned, it can be concluded that, when digits in the initial quintuplet can be repeated, the largest number of required questions is eight, and this happens only when initial sequence has only one digit repeated twice (case 5.). This is considerably smaller than the number of moves a human needs to do the same search.

A program that emulates this procedure has the form:  
short InitialFive[5], MyMoves[50][5], MyResult[50][2],  
OpponentsResult[50][2], OpponentsMoves[50][5];

```
int main() {
    int i, j, DigitNumber;
    int variations[120][5]
    int One[10], Two[45][2], Three[120][3],
        Four[210][4], Five[252][5];
```

```
    FormQuintuplet();
    DigitNumber = DetermineDigitNumber();
    switch(DigitNumber) {
    case 1: DetermineDigits(1);
            return 1;
    case 2: DetermineDigits(2);
            DetermineFrequencies();
            break;
    case 3: DetermineDigits(3);
            DetermineFrequencies();
            break;
    case 4: DetermineDigits(4);
            DetermineFrequencies();
            break;
    case 5: DetermineDigits(5);
            break;
    default: error();
            return 1;
    }
    EliminateVariations();
    return 0;
}
```

### III. CONCLUSION

The procedure described here shows that the far better efficiency can be achieved when the search is done for groups of elements as compared to approach using uniformed search. Also, the algorithm becomes much more effective if the symmetry and other properties are taken into account in decision making or selection performing. Whenever it is possible, different actions specified by decomposition should be partially overlapped. As one consequence, the efficiency of the algorithm is increased.

This algorithm performs more efficient than a human, in solving the same task - searching. Intended for our future research is to compare the results of the game played between two humans, with those obtained when a human player opposes the algorithm here described.

### REFERENCES

- [1] A. Srdanov, N. Ratković Kovačević, S. Vasić, and D. Milovanović, "Emulation of artificial intuition using random choice and logic," presented at and in *Proc. the 13th Symposium on Neural Networks (NEUREL)*, Belgrade, Serbia, pp. 57-62, Nov. 2016.
- [2] S. Kirkpatrick, C. D. Gelatt Jr, and M. P. Vecchi, "Optimization by Simulated Annealing," *Science*, vol. 220 (4598), pp. 671-680, May. 1983.
- [3] David E. Goldberg, Kalyanmoy Deb, and James H. Clark, "Genetic Algorithms, Noise, and the Sizing of Populations," *Complex Systems*, vol. 6 (4), pp. 333-362, Jan. 1992.
- [4] Sung-Hyuk Cha, and Charles Tappert, "A Genetic Algorithm for Constructing Compact Binary Decision Trees," *J. of Pattern Recognition Research*, vol. 1, pp. 1-13, Feb. 2009.
- [5] R. M. Karp, "An introduction to randomized algorithms," *Discrete Applied Mathematics*, vol. 34, pp. 165-201, 1991.
- [6] Xiaohui Bei, Ning Chen, and Shengyu Zhang. (2013. April). On the Complexity of Trial and Error. *arXiv: 1205.1183v2* [cs.CC]. Available: <http://arxiv.org/pdf/1205.1183v2.pdf>

# Alat Unity3D u konceptu proširene realnosti

Stevan Čakić, Aleksandar Popović, *Prirodno-matematički fakultet, Univerzitet Crne Gore*

**Sadržaj** — Koncept proširene realnosti danas je sve popularniji i zastupljeniji u svakodnevicu. U ovom radu će biti opisan sam koncept, alati koji će biti korišćeni pri izradi aplikacije koja sadži osnovne elemente ovog koncepta, kao i sama aplikacija. Pri izradi aplikacije korišćen je alat Unity3D, kao i dodatak za ovaj alat, Vuforia. Sami kod pisan je u programskom jeziku C#. Jedan od osnovnih ciljeva rada je približiti koncept proširene realnosti čitaocu na što jednostavniji način.

**Ključne riječi** — Proširena realnost, Unity3D, Vuforia

## I. UVOD

Stalni razvoj informatike nas dovodi u situaciju da smo gotovo svakodnevno zasuti informacijama o novim konceptima i dostignućima u ovoj oblasti. Jedan od takvih koncepata je proširena realnost (eng. *augmented reality*). Proširena realnost nam omogućava da pomoću dodatne opreme kao što su monitor, posebne naočare, kamera, itd. čujemo i vidimo, osim realnosti, stvari koje nisu dio stvarnog okruženja [1]. Ova tehnologija je bila dugo zapostavljena, a glavni razlozi su nedostatak opreme, sama složenost koncepta, spore obrade grafičkih modela i drugo. Kako u novijem periodu nije teško nabaviti računare dovoljno dobrih performansi za obradu složenih grafičkih modela, a i ostalu opremu neophodnu za realizaciju ovog koncepta, sve više ljudi se uključuje u istraživanje i primjenu ove oblasti.

Alat koji je danas najzastupljeniji za realizaciju koncepta proširene stvarnosti je Unity3D [2]. Ovaj alat se koristi prevashodno za izradu igara. Neke od najsavremenijih igara su pravljene baš pomoću ovog programa. Upotreba alata Unity3D je dosta olakšala kreiranje igara, posebno onih kompleksnijih u kojima je potrebno koristiti napredne matematičke discipline i zakone fizike. Međutim, da bi se na jednostavan način realizovao koncept proširene realnosti nije nam samo dovoljan alat Unity3D. Osim opreme (bar kamera za najjednostavnije primjere) potrebna je dodatna biblioteka, tj. paket za pomenuti alat, Vuforia.

I danas je koncept proširene realnosti dosta neistražen i još nije u potpunosti ušao u sastavni dio naših života. Međutim, prema nekim procjenama predviđa se da će ovaj koncept naglo početi da se primjenjuje u svakodnevnom životu i to već u skorijem periodu uz propratne tehnologije [3]. Jedan od projekata iz ove oblasti koji je doživio veliku popularnost poslednjih mjeseci jeste igra *Pokemon GO!*. Ova igra je u potpunosti realizovana pomoću alata

Unity3D i pravi je reprezent šta je to proširena realnost.

U ovom radu biće opisan koncept proširene stvarnosti, alat Unity3D, način realizacije ovog koncepta primjenom programa Unity3D i dodatnog paketa Vuforia. Takođe, sve ovo će biti sprovedeno kroz jedan konkretan primjer. Programski jezik C# je korišten pri izradi praktičnog dijela.

## II. OPIS KONCEPTA AR

Proširena realnost (eng. *Augmented Reality – AR*) je koncept kod kojeg se virtuelni elementi dodaju u stvarni svijet, na način da ti elementi izgledaju kao da su dio stvarnosti. Ključni cilj ovog koncepta je kreirati takav sistem gdje korisnik ne bi mogao da uoči razliku između virtuelnog - izmišljenog i stvarnog. Glavne karakteristike ovog koncepta su: spajanje stvarnog i virtuelnog, interakcija u realnom vremenu i percepcija tri dimenzije.

Godine 1968. Ivan Sutherland kreira prvi displej sistem koji se stavlja na glavu poput naočara. 1975. godine Myron Krueger sa saradnicima formira projekat *Videoplace*, a ideja koja stoji iza ovog projekta je kreiranje vještačke realnosti koja treba da okruži korisnika i da odgovara na njihove pokrete i akcije bez korišćenja specijalne opreme. Godine 1990. Thomas Caudell uvodi pojam proširena realnost dok je opisivao displej, koji su koristili električari aviona, gdje se miješala stvarnost sa virtuelnim elementima. Nakon dvije godine, 1992., Louis Rosenberg kreira tkz. virtuelni učvršćivač, a za njega se može reći da je to prvi AR sistem koji je kreiran. Od ove godine koncept se sve brže razvijao, već 2004. grupa njemačkih istraživača kreirala je AR sistem za korišćenje na mobilnim telefonima. AR igra Ingress kreirana 2014. godine, a dvije godine kasnije, 2016. godine, kreirana igra *Pokemon GO* [4].

Tri osnovna problema kod proširene realnosti su:

- Miješanje i prikaz slike
- Poravnanje i praćenje
- Prikupljanje podataka

Najbitniji ulazni i izlazni uređaji sistema proširene realnosti su vizuelni uređaji (npr. kamera). Oni se zasnivaju na konceptu miješanja virtuelne i stvarne slike. Tipovi miješanja virtuelne i stvarne slike su optičko miješanje, video miješanje i projekciono miješanje. Optičko miješanje koristi poluprovodno ogledalo ispred korisnika, po jedno za svako oko, na koje se projektuje slika sa monitora (optički kombinator). Kod video miješanja koriste se kamere koje snimaju stvarni svijet, a slika se miješa na računaru i šalje korisniku. Kod projekcionog miješanja slika se projektuje na predmete u stvarnoj okolini. Prikaz slike je usko vezan sa načinom

miješanja slike. Najčešći načina za prikaz slike su: displej na glavi, u oku, u prostoru, u ruci i projekcioni prikaz [6].

Poravnavanje realnih i virtuelnih predmeta je veoma bitan segment proširene realnosti. Da bi se postiglo poravnavanje, virtuelna scena se kreira u koordinantnom sistemu koji odgovara onom iz realnog svijeta. Ovo omogućava da se položaji stvarnih predmeta prebacuju u virtuelne scene na ispravan način. Postojanje ovakvog koordinantnog sistema omogućava da se virtuelni elementi postave u odnosu na stvarne. Da bi se ovo sprovelo, potrebno je znati položaj posmatrača, položaj kamere u stvarnom svijetu, položaj svih stvarnih predmeta na sceni, kao i položaj ostalih elemenata u virtuelnom svijetu. Ovim se postiže da se pomoću računara ostvari interakcija između stvarnih i virtuelnih stvari i dobija se poklapanje virtuelne i stvarne scene. Potrebno je ostvariti veliku preciznost jer ljudsko oko vrlo uspješno i brzo detektuje i najmanja pomjeranja. Podaci koji su potrebni za poravnavanje u AR sistemu dobijaju se praćenjem. Praćenje je postupak dobijanja položaja i orijentacije predmeta i posmatrača u realnom vremenu. Sistemi praćenja su prvo razvijeni za virtuelnu realnost. Ovi sistemi su obično nedovoljno precizni jer VR ima niže zahtjeve. Iz tog razloga su razvijene nove tehnike za AR. Neke od tehnika praćenja su: magnetno praćenje, ultrazvuk, mehaničko praćenje, optičko praćenje, GPS, hibridno praćenje. Razlikujemo dva tipa grešaka kod poravnanja: statičke greške i dinamičke greške. Grupi statičkih grešaka pripadaju: greške pri praćenju, pogrešni parametri virtuelne kamere, optička izobličenja, mehaničke nepreciznosti opreme. S druge strane, grupi dinamičkih grešaka pripadaju: one koje su nastale zbog kašnjenja, one nastale pri praćenju, prenosu podataka i greške pri iscertavanju. Optičko praćenje radi na principu detektovanja posebnih markera na slici koje računar, preko kamere, prima iz stvarnog svijeta. Marker su najbitniji elementi kod ovog tipa praćenja, mogu biti različitih boja, oblika, mogu biti i nevidljivi oku, mada u tom slučaju se koriste posebne kamere za njihovu detekciju. Slika koja stigne od kamere i koja se potom prenosi na računar se obrađuje i pretražuje se marker. Pomoću dobijene pozicije markera, računar dobija podatke o stvarnom svijetu, npr. lokacija predmeta, posmatrača i slično [6].

Prikupljanje podataka je zajedničko ime za tehnike dobijanja dodatnih podataka za prikaz u proširenoj realnosti. To su: medicinske slike i 3D rekonstrukcije, karte za precizno prekrivanje stvarnih i virtuelnih predmeta, prikaz za baze podataka.

Koncept proširene realnosti primjenjuje se u raznim oblastima i za razne svrhe: medicina (npr. koristi se za virtuelni rendgen), arhitektura (za dizajn enterijera, za prikaz struktura, instalacija), robotika (npr. koristi se za planiranje ugla rotiranja elementa robota), za obuku, za edukaciju, u vojne svrhe (kod pilota za dobijanje dodatnih informacija navođenja, prikaz ciljeva, dodatne taktičke informacije), marketing (za prikaz virtuelnih proizvoda u kombinaciji sa stvarnošću), industrija zabave (video igre).

Prednosti koncepta proširene realnosti su: informacije u okruženju postaju interaktivne i moguća je digitalna manipulacija sa njima, opis realnosti sa dodatnim informacijama koje nisu standardno dostupne, performanse

i preciznost, dok su mane sledeće: zadiranje u privatnost, skupa realizacija i održavanje, ljudi gube osjećaj šta je ispred njih.

Osim ovog koncepta danas je zanimljiv i koncept virtuelne realnosti. Glavna razlika između ova dva koncepta je to što kod proširene realnosti imamo prisutne stvarne i virtuelne elemente, dok je kod virtuelne realnosti sve virtuelno. Takođe, Microsoft Hololens je danas sve popularniji koncept. Pripada konceptu proširene realnosti, a za prikaz koristi displej na glavi (posebne naočare).

### III. OPIS ALATA UNITY3D I DODATKA VUFORIA

Program Unity3D pripada grupi alata koji se koriste prevashodno za izradu igara (*eng. game engine*). *Game engine* [7] je program dizajniran da omogući lakšu, bržu, ekonomičniju, portabilnu i efikasniju izradu video igara za konzole, računare i mobilne uređaje. Glavne komponente *game engine*-a su: komponenta za prikaz 2D i 3D grafike, komponenta za fiziku, komponenta za zvuk, komponenta za programiranje, komponenta za animaciju, komponenta za vještačku inteligenciju, komponenta za mrežu, komponenta za upravljanje memorijom i komponente za niti. Pored Unity3D *game engine*-a poznati su i Unreal Engine, Cry Engine, Hero Engine, GameMaker i drugi.

Unity3D [8] je kreiran od strane Unity Technologies, najavljen je 2005. od strane Apple-a. Do sada je publikovan u 5 glavnih verzija. Podržane platforme za Unity3D su: Android, Apple TV, iOS, Nintendo, Linux, Xbox 360, PlayStation 4, Windows, Wii, Windows Phone, itd.

Neki od najbitnijih elemenata alata Unity3D su [9]:

- objekti igre (*eng. game objects*) su osnovni objekti ovog alata koji su pri kreiranju prazni, ali mogu im se dodati razne karakteristike, dekoracije i resursi. Njima se mogu dodati i tkz. kolajderi koji služe za detekciju interakcije između objekata;
- primitivni objekti su objekti koji su već spremni za scenu. Grupu primitivnih objekata čine: kocka, sfera, kapsula, cilindar, podloga, četvorougao;
- prefabi su korisna karakteristika ovog alata, putem kojih je omogućeno da objekti igre budu sačuvani zajedno sa svim komponentama i osobinama. Jedan prefab se ponaša kao šablon za kreiranje novih instanci objekta na sceni. Svaka izmjena na prefabu se reflektuje na sve kreirane instance, ali moguće je promijeniti karakteristike pojedinačne instance ili dodati joj nove komponente;
- kamera je objekat kojim se definiše pogled na prostor scene, pozicija objekta definiše gledište, dok Y i Z koordinate objekta definišu smjer pogleda i vrh ekrana. Takođe, komponentna kamera definiše veličinu i oblik oblasti koju pokriva. Podešavanjem tih parametara kamera može da prikaže šta je trenutno vidljivi dio na ekranu; kako kamera može da se rotira i pomjera. Prikazani prostor se takođe rotira i pomjera ka tome;
- asemi su resursi koji se koriste za izradu projekta. Alat posjeduje standardnu grupu asemi, to su: 2D objekti, kamere, karakteri, komande kroz razne platforme, efekti, okruženje, sistem za čestice, prototipovi, vozila;



- svijetla su dio svake scene koji definišu boje i ton 3D okruženja. Najčešće, na sceni se nalazi više od jednog svijetla na sceni;
- sistem čestica služi za prikaz efekata na sceni: plamen, oblak, magla, dim, itd. Može da se koristi i za generisanje i animiranje velikog broja manjih 2D slika na sceni (npr. zvijezde, itd.).
- sistem za korisnički interfejs omogućava brzo i intuitivno kreiranje korisničkog interfejsa. Kanvas je glavni element *UI* sistema.

Ponašanje objekata igre je kontrolisano komponentama koje su zakačene na njih. Ugrađene Unity3D komponente su prilično ograničene, pa je potrebno na neki način obezbijediti veću kontrolu nad tim objektima. To se postiže kreiranjem novih komponenti koristeći skripte. One omogućavaju način djelovanja na određene događaje, modifikovanje osobina komponenti tokom postojanja, odgovaranje na zahtjev korisnika, itd [10].

Vuforia je dodatak (framework) za Unity3D koji omogućava pojednostavljenu implementaciju proširene realnosti u Unity3D projektu ovog tipa [11]. Prvi korak je pri podešavanju ovog dodatka je integrisati ga sa Unity3D alatom, nakon toga treba dodati licencu po kreiranju naloga na glavnoj stranici dodatka. Potom je potrebno kreirati bazu markera u kojoj će biti čuvani markeri koji će biti korišćeni u aplikaciji, a bazu treba uvezati sa alatom. Glavna komponenta dodatka Vuforia je AR kamera, koju ne posjeduju Unity3D. Pomoću ove kamere, kamera kao hardverska komponenta može da detektuje markere na sceni na osnovu kreirane baze.

Prednosti Unity3D alata su: cijena (besplatna osnovna verzija koja je dovoljna za kreiranje jednostavnijih igara), jednostavno dodavanje novih resursa, programiranje je pojednostavljeno, intuitivan UI, redovno ažuriranje, detekcija kolizija bez matematike i portabilnost. Mane su sledeće: glomazan alat, nije podržan za Linux, skup ako želite da koristite sve njegove karakteristike, rad na velikim igrama zahtijeva puno optimizacije, ne podržava napredne funkcionalnosti kao što je to slučaj kod Unreal Engine-a.

#### IV. OPIS IGRE

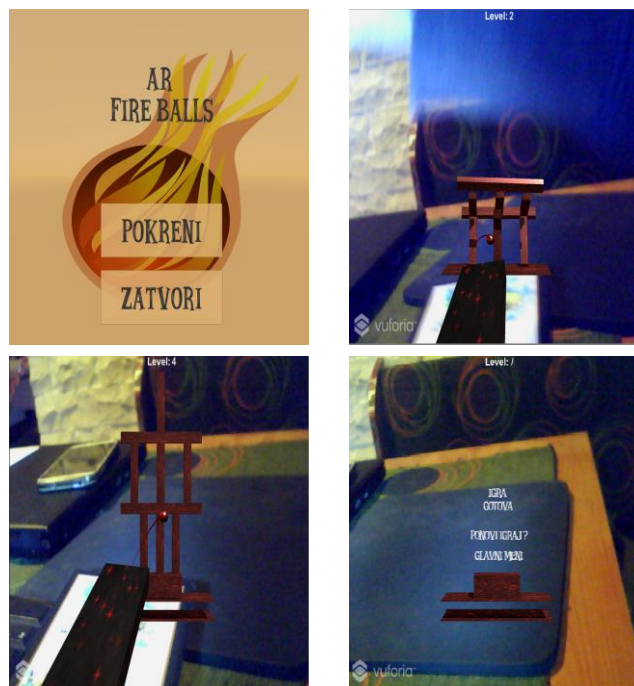
Igra je primarno namijenjena za telefone sa Android operativnim sistemom, mada se može instalirati i na desktop računaru. Nije predviđena za igranje na mreži, iako Unity pruža mogućnost prilagođavanja, gotovo automatizovanog, bilo koje aplikacije na više platformi. Takođe, potrebno je da vaš telefon, odnosno računar ako se odlučite da to bude vaša platforma, posjeduje kvalitetnu kameru. Kamera boljeg kvaliteta je neophodna radi lakšeg detektovanja objekta (markera - stvarni predmet) na osnovu kog se generiše ispaljivač (kvadar) za loptice.

Cilj igre, koja je primarno kreirana kako bi se putem nje prikazali glavni elementi opisanog alata i koncept proširene realnosti, je preći sve nivoe, a ima ih četiri. Novi nivo se prelazi tako što ispaljivanjem loptica obarate blokove. Igra se sastoji od glavnog menija i prostora igre. Pri pokretanju aplikacije korisniku se nudi mogućnost da startuje ili zatvori aplikaciju. U slučaju da korisnik pritisne dugme *Pokreni* igra se pokreće, a u slučaju da

pritisne/dodirne dugme *Zatvori* korisnik napušta aplikaciju. Na početku, marker (Sl. 2.) treba okrenuti ka kameri zbog detekcije i vezivanja sa objektom koji predstavlja ispaljivač loptica. Loptice se ispaljuju lijevim klikom miša (PC/Laptop) ili dodirnom ekrana telefona (kod tkz. pametnih telefona). Svaki put kad je nivo pređen automatski se pokreće novi nivo sa drugačijim rasporedom blokova. Nakon što korisnik ove aplikacije pređe sve nivoe otvara se prozor iz kog je moguće odabrati dvije opcije: Ponovo igraj (vraća vas na prvi nivo) i Glavni meni (vraća vas na početni meni).

Pri izradi projekta korišćeni su sljedeći alati: Unity3D, Vuforia i C# (integrisan u okviru Unity3D alata) za pisanje skripti. Projekat je kreiran da pokrije što više elemenata pomenutog alata kao i da se kroz njega prikaže koncept proširene realnosti. Grafika same aplikacije je prilično jednostavna jer nije bio cilj prikazati što atraktivniju grafiku, već najosnovniji način funkcionisanja proširene realnosti.

Najveći izazov pri izradi ovog projekta je bio odabir dobrog markera. U prvom dijelu rada je opisano šta je marker i koja mu je svrha. Za njega je bitno pomenuti i to da se mora voditi računa pri njegovom odabiru, bilo kakva simetričnost nije poželjna. Razlog je taj što pri detekciji markera od strane kamere može da se desi da kamera ne može da odluči u kom smjeru je okrenut marker. Na sceni se može naći jedan ili više markera. Ovdje je korišćen samo jedan marker i to za kreiranje primitivnog objekta (kvadar) koji se veže na sami marker pri detekciji od strane kamere.



Sl. 1. Aplikacija AR Fire Balls

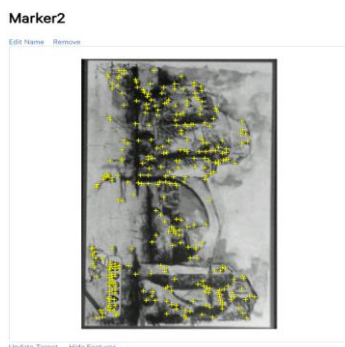
Aplikacija sadrži glavne komponente Unity3D alata i osnovne koncepte proširene realnosti.

Prvi korak pri izradi ovog projekta je bio instalacija alata Unity3D, kao i njegovo integrisanje dodatka Vuforia za proširenu realnost. Nakon podešavanja AR kamere i licence za korišćenje Vuforia dodatka potrebno je kreirati

bazu sa markerima. Baza se integriše sa Unity3D alatom, a potom se taj marker dodaje na scenu. Projekat sadrži tri scene: *PreStart*, *MainMenu* i *GameScreen*. Prva scena (*PreScreen*) služi za pripremu početka igre i ona je nevidljiva korisniku. Druga scena (*MainMenu*) služi za pokretanje ili zatvaranje aplikacije, a treća scena (*GameScreen*) sadrži glavne elemente igre kao što su blokovi koje treba oboriti, loptice, itd.

Za potrebe igre kreirano je šest skripti:

- *Block* – Služi za provjeru da li je blok oboren;
- *CompletedGameManager* – po prelasku igre omogućava korisniku vraćanje na početni meni ili pokretanje igra od prvog nivoa;
- *GameManager* – najbitnija skripta, provjera da li je nivo pređen, kreira potrebne elemente za novi nivo, omogućava prelazak na sledeći nivo, uklanja oborene blokove sa scene;
- *Gun* – kontroliše kretanje kuglice/metka nakon i detektuje događaj da neko želi da ispali metak/kuglicu;
- *MainMenuManager* – upravlja početnim menijem
- *TimedObjectDestructor* – uništava metak/kuglicu posle izvjesnog vremena (odabere se pomoću javne promjenljive koja je vidljiva Unity alatu na objektu koji sadrži ovu skriptu, a to je objekat *Gun*)



Sl. 2. Marker koji prepoznaje kamera

Moguće je uvesti brojna poboljšanja aplikacije kao što je npr. dodavanje atraktivnije grafike, izmjena interfejsa, dodavanje više nivoa, proširiti igru za novim funkcijama itd. Ipak, svrha izrade aplikacije nije komercijalna, već da se kroz nju vide osnovni elementi AR koncepta, kao i da se na što jednostavniji način opišu osnovne komponente Unity3D alata.

## V. ZAKLJUČAK

Postoje brojne primjene AR, ali ova tehnologija je i dalje u procesu razvoja i prilično je mlada, tako da nema puno naprednih aplikacija koje možete vidjeti, a da su rađene u AR-u. AR je nešto potpuno novo u informatici i prema riječima stručnjaka ovaj koncept bi uskoro mogao da zaživi kroz primjene u različitim domenima, od zabave do obuke iz raznih oblasti.

Današnji *game engine* alati olakšavaju izradu veoma kompleksnih aplikacija kao što su VR, AR, 3D igre. Unity3D je vrlo moćan predstavnik ovih alata i danas je veoma popularan. Gotovo svako ko krene putem izrade igara se susretne sa njim. Koliko je moćan moglo se vidjeti na primjeru ove aplikacije i to kroz kretanje loptica i detekciju sudara loptica i blokova. Pozivom nekoliko prethodno implementiranih metoda za kretanje i izmjenom par parametara može se kreirati prilično glatko i ispravno kretanje koje poštuje sve zakone fizike. Pored toga, može se izvršiti detekcija sudara dodavanjem kolajdera na one elemente koji mogu da se sudare sa drugim.

Danas postoji veliko "takmičenje" između koncepta AR i VR. U budućnosti ćemo saznati ko će pobijediti. Pretpostavlja se da će oba koncepta zaživjeti i to uskoro, ali koji će biti prisutniji u našoj svakodnevici ostaje da se vidi.

## LITERATURA

- [1] D.W.F van Krevelen, R. Poelman, "A Survey of Augmented Reality, Technologies, Applications and Limitations", 2010, The International Journal of Virtual Reality.
- [2] A. Blick, "Introduction to Augmented Reality with Unity3D", 2012.
- [3] J. R. Vallino, "Interactive Augmented Reality", 1998, University of Rochester, New York
- [4] [https://en.wikipedia.org/wiki/Pok%C3%A9mon\\_Go](https://en.wikipedia.org/wiki/Pok%C3%A9mon_Go), 2016
- [5] [https://en.wikipedia.org/wiki/Augmented\\_reality](https://en.wikipedia.org/wiki/Augmented_reality)
- [6] S. Siltanen, "Theory and applications of marker-based augmented reality", 2012
- [7] [https://en.wikipedia.org/wiki/Game\\_engine](https://en.wikipedia.org/wiki/Game_engine)
- [8] [https://en.wikipedia.org/wiki/Unity\\_\(game\\_engine\)](https://en.wikipedia.org/wiki/Unity_(game_engine))
- [9] Unity3D manual Available at : <https://docs.unity3d.com/Manual/index.html>
- [10] <https://docs.unity3d.com/ScriptReference/>
- [11] Vuforia Manual, Available at: <https://developer.vuforia.com/>

## ABSTRACT

Today, augmented reality concept is getting more popular, and more present in everyday life. In this article we will explain the AR concept, tools which are used to create an application containing basic elements of this concept, and the application itself. For creating this application Unity3D tool was used, as well as additional framework named Vuforia. Scripts are written in C# programming language. One of the main goals this article is to bring closer the AR concept to a reader in a very simple way.

## UNITY3D TOOL IN THE AUGMENTED REALITY CONCEPT

Stevan Čakić, Aleksandar Popović, *Prirodno-matematički fakultet, Univerzitet Crne Gore*

# Geo-lokacijski servis za javni prevoz

Lazar Radinović, Aleksandar Popović, *Prirodno-matematički fakultet, Univerzitet Crne Gore*

**Sadržaj** — Planiranje kako najbrže stići od lokacije A do lokacije B, odnosno pronalaženje najkraćeg puta, jedan je od svakodnevnih problema sa kojima se susreću kako stanovnici nekog grada tako i posjetioci. Jedno od rešenja ovog problema može da bude servis koji će odgovarati na zahtjeve i sadržati dodatne informacije o javnom prevozu toga grada na jednom mjestu. Rad opisuje implementaciju jednog ovakvog servisa koji koristi više savremenih tehnologija i koji se sastoji iz više komponenti. Veći dio softvera pisan je u programskom jeziku PHP5 koristeći Laravel framework, dok je dio zadužen za pronalaženje puta napisan u programskom jeziku C. Takođe, cilj rada je i da predstavi kako upotrebljavati savremene tehnologije kako bi razvili robustnu aplikaciju.

**Ključne riječi** — GIS, Laravel, Leaflet, PostGis

## I. UVOD

Krajem dvadesetog vijeka, potrebe za prostornim podacima koje se odnose na topografiju i specifične teme na zemljinoj površini, kao što su prirodni resursi, naglo su se uvećale. Razvoj daljinske detekcije, kao i ostalih metoda prikupljanja prostornih podataka, omogućili su kartiranje velikih površina sa izuzetnom preciznošću. Narastajuće potrebe za prostornim podacima, kao i alatima i tehnikama za njihovu analizu, mogu se uspješno riješiti primjenom računara [1]. Potrebe za prostornim podacima i prostornim analizama nijesu samo vezane za eksperte i istraživače koje se bave Zemljom. Ova vrsta podataka može poslužiti i za rješavanje svakodnevnih praktičnih zadataka i problema kao što je dobijanje informacija o gradskom prevozu i to u realnom vremenu. Ovo je omogućeno razvojem Internet servisa, odgovarajućih programskih alata i prostornih baza podataka. U ovom radu biće opisana implementacija jednog takvog servisa koji objedinjuje više savremenih tehnologija u ovoj oblasti.

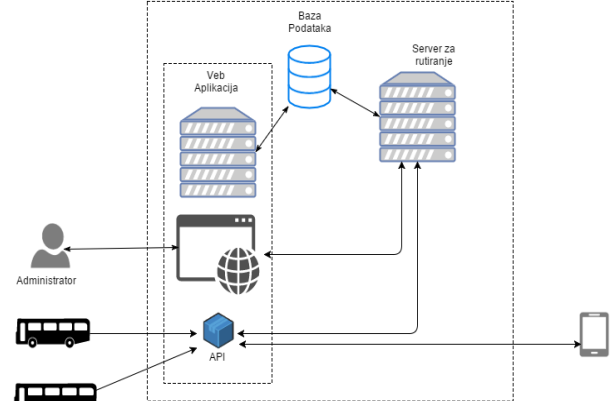
Savremeni sistemi za upravljanje bazama podataka (SUBP) pored tradicionalnih alfa-numeričkih podataka u stanju su da prihvataju i čuvaju specijalizovane podatke kao što su slike, multimedijalni zapisi i dr. Između ostalih, prostorni podaci izdvojili su se kao jako važni za mnoge programe i aplikacije. Za relativno kratko vrijeme najzastupljeniji SUBP-ovi su prošireni tako da mogu efikasno da prihvataju i čuvaju prostorne podatke [2]. Takođe, metode indeksiranja i upitni jezici su prošireni u cilju efikasnog procesiranja upita za tu vrstu podataka. Jedan od takvih SUBP-ova jeste i Postgres, čija komponenta PostGis predstavlja implementaciju navedenih zahtjeva ([3], [4]). Pored toga, PostGis poštuje i standarde propisane od strane OGC-a (Open GIS Consortium) [5]. Upravo ovaj SUBP je korišćen za

čuvanje prostornih podataka u okviru servisa koji je prezentiran u ovom radu.

Dobijanje informacija o gradskom prevozu i najkraćim rutama jedan je od svakodnevnih problema sa kojim se susreću kako stanovnici nekog grada tako i posjetioci. Jedno od rješenja ovog problema može da bude servis koji će odgovarati na zahtjeve upućene od strane građanina ili turista, i upravo to je tematika ovog rada. Rad opisuje implementaciju jednog takvog servisa, koji se sastoji iz tri komponente: administrativna komponenta, API komponenta i server za rutiranje. Administrativna komponenta omogućava unos i održavanje podataka u sistemu, API komponenta omogućava trećoj strani da dobije željene podatke dok treća komponenta, server za rutiranje, vrši pretragu najkraćeg puta. Administrativna i API komponenta realizovane su kao troslojna web aplikacija i u daljem tekstu za ovaj dio sistema ćemo koristiti termin Web Aplikacija. Server za rutiranje nezavisan je od navedenih komponenti i sa njima komunicira putem C soketa.

## II. ARHITEKTURA RJEŠENJA

U ovom sistemu možemo uočiti više entiteta: stanica, autobuska linija, trasa, prevoznik i autobus. Ovi entiteti zajedno čine javni prevoz jednog grada, pa je zbog toga uveden još jedan entitet projekat. Prostorni podaci koriste standardizovani model WGS84 za mapiranje podataka na zemaljskoj kugli. Stanice su u sistemu predstavljene kao tačke (eng. *Point*) koje imaju atribute geografske širine i dužine. Trase predstavljaju put koji autobus pređe u jednom smjeru, pa svaka autobuska linija ima dvije trase. Trase su predstavljene izlomljenim linijama (eng. *Polyline*). Autobusi su predstavljeni kao tačke. Entitet prevoznik nema prostorne podatke. Model uključuje i entitete za predstavljanje opština (regiona) i država koji su predstavljeni pomoću više poligona (eng. *Multi-Polygon*).



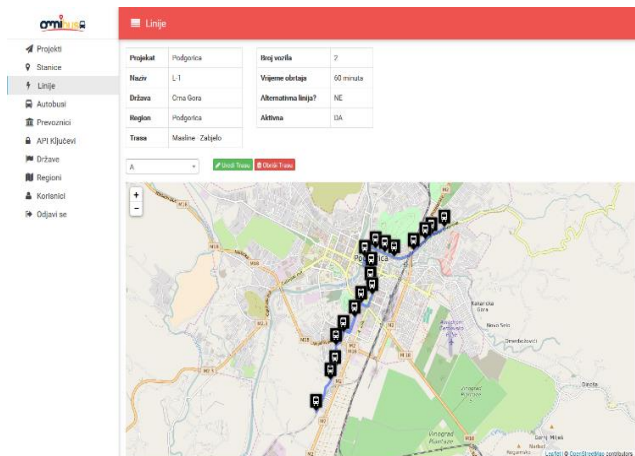
Sl. 1. Arhitektura rješenja

Pored ovih entiteta model takođe uključuje i entitete koji služe za upravljanje servisom: korisnik i API ključ.

Korisnik je entitet koji omogućava osobi da unosi i uređuje podatke dok API ključ je entitet koji omogućava sigurnu razmjenu koordinata autobusa i servisa. Kao što je rečeno, servis se sastoji iz tri komponente: administrativna i API komponenta, koje zajedno čine Web Aplikaciju, i server za rutiranje (Sl.1.).

Administrativna komponenta je izdijeljena na više cijelina tzv. modula. Svaki modul je skup prikaza i formi koje služe za pregled, unos i ažuriranje podataka. Implementirani su moduli: **Projekti, Stanice, Linije, Autobusi, Prevoznici, API Ključevi, Države, Regioni i Korisnici**. Interfejs je urađen pomoću HTML5 tehnologije, stilskog jezika CSS3 i JavaScript i koji se sastoji iz tri cijeline: traka menija (eng. *Sidebar*), zaglavlja stranice (eng. *Header*) i sadržaja stranice (eng. *Content*). Podaci iz formi za unos i uređivanje se većinom šalju serveru putem AJAX-a, dok server većinu podataka šalje nazad putem JSON formata. Na slici 2 prikazan je modul za autobuske linije i jedna od unesenih linija. Komponenta je namijenjena onim korisnicima servisa koji su zaduženi za uređivanje autobuskih linija i autobusa.

API komponenta je sastavni dio Web Aplikacija i dalje se može podijeliti na javni i privatni dio odnosno **Public API** i **Private API**. API komponenti se može pristupiti putem linka <http://api.omnibusapp.com> (navedeni domen nije operativan i naveden je samo kao primjer). Trenutna aktivna verzija je **v1** kojoj se pristupa tako što se string **v1** nadoveže na link za pristup: <http://api.omnibusapp.com/v1>. Cilj ovakvog upravljanja verzijama je da se omogućiti određeni vremenski period da sve aplikacije migriraju na novu verziju API-ja.

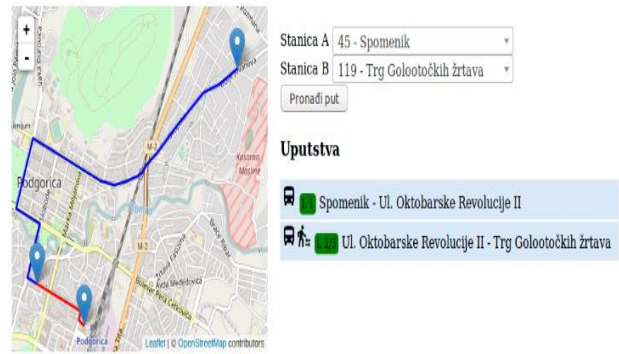


Sl. 2. Prikaz jedne od trasa sa stanicama

Javni dio komponente samo omogućava čitanje informacija. Njegova uloga je da omogućiti trećoj strani, na primer android aplikaciji, da prikazuje detalje o javnom prevozu grada u kome se korisnik trenutno nalazi. Podaci se razmjenjuju putem JSON formata, a u slučaju da su u pitanju prostorni podaci onda se oni razmjenjuju u GeoJSON formatu [6]. Javni dio uključuje izlistavanje stanica i autobuskih linija kao i lokaciju autobusa na trasama u realnom vremenu.

Na slici tri prikazan je primjer aplikacije koja koristi opisani servis. Aplikacija je zatražila od servisa da pronade put za zadate dvije stanice i dobila odgovor u GeoJSON formatu. Zatim je taj put prikazan izlomljenom

linijom. Autobuske linije su predstavljene različitom bojom.



Sl. 3. Primjer treće aplikacije

Privatni dio, t.j. Private API, je namijenjen vlasnicima projekta i prevoznicima i za sada ima samo jedan metod koji služi da autobus javi svoju lokaciju servisu. Za razliku od javnog API-ja, svim privatnim metodama potrebno je slati dodatna dva parametra: javni dio API ključa odnosno ključ pristupa i potpis zahtjeva. Razlog ovakvom ponašanju je obezbjeđivanje sigurnog prenosa koordinata bez uticaja napadača. Vlasnici projekta i prevoznici su dužni da tajni dio API ključa drže u tajnosti.

Ovaj servis uključuje samo bazu podataka i administrativni dio sa serverom za rutiranje ali ne i aplikaciju za korisnike javnog prevoza. Treća strana, ako želi može da razvije, na primer, android aplikaciju koja će komunicirati sa ovim servisom.

### III. SERVER ZA RUTIRANJE

Treća komponenta ovog sistema je server za rutiranje koji je napisan u programskom jeziku C i stoji odvojeno od baze podataka i Web Aplikacije. Pretraživanje grafa je zahtjevan proces pa je zbog performansi odabran jezik C [7]. Sa druge strane, ako bi ovaj dio sistema bio implementiran u PHP-u, kao i ostatak, tada bi pri svakom zahtjevu za pretragu najkraćeg puta između dvije tačke bilo potrebno kreiranje grafa što dodatno troši vrijeme ako graf ima veliki broj čvorova. Iz ovih razloga napisan je poseban program u jeziku C koji pri pokretanju učitava sve stanice i autobuske linije i pravi odgovarajuće grafove po regionima i čuva ih u memoriji. Pored toga, ovaj program otvara i socket za komunikaciju sa Web Aplikacijom (odnosno API komponentom). Prilikom izmjene podataka u bazi potrebno je poslati komandu za ponovno kreiranje grafa.

Komunikacija se obavlja putem soketa. Soket je generalni komunikacioni kanal između dva procesa, u ovom konkretnom slučaju između niti koja opslužuje trenutni HTTP zahtjev i servera za rutiranje [8].

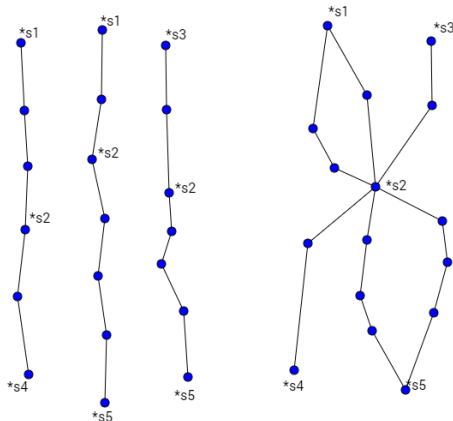
U API komponenti postoji metoda `GET /v1/routing` koja omogućava krajnjem korisniku da zatraži najkraći put između dvije stanice. Argumenti ovog poziva su identifikacioni brojevi početne i krajnje stanice i trenutna lokacija krajnjeg korisnika. API komponenta prihvata zahtjev, čita ulazne argumente i šalje sljedeću komandu serveru za rutiranje preko komunikacionog kanala:

```
search $regionId $start $end.
```

Prvi argument je identifikacioni broj regije u kojoj se korisnik nalazi, a on se izvuče iz baze prije samog slanja komande na osnovu trenutne lokacije krajnjeg korisnika. Potom API komponenta čeka da server za rutiranje pronađe najkraći put i na samom kraju šalje odgovor u JSON format.

Kreiranje grafova se vrši pri pokretanju servera. Za svaki od regiona kreira se poseban graf i čuva se u memoriju pod njegovim identifikacionim brojem. Za skladištenje grafova koristi se uređeni niz. Graf se pretražuje metodom binarnog traženja, pa je graf moguće pronaći za  $O(\log n)$  vrijeme [9], [10]. Podaci za kreiranje grafova se učitavaju pomoću biblioteke **libpq** [11]. Ova biblioteka sadrži sve potrebne funkcije za komunikaciju sa PostgreSQL serverom. Prostorni podaci koji se čitaju iz baze su u binarnom zapisu pa je potrebna transformacija podataka u format pogodan za dalje korišćenje. Graf je u memoriji predstavljen pomoću tri C-strukture: struktura za čvor, struktura za granu i struktura za graf. Grane i čvorovi se čuvaju u uređenim nizovima prema identifikacionim brojevima kako bi njihova pretraga bila ubrzana.

Proces kreiranja grafa se sastoji iz dva koraka. Prvi korak je prolazak kroz sve trase unutar jednog regiona. Za svaku trasu kreiraju se čvorovi i grane koje joj odgovaraju. Za pojedine čvorove vezuje se stanica koja im odgovara. Prilikom ovih iteracija svakoj stanici pridružuje se više čvorova. Ovaj korak rezultira grafom prikazanim na Sl. 4.a. Nad ovakvih grafom ne možemo da pronađemo put koji ima presijedanja. Ovaj problem rješava se korakom dva. Drugi korak prolazi kroz svaku stanicu i gleda koji su joj čvorovi pridruženi. Za svaku stanicu kreira se novi čvor i sve ulazne grane trenutno pridruženim čvorovima se usmjeruju ka novom čvoru i svim izlaznim granama se novi čvor postavlja kao početni. Na kraju ovaj čvor se dodaje u graf, a ostali čvorovi se brišu. Posle ovog koraka imamo graf prikazan na Sl. 4.b.



Sl. 4. a) graf posle prvog koraka b) graf posle drugog koraka

Pronalazak najkraćeg puta se vrši tako što se prvo pronađe graf koji odgovara identifikacionom broju regiona, a zatim se u tom grafu pronađu čvorovi kojima odgovaraju početna i krajnja stanica. Potom se nad grafom primjenjuje algoritam za pretragu najkraćeg puta i u ovom slučaju odabran je Dijkstrin algoritam. Najkraći put u

našem algoritmu definišemo na osnovu dva parametra: dužina puta izražena u metrima i broj presijedanja na tom putu. U toku pretrage konstruiše se i put. Prilikom konstrukcije potrebno je voditi računa da li je došlo do presijedanja, a to se postiže tako što se posmatraju čvorovi  $v_{i-1}$ ,  $v_i$  i  $v_{i+1}$ . Čvorovi koji leže na stanicama su označeni da ne pripadaju nekoj trasi. Treba provjeriti da li čvorovi  $v_{i-1}$  i  $v_{i+1}$  leže na istoj trasi i ako to nije slučaj znamo da je na stanici koja leži na čvoru  $v_i$  došlo do presijedanja. Odgovor se vraća API komponenti u JSON formatu, koja taj odgovor dalje prosljeđuje krajnjem korisniku.

Zbog ograničenosti prostora dat je samo pseudo kod primjene Dijkstrinog algoritma nad grafom opisanim u prethodnom dijelu teksta:

```
function omnibus_graph_dijkstra(G, source, end)
for each vertex v in G:
  dist[v->id] := INFINITY, prev[v->id] := UNDEFINED
  visited[v->id] := 0

create heap Q, path_exists := 0
dist[start->id] := 0, min_path := 0
add {start, 0, 0} to Q

while Q is not empty and visited[end->id] equals to 0:
  start, cost, nswap := extract(Q)
  min_path := entry->cost
  current_vertex := entry->current
  prev[current_vertex->id] := entry->prev[0]

  if current_vertex->id equals to end->id :
    path_exists = 1
    break

  visited[current_vertex->id] := 1
  for each neighbour v of current_vertex:
    if visited[v->id] equals to 0:
      dx := distance(current_vertex, v)
      if dx + min_path less then dist[v->id]:
        dist[v->id] := dx + min_path
        new_entry := {v, dist[v->id], nswap}
        new_entry->prev[0] := current_vertex->id
        new_entry->prev[1] := entry->prev[0]

        t0 := G->verticies[new_entry->prev[0]]
          ->trace_id
        t1 := G->verticies[new_entry->current->id]
          ->trace_id
        t2 := G->verticies[new_entry->prev[1]]
          ->trace_id

        if (t0 equals to 0) and (t1 not equal to t2):
          new_entry->nswap++

        add new_entry to Q

  if path_exists:
    P := omnibus_path_create(G, min_path, prev, end->id)

  return P;
end
```

Iako su čvorovi u grafu čuvaju u uređenim nizovima to ne utiče na performanse Dijkstrinog algoritma. Razlog smještanja čvorova u uređene nizove je direktan pristup čvoru i opisnim poljima koji su izvučeni iz baze.

Za naredne iteracije razvoja ovog servisa planirano je da se algoritam proširi tako da uzima u obzir trenutnu lokaciju autobusa i da pronalazi puteve na kojima stanice nisu povezane odnosno da će korisnik morati prepješačiti jedan dio puta.

#### IV. ALATI

U realizaciji rješenja upotrijebljeno je više savremenih tehnologija i alata. Alati i tehnologije su bili pažljivo birani kako se ne bi došlo do narušavanja performansi. Za pozadinski dio Web Aplikacije korišteni su web server

(apache ili nginx), PHP jezik, Laravel 5.1 ([12], [13]). Korisnički interfejs je rađen pomoću HTML, CSS i JavaScript jezika. JavaScript je dopunjen bibliotekom *jQuery* koja ubrzava razvoj i Leaflet koja omogućava interakciju sa kartama [14].

Za bazu podataka koristi se PostgreSQL i dopuna za geografske informacione sisteme PostGIS. Server za rutiranje je pisan u programskom jeziku C, i dodatno koristi biblioteku libpq za komunikaciju sa PostgreSQL serverom. Za dodatne algoritme i strukture (graf, vektor, prioritetni red, itd.) nisu korištene pomoćne biblioteke.

Laravel 5.1 je PHP framework odnosno skup biblioteka koji olakšava pisanje web aplikacija. Enkapsulira i abstrahuje kompleksne funkcionalnosti i koristi MVC (eng. *Model View Controller*) šablon za izradu aplikacija. Nastao je kao klon popularnog **Ruby On Rails**, a zatim je unaprijeđen na osnovu iskustva programera i sa sigurnošću u prvom planu. On diktira poslovnu logiku i strukturu koda. Pošto se Laravel nalazi u višim slojevima ove web aplikacije, to omogućava da se sigurnost implementira već u prvom sloju. Laravel rutira HTTP zahtjeve, parsira ih i odgovara odgovarajućim sadržajem. U izradi ovog projekta korištena je verzija Laravel 5.1 LTS. Iako u trenutku pisanja ovog rada postoje savremenije verzije, ova verzija će biti podržana narednih pet godina i svi propusti će biti dopunjeni, a pogotovo oni koji su vezani za sigurnost.

Leaflet je JavaScript biblioteka, koja je vodeća biblioteka otvorenog koda i omogućava korisnicima interakciju sa kartama na web stranicama. S obzirom da je njen kod otvoren, to omogućava programerima da proširuju i dopunjuju ovu biblioteku prema svojim potrebama kao i doprinose zajednici. Leaflet podržava skoro sve mobilne i desktop platforme kao i HTML5 i CSS3.

Za razliku od njene alternative OpenLayers, Leaflet biblioteka je razvijana sa ciljem da bude što jednostavnija, da nije prenatrpana velikim brojem mogućnosti, da je brza i da se lako koristi. Sadrži samo osnovne funkcionalnosti a dodatne se funkcionalnosti se naknadno instaliraju po potrebi. Jedan takav dodatak, koji je korišten u ovom projektu je **Leaflet.Draw** [15]. Ovaj dodatak omogućava jednostavnije crtanje objekata na kartama, prvenstveno onih koji su potrebni za predstavljanje autobuskih trasa i stanica na kartama.

## V. ZAKLJUČAK

Realizovan je servis koji građanima ili posjetiocima nekog grada pomaže u korištenju javnog gradskog prevoza (uz pomoć neke treće aplikacije, android, web aplikacije, itd.) Sve informacije su dostupne na jednom mjestu i moguće je izvršiti upit kako stići od stanice A do stanice B putem API komponente. Ovaj sistem je dalje moguće

unapređivati, kako dodavanjem novih mogućnosti tako i optimizacijom postojećih.

Iz izloženog možemo vidjeti da servis nije usko vezan samo za jedan grad već je moguće koristiti ga za više gradova, pa čak i država. Iako je za izradu potrebno dosta savremenih tehnologija, uviđamo da to ne predstavlja nikakav problem, čak nam dodatno olakšavaju posao.

Informatičko doba koje nam predstoji nudi veliki broj moćnih tehnologija sa kojima će biti moguće realizovati različite softvere koji će pomoći društvu.

## LITERATURA

- [1] P. Burrough, R.McDonnell, "Principles of Geographical Informations Systems", 1998, Oxford University Press.
- [2] P. Rigaux, M. Scholl, and A. Voisard, "Spatial Databases: With Application to GIS", 2002, Morgan Kaufmann Publishers.
- [3] PostgreSQL, Available: <https://www.postgresql.org/>
- [4] PostGIS, Available: <http://postgis.net/>
- [5] OGC Standards, Available: <http://www.opengeospatial.org/docs/is>
- [6] GeoJSON Specification, Available: <http://geojson.org/geojson-spec.html>
- [7] M. Čabarkapa, "C – osnovi programiranja", 1996. Krug, Beograd.
- [8] Brian Hall, "Beej Guide to Network Programming", Brian "Beej Jorgensen" Hall, 2016.
- [9] M. V. Tomašević, "Algoritmi i strukture podataka", 2008, Akademski misao, Beograd.
- [10] Robert Sedgewick, "Algorithms in C", 1990, Addison-Wesley Publishing Company.
- [11] PostgreSQL: Documentation 9.1: libpq – C Library, Available: <https://www.postgresql.org/docs/9.1/static/libpq.html>
- [12] PHP Manual, Available: <http://php.net/manual/en/index.php>
- [13] Laravel 5.1 LTS, Available: <https://laravel.com/docs/5.1/>
- [14] Leaflet Documentation, Available: <http://leafletjs.com/reference.html>
- [15] Leaflet.Draw Documentation, Available: <https://github.com/Leaflet/Leaflet.draw>

## ABSTRACT

Planning how to get from location A to location B in particular city, respectively finding shortest path between two given points, is an everyday problem one (citizen or tourist) have to think about. One solution to this problem can be a service which will hold all information about public transport and answer to queries from clients. One such a solution is described in this paper, and it uses many present-day technologies and consists of few tightly coupled components. Most of the code is written in PHP5 using Laravel framework and the part of service which handles queries for finding shortest paths is written in C language. Apart from giving solution to the problem, this paper also aims to show how to use modern technologies in order to develop a robust Web application.

## GEOLOCATION SERVICE FOR PUBLIC TRANSPORT

Lazar Radinović, Aleksandar Popović, *Prirodno-matematički fakultet, Univerzitet Crne Gore*

# Deep learning techniques for classification of handwritten digits

Pavlović Kosta and Šuković Goran, *Member, IEEE*

**Abstract** — This paper discusses motivation and principles of constructing learning algorithms for deep learning neural networks. Performances of deep neural network which consists of stacked autoencoder and softmax classifier and performances of convolutional neural network have been compared. Performances were tested on the problem of classification of handwritten digits. Handwritten digits used for training and testing algorithms are taken from MNIST database. All applications are written in MATLAB, on computer with AMD A8 (2.1 GHz) processor with 8 GB of RAM.

**Keywords** — Autoencoder, convolutional neural networks, deep learning, softmax classifier.

## I. INTRODUCTION

CLASSIFICATION of handwritten digits is one of the most important tasks in computer vision. A large number of state of the art methods such as k-nearest neighbors (KNN), linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), Gaussian Mixture Models (GMM), and Support Vector Machines (SVM) have been used of classification of handwritten digits. An overview of the results obtained by these methods using MNIST database is provided in [1]. In this paper, algorithms of deep learning have been used in solving the same task.

Deep learning algorithms have proven to be very successful in the fields of text, images and speech recognition. A large number of ideas for development of deep learning algorithms came from biology and the way human brain works.

Depth of biological neural networks has inspired scientists and researchers who have spent years trying to train deep neural networks, but there was no success until 2006. A turning point came in 2006 when Geoffrey E. Hinton, Simon Osindero i Yee-Whye Teh introduced Deep Belief Networks (DBNs) [2], with greedy algorithms that trains one by one level in neural network. Soon after this, similar algorithms based on autoencoders have been developed ([3], [4]). The key principle is to combine algorithms of unsupervised and supervised learning.

Pavlović Kosta – Faculty of Mathematics and Natural Sciences, Podgorica – University of Montenegro, George Washington st. bb, Podgorica 20000, Montenegro (phone:+38269-528-960, e-mail: [kosta.pavlovic994@gmail.com](mailto:kosta.pavlovic994@gmail.com))

Šuković Goran – Faculty of Mathematics and Natural Sciences, Podgorica – University of Montenegro, George Washington st. bb, Podgorica 20000, Montenegro, (phone: +38269-376-669, e-mail: [ai@rc.pmf.ac.me](mailto:ai@rc.pmf.ac.me))

Unsupervised learning is used for initial training (pre-training), and supervised learning is used for fine-tuning.

Performances of deep learning algorithms greatly depend on the representation of input data. Varieties of algorithms are used to learn data representation. Using these algorithms we attempt to learn data representations at a high level of abstraction. Then we use these representations for solving desired task.

Ideas of deep learning are usually implemented over neural networks. It is said that by the introduction of deep learning, rebranding of neural networks has been performed.

## II. MNIST

MNIST (Mixed National Institute of Standards and Technology database) is a database of handwritten digits which can be found on Internet [5]. It contains training set with 60 000 examples and test set with 10 000 examples. This database is actually a subset of larger set published by NIST (National Institute of Standards and Technology) which had to be rearranged for number of reasons. Size of images in database is 28 x 28..

## III. AUTOENCODER

Basic example of an algorithm that learns data representation is autoencoder ([3], [4]). Autoencoders are trained to learn good data representation using as less memory as possible. Different types of autoencoders manage to learn different properties of the input data and they are common in deep learning neural networks.

Autoencoder is unsupervised learning algorithms that applies backpropagation and tries to find an approximation of the input data. Although this function seems easy to learn, by placing some restrictions on the network, as limiting the number of hidden nodes, interesting properties of the data can be discovered or we can come up with some compressed representation of the input data. There are other types of interesting restrictions, as for example request for a small number of active neurons (sparsity). This constraint is motivated by the way human brain works, where at any given moment only small fractions of neurons are activated. Intention is to achieve the same effect with autoencoder.

We wish the fault of hypothesis, given by autoencoder, to be as low as possible so we minimize the cost function:

$$I_{\text{sparse}}(W, b) = \frac{1}{2m} \sum_{i=1}^m \|h_{W,b}(x^{(i)} - y^{(i)})\| + \frac{\lambda}{2} \sum_{l=1}^{n_l-1} \sum_{i=1}^{S_l} \sum_{j=1}^{S_{l+1}} (W_{ij}^{(l)})^2 + \beta \sum_{i=1}^m \left( p \log \frac{p}{\hat{p}_j} + (1-p) \log \frac{1-p}{1-\hat{p}_j} \right) \quad (1)$$

Sigmoid function is chosen to be activation function of neurons  $f$  and gradient descent method is used for training.

The first addend in the formula is obviously average of cost of individual examples, while the second addend is regularization term, whose role is to prevent phenomenon of overfitting. Parameter  $\lambda$  is called **weight decay parameter** and it controls the “importance” of the second addend. Parameter  $\beta$  controls the impact of third addend which “punishes” derogation from sparsity which is required,  $p$  is **sparsity parameter** ( $p$  has small value, e.g.  $p=0.05$ ). Autoencoder which has demand for minimization of the number of active neurons is called sparsity autoencoder.

- $x^{(i)}$  – input vector. In our case it is vector with  $28*28=784$  coordinates, each of which corresponds to one pixel of the image.
- $y^{(i)}$  – output vector for input  $x^{(i)}$ . Autoencoder tries to learn input data, so  $y^{(i)}=x^{(i)}$ .
- $h_{W,b}(x^{(i)})$  – hypothesis made by our network. With *autoencoder* we try to learn  $h_{W,b}(x) \approx x$
- $n_l$  – number of layers of the neural network
- $L_1, \dots, L_{n_l}$  – layers
- $S_l$  – number of nodes in layer  $L_l$
- $W_{ij}^{(l)}$  – weight of connection between node  $j$  in the layer  $L_l$  and node  $i$  in the layer  $L_{l+1}$
- $b_i^{(l)}$  – bias of node  $i$  in the layer  $L_{l+1}$
- $a_i^{(l)}$  – activation (output) of node  $i$  in the layer  $L_l$  (for  $l=1, a_i^{(l)}=x_i$ )
- $m$  – number of examples in training set
- $\hat{p}_j$  – average activation of hidden nodes

$$\hat{p}_j = \frac{1}{m} \sum_{i=1}^m (a_j^2(x^{(i)})) \quad (2)$$

In the context of this paper sparse autoencoder was implemented. It is used to create deep neural network in Section IV.

#### IV. DEEP NETWORK

Networks with multiple hidden layers are able to learn hidden features of input data. Each of layers learns nonlinear function of its inputs. In the case of images recognition, first layer performs edge detection, second layer could group those edges to detect more complex contours or some simple parts of objects and deeper layers could, by grouping contours, learn very complex shapes. For example, sparse autoencoder described in previous section manages to learn features of images that look like pencil strokes (Figure 1).

Although deep networks are present in theory for a long period of time, until recently there was no success in their training. Traditional approach to start with random set of

weights, and then to minimize error using gradient descent or L-BFGS [6] (Limited-memory Broyden–Fletcher–Goldfarb–Shanno) algorithm, has not lead to significant results.

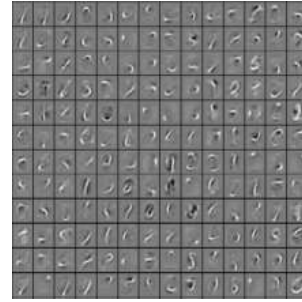


Fig. 1. Pencil strokes that sparse autoencoder manages to learn

Greedy layer-wise method ([2], [7]) learns parameters of each layer in network separately, while freezing the rest of the model’s parameters. Training can be supervised, but it is more often unsupervised (as with autoencoder). At the end of this process, weights  $W^{(l)}$  from all layers are used to initialize weights in the final network and then fine-tuning is performed.

Fine-tuning ([6], [8]) is a procedure that is essentially an application of the gradient descent from the point which corresponds to the current values of the parameters and it is performed over the labeled set. It is empirically shown that the gradient descent from this point converges to good local minimum with higher probability.

Stacked autoencoder ([6], [8]) is a neural network consisting of multiple layers of sparse autoencoder in which the output of each layer is connected to the input of the next layer. Stacked autoencoder has all characteristics of a deep neural network. Training of stacked autoencoders is performed using greedy layer-wise method.

For classification of images from MNIST database we constructed stacked autoencoder with two hidden layers. The last layer of autoencoder is connected with softmax classifier (Figure 2).

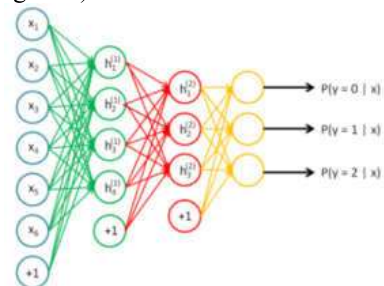


Fig. 2. A simplified scheme of classifier with reduced number of nodes (source [6])

*Softmax regression* is supervised learning algorithm which is a generalization of logistic regression algorithm. It is used for classification problems when there are more than two classes [9].

The last step in the training of this network is to fine-tune the classifier. This significantly improves the performance of stacked autoencoders. Fine-tuning considers all the layers in stacked autoencoders together, and in one iteration changes all parameters of model.

The form of cost function in this algorithm is as follows:



$$J(W, b) = \frac{\lambda}{2} \sum_{i=1}^{k-1} \sum_{j=0}^n \theta_{ij}^2 - \frac{1}{m} \sum_{i=1}^m \sum_{j=1}^k 1\{y^{(i)} = j\} \log \frac{e^{\theta_j^T a^{(n)}(x^{(i)})}}{\sum_{p=1}^k e^{\theta_p^T a^{(n)}(x^{(p)})}} \quad (3)$$

- $x^{(i)}$  – input vector (in this case image of size 28x28)
- $y^{(i)}$  – class label which the input  $x^{(i)}$  belongs (in this case these are digits from 0 to 9)
- $1\{y^{(i)} = j\}$  – a term that is equal to 1 if the label of  $i$ -th element in the training set is  $j$  ie. if it belongs to class  $j$ . Otherwise, this value is equal to 0.
- $\theta_{ij}$  – parameters of softmax classifier
- $\lambda$  – weight decay parameter

This algorithm has been implemented in MATLAB and tested using MNIST database. Code vectorization has been performed, and thus the algorithm has been further accelerated. All the advantages of MATLAB built-in operations of linear algebra have been used (matrix addition, matrix multiplication, vector and matrix multiplication, etc.). Values of parameters  $\lambda$ ,  $\beta$  i  $p$  have been taken according to recommendations from [6]. Training has been performed on a set of 60 000 examples through 400 iterations. Hidden layers of both autoencoders contain 200 nodes. Time to train first autoencoder is 120 minutes. Training of second autoencoder lasts 40 minutes. Time for training the softmax classifier is short and does not affect total time for training the network. Before fine-tuning process classifier has been tested on a set of 10 000 examples. The results are presented in Table 1.

TABLE 1: REVIEW OF THE RESULTS

Network	Training time	Predictive accuracy
Two-layer stacked autoencoder with softmax classifier without fine-tuning	160 min	87.69%
Two-layer stacked autoencoder with softmax classifier with fine-tuning	220 min	97.61%

V. CONVOLUTIONAL NEURAL NETWORKS

The network built in the previous section had completely linked layers. Number of parameters that have to be learned in these kinds of networks is quite large. With relatively small images from MNIST database (size 28x28), the training of the network could be done within a reasonable period of time. However, for the problems that have more complex input data, usage of networks with exclusively fully connected layers is not suitable. In order to solve this problem, locally connected layers are introduced. Neurons of locally connected layers are connected only with a small group of neighboring neurons in the preceding layer. The idea for the introduction of locally connected layers originates from biology, ie. from

the way visual system is organized. Neurons in the visual cortex have localized fields of receptors and respond only to stimuli in a particular location [10].

Using convolution we can learn some features of an image. Features of images that are obtained by convolution in combination with a fully connected network such as the softmax classifier can be used for classification. Layers whose neurons perform convolution of the output of the preceding layer and the appropriate filter are called convolutional layers. Activations of neurons in the convolutional layer are calculated as follows:

$$a^{(l)} = f(W^{(l)} * a^{(l-1)} + b^{(l)}) \quad (4)$$

Matrix  $W^{(l)}$  determines filter,  $a^{(l-1)}$  is vector of activations of neurons in the layer preceding the convolutional. Operation „\*“ represents convolution (in this case it is a so-called valid convolution [9], that ignores points where the filter exceeds borders of the image with which it is being convoluted). Activation function  $f$  is sigmoid function,  $b^{(l)}$  are bias inputs of convolutional layer. Vector of activations of neurons in convolutional layer is  $a^{(l)}$  and it has  $k \times (m-f+1) \times (m-f+1)$  coordinates, because we assumed input dimensions  $m \times m$  and that we have  $k$  filters of size  $f \times f$ . In order to train the convolutional layer, it is necessary to learn the parameters  $W^{(l)}$  i  $b^{(l)}$ . Number of parameters is reduced comparing to networks with fully connected layers. Neurons that use same filter for convolution have the same parameters. Therefore, the parameters are related to the filters which are small in size.

If we want to attach classification layer to convolutional layer, it means that we have to train a classifier with a large number of input nodes. This architecture is unwieldy and prone to overfitting ([10], [11], [12]), so in the case of large input data natural approach is to perform sub-sampling [10]. New layers that perform sub-sampling are introduced. They split the output of preceding layer in regions, and then apply an aggregate function over each of these regions. Aggregate functions can be for example the mean or maximum in a given region. In this way we improve network’s performance and we reduce training time. In this paper mean of activations of neurons in particular region has been used as aggregate function for sub-sampling. Figure 3 shows how sub-sampling works.

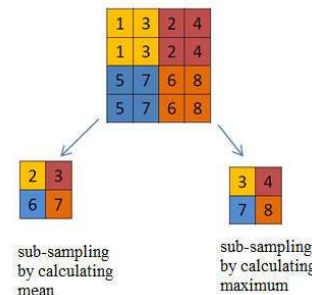


Fig. 3. Two ways to perform sub-sampling (source [13])

Convolutional neural networks ([10], [11], [12]) consist of one or more convolutional layers with layers for sub-sampling, followed by one or more fully connected layers. Convolutional networks are designed to take advantage of two-dimensional images, or other two-dimensional signals that they classify. Another advantage of convolutional

neural networks is that they have many fewer parameters than the networks with fully connected layers, so they are trained faster. Figure 4 is a sketch of a convolutional network.

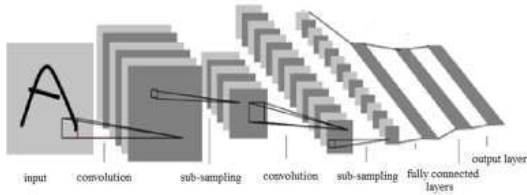


Fig. 4. Scheme of convolutional neural network ([11])

Methods such as L-BFGS which, in one step, use the entire training set for modifying parameters, converge to good local minima. However, in practice, the calculation of gradients using the complete training set is very slow. Also, we can not easily expand training set for networks that are trained by these methods. The method of stochastic gradient descent [10] updates parameters using a small fraction of the training set. Stochastic gradient descent overcomes all mentioned problems and leads to faster convergence to the local minimum. For the method of stochastic gradient descent it is important in which order data is taken from the training set. A reasonable distribution of data can adversely affect the gradients and disrupt convergence. Therefore, the data in the training set must be shuffled at the beginning. Learning through stochastic gradient descent can initially be very unstable because it depends on the parts of training set which can be very different, but it stabilizes by time.

Convolutional network for classification of handwritten digits has been implemented. The network consists of one convolutional layer and sub-sampling layer which are followed by a fully connected layer that performs softmax regression. Convolutional layer has 20 filters of size  $9 \times 9$ , and size of region for sub-sampling is  $2 \times 2$ . Training time and accuracy are showed in Table 2.

TABLE 2: COMPARISON OF CONVOLUTIONAL NETWORK AND SOFTMAX CLASSIFIER

<i>Network</i>	<i>Training time</i>	<i>Predictive accuracy</i>
Two-layer stacked autoencoder with softmax classifier with fine-tuning	220 min	97.61%
Convolutional neural network	30 min	97.21%

Results clearly show advantages of convolutional networks. Training time for convolutional neural network is much shorter, and almost same predictive accuracy is achieved.

## VI. CONCLUSION

In this paper two classifiers for handwritten digits have been implemented. The first is a deep neural network consisting of a two-layer stacked autoencoder and softmax classifier, and the other is a convolutional neural network.

Although this paper focused on deep neural networks, the ideas of deep learning should be explored outside of the domain of neural networks (eg. with decision trees and decision networks). Also, it is desirable to try some other deep learning algorithms, such as Deep Boltzmann machines. It is also possible to use genetic algorithm for training deep neural networks.

## REFERENCES

- [1] Ming Wu, Zhen Zhang. "Handwritten Digit Classification using the MNIST Data Set" *Pattern Classification & Analysis, MSU, Spring 2010*.
- [2] G. E. Hinton, S. Osindero, and Y. Teh, "A fast learning algorithm for deep belief nets," *Neural Computation*, vol. 18, pp. 1527–1554, 2006.
- [3] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, "Extracting and composing robust features with denoising autoencoders," in *Proceedings of the Twenty-fifth International Conference on Machine Learning (ICML'08)*, pp. 1096–1103, 2008.
- [4] Dumitru Erhan, Yoshua Bengio, Aaron Courville, Pierre-Antoine Manzagol, Pascal Vincent, and Samy Bengio, "Why does unsupervised pre-training help deep learning?" *Journal of Machine Learning Research*, vol. 11 pp. 625–660, 2010..
- [5] MNIST database <http://yann.lecun.com/exdb/mnist/>
- [6] Andrew Ng, "CS294A/CS294W Deep Learning and Unsupervised Feature Learning", 2011.
- [7] Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle "Greedy layer-wise training of deep networks." *NIPS. MIT Press*, 2007.
- [8] Y. Bengio, "Learning Deep Architectures for AI" *Foundations and Trends in Machine Learning* vol. 2, pp. 1–127, 2009.
- [9] Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani, "An Introduction to statistical learning" *Springer Science+Business Media New York*, 2013.
- [10] "[Convolutional Neural Networks \(LeNet\) – DeepLearning 0.1 documentation](#)" *DeepLearning 0.1. LISA Lab*, 2013
- [11] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. "Gradient based learning applied to document recognition." *Proceedings of the IEEE*, pp. 2278–2324, 1998.
- [12] F.-J. Huang and Y. LeCun. "Large-scale learning with svm and convolutional nets for generic object categorization." *CVPR. IEEE Press*, 2006
- [13] [blogqpot.com/images/matconvnet](http://blogqpot.com/images/matconvnet)

# Applying NEAT on the game 2048

Boris Tuponja and Goran Šuković, *Member, IEEE*

**Abstract** — This paper introduces neuroevolutionary algorithm NEAT and its application on game 2048, with aim of learning neural networks how to play it. Over the course of several hundred generations, the network evolves and learns how to play game.

**Keywords** — Neuroevolution, NEAT, 2048, neural networks, evolutionary algorithms.

## I. INTRODUCTION

THOUGHT history games, beside their entertainment factor, represented sort of mental competitions and challenges, such as chess and Go. From the earliest days of computing, there is desire for creation of artificial agent that would be able to learn and surpass humans in their own games. Motivated by this, great number of methods from the field of Computational Intelligence has been applied to games. Even though this field is well established, it is still rapidly developing. Neuroevolution [1]-[4] is one of popular tools. It is a group of methods that banded together neural networks and evolutionary algorithms. Neuroevolution of augmenting topologies (NEAT) [5] is member of this group. This paper represents application of this algorithm on the game 2048 [6]. 2048 is puzzle game made of board with 16 fields, which can have tiles with some values (Figure 1). Aim of the game is to merge two neighbouring tiles with same values to create one with double the value and receive some points. In the game, there are 4 available actions, each moves all tiles on the board in one of following directions: up, down, left, right. After every move, on a free field a new tile emerges with a value 2 or 4. The game ends when there are not any possible moves. At the start, there are just 2 tiles on the board.

We wanted to create artificially intelligent agent that could learn to play this game. Inspiration for this application came from the fact that neuroevolution is important method that has seen continued popularity since its inception and there are successful applications of neural network on game Super Mario [3]-[4].

NEAT was effectively used in various games. Examples are NERO (NeuroEvolving Robotic Operatives) [7]-[9]

Tuponja, Boris, Faculty of Mathematics and Natural Sciences, Podgorica – University of Montenegro, George Washington st. bb, Podgorica 20000, Montenegro (phone: +382-67-863-811, e-mail: [boris@mnet.me](mailto:boris@mnet.me))

Šuković, Goran, Faculty of Mathematics and Natural Sciences, Podgorica – University of Montenegro, George Washington st. bb, Podgorica 20000, Montenegro (phone: +382-69-376-669, e-mail: [goran.sukovic@gmail.com](mailto:goran.sukovic@gmail.com))

and Galactic Arms Race [10]-[12]. Also, NEAT was used in game Go [13].

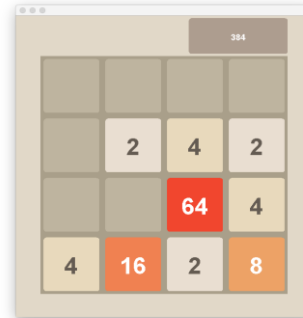


Fig. 1 – Game board after few played moves

## II. NEAT

Idea behind NEAT is to make the most of minimizing the dimensionality of the search space of connection weights [5]. Most algorithms involving neural networks work with previously specified topology and change only network weights. Usually, it is unclear how network should look in order to show acceptable performance on required task.

To enhance network performance, NEAT algorithm tries to start with minimal network and to augment it only when and where it leads to better results. This kind of neural networks are known as TWEANNs (Topology and Weight Evolving Artificial Neural Networks).

Another important fact about TWEANNs is type of encoding that they use. There are two types of encodings - direct encoding and indirect encoding. Direct encoding uses one-on-one mapping from genotype to phenotype (network), i.e. for every node of network there is one gene in genotype. Indirect encoding, on the other hand, has more complex rules for these mappings. NEAT uses direct encoding.

Features that make NEAT great are ones that were not used in other NE algorithms. These features are historical origins and speciation [5].

Historical origins are important for problem known as *Permutation Problem* [5]: two networks that represent same solution have different distinct genome encoding. It is the problem, because reproducing these networks could produce flawed descendant incapable of solving intended task. To avoid this, NEAT uses historical origins. If two genes have same historical origin, then these genes have same ancestor and only one of them would be transferred to an offspring, resulting in correct reproduction.

Tracking historical origins is fairly easy. When a new gene appears through topological mutation, global innovation counter is incremented and value is assigned to the new gene. This is the straightforward method, which

enables following order of gene appearance and eases process of reproduction.

During reproduction of two organisms, innovation numbers line up their genes. The genes that have a same innovation number are called matched genes and one of them, randomly chosen, will be transferred to an offspring. The genes whose innovations numbers do not match are either excess or disjoint, depending on their positions. If they appear within the range of the other parent's innovation numbers, then they are disjoint. Otherwise they are excess. They represent structure not present in the other parent. Both disjoint and excess genes are included in an offspring from the fit parent.

Through this kind of reproduction, it is enabled to create different topological structures, but adding new nodes usually results in the network's initial decrease of the fitness. Moreover, larger networks need more time than smaller networks to be optimized, so newly augmented networks have little chance to survive more than one generation, even though these new structures could be crucial in long run.

To prevent premature extinction NEAT uses speciation. The network competes primary with the networks in its specie instead against the whole population. Idea is to protect unseen topology structures in new specie and allow them to optimize through competition only in that specie. This way population is divided in groups in which members are similar. For calculating compatibility of networks NEAT uses linear combination of matching, disjoint and excess genes as following [5]:

$$\delta = \frac{c_1 E}{N} + \frac{c_2 D}{N} + c_3 \overline{W} \quad (1)$$

where  $\delta$  is distance between two networks,  $E$  is number of excess genes,  $D$  is number of disjoint genes,  $\overline{W}$  is average weight difference of matched genes and  $N$  is number of genes in larger genome. Coefficients  $c_1$ ,  $c_2$ ,  $c_3$  allow us to adjust how big impact has each factor. If  $\delta$  is less than some threshold  $\delta_t$ , we can assume that those genomes belong to same species.

NEAT keeps an ordered list of the species. In each generation, the new genomes are inserted in the first species that satisfies the threshold. Each existing specie is represented with one randomly chosen genome in the specie from previous generation. If the new genome does not find appropriate specie, new specie with that genome is created.

Using this mechanism there is risk that one species became too big and dominate whole population. To avoid this problem, NEAT uses explicit fitness sharing. These values are used during genome reproduction. The adjusted fitness for every genome is calculated by following equation [5]:

$$f'_i = \frac{f_i}{N_i} \quad (2)$$

where  $f_i$  is adjusted fitness of  $i$ th genotype and  $N_i$  is the number of genotypes in its specie.

During reproduction, every species is assigned a potentially different number of offspring in proportion to the sum of adjusted fitness of its members. After the numbers of offspring are defined, the least fit organisms are removed from population, but only if they are older than some fixed number of generations, because they had some time to adjust to problem and failed. After filtering which organisms stay in population, they receive part of the offspring granted to its specie in proportion to the original fitness value.

Most TWEANN algorithms start with number of randomly generated networks, but NEAT have different approach. All starting networks do not have hidden nodes. Differences between networks are the weights of the edges. Complexity of the neural network is increased only through topology mutations. This is an advantage because parameter search field is minimized. Combination of all these features made NEAT very powerful and attractive tool. It recorded better results than any other neuroevolutionary algorithm on a set of experimental problems. Considering all of these features and results, we decided to apply NEAT on the game 2048.

### III. RECURRENT EDGES

Most neural networks take inputs as independent entities, but sometimes this is not a correct approach. For example, in systems that predict next word, it is crucial to know which word had come before it. To adjust for this, sometimes it is needed for previous network outputs affect current network state. It could be said the network needs some kind of memory where it could store the information. Recurrent edges come to rescue (Figure 2). A recurrent edge is same as any other edge with distinction that it has impact on next input of neuron. Theoretically, the networks could keep great number of previous states, but in practice only few are remembered.

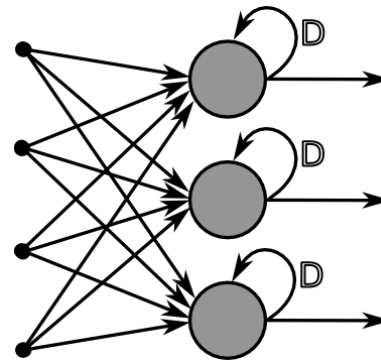


Fig. 2 – Fully connected recurrent neural network

### IV. NEAT AND 2048

For experimenting with this algorithm for game 2048, we used library ANJI (Another Java NEAT Implementation) [14]. This library offers some great tools for logging and statistics and because of that we decided to use it over native NEAT implementation written by Stanley and Miikkulainen ([5]). Also, for easier tracking of

game, we wrote GUI for the game, which was presented on Figure 1.

We created multiple experiments that included changes in parameters, neuron functions, and mappings from board to the network inputs. Here are presented just a few of our results.

All of the experiments were implemented on following hardware: Intel i7 4790 3.6 GHz processor with 4 cores and 32 GB DDRAM III. During training, the executions were headless, i.e. without GUI. The neural networks were made of 16 input neurons for the fields on the board and 4 output neurons for possible moves. Some parts of the experiments were executed in parallel, which led to CPU load of 70% and 40% of RAM.

During evolution, every neural network was tested against 10 randomly created boards and the fitness was average score of those games.

Parameters for presented experiments are given in table 1. We used normal distribution of initial weights. Input neurons used linear activation function i.e. they just outputted what they received, but the hidden and the output neurons used sigmoid function.

TABLE 1: FIRST GROUP EXPERIMENT PARAMETERS.

Parameter	Value
Number of generations	5000
Population size	200
Add edge mutation probability	0.08
Add neuron mutation probability	0.05
Weight mutation probability	0.8
Coefficient of excess neurons	0.5
Coefficient of disjoint neurons	0.5
Coefficient of common neurons	0.2
Speciation threshold	0.3
Weight of edges	[-500, 500]
Standard deviation	1.5

For the network input we did not use values shown on board, but we calculated  $\log_2$  of those values, then normalized them before inputting them. Because all tile values are powers of 2, we wanted to reduce the exponential growth to linear, hereof using  $\log_2$ . Moreover, we normalized all input values to  $[0,1]$ .

During game playing, the network played move related to the output neuron with the highest output. If two or more neurons had same value, we chose randomly between them. The game ends when there are no more possible moves or the outputted move is invalid. By ending the game on the first invalid move, we wanted to push neural networks to perform the best they could and know that wanted move could not be performed.

The results of this experiment are shown on Figure 3. Green line presents maximum average score of 10 played games that some network recorded at the specific generation. Average (blue) and minimum (red) scores are similar

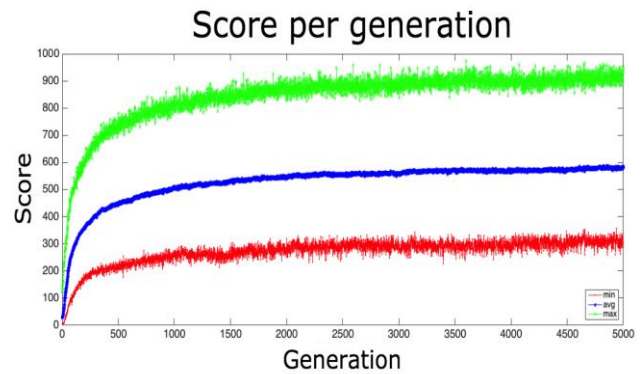


Fig. 3 – Graph of scores per generation

On the figure 4 is presented a graph with number of species per generation, but only for the first 200 generations for clarity, because there are no changes in later generations. The number of species was around 100, which means that on average 2 networks were in same species.

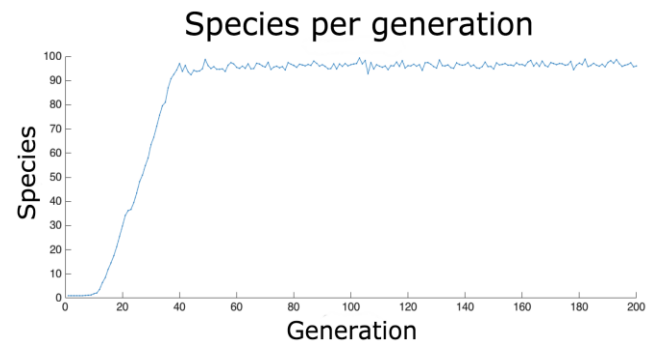


Fig. 4 – Graph of number of species per generation

The figure 5 presents average complexity of the neural networks. The initial complexity is 84, which stands for 20 neurons and 64 edges between them, because the initial neural networks were fully connected. The complexity increases linearly.

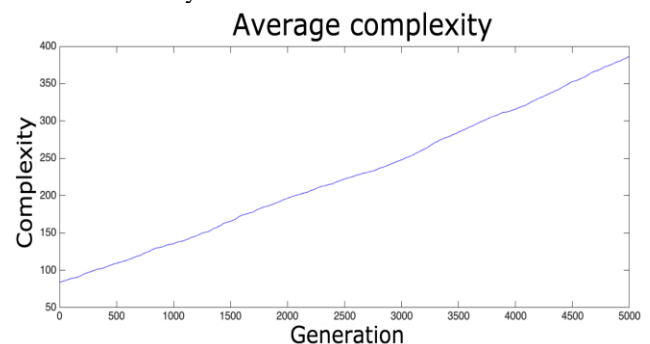


Fig. 5 – Complexity of neural networks

In the other experiment, we used same settings as before, but introduced recurrent edges that could remember one previous state of the board. The number of species and the complexity was the same, but the result increased almost 2 times, as shown in figure 6.

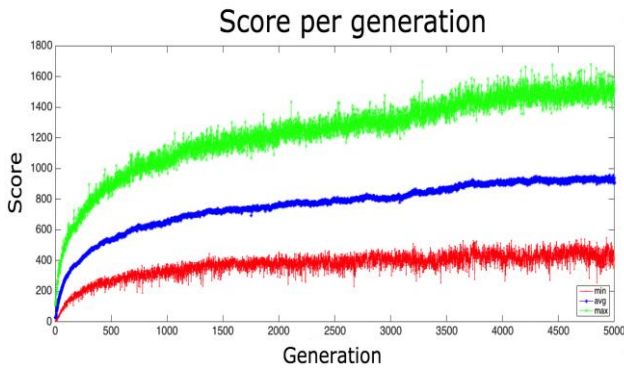


Fig. 6 - Graph of scores per generation for the experiment with recurrent edges

Figure 6 shows that the best-recorded score was around 1600 points, but it had slight tendency of future increase. This is still number of points that human could achieve from without much difficulty, but it is a great improvement for the neural networks. We try to increase number of cycles in order to find networks that could remember multiple previous states, but we did not have much improvement.

After executing experiments, we come across with similar simulation [15]. Simulation presented in [15] uses NEAT algorithm with different kind of network encoding.

## V. CONCLUSION

Paper presents first experiments with NEAT algorithm in order to evolve recurrent neural network capable to learn how to play game 2048. Stochastic nature of the game 2048 and random choosing of neuron when there are multiple outputs with the same greatest value creates very challenging environment for evolution of the neural networks. On the other hand, in the Super Mario level is always the same, so the evolving networks have a lot of space to optimize.

Even though we were not successful to achieve score that would come close to a human score or surpass it, this was just a first step towards that. There are other versions of this algorithm, which use different type of encoding that could better adapt to complex situations, such as the game 2048. Our further steps would be trying different parameter combinations, using other kind of encodings, combination of multiple neural networks etc.

## REFERENCES

- [1] Floreano, D.; Dürr, P.; Mattiussi, C.: *Neuroevolution: from architectures to learning*, C. Evol. Intel. (2008) 1: 47. doi:10.1007/s12065-007-0002-4
- [2] Sher, Gene I.: *Handbook of Neuroevolution Through Erlang*, Springer, 2013
- [3] Togelius, J.; Shaker, N.; Karakovskiy, S.; Yannakakis N.G. *The Mario AI Championship 2009-2012*. AI Magazine, 34(3), 89-92, 2013
- [4] Togelius, J.; Risi, S.: *Neuroevolution in Games: State of the Art and Open Challenges*, arXiv:1410.7326
- [5] Stanley, Kenneth O.; Miikkulainen, Risto: *Evolving Neural Networks through Augmenting Topologies*, MIT Press, 2002
- [6] Game 2048: [https://en.wikipedia.org/wiki/2048\\_\(video\\_game\)](https://en.wikipedia.org/wiki/2048_(video_game))
- [7] Stanley, Kenneth O.; Bryant, Bobby D., Miikkulainen, Risto: *Evolving Neural Network Agents in the NERO Video Game*, Proceedings of the IEEE 2005 Symposium on Computational Intelligence and Games (CIG'05), 2005
- [8] Stanley, Kenneth O.; Bryant, Bobby D., Miikkulainen, Risto: *Real-time Neuroevolution in the NERO Video Game*, IEEE Transactions on Evolutionary Computation, 653-668, 2005
- [9] NERO. Available: <http://nerogame.org/>
- [10] Hastings, Erin J.; Guha, Ratan K.; Stanley, Kenneth O.: *Evolving Content in the Galactic Arms Race Video Game*, Proceedings of the IEEE Symposium on Computational Intelligence and Games (CIG09), 2009
- [11] Hastings, Erin J.; Guha, Ratan K.; Stanley, Kenneth O.: *Automatic Content Generation in the Galactic Arms Race Video Game*, IEEE Transactions on computational intelligence and AI in games, 2009
- [12] Galactic Arms Race. Available: <http://galacticarmsrace.blogspot.com>
- [13] Stanley, Kenneth O.; Miikkulainen, Risto: *Evolving a Roving Eye for Go*, Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2004), 2004
- [14] ANJI library. Available: <http://anji.sourceforge.net/>
- [15] YouTube video. Available: <https://youtu.be/jsVnuw5Bv0s>

# Četiri ekstrema vezana za sudoku

mr Aleksa S. Srdanov, dr Vladan R. Đulaković, Dragan M. Milovanović, Visoka tehnička škola strukovnih studija Požarevac.

**Sadržaj** — Prilikom rešavanja igre sudoku nameću se barem četiri osnovna problema. Dva problema su vezana za minimalnu i maksimalnu početnu postavku u smislu broja početno datih cifara. Pri tome početna postavka predstavlja onu u okviru koje se ne može naći manja postavka koja bi takođe dovela do istog jedinstvenog rešenja. Treći problem je vezan za minimalan broj potrebnih strategija za rešavanje proizvoljnog sudokua, a četvrti za minimalan broj poteza koje je potrebno napraviti da bi se do rešenja došlo. Ovaj rad daje odgovore na sva četiri pomenuta pitanja.

**Ključne reči** — sudoku, ekstremi sudokua.

## I. UVOD

Za rešavanja sudokua moguće je koristiti relativno veliki broj strategija. Mnoge od njih su u suštini identične i samo prividno različite. U ovom radu se problem rešavanja sudokua razmatra sa stanovišta računara. Neke strategije potrebne za rešavanje sudokua uz pomoć računara za čoveka i nisu strategije. Rešavanje sudokua počinje zadavanjem svih mogućih vrednosti u poljima koja nisu određena početnom postavkom. Zatim bi se primenila strategija kod koje se korišćenjem početne postavke, zbog uslova jedinstvenosti sudokua, uklone sve već zadate cifre u svim odgovarajućim: vrstama, kolonama i kvadratima. Tada se na nekoj poziciji može pojaviti jedinstveni kandidat. Za čoveka je to završetak dalje pretrage, jer ta cifra postaje odmah rešenje na toj poziciji. Ali da bi računar izveo zaključak da se na nekom mestu pojavio takav slučaj mora dodatno izvršiti odgovarajuću proveru. Zbog toga, broj potrebnih strategija za računar biće nešto veći nego broj strategija koje bi za rešavanje sudokua koristio čovek. Zato, sa stanovišta pisanja računarskog algoritma, opravdano se postavlja pitanje: „Da li postoji minimalni skup strategija kojima se svaki sudoku može rešiti i koje bi, kao takve, bilo potrebno implementirati u opšti algoritam?“

Prilikom rešavanja sudokua mnogi algoritmi proveravaju iste cifre na istim pozicijama više puta u toku rada algoritma. Tako se gomila broj poteza koje treba uraditi da bi se došlo do krajnjeg rešenja. S toga je opravdano pitanje: „Da li je moguće sastaviti opšti

algoritam za rešavanje sudokua koji se odvija u minimalanom broju operacija?“ Dakle, sastaviti algoritam koji koristi uslov jednoznačnosti i ne proverava grane koje ne vode ka rešenju.

Poznato je da sudoku ukrštenice imaju različite polazne postavke u smislu broja datih cifara. Interesantno je da postoje sudoku rebusi koji imaju: recimo, početno dato, sedamnaest ili dvadeset osam cifara i imaju jedinstveno rešenje. Ali ako se iz takve početne pozicije ukloni bilo koja od datih cifara sudoku više nema jedinstveno rešenje. Postavlja se opravdano pitanje: „Koji je to minimalan, a koji maksimalan broj početno datih cifara?“

## II. MINIMALAN BROJ STRATEGIJA.

U radovima [1] i [2] izložene su strategije potrebne za rešavanje sudokua. Posle učitavanja početno date pozicije neophodno je pomoću nekih metoda ukloniti cifre na pozicijama koje nisu razrešene početnom pozicijom. Zbog jednoznačnosti sudokua posle unosa početno datih vrednosti treba eliminisati sve te vrednosti iz svih: kolona, vrsta i kvadrata. Takve strategije možemo zajednički imenovati: strategije eliminacije. Njih je dodatno moguće poredati u nivoe na sledeći način.

Na nultom nivou je strategija koja posle određivanja bilo koje nove cifre vrši eliminaciju iste iz svih preostalih pozicija svoje vrste, kolona ili kvadrata. Kada se unutar nekog kvadrata može uočiti da se unutar neke njegove vrste ili kolone javlja cifra koje nema u preostalim vrstama ili kolonama tog kvadrata, tada se duž te cele vrste ili kolone, gde se ta cifra javlja, mogu eliminisati sve cifre koje su njoj jednake. Ovakvu eliminaciju je pogodno imenovati sa: eliminacija nivoa jedan.

Kada se u okviru iste vrste, kolone ili kvadrata, na nerazrešenim pozicijama, mogu uočiti dva jednaka kandidata tada možemo razlikovati dva bitno različita podslučaja. Prvi, u okviru te dve pozicije to su jedini kandidati, tada se sa preostalih nerešenih pozicija iste vrste, kolone ili kvadrata mogu ukloniti svi kandidati koji su njima jednaki. Druga, na tim pozicijama ima i drugih kandidata, ali u celoj odgovarajućoj vrsti, koloni ili kvadratu oni se više ne pojavljuju. Tada se u okviru te dve pozicije mogu izbrisati svi kandidati osim ta dva jednaka. Potpuno analogna metoda može da se formuliše i za nivo tri. Praktično se pokazalo da strategije eliminacije označene kao: nula, jedan, dva i tri čine minimalni i dovoljan skup potrebnih strategija za eliminaciju i rešavanje svakog sudokua 9x9. Za određivanje konačnih vrednosti na nekoj poziciji mogu se koristiti dve strategije.

mr Aleksa S. Srdanov, Visoka tehnička škola strukovnih studija Požarevac, Nemanjina 2, 12000 Požarevac, Srbija (telefon: +381-12-531-667; e-mail: aleksa.srdanov@vts-pozarevac.edu.rs).

dr Vladan R. Đulaković, Visoka tehnička škola strukovnih studija Požarevac, Nemanjina 2, 12000 Požarevac, Srbija (telefon: +381-12-531-667; e-mail: vladan.djulakovic@vts-pozarevac.edu.rs).

Dragan M. Milovanović, Visoka tehnička škola strukovnih studija Požarevac, Nemanjina 2, 12000 Požarevac, Srbija (telefon: +381-12-531-667; e-mail: dragan.milovanovic@vts-pozarevac.edu.rs).

Prva strategija izbora je, već spomenuta, ako posle primene strategija eliminacije na nekoj poziciji preostane jedini kandidat, onda je on rešenje na toj poziciji. Drugi strategija je da u okviru neke vrste, kolone ili kvadrata u svakom nerazrešenom polju postoji više različitih kandidata ali za celu vrstu, kolonu ili kvadrat u nekom nerazrešenom polju egzistira samo jedan jedinstven kandidat. Njega treba uzeti kao rešenje na toj poziciji. Načini određivanja konačne vrednosti po prvoj i drugoj strategiji su suštinski različite. Na primerima je moguće pokazati da korišćenje samo jedne od njih ne dovodi do rešenja.

5				3	1	6	
6	1		8		3	4	
			1				8
		5			6	9	7
8		7			4		3
1	3	6			8		
2				9			
	6	9		8		2	1
	8	4	2				6

Sl. 1. Postavka 1.

Na Slici 1. Postavku 1. je moguće razrešiti samo sa jednom strategijom izbora i jednom strategijom eliminacije. Postavku 2. je moguće razrešiti samo jednom strategijom izbora i dve strategije eliminacije. Postavka 3. zahteva dve strategije izbora i tri strategije eliminacije.

		4	5				
5		3	7				
	9		3			5	
7							3
2			4	7			8
4					1		7
	4			6		3	
			5		7		2
				8	4		

Sl. 2. Postavka 2.

						1	
4							
	2						
			5		4	7	
		8			3		
		1	9				
3		4			2		
	5		1				
		8	6				

Sl. 3. Postavka 3.

Sa ovim strategijama se mogla rešiti svaka sudoku brojčanica sve dok se nije oglasio finski matematičar Arto Inkala sa "najtežim sudokuom na svetu" [4].

Rešenje sudokua je jedinstveno. Otuda uvek postoji strategija koja to jedino rešenje odvaja iz skupa svih mogućih. Za ovakve sudokue potrebna je još jedna

dodatna strategija. Pokazaće se da će to biti treća strategija izbora.

8							
		3	6				
	7			9	2		
	5				7		
				4	5	7	
			1				3
		1				6	8
		8	5				1
	9				4		

Sl. 3. "Najteži sudoku na svetu".

Kada se primene sve navedene strategije eliminacije onda u praznim pozicijama preostanu samo mogući kandidati. Ako iz iste vrste, kolone ili kvadrata iz svih nepopunjenih pozicija uklonimo sve zajedničke cifre koje se nalaze u svakoj od njih može se na nekom mestu pojaviti jedinstven kandidat. Takav preostali kandidat je rešenje na toj poziciji. Kada se ova strategija primeni na „najteži sudoku“ ona razreši čak šest nepoznatih mesta. Posle toga sudoku je rešiv samo sa prethodno navedenim strategijama. Ova strategija je potpuno prirodna nadogradnja i uopštenje prve strategije i nije uočena, jer nije postojao primer za primenu. Osim toga, kada uopštimo drugu strategiju izbora na isti način dobija se potpuno ista strategija.

Zato je minimalni broj strategija izbora jednak sedam: četiri strategije eliminacije i tri za izbor kandidata za sve sudokue 9x9.

### III. MINIMALAN BROJ POTEZA.

Strategije opisane u prethodnom odeljku samo eliminišu kandidate. Kada je neki kandidat (cifra) uklonjena algoritam se sa njim više ne bavi. U nerazrešenom sudokuu postoje 81 mesto na koje može doći bilo koja od 9 cifara. Definišimo polupotez kao uklanjanje neke cifre sa nekog mesta, a potez kao razrešenje nekog polja. Maksimalan broj polupoteza da se razreši sudoku je  $81 \times 8 = 648$ . (U svakom polju mora ostati samo jedna cifra.) Kada se u prazan sudoku uvrsti prva cifra (jedan potez) ona smanji broj polupoteza za  $8 + 8 + 8 + 4 = 28$  (eliminiše: 8 iz istog polja, 8 iz iste vrste, 8 iz iste kolone i 4 iz istog kvadrata). U sledećem potezu je taj broj manji ili jednak od 28, zavisno da li je sledeći potez sa istom cifrom, samo u drugom kvadratu, ili prvi put postavljena neka druga cifra u isti kvadrat, tada je taj broj 25. Sa svakim novim razrešenjem broj mogućih polupoteza se smanjuje unutar tog kvadrata. Ako u početnoj postavci postoji zadato n cifara broj potrebnih „kućica“ koje treba razrešiti je  $81 - n$ , za svaki sudoku koji ima rešenje. Predložene strategije ne prave pretpostavke, već koriste jednoznačnost sudokua za njegovo rešavanje. Otuda se predloženim strategijama postiže da se sa minimalnim brojem 648 polupoteza i  $81-n$  poteza razreši svaki sudoku.



## IV. MINIMALNA POSTAVKA SUDOKUA.

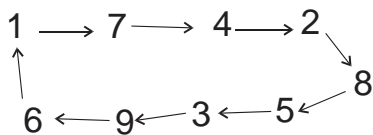
Poznato je da se u početnim postavkama sudokua pojavljuje različit broj cifara. Opravdano se postavlja pitanje: „Koji broj cifara predstavlja minimalnu moguću početnu postavku?“. Sa ovim pitanjem je u vezi veliki broj drugih pitanja u vezi sudokua, kao na Slikama Sl. 5., Sl. 6. i Sl. 7.

		4				6		
		6				4		

Sl. 5.

Na slici 5. može se videti da ako u polaznoj postavci ne bi bio dat ni jedan od upisana 4 broja sudoku ne bi imao jedinstveno rešenje. Ako bi sve četiri pozicije bile prazne jedno rešenje je već dato. Drugo se dobija kada cifre 4 i 6 međusobno zamene mesta. Može se zaključiti, da svaka ovakva pozicija obavezno zahteva barem jednu svoju cifru u početnoj postavci. Zato je jasno da u minimalnoj postavci treba očekivati što manje ovakvih pozicija. Položaj ovakvog pravougaonika ne može biti proizvoljan unutar sudokua. To može biti samo unutar označenih kvadrata. Broj takvih pravougaonika je  $3 \times 9 \times 9 = 243$ .

1	2	3	4	5	6	7	8	9
7	8	9	2	3	1	4	5	6
4	5	6	7	8	9	3	1	2

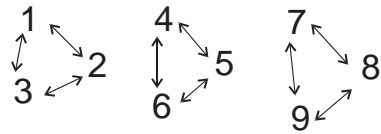


			4					
7								
			8	3				

Sl. 6.

Na slikama 6. i 7., prikazana su samo tri popunjena kvadrata nekih sudoku brojčanica. Ako brojeve unutar bilo koje dve vrste posmatramo kao preslikavanje jedne vrste na drugu, možemo primetiti dva oblika takvih preslikavanja. Na slici 6. to je pun ciklus, a na slici 7. to su tri nezavisna ciklusa. Pod pretpostavkom da su svi brojevi u ostalim kvadratima poznati onda bi za rešavanje sudokua koji je tipa slike 6. ili 7., bilo potrebno različit broj početno zadatih cifara. Četiri za pun ciklus na slici 6. i čak 6 za slučaj tri trojna ciklusa na slici 7. Za bilo koje dve vrste ili kolone broj ciklusa može biti samo: jedan, dva ili tri.

1	7	4	2	8	5	3	9	6
2	8	5	3	9	6	1	7	4
3	9	6	1	7	4	2	8	5



1								
	8		3					4
		6		7				

Sl. 7.

Jasno je da najmanje 8 različitih brojeva mora biti sadržano u početnoj poziciji. U protivnom, ako bi postojala dva broja koja nisu zadata na početku tada bi postojala bar dva rešenja. Jedno, kada se ta dva broja upišu tako da se dobije jedno rešenje, a drugo se dobija kada u istom rasporedu te dve cifre zamene mesta u svim pozicijama gde se nalaze.

Na praznoj tabli za određivanje pozicije bilo koje sledeće nepoznate cifre potrebno je minimalno 4 iste takve cifre na poznatim mestima (Sl. 8.).

Jedno od važnih pitanje je: „Koliko kvadrata može biti bez cifara u početnoj postavci?“. Nije teško zaključiti da je taj broj tri ali ne bilo koja tri kvadrata. Pre svega, ta tri mogu biti svi kvadrati duž dijagonala. Ako su sve pozicije u ostalim kvadratima razrešene sve cifre unutar dijagonalnih kvadrata su time jedinstveno određene. Prazni kvadrati ne mogu biti susedni jer čak i da su sve ostale pozicije razrešene cifre unutar njih nisu jedinstveno određene. Kada se razmotre i ostali slučajevi može se izvesti zaključak da mogu postojati najviše tri prazna kvadrata ili duž dijagonala ili bilo gde drugde unutar sudokua ali da nikoja dva nemaju zajedničku stranicu.

Potpuno analogan zaključak može se izvesti i za vrste i kolone. U početnoj poziciji mogu postojati najviše tri prazne vrste (kolone) ali takve da nikoje dve nisu susedne. Osim toga i sve strategije za rešavanje su ograničene do primene na trojke. Teorijski nije teško postaviti situaciju gde se neki zaključak može izvesti i za četvorke. Problem je u tome što ako biste nastavili dalje rešavanje suočili biste se sa činjenicom da sudoku nije jedinstveno određen.

	1						1	
			1					

Sl. 8.

U [3] je dokazano da ne postoji postavka sudokua sa 16 početno datih cifara. Takođe je u istom radu potvrđeno da, za sada, postoji samo jedan sudoku koji ima 29 različitih

postavki sa 17 cifara, što navodi na zaključak da ne postoji početna postavka sa 16 cifara jer bi u tom slučaju moralo postojati daleko više različitih sedamnaestocifrenih početnih postavki.

#### V. MAKSIMALNA JEDINSTVENA POSTAVKA SUDOKUA.

Osnovni problem ovde je da se nađe početna postavka, sa što više početno zadatih cifara, ali tako da ako bi se neka od njih uklonila iz početne postavke sudoku više ne bi imao jedinstveno rešenje. Postavka koja je prezentovana u radu [2], a koja je prikazana na Sl. 9, se ne može smatrati potencijalno takvom postavkom, jer je očigledno da se može ukloniti 7 iz prvog reda i 8 iz poslednjeg, a da je rešenje sudokua i dalje jedinstveno. Zadatak je dakle pronaći maksimalnu „minimalnu“ jedinstvenu postavku sudokua.

				7		9	4	
	7			9				5
3					5		7	
	8	7	4			1		
4	6	3		8				
					7		8	
8			7					
7							2	8
	5		2	6	8			

Sl. 9.

U otkrivanju takve postavke najjednostavnije je krenuti od već popunjenog sudokua i eliminisati samo one cifre koje su svoje mesto dobile sa već pethodno smeštenim. Tj. poči obrnutim putem i vratiti se do najbrojnije početne postavke. Ako iz popunjenog sudokua isključimo sve cifre sadržane u sva tri kvadrata neke od dijagonala sudoku ostaje jedinstveno rešiv. U praznim vrstama i kolonama svakog od kvadrata preostaju po tri cifre iz njihovih kompletnih vrsta i kolona. U presek dolazi kao jedini kandidat zajednička cifra odgovarajućih trojki iz vrsta i kolona. U svakom popunjenom kvadratu je dodatno moguće ukloniti samo po jednu proizvoljnu cifru i sudoku je i dalje jedinstveno rešiv. Ta cifra je određena pripadnošću istom kvadratu. Neka je to uvek ista cifra. Ako bi taj postupak ponovili još sa jednom cifrom eliminacijom iz svih 6 preostalih kvadrata sudoku postaje neodređen. Zato se moramo vratiti i u proizvoljni od 6 kvadrata upisati drugu cifru samo jedanput. To isto je moguće uraditi sa sledećom cifrom ali za korak manje itd.

5				6	3			
7	4	2		5				
	1				9			
		5				4	9	8
8	6	9					7	
4							5	6
				8	7			
			5	9				7
					6	9	8	

Sl. 10.

Tako se dolazi do maksimalno minimalne postavke jedinstvenog sudokua sa  $7+6+5+4+3+2+1=28$  početno zadatih cifara. Jedan primer je na slici Sl. 10. gde su uklonjene sve jedinice i dvojke i ostavljena jedna dvojka u prvom kvadratu itd.

Kako je maksimalno jednoznačna postavka data sa 28 cifara, a minimalna sa 17 cifara to sledi da je za razrešenje sudokua maksimalno potrebno naći 11 cifara tada je dobijeni sudoku jedinstveno razrešiv samo popunjavanjem jedinih preostalih kandidata u svakom nerazrešenom polju.

#### VI. ZAKLJUČAK

Može se zaključiti da i ako postoji relativno veliki broj različitih sudokua svi oni u suštini predstavljaju samo 11 različitih klasa, zavisno od potrebnog broja poteza do jedinstvenog automatskog razrešenja. Postavke koje sadrže 22, 23 ili 24 početno date cifre su najsloženije za rešavanje jer one koje sadrže 17 ili 28 se sastoje samo u popunjavanju unapred jednoznačno određenih praznih polja.

S obzirom na relativno veoma mali broj potrebnih različitih strategija za rešavanje, ova igra i ako veoma popularna, može se svrstati u kombinatorno jednostavne zabave.

#### LITERATURA

- [1] Aleksa Srdanov, Dragan Milovanović. Dekompozicija algoritma za rešavanje sudokua. IT Zabjak 2016.
- [2] The Mathematics of Sudoku. Tom Davis, tomrdavis@earthlink.net, <http://www.geometer.org/mathcircles> (Preliminary) September 13, 2012.
- [3] There is no 16-Clue Sudoku: Solving the Sudoku Minimum Number of Clues Problem. Gary McGuire, Bastian Tugemann, Gilles Civario, January 1, 2012. [http://www.math.ie/McGuire\\_V1.pdf](http://www.math.ie/McGuire_V1.pdf)
- [4] Pillay Nelishia. Finding Solutions to Sudoku Puzzles Using Human intuitive Heuristics. Sout African Compute Journal sep.2012. vol. 49 p.25-34.
- [5] Deng Xu Qion, Li Yjong Da. A novel hybrid genetic algorithm for solving Sudoku puzzles. Optimizatin Letters 2013. vol 7 issue 2. p 241.
- [6] (Journal Online Sources style) Nick Collins. (2012., jun). World's hardest sudoku: can you crack it? The Telegraph/ . Available: <http://www.telegraph.co.uk/news/science/science-news/9359579/Worlds-hardest-sudoku-can-you-crack-it.html>

#### ABSTRACT

There are a number of algorithms for the solution of a very popular game sudoku. For all of them are imposed at least four basic problems. The first two are related to the minimum and maximum initial setting in terms of number of given initial digits. Initial setting in the sense represent one that within it you can not find a lower setting that would also lead to a unique solution. The third is related to the minimum number of necessary strategies for solving any sudoku, and the fourth for the minimum number of moves needed to do it, in order to come up with a solution. This article provides answers to all four questions mentioned.

#### FOUR EXTREMES RELATED TO SUDOKU

Aleksa S. Srdanov, Vladan R. Đulakovićm, Dragan M. Milovanović, Technical College of Vocational Studies, Požarevac

# Client based selection algorithm for Hybrid Wireless Networks

Aleksandar Tudzarov<sup>1</sup>, Goce Stefanov<sup>2</sup>, Saso Gelev<sup>3</sup>

*Faculty of Electrical Engineering, Goce Delcev University – Stip, R.Macedonia*

**Abstract** — This paper is focused on development of algorithm for network selection in Hybrid wireless environment based on user-side perspective. Client terminal in the proposed hybrid access design has possibility to change the Radio Access Technology - RAT based on different criteria. The selection of the RAT is performed by the wireless terminal using the user agent algorithm for decision making based on the experience from the performance measurements of RATs conducted by the client terminal in the past. We present results of simulation analysis of the proposed intelligent wireless RAT selection algorithm on the client-side for scenario with basic RAT technologies, WWAN and WLAN. The results showed that the presented algorithm gives better results compared to the traditional RAT selection algorithms and as such can play valuable role in designing of RAT selection algorithms for Next Generation Networks.

**Keywords** — Hybrid wireless networks, Quality of Service (QoS), Fuzzy Logic, Particle Swarm Optimization, MCDM.

## I. INTRODUCTION

In this paper we provide complete functional overview of selection based algorithms in Hybrid wireless networks. The main assumption in our approach is that the user terminal will have the possibility to access different RATs (Radio Access Technologies) from single mobile device at the same time, which is reality even today. Then, we propose adding of new network nodes for policy-based routing between IP tunnels to mobile user via different RATs, which are placed in service stratum of the network. The focus in this paper is placed on designing of network selection algorithm that will offer the best selection mechanism for appropriate wireless network according to user preferences, network conditions and other input parameters.

Usually access network selection algorithms that can be found in literature are concentrated on single selection criteria with focus on radio signal strength (RSS) thresholds for each RAT as criteria upon which selection of RAT is based. Mobile terminal compares the RSS of analyzed RAT with the signal thresholds and decides if it should start handoff procedure. The ways that these algorithms operate have certain limitation regarding its reaction to the changing environment conditions.

Usually this algorithm does not take in consideration different viewpoints and goals of the operators, users, and QoS requirements which make them inefficient for a multicriteria problem such as access network selection problem.

Access network selection based on Fuzzy Logic is proven to provide better results and to be more robust when compared to random-based selection algorithms. Furthermore usage of artificial intelligence algorithms inspired by nature for optimization of Fuzzy Logic Controllers (FLC) and Multi-Criteria Decision Making systems (MCDM) are used in order to incorporate past knowledge of wireless networks behavior. All of these algorithms are generally based on learning capabilities provided from measured data. The main contribution of this paper is the developments of algorithm that enhances the way the Fuzzy Logic Controllers are build in a manner that optimizes Fuzzy Logic (FL) decision, generated and optimized by PSO (Particle Swarm Optimization) nature inspired algorithms.

The remainder of this article is structured as follows. Section 2 provides overview of defined radio selector architecture in the Hybrid architecture. Section 3 provides simulation results of proposed selector architecture. Finally, Section 4 concludes the paper.

## II. RADIO ACCESS TECHNOLOGY SELECTOR

In following part of the text we provide description of a novel algorithm for radio networks selection in hybrid wireless environment, which is created using artificial intelligence inspired algorithms.

The algorithm consists of three building components as shown in Figure 1. First component or module is a set of parallel Fuzzy Logic (FL) controllers that are using input data gathered from the measurements for different selection criteria, including: user requirements, QoS requirements, service policies, as well as radio link conditions in different wireless technologies.

The second module is multi-criteria decision mechanism algorithm, which uses as inputs the outputs of the FL controllers form the first module. That is, each criterion can have different weight, which depends upon the assumption of its impact on the best network selection process (i.e. the decision).

The Third module is Particle Swarm Optimization (PSO) [1], [2], [3] mechanism which dynamically modifies the functions of Fuzzy Logic controllers in the first module (shown in Figure 1). This algorithm is used to develop adaptive, flexible, and scalable Network Selection system that can utilize hybrid parallel FL decision-making systems and MCDM systems. FL systems are used in this algorithm in order to introduce flexibility and to deal with

<sup>1</sup> Aleksandar Tudzarov, Faculty of Electrical Engineering, Goce Delcev University – Stip, R.Macedonia, E-mail: aleksandar.tudzarov@ugd.edu.mk

<sup>2</sup> Goce Stefanov, Faculty of Electrical Engineering, Goce Delcev University – Stip, R.Macedonia, E-mail: goce.stefanov@ugd.edu.mk

<sup>3</sup> Saso Gelev, Faculty of Electrical Engineering, Goce Delcev University – Stip, R.Macedonia, E-mail: saso.gelev@ugd.edu.mk

uncertainty that arises from constant changes of radio conditions that impacts radio technologies as well as to be able to use selection mechanism based on different types of inputs that are normally incomparable due to their incompatibility by the nature and behavior. By introducing FL in the selection algorithm possibility for using of different input parameters by their nature arises and increases flexibility and overall usability of the algorithm.

Considering this fact, input variables are in practice normalized in dimensionless values which magnitude represents their impact in selection criteria for each RAT. All of the outputs are normalized and receive values from [0 to 1] interval. In order to simplify implementation of FL and to give some clearance parallel FLC are used. Implementation of Parallel FLC reduces the complexity of inference rules used in the fuzzy-based solutions. Furthermore use of nature inspired algorithms like PSO – “Particle Swarm Optimization” for optimization of FLC and MCDM systems in order to incorporate past knowledge of network behaviors in these systems.

Proposed scheme to solve the network selection problem as well as Software assistant (SA) based on the proposed scheme in coexisted WWAN (Wireless Broadband Technologies 3G/4G) – WLAN (802.11 b/g/n/ac) based systems is described in Simulation below. [5], [6], [7], [8]

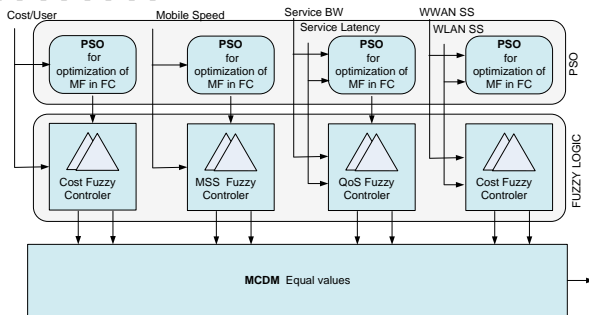


Fig.1. Network selection scheme

The selection procedure can be divided into three phases, initiation phase, decision phase, and execution phase. The objective of the initiation phase is to recognize the need for selection change and subsequently initiate it and find out the required information and measurements for the decision phase. In the decision phase, a comparison of the information and the measurements calculated from a variety of sources including networks measurements, QoS requirements, user preferences, and operator policies is done. This comparison leads to the identification of the best available Wireless Network according to the defined performance evaluation metrics. The scheme decision phase can be described in more detail as follows:

(i) The heterogeneous wireless environment contains up to  $n$  RATs (RAT1, RAT2... RATn) and the framework has to select the most promising one or to rank the RATs according to their suitability.

(ii) The selection depends on multiple criteria up to  $i$  ( $c_1, c_2, \dots, c_i$ ). Different type of criteria can be measured from different sources to cover the different viewpoints of the users, the operators, the applications, and the network conditions. Each criterion is measured then passed to its FL-based control subsystem in the first component.

(iii) Using the PSO system FLC optimization of Fuzzy MF is conducted based on measured inputs and wanted output behavior in order to create most suitable FLC that corresponds to input variable ranges and values. Optimization is conducted by minimizing of MSE between wanted and calculated outputs from Fuzzy Controller that is being optimized.

(iv) Every FL-based subsystem produced by PSO gives an initial score for each RAT that reflects the suitability of that RAT according the FL subsystem criterion. The different sets of scores ( $d_1, d_2, \dots, d_i$ ) are sent to the MCDM in the third component.

(v) Using the initial scores coming from the first component and the weights that are assigned manually (they are equal for each FLC input) the MCDM will select the most promising access network or will rank the available RATs according to their suitability.

Proposed algorithm utilizes the advantages of parallel FL control, PSO optimization and MCDM, scheme is presented on Figure 1.

The idea of the parallel FLC reduces the complexity of the inference rules used in the fuzzy-based solutions. Parameters of FLC that have to be considered while constructing the fuzzy logic are shape of Membership Functions (MF) and type of fuzzy inference system (FIS). In general there are two types of FIS which main difference are in aggregation and defuzzification process (Sugeno & Mamadani). In this work after short analyzes it is determined that it would be best if FLC are based on Mamadani FIS and triangular MF.

Particle swarm optimization (PSO) is a population based stochastic optimization technique inspired by social behavior of bird flocking or fish schooling. The system is initialized with a population of random solutions and searches for optima by updating generations.

Network selection problem is a multicriteria problem by nature, therefore flexible and complementary network selection multicriteria mechanism have to be used in order to provide a solution that can cope with the different viewpoints and goals. [4] Network selection MCDM solution enhanced version of simple multiattribute rating technique (SMART) has been used. SMART is one of the simplest and most efficient MCDM methods. The ranking value  $x_j$  of alternative  $A_j$  is obtained simply as the weighted algebraic mean of the utility values associated with it, that is,  $a_{ij}$  according to (1):

$$X_j = \frac{\sum_{i=1}^m w_i a_{ij}}{\sum_{i=1}^m w_{ij}}, \quad j=1,2,\dots,n \quad (1)$$

SMART employs relatively uncomplicated and straightforward manipulation method, which makes it stronger and easier to use in a hybrid and more complex models such as the proposed one in this paper. In proposed algorithm, there are two alternatives for the MCDM, one is a WWAN based network and the other is a WLAN based network. The input criteria of the MCDM are the outputs of the FL-based control subsystems in the first component. Weights  $W_i$  for criteria  $i$  are assigned to reflect their relative importance. The criteria with more importance to

the operator and user can be assigned higher weight. Since all the outputs of FL subsystems are in the range [0, 1], there is no need to scale the criteria performance against alternatives, and there for all weights in this case are set to equal values. The ranking value of WWAN network  $X_{wwan}$  and the ranking value of WLAN network  $X_{wlan}$  can be calculated as follows:

$$x_{wwan} = \frac{MSS_{c1} * W_v + RSS_{c1} * W_s + STT_{c1} * W_t + UPP_{c1} * W_u}{TW}$$

$$x_{wlan} = \frac{MSS_{c2} * W_v + RSS_{c2} * W_s + STT_{c2} * W_t + UPP_{c2} * W_u}{TW} \quad (2)$$

where  $W_v$  is the assigned weight for the mobile station speed criteria.  $W_s$  is the assigned weight for the received signal strength criteria.  $W_t$  is the assigned weight for the service type criteria.  $W_u$  is the assigned weight for the user preferred price criteria.  $TW$  is the total weight and is calculated using:

$$TW = W_v + W_s + W_t + W_u. \quad (3)$$

The weights of the input criteria  $\{W_v, W_s, W_t, W_u\}$  as previously mentioned are set to equal values (0,25)

### III. SIMULATION OF SELECTOR MODEL

Proposed selection algorithm is evaluated using the simulation approach. MATLAB mathematical software and a set of functions called RUNE [9] "RUDimentary Network Emulator" have been used for the simulation.

The system model considers the coexistence of WWAN network with seven macro cells with omnidirectional antenna and cell radius = 1000 m and a WLAN network with hundred and eight microcells with omnidirectional antenna and cell radius = 250 m. In the system environment, each mobile has a velocity and is moved with a random distance and a random direction at defined time steps. The velocity is a vector quantity with magnitude and direction. The velocity of the  $i$ -th mobile is updated according to (4);

$$v_i = v_{i-1} * P + \sqrt{1 - P^2} * v_{mean} * X \quad (4)$$

Where  $v_i$  is the complex speed [m/s].  $P$  is the correlation of the velocity between time steps. It depends on both  $a_{mean}$  that is the mean acceleration of the mobile user and  $v_{mean}$  which is mean velocity of mobile user.  $P$  is calculated according equation (5):

$$P = \left( \frac{-dt * a_{mean}}{v_{mean}} \right) \quad (5)$$

$X$  is Rayleigh distributed magnitude with mean 1 and a random direction.  $v_{mean}$  is the mean speed of mobiles.  $v_{mean}$ , was set to 10 km/h and  $a_{mean}$  has been set to 1 km/h<sup>2</sup>.

Four types of services are considered in the simulation and they are equally distributed among the users: voice calls, low bit rate real-time video telephony, high bit rate video and the nonreal-time data traffic.  $N$  mobile users are

created and they are randomly distributed among the defined service types.

In general their requirements from the wireless network are simulated as pair of values defined as: [**Latency**, **Bandwidth**]. In that sense following 4 pairs were considered  $\{[100,64]; [200,128]; [400,256]; [800,512]\}$ .

The traffic is modeled according to Poisson process. The main holding time is assumed to be 50 seconds. For the purpose of simulation FLC where designed that fits the FLC parallel scheme. Considering that two RAT technologies are analyzed in the scenario two outputs from every FLC are taken. Outputs from the FLC present degree of membership of each RAT scaled to input variables regarding FL rules. Taking into account that most of ANS algorithm are based on Radio Signal Strength and that this parameter has the biggest impact on final decision in ANS, special Particle Swarm Optimization (PSO) algorithm described above is used in order to produce FLC which membership function are tuned to measured signal strengths from simulation.

PSO algorithms uses Swarm size of 50 particles and maximum number of iteration = 50. Evaluation function is based on minimizing the mean square error (MSE) comparing to expected predefined values. Expected values are defined as values taken from humanly decision that would be made if access network selection is done by human for every point in time and separately for each analyzed criteria.

This subsection shows some simulation results and compares the performance of proposed solution to two different reference ANS (Access network selection) algorithms.

The first algorithm is a service-type based selection algorithm where high bit services with low propagation delay requirements are sent to the WLAN and the low bit rate services with the high propagation delay requirements are sent to the WWAN. The second algorithm is a random-based selection algorithm where the users are assigned randomly to the two networks. All solutions have been simulated, evaluated, and compared for the same objective optimization and that is to maximize the percentage of assigned to the networks with stronger signal strength ( $P_q$ ). Several runs of simulation have been carried out for different number of users in simulation scenario (from 100 to 1000 with step of 100). Results are given on Figure2 and in Table 1.

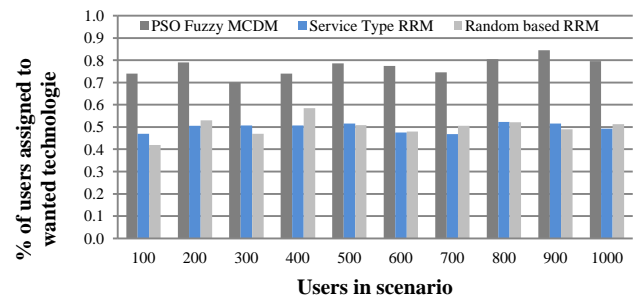


Fig. 2. Percentage of satisfied users

TABLE 1: PERCENTAG OF SATISFIED USERS

Nr. Users	PSO-Fuzzy MCDM ANS	Service Type based ANS	Random based ANS
100	0,740	0,470	0,420
200	0,790	0,505	0,530
300	0,700	0,507	0,470
400	0,740	0,508	0,585
500	0,786	0,516	0,508
600	0,775	0,475	0,480
700	0,746	0,469	0,506
800	0,805	0,523	0,521
900	0,844	0,516	0,490
1000	0,796	0,493	0,513

For example, with 900 users in the scenario Pq for proposed algorithm is 0,844 that means that 84% of the users are assigned to RAT with better signal strength and by that means with better QoS, while in that same scenario reference algorithms can achieve Pq factor of around 50%. In average, proposed solution achieves around 27% enhancement over the random-based selection algorithm.

#### IV. CONCLUSION

A novel scheme to solve the network selection problem has been presented in this paper. The scheme can cope with the different and contrast view points and goals of the operator and users. The proposed scheme has been used to present and design a multicriteria network selection solution that considered the user, the QoS, and the operator view points. The simulation results show that the proposed solution has a better and more robust performance over the reference solutions.

#### REFERENCES

- [1] J. Kennedy, R.C. Eberhart, Particle swarm optimization, in: Proceedings of the IEEE International Conference on Neural Networks, vol. 4, pp. 1942–1948, 1995.
- [2] S. Doctor, V. K. Venayagamoorthy, “Navigation of mobile sensors using PSO and embedded PSO in a fuzzy logic controller”, IEEE Industry Applications Conference, vol. 2, pp. 1200-1206, Oct. 2004.
- [3] M. Clerc, J. Kennedy, The particle swarm explosion, stability, and convergence in a multidimensional complex space, IEEE Transaction on Evolutionary Computation, vol.6, pp. 58–73, 2002.
- [4] J. P’erez-Romero, O. Sallent, R. Agust’1, and M. A. D’iaz-Guerra, RadioResourceManagement Strategies in UMTS, JohnWiley & Sons, New York, NY, USA, 2005.
- [5] L. Giupponi, R. Agust’1, J. P’erez Romero, and O. Sallent, “A novel joint radio resource management approach with reinforcement learning mechanisms,” in Proceedings of the 24th IEEE International Performance, Computing, and Communications Conference (IPCCC ’05), pp. 621–626, Phoenix, Ariz, USA, April 2005. J. Wang, “Periodical style—Submitted for publication,” IEEE J. Quantum Electron., submitted for publication.
- [6] R. Agust’1, O. Sallent, J. P’erez-Romero, and L. Giupponi, “A fuzzy neural based approach for joint radio resource management in a B3G framework,” in Proceedings of the 1st International Conference on Quality of Service in Heterogeneous Wired/Wireless Networks (QSHINE ’04), pp. 216–224, Dallas, Tex, USA, October 2004.
- [7] A. Wilson, A. Lenaghan, and R. Malyan, “Optimizing wireless access network selection to maintain QoS in heterogeneous wireless environments,” in Proceedings of the Wireless Personal Multimedia Communications (WPMC ’05), Aalborg, Denmark, September 2005.
- [8] M. Alkhwilani and A. Ayesh, “Access Network Selection Based on Fuzzy Logic and Genetic Algorithms”, Advances in Artificial Intelligence, Volume 8, Issue 1, January 2008.
- [9] J. Zander and S. Kim, Radio Resource Management for Wireless Networks, Artech House, Boston, Mass, USA, 2001.

# Rješenja za poboljšanje D2D komunikacija u 5G mrežama

Uglješa Urošević, Zoran Veljović

**Sadržaj** — Naredna generacija mobilnih radio komunikacionih sistema treba da omogući prenos podataka uz značajno veće brzine prenosa, mnogo veći broj konekcija i konektovanih uređaja, veću spektralnu efikasnost, veću energetska efikasnost, poboljšan kvalitet servisa u pogledu kašnjenja prilikom prenosa podataka, kapaciteta, pouzdanosti i sigurnosti. Glavne tehnologije koje trebaju to da omoguće su masivni MIMO (Multiple-Input Multiple Output) sistemi, korišćenje frekvencija iz viših opsega, posebno mmWave (millimeter-Wave) frekvencija, P2P (Peer-to-Peer) komunikacije, heterogene mreže (HetNet - Heterogeneous Networks) itd. U ovom radu su data neka rješenja za poboljšanje performansi D2D komunikacija primjenom distribuiranih MIMO tehnika.

**Ključne reči** — BER, D2D, QOSTBC, 5G.

## I. UVOD

EKSPONENCIJALNI rast bežičnih servisa i količine podataka, kojeg pokreću mobilni internet i pametni uređaji, je motivisao istraživanja o narednoj generaciji mobilnih radio komunikacionih sistema, odnosno o 5G mobilnim mrežama. Buduće mreže će morati da podrže multimedijalne aplikacije uz ispunjavanje određenih zahtjeva ako što su: povećanje brzine prenosa korisničkih podataka do 100 puta, *end-to-end* kašnjenje manje od 5ms, poboljšanu *indoor* pokrivenosti, povećanje spektralne i energetske efikasnosti, i tako dalje, [1], [2].

Postoji dosta različitih tehnologija koje obećavaju ispunjavanje zahtjeva 5G bežičnih komunikacionih sistemima, kao što su masivni MIMO (*Multiple Input Multiple Output*) sistemi, prenos na milimetarskim talasima (*mm-wave*), spektralno efikasne komunikacije, odnosno kognitivne radio mreže, zgušnjavanje postojećih mobilnih mreža sa masovnim dodavanjem malih ćelija, P2P (*Peer-to-Peer*) komunikacije kao što su D2D (*Device-to-Device*), M2M (*Machine-to-Machine*), V2V (*Vehicle-to-Vehicle*), heterogene mreže (HetNet), istovremeni prenos i prijem (*full duplex* komunikacija), prikupljanje energije (*energy harvesting*), C-RAN (*Cloud-based Radio Access Network*), SDN (*Software Defined Networks*), virtuelizacija bežičnih resursa, itd., [3].

D2D komunikacije se odnose na radio tehnologiju koja omogućava uređajima da direktno komuniciraju, bez korišćenja mrežne infrastrukture. D2D komunikacije poboljšavaju spektralnu efikasnost, korisničko iskustvo, smanjuju kašnjenje prilikom prenosa podataka, povećavaju brzinu prenosa podataka...

U D2D komunikaciji korisnički podaci se direktno prenose između korisnika bez rutiranja kroz jezgro mreže. Osim toga, resursi između D2D korisnika se mogu ponovo koristiti. Sve ovo vodi ka povećanju spektralne efikasnosti, kapaciteta i smanjivanju kašnjenja usled prenosa podataka. Takođe, može doći do kolapsa komunikacionog sistema ukoliko se desi oštećenje jezgra mreže ili pristupnih mrežnih elemenata. D2D omogućava komunikaciju terminala i u tim scenarijima. Ako je bežična infrastruktura oštećena ili korisnici nisu u zoni pokrivanja bežičnom mrežom, *multi-hop* D2D može omogućiti komunikaciju bez učešća jezgra mreže ili čak pristup jezgru mreže. Na ovaj način se može proširiti broj bežičnih aplikacija. Mogućnosti primjene D2D uključuju lokalne usluge, hitne komunikacije, javnu bezbednost, *proximity based* igre i društveno umrežavanje, oglase za prolaznika, kontrolu saobraćaja i bezbjednosti, inteligentne transportne sisteme, IoT (*Internet of Things*), itd., [4] - [6].

Ovdje su data rješenja za poboljšanje D2D komunikacija primjenom distribuiranih MIMO sistema. Simulacioni rezultati pokazuju benefite razmatranog pristupa.

Rad je organizovan na sledeći način. Nakon predloženih rješenja slijede simulacioni rezultati. Na kraju je dat zaključak.

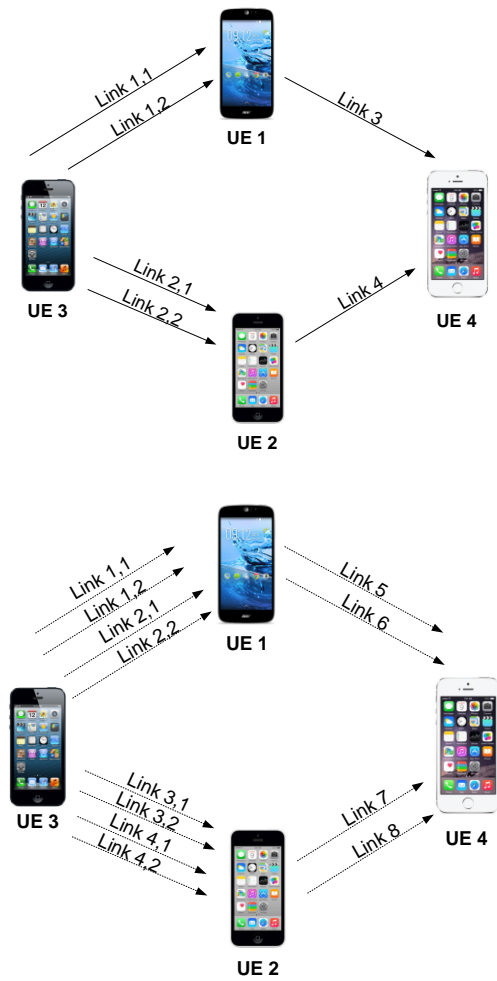
## II. PREDLOŽENA RJEŠENJA

Ovdje se razmatraju dva D2D scenarija i predlažu rješenja za virtuelnu implementaciju QOSTBC (*Quasi Orthogonal Space-Time Block Code*).

U prvom scenariju je pretpostavka da  $UE_3$  (*User Equipment*) preko dva UE prenosi podatke ka  $UE_4$ , slika 1.

Uglješa Urošević, Elektrotehnički fakultet u Podgorici, Univerzitet Crne Gore, Džordža Vašingtona bb, 81000 Podgorica, Crna Gora (telefon: 382 20 245 839; e-mail: ugljesa@ac.me).

Zoran Veljović, Elektrotehnički fakultet u Podgorici, Univerzitet Crne Gore, Džordža Vašingtona bb, 81000 Podgorica, Crna Gora (telefon: 382 20 245 839; e-mail: veljovic@ac.me).



Sl. 1. Dva D2D scenarija

UE<sub>3</sub> koristi dvije predajne antene dok UE<sub>1</sub>, UE<sub>2</sub> i UE<sub>4</sub> koriste po jednu antenu. QOSTBC za klasični MIMO sistem sa četiri predajne antene je dat sa [7] - [11]:

$$\text{QOSTBC} = \frac{1}{\sqrt{4}} \begin{bmatrix} a_1 & -a_2^* & -a_3^* & a_4 \\ a_2 & a_1^* & -a_4^* & -a_3 \\ a_3 & -a_4^* & a_1^* & -a_2 \\ a_4 & a_3^* & a_2^* & a_1 \end{bmatrix} \quad (1)$$

Ovdje je predložena sledeća kodna procedura za distribuiranu realizaciju. Za svaka četiri simbola  $a_i$ ,  $i=1, \dots, 4$ , UE<sub>3</sub> generiše dvije sekvence simbola:

$$\mathbf{UE}_{3,1} = \frac{1}{\sqrt{2}} \begin{bmatrix} a_1 & -a_2^* & -a_3^* & a_4 \end{bmatrix} \quad (2)$$

$$\mathbf{UE}_{3,2} = \frac{1}{\sqrt{2}} \begin{bmatrix} a_2 & a_1^* & -a_4^* & -a_3 \end{bmatrix} \quad (3)$$

Srednja snaga po simbolu je fiksirana skalirajućim faktorom  $1/\sqrt{2}$ . Pretpostavljeno je da su svi simboli u jednoj sekvenci pod istim uticajem fedinga.

Koeficijent kanala između  $i$ -tog,  $i=1,2$ , UE i  $j$ -te,  $j=1,2$ , antene na UE<sub>3</sub> je  $h_{ij}$ . Koeficijenti kanala na UE<sub>1</sub>-UE<sub>4</sub> i

UE<sub>2</sub>-UE<sub>4</sub> linkovima su  $h_3$  i  $h_4$ , respektivno. Prijemna sekvenca simbola na  $i$ -tom,  $i=1, 2$ , UE je:

$$\mathbf{UE}_{i,\text{in}} = \sum_{j=1}^2 \mathbf{UE}_{3,j} \cdot h_{ij} + \mathbf{N}_{\text{UE},i} \quad (4)$$

$\mathbf{N}_{\text{UE},i}$  je matrica za bijeli Gausov šum (AWGN) na  $i$ -tom,  $i=1,2$ , UE:

$$\mathbf{N}_{\text{UE},i} = [n_{\text{UE},i,1} \quad n_{\text{UE},i,2} \quad n_{\text{UE},i,3} \quad n_{\text{UE},i,4}] \quad (5)$$

UE<sub>1</sub> i UE<sub>2</sub> generišu sledeće sekvence:

$$\mathbf{UE}_{1,\text{out}} = \frac{1}{\sqrt{2}} \mathbf{UE}_{1,\text{in}} \quad (6)$$

$$\mathbf{UE}_{2,\text{out}} = \frac{1}{\sqrt{2}} \begin{bmatrix} -\mathbf{UE}_{2,\text{in}}^*(3) & -\mathbf{UE}_{2,\text{in}}^*(4) & \mathbf{UE}_{2,\text{in}}^*(1) & \mathbf{UE}_{2,\text{in}}^*(2) \end{bmatrix} \quad (7)$$

Prijemna sekvenca simbola na UE<sub>4</sub> je:

$$\mathbf{UE}_{4,\text{in}} = \mathbf{UE}_{1,\text{out}} \cdot h_3 + \mathbf{UE}_{2,\text{out}} \cdot h_4 + \mathbf{N}_{\text{UE4}} \quad (8)$$

$\mathbf{N}_{\text{UE4}}$  je AWGN matrica na MS:

$$\mathbf{N}_{\text{UE4}} = [n_{\text{UE4},1} \quad n_{\text{UE4},2} \quad n_{\text{UE4},3} \quad n_{\text{UE4},4}] \quad (9)$$

Uz dati kooperativni prenos kreira se virtuelni 4x1 MISO kanal između UE<sub>3</sub> i UE<sub>4</sub>, odnosno:

$$\mathbf{UE}_{4,\text{in}} = \frac{1}{\sqrt{4}} \begin{bmatrix} a_1 & -a_2^* & -a_3^* & a_4 \\ a_2 & a_1^* & -a_4^* & -a_3 \\ a_3 & -a_4^* & a_1^* & -a_2 \\ a_4 & a_3^* & a_2^* & a_1 \end{bmatrix}^T \begin{bmatrix} H_1 \\ H_2 \\ H_3 \\ H_4 \end{bmatrix} + \mathbf{N} \quad (10)$$

gdje je  $H_1 = h_{11}h_3$ ,  $H_2 = h_{12}h_3$ ,  $H_3 = h_{21}^*h_4$ ,  $H_4 = h_{22}^*h_4$ .

U drugom scenariju je pretpostavka da su UE<sub>1</sub> i UE<sub>2</sub> opremljeni sa po dvije antene.

Ovdje se predlaže rješenje za kooperativni prenos za kreiranje virtuelnog 8x1 MISO kanala.

QOSTBC za klasični MIMO sistem sa četiri predajne antene je dat sa [7] - [11]:

$$\text{QOSTBC}_8 = \frac{1}{\sqrt{8}} \begin{bmatrix} a_1 & -a_2^* & -a_3^* & a_4 & -a_5^* & a_6 & a_7 & -a_8^* \\ a_2 & a_1^* & -a_4^* & -a_3 & -a_6^* & -a_5 & a_8 & a_7^* \\ a_3 & -a_4^* & a_1^* & -a_2 & -a_7^* & a_8 & -a_5 & a_6^* \\ a_4 & a_3^* & a_2^* & a_1 & -a_8^* & -a_7 & -a_6 & -a_5^* \\ a_5 & -a_6^* & -a_7^* & a_8 & a_1^* & -a_2 & -a_3 & a_4^* \\ a_6 & a_5^* & -a_8^* & -a_7 & a_2^* & a_1 & -a_4 & -a_3^* \\ a_7 & -a_8^* & a_5^* & -a_6 & a_3^* & -a_4 & a_1 & -a_2^* \\ a_8 & a_7^* & a_6^* & a_5 & a_4^* & a_3 & a_2 & a_1^* \end{bmatrix} \quad (11)$$

Predloženo je da za svakih osam simbola podataka  $a_i$ ,  $i=1, \dots, 8$ , UE<sub>3</sub> generiše dvije sekvence simbola:

$$\mathbf{UE3}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} a_1 & -a_2^* & -a_3^* & a_4 & -a_5^* & a_6 & a_7 & -a_8^* \end{bmatrix} \quad (12)$$

$$\mathbf{UE3}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} a_2 & a_1^* & -a_4^* & -a_3 & -a_6^* & -a_5 & a_7 & a_8^* \end{bmatrix} \quad (13)$$



Koeficijent kanala između  $i$ -te,  $i=1,2$ , antene na  $k$ -tom UE,  $k=1,2$ , i  $j$ -te,  $j=1,2$ , antene na UE $_j$  je  $h_{k,ij}$ .

Prijemna sekvenca simbola na  $i$ -toj anteni na  $k$ -tom UE je:

$$\mathbf{UE}_{k,i,\text{in}} = \sum_{j=1}^2 \mathbf{UE}_{3j} \cdot h_{k,ij} + \mathbf{N}_{\mathbf{UE}_{k,i}} \quad (14)$$

$\mathbf{N}_{\mathbf{UE}_{k,i}}$  je AWGN matrica na  $i$ -toj anteni na  $k$ -tom UE:

$$\mathbf{N}_{\mathbf{UE}_{k,i}} = [n_{\mathbf{UE}_{k,i}1} \ n_{\mathbf{UE}_{k,i}2} \ \dots \ n_{\mathbf{UE}_{k,i}8}] \quad (15)$$

Ovdje,  $n_{\mathbf{UE}_{k,i}s}$ ,  $s=1, \dots, 8$  predstavlja AWGN na  $i$ -toj anteni na  $k$ -tom UE u  $s$ -tom vremenskom slotu.

Na prvoj anteni UE $_1$  generiše sekvencu:

$$\mathbf{UE}_{1,1,\text{out}} = \frac{1}{\sqrt{4}} \mathbf{UE}_{1,1,\text{in}} \quad (16)$$

Na drugoj anteni UE $_1$  permutuje prijemnu sekvencu simbola i generiše izlaznu sekvencu:

$$\mathbf{UE}_{1,2,\text{out}} = \frac{1}{\sqrt{4}} \begin{bmatrix} -\mathbf{UE}_{1,2,\text{in}}^*(3) & -\mathbf{UE}_{1,2,\text{in}}^*(4) & \mathbf{UE}_{1,2,\text{in}}^*(1) \\ \mathbf{UE}_{1,2,\text{in}}^*(2) & -\mathbf{UE}_{1,2,\text{in}}^*(7) & -\mathbf{UE}_{1,2,\text{in}}^*(8) & \mathbf{UE}_{1,2,\text{in}}^*(5) & \mathbf{UE}_{1,2,\text{in}}^*(6) \end{bmatrix} \quad (17)$$

Na prvoj i drugoj anteni UE $_2$  generiše sekvence simbola:

$$\mathbf{UE}_{2,1,\text{out}} = \frac{1}{\sqrt{4}} \begin{bmatrix} -\mathbf{UE}_{2,1,\text{in}}^*(3) & -\mathbf{UE}_{2,1,\text{in}}^*(4) & -\mathbf{UE}_{2,1,\text{in}}^*(1) \\ -\mathbf{UE}_{2,1,\text{in}}^*(2) & \mathbf{UE}_{2,1,\text{in}}^*(7) & \mathbf{UE}_{2,1,\text{in}}^*(8) & \mathbf{UE}_{2,1,\text{in}}^*(5) & \mathbf{UE}_{2,1,\text{in}}^*(6) \end{bmatrix} \quad (18)$$

$$\mathbf{UE}_{2,2,\text{out}} = \frac{1}{\sqrt{4}} \begin{bmatrix} \mathbf{UE}_{2,2,\text{in}}(3) & \mathbf{UE}_{2,2,\text{in}}(4) & -\mathbf{UE}_{2,2,\text{in}}(1) \\ -\mathbf{UE}_{2,2,\text{in}}(2) & -\mathbf{UE}_{2,2,\text{in}}(7) & -\mathbf{UE}_{2,2,\text{in}}(8) & \mathbf{UE}_{2,2,\text{in}}(5) & \mathbf{UE}_{2,2,\text{in}}(6) \end{bmatrix} \quad (19)$$

Koeficijenti kanala na UE $_1$ -UE $_4$  i UE $_2$ -UE $_4$  linkovima su  $h_5$ ,  $h_6$ ,  $h_7$ ,  $h_8$ , respektivno. Prijemna sekvenca simbola na UE $_4$  je:

$$\mathbf{UE}_{4,\text{in}} = \mathbf{UE}_{1,1,\text{out}} \cdot h_5 + \mathbf{UE}_{1,2,\text{out}} \cdot h_6 + \mathbf{UE}_{2,1,\text{out}} \cdot h_7 + \mathbf{UE}_{2,2,\text{out}} \cdot h_8 + \mathbf{N}_{\mathbf{UE}_4} \quad (20)$$

$\mathbf{N}_{\mathbf{UE}_4}$  je AWGN matrica na UE $_4$ :

$$\mathbf{N}_{\mathbf{UE}_4} = [n_{\mathbf{UE}_4,1} \ n_{\mathbf{UE}_4,2} \ \dots \ n_{\mathbf{UE}_4,8}] \quad (21)$$

Uz dati kooperativni prenos kreira se virtuelni 4x1 MISO kanal između UE $_3$  i UE $_4$ , odnosno:

$$\mathbf{UE}_{4,\text{in}} = \frac{1}{\sqrt{8}} \begin{bmatrix} a_1 & -a_2^* & -a_3^* & a_4 & -a_5^* & a_6 & a_7 & -a_8^* \\ a_2 & a_1^* & -a_4^* & -a_3 & -a_6^* & -a_5 & a_8 & a_7^* \\ a_3 & -a_4^* & a_1^* & -a_2 & -a_7^* & a_8 & -a_5 & a_6^* \\ a_4 & a_3^* & a_2^* & a_1 & -a_8^* & -a_7 & -a_6 & -a_5^* \\ a_5 & -a_6^* & -a_7^* & a_8 & a_1^* & -a_2 & -a_3 & a_4^* \\ a_6 & a_5^* & -a_8^* & -a_7 & a_2^* & a_1 & -a_4 & -a_3^* \\ a_7 & -a_8^* & a_5^* & -a_6 & a_3^* & -a_4 & a_1 & -a_2^* \\ a_8 & a_7^* & a_6^* & a_5 & a_4^* & a_3 & a_2 & a_1^* \end{bmatrix}^T \begin{bmatrix} H_1 \\ H_2 \\ H_3 \\ H_4 \\ H_5 \\ H_6 \\ H_7 \\ H_8 \end{bmatrix} + \mathbf{N}^T \quad (22)$$

gdje je:

$$H_1 = h_{1,11}h_5, H_2 = h_{1,12}h_5, H_3 = h_{1,21}^*h_6, H_4 = h_{1,22}^*h_6, \\ H_5 = h_{2,11}^*h_7, H_6 = h_{2,12}^*h_7, H_7 = h_{2,21}h_8, H_8 = h_{2,22}h_8.$$

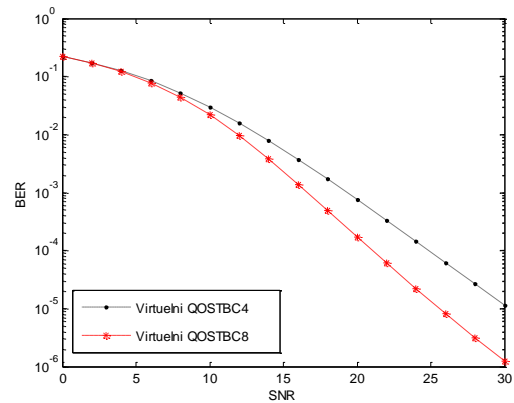
Treba napomenuti da se predložena rješenja mogu proširiti na veći broj antena što je posebno značajno sa stanovišta masivnih MIMO sistema.

Takođe, kod razmatranih QOSTBC kodova broj prijemnih antena može biti proizvoljan, odnosno ne utiče na transmisionu šemu. Naravno, veći broj prijemnih antena znači veći red *diversity*-ja a samim tim i manju vjerovatnoću greške prilikom prenosa podataka.

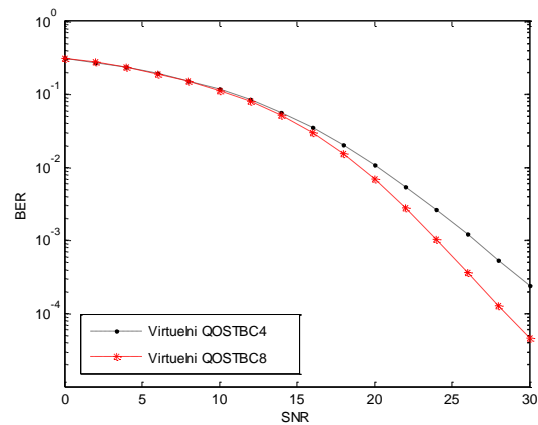
### III. SIMULACIONI REZULTATI

U ovoj sekciji se predložene šeme za kreiranje virtuelnog MISO kanala porede sa stanovišta vjerovatnoće greške (BER-a). Simulacioni rezultati su dati pod pretpostavkom da je mobilni radio kanal definisan *Rayleigh*-evim fading statistikom. Takođe, idealna i neidealna estimacija koeficijenata kanala je razmatrana.

Za QPSK i 64QAM modulaciju BER performanse u funkciji odnosa signal/šum (SNR) su date na slikama 2 i 3. Kao što se očekivalo virtuelna QOSTBC $_8$  šema pokazuje manju vjerovatnoću greške u poređenju sa virtuelnom QOSTBC $_4$  šemom.



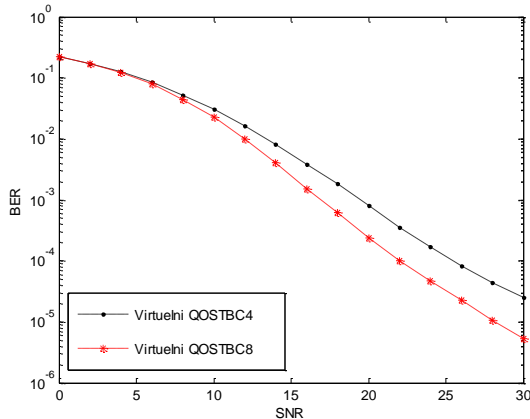
SI 2. BER performanse virtuelnih QOSTBC $_8$  i QOSTBC $_4$  šema, QPSK modulacija



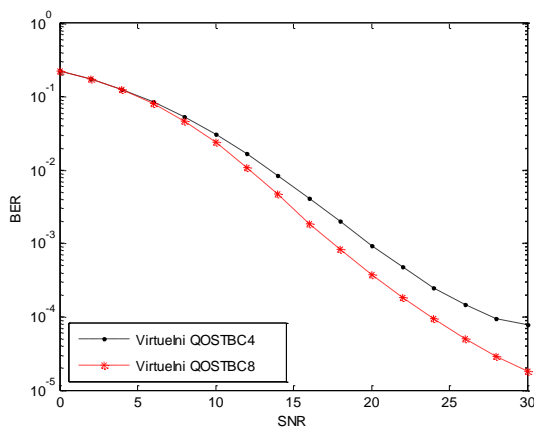
SI 3. BER performanse virtuelnih QOSTBC $_8$  i QOSTBC $_4$  šema, 64QAM modulacija

U slučaju QPSK modulacije i BER vrijednosti od  $10^{-5}$  predložena virtualna QOSTBC<sub>8</sub> šema zahtijeva za 4.7dB manji SNR u poređenju sa virtualnom QOSTBC<sub>4</sub> šemom. Za 64QAM modulaciju i BER vrijednost od  $10^{-3}$  zahtijevani SNR je za 2.4dB manji kod virtualne QOSTBC<sub>8</sub> šeme u poređenju sa virtualnim QOSTBC<sub>4</sub>.

Osim idealne estimacije kanala razmatrana je i realna situacija sa neidealnom estimacijom, pri čemu su date BER performanse za QPSK modulaciju i MSPE (*Mean Square Percentage Error*) vrijednosti od 3% i 5%. Kao što se može uočiti što je veća greška prilikom estimacije lošije su BER performanse.



SI 4. BER performanse virtualnih QOSTBC<sub>8</sub> i QOSTBC<sub>4</sub> šema, QPSK modulacija, MSPE=3%



SI 5. BER performanse virtualnih QOSTBC<sub>8</sub> i QOSTBC<sub>4</sub> šema, QPSK modulacija, MSPE=5%

#### IV. ZAKLJUČAK

Ključne tehnike koje će omogućiti ispunjavanje zahtjeva mobilnih radio komunikacionih sistema naredne generacije su masivni MIMO sistemi, prenos u mmWave frekvencijskom opsegu, odnosno korišćenje većih učestanosti, P2P komunikacije tipa D2D, V2V i sl., heterogene mreže... U ovom radu su data rješenja za poboljšanje D2D komunikacija kroz distribuirane MIMO sisteme. Novi pristup za poboljšanje performansi virtualnih MIMO komunikacija je objašnjen. Takođe, bitno je što se

razmatrani koncept može primijeniti na veći broj antena, odnosno u scenariju gdje se primjenjuje masivni MIMO.

#### LITERATURA

- [1] E. Hossain, M. Hasan, "5G Cellular: Key Enabling Technologies and Research Challenges," IEEE Instrumentation & Measurement Magazine, Volume 18, Issue 3, pp. 11 – 21, June 2015.
- [2] J. Guey, P. Liao, Y. Chen, A. Hsu, C. Hwang, G. Lin, "On 5G Radio Access Architecture and Technology," IEEE Wireless Communication, Volume 22, Issue 5, pp. 2 – 5, October 2015.
- [3] D. Soldani, A. Manzalini, "Horizon 2020 and Beyond: On the 5G Operating System for a True Digital Society," IEEE Vehicular Technology Magazine, Volume 10, Issue 1, pp. 32 - 42, March 2015.
- [4] K. Choi, Z. Han, "Device-to-Device Discovery for Proximity-Based Service in LTE-Advanced System," IEEE Journal on Selected Areas in Communications, Volume 33, Issue 1, pp. 55 – 66, January 2015.
- [5] M. Jo, T. Maksymyuk, B. Strykhalyuk, C. Cho, "Device-to-device-based Heterogeneous Radio Access Network Architecture for Mobile Cloud Computing," IEEE Wireless Communications, Volume 22, Issue 3, pp. 50 – 58, June 2015.
- [6] W. Zhong, Y. Fang, S. Jin, K. Wong, "Joint Resource Allocation for Device-to-Device Communications Underlying Uplink MIMO Cellular Networks," IEEE Journal on Selected Areas in Communications, Volume 33, Issue 1, pp. 41 – 54, January 2015.
- [7] U. Urošević, Z. Veljović, M. Pejanović-Djurisic, "MIMO Solution for Performance Improvements of OFDM-CDMA System with Pilot Tone," Wireless Networks, Volume 19, Issue 8, pp. 2021-2028, Springer, November 2013.
- [8] Kühn, Wireless Communications over MIMO Channels - Applications to CDMA and Multiple Antenna Systems, pp. 283–288, Wiley, 2006.
- [9] T. Yamaoka, Y. Hara, N. Fukui, H. Kubo and T. Yamazato, "A Simple Cooperative Relaying with Alamouti Coded Transmission," IEICE Trans. Commun., Vol. E95–B, No.2, pp. 643-646, February 2012.
- [10] U. Urošević, Z. Veljović, M. Pejanović-Djurisic, "A New Solution for Simple Cooperative Relaying", Wireless Personal Communications, Vol. 75, Issue 2, pp. 1235-1250, Springer, March 2014.
- [11] V. Tarokh, H. Jafarkhani, A. R. Calderbank, "Space-Time Block Codes from Orthogonal Designs," IEEE Transactions on Information Theory, vol. 45, no. 5, pp. 1456-1467, July 1999.

#### ABSTRACT

The future generation of wireless networks is expected to support a significantly large amount of mobile data traffic, massive number of wireless connections and devices, increased energy and spectral efficiency, improved quality of service (QoS) in terms of communication delay, capacity, reliability and security. The main techniques that will enable these features are massive multiple-input, multiple-output (MIMO), utilization of higher frequencies, particularly millimeter-wave (mmWave) frequencies, peer-to-peer (P2P) communications, heterogeneous network (HetNet) implementation, etc. Here we present new solutions for D2D communications improvement through distributed MIMO techniques.

#### SOLUTIONS FOR D2D IMPROVEMENTS IN 5G NETWORKS

Ugljesa Urošević, Zoran Veljović

# Zaštita od DoS napada primenom Remotely Triggered Black Hole tehnike filtriranja saobraćaja

Jelena T. Šuh and Snežana Lj. Stefanovski Pajović

**Sadržaj** — Računarske mreže su kompleksni sistemi koji imaju zadatak da obezbede normalno funkcionisanje velikog broja servisa i aplikacija. Pored standardnih zahteva u pogledu dostupnosti i skalabilnosti, bezbednost u računarskim mrežama predstavlja jednu od najznačajnijih komponenti u procesu projektovanja. Pružanje bezbednosti u mreži je proces primene različitih hardverskih i softverskih rešenja sa ciljem da se mrežna infrastruktura i resursi zaštite od neautorizovanog pristupa i zloupotrebe. Cilj ovog rada je da se ukaže na problem bezbednosti u računarskim mrežama, pri čemu je posebno istaknut značaj zaštite od DoS (*Denial of Service*) napada. U radu je opisano RTBH (*Remotely Triggered Black Hole*) rešenje, kao primer efikasne tehnike za mitigaciju DoS napada. Pravovremena detekcija i akcija u cilju sprečavanja malicioznog napada na određene mrežne resurse doprinosi očuvanju kvaliteta funkcionisanja mreže i servisa realizovanih u njoj.

**Ključne reči** — Bezbednost, *Denial of Service*, Računarske mreže, *Remotely Triggered Black Hole*.

## I. UVOD

PROJEKTOVANJE računarskih mreža predstavlja kompleksan proces u kome bezbednost zauzima izuzetno važno mesto. Bezbednost se može definisati kao sposobnost zaštite ili odbrane sajber prostora od napada [1]. Da bi se ostvario ovaj zadatak neophodno je primeniti adekvatne mere sa ciljem da se obezbedi normalno funkcionisanje kritičnih funkcija u mreži uprkos postojanju rizika od različitih napada [2]. Krajnji cilj je da se primenom bezbednosnih polisa i određenih procedura obezbedi poverljivost, integritet i dostupnost (*confidentiality, integrity, and availability* – CIA) informacionog sistema u slučaju postojanja malicioznih napada koji mogu da ugroze određeni mrežni resurs [3].

Mrežna infrastruktura servis provajdera je kompleksan sistem, koji ima zadatak da obezbedi veliki broj različitih servisa rezidencijalnim i poslovnim korisnicima. Iz tog razloga informacioni sistemi servis provajdera predstavljaju posebno kritična mesta pošto su često meta različitih napada. Na osnovu izveštaja velikih servisa i

*content* provajdera, kao i proizvođača opreme, dominantan tip napada kada su u pitanju provajderska okruženja je DoS (*Denial of Service*) napad. Kada je u pitanju ukupan broj DoS napada, istraživanja pokazuju da postoji povećanje od 71% u odnosu na prethodnu godinu [4].

U ovom radu su opisane osnovne karakteristike DoS napada. Prikazano je RTBH (*Remotely Triggered Black Hole*) rešenje za filtriranje saobraćaja, kao primer tehnike za mitigaciju DoS napada. Opisana su dva načina rada, *source-based* i *destination-based* RTBH, a prikazan je i unapređeni koncept ABH (*Advanced BlackHoling*).

## II. DoS NAPADI

Osnovna ideja DoS napada je da se legitimnim korisnicima onemogući normalno korišćenje servisa i pristup informacijama. Najčešće se ovakav vid napada realizuje slanjem velike količine podataka (*flooding*) kako bi se mreža „zagušila“ i bila sprečena da normalno funkcioniše. Ukoliko se napad realizuje sa više izvora tada je reč o distribuiranom DoS napadu (*Distributed Denial of Service* – DDoS).

Postoji veliki broj različitih DoS i DDoS napada [5], a upravo zbog njihove različitosti i konstantne pojave novih napada nije moguće napraviti jedinstvenu metodologiju za rešavanje ovog problema. Ipak, svi DDoS napadi se mogu svrstati u jednu od sledeće tri kategorije:

- Volumetrijski napadi – imaju za cilj da zauzmu propusni opseg u mreži ili između mreže i Interneta kako bi usled zagušenja bio onemogućen normalan rad servisa.
- TCP *State-Exhaustion* napadi – imaju za cilj da „zauzmu“ sve konekcije koje postoje u mnogim infrastrukturnim uređajima, kao što su serveri, *firewall* uređaji, *load balancer* uređaji itd.
- Napadi na aplikativnom sloju – imaju za cilj da onemoguće rad aplikacije ili servera korišćenjem njihovih nedostataka i slabih tačaka. Ovi napadi su najsofisticiraniji i teško se detektuju pošto je za realizaciju napada potreban mali broj uređaja, a saobraćaj koji se pri tome generiše može izgledati kao legitiman saobraćaj u mreži.

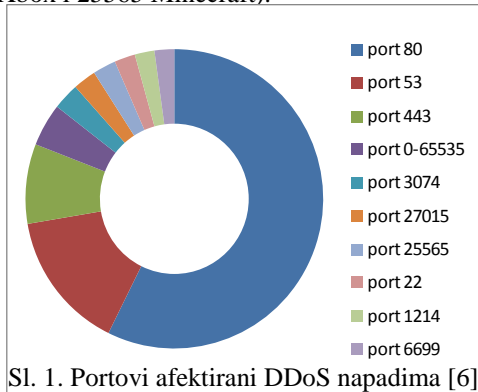
O važnosti problema DoS napada govori i činjenica da je u 2016. godini u provajderskom okruženju zabeležen veliki porast količine saobraćaja prilikom ovakvog napada. Prema izveštaju [6] prijavljen je napad od 500Gb/s, dok izveštaj [4] navodi da je zabeležen napad od čak 623Gb/s.

Na Sl. 1. su prikazani portovi, odnosno aplikacije, koje su afektirane DDoS napadima. Na prvom mestu je HTTP

Jelena T. Šuh (autor za kontakte), Telekom Srbija a.d, Katićeva 14, 11000 Beograd, Srbija (telefon: 381-11-3069110, e-mail: jelenasuh@telekom.rs).

Snežana Lj. Stefanovski Pajović, Telekom Srbija a.d, Bulevar umetnosti 16a, 11000 Beograd, Srbija (telefon: 381-11-3200330, e-mail: snezanas@telekom.rs).

saobraćaj, a zatim slede DNS, kao i HTTPS saobraćaj, koji su prisutni u gotovo svim mrežama. Pored toga, među prvih deset najugroženijih portova su i *gaming* aplikacije (3074 Xbox i 25565 Minecraft).



Sl. 1. Portovi afektirani DDoS napadima [6]

### III. TEHNIKE ZA MITIGACIJU DOS NAPADA

Navedeni podaci o DoS napadima jasno pokazuju da je u pitanju ozbiljan problem i da je u računarskim mrežama, a posebno u provajderskom okruženju, neophodno primeniti određene tehnike za mitigaciju ovih napada.

Pristupne liste (*Access Control List* – ACL) predstavljaju jedno od najčešće korišćenih rešenja u borbi protiv DoS napada. ACL čini niz pravila koji se konfigurišu na ruteru ili *firewall* uređaju sa ciljem da se određeni saobraćaj u mreži filtrira [7]. U pitanju je rešenje, koje može biti zahtevno u pogledu realizacije u kompleksnim mrežama.

*FlowSpec* predstavlja metod za distribuciju informacije o toku podataka (*flow*) korišćenjem BGP protokola [RFC5575]. Osnovni motiv za primenu ovog rešenja je da se obezbedi distribucija pravila za filtriranje saobraćaja unutar i između provajderskih okruženja sa ciljem da se filtriraju DoS i DDoS napadi.

Postoje i mnoga specifična vendorska rešenja za mitigaciju DoS napada, koja se razlikuju u stepenu kompleksnosti. Njihov najveći nedostatak je što su u pitanju *proprietary* rešenja, koja uglavnom zahtevaju velika finansijska ulaganja.

Tabela 1 daje uporedni prikaz karakteristika rešenja za mitigaciju DoS napada.

TABELA 1: TEHNIKE ZA MITIGACIJU DOS NAPADA.

	<i>Skalabilnost</i>	<i>Troškovi</i>	<i>Primena</i>
ACL	loša	nema	kompleksna
FlowSpec	zavisi	nema	zavisi
Vendorska rešenja	zavisi	visoki	jednostavna

Na osnovu prikazanih karakteristika jasno je da ne postoji jednoznačan odgovor na pitanje koji mehanizam zaštite od DoS napada treba primeniti. Takođe, za uspešnu realizaciju zaštite mreže neophodna je saradnja između vendora, servis provajdera i korisnika [8].

### IV. RTBH

Prilikom DoS napada vreme predstavlja kritičan faktor, pa je zato potrebno primeniti rešenje koje zahteva intervenciju na jednom mestu u mreži, a saobraćaj generisan prilikom napada odbacuje na obodu mreže kako

bi se sprečio prolazak ovog saobraćaja kroz celu mrežu. RTBH rešenje radi po ovom principu. U praksi je utvrđeno da je ovo jedno od češće primenjenih rešenja u provajderskom okruženju, kao i da predstavlja osnovu za druge tehnike za mitigaciju DoS napada.

*Blackhole* filtriranje se nalazi u osnovi RTBH rešenja i predstavlja proces prosleđivanja paketa na logički *null0* ili *discard* interfejs rutera (u zavisnosti od proizvođača), odnosno odbacivanje određenih paketa.

RTBH je tehnika koja koristi *update* poruke protokola rutiranja za unos zapisa u tabelu rutiranja na obodu mreže tj. na PE (*Provider Edge*) ruteru ili bilo gde u mreži, kako bi neželjeni saobraćaj bio odbačen pre ulaska u mrežu. Filtriranje saobraćaja može da se vrši na osnovu izvorne ili odredišne IP adrese, pa tako postoji *source-based* i *destination-based* RTBH. Bez obzira na tip RTBH rešenja prilikom implementacije se izdvajaju dva tipa uređaja:

- Triger uređaj – Ruter koji pokreće *blackhole* filtriranje i vrši oglašavanje mrežnih prefiksa. Preporuka je da bude poseban uređaj u mreži, mada se može koristiti i neki od postojećih rutera provajderske mreže.
- PE ruter – Uređaj koji vrši filtriranje saobraćaja slanjem ka *null0* ili *discard* interfejsu.

#### A. *Destination-based* RTBH

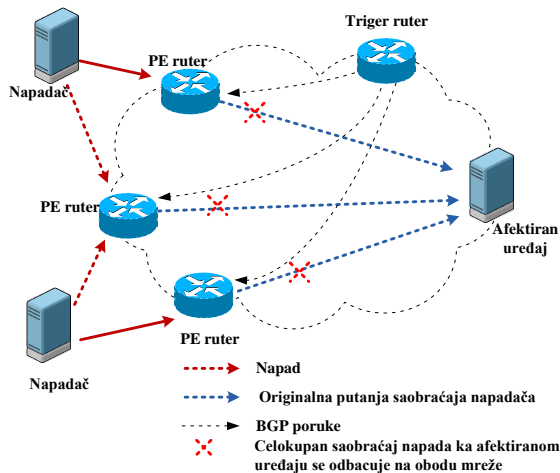
Kod *destination-based* RTBH opcije, blokira se sav saobraćaj koji je namenjen određenoj IP adresi, tako što se vrši filtriranje na obodu mreže. Na ovaj način se vrši mitigacija DoS napada, ali uz napomenu da ni koristan saobraćaj neće biti isporučen toj odredišnoj adresi, što je nedostatak ovog rešenja. Primarni zadatak *destination-based* RTBH rešenja je da se brzo odbaci neželjeni saobraćaj, a to se postiže definisanjem statičke rute za taj saobraćaj na triger ruteru. Triger ruter je zadužen da tu informaciju prosledi ostalim PE ruterima koji učestvuju u RTBH putem iBGP (*internal Border Gateway Protocol*) protokola, tako što ovu rutu oglašava sa određenom *next-hop* IP adresom ili BGP *community* atributom da PE ruter zna da saobraćaj treba da odbaci [RFC 3882]. Triger ruter treba da preko iBGP protokola komunicira sa svim ruterima u mreži, a ako u mreži postoje RR (*Route Reflector*) uređaji, triger ruter treba da ima iBGP sesije sa njima. U slučaju da se koristi *next-hop* IP adresa, ona treba da ukazuje na drugu prekonfigurisanu statičku rutu, koja saobraćaj usmerava na *null0* ili *discard* interfejs. Sl. 2. prikazuje *destination-based* RTBH princip funkcionisanja.

Implementacija *destination-based* RTBH rešenja sa *next-hop* IP adresom se vrši u tri koraka [9]:

- Priprema – Pored uspostavljanja iBGP sesija između triger rutera i ostalih PE rutera u mreži (ili RR rutera) potrebno je na svim PE ruterima konfigurirati statičku rutu za rezervisan IP adresni opseg koja ukazuje na *null0* interfejs. Prilikom definisanja ove rute najčešće se koriste IP adrese iz TEST-NET-1 [RFC5737] opsega.
- Aktivacija – Kada dođe do DoS napada mrežni administrator dodaje statičku rutu na triger ruter, koji je preko BGP protokola šalje svim iBGP peer-ovima, pri čemu za odredišnu adresu, koja je pod napadom, podešava *next-hop* IP adresu na adresu iz TEST-NET-1 opsega. PE

ruteri dobijaju iBGP poruku i podešavaju *next-hop* IP adresu za ugroženu odredišnu IP adresu na adresu iz TEST-NET-1 opsega. Na PE ruteru postoji statička ruta za ovu TEST-NET-1 IP adresu koja ukazuje na *null0* interfejs. Ovako se sav saobraćaj usmeren na odredišnu IP adresu koja je pod napadom odbacuje na PE ruteru.

- **Brisanje** – Po prestanku DoS napada administrator mora da ukloni iz konfiguracije trigger rutera prethodno unetu statičku rutu. Nakon toga i PE ruteri će iz tabele rutiranja izbaciti rutu koja saobraćaj ka odredišnoj IP adresi, koja je bila pod napadom, usmerava na TEST-NET-1 IP adresu i ubaciti rutu na osnovu internog protokola rutiranja.



Sl. 2. Destination-based RTBH [9]

### B. Source-based RTBH

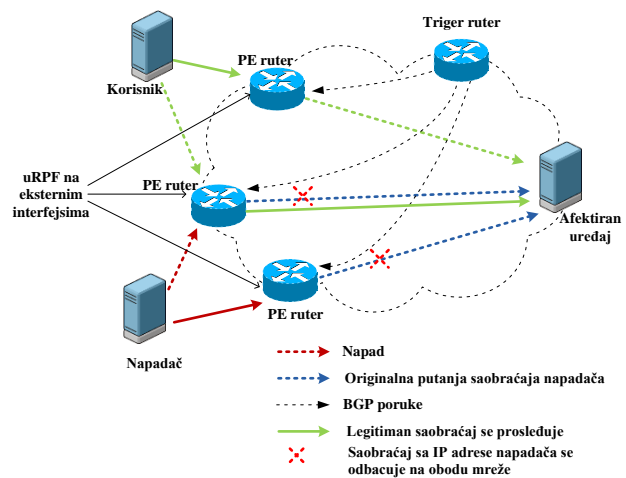
*Source-based* RTBH opcija omogućava filtriranje saobraćaja na obodu mreže na osnovu određene izvorne IP adrese ili opsega IP adresa sa kojih se vrši DoS napad [RFC 5635]. Kod *destination-based* RTBH rešenja u slučaju DoS napada celokupan saobraćaj ka odredišnoj IP adresi se filtrira, tako da *source-based* RTBH rešenje predstavlja bolju opciju, s obzirom da dopušta da legitiman saobraćaj sa ostalih izvorišnih IP adresa prolazi bez problema. Primena *source-based* RTBH rešenja je zasnovana na uRPF (*Unicast Reverse Path Forwarding*) mehanizmu, i to najčešće u *loose* modu. *Loose* uRPF proverava dolazni paket i prosleđuje ga ako u FIB (*Forwarding Informatin Base*) tabeli postoji zapis o izvornoj IP adresi tog paketa. Ukoliko ruter nema taj zapis u FIB tabeli ili ako ukazuje na *null0* interfejs, uRPF provera ne prolazi i paket se odbacuje.

Na Sl. 3. je prikazan princip funkcionisanja *source-based* RTBH rešenja. Implementacija ovog rešenja sa *next-hop* IP adresom se vrši u tri koraka [9]:

- **Priprema** – Pored uspostavljanja iBGP sesija između trigger rutera i ostalih PE rutera u mreži (ili RR rutera) potrebno je na svim PE ruterima konfigurisati statičku rutu za rezervisan IP adresni opseg (TEST-NET-1 opseg) koja ukazuje na *null0* interfejs. Na svim eksternim interfejsima na PE ruterima treba da se konfigurirše *loose* uRPF.
- **Aktivacija** – Kada dođe do DoS napada mrežni administrator dodaje statičku rutu na trigger ruter, koji je preko BGP protokola šalje svim iBGP *peer*-ovima, pri

čemu za izvornu adresu napada podešava *next-hop* IP adresu na adresu iz TEST-NET-1 opsega. PE ruteri dobijaju iBGP poruku i podešavaju *next-hop* IP adresu za izvornu IP adresu napada na adresu iz TEST-NET-1 opsega. Na PE ruteru postoji statička ruta za ovu TEST-NET-1 IP adresu koja ukazuje na *null0* interfejs. Na ovaj način saobraćaj koji dolazi sa navedene izvorne IP adrese ne prolazi uRPF proveru na PE ruteru i odbacuje se.

- **Brisanje** – Po prestanku DoS napada mrežni administrator mora da ukloni iz konfiguracije trigger rutera prethodno unetu statičku rutu. Nakon toga i PE ruteri će iz tabele rutiranja izbaciti rutu koja saobraćaj sa izvorne IP adrese napada usmerava na TEST-NET-1 IP adresu i ubaciti rutu na osnovu određenog internog protokola rutiranja. Ako je nova ruta ispravna, *loose* uRPF provera će biti uspešna i saobraćaj sa prethodno blokiranе izvorne IP adrese će se prosleđivati regularno.



Sl. 3. Source-based RTBH [9]

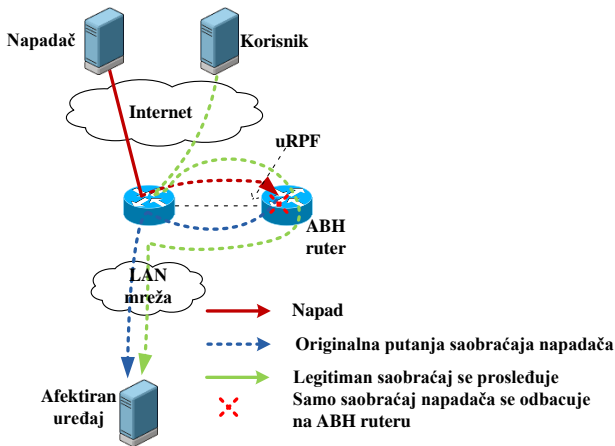
### C. Diskusija

Prednosti primene RTBH rešenja su višestruke, a ogledaju se u činjenici da je u slučaju DoS napada potrebno na samo jednom ruteru (trigger ruter) uneti odgovarajuću statičku rutu, čime se problem rešava. Nije potrebna izmena konfiguracije većeg broja uređaja niti ažuriranje velikog broja pristupnih lista, što je od značaja kada su u pitanju velike mreže. Zato je ovo rešenje adekvatno primeniti u slučaju čestih promena, odnosno dinamičkih napada ili kada su u pitanju višestruki napadi na veći broj korisnika. Međutim, ovo rešenje ima izvesna ograničenja, a ona su zasnovana na činjenici da se filtriranje saobraćaja radi na trećem sloju OSI modela odnosno na osnovu IP adrese. Problem može biti i to što bez primene dodatnih mehanizama dolazi do odbacivanja celokupnog saobraćaja sa ili ka određenoj IP adresi.

### V. ABH

Kako bi se prevazišli nedostaci klasičnog RTBH rešenja, koje vrši filtriranje celokupnog saobraćaja za određenu IP adresu, postoji ideja da se *source-based* RTBH koncept izmeni u skladu sa principima koji se primenjuju u vendorskim rešenjima za mitigaciju DoS napada. Kod vendorskih rešenja se saobraćaj u mreži preusmerava ka

određenom uređaju, koji na osnovu primljenog saobraćaja i određenih pravila uočava anomalije, vrši filtriranje saobraćaja i „očišćen“ saobraćaj vraća u mrežu tj. ka korisnicima. Kombinacija *source-based* RTBH koncepta i vendorskog načina mitigacije napada predstavlja osnovnu ideju ABH koncepta [10], koji je prikazan na Sl. 4.



Sl. 4. ABH koncept [10]

ABH koncept se može ukratko opisati na sledeći način: U slučaju DoS napada saobraćaj se preusmerava ka ABH ruteru na kome se aktivira uRPF provera. Zatim se izvorna IP adresa napadača preusmerava na *null0* interfejs tj. vrši se odbacivanje samo saobraćaja sa IP adrese napadača.

Da bi se postigla veća fleksibilnost rešenja može se ABH ruter povezati sa *injector* uređajem (što može biti i server), koji ima zadatak da obezbedi informacije o IP adresama napadača na osnovu dostupnih baza napada, *netflow* statistike, geografskih podataka i sl.

Postoji *proof of concept* za ABH rešenje, međutim, još uvek ne postoje izveštaji servis i *content* provajdera o primeni i efikasnosti ovog rešenja, ali se oni očekuju u narednom periodu.

## VI. ZAKLJUČAK

Cilj ovog rada je da ukaže na potrebu da se pitanju bezbednosti u računarskim mrežama posveti veća pažnja. Ukazano je na ozbiljnost problema DoS i DDoS napada sa kojim se servis provajderi i korisnici svakodnevno suočavaju. Iako postoji niz alata i tehnika za mitigaciju ovih napada [11], zbog njihove prirode i činjenice da konstantno evoluiraju, nije moguće dati jednoznačan odgovor i preporuku kako zaštititi mreže od ovih napada. U radu je opisano *source-based* i *destination-based* RTBH rešenje, kojim se vrši mitigacija DoS napada, ali uz određena ograničenja. U cilju prevazilaženja ovih nedostataka predstavljen je ABH koncept, od koga se očekuje da u narednom periodu pokaže efikasnost i posluži kao ideja za razvoj novih tehnika mitigacije DoS napada.

## LITERATURA

- [1] "NIST Cybersecurity Framework Glossary." [Online]. Available: [https://www.nist.gov/sites/default/files/documents/itl/draft\\_framework\\_ork\\_glossary.pdf](https://www.nist.gov/sites/default/files/documents/itl/draft_framework_ork_glossary.pdf).
- [2] "Committee on National Security Systems (CNSS) Glossary CNSSI No. 4009," 2015. [Online]. Available: <https://cryptosmith.files.wordpress.com/2015/08/glossary-2015-cnss.pdf>.
- [3] P. Bowen, J. Hash, and M. Wilson, "Information Security Handbook: A Guide for Managers," *NIST Special Publication 800-100*, 2006. [Online]. Available: <http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-100.pdf>.
- [4] "Akamai's [state of the internet] / security Q3 2016 executive summary," 2016. [Online]. Available: <https://www.akamai.com/us/en/multimedia/documents/state-of-the-internet/q3-2016-state-of-the-internet-security-executive-summary.pdf>.
- [5] J. Mirkovic and P. Reiher, "A taxonomy of DDoS attack and DDoS defense mechanisms," *Computer Communication Review*, vol. 34, no. 2, pp. 39–53, 2004.
- [6] "Arbor Networks Worldwide Infrastructure Security Report Volume XI," 2016. [Online]. Available: [https://www.arbornetworks.com/images/documents/WISR2016\\_EN\\_Web.pdf](https://www.arbornetworks.com/images/documents/WISR2016_EN_Web.pdf).
- [7] O. Santos and J. Stuppi, *CCNA Security 210-260 Official Cert Guide*. Cisco Press, 2015.
- [8] S. Mansfield-Devine, "DDoS: threats and mitigation," *Network Security*, vol. 2011, no. 12, pp. 5–12, 2011.
- [9] "White paper Remotely Triggered Black Hole Filtering-Destination Based and Source Based," *Cisco Systems*. [Online]. Available: [http://www.cisco.com/c/dam/en\\_us/about/security/intelligence/blackhole.pdf](http://www.cisco.com/c/dam/en_us/about/security/intelligence/blackhole.pdf).
- [10] F. Contat, "Advanced BlackHoling - ABH," in *RIPE 73*, 2016.
- [11] A. Cardigliano, L. Deri, and T. Lundstrom, "Commoditising DDoS mitigation," in *2016 IEEE International Wireless Communications and Mobile Computing Conference (IWCMC)*, 2016, pp. 523–528.

## ABSTRACT

Computer networks are complex systems having the task to provide regular operation of large number of services and applications. In addition to the standard requirements in terms of availability and scalability, security in computer networks represents one of the most important components of the network design. Providing network security is a process of applying different hardware and software solutions in order to protect network infrastructure and resources from unauthorized access and misuse. The aim of this paper is to highlight the problem of security in computer networks, where the importance of protection against DoS (*Denial of Service*) attacks is particularly emphasized. The paper describes RTBH (*Remotely Triggered Black Hole*) solution, as an example of efficient technique for mitigation of DoS attacks. Timely detection and action in order to prevent malicious attacks on specific network resources lead to the preservation of the quality of network operation and the services implemented in that network.

## PROTECTION AGAINST DOS ATTACKS USING THE REMOTELY TRIGGERED BLACK HOLE TECHNIQUE FOR TRAFFIC FILTERING

Jelena T. Šuh and Snežana Lj. Stefanovski Pajović

# Index modulation for performance improvement of OFDM based systems

Jovana Mrkić, Enis Kočan, *Member, IEEE*

**Abstract** — Orthogonal frequency division multiplexing (OFDM) is modulation technique proven to enable high data rates even in frequency selective wireless fading channels. Recently, OFDM with index modulation (OFDM-IM) is proposed for OFDM performance improvement in terms of spectral and energy efficiency, but at the price of bit error rate (BER) degradation. In OFDM-IM, fixed number of subcarriers is chosen to carry constellation symbols. At the same time, indices of the chosen subcarriers carry additional information bits. In this paper we examine BER performance of OFDM-IM system in case of frequency selective fading channel, with Rayleigh fading statistics. Through Monte Carlo simulations, we compare BER performances of OFDM-IM and OFDM systems, analyzing trade-off between BER performance degradation due to implementation of index modulation on one side, and spectral and energy efficiency improvement attained through its deployment on the other side. The obtained results have shown that for high signal-to-noise ratios (SNR), OFDM-IM reaches, or even outperforms OFDM system in terms of BER performance.

**Keywords** — OFDM, index modulation, spectral efficiency, BER, energy efficiency.

## I. INTRODUCTION

Multicarrier transmission techniques are now widely accepted in high speed wireless communications systems. Among them, OFDM (Orthogonal Frequency Division Multiplexing) represents a dominant and well-established multicarrier technique, due to its ability to deal with deteriorate effects of multipath propagation, as well as due to its high spectral efficiency. Namely, OFDM is very robust to frequency selectivity of multipath fading channel, as it turns this type of channel in a set of parallel frequency-flat channels, and it can efficiently cope with inter-symbol interference (ISI) by introducing guard interval (GI), thus providing high speed communications in highly unfavorable channel conditions. As such, OFDM is integrated in many contemporary wireless communication standards, such as LTE, LTE-Advanced, IEEE 802.11a/g/n/ac WLAN (Wireless Local Area Network) standards, WiMAX, DAB (Digital Audio Broadcasting), DVB-T (Digital Video Broadcasting – Terrestrial) standards, etc., [1]. There is still strong

research interest for performance improvement of OFDM modulation, in order to enable its implementation in different upcoming communication systems, and to provide smooth transition towards new generation of wireless communication systems.

One of the newly proposed concept is OFDM with Index Modulation (OFDM-IM), which is introduced as an extension of spatial modulation (SM) principle, originally referring to multiple-input-multiple-output (MIMO) systems. SM represents an alternative to existing MIMO transmission techniques. Essence of SM is that information is conveyed by means of standard amplitude/phase modulations, as well as by selection of indices of active antennas, [2]. The same principle has been proposed for OFDM subcarriers. Namely, information is transmitted by  $M$ -ary signal constellations, in the same way it is transmitted in classical OFDM system, as well as by combination of active subcarriers. Each subcarrier has its own index and transmission subcarriers are selected by their indices, depending on incoming information bits. In this way, the system spectral efficiency is increased, as additional information are assumed to be transmitted to destination, just through selection of certain subcarriers. Thus, for example, in OFDM-IM system having  $2^N$  subcarriers, let us assume that only one subcarrier is used for information transmission. Each subcarrier index can be represented by combination of  $N$  bits. Information bits are splitted into two flows, where one flow is used for creating  $M$ -ary symbol constellation to be transmitted, and the other flow, having  $N$  bits, denotes which subcarrier should be selected for transmission of the created symbol.

OFDM-IM has been first time considered in [3], and after that, many researchers have analyzed its performance, as well as different modifications have been proposed, [4]-[10]. According to [4], OFDM-IM uses frequency selectivity as a benefit, boosting in this way performances of the system in time/doubly dispersive fading channels. If compared to classical OFDM system, spectral efficiency can be increased due to additional bits transmitted, as well as energy efficiency, because only a part of subcarriers is being used. Increase in spectral efficiency of OFDM-IM system has been obtained at the expense of slightly worse BER performance, compared to OFDM. However, for high signal-to-noise ratios (SNR) values, OFDM-IM can reach, or even outperform classical OFDM system, in terms of BER performance [5].

In [4], two different ways for mapping information bits into groups of active subcarriers were proposed. The first method is a simple look-up table, which contains connections between specific combinations of incoming

Jovana Mrkić, Crnogorski Telekom (phone: 382-69-183507 e-mail: jvnmrkić94@gmail.com).

Enis Kočan, Faculty of Electrical Engineering, University of Montenegro, Džordža Vasiingtona bb, 81000 Podgorica, Crna Gora (phone: 382-66-152119; e-mail: enisk@ac.me).

bits and corresponding indices that are activated for transmission. Look-up table is implemented together with Maximum-Likelihood (ML) detector at the receiver. In case of a larger group of OFDM subcarriers, look-up table should contain a huge number of registrations, which makes this method impractical. Therefore, another technique based on combinatorial number theory has been proposed, which determines combination of active indices of subcarriers on the basis of input bits, whereas Log-Likelihood Ratio (LLR) detector is employed at the receiver side. In [6], generalization schemes for OFDM-IM have been presented, OFDM with generalized index modulation 1 (OFDM-GIM1) and OFDM with generalized index modulation 2 (OFDM-GIM2). The idea of OFDM-GIM1 is that number of active subcarriers in OFDM subblocks can be different and not constant, as it was the case in OFDM-IM. Another generalization of OFDM IM, called OFDM-GIM2 takes into consideration separation of „in-phase“ and „quadrature“ components of subcarriers, so that index modulation is applied independently to these two components. Further on, in [7], an interleaved subcarrier index modulation technique has also been proposed to mitigate poor BER characteristics for lower SNR values with OFDM IM.

As it can be seen from the given overview, OFDM-IM has attracted lately a lot of attention of researchers and some possible applications of this concept have already been suggested. OFDM-IM can be interesting candidate for low-power, wide range communications, as only few subcarriers can be used for information transmission in different IoT (Internet of Things) based applications, [10]. It is also interesting candidate for 5G mobile communication systems, [10]. Some authors have also explored and suggested the possibility of application of OFDM-IM to V2X communications, [11]. In [12], OFDM-IM has been proposed for Underwater Acoustic (UWA) communications. OFDM has already been recognized as a strong candidate for UWA communications, because of its robustness to intersymbol interference. However, OFDM-IM system can be regarded as even stronger candidate, because certain number of subcarriers remains inactive, which also mean a lower level of intercarrier interference (ICI).

In our research work, we have created simulation model of OFDM-IM system, and we examined BER performance in case of Rayleigh fading channel, for different realizations of OFDM index modulation, assuming LLR detection. The obtained BER performances are compared to BER performance of classical OFDM system in the same channel conditions, and we also provide insight in spectral efficiency enhancement achieved with OFDM-IM. The paper is organized as follows. Section 2 provides description of OFDM IM system model. Section 3 gives BER simulation results of our model and their discussion, while Section 4 concludes the paper.

## II. SYSTEM MODEL

Figure 1 present block-scheme of OFDM-IM transmitter.

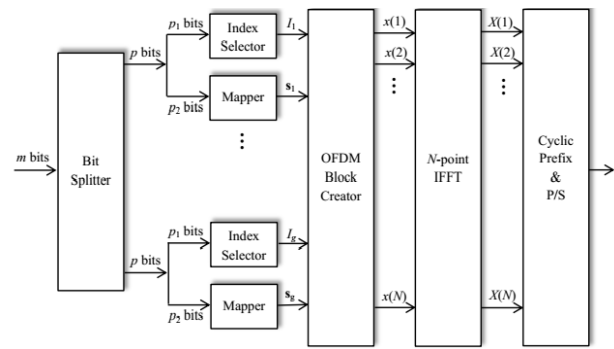


Figure 1. Block-scheme of OFDM-IM transmitter

In OFDM-IM transmitter, a total number of  $N$  OFDM subcarriers are divided into  $g$  groups, each having  $n=N/g$  subcarriers. Information entering the system are splitted into  $g$  groups of  $p$  bits in the bit splitter, thus having  $m=p \cdot g$ . After bit splitter block, in each of the  $g$  branches,  $p$  bits are again subdivided into two groups having  $p_1$  and  $p_2$  bits, where  $p_1+p_2=p$ . As it is earlier explained, in OFDM-IM system, information are transferred to receiver not only by data symbols, but also through selection of appropriate subcarriers to be used for symbol transmission, as each subcarrier has its own index. In that manner,  $p_2$  information bits in each subbranch are mapped to  $M$ -ary symbol constellations, while  $p_1$  bits are going into index selector block, and are used for selecting  $k$  active subcarriers, which will be used to transmit  $M$ -ary symbols, from total number of  $n$  subcarriers contained in each of  $g$  groups.

Number of bits used for selection of indices can be presented as:

$$p_1 = \log_2 (C(n, k)), \quad (1)$$

where  $C(n, k)$  is a binomial coefficient. Remaining  $p_2$  bits are being mapped into  $M$ -ary symbols that are going to be transmitted by  $k$  active subcarriers. From this, it follows that:

$$p_2 = k \cdot \log_2 M. \quad (2)$$

At the level of the entire OFDM block,

$$g \cdot p_1 = g \cdot \log_2 (C(n, k)) \quad (3)$$

bits are transmitted by selecting the active indices, and

$$g \cdot p_2 = g \cdot k \cdot \log_2 M \quad (4)$$

are transmitted by  $M$ -ary symbols.

The created  $M$ -ary symbols are going to be transmitted by selected subcarriers, while the rest of subcarriers are remaining inactive, i.e. they are set to zero. This process is carried out by OFDM block creator. After this block, standard OFDM modulation is performed, which means that signal goes through  $N$ -point IFFT, then parallel-to-serial conversion and cyclic prefix addition.

Index selector can be implemented in two ways, [4]. One way is realization by employing a simple look-up table, which is used on both transmitter and receiver sides. At the transmitter side, the active indices are read from the table depending on incoming  $p_1$  bits, while at the reception the opposite action is taking place. This method is simple but efficient only for small values of  $n$  and  $k$ , because it is not practical to work with large look-up tables. At the



reception side maximum likelihood (ML) detector is used, because it needs to have an information about all possible indices combinations. Log-likelihood Ratio (LLR) detector can not be used with a look-up table, because it can not make a decision, if detected combination of indices does not exist in the table. Table 1 presents an example of look-up table.

TABLE 1: EXAMPLE OF LOOK-UP TABLE.

Bits	Indices
[0 0]	{1,2}
[0 1]	{1,3}
[1 0]	{2,3}
[1 1]	{1,4}

In the given example,  $n=4$ ,  $k=2$ ,  $p_1=2$ . Since two bits ( $p_1$ ) can map four combinations of active indices, two combinations (2,4), (3,4) will not be used. ML detector takes into consideration all possible realizations and makes a joint decision about active indices and constellation symbols for every subblock.

The other way of index selection is the method of combinatorial numbers, which represents mapping of natural numbers into strictly descending sequence of active indices. Bits entering the index selector, first get converted into decimal number  $Z$ , by means of binary-to-decimal convertor. The sequence of active indexes  $J(c_k, c_{k-1}, \dots, c_1)$ ,  $c_k \geq c_{k-1} \geq \dots \geq c_1$ ,  $c \in [0, \dots, n-1]$ , can be calculated using the equation:

$$Z = C(c_k, k) + C(c_{k-1}, k-1) + \dots + C(c_1, 1) \quad (5)$$

In (5),  $c_k$  values are calculated by finding the maximal  $c_k$  satisfying the condition:  $C(c_k, k) \leq Z$ , and after that finding maximal  $c_{k-1}$ , for which it is satisfied  $C(c_{k-1}, k-1) \leq Z - C(c_k, k)$  and so on. Actual indices of subcarriers to be used for transmission are then obtained by adding 1 to obtained  $c_k$  values, because the subcarriers set of indices are  $[1, \dots, n]$ . The same algorithm is used at the receiver side for obtaining number  $Z$  from active indices and afterwards for converting it into the sequence of bits, by means of decimal-to-binary converter.

The described method is implemented with LLR detector. LLR detector determines the logarithmic ratio between a-posteriori probabilities that on the given subcarrier is transmitted a symbol or zero. If  $y_f$  represents the received signal, and  $x$  represents the sent signal in frequency domain and  $s_i$ ,  $i \in [1, \dots, M]$  is modulation alphabet, then for each subcarrier  $\alpha$ , the value of  $\lambda(\alpha)$  is:

$$\lambda(\alpha) = \ln \frac{\sum_{i=1}^M P(x(\alpha) = s_i | y_f(\alpha))}{P(x(\alpha) = 0 | y_f(\alpha))} \quad (6)$$

The higher value of  $\lambda(\alpha)$  denotes the higher probability that subcarrier  $\alpha$  was active in the transmission. In each OFDM subblock, out of  $n$  subcarriers, the highest  $k$  values  $\lambda(\alpha)$  marks  $k$  subcarriers, that were selected to be active at the transmitter side. After that, evaluation of first  $p_1$  transmitting bits is provided from estimated active indices, by means of algorithm opposite to the one at transmitter side. The remaining  $p_2$  of transmitted bits are discovered by demodulation of constellation symbols on active

subcarriers.

### III. SIMULATION RESULTS

In this section, performances of OFDM-IM system in terms of bit error rate (BER) are evaluated. Active subcarriers are modulated using binary phase shift keying (BPSK), with different system parameters, while LLR detector is implemented at receiver side.

#### A. Simulation model

The subsequent presented simulation results assume perfectly synchronized OFDM-IM system. Simulation results are obtained through Monte Carlo simulations. Only the frequency domain part of the OFDM-IM system is taken into consideration, which is possible approach, as perfectly time and frequency synchronization is assumed. Subcarriers' channel transfer functions are generated as zero mean, independent, circularly symmetric complex Gaussian random variables, with variances 1/2, so their magnitudes are Rayleigh distributed.

#### B. BER performance assessment

We have chosen to present BER results of OFDM-IM system that achieve the highest spectral efficiency improvement compared to classical OFDM system. If OFDM system has  $g$  groups with  $n=16$  subcarriers in each group, the highest spectral efficiency is obtained through OFDM-IM if  $k=11$ , i.e. 11 subcarriers are active for transmission (Fig. 2). The spectral efficiency of this OFDM-IM scheme is directly proportional to the total number of bits transmitted by the given system,  $m$ . While classical OFDM system using BPSK with 16 subcarriers transmit 16 bits, OFDM-IM system transmit 12 bits just for selection of 11 active subcarriers, and 11 bits by constellation symbols carried by those active subcarriers, which is 23 bits in total. In case that OFDM has 8 subcarriers, in terms of spectral efficiency, the result is the best when  $k=4$ ,  $k=5$ ,  $k=6$  or  $k=7$  (Fig. 2). When  $k=4$ , system transmits  $p_1=6$  bits for selecting 4 active subcarriers, and  $p_2=4$  bits are transmitted by constellation symbols. For  $k=5$ ,  $p_1=5$ ,  $p_2=5$ , for  $k=6$ ,  $p_1=4$ ,  $p_2=6$ , for  $k=7$ ,  $p_1=3$ ,  $p_2=7$ , which is 10 bits in total in all those cases, while with classical OFDM there are 8 bits transmitted in total.

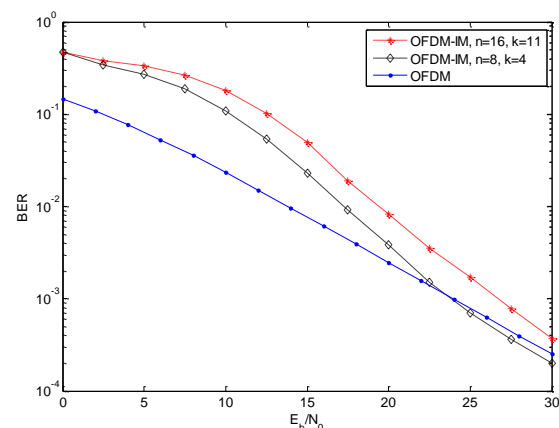


Figure 2. BER performances of OFDM-IM and classical OFDM

At Fig. 2, BER performance of classical OFDM system OFDM-IM systems having  $n=16$  and  $k=11$ , and  $n=8$  and  $k=4$ , are compared. As it can be seen from the figure, for BER value of  $10^{-3}$ , OFDM-IM system ( $n=16$  and  $k=11$ ) has about 3 dB worse performance than classical BPSK modulated OFDM. At the price of this BER performance decreasing, OFDM-IM with this configuration provides 31% better spectral efficiency than OFDM. It can be concluded that for  $E_b/N_0$  values above 30dB this OFDM-IM system will attain the same BER performances like classical OFDM system. On the other side, for  $E_b/N_0$  values higher than 24 dB, OFDM-IM with  $n=8$  and  $k=4$  can even achieve better BER performances than OFDM. This system provides 20 % better spectral efficiency than classical OFDM system.

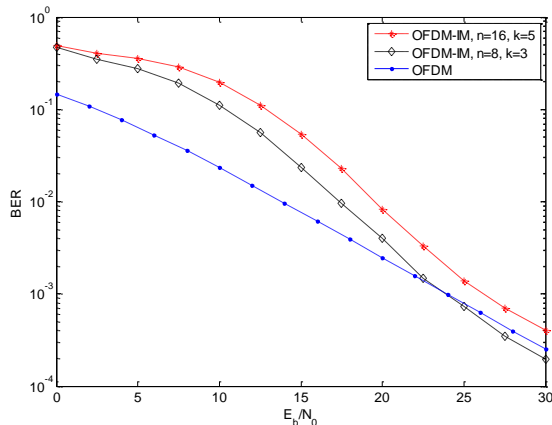


Figure 3. BER performances of OFDM-IM and classical OFDM schemes with approximately same spectral efficiency

Fig. 3 presents BER comparison between classical OFDM and OFDM-IM systems, with configurations that achieve the same, or approximately the same spectral efficiency as OFDM. OFDM-IM with  $n=16$  and  $k=5$  transmits 12 bits for selecting active indices and 5 bits to be carried by constellation symbol, which is 17 bits in total, while OFDM transmits 16 bits to be carried on 16 subcarriers. However, as OFDM-IM uses only 5 active subcarriers, it provides more than 68% energy efficiency improvement, at the price of worse BER performance for lower  $E_b/N_0$  values.

OFDM-IM scheme with  $n=8$  and  $k=3$  transmits 5 bits for selecting active indices and 3 bits for creating symbol constellations to be transmitted, which is 8 bits in total, same as classical OFDM with 8 subcarriers. It also can be seen from in Fig. 3 that this OFDM-IM configuration, for BER values of  $10^{-2}$ , has about 3 dB worse performance than OFDM, but for  $E_b/N_0$  higher than 24 dB, it can even achieves better BER performances than OFDM. In this case, it should be taken into consideration that OFDM-IM uses only 3 subcarriers for transmission, which means that it has 62% better energy efficiency than classical OFDM system.

#### IV. CONCLUSIONS

Index modulation, applied through spatial modulation, and/or in OFDM based systems, represents very interesting

solution for improvement of spectral and energy efficiency of future wireless communication systems. In OFDM-IM system, besides the information symbols transmitted on a chosen number of active subcarriers, additional information are transferred to destination by means of indices of the active subcarriers. This interesting solution has attracted attention for implementation in various systems, starting from 5G, over V2X systems, up to underwater acoustic communications (UWA).

In this paper we have examined BER performance of BPSK modulated OFDM-IM system, for different configurations, in terms of number of total subcarriers per group, and in number of active subcarriers., ali in case of Rayleigh fading channel. We compared simulation obtained BER results for OFDM-IM with BER performance of classical OFDM system. The obtained results have shown that for small  $E_b/N_0$  values, OFDM system has significantly better BER performance. However, for  $E_b/N_0$  values of 24 dB and above (depending on the number of active indices), OFDM-IM attains the same BER performance, or even outperforms OFDM system. Having in mind energy efficiency and/or spectral efficiency improvement that index modulation brings, it can be said that it presents very interesting solution for implementation in future wireless communication systems.

#### REFERENCES

- [1] P. Bharti, S. Malik, S. Malik, "Transmission Technology OFDM: Concept, Scope & its Applications", International Journal of Enhanced Research in Science Technology & Engineering, Vol. 4 Issue 6, pp. 305-309, June-2015.
- [2] P. Yang, M. D. Renzo, Y. Xiao, S. Li, and L. Hanzo, "Design Guidelines for Spatial Modulation", IEEE Communications Surveys & Tutorials, Vol. 17, no. 1, pp. 6 – 26, May 2014.
- [3] R. Abu-Alhiga and H. Haas, "Subcarrier-index modulation OFDM", in Proc. IEEE Int. Sym. Personal, Indoor and Mobile Radio Commun., pp. 177-181, Tokyo, Japan, Sept. 2009.
- [4] E. Basar, U. Aygolu, E. Panayirci, and H. V. Poor, "Orthogonal Frequency Division Multiplexing with Index Modulation", IEEE Transactions on Signal Processing, pp. 5536 – 5549, vol. 61, no. 22, Nov. 2013;
- [5] S. G. Domouchtsidis, G. D. Ntouni, V. M. Kapinas, and G. K. Karagiannidis, "OFDM-IM vs FQAM: A Comparative Analysis", International Conference on Telecommunications, Thessaloniki, Greece, May 16-18, 2016
- [6] R. Fan, Y. Jun Yu, and Y. L. Guan, "Generalization of Orthogonal Frequency Division Multiplexing With Index Modulation", IEEE Trans. on Wireless Comm., Vol. 14, no. 10, pp. 5250-5359, Oct. 2015.
- [7] Y. Xiao et al., "OFDM with interleaved subcarrier-index modulation", IEEE Commun. Lett., vol. 18, no. 8, pp. 1447-1450, Aug. 2014.
- [8] M. Wen, X. Cheng, M. Ma, B. Jiao, and H. V. Poor, "On the achievable rate of OFDM with Index Modulation", IEEE Transactions on Signal Processing, Vol. 64, No. 8, April 15, 2016
- [9] E. Basar, "Multiple-Input Multiple-Output OFDM with Index Modulation", IEEE Signal Processing Letters, vol. 22, no. 12, pp. 2259 – 2263, Dec. 2015
- [10] E. Basar, "Index Modulation Techniques for 5G Wireless Networks", IEEE Commun. Magazine, July 2016, pp. 168-175
- [11] X. Cheng, Miaowen Wen, Liuqing Yang, and Yuke Li, "Index Modulated OFDM with Interleaved Grouping for V2X Communications", in Proc. IEEE 17th Int. Conf. ITSC, 2014.
- [12] M. Wen, X. Cheng, L. Yang, Y. Li, X. Cheng, and Fei Ji, "Index Modulated OFDM for Underwater Acoustic Communications", IEEE Communications Magazine, pp. 132-137, May 2016.

# Meeting 5G Requirements through Millimeter Wave Communications

Aleksandra Lopušina, Enis Kočan, *Member, IEEE* and Milica Pejanović-Đurišić, *Member, IEEE*

**Abstract** — It is expected that 5G communication systems will bring significant performance improvement over 4G. For example, peak data rates are foreseen to be 10 times higher than the ones in 4G (i.e. up to 10Gb/s), and radio link latency should be 10 times lower (less than 1ms). Among other proposed solutions for meeting these requirements is implementation of millimeter wave (mmWave) communications, which can provide channel bandwidths in the order of GHz. In this paper we have tested performance of mmWave communication systems operating at 28GHz and 73GHz, in terms of achievable data rates, for three real case scenarios. We use *ns-3* simulator, with module that incorporates channel parameters for these two mmWave bands, obtained through series of measurements in urban environment. The obtained results have confirmed that mmWave communications can meet 5G expectations for providing services to low mobile users in outdoor environment, at distances up to few hundred meters, but in case where no significant obstacles exist between transmitter and receiver.

**Keywords** — 5G, millimeter wave, data rate, *ns-3*, 28 GHz, 73 GHz

## I. INTRODUCTION

RAPID growth of data traffic and overall number of Internet connected devices over mobile cellular networks is a certain sign of technology advancement, but it sets up various challenges and raises expectations for future wireless communication systems. In many developed countries, having great number of “heavy data users” performance of mobile networks are pushed to their limits, which accentuates the necessity of revolution in network design: starting from new solutions on physical layer, up to radical changes in network topology. According to *Cisco's 2020 Mobile Forecast Highlights*, [1], mobile data traffic will grow 8 times from 2015 to 2020 (from 3.7 exabytes per month in 2015 to 30.6 exabytes per month in 2020) and in 2020 it will be equivalent to 15x the traffic of the entire Global Internet in 2005. This immense forecasted growth comes with higher expectations from users as well: in terms of availability, reliability, latency and data rates. Latency requirements are as low as 1 ms on radio link; peak data rates in the order of Gbps and data rates at cell edge in the order of

tens of Mbps with very high reliability, [2]. High reliability means that moderate rates should be sustained even in crowded locations, in rural areas, or at high mobility. Some of the most promising techniques contributing to the fulfillment of these goals are: network densification, wide channel bandwidths at millimeter wave bands, massive Multiple Input Multiple Output (MIMO) systems with 256x256 antenna elements, narrowband transmission like Filter-Bank Multi-Carrier (FBMC), multiple access scheme like Sparse Code Multiple Access, more prominent use of Time Division Duplex (TDD) mode in a dynamic manner, coordination among cells with reduced Transmission Time Interval (TTI), etc, [2,3]. With commercial rollout of 5G planned for 2020, leading vendors and operators are working intensely on defining new solutions. Although there are still not many confirmed details, there are few indicators of general direction in which will 5G revolution move. Frequency bands of interest span from just few hundreds of MHz up to 100 GHz. Large amounts of spectrum, in the order of 1 GHz or more, are only available in the mmWave range, from 30 GHz to 300 GHz (suitable for providing high peak data rates in specific areas where traffic demands are very high).

Propagation properties of millimeter waves are suitable for 5G requirements: the smaller wavelength of mmWave signals means that more antennas can be fit into the same physical area which enables greater antenna gain for the same physical area. And, although path loss increases with the frequency, highly directional beamforms with large antenna array gain are key in combating those losses.

Reflection mechanism is fairly consistent for all frequencies in the mmWave band and hence is the most reliable means of receiving signals in non-line of sight (NLOS) conditions. Foliage loss increases with frequency and will be a detriment in mmWave communications, but it can be overcome with reflections and/or rapid rerouting to a different access point. Effects of diffraction around obstacles are mostly negligible in this range.

In higher ranges, few frequency bands have already stood out as eligible: 28 GHz and 73 GHz, which have already been intensively tested in urban environment of Manhattan, New York, [4]. Both frequency bands show good propagation characteristics, with signal being detected at least 100 to 200 meters from the potential cell site, even in absence of line of sight (LOS) connectivity. Based on this statistical channel model, New York University's Wireless team has created first 5G millimeter wave simulator, [5], implemented as a separate module in *ns-3* simulator, [6]. Using this new *ns-3* module, we

A. Lopušina, Crnogorski Telekom, Moskovska 29, 81000 Podgorica, Crna Gora (phone: 382-67-457167; e-mail: lopusina@t-com.me).

E. Kočan, Faculty of Electrical Engineering, University of Montenegro, Džordža Vašingtona bb, 81000 Podgorica, Crna Gora (phone: 382-20-245839; e-mail: [enisk@ac.me](mailto:enisk@ac.me))

M. Pejanović-Đurišić, Faculty of Electrical Engineering, University of Montenegro, Džordža Vašingtona bb, 81000 Podgorica, Crna Gora (phone: 382-20-245839; e-mail: [milica@ac.me](mailto:milica@ac.me))

analyze whether mmWave communications, in the specified frequency bands, can meet 5G requirements in terms of data rate for different real-case scenarios assuming outdoor propagation environment.

The rest of the paper is organized as follows: Section II introduces the *ns-3* simulator, as well as the millimeter wave module with its abilities and limitations. Section III describes different scenarios that were tested. In section IV, simulation results are presented, for both mmWave frequencies of interest and both LOS/NLOS propagation models. Section V summarizes the results and compares them to 5G expectations.

## II. SIMULATION TOOL

Simulation tool used in this research paper is *ns-3*, C++ based network simulator. 5G millimeter-wave support is implemented as a separate module and it is developed by New York University's Wireless team led by Theodore Rappaport [7].

Lack of defined standard for 5G physical layer caused the simulator to heavily reside on LTE, primarily on its Orthogonal Frequency Division Multiplexing (OFDM) model (which will probably suffer major changes for 5G, if chosen at all), but with significant modifications which facilitate achieving 5G requirements. Since it is expected that 5G will be directed towards TDD operations, in order to reduce the latency over the air interface, the implemented TDD frame structure is fully customized - from symbol length to guard interval and carrier frequency, as given in Table 1.

TABLE 1: PARAMETERS FOR CONFIGURING MMWAVE FRAME STRUCTURE

Parameter Name	Default Value	Description
<i>SymbolPerSlot</i>	30	Number of OFDM symbols per slot
<i>SymbolLength</i>	4.16 $\mu$ s	Length of one OFDM symbol in $\mu$ s
<i>SlotsPerSubframe</i>	8	Number of slots in one subframe
<i>SubframePerFrame</i>	10	Number of subframes in one frame
<i>NumReferenceSymbols</i>	6	The number of reference OFDM symbols per slot
<i>TDDControlDataPattern</i>	"ccddddd"	The control (c) and data(d) pattern
<i>SubcarriersPerSubband</i>	48	Number of subcarriers in each sub-band
<i>SubbandsPerRB</i>	18	Number of sub-bands in one resource block
<i>SubbandWidth</i>	13.89e6	The width of one sub-band in Hz
<i>NumResourceBlock</i>	4	Number of resource blocks in one slot
<i>CenterFreq</i>	28 GHz	The carrier frequency in Hz

Allocated bandwidth is 1 GHz in both analyzed frequency bands, which, given the default parameters in Table 1, comes down to 103 680 subcarriers available to user data.

In order to properly capture characteristics of mmWave propagation, several other features of the physical layer have been implemented. Radio characterization includes small and large scale channel variations. Link budget is given by, [7]:

$$P_{RX} = P_{TX} + G_{BF} - PL - SW \quad (1)$$

where  $P_{RX}$  is the total received power in dBm,  $P_{TX}$  is the total transmit power,  $G_{BF}$  is the beamforming gain, and finally  $PL$  and  $SW$  represent the path loss and shadowing, respectively. Transmit power used in simulations is 30 dBm.

Path loss and shadowing (expressed in dB) are calculated using the formula, [7]:

$$PL(d) = \alpha + \beta * 10 * \log_{10}(d) + \zeta, \zeta \sim N(0, \sigma^2) \quad (2)$$

where  $\zeta$  represents shadowing,  $d$  is the distance from receiver to transmitter (in meters), and the values of  $\alpha$ ,  $\beta$ , and  $\zeta$  are given in [3], and are presented in Table 2.

TABLE 2: PATH LOSS PARAMETERS

	LOS at 28 GHz	NLOS at 28 GHz	LOS at 73 GHz	NLOS at 73 GHz
$\alpha$	61.4	72	69.8	82.7
$\beta$	2	2.92	2	2.69
$\zeta$	5.8	8.7	5.8	7.7

mmWave signals are extremely susceptible to shadowing effect. For example, materials such as brick can attenuate signals 40–80 dB and attenuation through the human body can result in a 20–35-dB loss, [8, 9].

The beamforming gain from transmitter  $i$  to receiver  $j$  is given as, [7]:

$$G(t, f)_{ij} = \left| w_{rxij}^* * H(t, f)_{ij} * w_{txij} \right|^2 \quad (3)$$

where:  $H(t, f)_{ij}$  is the channel matrix of  $ij_{tx}$  link,  $w_{txij}$  is the beamforming vector of transmitter  $i$ , when transmitting to receiver  $j$  and  $w_{rxij}$  is the beamforming vector of receiver  $j$ , when receiving from transmitter  $i$ . Small-scale fading is generated based on the number of clusters, number of sub-paths per cluster, Doppler shift, power spread, delay spread and angle of arrival. Decoding error model is also implemented at the receiver. Interference is calculated in a way which takes into account the beamforming directions associated with each link i.e. treating other base stations as source of interference. Channel Quality Index (CQI) Feedback loop is also included, similar to the one used in existing LTE systems, same as CQI to modulation and coding scheme mapping. Medium Access Control (MAC) layer is also implemented in the simulation tool, taking care of scheduling and resource allocation, as well as adaptive modulation and coding.

Apart from the physical and link level details listed above, there are several other parameters than can be adjusted: propagation loss model, channel model, mobile station moving pattern, transmitter's power, receiver's noise figure, number of antennas in base station and mobile station, enabling small scale fading, enabling Hybrid automatic repeat request (HARQ), requested Bit Error rate, etc. Environment surrounding the network is also customizable in terms of possibility to: add obstacles, define number of floors and rooms in building, define its type and material used.

## III. SIMULATION SCENARIOS

Due to the limited range of mmWave signals, most of the cellular applications for mmWave systems focus on small-cells, assuming outdoor deployments with LOS connectivity. In order to evaluate performance of the

system described in Section II and its ability to deliver data rates expected from 5G systems, we created three real case scenarios.

As peak data rates advertised for 5G are primarily targeting low mobile users (stationary or walking), two scenarios include those user types and the third one tests performance of the system for medium mobile users. All scenarios were tested for both considered frequency bands.

Assumed scenarios are set in hypothetical city square, with signal coverage from two base stations. Base stations' positions, as well as mobile stations' starting positions, are depicted in Figure 1. Simulated data transfer time in all three scenarios is 80 seconds.

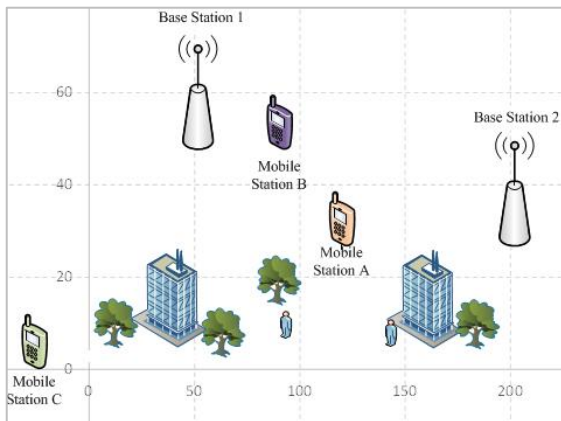


Figure 1: Test network scenarios (A, B, C)

The following scenarios are discussed:

#### A. Scenario A

Scenario A implements *Random Walk* model, i.e. pedestrian walk. It includes mobile station moving at constant velocity of 1.5 m/s at the defined area at random directions. During 80 seconds of simulation, the position is changed 120 times. Starting position of MS that follows this model is given in Figure 1 and is labeled as "*Mobile station A*".

#### B. Scenario B

Scenario B implements *Constant Velocity* model, i.e. MS moving at constant velocity of 36 km/h along  $x$  axis. Given that the simulation time was 80 seconds, the MS in this scenario travelled 800 meters. Starting position of MS that follows this model is given in Figure 1 and is labeled as "*Mobile station B*".

#### C. Scenario C

Scenario C implements *Constant Position* model, defined in such way that every 2 seconds static MS is moved 5 meters along  $x$  axis. This model also includes obstacles: two buildings and six obstacles which model people and trees. Both buildings are defined as residential, with 4 floors and 1 room per floor. Exterior walls are defined as *Concrete With Windows*. Starting position of MS that follows this model is given in Figure 1 and is labeled as "*Mobile station C*".

## IV. SIMULATION RESULTS

**Scenario A:** Simulation results obtained in this scenario are showing that in case of low mobile users, with good coverage and LOS propagation, achieved average data rates do not change as carrier frequency increases. In fact, in both cases, data rates are constant and maximal possible for this simulation tool. This is because in LTE, for all Signal-to-(Interference+Noise) Ratio (SINR) values higher than 20.5 dB, the system uses highest available signal constellations, thus achieving maximal data rates. For NLOS connectivity case, average data rates are also constant, but significantly lower. Also, propagation characteristics at different frequency bands came to light, so the average SINR at 73 GHz is 55% smaller than the one obtained at 28 GHz. The results are encouraging, as this is the primary scenario which 5G mmWave frequencies are focusing on: outdoor open spaces with few to none obstacles, but with lot of connected devices, such as main squares, concert and sport events, etc. Results are given in Figure 2.

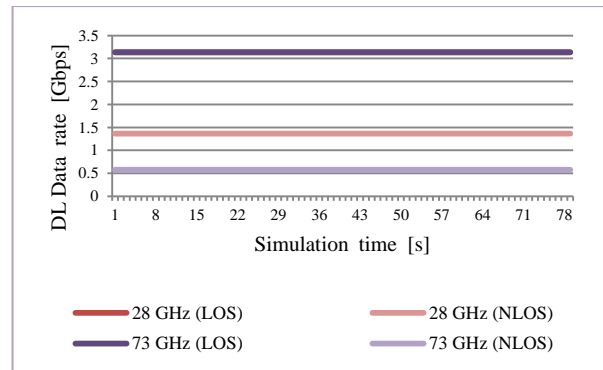


Figure 2: Scenario A results

**Scenario B:** This scenario shows the impact of MS-BS distance on average downlink data rate. As depicted in Figure 3, data rates for both 28 GHz and 73 GHz are constant only while that distance is less than 100m. After that, data rate decreases, reaching 60% of initial data rate at the ending position, 830m from the BS. At 73 GHz, data rate degradation pace is higher, resulting in 70% reduction. But, even at 830 meters from the BS, users could still achieve significant data rates, around 2 Gbps at 28 GHz and 1 Gbps at 73 GHz. For NLOS connections measurement results from [4] are confirmed: signals were detected even further than 100m from the potential cell site (230m for 28 GHz and 130m for 73 GHz).

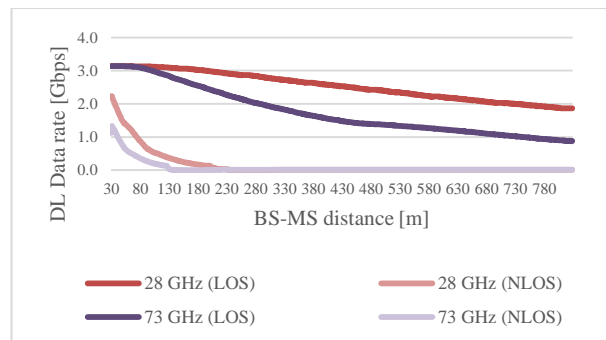


Figure 3: Scenario B results

TABLE 3: SIMULATION RESULTS

Scenario	Propagation	28 GHz		73 GHz	
		Average DL data rate [Gbps]	Average SINR [dB]	Average DL data rate [Gbps]	Average SINR [dB]
Scenario A (Random Walk Model)	LOS	3.13	39.89	3.13	31.49
	NLOS	1.36	11.22	0.58	5.04
Scenario B (Constant Velocity Model)	LOS	2.53	21.30	1.75	14.29
	NLOS	0.16	-11.52	0.07	-16.34
Scenario C (Fixed Position Model + Obstacles)	N/A	2.25	24.54	2.00	17.05

**Scenario C:** This is the only scenario considered, in which obstacles were included. Results show that any kind of obstacle significantly deteriorates performance: first degradation of achieved data rate is due blockage by building, but the second blockage (at 55m) occurs because of trees and people present.

These results are in line with previously described propagation characteristics of mmWave bands, proving itself as a very susceptible to any kind of blockage, due to the short wavelengths.

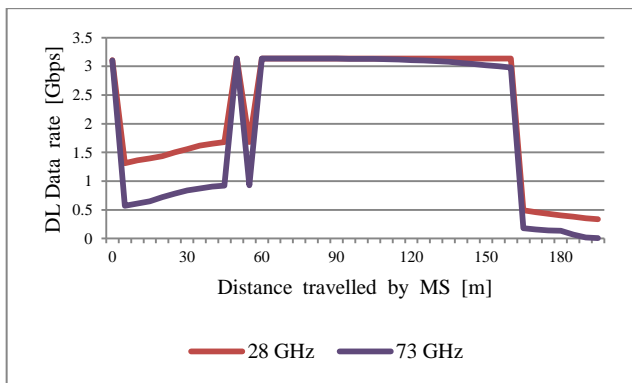


Figure 4: Scenario C results

In Table 3, we summarize achieved results for all three scenarios in both assumed frequency bands, in terms of average SINR values and average data rates in LOS and NLOS conditions. We observe that increasing carrier frequency has had almost the same effect in Constant Velocity model and Random Walk model, reducing data rates for more than a half for NLOS connections. For LOS connections, these degradations were much lighter, from 10% to 30%. From these results, we can conclude that in NLOS Constant Velocity scenario, mmWave model cannot fulfill 5G required performance targets. Its results are in the order of LTE data rates, which confirms earlier statement that mmWave communications are primarily intended for excellent and stable radio conditions, not for mobile users. On the other hand, all tested LOS scenarios showed excellent performance, proving that in such conditions, 5G requirements can be met through mmWave communications.

## V. CONCLUSIONS

With the potential to offer up to ten times greater capacity compared to 4G mobile communication systems, mmWave communications became a promising candidate

for the 5G mobile networks. Due to high propagation loss at high frequency bands, mmWave communications will be primarily deployed in outdoor environment and focused on static and low mobile users with LOS connectivity, so that the promised peak data rates could be delivered.

In this paper, we tested performances of mmWave communication systems operating at 28 GHz and 73 GHz frequency bands, for three real case scenarios. It is shown that the basic 5G system employing legacy OFDM is capable of delivering average user data rates which are more than 10 times better than the ones in LTE-Advanced, in LOS environment, with only 1W of transmit power. It should be noticed that this transmit power is far much lower than transmit powers used for outdoor base stations in LTE and LTE-Advanced systems.

This paper provides useful insight into the achievable performance of simplified mmWave network. Our survey shows that, due to particular features of the mmWave channel, these systems generally will need high base station density to achieve defined milestones, such as 10 Gbps per low mobility users and 100 Mbps at cell edge.

## REFERENCES

- [1] Cisco, *VNI Complete Forecast Highlights Tool / Mobile Highlights* Available at: <http://www.cisco.com>
- [2] H. Tullberg, P. Popovski, Z. Li, Mikko A. Uusitalo, A. Höglund, Ö. Bulakci, M. Fallgren, and J. F. Monserrat "The METIS 5G System Concept – Meeting the 5G Requirements", IEEE Comm. Magazine, vol. 54, no. 12, pp. 132 – 139, Dec. 2016
- [3] J. G. Andrews, S. Buzzi, W. Choi, S. Hanly, A. Lozano, A. C.K. Soong, J. C. Zhang, "What Will 5G Be?" IEEE Journal on Selected Areas in Comm., vol.32, no. 6, pp. 1065-1082, Sept. 2014.
- [4] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, E. Erkip, "Millimeter Wave Channel Modeling and Cellular Capacity Evaluation," IEEE Journal on Selected Areas in Comm., vol. 32, no. 6, pp. 1164–1179, June 2014.
- [5] mmWave module for ns-3; Available at: <https://github.com/nyuwireless/ns3-mmWave>
- [6] ns-3 Network Simulator; Available at: <https://www.nsnam.org>
- [7] M. Mezzavilla, S. Dutta, M. Zhang, M. R. Akdeniz, S. Rangan. "5G mmWave Module for the ns-3 Network Simulator" In Proc. of the 18th ACM Intern. Conf. on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM '15). ACM, NY, USA
- [8] S. Rangan, T. S. Rappaport, E. Erkip, "Millimeter-Wave Cellular Wireless Networks: Potentials and Challenges" Proceedings of the IEEE , vol.102, no.3, pp.366-385, March 2014
- [9] Y. Niu, Y. Li, D. Jin, L. Su, A. V. Vasilakos, "A Survey of Millimeter Wave (mmWave) Communications for 5G: Opportunities and Challenges", Wireless Networks, vol 21, no. 8, pp. 2657-2676, Nov. 2015.

# Arduino-based system for soil moisture measurement

Vuk Radman, Milutin Radonjić, *Member, IEEE*

**Abstract** — A substantial number of modern systems for soil moisture measurement exploit two properties of soils: moisture content and water potential. Sensors for measuring these parameters vary substantially in precision and cost. Moisture content sensors provide data that is more suitable for human reading. However, these sensors are less reliable or very expensive. In this paper we describe the system that utilizes inexpensive and precise sensors, which provide a convenient way of data presentation. For this purpose, the Watermark 200SS granular matrix water potential sensor is used, supported by a scaled MV-VG model for soil moisture retention curve prediction. This model provides a correlation between the soil moisture content and water potential. Since our proposed system relies on inexpensive components, it is a suitable alternative to similar expensive systems, especially for multiple monitoring locations.

**Keywords** — soil moisture measurement, Arduino UNO, Watermark 200SS, soil moisture retention curve.

## I. INTRODUCTION

DEVELOPMENTS in technology, science and medicine have allowed a much better quality and longevity of a man's life, leading to a continuous increase in population. Consequently, requirements for adequate quantities of food and drinking water has increased, as well as the need to improve its quality, especially in third world countries where each of these recourses are limited. Through the usage of the "smart" irrigation systems and monitoring of soils we are heading in the right direction to meet these requirements.

Soil moisture measurement is of great importance for applications in the field of agriculture, allowing farmers to more efficiently manage irrigation systems. Knowing the exact properties of the soil, farmers are not only able to use less water to grow crops, but also increase production and quality, irrigating the plants in the critical periods of its growth. Moisture sensors are commonly used in combination with irrigation system controllers. Connecting a moisture sensor to an ordinary timer irrigation system, it gets some "smart" features, stopping irrigation when the soil is already wet enough. Soil moisture sensors also found diverse applications in scientific research of agriculture and horticulture, climate and environment related research and many other areas of science. For example, on golf courses these sensors are used to prevent

excessive soaking, as well as to prevent runoff of fertilizers and other chemicals into the deeper layers of the soil.

Purpose of this work was to design a simple system for soil moisture monitoring, using affordable components that provide reliable and accurate measurement. The system is intended primarily for application in agricultural activities. Although there are a number of soil moisture monitoring systems on the market, numerous disadvantages limit the scope of their application. Systems that provide data in the form suitable for agricultural applications (soil moisture volume percentage / soil moisture content) are very expensive and require sophisticated electronic equipment for their operation. On the other hand, affordable systems usually display data in unsuitable form (voltage, resistivity, frequency, suction force, ...).

The sensor used in our system measures suction force (water potential), while the water content mostly depends on physical characteristics of the soil. To draw a correlation between these two parameters (water potential and moisture content), we need to determine the soil moisture retention curve (SMR). The accurate SMR determination is usually performed by laboratory analysis, which we will substitute with a complex mathematical model. Due to certain model approximations and the specificities of different soil types, additional calibrations are necessary for very precise measurements.

Paper is organized as follows. Determination of SMR is explained in Section II. Hardware design of proposed system is presented in Section III. Calibration of obtained results is elaborated in Section IV. Paper is concluded in Section V.

## II. SOIL MOISTURE RETENTION CURVE

Measuring the soil parameters, such as the soil moisture retention curve (SMR) is expensive and time consuming, very difficult and requires laboratory conditions and equipment. Therefore, many scientists have made an effort to develop indirect methods as an alternative. These indirect methods are classified into empirical, conceptual and semi-physical.

Semi-physical methods are based on curve similarity between the SMR and the particle size distribution curve (PSD), implying that the pore size distribution (PoSD) is closely related to the PSD. This PoSD-PSD relation is determined by scaling the pore length using a scaling factor  $\alpha$ , of an average value of 1.38. However, it is later revealed that  $\alpha$  value varies between 1.02 and 2.97 for

V. Radman is with the Faculty of Electrical engineering, University of Montenegro, Podgorica (e-mail: [vuk.radman@gmail.com](mailto:vuk.radman@gmail.com))

M. Radonjić is with the Faculty of Electrical engineering, University of Montenegro, Podgorica (e-mail: [mico@ac.me](mailto:mico@ac.me))

fine- and coarse- textured soils, respectively [1]. Even the slightest error in the estimation of  $\alpha$  may result in immense error while estimating the SMR. The essential problem of the semi-physical methods is leading to the dry range underestimation of the SMR.

To predict the continuous SMR from the PSD data and porosity, we used the parameterized Van Genuchten model (VG), which was integrated with the Mohammadi and Meskini-Vishkae (MV) model. Model was tested on 82 soils samples from the UNSODA database, which has shown that the scaling approach improves the MV-VG model estimations on average by 30% [2]. Results show that the scaled method properly estimates the SMR for all soil samples and may be used at larger field scale for predicting continuous SMR from basic soil data.

### A. Theory

Model of the SMR developed by the Van Genuchten [3] is widely used and mathematically defined as:

$$S_e = \left[ \frac{1}{1+(\alpha h)^n} \right]^m, \quad (1)$$

$$S_e = \frac{\theta_i - \theta_r}{\theta_s - \theta_r}, \quad (2)$$

where  $S_e$  is effective saturation degree,  $\theta_i$  is the soil moisture content,  $\theta_s$  and  $\theta_r$  are saturated and residual soil moisture contents, respectively. Parameter  $h$  represents the suction head,  $n$  represents pore-size distribution index,  $\alpha$  is the pore-length scaling factor and  $m$  is a fitting coefficient.

The suction head,  $h_i$  is given as:

$$h_i = \frac{0.543 \cdot 10^{-4}}{R_i} \xi, \quad (3)$$

and corresponds to the particle radius of the  $i$ -th fraction  $R_i$  [4]. Parameter  $\xi$  is defined as:

$$\xi = \frac{1.9099}{1+e}, \quad (4)$$

and represents a coefficient depending on the state of soil particle packaging. Parameter  $e$  is the void ratio given by:

$$e = \frac{\rho_s - \rho_b}{\rho_s}, \quad (5)$$

where  $\rho_s$  is soil particle density and  $\rho_b$  is bulk density.

Because of the similarity between the SMR and PSD curves, many scientists expressed a SMR model in terms of a PSD model. Arya and Paris (AP) [5] suggested that the soil moisture content  $\theta_i$  can be obtained from PSD and  $\theta_s$  as

$$\theta_i = \theta_s \sum_{j=1}^{j=i} \omega_j; \quad i = 1, 2, 3, \dots, k, \quad (6)$$

where  $\omega_j$  is the mass fraction of particles in the  $j$ -th particle size fraction. Let us now consider that

$$P_i = \sum_{j=1}^{j=i} \omega_j, \quad (7)$$

where  $P_i$  is the cumulative mass fraction of soil particles, would result in

$$\theta_i / \theta_s = S, \quad (8)$$

where  $S$  is the saturation degree. If  $\theta_r = 0$ , then  $S_e = S$  and subsequently  $S = P_i$ . However, the AP model ignores the residual moisture content, which has considerable values for many types of soil. Combining (1) and (3) and invoking them in (7) yields

$$P_i = \left[ \frac{1}{1+(\alpha \frac{0.543 \cdot 10^{-4}}{R_i} \xi)^n} \right]^m. \quad (9)$$

Fitting (9) to the PSD data enables us to predict the

continuous SMC parameters such as  $n$ ,  $m$  and  $\alpha$ .

Since the MV-VG model assumes that the residual moisture content is zero, leading to an underestimation in dry range of the SMR, especially in clayey soils, we used a conceptual scaling approach to reduce the model bias.

### B. Scaling Approach

Let us assume that the reference soil consists of particles, which are spherical and uniform in size and are arranged in random close packing state leading to minimal porosity. This minimal porosity is known as the "Kepler conjecture" with value of 0.259 for this packing state or, in our case, for the reference soil. On the other hand, the maximum value of porosity  $\xi_{max}$  for the reference soil would equal to 1.41432. Thus we can suggest the scaling factor  $\lambda$ , for each soil sample as:

$$\lambda = \frac{\xi}{\xi_{max}}. \quad (10)$$

Now being that the packing parameter  $\xi$  and pore-size distribution index  $n$  are small for the fine-textured and large for the coarse-textured soils, F. Meskini-Vishkae, M. H. Mohammadi and M. Vancloster suggest that the scaling factor  $\lambda$  can scale the index  $n$  as follows:

$$n^* = \lambda * n, \quad (11)$$

where  $n^*$  is scaled to the PoSD index in the VG model [2]. Finally, the scaled model is determined as:

$$\theta_i / \theta_s = \left[ \frac{1}{1+(\alpha h)^{n^*}} \right]^m. \quad (12)$$

## III. HARDWARE DESIGN

Our design is based on the Arduino UNO microcontroller platform, equipped with the soil moisture sensor Watermark 200SS and the GSM/GPRS SIMCOM 900 mobile communication module (Fig. 1). The sensor is connected to a built-in 10-bit analog-to-digital converter (ADC) of the Arduino platform. The platform measures output voltage of the moisture sensor, processes obtained data and sends it over the GPRS connection to the cloud.

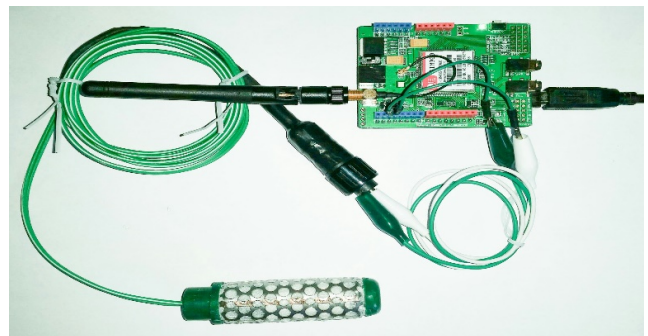


Fig. 1. Experimental prototype of the system.

In order to properly measure the water potential using the Watermark sensor, it is necessary to provide an alternating current (AC) power source, because long exposure to the direct current polarize sensor electrodes. This has a negative impact on the measured results, and leads to the degradation of the sensor over time. The AC power source is emulated by digital output ports of the microcontroller, as shown in Fig. 2. Alternating the status



of the digital pins D0 and D1 to HIGH (5V) and LOW (0V) digital levels, rectangular signal is generated. Initially, D0 is set to HIGH, while D1 is set to LOW, so the current flows in one direction. We then measure the voltage drop on the sensor and perform the necessary calculations. The voltage polarity is then reversed, putting D0 to LOW and D1 to HIGH. The current now flows in the opposite direction and the measurement starts again.

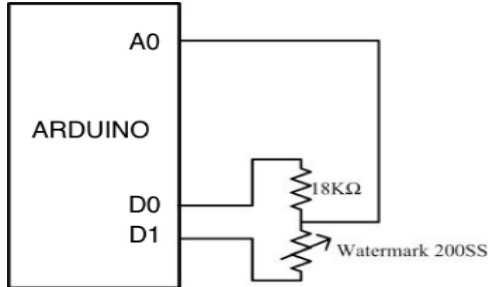


Fig. 2. Watermark 200SS sensor connection to Arduino.

Voltage drop (resistance of the Watermark 200SS sensor) is measured using A0 input of the Arduino ADC. Arduino uses the 10-bit ADC converter which has a 4.88 mV resolution. This means that the final resistance value is to be calculated with a certain error due to the imprecision of the ADC. The sensor voltage is calculated according to the voltage divider as follows:

$$V = 5 * \frac{R_w}{R_w + 18000} \quad (13)$$

where V is the sensor voltage and  $R_w$  is sensor resistance.

After calculation of the resistance using (13), we proceed to the soil moisture potential calculation using (14), (15), and (16). Finally, after applying the moisture potential into scaled MV-VG model using (12), we acquire the value of the moisture content for the desired soil type. The final value is then transmitted using the GPRS module to the desired location.

We should take into consideration a flaw of this method of measurement. Due to the large difference of the unknown (sensor) resistance and the reference resistors (18kΩ), the result can be fairly inaccurate. This problem arises due to the entirety of voltage drop occurs across the resistor of larger value, so consequently the unknown voltage carries far less useful information during AD conversion. This problem can be resolved by an auto-ranging voltage divider.

#### IV. CALIBRATION OF RESULTS

In order to perform the required conversion of the sensors electrical resistance into the soil moisture potential it is necessary to perform the calibration. For this purpose, a number of calibration points, at a temperature of 24°C, are provided by the manufacturer [6] (Table 1).

Based on calibration points given in Table 1 it is possible to draw the transfer function of the Watermark 200SS sensor (Figure 3). The most efficient way for reproducing the curve in Fig. 3 is to divide the resistance domain into three ranges and project the corresponding curves for each of them. The first section of the curve covers the range from 0Ω to 1000Ω, which is estimated by

the linear function:

$P = -20 * (R * (1 + 0.018 * (T - 24))) - 0.55$  (14) where  $P$  (kPa) is a potential humidity,  $R$  (kΩ) is the sensor resistance and  $T$  (°C) is the soil temperature. The value  $0.018*(T-24)$  represents the 1.8% resistance change for each degree of Celsius change in soil temperature from the reference value of 24°C.

TABLE I. CALIBRATION POINTS OF THE WATERMARK 200SS

SENSOR	
Moisture potential [kPa/cbar]	Resistivity [Ω]
0	550
9	1000
10	1100
15	2000
35	6000
55	9200
75	12200
100	15575
200	28075

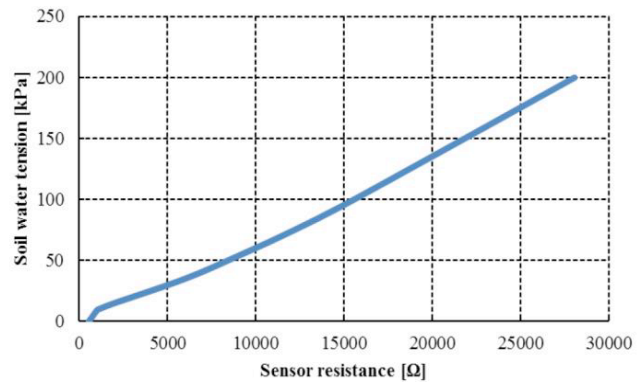


Fig. 3. Transfer function of the Watermark 200SS sensor.

The second segment projection, which covers the range from 1000Ω to 8000Ω, was developed by Shock [7]. This equation, developed from purely empirical data, deviates from the values in Table 1 for moisture potential values greater than 100 kPa :

$$P = \frac{-3.213 * R - 4.093}{1 - 0.009733 * R - 0.01205 * T} \quad (15)$$

Finally, for scope over 8kΩ, quadratic equation developed by Rick Allen [8] is used:

$$P = -2.246 - 5.239 * R * (1 + 0.018 * (T - 24)) - 0.06756 * R^2 * (1 + 0.018 * (T - 24))^2 \quad (16)$$

Equation (16) was determined by the simple square regression of the values taken from Table 1 in the range of 10kPa to 200kPa. The equation has a coefficient of determination  $r^2=0.9996$  and a standard error of 1.07kPa [8].

Although the temperature sensor calibration is necessary for serious research of soil moisture, our design does not implement the temperature sensor. The system was originally designed for agricultural uses and in the growing season changes in soil temperature rarely vary from the standard 24°C. However, the possibility of a temperature calibration is included in the calculations and may be implemented in the future works.

Typical examples of the measured and predicted SMRs values with the MV-VG model, scaled approach and

Rosetta software (computational modeling and analysis software based on a purely empirical data) for clay and silt loam textures, are shown in Fig. 4 and Fig. 5 [2]. For both textures, scaled approach well corresponds to the measured data and gives better prediction than the MV-VG model in the whole range of the SMR. For silty loam (Fig. 5) scaled approach overestimates the measured data slightly while MV-VG model underestimates the wet part of the SMR [2]. The graphs also shows that Rosetta software performs very badly while predicting the wet parts of the SMR.

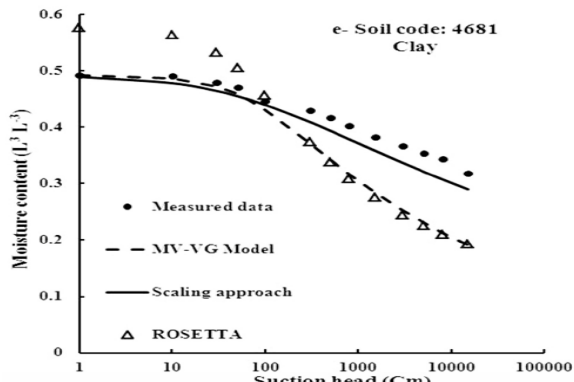


Fig. 4. Example of measured vs. predicted SMRs using integrated MV-VG model, scaling approach and Rosetta software for clay textured soils.

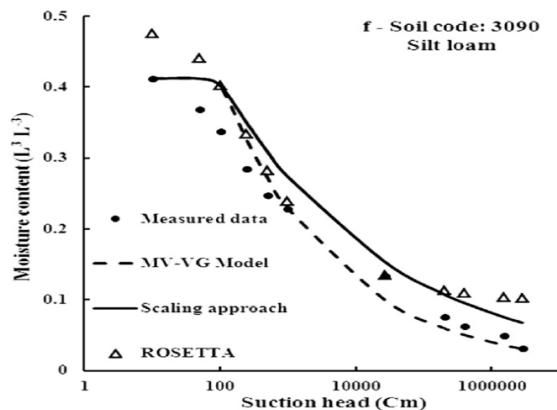


Fig. 5. Example of measured vs. predicted SMRs using integrated MV-VG model, scaling approach and Rosetta software for silty loam textured soils.

The proposed system has been tested in a sample of clay soil in a large flower pot at a depth of 20 cm. The results shown in Fig. 6 were obtained during a short test of tens of minutes of the previously sprinkled soil, with periods of measurement of a little more than a minute. The initial soil moisture was at 50.9%. The soil is then briefly moistened and allowed to dry over a period of several minutes. Increase in soil moisture (from first to fifth point) can be noticed until the soil has reached the maximum humidity saturation of 51%. The excess water is then removed and slight pressure is applied to the soil to remove a certain amount of water for the purposes of demonstration (fifth to tenth point). The rest of the graph shows a constant value since the drying of clay soils is a time consuming process.

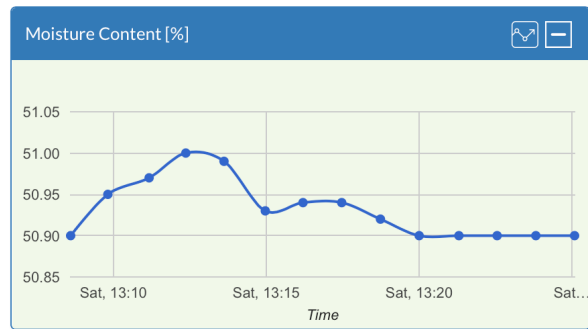


Fig. 6. Measured moisture content for clay soil.

## V. CONCLUSION

In this paper we described design of the microcontroller system for soil moisture measurement using Watermark 200SS soil moisture sensor. The idea was to develop a modular, flexible system using affordable components. Proposed solution is designed to measure water potential from 0kPa to 200kPa, and to be used for agricultural purposes. System is flexible, low cost, simple to use and offers real-time monitoring. It is a suitable alternative to similar expensive systems, especially in case of multiple monitoring locations.

Future work on the system will include several different activities. Main point of interest will be to evaluate the recorded data and compare it to high accuracy systems. Attention will also be focused to autonomy of the system, which can be achieved through autonomous power supply consisting of solar energy or batteries and implementation of attention (AT) commands. This greatly improves flexibility and manageability of the system as it allows us to control the system from a central point wirelessly, without the need of field work. Another important issue is to implement a temperature sensor for increased precision and calibration of the moisture sensor.

## LITERATURE

- [1] A. Basile and G. D'Urso, "Experimental corrections of simplified methods for predicting water retention curves in clay-loamy soils from particle-size determination", *Soil Technol.*, 10, 261–272, 1997.
- [2] F. Meskini-Vishkaee, M. H. Mohammadi, and M. Vanclouster. *Hydrol. Earth Syst. Sci.*, 18, 4053–4063, 2014. "Predicting the soil moisture retention curve, from soil particle size distribution and bulk density data using a packing density scaling factor".
- [3] Van Genuchten, M. Th. 1980. "A closed-form equation for predicting the hydraulic conductivity of unsaturated soils". *Soil Sci. Soc. Am. J.* 44:892-898.
- [4] Mohammadi, M. H. and Vanclouster, M.: Predicting the soil moisture characteristic curve from particle size distribution with a simple conceptual model, *Vadose Zone J.*, 10, 594–602, 2011.
- [5] Arya, L. M. and Paris, J. F.: "A physicoempirical model to predict the soil moisture characteristic from particle-size distribution and bulk density data". *Soil Sci. Soc. Am. J.*, 45, 1023–1030, 1981.
- [6] From Irrometer Company, Inc., Riverside, CA via M.K. Hansen Company, E.Wanatchee, WA, October, 2000.
- [7] Shock, C.C., J.M. Barnum, and M. Seddigh. 1998. Calibration of Watermark soil moisture sensors for irrigation management. *Proceedings of the 1998 Annual Meeting of the Irrigation Association.* p. 139-146.
- [8] Rick Allen (October 13, 2000). "Calibration for the Watermark 200SS Soil Water Potential Sensor to fit the 7-19-96 "Calibration #3" Table from Irrometer". University of Idaho.

CIP - Каталогизација у публикацији  
Национална библиотека Црне Горе, Цетиње

ISBN 978-86-85775-20-8  
COBISS.CG-ID 32715536

ISBN 978-86-85775-20-8



9 788685 775208 >